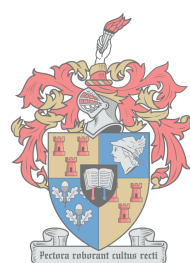


# Methods for aggregating microeconomic data: applications to art prices, business sentiment and historical commodity prices

by Laurie H. Binge



UNIVERSITEIT  
iYUNIVESITHI  
STELLENBOSCH  
UNIVERSITY

100  
1918 - 2018

*Dissertation presented for the degree of Doctor of Philosophy in Economics in the Faculty of Economic and Management Sciences at Stellenbosch University.*

SUPERVISOR: PROFESSOR WILLEM H. BOSHOFF

March 2018

## Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2018

Copyright © 2018 Stellenbosch University

All rights reserved

Declaration

---



P.O. Box 263  
Constantia  
7848

**19 September 2017**

**University of Stellenbosch, Dept. of Economics**

Private Bag X1  
Matieland  
7599

**Attention:** Mr Laurie Binge

Dear Laurie,

**CERTIFICATION: COPY-EDITING OF YOUR ECONOMICS PHD DISSERTATION**

I wish to certify that I have copy-edited the economics PhD dissertation of Mr Laurie Binge.

Further, in terms of my background and experience, wish to point out the following.

I founded WriteArt in 2004 and have been working as a full-time professional copy-editor and proofreader since then.

Yours is among one of the many theses and dissertations that I have copy-edited for students of many universities, in South Africa and abroad. In addition, I have copy-edited numerous articles and papers for academics, also from many universities worldwide.

I also copy-edit other types of publications for a range of clients.

Please contact me if you have any queries.

Yours sincerely,

Russell de la Porte

## Abstract

In the modern world, large microeconomic datasets are becoming increasingly available due to technological developments. These datasets provide an opportunity to improve the measurement of a range of economic phenomena and to bolster economic research. One particular way in which these large datasets can aid economic analysis is to allow the creation of macroeconomic indicators from aggregated microeconomic data. Yet, there are often challenges in aggregating these large datasets and in identifying the underlying pattern in the data.

The aim in this dissertation is to explore aggregation methods that overcome specific challenges in aggregating three relatively large microeconomic datasets to create time-series indicators. The first case explores aggregation methods for estimating South African art price indices (2000Q1-2015Q4), using a large database of art auction prices. The challenge in aggregating this dataset is that artworks are by and large unique and infrequently traded, which means that the composition of items sold is not constant over time. To address this challenge, central tendency, hedonic and hybrid repeat sales methods are used to estimate quality-adjusted South African art price indices.

The second case explores aggregation methods for estimating indicators of business confidence and uncertainty for South Africa (1992Q1-2016Q3), using the Stellenbosch University Bureau for Economic Research's (BER) business tendency surveys. The challenge in aggregating this dataset is to measure these concepts by identifying a pattern in the disparate views of individual agents. To address this challenge, aggregation methods for capturing the full distribution of the qualitative survey responses are explored. The cross-sectional weighted first and second moments of the distribution of responses are calculated to create new indicators of business confidence and uncertainty for South Africa.

The third case explores aggregation methods for estimating monthly commodity price indices for the Cape Colony (September 1889 - July 1914), using two newly digitised datasets of commodity prices for various towns in the Colony. The challenge in aggregating these datasets is that both sets of records are incomplete, in terms of the coverage of both products and towns. The repeat sales method is used to aggregate the incomplete price series for various towns from both sources, to create more complete monthly commodity price indices for the Cape Colony.

Testing specific hypotheses is useful, both in demonstrating the potential research applications for the aggregated indicators, and in assessing the validity of the proposed aggregation methods. The dissertation therefore uses the time-series indicators to test a specific hypothesis in each case. The first case examines the estimated South African art price indices for evidence of a bubble. The hypothesis that South African art prices exhibited mildly explosive behaviour between 2000 and 2015 is tested. The second case examines the relationship between business sentiment and real activity in South Africa, by testing the hypothesis that there was significant comovement between the sentiment indicators and real GDP growth. The third case examines the commodity price indices, as well as indicators of price dispersion, for evidence of increasing internal market integration in the

---

## Abstract

---

Cape Colony. The hypotheses of price convergence between towns and cointegration of regional price indices are tested.

This dissertation is a contribution to the literature in that it demonstrates suitable aggregation methods to overcome some of the challenges in aggregating relatively large microeconomic datasets. These aggregation challenges relate to (i) estimating quality-adjusted price indices for unique and infrequently traded items, (ii) developing aggregate measures of sentiment based on the disparate views of a large number of respondents, and (iii) estimating complete price indices from data that is incomplete. These aggregation methods may prove useful in a variety of settings where there are similar challenges. The estimated time series may prove useful for further research in each of the relevant fields and are reported in the chapter appendices.

## Opsomming

In die moderne wêreld lei tegnologiese ontwikkelings tot die toenemende beskikbaarheid van groot mikro-ekonomiese datastelle. Hierdie datastelle bied die geleentheid om verskeie ekonomiese verskynsels beter te meet en om sodoende ekonomiese navorsing te verbeter. Een wyse waarop hierdie groot datastelle ekonomiese ontleding kan ondersteun is in die skep van makro-ekonomiese aanwysers op grond van geaggregeerde mikro-ekonomiese data. Tog is daar dikwels uitdagings rondom die aggregasie van hierdie datastelle en die identifisering van die onderliggende patroon in die data.

Die doel van hierdie proefskrif is om metodes te ondersoek wat spesifieke uitdagings in die aggregasie van drie relatief groot mikro-ekonomiese datastelle kan oorkom, ten einde tydreeksaanwysers te beraam. Die eerste geval behels 'n studie van aggregasiemetodes om Suid-Afrikaanse kunsprijsindekse te bereken (2000K1 tot 2015K4), met behulp van 'n groot databasis van kunsveilingpryse. Die uitdaging vir die aggregasie van hierdie datastel is dat die kunswerke merendeels uniek is en nie gereeld verhandel word nie, wat beteken dat die komposisie van die items wat wel verkoop word nie konstant is oor tyd nie. Sentrale-tendens-, hedoniese en gemeng-herhaalde-verkope-metodes word ondersoek ten einde kwaliteit-aangepaste Suid-Afrikaanse kunsprijsindekse te beraam.

Die tweede geval behels 'n ondersoek na aggregasiemetodes om aanwysers van sakevertroue en onsekerheid vir Suid-Afrika te beraam (1992K1 tot 2016K3), op grond van saketendens-opnames van die Buro vir Ekonomiese Onderzoek aan die Universiteit Stellenbosch. Die uitdaging vir die aggregasie van hierdie datastel is om hierdie konsepte te meet deur 'n onderliggende patroon te identifiseer in die uiteenlopende vooruitsigte van individuele firmas. Om hierdie uitdaging aan te spreek word aggregasie metodes ondersoek om die volle verdeling van die kwalitatiewe antwoorde vas te vang. Die kruissnit-geweegde-eerste en -tweede momente van die verdeling word bereken om aanwysers van sakevertroue en onsekerheid vir Suid-Afrika te beraam.

Die derde geval ondersoek aggregasiemetodes om maandelikse kommoditeitsprijsindekse vir die Kaapkolonie te beraam (September 1889 tot Julie 1914), met behulp van twee nuwe datastelle van kommoditeitspryse vir verskeie dorpe in die Kolonie. Die uitdaging by die aggregasie van hierdie twee datastelle is dat beide stelle rekords onvolledig is, met wesenlike gapings in die individuele reekse. Die herhaalde-verkope metode gebruik om die onvolledige rekords saam te voeg om meer volledige maandelikse kommoditeitsprijsindekse vir die Kaapkolonie te beraam.

Die toets van spesifieke hipoteses is nuttig, beide om potensiële toepassings vir die aanwysers te demonstreer, en om die geldigheid van die aggregasiemetodes te evalueer. Die proefskrif gebruik dus die tydreekse om 'n spesifieke hipotese te toets in elke geval. Die eerste geval ondersoek die beraamde Suid-Afrikaanse kunsprijsindekse vir bewyse van 'n prysborrel. Die hipotese van eksplosiewe gedrag in Suid-Afrikaanse kunsprijsindekse gedurende die periode word getoets. Die tweede geval ondersoek die verhouding tussen sakesentiment en reële ekonomiese aktiwiteit in Suid-Afrika. Die hipotese dat sentimentaanwysers beduidend met reële BBP groei saambeweeg het word getoets. Die derde

---

## Opsomming

---

geval ondersoek die beraamde kommoditeitsprysindekse, asook aanwysers van prysverspreiding, vir bewyse van toenemende interne mark-integrasie in die Kaapkolonie. Die hipoteses word getoets dat pryse tussen dorpe in die Kaapkolonie gedurende die steekproefperiode konvergeer het, en dat daar ko-integrasie tussen streeksprysindekse bestaan.

Hierdie proefskrif lewer 'n bydrae tot die literatuur deur tegnieke te demonstreer wat spesifieke uitdagings by die aggregasie van relatief groot mikro-ekonomiese datastelle kan oorkom. Die uitdagings behels (i) die beraming van kwaliteit-aangepaste prysindekse vir items wat uniek is en nie gereeld verhandel word nie, (ii) die ontwikkeling van geaggregeerde aanwysers van sentiment gebaseer op die uiteenlopende vooruitsigte van 'n groot aantal respondente, en (iii) die beraming van volledige prysindekse op grond van onvolledige data. Die aggregasietegnieke mag nuttig wees in ander gevalle waar soortgelyke uitdagings bestaan. Die beraamde aanwysers mag nuttig wees vir toekomstige navorsing in elk van hierdie velde en word in die bylaag van elke hoofstuk voorsien.

## Acknowledgements

My supervisor, Wimpie Boshoff, provided the initial impetus for this endeavour. Without his valuable insight and optimism throughout, this dissertation would not have been possible. I am grateful to a number of people for their assistance with the underlying data. George Kershoff assisted in providing and analysing the BER survey data, Johan Fourie provided much of the historical data, Nicola Theron and Cobus Venter organised the auction data, and Rick Walker helped with some of the financial data. Jaco Franken made the ‘process’ as smooth as possible and my fellow GEM students offered valuable feedback and friendship. My parents-in-law were always ready with a kind word and a cup of tea. Finally, I would like to thank my brother and my parents for their unwavering encouragement and support, and Elize, for being my *fish*.



## Table of Contents

**Contents**

<b>Declaration</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Opsomming</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>Table of Contents</b>	<b>viii</b>
<b>List of Figures</b>	<b>xiv</b>
<b>List of Tables</b>	<b>xvi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Art Prices . . . . .	2
1.2 Business Sentiment . . . . .	5
1.3 Historical Commodity Prices . . . . .	7
1.4 Objectives and Contributions . . . . .	9
<b>2 Methods for Estimating Quality-Adjusted Price Indices: An Application to South African Art Prices</b>	<b>12</b>
2.1 Introduction . . . . .	12
2.2 Methodologies for Constructing Art Price Indices . . . . .	14
2.2.1 Central Tendency Methods . . . . .	14
2.2.2 The Hedonic Method . . . . .	15
2.2.3 The Repeat Sales Method . . . . .	18
2.2.4 Hybrid Methods . . . . .	20
2.2.5 South African Art Price Literature . . . . .	22
2.3 Data: South African Art Auction Prices . . . . .	23
2.3.1 Public and Private Art Market Prices . . . . .	23
2.3.2 South African Art Auction Prices . . . . .	25
2.3.3 Artwork Characteristics . . . . .	26
2.3.4 Continuous Artist Reputation Variable: Two-Step Hedonic Approach . . . . .	29
2.4 Index Results . . . . .	30
2.4.1 Central Tendency Indices . . . . .	31
2.4.2 Hedonic Indices . . . . .	32
2.4.3 Repeat Sales and Hybrid Indices . . . . .	35
2.4.4 Market Segments and Artist Indices . . . . .	39
2.5 Validity Tests and Evaluation . . . . .	43

## Table of Contents

2.5.1	Internal Validity Test . . . . .	44
2.5.2	External Comparison: Traditional South African Assets . . . . .	46
2.5.3	Evaluation: Smoothness . . . . .	49
2.6	Bubble Detection . . . . .	52
2.6.1	Bubble Detection Framework . . . . .	54
2.6.2	Bubble Detection Results . . . . .	56
2.6.3	Market Segments . . . . .	59
2.6.4	Discussion . . . . .	60
2.7	Conclusion . . . . .	63
2.8	Appendix . . . . .	64
2.8.1	Rolling Coefficient Estimates . . . . .	64
2.8.2	Index Confidence Intervals . . . . .	65
2.8.3	Weighted Bubble Detection Tests . . . . .	69
2.8.4	South African Art Price Indices . . . . .	71
<b>3</b>	<b>Methods for Aggregating Disparate Qualitative Survey Responses: An Application to Business Sentiment in South Africa</b>	<b>73</b>
3.1	Introduction . . . . .	73
3.2	Confidence . . . . .	75
3.2.1	Theory on Confidence and Economic Outcomes . . . . .	75
3.2.2	Empirical Findings . . . . .	77
3.3	Uncertainty . . . . .	82
3.3.1	Theory on Uncertainty and Economic Outcomes . . . . .	82
3.3.2	Empirical Findings . . . . .	84
3.4	Data: Business Tendency Surveys . . . . .	89
3.4.1	The BER Business Tendency Surveys . . . . .	89
3.4.2	The BER Business Confidence Indicator . . . . .	93
3.5	Methodology . . . . .	94
3.5.1	Confidence . . . . .	95
3.5.2	Uncertainty . . . . .	98
3.5.3	Weighing the Survey Responses . . . . .	104
3.6	Results: Confidence . . . . .	107
3.6.1	Confidence Indicators . . . . .	108
3.6.2	Validity Tests and Evaluation . . . . .	110
3.7	Results: Uncertainty . . . . .	116
3.7.1	Uncertainty Indicators . . . . .	116
3.7.2	Validity Tests and Evaluation . . . . .	119
3.8	The Relationship between Business Sentiment and Real Economic Activity . . . . .	124
3.8.1	Granger Causality Tests . . . . .	125
3.8.2	VAR Analysis . . . . .	126

## Table of Contents

3.9	Conclusion . . . . .	137
3.10	Appendix . . . . .	139
3.10.1	Confidence Intervals for the Confidence Indicators . . . . .	139
3.10.2	Stable Sample Indicators . . . . .	139
3.10.3	Unit Root Tests . . . . .	145
3.10.4	IRFs for the Confidence Indicators . . . . .	145
3.10.5	Sentiment Indices . . . . .	147
<b>4</b>	<b>Methods for Aggregating Incomplete Time Series from Various Sources: An Application to Commodity Prices in the Cape Colony, 1889-1914</b>	<b>150</b>
4.1	Introduction . . . . .	150
4.2	Economic History in the Cape Colony: 1886-1914 . . . . .	152
4.2.1	Overview of Economic Activity in the Cape Colony . . . . .	153
4.2.2	The Second South African War (1899-1902) . . . . .	157
4.2.3	Overview of Existing Research on Prices and Inflation . . . . .	160
4.3	Data: Commodity Prices . . . . .	162
4.3.1	Commodity Prices in the Cape Colony . . . . .	163
4.3.2	Pricing Dynamics in the Monthly Records . . . . .	165
4.4	Methodology . . . . .	168
4.4.1	Limitations of Traditional Techniques . . . . .	169
4.4.2	The Repeat Sales Method . . . . .	171
4.4.3	Creating Indices for Specific Commodities . . . . .	173
4.4.4	Weighing the Commodity Indices . . . . .	175
4.5	Index Results . . . . .	177
4.5.1	Repeat Sales Price Indices . . . . .	178
4.6	Validity Tests . . . . .	181
4.6.1	Internal Validity . . . . .	181
4.6.2	External Validity: Consumer Price Indices . . . . .	183
4.7	Evaluation: The Money Supply . . . . .	184
4.7.1	The Monetary and Banking System in the Cape Colony . . . . .	185
4.7.2	Money Supply and Inflation . . . . .	188
4.8	Market Integration . . . . .	190
4.8.1	Development of the Railways . . . . .	190
4.8.2	Market Integration in the Cape Colony . . . . .	192
4.8.3	Methodology: Tests of Market Integration . . . . .	194
4.8.4	Market Integration Results . . . . .	199
4.9	Conclusion . . . . .	205
4.10	Appendix . . . . .	206
<b>5</b>	<b>Conclusion</b>	<b>208</b>

Table of Contents

---

5.1	Art Prices . . . . .	209
5.2	Business Sentiment . . . . .	210
5.3	Historical Commodity Prices . . . . .	211
	<b>References</b>	<b>213</b>

---

## List of Figures

2.1	Total number of auction lots sold, by auction house (2000Q1-2015Q4)	26
2.2	Boxplot of (log) hammer prices	27
2.3	Relationship between prices and artwork sizes, by medium	28
2.4	Relationship between prices and the reputation index	30
2.5	Central tendency indices of South African art prices (2000Q1=100)	31
2.6	Chain-linked 2-year adjacent-period art price index	34
2.7	Hedonic indices of South African art prices (2001=100)	35
2.8	Repeat sales index of South African art prices (2000Q4=100)	36
2.9	Pseudo-repeat sales samples	37
2.10	Pseudo-repeat sales indices of South African art prices (2001=100)	38
2.11	Histogram of (log) art prices	40
2.12	Art price indices for different price segments (2000Q1=100)	40
2.13	Histogram of (log) average art prices (2000-2004)	41
2.14	Art price indices for value and medium segments (2000Q1=100)	42
2.15	Art price indices for individual artists (2000=100)	43
2.16	Comparing South African art price indices (2000Q1=100)	44
2.17	Comparing South African real asset price indices (2001=100)	47
2.18	Test statistics and critical values for models without drift	58
2.19	Test statistics and critical values for models with drift	58
2.20	Coefficient estimates from the 1-year adjacent-period models	65
2.21	Coefficient estimates and adjusted R-squared from the 1-year adjacent-period models	66
2.22	Hedonic indices with the full hedonic model confidence intervals	67
2.23	2-year adjacent-period hedonic model coefficients with confidence intervals	67
2.24	Pseudo-repeat sales indices with larger sample (ps-RS2) model confidence intervals	68
2.25	Art price indices with the 1-year adjacent-period hedonic model confidence intervals	68
2.26	Art price indices with ps-RS2 model confidence intervals	69
2.27	Test statistics and critical values for weighted models without drift	70
2.28	Test statistics and critical values for weighted models with drift	70
3.1	Number of respondents over time, by sector (1992Q1-2016Q3)	91
3.2	Approximate response rates over time per sector (1992Q1-2016Q3)	92
3.3	Sample sizes of the forecast errors, by sector (1992Q1-2016Q3)	101
3.4	Subsector weights applied in the manufacturing sector	104
3.5	Subsector weights applied in the construction sector	105
3.6	Subsector weights applied in the trade sector	106
3.7	Subsector weights applied in the services sector	106
3.8	GDP share weights applied to the main sectors	107
3.9	Weighted sectoral confidence indicators on current and expected conditions	109
3.10	Weighted confidence indicators on current and expected conditions	109

## List of Figures

3.11	Confidence indicators compared to real GDP growth . . . . .	111
3.12	Cross-correlograms of the confidence indicators and real GDP growth . . . . .	112
3.13	Sectoral confidence indicators compared to real sectoral GDP growth . . . . .	113
3.14	Cross-correlograms of the indicators and real GDP growth in the manufacturing sector	113
3.15	Confidence indicator turning points compared to the official SARB turning points .	115
3.16	Weighted sectoral indicators of dispersion . . . . .	117
3.17	Weighted sectoral indicators of aggregate forecast error uncertainty . . . . .	117
3.18	Weighted sectoral indicators of idiosyncratic forecast error uncertainty . . . . .	118
3.19	Weighted indicators of dispersion, aggregate error and idiosyncratic error uncertainty	120
3.20	Indicators of economic policy, financial market, and combined uncertainty . . . . .	121
3.21	Cross-correlograms of uncertainty indicators and real GDP growth . . . . .	123
3.22	Cross-correlograms of manufacturing uncertainty indicators and real manufacturing GDP growth . . . . .	124
3.23	IRFs of confidence (current conditions) and real GDP growth . . . . .	128
3.24	FEVDs of confidence (current conditions) and real GDP growth . . . . .	129
3.25	IRFs of confidence (current) and real GDP growth in the manufacturing sector . . .	130
3.26	FEVDs of confidence (current) and real GDP growth in the manufacturing sector . .	130
3.27	IRFs of dispersion and real GDP growth . . . . .	131
3.28	IRFs of uncertainty (combined) and real GDP growth . . . . .	132
3.29	FEVDs of uncertainty (combined) and real GDP growth . . . . .	133
3.30	IRFs of uncertainty (combined) and real GDP growth in the manufacturing sector .	133
3.31	FEVDs of uncertainty (combined) and real GDP growth in the manufacturing sector	134
3.32	Confidence (current) and uncertainty (combined) . . . . .	135
3.33	IRFs of real GDP growth to confidence and uncertainty in the three-variable VAR .	135
3.34	FEVDs of the three-variable VAR . . . . .	136
3.35	IRFs of real GDP, production and investment growth to confidence and uncertainty	137
3.36	Confidence indicators with 95% confidence intervals . . . . .	140
3.37	Sectoral confidence indicators with 95% confidence intervals . . . . .	140
3.38	Confidence indicators based on the full and forecast error samples . . . . .	141
3.39	Uncertainty indicators based on the full and forecast error samples . . . . .	142
3.40	Number of respondents in the stable samples . . . . .	142
3.41	Confidence indicators based on the full and stable samples . . . . .	143
3.42	Uncertainty indicators based on the full and stable samples . . . . .	144
3.43	IRFs of confidence (expected conditions) and real GDP growth . . . . .	146
3.44	IRFs of BER BCI and real GDP growth . . . . .	146
3.45	IRFs of SACCI BCI growth and real GDP growth . . . . .	147
4.1	Real GDP in the Cape Colony (1886-1909) . . . . .	154
4.2	Total imports and exports in the Cape Colony (1886-1909) . . . . .	156
4.3	Cape Colony domestic retail price indices (1910=100) . . . . .	161

## List of Figures

---

4.4	Consumer price indices for the Cape Colony . . . . .	162
4.5	Total number of monthly observations by commodity . . . . .	164
4.6	Monthly wheat prices by town in the Agricultural Journals . . . . .	164
4.7	The size of price changes . . . . .	166
4.8	The frequency of price changes . . . . .	167
4.9	Decomposition of inflation (12-month moving averages) . . . . .	168
4.10	Wheat price indices for the Cape Colony (Nov-1889=100) . . . . .	174
4.11	Individual commodity price indices . . . . .	179
4.12	Commodity group indices (1) . . . . .	180
4.13	Commodity group indices (2) . . . . .	180
4.14	The total commodity price index for the Cape Colony . . . . .	181
4.15	Comparing individual commodity price indices . . . . .	182
4.16	Comparing total commodity price indices for the Cape Colony (1890=100) . . . . .	182
4.17	Comparing the total price index to the CPIs for the Cape Colony . . . . .	184
4.18	Money supply (M1) in the Cape Colony (1886-1909) . . . . .	186
4.19	Comparing the total commodity price index to the monetary aggregates . . . . .	189
4.20	Development of the railways in South Africa . . . . .	191
4.21	Absolute deviations over time for selected commodities . . . . .	200
4.22	Coefficient of variation over time for selected commodities . . . . .	200
4.23	Regional commodity price indices . . . . .	203

---

**List of Tables**

2.1	Hedonic regression results . . . . .	33
2.2	Correlations between art price indices (in growth rates) . . . . .	45
2.3	Correlations between real asset returns . . . . .	48
2.4	Smoothness indicators . . . . .	52
2.5	Dates of explosive price behaviour . . . . .	59
2.6	Dates of explosive price behaviour in the different market segments . . . . .	60
2.7	Dates of explosive price behaviour (weighted models) . . . . .	71
2.8	Art price indices (2000Q1=100) . . . . .	72
3.1	Sample characteristics . . . . .	91
3.2	Survey questions used, by sector . . . . .	98
3.3	Possible forecast errors . . . . .	100
3.4	Comparing sample characteristics in terms of firm size . . . . .	102
3.5	Correlations between confidence indicators and real GDP growth . . . . .	111
3.6	Correlations between sectoral confidence and real sectoral GDP growth . . . . .	112
3.7	Concordance statistics with the SARB business cycle . . . . .	116
3.8	Factor loadings for the first principal component . . . . .	120
3.9	Correlations between the uncertainty indicators . . . . .	122
3.10	Correlations between the sectoral uncertainty indicators and real GDP growth . . . . .	122
3.11	Granger causality tests: confidence . . . . .	125
3.12	Granger causality test statistics: sectoral confidence . . . . .	125
3.13	Granger causality tests: uncertainty . . . . .	126
3.14	Granger causality test statistics: sectoral uncertainty . . . . .	126
3.15	Correlations between indicators based on the full and forecast error samples . . . . .	141
3.16	Sample characteristics in terms of firm size (full and stable samples) . . . . .	143
3.17	Correlations between indicators based on the full and stable samples . . . . .	145
3.18	Unit root test statistics . . . . .	145
3.19	Sentiment indicators . . . . .	148
4.1	Business cycle turning points . . . . .	153
4.2	Census commodity production numbers for the Cape Colony . . . . .	159
4.3	Pricing dynamics . . . . .	168
4.4	Repeat sales example with wheat prices . . . . .	169
4.5	Correlations in overlapping average annual commodity prices . . . . .	175
4.6	Commodity classification . . . . .	176
4.7	Correlations between repeat sales and median indices . . . . .	183
4.8	Correlations between the price indices . . . . .	184
4.9	Banking statistics in South Africa (in £'000s) . . . . .	187
4.10	Correlations between the price indices and monetary aggregates (growth rates) . . . . .	189
4.11	Price dispersion in each region . . . . .	198



List of Tables

---

4.12	Correlations between ADs and CVs . . . . .	201
4.13	Coefficients of time trends . . . . .	202
4.14	Number of significant negative time trends . . . . .	202
4.15	Correlations between regional price indices (growth rates) . . . . .	203
4.16	Number of significant positive correlations . . . . .	203
4.17	Cointegration trace test statistics . . . . .	204
4.18	Number of commodities with one significant cointegrating relationship . . . . .	204
4.19	Total commodity price index . . . . .	207

# 1 Introduction

Technological developments are making it possible to collect increasing amounts of data on many aspects of our lives. These aspects include online spending, interests, opinions, political leanings, and daily activity and exercise. Large microeconomic datasets provide an opportunity to improve the measurement of a range of economic phenomena and to bolster economic research (Schutt and O’Neil, 2013). Microeconomic datasets are of little use if it cannot be properly aggregated, analysed and communicated to decision makers (Tissot, 2016). The abundance of inexpensive computing power enables the processing of large microeconomic datasets to provide useful information, and to distinguish the signal from the noise.

One way in which these microeconomic datasets can aid economic analysis is to allow the construction of better macroeconomic<sup>1</sup> indicators. The richness of microeconomic data can assist in improving the quality and the accuracy of macroeconomic indicators (Tissot, 2016). The improvement in quality and accuracy is dependent, *inter alia*, on the aggregation methods used in constructing the indicators. There are often challenges in aggregating information and in identifying the underlying pattern in the data.

This dissertation considers various challenges in aggregating microeconomic datasets, and demonstrates methods that are useful in dealing with these challenges. The dissertation uses three relatively large microeconomic datasets to create macroeconomic time-series indicators. The three datasets each present specific challenges in estimating time-series indicators.

The first case entails an exploration of aggregation methods for estimating South African art price indices (2000Q1-2015Q4), using a large database of art auction prices. The challenge in aggregating this microeconomic dataset is that artworks are mostly unique and infrequently traded, which means that the composition of items sold is not constant over time.

The second case entails an exploration of aggregation methods for estimating indicators of business confidence and uncertainty for South Africa (1992Q1-2016Q3), using the Stellenbosch University Bureau for Economic Research’s (BER) business tendency surveys. The challenge in aggregating this microeconomic dataset is to measure these concepts by identifying a pattern from the disparate qualitative survey responses.

The third case provides an exploration of aggregation methods for estimating monthly commodity price indices for the Cape Colony (1889/09-1914/07), using two newly digitised datasets of commodity prices for various towns in the Colony. The challenge in aggregating this dataset is that both sets of records are incomplete, in terms of the coverage of both products and towns, with substantial gaps in the individual series.

---

<sup>1</sup>In this dissertation the term ‘macroeconomic’ is used broadly to refer to aggregated indicators, such as South African art price indices, business sentiment indicators for South Africa, and commodity price indices for the Cape Colony.

The aim in this dissertation is to demonstrate aggregation methods that overcome these specific challenges. The first case uses the central tendency, hedonic and hybrid repeat sales methods to estimate quality-adjusted South African art price indices that reflect the mean in the distribution of growth rates. The second case uses aggregation methods to capture the full distribution of the qualitative survey responses, by calculating the cross-sectional weighted first and second moments (mean and variance) of the distribution of responses. In the third case, the repeat sales method is used to aggregate the incomplete price series from both sources, to create more complete monthly commodity price indices for the Cape Colony, which reflect the mean in the distribution of growth rates.

Testing specific hypotheses is useful, both in demonstrating the potential research applications for the aggregated indicators, and in assessing the validity of the proposed aggregation methods. The dissertation uses the indicators to test a specific hypothesis in each case. The first case examines the art price indices for evidence of a bubble in South African art prices. The hypothesis is tested that South African art prices exhibited mildly explosive behaviour over the sample period. The second case examines the relationship between business sentiment and real activity in South Africa. The leading indicator properties of the indicators are investigated, and the hypothesis is tested that there was significant comovement between the sentiment indicators and real GDP growth. The third case investigates the estimated price indices, as well as the cross-sectional moments in the distribution of prices, for evidence of increasing internal market integration. This involves testing the hypotheses that price dispersion between towns was declining and that regional price indices were cointegrated over the period concerned.

The three chapters follow the same structure. The chapters first demonstrate the use of techniques for overcoming the data challenges in aggregating relatively large microeconomic datasets. Next, a set of macroeconomic time-series indicators are estimated and presented: art price indices, sentiment indices, and commodity price indices. The internal validity of the indicators is assessed by comparing the measures calculated with different methodologies. The external validity of the indicators is assessed by testing their conformity to similar existing measures, where available. Each set of indicators is then evaluated according to a specific criterion appropriate in each setting, to identify the most accurate indicators. Finally, the estimated indicators are used to test a specific hypothesis. The following sections present brief chapter summaries.

## 1.1 Art Prices

Contemporary African art has experienced a surge in popularity over the last few decades. The South African art market in particular has grown markedly, both in terms of the number of transactions and total turnover (Fedderke and Li, 2014). Yet, to date there has been little research on the South African art market, and particularly on trends in art prices. This is, at least in part, due to a lack of data on art prices (Campbell, 2009). It is important to analyse price movements over time in

order to understand developments in the art market.

Estimating accurate price indices for unique items such as artworks is challenging (Jiang, Phillips and Yu, 2015). Artworks are typically sold infrequently, which means that only a small sample of market prices are observed at any given time. Artworks are also typically unique, which implies that the composition or quality-mix of artworks observed will generally differ between periods (Hansen, 2009). These features make it difficult to compare prices over time and present challenges in measuring the state of the market over time.

The aim in Chapter 2 is to explore methods for constructing quality-adjusted South African art price indices, using a large database of auction prices. This is the first price index for South African art in the academic literature, of the type often estimated internationally (e.g. Mei and Moses (2002), Renneboog and Van Houtte (2002), and Kräussl and Lee (2010)). The indices reflect the mean of the distribution of the art price growth rates, while adjusting for compositional changes over time.

The chapter applies three methodologies for developing price indices for South African art: central tendency, hedonic and hybrid repeat sales methods. Simple central tendency indices are estimated as a baseline for comparing the results from the different methodologies. In the chapter it is argued that central tendency measures do not adequately control for compositional changes over time. Various indices are estimated with the hedonic regression method, which is able to control more adequately for quality-mix changes, by attributing implicit prices to a set of characteristics. The hedonic indices paint a consistent picture of the trend in South African art prices over time. A shortcoming of indices based on the hedonic method is that they may suffer from potential omitted variable bias.

The repeat sales method provides an alternative approach for estimating quality-adjusted price indices. Repeat sales indices suffer less from potential omitted variable bias, but have the shortcoming of potential sample selection bias. The repeat sales method controls for quality-mix changes by tracking the same asset over time. Hence, it utilises only artworks that have been traded more than once. The scarcity of repeat sales observations in the database limits the usefulness of the classical repeated sales approach in this case. The chapter proposes a new hybrid repeat sales method as an alternative, in the spirit of the pseudo-repeat sales method in Guo *et al.* (2014). This method addresses the problem of limited repeat sales observations and, to some extent, the potential omitted variable bias inherent in the hedonic method.

The internal validity of the indicators is assessed by comparing the indices calculated with the different methods. In this way, the indices estimated with the pseudo-repeat sales approach acts as a type of internal validity test, to check that the results are not driven by the inherent biases of a specific method. The indices estimated with the hedonic and pseudo-repeat sales methods point to the same general trend in South African art prices, with a large increase in the run-up to the Great Recession and a flat trend after 2009. This implies that potential omitted variable bias inherent

---

## Introduction

---

in the hedonic method, and the sample selection bias inherent in the repeat sales method, are not dictating the results (Calomiris and Pritchett, 2016). The indices estimated with the hedonic and pseudo-repeat sales methods differ markedly from the central tendency indices, which demonstrates the importance of adjusting for quality-mix changes when producing price indices. The indices for the different market segments indicate that the large price increases occurred especially in the more expensive or higher-end segments of the art market, and for oil paintings.

In the absence of another South African art price index for an external validity test, the indices are compared with the price indices of traditional South African assets (such as stocks and real estate). The art price indices exhibit a similar rapid appreciation between 2005 and 2008 to equity prices, and much higher returns than the other assets. The decrease in art prices during the Great Recession was as marked as the decrease in equity prices, and more so than the declines in the other asset prices. After 2009, art prices have not kept pace with the conventional asset prices.

The indices are subsequently evaluated directly in terms of smoothness, or signal-to-noise metrics, in order to assess which index provides the most accurate measure of South African art prices over time. The smoothness metrics suggest that the 1-year adjacent-period hedonic index is the most accurate, although all the versions of the hedonic and pseudo-repeat sales indices provide similar results.

In order to demonstrate how the estimated art price indices may be useful in exploring particular price patterns in the South African art market, this chapter studies the indices for evidence of a bubble in South African art prices. The chapter follows Kräussl, Lehnert and Martelin (2016) in using the bubble detection tests developed by Phillips, Wu and Yu (2011). The approach is based on a right-tailed augmented Dickey-Fuller unit root test, which can detect explosive behaviour directly in time series. This approach is used to test the hypothesis that South African art prices exhibited mildly explosive behaviour over the period.

The results indicate that there is evidence of bubble-like behaviour in all of the hedonic and pseudo-repeat sales art price indices. These indices provide reasonably consistent results in terms of the explosive periods in the South African art market, with a potential bubble most likely beginning in 2006 and ending in 2008. The bubble detection tests performed on the different market segments suggest that the bubble process occurred mainly for higher-end artworks, and for oil and watercolour paintings.

In the chapter it is argued that the hedonic and pseudo-repeat sales methods are useful aggregation techniques for unique items that are infrequently traded, where the quality-mix of items is not constant over time. These methods produce indices that improve upon the more common central tendency indices. The hedonic and pseudo-repeat sales methods may be useful in constructing indices for other unique assets, such as real estate, antiques, coins, and wine, where the quality-mix of items differs over time, and there is a lack of repeat sales. The quality-adjusted art price indices estimated in this chapter may be useful for future research on the South African art market.

## 1.2 Business Sentiment

The global financial crisis and subsequent Great Recession were associated with exceptionally low levels of confidence and heightened uncertainty (European Central Bank, 2013). The idea that weak business sentiment influenced economic activity has motivated a large body of literature investigating the impact of changes in business sentiment, and especially uncertainty, on investment and output decisions. To date there has been little research on business sentiment in South Africa. It is important to confirm the existence and nature of the relationship between sentiment and economic activity in developing countries, such as South Africa, which often face higher uncertainty than do developed countries (Bloom, 2014).

Business sentiment refers to two separate concepts, namely *confidence* and *uncertainty*. Business confidence refers to agents' perceptions of current and future business conditions (Mendicino and Punzi, 2013). Business uncertainty refers to their inability to forecast the probability of future events occurring (Knight, 1921). It is challenging to measure these concepts (Santero and Westerlund, 1996), as both are not directly observable and their definitions are difficult to operationalise. Although measuring business sentiment is not straightforward, survey-based indicators can be helpful in discovering agents' opinions on economic developments (Organisation for Economic Co-operation and Development, 2003). Survey-based measures have the advantage that they are derived from the opinions of key economic agents (Girardi and Reuter, 2017), are available earlier than official statistics, and are usually not subject to revision (European Central Bank, 2013). In South Africa, the BER conducts regular quarterly business tendency surveys. The challenge in aggregating the microeconomic data from the BER business tendency surveys lies with fully exploiting the disparate views of individual agents, in order to identify an underlying pattern of confidence and uncertainty.

The aim in Chapter 3 is to explore aggregation methods to develop proxies for business confidence and uncertainty in South Africa from these qualitative survey responses. The information in the qualitative survey responses is aggregated, by calculating the first and second weighted cross-sectional moments of the distribution. The composite indicators of business confidence and uncertainty are based on these moments. Two composite confidence indicators are calculated in the chapter. The first is the cross-sectional mean of responses to questions on current business conditions, and the second is the cross-sectional mean of responses to questions on expected future business conditions (Organisation for Economic Co-operation and Development, 2003). Three composite uncertainty indicators are calculated. The first is the scaled cross-sectional standard deviation of forward-looking responses (Girardi and Reuter, 2017). The second is the cross-sectional mean of individual firm forecast errors, and the third is the cross-sectional standard deviation of these forecast errors (Bachmann, Elstner and Sims, 2013; Arslan *et al.*, 2015).

As far as research on confidence is concerned, only two business confidence indicators are regularly published for South Africa: the South African Chamber of Commerce and Industry Business Confidence Index (SACCI BCI) and the BER Business Confidence Index (BER BCI). The SACCI

BCI is a composite measure of economic activity, rather than a confidence indicator in the way it is defined in the literature. The BER BCI is a measure of confidence derived from the BER's business tendency surveys (Kershoff, 2015). The BER BCI is based on a single question on current conditions. The survey responses are weighted in an ad hoc manner, and the services sector survey is excluded from the calculation. As far as research on uncertainty is concerned, only a few studies have created proxies for uncertainty in South Africa (e.g. Redl (2015) and Hlatshwayo and Saxegaard (2016)). The existing proxies capture different aspects of uncertainty, but none has fully exploited the information contained in the BER business tendency surveys to measure business uncertainty in South Africa.

The new composite indicators attempt to improve on the existing measures of sentiment for South Africa, by incorporating more of the information from the surveys than is currently used by the BER. The composite indicators incorporate the survey responses from questions on general business conditions, output, employment, orders placed and profitability. The responses are weighted to produce sectoral and aggregate indicators, which are analysed separately. The indices also incorporate the services sector, and a number of construction subsectors, which are not included in the BER BCI. In Chapter 2 various quality-adjusted measures of the mean of the distribution of growth rates in art prices are calculated. In Chapter 3 various measures of the mean and variance of the distribution of survey responses are calculated, as proxies for confidence and uncertainty.

The validity of the indicators is assessed by testing their conformity to events that were thought to have caused large changes in confidence and uncertainty in South Africa, as well as to existing sentiment indicators. The two confidence indicators are compared with the BER BCI and the SACCI BCI. The three composite uncertainty indicators are compared with a measure of financial market uncertainty (the SAVI) and the economic policy uncertainty indicator created by Hlatshwayo and Saxegaard (2016). In addition, a combined overall measure of uncertainty is constructed, which combines the survey-based uncertainty indicators with the measures of financial market and economic policy uncertainty.

The sentiment indicators are subsequently evaluated according to their correlation with real GDP growth, to assess whether they are able to improve upon the existing indicators. The new composite confidence indicators exhibit a significant positive correlation with real GDP growth and the measure based on current conditions exhibits a higher correlation with real GDP growth than existing measures. The leading indicator properties of the confidence indices are also evaluated, in terms of the timing of turning points and concordance with the official SARB business cycle. The composite confidence indicator exhibits a higher concordance statistic with the official business cycle than existing indicators. The new indicators therefore provide advance warning of turning points in the business cycle, which implies that they may be useful for monitoring economic developments in real time.

The composite uncertainty indicators generally exhibit a significant negative correlation with real GDP growth rates. The most plausible proxy for uncertainty is the combined indicator, as it seems

to reflect key economic events and exhibits a higher negative correlation with real GDP growth than do the individual indicators.

To demonstrate the usefulness of the aggregation methods and the estimated indicators, as well as provide an additional validity test, the chapter further examines the relationship between business sentiment and real economic activity in South Africa. There has been little analysis of this relationship in the South African context (e.g. Pellissier (2002) and Redl (2015)). The relationship between the indicators and real GDP growth is therefore investigated, including the timing of the relationship, and the extent to which correlation is conditional on other economic variables. The hypothesis that there was significant comovement between the sentiment indicators and real GDP growth is tested, using the standard VAR framework.

The results suggest that the survey-based confidence indicators contain relevant information for predicting output growth. The impulse-response analysis suggests that a positive shock to the confidence indicators is followed by a significant increase in real GDP growth. This is the case for the aggregate indicators as well as the sectoral indicators, implying that the confidence indicators are potentially useful for forecasting future economic activity. The impulse-response analysis shows that a positive shock to uncertainty is followed by a decrease in real GDP growth. The combined composite uncertainty indicator, which incorporates multiple sources of uncertainty, exhibits a stronger relationship with real GDP growth than any of the individual measures. The findings are robust to the inclusion of additional economic variables. Moreover, the impact of shocks to the sentiment indicators are larger on real production and investment growth than on real GDP growth, as theory predicts (Bloom, 2009).

In the chapter it is argued that the first and second weighted cross-sectional moments of the survey responses are useful in identifying an underlying pattern from the disparate views of individual agents. This approach produces composite indicators of business confidence and uncertainty that improve upon the existing measures for South Africa. The weighted cross-sectional moments employed in this chapter may be useful in other applications with qualitative survey responses, such as consumer surveys, where there are challenges in capturing the full richness in the data. The new sentiment indicators created in this chapter may be useful for a range of applications, including forecasting and nowcasting exercises, and as improved leading indicators of the business cycle.

### **1.3 Historical Commodity Prices**

Over the past two decades, the increase in access to online resources and inexpensive computing power have sparked a data revolution in African economic history (Fourie, 2016). Scholars have been able to capture and analyse historical statistics on a much larger scale than before. Historical statistics, and higher-frequency records in particular, often contain substantial gaps, which makes aggregation challenging (Klovland, 2014).



---

## Introduction

---

The aim in Chapter 4 is to explore aggregation methods for calculating a high-frequency commodity price indices, using incomplete historical records. The indices are based on two newly digitised sets of historical price records. The first is an expanded version of the historical dataset used in Boshoff and Fourie (2017). The records consist of monthly prices in various towns in the Cape Colony from 1889 to 1914, reported in the Agricultural Journals of the Department of Agriculture. The second dataset consists of annual market prices in various towns reported in the Cape Colony Blue Books. The challenge in aggregating the datasets is that both sets of records are incomplete, in terms of the coverage of both products and towns.

Klovland (2014) suggested that the repeat sales method, which is typically used to create indices for unique and infrequently traded items, such as artworks, may be used to aggregate incomplete data in this context. The chapter demonstrates that the repeat sales method provides a consistent way to aggregate the price data and produces indices with substantially fewer gaps than there are in the individual series. In Chapter 2, various quality-adjusted measures of the mean of the distribution of growth rates in art prices are estimated with the hedonic and pseudo-repeat sales methods. The classical repeat sales method used in Chapter 4 is related to those methods, as it also calculates the mean in the distribution of growth rates. The use of the repeat sales method in the context of estimating complete series from incomplete historical records is a novel application of this method, where each commodity price series (for a particular quality description) in each town is treated as a specific item to track over time.

Monthly wholesale commodity price indices are estimated for several individual products, which are intended to shed light on the demand and supply factors that influenced product prices. These indices are then aggregated to form a total commodity price index, with weights based on the production values reported in the 1904 census, and on import values reported in the Blue Books. The total commodity price index is intended to provide a clearer picture of the inflation history of the period. The price indices seem to correspond well with the economic history of the Cape Colony, with a large increase before and during the Second South African War (1899-1902) and a large decrease prior to unification (1910).

As an internal validity test, the repeat sales indices are compared to simple median indices of the prices in the Agricultural Journals. The external validity of the indices is assessed by testing their correlations with the annual consumer price indices calculated by De Zwart (2011). The price indices are then evaluated according to their conformity with the path implied by available monetary aggregates. The total commodity price index seems to better reflect the monetary history of the Cape Colony than existing price indices.

Higher frequency records, such as these, allow for a more detailed investigation of price histories, business cycles, and market behaviour in general (Mitchell, Solomou and Weale, 2012). In addition to providing a clearer picture of demand and supply conditions and the inflation history of the period, the high-frequency price records may be used to investigate market integration in the Cape Colony around the turn of the 20th century. During this period the mineral revolution led to the rapid

expansion of the railway network and created a potentially large internal market (Schumann, 1938; Herranz-Loncán and Fourie, 2017). There is a large international literature on market integration, but very few studies have considered internal market integration in the Cape Colony over this period (e.g. Boshoff and Fourie (2017)). This chapter therefore explores further methods to aggregate the high-frequency price records, in order to investigate internal market integration in the Cape Colony, which demonstrates the usefulness of the aggregation methods and the estimated time series.

Market integration concerns the convergence of prices for the same goods throughout the economy (Andrabi and Kuehlwein, 2010), as well as the efficiency with which price gaps return to equilibrium after a shock (Federico and Sharp, 2013). Higher-frequency data can aid the investigation along both dimensions (Ejrnæs, Persson and Rich, 2008). Two aggregation methods are used to examine these two dimensions of market integration. The first is to calculate measures of cross-sectional price dispersion between the towns in the Cape Colony, such as the coefficient of variation, which are used to investigate price convergence. They are based on the first and second cross-sectional moments of the distribution of commodity prices. These measures are related to the cross-sectional mean and standard deviation of responses and forecast errors calculated in Chapter 3. The second method is to calculate repeat sales commodity price indices at the regional level, which are used to investigate market efficiency.

The hypotheses that price dispersion between towns was declining over the period and that regional price indices were cointegrated are tested. According to these tests, many of the commodity prices between towns were converging over the period, and various regional commodity price indices were cointegrated, with the number of cointegrated series increasing in the latter part of the sample period. This implies that there was increasing market integration in the Cape Colony over the sample period.

In the chapter it is argued that the repeat sales approach is a useful technique for aggregating incomplete information from various sources. This approach produces commodity price indices that provide insight into the supply and demand conditions for individual products, as well as a proxy for inflation that is an improvement on the existing measures. The new commodity price indices estimated in this chapter may be useful for future research on this period of the Cape Colony's history. The repeat sales method employed in this chapter may be useful in a diverse range of applications where the information is incomplete. This is likely to be a more common problem in future as more historical records are digitised and become available for analysis.

## 1.4 Objectives and Contributions

This dissertation makes a contribution to the literature by demonstrating useful methods for overcoming specific challenges in aggregating relatively large microeconomic datasets to estimate macroeconomic time-series indicators. These aggregation challenges relate to (i) estimating quality-adjusted price indices for unique and infrequently traded items, (ii) developing aggregate measures

of sentiment based on the disparate qualitative survey responses of a large number of respondents, and (iii) estimating complete price indices from incomplete historical records. In some cases, these methods are applied to the South African setting for the first time.

These aggregation methods may prove useful in a variety of research settings that present similar challenges. The hedonic and hybrid repeat sales methods employed in this dissertation would be useful in creating quality-adjusted price indices for other unique assets, such as real estate, antiques and wine, where there are a lack of repeat sales and the quality-mix of items differs over time. International real estate price indices are often calculated with the repeat sales or hedonic methods. The repeat sales method, for instance, is used to calculate the S&P/Case-Shiller Home Price Indices in the US. In South Africa, ABSA, FNB and Standard Bank currently only use stratified central tendency methods to calculate their property price indices. As demonstrated below, the central tendency method does not adequately control for quality-mix changes. Using the repeat sales or hedonic methods would substantially improve upon real estate price indices in South Africa (Els and Von Fintel, 2010). Moreover, these techniques may be used to calculate the first quality-adjusted price indices for other unique items, such as wine and antiques, as more comprehensive microeconomic datasets become available.

The weighted cross-sectional moments of qualitative responses employed in this dissertation would be useful in other applications with qualitative survey responses, such as consumer surveys, where there are challenges in identifying an underlying trend from the disparate views of individual agents. Consumer confidence measures are popular indicators that are calculated all over the world, often using qualitative surveys. Popular international confidence indicators include the European Commission's Economic Sentiment Index and the University of Michigan Consumer Sentiment Index for the US. The BER calculates consumer confidence in South Africa using their consumer tendency surveys. Using the techniques demonstrated below, it would be possible to improve on the existing measures of consumer confidence, by capturing more of the richness in the data. Moreover, it would be feasible to create new measures of consumer uncertainty, by calculating the weighted cross-sectional dispersion of responses. These indicators may be useful for forecasting and as leading indicators of the cycles in real economic activity.

The repeat sales method employed in this dissertation may prove useful in a diverse range of applications where the information from various sources is incomplete. Historical data, and higher-frequency records in particular, often suffer from missing observations, which makes it challenging to form complete time series. The repeat sale method provides a simple and consistent way to overcome some of the problems with aggregating incomplete information from various sources. It may be useful to aggregate information on, for example, wages or incomes, especially when the information from various sources can be combined to create a complete series. These aggregation techniques may become increasingly useful as more historical information is digitised and becomes available for quantitative research.

The new indices estimated in this dissertation offer the following improvements: the art price

indices are the first quality-adjusted price indices for South African art; the business confidence and uncertainty indicators improve upon existing measures for South Africa, in the sense that they capture more of the information in the survey responses, and exhibit higher correlations with real GDP growth than existing indicators; the historical commodity price indices are the first monthly price indices for the Cape Colony over the sample period (1889-1914), and the total commodity price index provides an improved measure of inflation over that period, exhibiting a higher correlation with available monetary aggregates than existing indices. To aid further research in each field, the time-series indicators are reported in the chapter appendices.

Three questions are addressed in each of the chapters. The first relates to the methods that may be used to overcome the specific challenges in aggregating the microeconomic datasets. The second involves a description of the estimated time series, which helps to establish stylised facts in each of the three fields. The third involves testing a specific hypothesis in each case, which serves to demonstrate the usefulness of the aggregation methods and the estimated time series, and provides a further validity check of the indicators. The specific questions addressed in the dissertation are the following:

1. Art Prices:
  - 1.1. How can auction prices for unique and infrequently traded artworks be aggregated to construct quality-adjusted art price indices?
  - 1.2. What has happened to South African art prices over time?
  - 1.3. Is there evidence of a bubble in art prices over the sample period?
2. Business Sentiment:
  - 2.1. How can the disparate qualitative survey responses of a large number of respondents be aggregated to construct sentiment indicators?
  - 2.2. What has happened to business confidence and uncertainty in South Africa over time?
  - 2.3. Is there evidence of significant comovement between the sentiment indicators and real activity?
3. Historical Commodity Prices:
  - 3.1. How can incomplete price series for various towns from two sets of historical records be aggregated to construct complete monthly commodity price indices?
  - 3.2. What happened to commodity prices in the Cape Colony over time?
  - 3.3. Is there evidence of increasing market integration in the Cape Colony around the turn of the 20th century?

## 2 Methods for Estimating Quality-Adjusted Price Indices: An Application to South African Art Prices

### 2.1 Introduction

Contemporary African art has experienced a surge in popularity over the last few decades. The South African art market in particular has received a lot of attention, and has grown markedly over the last two decades, both in terms of the number of transactions and total turnover (Fedderke and Li, 2014). Artworks by South African artists have reached record prices at local and international auctions, both for the country's 'masters', including Irma Stern, JH Pierneef and Walter Battiss, and contemporary artists like William Kentridge (Naidoo, 2013). In 2011, Stern's *Two Arabs* was sold by Strauss & Co. for a hammer price of R19 million, a record for a South African auction. Also in 2011, Bonhams in London sold Irma Stern's *Arab Priest* for a hammer price of £2.7 million, a world record for a South African artwork at auction. The increase in interest in South African art, both locally and abroad, has sparked a vibrant market for collectors.

The increase in the popularity of South African art, at least partly as an investment vehicle, is commensurate with international trends, where fine art has become an important asset class in its own right. In 2010, around 6% of total wealth was held in passion investments, which include art, antiques, wines and jewellery (Renneboog and Spaenjers, 2015). In 2013, art made up around 17% of high net worth individuals' allocations to passion investments (Capgemini, 2013). Of all these passion investments, art is the most likely to be purchased for potential value appreciation (Capgemini, 2010).

To date there has been little research on the South African art market and particularly trends in art prices. This is due, at least in part, to a lack of data on art prices (Campbell, 2009). It is important to analyse price movements over time in order to understand the dynamics of the market and to be able to answer questions about developments in the market.

The aim in this chapter is to explore methods for constructing quality-adjusted South African art price indices. It can be challenging to estimate accurate price indices for unique items such as artworks (Jiang, Phillips and Yu, 2015). Artworks have a low transaction frequency, which means that only a small part of the overall market is traded at any given time, while the prices of non-transacted items are unobservable. Artworks are typically unique, or heterogeneous, which means that the quality of items sold is not constant over time. The composition of items sold will generally differ between periods, making it difficult to compare prices over time (Hansen, 2009). These features present challenges for the measurement of the state of the market over time, and necessitate a different approach than is used for indices of traditional assets.

In this chapter three methodologies are used to develop price indices for South African art: central tendency, hedonic and hybrid repeat sales methods. Simple central tendency indices are estimated

as a baseline to compare the results from the different methodologies. In this chapter it is argued that central tendency measures do not adequately control for quality-mix or compositional changes over time. Various indices are estimated with the hedonic regression method, which is able to control more adequately for quality-mix changes, by estimating implicit prices for a set of item attributes. A shortcoming of indices based on the hedonic method is that they may suffer from potential omitted variable bias.

The repeat sales method is an alternative estimation method for quality-adjusted price indices. Repeat sales indices suffer less from potential omitted variable bias, but have the shortcoming of potential sample selection bias. The repeat sales method controls for quality-mix changes, by tracking the same asset over time. Hence, only artworks that have sold more than once are utilised. The scarcity of repeat sales observations in the database limits the usefulness of the classical repeated sales approach in this case. In this chapter a simple hybrid repeat sales method is proposed for estimating alternative price indices for South African art. This approach addresses the problem of the scarcity of repeat sales observations and, to some extent, the potential omitted variable bias inherent in the hedonic method.

The internal validity of the indicators is assessed by comparing the indices calculated according to the different methodologies. In this way, the indices estimated with the hybrid repeat sales approach act as an internal validity test, to check that the results are not driven by the inherent biases of a specific method. The indices estimated with the hedonic and pseudo-repeat sales methods point to the same general trend in South African art prices, with a large increase in the run-up to the Great Recession and a flat trend after 2009. The indices for the different market segments indicate that the large price increases occurred in especially for more expensive artworks and for oil paintings. As there is no other South African art price index for an external validity test, the indices are compared to the price indices of traditional South African assets. The indices are then evaluated directly in terms of smoothness or signal-to-noise metrics, in order to assess which index provides the most accurate measure of South African art prices over time.

In order to demonstrate how the estimated art price indices may be useful in exploring particular price patterns in the South African art market, this chapter studies the indices for evidence of a bubble in South African art prices between 2000 and 2015. During this period many commentators claimed that the market was overheating and suggested the possibility of a bubble in the market (e.g. Rabe (2011); Hundt (2010); Curnow (2010)). A reduced-form bubble detection method is used to test for periods of mildly explosive behaviour in the art price indices. The evidence points to consistent evidence of an explosive period between 2006 and 2008. The bubble detection tests performed on the different market segments indicate that the bubble process occurred especially in the higher-end, oil and watercolour segments of the market.

## 2.2 Methodologies for Constructing Art Price Indices

The recent increase in the availability of data on art prices has increased the interest in art as an asset class (Campbell, 2009). A large number of academic studies have constructed art price indices for various art markets around the world (e.g. Anderson (1974); Renneboog and Van Houtte (2002); and Kräussl (2015)). These studies have relied almost exclusively on publicly available auction prices, and have typically been interested in evaluating the risk-adjusted returns to art, in order to investigate whether they provide potential diversification benefits for an investment portfolio.

The construction of price indices for unique assets is challenging for at least two reasons (Jiang, Phillips and Yu, 2015). Firstly, the low frequency of trading means that only a subset of the market is traded at a given time, while the prices of non-transacted items are unobservable. Secondly, the heterogeneity of these items means that the quality of assets sold is not constant over time. Thus, the composition of items sold will generally differ between periods, making it difficult to compare prices over time (Hansen, 2009). Constructing an index for unique items, like artworks, therefore requires a different approach than for indices of traditional assets such as stocks and bonds. Four broad measurement techniques have been used to construct these indices (Eurostat, 2013):

- a) Central tendency methods
- b) Hedonic methods
- c) Repeat sales methods
- d) Hybrid methods

The following sections provide a brief introduction to these methodologies. The literature does not provide an a priori indication of the most appropriate method and, in practice, the data dictates the choice of method.

### 2.2.1 Central Tendency Methods

The simplest method for constructing a price index is to calculate a measure of the central tendency of the price distribution. As price distributions are generally positively skewed, the mean of the prices in logs or the median of the prices in levels are often calculated as measures of the central tendency. This may be due to the zero lower bound on transaction prices, positively skewed income distributions, and the unique nature of these assets (Hansen, 2009). These average measures have the advantages of being simple to construct and not requiring detailed data.

Despite its advantages, an index based on average prices does not account for the difficulties mentioned above. For assets such as artworks, central tendency indices may depend more on the composition, or quality-mix, of assets sold than on changes in the underlying market. For instance, if there is an increase in the share of higher quality assets, a measure based on averages will show an increase in prices, even if the market prices remained constant. Hence, such a measure may not be representative of the price movements of all the assets in the market. If there is a



correlation between compositional changes and turning points in asset price cycles, the average could be especially inaccurate (Hansen, 2009).

Stratified central tendency measures can control, to some extent, for compositional changes in assets sold over time, by dividing the sample into subgroups according to item attributes such as artist and medium. The central tendency for each subgroup is calculated, and the aggregate quality-adjusted index is then calculated as a weighted average of the indices for the subgroups. The Fisher index, which is the geometric mean of the Laspeyres and Paasche indices<sup>2</sup>, is often recommended (Eurostat, 2013).

Stratified central tendency methods are currently used by several South African financial institutions to construct property price indices for South Africa, based on the finance applications they receive. ABSA and FNB publish mean property price indices, while Standard Bank publishes a median property price index.<sup>3</sup> The ABSA House Price Index, for instance, is based on the mean sales prices of properties categorised by house size and price segment (Aye *et al.*, 2014).

However, scholarly work rarely employs stratified central tendency indices, as these measures adjust only for the variation in the composition of assets across the subgroups. The ABSA House Price Index, for instance, does not control for changes in the composition of properties unrelated to size and price segment. The number of subgroups may be increased to reduce the quality-mix problem, if the data permits this, although some compositional changes will likely remain (Hansen, 2009). However, this will reduce the average number of observations per subgroup and raise the standard error of the overall index (Eurostat, 2013). If the subgroups become very small, small changes can have a large impact on the index. As a consequence of these difficulties, the repeat sales and hedonic methods have dominated the international literature, especially with regard to art price indices.

### 2.2.2 The Hedonic Method

The hedonic method is based on hedonic price theory, which is useful for the analysis of differentiated good pricing (Griliches, 1961). The hedonic method is derived from the microeconomic theory of implicit prices, which supposes that utility is derived from the characteristics or attributes of goods (Lancaster, 1966). Each good  $i$  is described by a vector  $x$  of  $J$  quantifiable and inseparable attributes that determines its price:  $x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{iJ})$ . In the context of art, attributes may include physical (e.g. medium) and non-physical attributes (e.g. artist reputation). According to this theory, goods offer buyers distinct packages of attributes. When consumers purchase a particular good  $i$ , they have chosen a particular vector  $x$  of attributes (Rosen, 1974).

<sup>2</sup>The Laspeyres price index  $P_L$  compares prices in the comparison period  $P^1$  to prices in the base period  $P^0$  for each stratum  $m$ , while holding quantity weights  $Q$  fixed in the base period 0:  $P_L = \sum_{m=1}^M P_m^1 Q_m^0 / \sum_{m=1}^M P_m^0 Q_m^0$ . The Paasche price index  $P_P$  compares prices in the comparison period  $P^1$  to prices in the base period  $P^0$  for each stratum  $m$ , while holding the quantity weights  $Q$  fixed in the comparison period 1:  $P_P = \sum_{m=1}^M P_m^1 Q_m^1 / \sum_{m=1}^M P_m^0 Q_m^1$ .

<sup>3</sup><http://propertywheel.co.za/wp-content/uploads/2016/07/HPI-Jun-2016.pdf>;

<https://blog.fnb.co.za/2016/07/fnb-property-barometer-house-price-index-jun-2016/>;

<http://propertywheel.co.za/wp-content/uploads/2016/05/April-House-Price-Index-Standard-Bank-1.pdf>



The price of the good is determined by the specific combination of attributes. The price of good  $p_i$  is a function of its attributes  $x$ :  $p_i = p(x_i) = p(x_{i1}, x_{i2}, x_{i3}, \dots, x_{iJ})$ . The hedonic price function  $p(x_i)$  specifies how the market price of a good varies as its attributes vary (Epple, 1987).

Rosen (1974) offers a theoretical framework in which  $p(x)$  emerges from the interaction between buyers and sellers. Buyers and sellers base their locational and quantity decisions on maximising behaviour and are in equilibrium along the hedonic price function. The solution to the maximisation problem produces a set of implicit (or shadow) prices for the attributes (Anderson, 1974).

The implicit prices for each attribute  $j$  of good  $i$  may be represented as  $p_j(x_i) = \frac{\partial p}{\partial x_j}$ . This  $p_j$  is considered an implicit price, as there is no direct market for the attributes, and their prices are not independently observed. One could infer that this price represents the value added to a good for a unit increase of a given attribute. The demand and supply for the goods implicitly determine the marginal contributions of the attributes to the prices of the goods. Implicit prices are revealed to agents from the observed prices of differentiated goods and the attributes associated with them (Eurostat, 2013).

The hedonic method controls for compositional changes by assigning implicit prices to a set of value-adding attributes of an individual item. Thus, the hedonic approach can circumvent the problems of changes in composition or quality-mix over time (Hansen, 2009).<sup>4</sup> Hedonic regressions control for the observable attributes of an item to produce a price index for the ‘standard asset’ (Renneboog and Van Houtte, 2002). The approach entails regressing the logarithm of the sales price on the relevant attributes. The standard hedonic model usually takes the following form:

$$\ln P_{it} = \sum_{t=1}^T \delta_t D_{it} + \sum_{j=1}^J \beta_{jt} X_{jit} + \sum_{k=1}^K \gamma_{kt} Z_{kit} + \epsilon_{it},$$

where  $P_{it}$  represents the price of item  $i$  at time  $t$  ( $t = 1, \dots, T$ );  $D_{it}$  is a time dummy variable taking the value of 1 if item  $i$  is sold in period  $t$  and 0 otherwise;  $X_{jit}$  is a set of  $j$  ( $j = 1, \dots, J$ ) observed attributes of item  $i$  at time  $t$ ;  $Z_{kit}$  is a set of  $k$  ( $k = 1, \dots, K$ ) unobserved attributes that also influence the price; and  $\epsilon_{it}$  is a random (white noise) error term.

The coefficients of the time dummies  $\delta_t$  provide an estimate of the average increase in prices between periods, holding the change in any of the measured quality dimensions constant (Griliches, 1961). In other words, they capture the ‘pure price effect’ (Kräussl and Lee, 2010). The price index is then derived from the series of estimated coefficients:  $\hat{\delta}_1, \dots, \hat{\delta}_T$ .

The most common form of the hedonic equation assumes that the implicit prices (i.e. the coefficients  $\beta_t$  and  $\gamma_t$ ) are constant over the entire sample. However, when demand and supply conditions (e.g. tastes) change, the implicit prices of the attributes may change (Renneboog and Spaenjers,

<sup>4</sup>According to Hansen (2009), there are various weighting approaches. An equal weighting of art price inflation rates is appropriate when measuring the price changes of a representative artwork. A higher weight should be given to the price changes of higher-value artworks when measuring changes in the value of the art stock (or a representative portfolio). This chapter focuses on the pure price changes for a representative artwork, assuming an equal weighting.

2013). Another problem with this multi-period pooled model is that the coefficients are not stable when data from additional periods are added to the sample. An adjacent-periods or chained regression can allow for shifts in parameters (Triplett, 2004). Separate regressions are estimated for adjacent time periods and the sequence of shorter indices are then chain-linked together to form the continuous overall index (McMillen, 2012). The coefficients, and therefore the implicit prices assigned to the attributes, are allowed to vary in each regression (Triplett, 2004).

### **2.2.2.1 The hedonic method applied to art prices**

The majority of studies of art price indices have used hedonic models to construct the indices, due to the lack of repeat sales of artworks and the availability of information on many of their important attributes. Anderson (1974) first applied the hedonic method to art prices. More recent examples include Renneboog and Van Houtte (2002), who estimated an index for Belgian paintings; Kräussl and Lee (2010), who studied the prices of the top 500 artists in the world; Kräussl and Logher (2010), who analysed the performance of art in Russia, China and India; and Kräussl (2015) who analysed art from the Middle East and Northern Africa region.

In estimating art price indices, studies typically set up some form of selection criteria for which artists to include in the index calculation. A common criterion has been historical importance, measured as the frequency with which an artist has been mentioned in a collection of art literature. Kräussl and Van Elsland (2008) argued that availability and liquidity are better criteria from an investor's point of view, as the index will reflect artworks actually traded in the market. This implies that selection could be based on the number of sales, rather than historic relevance. Kräussl and Van Elsland (2008) developed a two-step hedonic approach, which allows the use of every auction record, instead of only those auction records that belong to a subsample of selected artists. This approach is discussed in more detail below.

The choice of the attributes in a hedonic regression is limited by data availability and involves subjective judgement. Hedonic models typically include attributes that are easily observable and quantifiable. The attributes include the artist, the auction house, the size, the medium, the theme, whether the artwork is signed, and the artist's living status (Kräussl and Logher, 2010).

If the functional form is misspecified or the omitted variables are correlated with sales timing, it will result in misspecification or omitted variable bias, which will bias the parameter estimates and therefore the indices (Jiang, Phillips and Yu, 2015). The primary difficulty with hedonic price indices is this potential omitted variable bias.<sup>5</sup> Nevertheless, the hedonic method may be especially appropriate for luxury consumption goods, where the willingness to pay for an item is

---

<sup>5</sup>According to Triplett (2004), even if the hedonic coefficients are biased, it is not necessarily the case that the hedonic index will be biased. If the cross-sectional correlations between included and omitted attributes are the same as the time-series correlations, the hedonic index may be unbiased, even though the hedonic coefficients are biased. Changes in omitted attributes between two periods may offset the error in estimating the implicit prices of included variables. The bias therefore becomes an empirical matter, as the effect on the price index is important, not just the effect on the hedonic coefficients.

often determined by a few key attributes. Relatively detailed data is available for art, which should capture a large part of the variation in sales prices. Omitted variable bias should therefore be less of a problem than for other unique assets such as real estate, and the omitted variable bias is often small in practice (Triplett, 2004; Renneboog and Spaenjers, 2013).

Bought-in lots (i.e. items that do not reach the reserve price and remain unsold) are always a problem when estimating these indices. Most studies lack data on buy-ins and are forced to ignore the problem. Collins, Scorcu and Zanola (2009) developed a hedonic index that corrected for sample selection bias from buy-ins. They argued that because auctions have high proportions of unsold lots (typically 30%-40%), price indices suffer from non-randomness in the data. Including only items that systematically exclude ‘less fashionable’ artworks potentially introduces bias in the sample. They used a Heckman selection model to address this issue. The results confirm a statistically significant sample selection problem, in line with similar studies on the property market.

### 2.2.3 The Repeat Sales Method

The repeat sales method provides an alternative approach for estimating quality-adjusted price indices, based on price changes in items sold more than once. It was initially proposed by Bailey, Muth and Nourse (1963) to calculate house price changes, was subsequently extended by Case and Shiller (1987), and is currently used to produce the S&P/Case-Shiller Home Price Indices in the US. Mei and Moses (2002) constructed the most influential repeat sales art price index.

The repeat sales method tracks the sale of the same item over time. It aggregates sales pairs and estimates the average growth rate in the distribution for each period (Kräussl and Lee, 2010). As a result, it does not require a measure of quality, only that the quality of each item be constant over time (Case and Shiller, 1987). Advocates of the repeat sales method argue that it controls more accurately for the attributes of goods, as well as for potential omitted variables (Jiang, Phillips and Yu, 2015).

The repeat sales model can be derived from the hedonic model if the hedonic model is differenced with respect to consecutive sales of items that have sold more than once (McMillen, 2012).<sup>6</sup> The standard model may be formulated as the change in the log of the sales price of item  $i$  that sold at time  $t$  and at an earlier time  $s$ :

$$\ln P_{it} - \ln P_{is} = \left( \sum_{t=1}^T \delta_t D_{it} - \sum_{s=1}^T \delta_s D_{is} \right) + \left( \sum_{j=1}^J \beta_{jt} X_{jit} - \sum_{j=1}^J \beta_{js} X_{jis} \right) + \left( \sum_{k=1}^K \gamma_{kt} Z_{kit} - \sum_{k=1}^K \gamma_{ks} Z_{kis} \right) + (\epsilon_{it} - \epsilon_{is})$$

If the attributes ( $X$  and  $Z$ ) of item  $i$  and the implicit prices ( $\beta$  and  $\gamma$ ) are constant between sales,

<sup>6</sup>While the repeat sales model can be derived as the differenced hedonic model, it can also stand on its own (Guo et al. 2014). Baily et al. (1963) saw their procedure as a generalisation of the chained-matched methodology, used previously in constructing real estate price indices.

the equation reduces to the standard estimating equation:

$$\ln \frac{P_{it}}{P_{is}} = \sum_{t=1}^T \delta_t G_{it} + u_{it},$$

where  $P_{it}$  is the purchase price for item  $i$  in time  $t$ ;  $\delta_t$  is the parameter to be estimated for time  $t$ ;  $G_{it}$  represents a time dummy equal to 1 in period  $t$  when the resale occurs, -1 in period  $s$  when the previous sale occurs, and 0 otherwise; and  $u_{it}$  is a white noise residual.

Thus, in the standard repeat sales model, the dependent variable is regressed on a set of dummy variables corresponding to time periods. The coefficients are estimated only on the basis of changes in asset prices over time. Again, the price index is derived from the series of estimated coefficients:  $\hat{\delta}_1, \dots, \hat{\delta}_T$ .

This estimating equation provides unbiased estimates of pure time effects without having to correctly specify the item attributes  $X$  or the functional form of the hedonic equation (Deng, McMillen and Sing, 2012). By differencing the hedonic equation, it also potentially controls for the omitted variables  $Z$ , as these are eliminated from the equation. Furthermore, it has the advantage of not being data intensive, as the only information required to estimate the index is the price, the sales date and a unique identifier (e.g. the address of the property). The repeat sales method has often been applied in creating real estate price indices (e.g. Bailey, Muth and Nourse (1963), Case and Shiller (1987), Hansen (2009), and Shimizu, Nishimura and Watanabe (2010)) where there is a lack of detailed information on each sale.

A disadvantage of the repeat sales method is the possibility of sample selection bias.<sup>7</sup> Items that have traded more than once may not be representative of the entire market. For example, if cheaper artworks sell more frequently than expensive artworks, but high-quality artworks appreciate faster, a repeat sales index will tend to have a downward bias (Eurostat, 2013). The size and direction of the bias will vary by the sample under investigation. The biggest problem with the repeat sales method in the current context is that single-sale data is discarded. This is problematic because the resale of a specific artwork may occur only infrequently, which might be related to the high transaction costs involved. This substantially reduces the total number of observations available.

### 2.2.3.1 The repeat sales method applied to art prices

A few studies have utilised the repeat sales method to estimate art price indices. These studies have typically relied on very large sales databases, due to the infrequency of repeat sales of individual artworks. Mei and Moses (2002) constructed the seminal repeat sales index of art prices for the period 1875-2000. Their methodology is currently used to produce the Mei Moses Art Index for Beautiful Asset Advisors. Goetzmann, Renneboog and Spaenjers (2011) used a long-term repeat

<sup>7</sup>The nature of sample selection bias is different in the hedonic and repeat sales approaches. The repeat sales method ignores single-sales observations, so that it may not be representative of the population. The hedonic method uses only sold items, so a bias may arise from unsold items.

sales art price index to investigate the impact of equity markets and top incomes on art prices. The analysis was based on over a million sales dating back to the 18th century.

Korteweg (2013) constructed a repeat sales index based on a large database of repeat sales between 1972 and 2010. He argued that standard repeat sale indices suffer from a sample selection problem, as sales are endogenously related to asset performance. If artworks with higher price increases were more likely to trade, the index would not be representative of the entire market. In periods with few sales, it would be possible to observe large positive returns, even if overall values were declining. A Heckman selection model, predicting whether an artwork actually sold, was used to correct for this bias. The correction decreased the returns to investments in art.

#### 2.2.4 Hybrid Methods

The major problem with the hedonic method is the potential for omitted variable bias, while the biggest problems associated with the repeat sales method are that it suffers from potential sample selection bias and that it discards single-sale observations. A number of hybrid models, which involve a combination of the two methods, have been proposed to address these problems. A combination of the two methods might lead to a quality-adjusted index that exploits all the sales data and reduces both sample selection and omitted variable bias (Jiang, Phillips and Yu, 2015).

In the context of real estate, for instance, Case and Quigley (1991) used single-sale and repeat sale properties to jointly estimate price indices using generalised least squares regressions. More recently, Nagaraja, Brown and Zhao (2011) suggested a model consisting of a fixed time effect and a random postal code effect, combined with an autoregressive component. The index was a weighted average of estimates from single-sale and repeat sales prices, with the repeat sales prices having a higher weight.

An interesting perspective, which is relevant to this chapter, is to view the repeat sales specification as an extreme solution to a matching problem. This is because the repeat sales approach requires an exact match to estimate the index. For example, the same Van Gogh *Wheat Field with Crows* is tracked over time to control for all the observable and unobservable attributes. The idea behind the imperfect matching method proposed by McMillen (2012) is that some items may be similar enough to control for many of the differences in (observable and unobservable) attributes. For example, Van Gogh's well-known *Sunflowers* series, of which there are five versions, might be similar enough to be treated as repeat sales. The objective is to match sales observations over time, by some criterion, to cancel out as many as possible of the differences in attributes (Guo *et al.*, 2014).

McMillen (2012) used a matching criterion based on propensity score matching. Each property sold in the base period was matched with one property sold in each subsequent period, based on propensity scores from a logit model. The procedure is a data preprocessing one that builds an estimation sample.

Guo *et al.* (2014) proposed a pseudo-repeat sales (ps-RS) method to estimate price indices for newly constructed homes in China. The ps-RS procedure was developed to deal with two features in the Chinese urban residential market. Firstly, new home sales accounted for a large share of total sales (87% in 2010). As a consequence, there were a limited number of repeat sales, similar to the South African art market. Secondly, new housing developments often included similar units within a residential complex. The idea was to match similar homes within each complex or building in order to construct a large pseudo-repeat sales sample.

As a matching criterion, Guo *et al.* (2014) used a distance metric to identify similar transactions across adjacent periods. The distance metric between two sales was defined as the absolute value of the difference between the two predicted prices from a hedonic equation, excluding time dummies (i.e. the non-temporal component). Pairs with a distance metric smaller than a certain threshold were selected as pseudo-pairs. The threshold was a trade-off between the within-pair homogeneity and sample size, and was set flexibly.

All pseudo-pairs were then used in a ps-RS regression model. The ps-RS model is similar to the differential hedonic regression above. For a given building  $b$ , home  $i$  in quarter  $t$  and home  $h$  in quarter  $s$  are adjacent transactions ( $t > s$ ), and make a matched pair:

$$\ln P_{itb} - \ln P_{hsb} = \sum_{j=1}^J \beta_j (X_{itbj} - X_{hsbj}) + \sum_{t=0}^T \delta_t G_{it} + \epsilon_{ithsb},$$

where  $G_{it}$  is again a time dummy equal to 1 if the later sale in the pair occurred in quarter  $t$ , -1 if the former sale in the pair occurred in quarter  $s$ , and 0 otherwise; and  $\epsilon_{ithsb}$  again represents a white noise residual.

Taking within-pair first differences cancels out any attributes that are the same between the two units, including both observable (e.g. number of bedrooms) and unobservable attributes (e.g. locational or neighbourhood effects). Only attributes that differ between the two units within a pair will be left as independent variables, in differenced form, reflecting the hybrid specification. The ps-RS indices were derived from the coefficients of the time dummies. Guo *et al.* (2014) found that the ps-RS method controlled more adequately for quality-mix differences over time and produced a smoother index, indicating less random estimation error.

The approach is a hybrid model that mitigates the problem of potential omitted variable bias with the hedonic method, by taking first differences between similar items. It mitigates the problems of small sample sizes and sample selection bias with repeat sales methods by using more of the observations (McMillen, 2012).

Calomiris and Pritchett (2016) used a similar procedure, based on the differential hedonic equation, in analysing slave price indices. They argued that while their hedonic model controlled for observable slave attributes, it may be sensitive to the presence of unobservable attributes. They created a matched sample that enabled the estimation of a repeat sales model for the changes in slave

prices. Because they observed repeated sales of the same slave, the unobserved attributes would be similar for both transactions. They also allowed for the possibility that the observable attributes had changed between the date of initial purchase and the subsequent sale. A differential hedonic regression was used to estimate the hybrid repeat sales model, which eliminated the time-invariant and unobserved effects. Calomiris and Pritchett (2016) found that their hybrid repeat sales index was similar to the hedonic price index, but with greater volatility. They argued that the similarities between the indices provided confidence that temporal variation in unobservable characteristics was not dictating the results.

As previously mentioned, there is no consensus regarding the preferred approach to constructing quality-adjusted price indices, either theoretically or empirically. However, there is reason to believe that more advanced measures may provide a better guide to pure price changes than simple central tendency methods (Hansen, 2009). The specific methodology adopted is dependent on the data available. Art price indices tend to employ some variant of the hedonic method, due to the availability of more detailed data on artwork characteristics and a lack of repeat sales.

### **2.2.5 South African Art Price Literature**

Few studies have investigated South African art prices. In an important contribution, Fedderke and Li (2014) studied the relationship between South Africa's two major fine art auction houses: Strauss & Co and Stephan Welz & Co. The analysis was based on a hand-coded dataset of auction prices. They developed a theoretical framework to consider the interaction between the market leader (Strauss) and the market follower (Stephan Welz). The model predicted that the market follower would be forced to issue excessive price estimates to attract sellers, at the cost of higher buy-in rates. The predictions were tested by employing a hedonic model to construct a counterfactual for auction prices. Both direct and indirect tests confirmed the predictions of the theoretical model.

Olckers, Kannemeyer and Stevenson (2015) related South African art auction prices (i.e. the economic value of art) to the cultural value of South African art. Art auction results (1996-2012) were obtained from AuctionVault's online database. An Art Critic Index was created as a proxy for cultural value, based on a survey of the South African art literature. Using a hedonic model, they found that the cultural value of art was positively correlated with economic value. Interestingly, they singled out and analysed some artists who were outliers in this relation.

Citadel, a wealth manager, has been publishing the Citadel Art Price Index (CAPI) since 2011. The CAPI is intended to outline general trends in the South African art market. It uses an adjacent-period hedonic regression model, based on the top 100 artists in terms of sales volumes, and a 5-year rolling window estimation period (Econex, 2012).<sup>8</sup>

Botha, Snowball and Scott (2016) used the CAPI to test the potential for art to be used to diversify

---

<sup>8</sup>The CAPI was estimated by the author on behalf of Citadel.



investment portfolios in the South African context. This might result if the art market exhibits different risk-return characteristics to conventional assets. To test this proposition, they used a VAR model, including the CAPI, the JSE All Share Index, the All Bond Index, and the ABSA House Price Index. They found that a positive shock to stock market returns was followed by a significant positive response in the CAPI in the following quarter. They concluded that South African art, as measured by the CAPI, would not aid portfolio diversification.

The price indices estimated in this chapter build on the methodology of the CAPI. They are the first price indices for South African art in the academic literature, of the type often estimated internationally (e.g. Mei and Moses (2002), Renneboog and Van Houtte (2002), and Kräussl and Lee (2010)). The indices may be useful for an improved understanding of developments in the South African art market. By estimating more accurate measures of average price trends, it is possible to investigate particular price trends in the South African art market, such as bubble behaviour and potential portfolio diversification benefits.

## **2.3 Data: South African Art Auction Prices**

In this chapter, quarterly price indices are estimated for South African art from 2000Q1 to 2015Q4. The art price indices are based on the database of auction prices recorded by AuctionVault. The following section presents a brief discussion of public and private art market prices. The subsequent sections present the auction data and the artwork characteristics used to estimate the South African art price indices.

### **2.3.1 Public and Private Art Market Prices**

The literature on estimating art price indices has relied almost exclusively on publicly available auction prices.<sup>9</sup> Art is also sold privately, either directly by artists or through dealers. However, dealers' sales records are generally not available. Releasing such information may be damaging to dealers' businesses, and they have an incentive to give the impression that there is high demand for their artworks (Olckers, Kannemeyer and Stevenson, 2015).

Nevertheless, it is generally accepted that auction prices provide a benchmark that is used in the private market (Renneboog and Spaenjers, 2013). Anecdotal evidence suggests that private dealers are very aware of auction prices and follow them closely. Differences between auctions and private markets in terms of institutional arrangements, transaction costs, and available information might lead to different price levels for the same or similar artworks. However, in constructing price indices, the focus is not on the price levels of individual artworks, but on the trend in art prices over time.

---

<sup>9</sup>Auctions account for around half of the art market according to The European Fine Art Fair Art Market Report 2014.



Auction prices and private prices are likely to be correlated over time for similar artworks, even if their levels are different (Olckers, Kannemeyer and Stevenson, 2015). This is due to demand-side and supply-side substitution. The intuition is that there is a limit to how far prices can deviate before either demand-side or supply-side substitution forces them back in line (Stigler and Sherwin, 1985). Products can compete despite being priced at different absolute levels, as demand-side substitution depends on the willingness of marginal consumers to switch from one product to another as relative prices change (Davis and Garcés, 2010). If auction markets and private markets are substitutes to a certain degree, their prices should be correlated over time, even if the price levels are different.

Where the auction houses sell a particular artwork, buyers may or may not have similar artworks available as substitutes in private markets (e.g. from dealers or the artist's studio). If similar artworks are available (e.g. artworks by the same artist in the same medium or prints from the same series), to some extent there are substitutes available for buyers. Buyers would then be able to purchase the similar artwork at auction or privately. Dealers, for instance, cannot charge exorbitant prices over time if similar works are selling for much less at auctions. The substitutes do not have to be perfect, nor do the prices have to be identical (Hoehn *et al.*, 1999). All that is necessary is that they provide a competitive constraint. From the seller's perspective, an artwork sold at auction may also be sold privately. If auction houses charge a very high commission or fail to attract the desired hammer prices, sellers can substitute towards private markets.

In this way, auction markets and private markets may constrain each other, based on imperfect demand-side and supply-side substitution. If this is the case, auction and private market prices should be correlated over time. The substitution and price convergence may not necessarily occur immediately, depending on market efficiency and the diffusion of information, but over the longer term, prices in these markets should move together. A caveat is that if private markets have less access to rare or high-end artworks than auction markets, and if high-end artworks are the drivers of art price inflation, prices in the two markets are not necessarily correlated over time.

There is some empirical evidence that auction prices and private prices are correlated. Candela and Scorcu (2001) is one of the few studies to have access to private art market information. They used the prices from a leading gallery (Prandi) in Italy to construct price indices for private sales for the period 1977-1998. The indices were then compared with price indices based on auction prices. They found a high correlation between the indices, as well as evidence of a cointegrating relationship. They argued that this was due to substitutability between private and auctions markets, arbitrage by dealers between the markets, and the existence of common fundamentals. They also found that auction prices Granger-caused private market prices, while the converse was not true. Their results suggest that auction prices represent a benchmark for private market prices.

### 2.3.2 South African Art Auction Prices

Auction prices are the only consistently available price data for the South African art market. This chapter therefore relies on publicly available auction prices, similar to virtually all other studies estimating art price indices. As explained above, private sales prices are likely anchored by auction prices and are likely to be highly correlated over time for similar artworks, even if their levels are different (Olckers, Kannemeyer and Stevenson, 2015).

Strauss & Co and Stephan Welz & Co are the two local auction houses that have handled the bulk of sales in recent years, with auctions in Cape Town and Johannesburg. Other local auction houses include Bernardi in Pretoria and Russell Kaplan in Johannesburg. Bonhams in London is currently the only major international auction house with a dedicated South African art department, although competition is emerging from Christie's and Sotheby's. Bonhams has two major South African art sales annually. The auction houses follow an open ascending auction, where the winner pays the highest bid. A sale is made only if the hammer price is above the secret reserve price; otherwise, the artwork is unsold and is said to be bought in (Fedderke and Li, 2014).

The indices are based on data recorded by AuctionVault. This data covers sales of South African art at eight auction houses<sup>10</sup> from 2000 to 2015. The database includes 52,059 sales by 4,123 different artists. The following characteristics are available for each auction record: auction house; date of auction; artist's name; title of work; medium; size; whether the artwork is signed, dated and titled; hammer price; and the number of distinct works in the lot. Like most studies, the database lacks information on buy-ins, and therefore the analysis is forced to disregard the potential sample selection problem.

Figure 2.1 illustrates the number of auction lots sold in the sample over the period (2000Q1-2015Q4) by auction house. The number of sales in the sample increased markedly over the period, especially in 2007 and 2011. This increase was due to an improvement in data collection from existing auction houses and the entry of auction houses such as Strauss & Co and Bonhams. These two auction houses now account for the bulk of turnover in the market. Total auction turnover echoed the increase in the number of lots over the period. At its peak in 2011, annual turnover in the sample had reached almost R400 million.

Figure 2.2 illustrates boxplots for the logarithm of the hammer prices for each year.<sup>11</sup> The sample is highly positively skewed, with the overall mean price of R49,824 being much higher than the median of R7,000. There were a number of outliers, including the hammer price of over R30 million for

---

<sup>10</sup>These are: 5th Avenue, Ashbeys, Bernardi, Bonhams, Christies, Russell Kaplan, Stephan Welz & Co and Strauss & Co.

<sup>11</sup>The boxplot displays the distribution of prices in each year. The central rectangle spans the first to the third quartile (i.e. the interquartile range or IQR). The middle line represents the median and the 'whiskers' above and below the box denote the minimum and maximum. Outliers are displayed separately and are defined as prices more than 1.5 times the IQR below the first quartile and above the third quartile. If outliers are present, the whiskers are taken to 1.5 times the IQR from the first and third quartile.

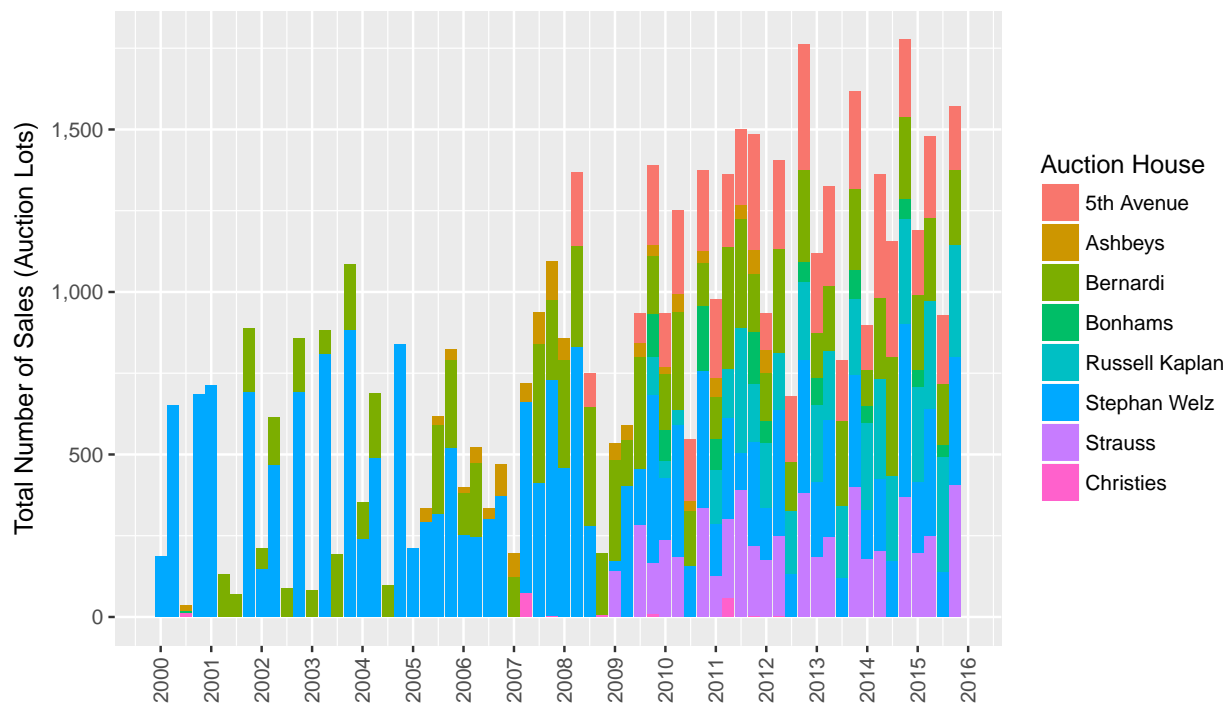


Figure 2.1: Total number of auction lots sold, by auction house (2000Q1-2015Q4)

Irma Stern's *Arab Priest* in 2011. Annual median sales prices increased substantially from R3,200 in 2003 to R10,000 at their peak in 2010.

### 2.3.3 Artwork Characteristics

Hedonic art price models typically include characteristics that are easily observable and quantifiable. This section briefly discusses the variables typically included in hedonic models of art prices.

*Artist reputation:* Hedonic models typically include dummy variables to control for the artists. However, often some artists have to be excluded from the estimation, due to a lack of degrees of freedom. Alternatively, a reputation variable can be constructed, either from the art literature, or from the auction data itself with a procedure like the two-step hedonic approach suggested by Kräussl and Van Elsland (2008). The models in this chapter are estimated using a continuous reputation variable, as explained below.

*Size:* The most common variable used to describe the physical characteristics of an artwork is its size or surface area. The models use the logarithm of the size of the artwork in  $cm^2$ . They also include size and medium interaction terms. This is particularly important for sculptures, as the size of a sculpture is usually only recorded in terms of its height (in cm). Figure 2.3 illustrates the positive relationship between artwork sizes and prices (the blue line shows the relationship

## Art Prices

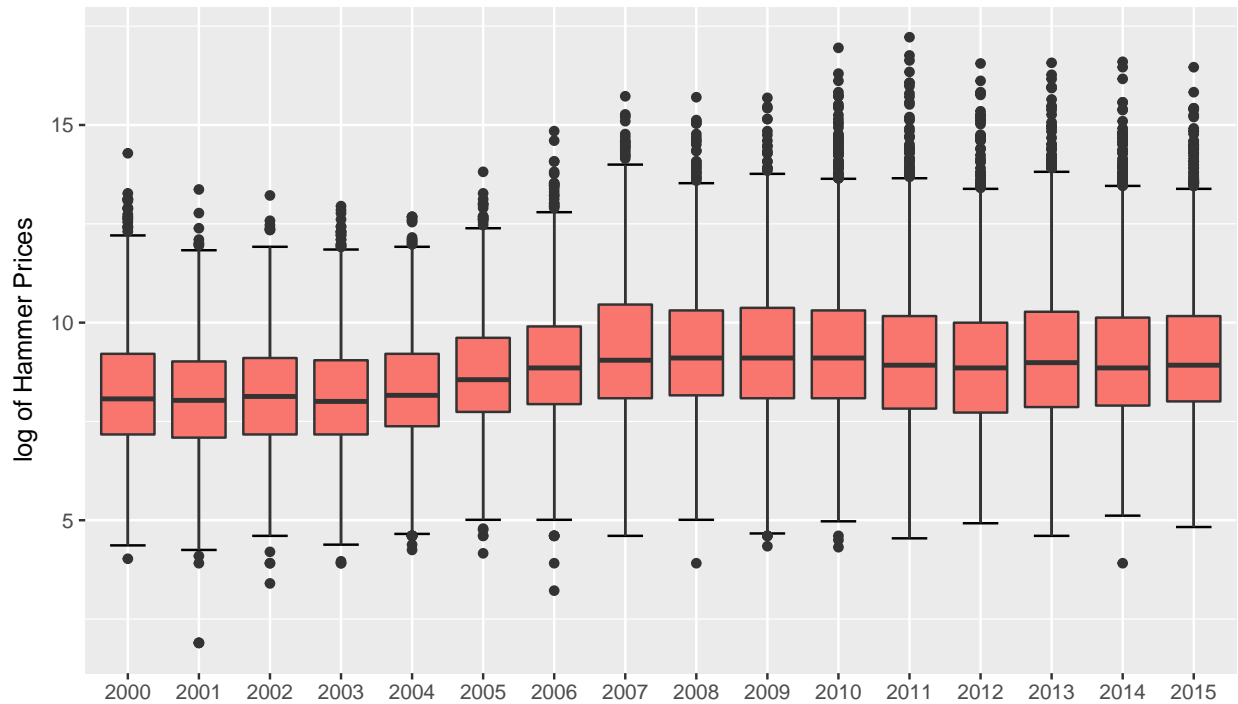


Figure 2.2: Boxplot of (log) hammer prices

for sculptures and the red line shows the relationship for the other mediums). Squared values are occasionally included to take potential non-linearities into account (Fedderke and Li, 2014). In this sample, however, the relationship does not seem to exhibit an inverted U-shape, and the squared term is positive and economically insignificant in the regression models.

*Auction house:* Dummy variables for the auction houses are typically included. The more prominent auction houses usually have a positive effect on prices. One reason might be that more renowned auction houses offer higher quality work (Kräussl and Logher, 2010). Thus, the variables might pick up otherwise unobservable quality differences and do not necessarily reflect auction house certification (Renneboog and Spaenjers, 2013). Moreover, different auction houses charge different commissions to both buyers and sellers. For example, Strauss & Co reported a buyer's premium of 10%-15%, while Bonhams charged premiums of up to 25% (Olckers, Kannemeyer and Stevenson, 2015). The hammer prices exclude these premiums and are therefore not a perfect measure of the buyer's cost or the seller's revenue. For the purposes of a price index, the auction house dummies should capture the different premiums charged by the auction houses.

*Mediums:* Average prices vary across mediums. This might be due to the durability of the medium, the production stage the medium is associated with (e.g. preparatory drawings), or in some cases, the value of the materials used (e.g. sculptures cast in bronze). Oil paintings traditionally earn the highest prices. The availability of copies may decrease the prices of prints and photographs relative to other mediums. Studies typically include dummy variables for the different mediums defined in

## Art Prices

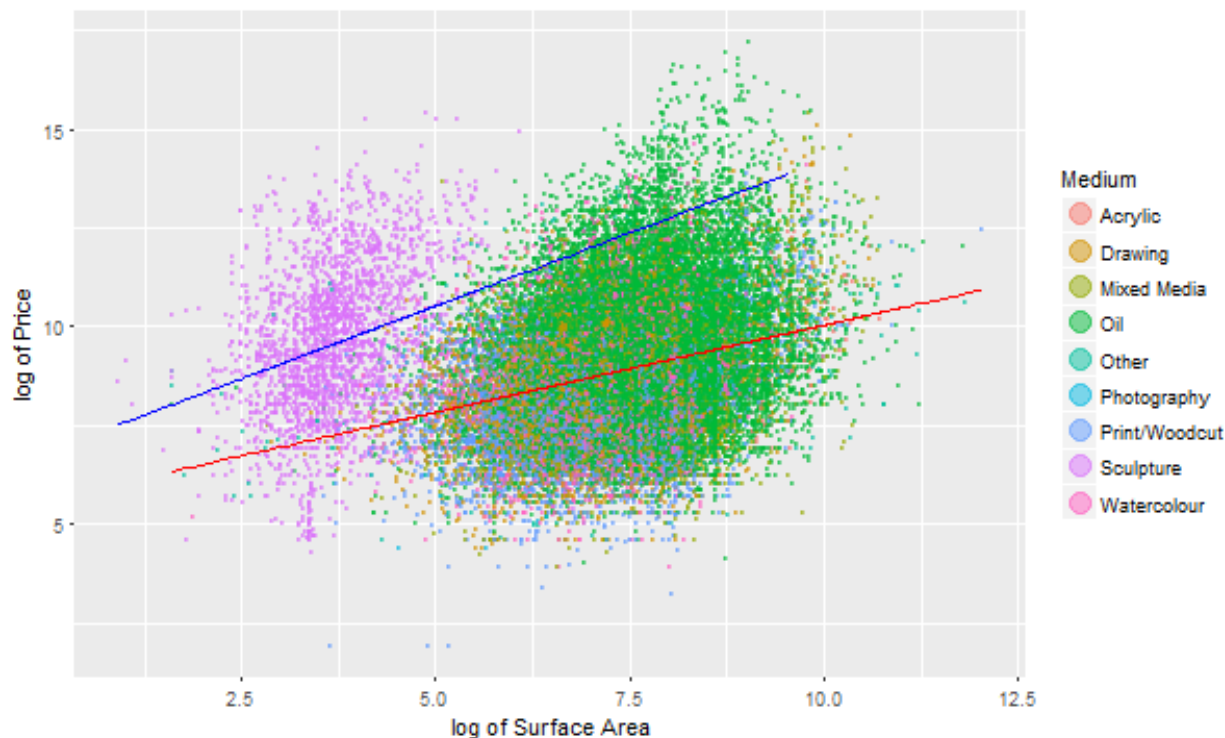


Figure 2.3: Relationship between prices and artwork sizes, by medium

their data (Kräussl and Logher, 2010). The models in this chapter use the nine mediums defined in the dataset; the same mediums were used by Olckers, Kannemeyer and Stevenson (2015).<sup>12</sup>

*Authenticity dummies:* Models often include dummies for whether the artwork is signed and dated. There might be a premium for these attributes, as there is less uncertainty about authenticity (Renneboog and Spaenjers, 2015). These dummies are included in the models below and are expected to have positive coefficients.

*Number of works in the lot:* The models below also control for cases in which more than one artwork was sold in the same auction lot. This is because the recorded size corresponds to each artwork separately and not to the group. Moreover, it is possible that lots including more than one artwork fetch a lower price per artwork than if they are sold separately.

*Date dummies:* The models below include time dummies at a quarterly frequency, which are used to estimate the indices. The exponentials of the time dummy coefficients represent the growth in

<sup>12</sup>The data do not include enough detail to differentiate between medium (e.g. oil) and material (e.g. canvas), or to identify the subject matter or theme of an artwork (e.g. portraits, landscapes, and abstract works). A few studies have included dummies to indicate whether an artist was alive. Artworks of artists who are no longer alive are generally thought to be more valuable, as production has ceased. However, artists who are no longer alive are not able to build on their reputations, which might result in lower sale prices in the long run (Kräussl & Lee 2010). Hence, it is not clear if the variable will be significant. Fedderke & Li (2014) found that the date of death and the age of the artist were statistically insignificant in their South African sample.

art prices in a specific period, relative to the common base period.<sup>13</sup>

### 2.3.4 Continuous Artist Reputation Variable: Two-Step Hedonic Approach

The number of artist dummy variables that can be included in the hedonic regression is limited by the degrees of freedom, which means that some artists usually get excluded from the sample. Kräussl and Van Elsland (2008) developed a two-step hedonic approach, which allows the use of every auction record, rather than only a selected subsample of artists. The approach involves estimating a continuous artist reputation variable, which is included in the regression instead of the artist dummy variables. The approach increases the sample size of artworks that can be included in the regression models and reduces selection bias.

Triplett (2004) showed that a hedonic function with a logarithmic dependent variable would yield an index equal to the ratio of the unweighted geometric means of prices in periods  $t$  and  $t + 1$ , divided by a hedonic quality adjustment, as below. The superscripts  $n$  and  $m$  indicate the generally unequal number of artworks sold per period. The hedonic quality adjustment is a measure of the mean change in the  $j$  characteristics of items sold in period  $t$  and  $t + 1$ , valued by their implicit prices ( $\beta_j$ ):

$$Index = \frac{\prod_{i=1}^n (P_{i,t+1})^{\frac{1}{n}}}{\prod_{i=1}^m (P_{i,t})^{\frac{1}{m}}} / \text{hedonic adjustment}$$

$$\text{hedonic adjustment} = \exp \left[ \sum_{j=1}^J \beta_j \left( \sum_{i=0}^n \frac{X_{ji,t+1}}{n} - \sum_{i=1}^m \frac{X_{ji,t}}{m} \right) \right]$$

Kräussl and Van Elsland (2008) argued that the same method could be used to adjust the average price of an artist's work for differences in quality. The resulting index yields the value of artworks by artist  $y$ , relative to the base artist 0:

$$\text{Artist reputation index} = \frac{\prod_{i=1}^n (P_{i,y})^{\frac{1}{n}} / \prod_{i=1}^m (P_{i,0})^{\frac{1}{m}}}{\exp \left[ \sum_{j=1}^z \beta_j \left( \sum_{i=0}^n \frac{X_{ji,y}}{n} - \sum_{i=1}^m \frac{X_{ji,0}}{m} \right) \right]},$$

where  $P_{i,y}$  is the value of painting  $i$  ( $i = 0, \dots, n$ ) created by artist  $y$ ;  $X_{ji}$  are the characteristics of the artworks, excluding the artist dummy variables.

The first step is to estimate the full hedonic model on a subsample of artists, using artist dummy variables, to obtain the characteristic prices ( $\beta_j$ ). Following Kräussl and Van Elsland (2008), the subsample includes the top 100 artists in terms of volume, representing 53% of records and 92% of the value in the sample. The coefficients are similar to those for the full pooled model, and it is

<sup>13</sup>Because of the log transformation prior to estimation, the index reflects the geometric mean, rather than the arithmetic mean, of prices over time. If it is assumed that the regression residuals are normally distributed in each period, a correction can be made by defining corrected index values as  $I_t = \exp \left[ \gamma_t + 1/2(\sigma_t^2 - \sigma_0^2) \right] * 100$ , where  $\sigma_t^2$  is the estimated variance of the residuals in period  $t$  (Renneboog & Spaenjers 2012). In practice, this adjustment is often negligible (Hansen 2009), which is also the case in this sample.

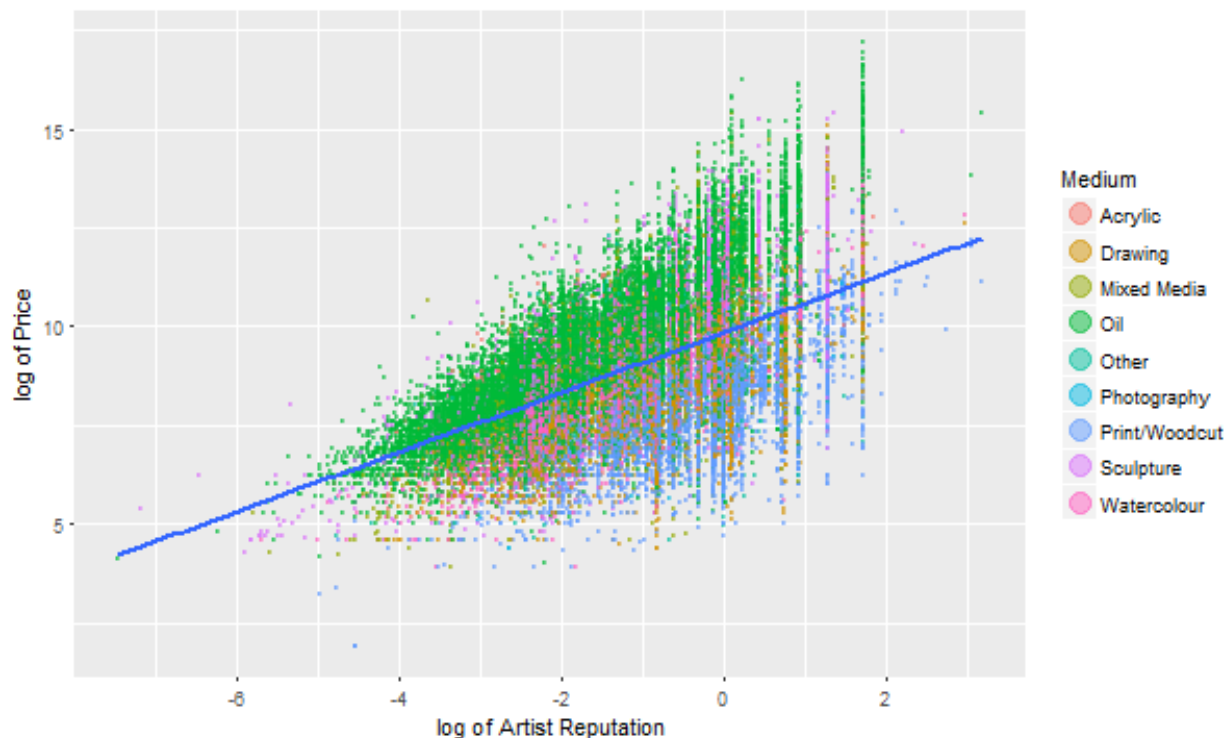


Figure 2.4: Relationship between prices and the reputation index

assumed that the characteristic prices are representative. In the second step, the artist reputation index is calculated for each artist relative to the base artist (Walter Battiss), i.e. the quality-adjusted prices for the works of artist  $y$  relative to artist 0. The reputation index is then used as a continuous proxy variable for artistic value in the hedonic models, instead of the multiple artist dummies.

Figure 2.4 illustrates the positive relationship between artwork prices and the reputation index. As a robustness check, the models are also estimated including all of the artist dummies, except for those artists who sold only one artwork over the sample period. The results are very similar, in line with the findings in Kräussl and Van Elsland (2008).

## 2.4 Index Results

This section presents the results for the three sets of quarterly art price indices, using central tendency, hedonic and hybrid repeat sales methods. It also examines different segments of the art market, in order to establish in which segments the marked price increases occurred, and estimates a few indices for individual artists.

## Art Prices

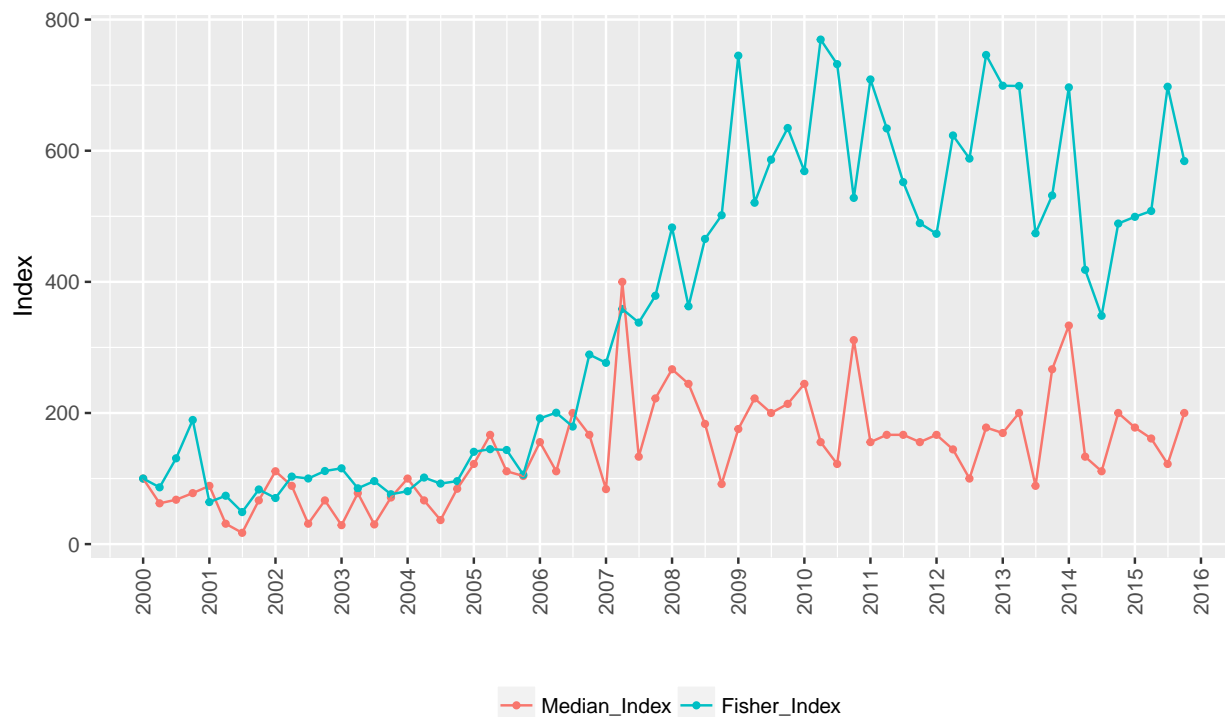


Figure 2.5: Central tendency indices of South African art prices (2000Q1=100)

### 2.4.1 Central Tendency Indices

Two central tendency price indices are estimated at a quarterly frequency to act as a baseline in comparing the indices resulting from the different methodologies. The median index is simply the median price for each quarter. The Fisher index is a mix-adjusted central tendency index, which stratifies the sample into subgroups by artist and medium. The Fisher index is calculated as the geometric mean of the Laspeyres and Paasche indices. The base periods were allowed to vary for each index point, and the index points were then chained together to form the overall chain-link index.

Figure 2.5 illustrates the two central tendency indices. The simple median index provides a noisy estimate of price changes, and no consistent picture emerges. The large variation is likely due to the large differences in the quality-mix or composition of the artworks sold between different periods.

The Fisher index also exhibits a large variation and implausibly large increases over the sample period. In this case, the stratification does not seem to be very effective. This is probably because the artist and medium categories capture only a small portion of the differences in the quality of artworks sold between periods. The quality-adjusted measure does not take account of any changes in the quality-mix of artworks sold that are unrelated to artist and medium type. The stratified index also does not account for changes in the composition of artworks sold within each subgroup, in this case changes in the quality-mix of artworks by a certain artist in a specific medium (Eurostat,



2013). Moreover, the subgroups become small when separated in this way, which means that small changes can have a large effect on the index.

The results illustrate that central tendency measures are deficient in this case and should be used with caution, echoing the findings in Els and Von Fintel (2010) for South African real estate. As a consequence, regression-based measures are generally preferred in the academic literature.

#### 2.4.2 Hedonic Indices

The full pooled sample regression results are reported in Table 2.1. Nearly all of the coefficients are significant (with the exception of the smaller medium categories) and have the expected signs. The size of the artwork is highly significant and positive. Bonhams and Strauss & Co are the auction houses with the highest average prices, after controlling for other factors, perhaps reflecting higher quality work. Oil is the most expensive medium category. The medium and size interaction terms are all negative and mostly significant, except for the sculpture size term, which is positive and significant. The authentication dummies are both positive and significant, as is the artist reputation variable. The number of works variable indicates that more than one artwork in a lot leads to slightly lower prices per artwork. The adjusted  $R^2$  is relatively high, suggesting that these variables capture a large part of the variation in sales prices.<sup>14</sup> The time dummy coefficients are used to calculate the full period pooled hedonic index.

To allow for shifts in the implicit prices, two adjacent-period or chain-linked indices were calculated by estimating separate models for adjacent subsamples. Selecting the length of the estimation window involves a trade-off. A shorter window decreases the likelihood of large breaks, but also reduces the number of observations used to estimate the parameters (Dorsey *et al.*, 2010). 1-year and 2-year estimation windows were selected, similar to Dorsey *et al.* (2010) in the context of real estate and Renneboog and Spaenjers (2013) in the context of art. This seems to be reasonable for the South African art market, where large auctions are held infrequently. The indices were then calculated by chain-linking the estimates together, as Figure 2.6 illustrates for the 2-year version of the index.

In the context of real estate, Shimizu, Nishimura and Watanabe (2010) suggests a so-called overlapping-periods hedonic regression method using multiple ‘neighbourhood periods’, allowing gradual shifts in the parameters. Parameters are estimated by choosing a specific estimation window size and shifting the period forward in rolling regressions. They argue that this method will be able to deal better with seasonal changes in parameters than adjacent-periods regressions. To apply this method, 5-year rolling regressions were run, which correspond to the rolling 5-year regressions used to estimate the Citadel Art Price Index. The estimation window was then shifted forward one year,

---

<sup>14</sup>The diagnostic tests indicate that there might be some problems with the assumptions of normality and homoscedasticity of the residuals. These assumptions are not crucial, however, as only the point estimates of the time dummy coefficients are of interest.

## Art Prices

Table 2.1: Hedonic regression results

	<i>Dependent variable:</i>
	log(Price)
log(Area)	0.478*** (0.029)
Auction House (ref = 5th Avenue)	
Ashbeys	0.111*** (0.027)
Bernardi	0.125*** (0.013)
Bonhams	1.164*** (0.026)
Christies	1.133*** (0.064)
Russell Kaplan	0.127*** (0.015)
Stephan Welz	0.582*** (0.013)
Strauss	1.100*** (0.015)
Medium (ref = Acrylic)	
Drawing	-1.341*** (0.251)
Mixed Media	-0.826*** (0.250)
Oil	0.668*** (0.239)
Other	0.404 (0.278)
Photography	-0.827 (0.669)
Print/Woodcut	-0.757*** (0.243)
Sculpture	1.233*** (0.248)
Watercolour	0.019 (0.255)
Signed	0.200*** (0.015)
Dated	0.048*** (0.007)
No. Works	-0.090*** (0.003)
log(Reputation)	0.950*** (0.003)
Constant	5.562*** (0.243)
Medium Size Interactions	Yes
Quarterly dummies	Yes
Observations	51,454
R <sup>2</sup>	0.789
Adjusted R <sup>2</sup>	0.789
Residual Std. Error	0.768 (df = 51362)
F Statistic	2,111.353*** (df = 91; 51362)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

allowing for gradual shifts in the parameters.

The coefficients from these models are similar in magnitude to the full pooled sample model and are significant in virtually all cases. For example, the coefficient associated with the size of the artwork is 0.426 in the standard full hedonic regression, while the average coefficients from the other regressions are 0.44, 0.43 and 0.42. The coefficient estimates from the 1-year adjacent-period hedonic models are illustrated in the chapter Appendix (section 2.8). The coefficient estimates are relatively stable over time, and there do not appear to be any large structural changes over the period. However, there are a few cases in which the estimated parameters fluctuate over time, indicating that it is useful to allow for shifts in the implicit prices over time. For example, the coefficient of the dummy for whether the artwork is signed varies between 0 in the model for 2005 and 0.5 in the 2008 model.

The adjusted- $R^2$  values for the 1-year adjacent-period hedonic models are also illustrated in the

## Art Prices

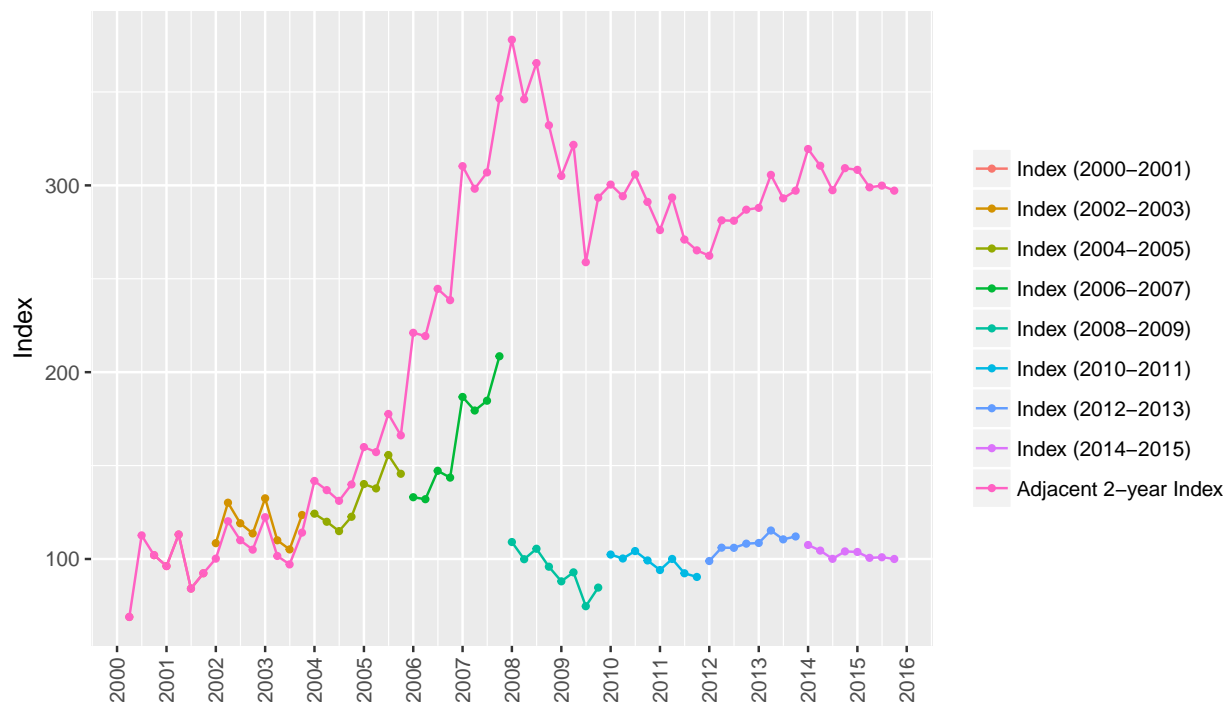


Figure 2.6: Chain-linked 2-year adjacent-period art price index

chapter Appendix. This gives an indication of how influential the hedonic features are at various points in time, and how strong the ‘signal’ of the resulting index is. The adjusted- $R^2$  is relatively high throughout the sample period, and varies between 0.63 in the 2000 model and 0.83 in the 2013 model.

Figure 2.7 illustrates the resulting quarterly art price indices from these four models. The hedonic indices follow a similar cyclical pattern over the period, which seems to reflect developments in economic activity, although the levels are slightly different. The indices seem more plausible than the central tendency measures, supporting the case for regression-based measures. All four of the indices display dramatic increases in auction prices of more than 200% between 2003 and 2008. All four indices peak in 2008Q1, which is before the peak in sales and annual median prices in the sample. This conforms to the idea of the formation of a so-called bubble, with a dramatic rise and subsequent decrease in prices. The indices are all relatively flat after 2009, even in nominal terms. The hedonic indices, along with the 95% confidence interval for the full hedonic model, are illustrated in the Appendix. The indices are within the confidence interval for the entire period, except for the 1-year adjacent-period model, which is below the confidence interval for the latter part of the period.

The hedonic price indices therefore display similar trends over the period, with large price increases in the run-up to the Great Recession. However, the hedonic indices may suffer from omitted variable or misspecification bias, due to unobservable (or difficult to measure) nuances that make

## Art Prices

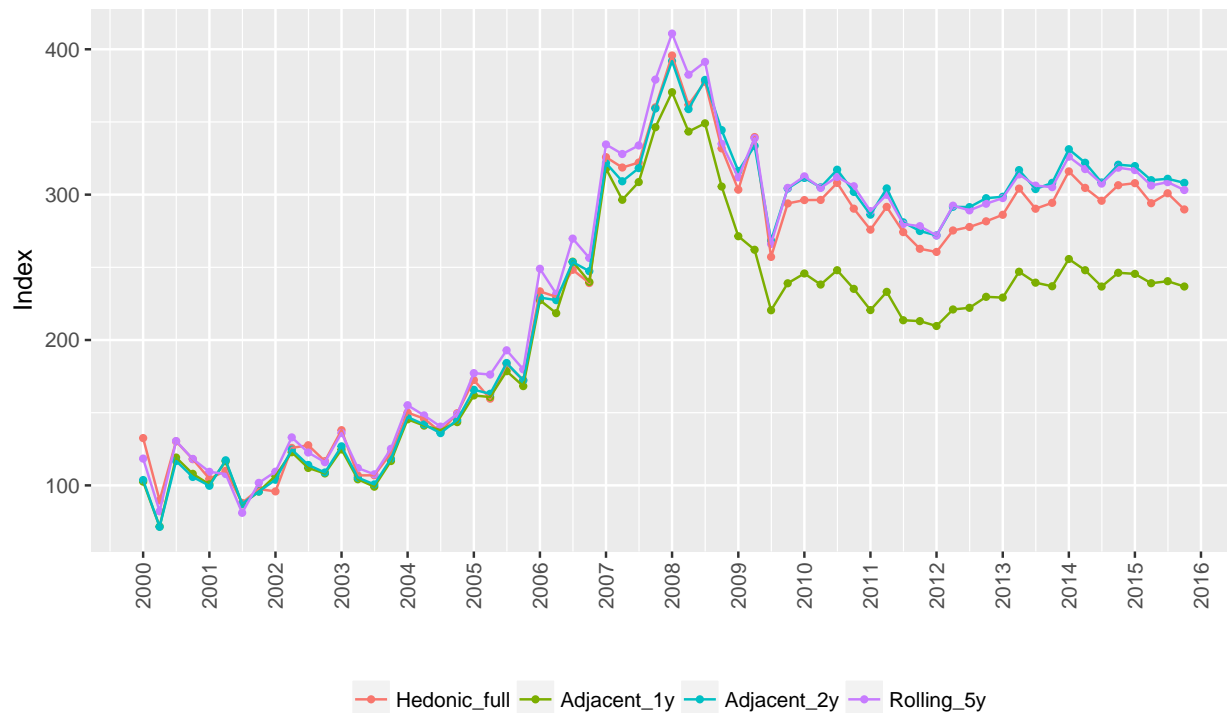


Figure 2.7: Hedonic indices of South African art prices (2001=100)

a given artwork unique and influence its price. The omitted variables might include, for instance, interaction terms (e.g. artist and medium combinations), squared terms, finer medium classifications (e.g. linocuts), or attributes such as material, theme and style (e.g. canvas or landscape). These omitted variables potentially bias the coefficients if they are correlated with sales timing, which in turn may bias the indices, although the bias is often small in practice (Triplett, 2004; Renneboog and Spaenjers, 2013).

### 2.4.3 Repeat Sales and Hybrid Indices

The repeat sales method is less prone to potential omitted variable bias than the hedonic method, as it tracks sales of the same item over time. Because the dataset does not uniquely identify each artwork, repeat sales of the same artwork were identified by matching sales records using the following attributes: artist name, artwork title, size, medium, the presence of a signature and a date, and the number of artworks in the lot. Only 515 true repeat-sales pairs could be identified in the sample. Figure 2.8 illustrates the index generated using the classical repeat sales approach. The index is volatile, with many missing values, and exhibits a large appreciation in prices over the period. The limited number of repeat sales observations therefore limits the usefulness of the classical repeated sales approach in this case.

In this chapter a simple new hybrid repeat sales model is proposed as an alternative. This procedure

## Art Prices

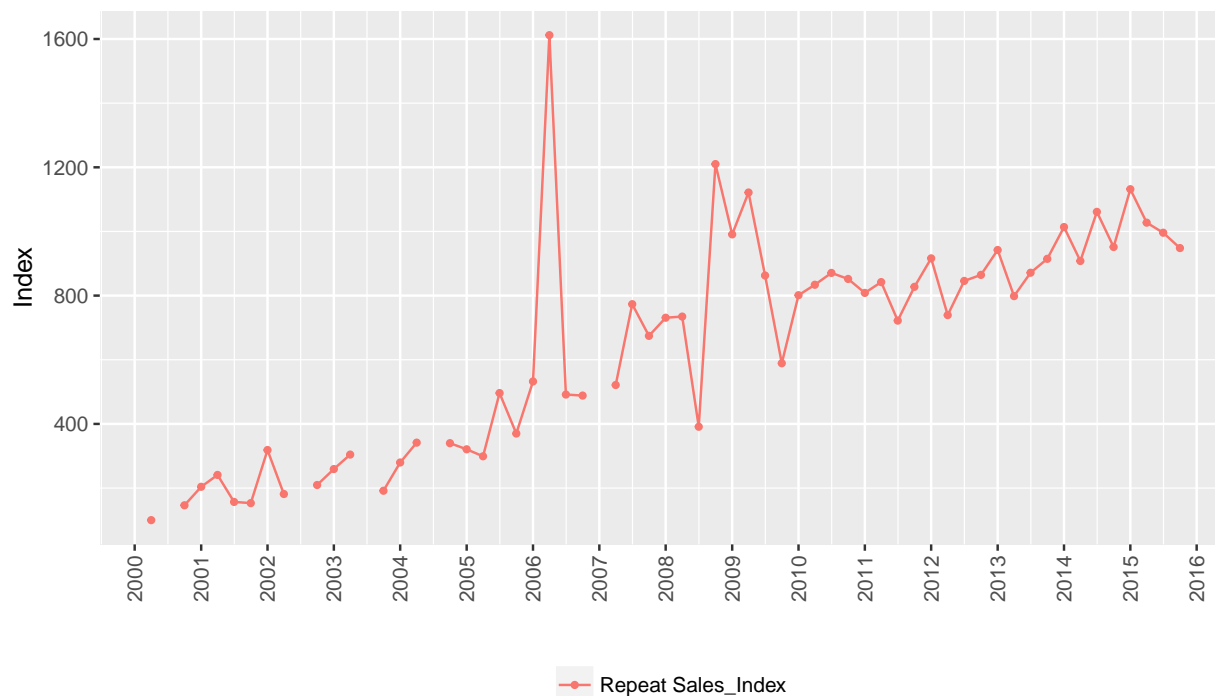


Figure 2.8: Repeat sales index of South African art prices (2000Q4=100)

is similar in spirit to the ‘pseudo repeat sales’ (ps-RS) procedure suggested by Guo *et al.* (2014). Instead of requiring exact matches to form sales pairs, very similar artworks may be treated as repeat sales pairs. In this way, the ps-RS method supplements the true repeat sales in the sample and mitigates the problem of small sample size. In so doing it allows for the estimation of a variant of the repeat sales index, which should address to some extent the potential omitted variable bias inherent in the hedonic method. The caveat is that even two artworks by the same artist of a similar size and in the same medium do not necessarily serve as close substitutes (Olckers, Kannemeyer and Stevenson, 2015).

The first ps-RS sample is created by matching artworks on all the hedonic attributes, except for the title of the artwork. The matched pairs therefore have the same hedonic attributes except for the title of the artwork. Matching by these criteria increases the number of repeat sales pairs to 6,539, which includes the 515 true repeat sales or exact matches. The second ps-RS sample allows the sample to increase further by matching on all the hedonic attributes except the title and the presence of a signature and date on the artwork, i.e. exact covariate matches except for the title and the authenticity dummies. This increases the pseudo repeat sales sample to 7,816 sales pairs. This involved a trade-off between the within-pair ‘similarity’ and the sample size. Higher similarity is good for mitigating bias, while a larger size is good for reducing standard errors (Guo *et al.*, 2014). Figure 2.9 illustrates the sample sizes of these two pseudo-repeat sales samples, which are similar. This smaller sample includes 13,290 of the original 52,059 observations, and the larger

## Art Prices

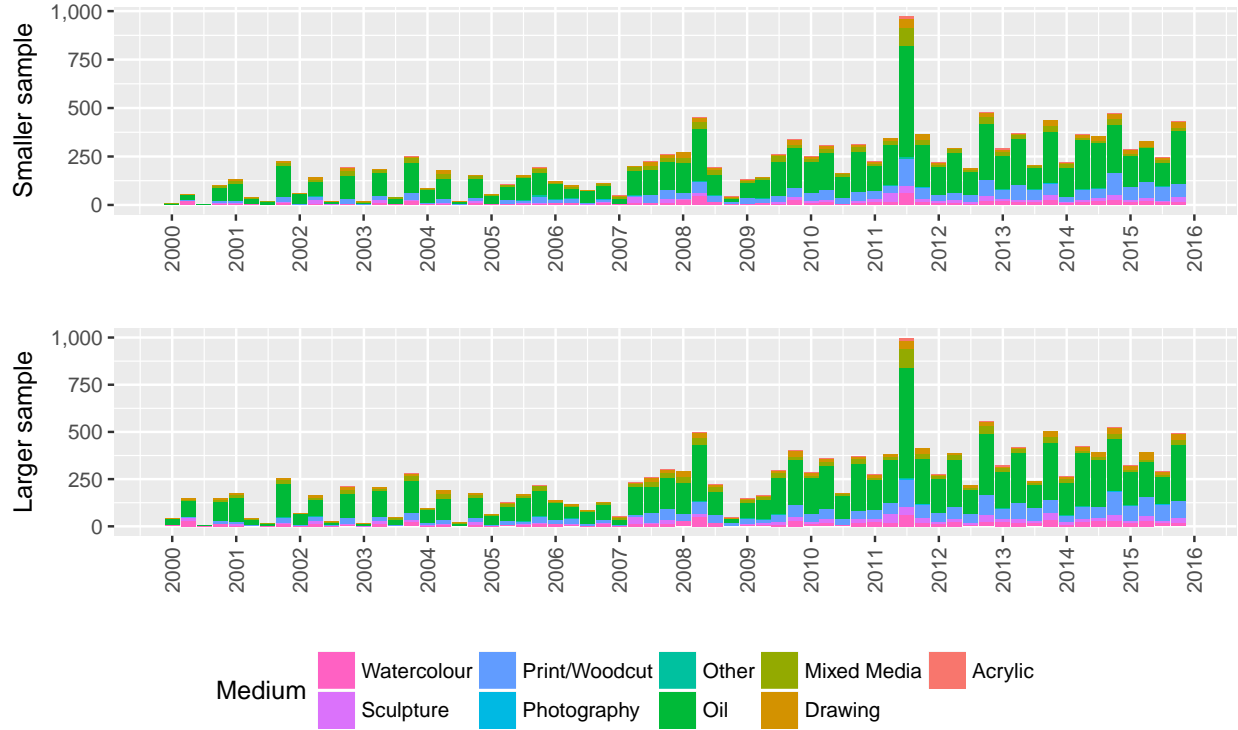


Figure 2.9: Pseudo-repeat sales samples

sample includes 15,341 observations.

The differential hedonic equation is then applied to the pseudo repeat sales samples, where artwork  $i$  in quarter  $t$  and artwork  $h$  in quarter  $s$  form a matched pair ( $t > s$ ):

$$\ln P_{it} - \ln P_{hs} = \sum_{j=1}^J \beta_j (X_{itj} - X_{hsj}) + \sum_{t=0}^T \delta_t G_{it} + \epsilon_{iths},$$

where  $G_{it}$  is again a time dummy equal to 1 if the later sale occurred in quarter  $t$ , -1 if the former sale in the pair occurred in quarter  $s$ , and 0 otherwise; and  $\epsilon_{iths}$  again represents a white noise residual.

For the first ps-RS sample, the only remaining independent variable is the difference in the auction house dummies ( $X_{it1} - X_{hs1}$ ). This takes account of possible differences in quality and commission structures. In the second ps-RS sample, the independent variables represent the differences in the auction house dummies and the differences in the two authenticity dummies. The independent variables therefore include these within-pair differentials in attributes, which are relatively small and easy to measure.

Thus, the ps-RS approach addresses the problem of lack of repeat sales data and to some extent the potential omitted variable bias inherent in the hedonic method. The pseudo repeat sales pairs include the 515 true repeat sales, or exact matches, where the model controls for all the observed and

## Art Prices

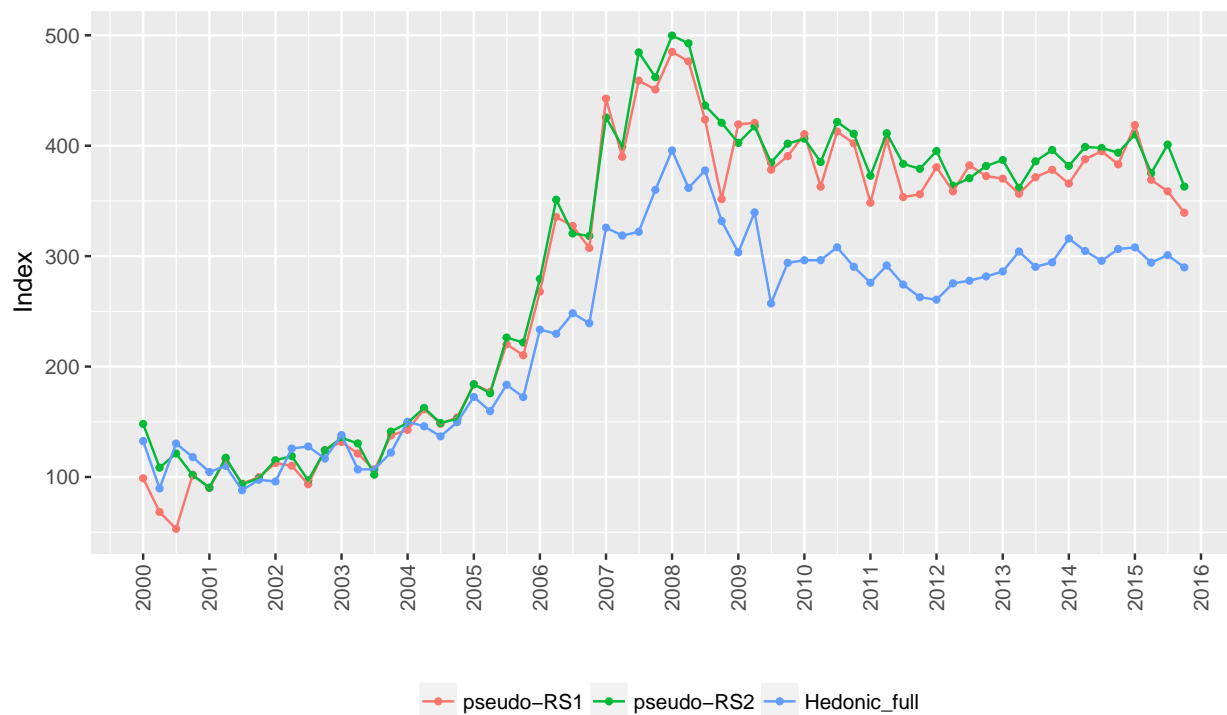


Figure 2.10: Pseudo-repeat sales indices of South African art prices (2001=100)

unobserved attributes by taking first differences. For the pseudo sales pairs, taking first differences will control for omitted variables when they are the same for the two items that form the pseudo sales pairs. For example, if Van Gogh's *Sunflowers* paintings are treated as repeat sales, taking first differences would control for attributes such as theme, style, material, prominence, and the stage of the artist's career. Other potentially significant variables might include an array of interaction and non-linear terms.

Figure 2.10 illustrates the two versions of the ps-RS indices, together with the full hedonic index. The indices point to similar cyclical trends in art prices over the sample period. Both ps-RS indices appreciated rapidly in the run-up to the Great Recession, peaked in 2008Q1, and were relatively flat after 2009. This is a similar trend to the full hedonic index, although there is an even more marked increase in South African art prices in 2006 and 2007. This provides more confidence in the results being robust to changes in methodology. These indices are illustrated in the Appendix, along with the 95% confidence interval for the larger sample model (ps-RS2). The smaller sample ps-RS index is within the confidence interval for the entire period. The full hedonic index is within the confidence interval for the first half of the sample period, but hovers around the lower bound for the second half of the period.

#### 2.4.4 Market Segments and Artist Indices

This section examines whether different segments of the South African art market experienced different price trends over time. The market may be segmented in a number of ways, such as by price, artist value, and medium category. The caveat is that slicing the data thinly results in small sample sizes and more volatile indices. This makes it more difficult to discern a pattern and to distinguish the signal from the noise. The indices should therefore be interpreted with caution.

Fedderke and Li (2014) suggested that the South African art market should be segmented into three price ranges and found different hedonic relationships for the three market segments. The art market may be segmented for a variety of reasons, such as the following: first, wealthy individuals may be less tempted to buy artworks at the lower end of the market that do not signal the same social status; second, small investors are typically unable to purchase more expensive artworks; and third, more expensive artworks may be more prone to speculation (Renneboog and Spaenjers, 2013).

To test the possibility that historical rates of appreciation varied across the price distribution, and particularly in the upper part of the distribution, separate indices for different parts of the price distribution are calculated. Figure 2.11 illustrates the distribution of the log of art prices for the entire sample period. The median and the 90th percentile are illustrated as vertical lines, which correspond to prices of R6,500 and R70,000. Separate hedonic indices<sup>15</sup> are estimated for each of these three segments - i.e. the lower half of the distribution (all artworks below the median of R6,500), a middle part of the distribution (artworks between R6,500 and R70,000), and the upper part of the distribution (artworks above R70,000). This allows the characteristic prices to vary across the price distribution.

Quantile regressions provide an alternative means to investigate different parts of the price distribution. Quantile regressions can characterise the entire distribution of the dependent variable, as opposed to OLS regressions, which provide estimates for conditional means. Quantile regressions are also more robust to potential outliers. Quantile regression and indices are estimated for the middle points of the three segments of the price distribution, i.e. the 25th percentile, the 70th percentile and the 95th percentile.

Figure 2.12 illustrates the two sets of indices for the three segments of the price distribution. The indices suggest that the dramatic price increases occurred in the upper part of the price distribution. In the first panel, this corresponds to the top 10% of the distribution (artworks with a hammer price of more than R70,000). In the section panel, it corresponds to a quantile regression for the 95th percentile (although this index is relatively volatile).

However, there is a potential weakness with this finding. It stands to reason that expensive artworks experienced a substantial increase in their prices to become expensive. There might be an inherent correlation or endogeneity between rapid growth in the price of the item and it being expensive. One

---

<sup>15</sup>The models are estimated with the full hedonic method. The adjacent-period hedonic and ps-RS models were used to confirm the results.



Art Prices

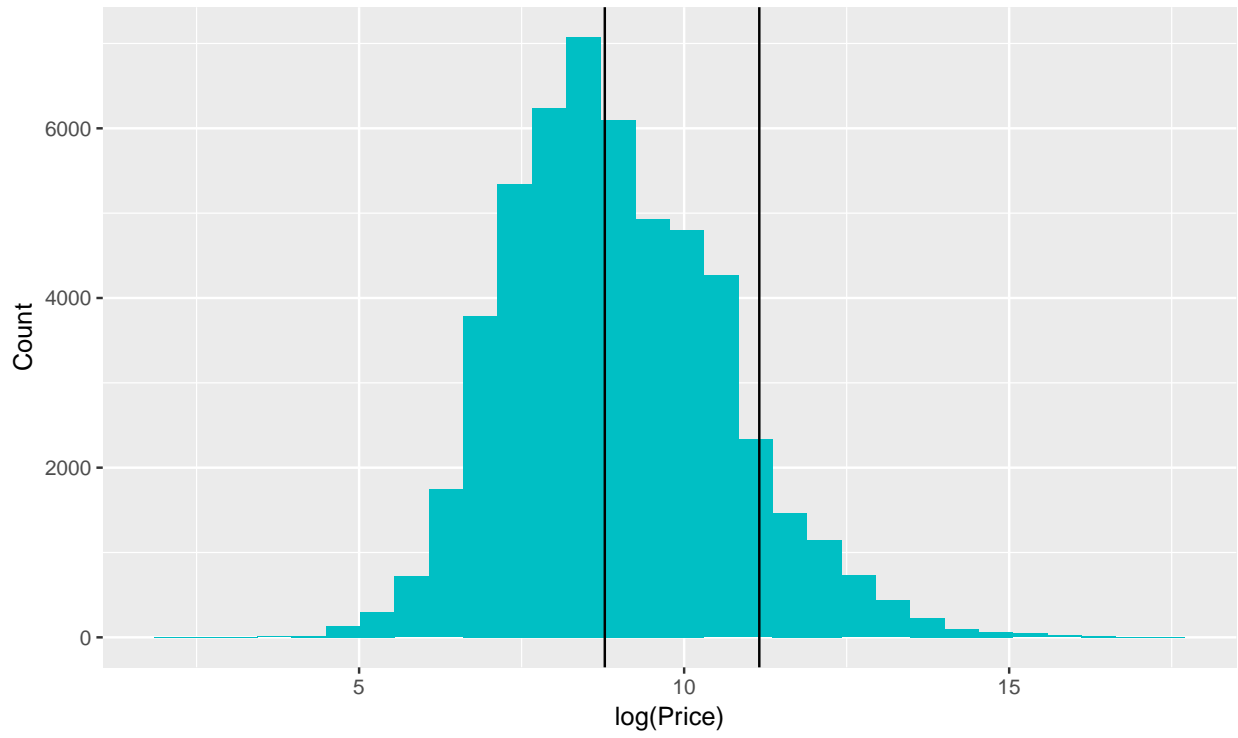


Figure 2.11: Histogram of (log) art prices

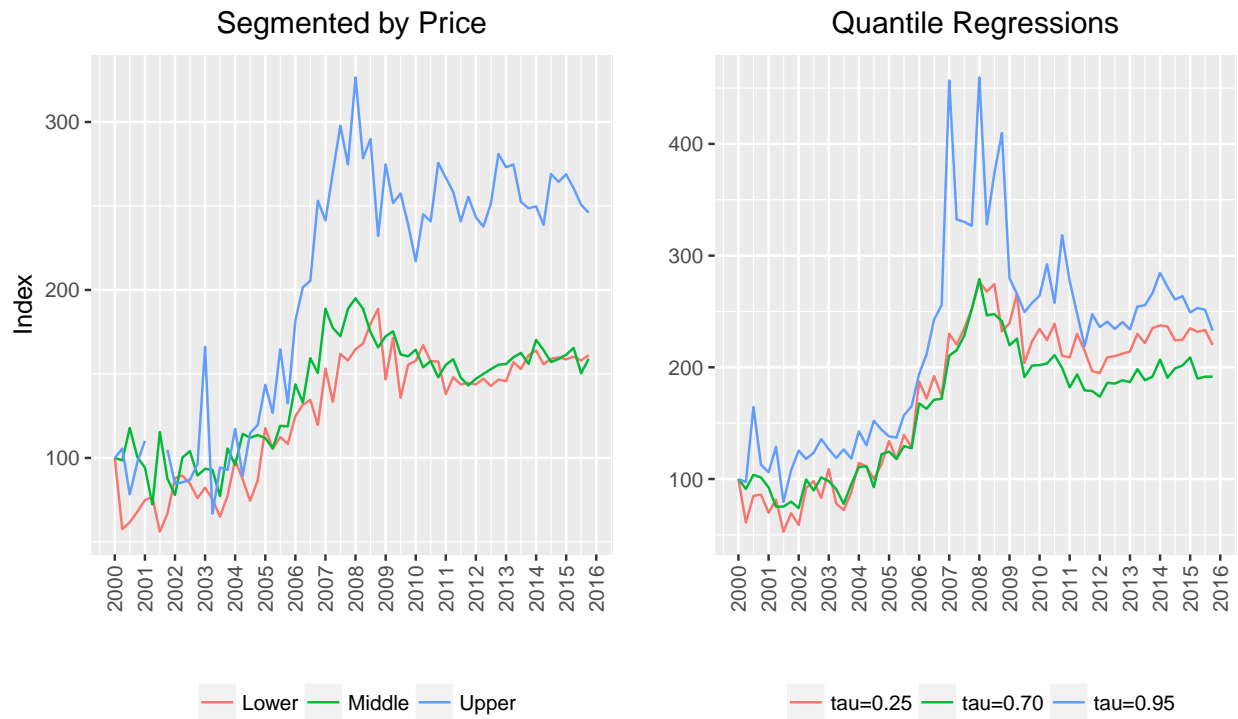


Figure 2.12: Art price indices for different price segments (2000Q1=100)

## Art Prices

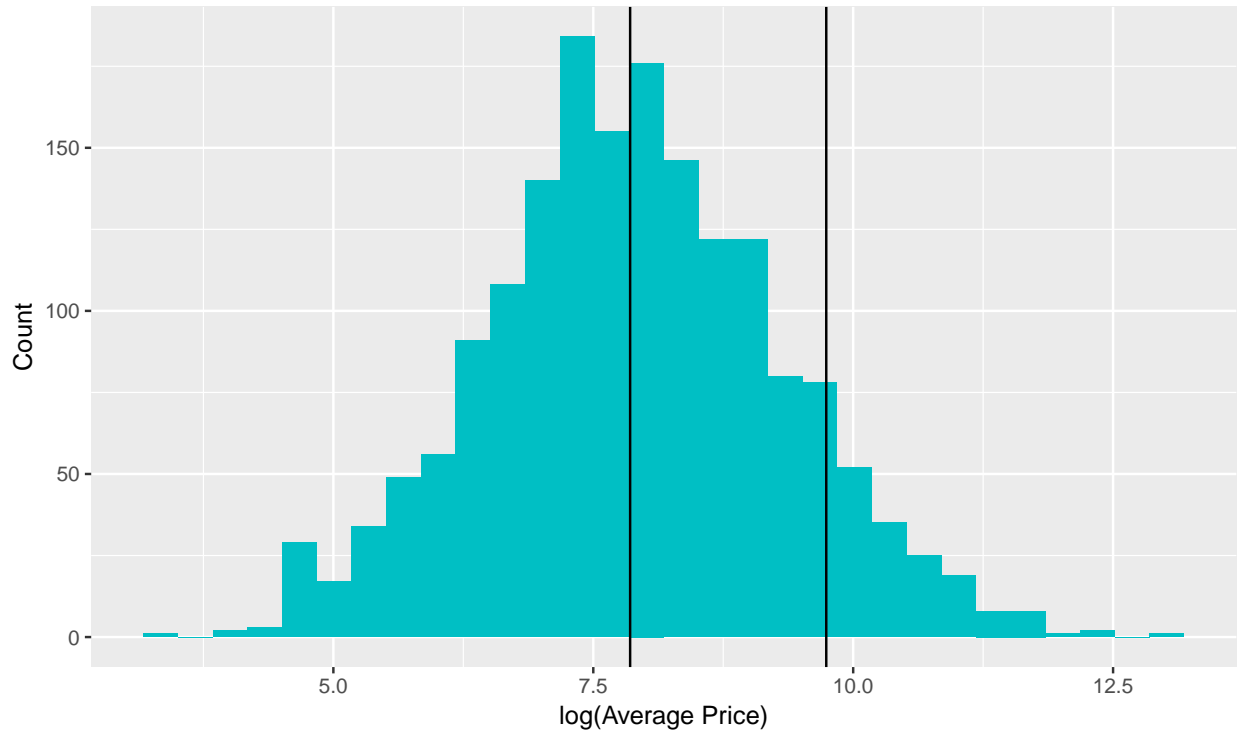


Figure 2.13: Histogram of (log) average art prices (2000-2004)

way to correct for this potential bias would be to track artworks that were expensive in the initial part of the sample period and then to calculate indices with the ps-RS method. Unfortunately, there are too few repeat sales in the sample to do this. To try to correct for this potential endogeneity problem, art prices for only those artists that were present at the beginning of the sample period are investigated, dividing them into separate value segments. This allows an examination of the price appreciation of the work of relatively expensive artists.

Average prices per artwork are calculated for the artists in the first five years of the sample period (2000-2004), before the large appreciation. Figure 2.13 illustrates the price distribution of these average prices. The median and the 90th percentile are illustrated as vertical lines, which correspond to average prices of R2,577 and R17,000. The artists are divided into three value segments, according to the average price of their artworks in the initial period. Separate hedonic regressions are then estimated for the three segments, by including only the artists that formed part of each segment in the initial period. In this way, for instance, the artists that were relatively expensive in the initial part of the sample period can be tracked, by estimating indices that include only ‘expensive artists’.

Figure 2.14 indicates that prices increased more dramatically for artworks by the more expensive artists, which in this case includes artists with an average price of more than R17,000 in the first five years of the sample period. The results are similar when the average values are determined using the reputation variable (i.e. segmenting the artists by their reputation variable in the first five years), although the indices are very volatile.

## Art Prices

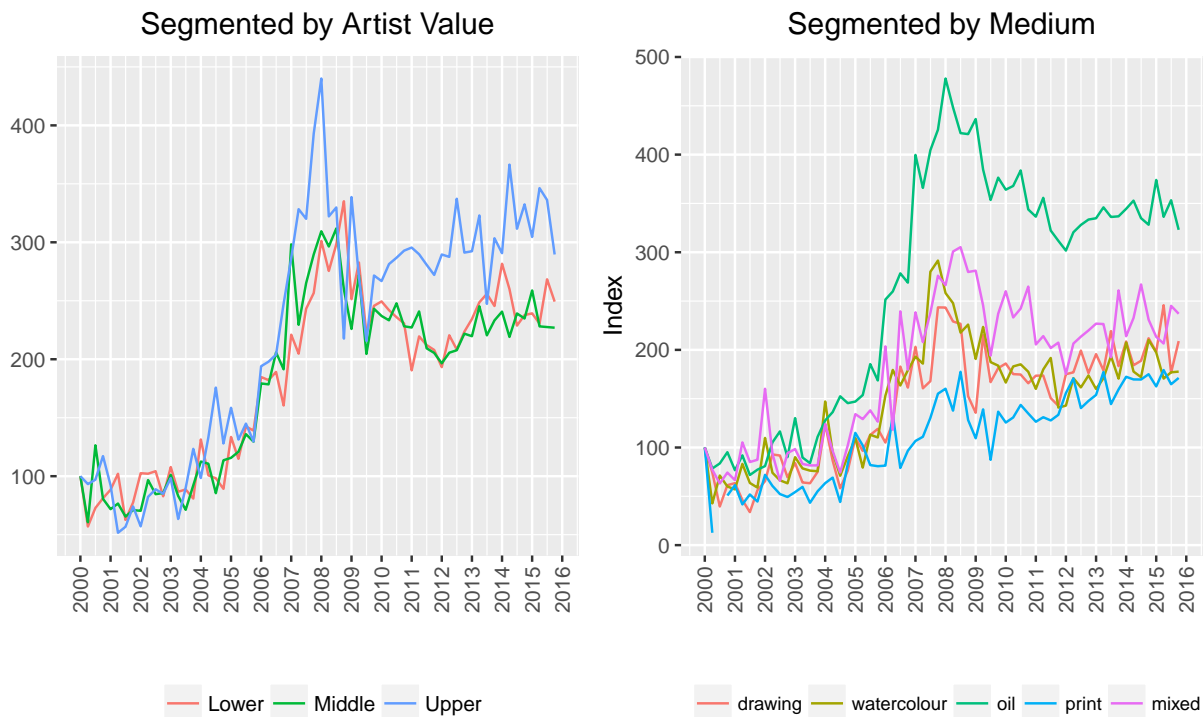


Figure 2.14: Art price indices for value and medium segments (2000Q1=100)

The market may also be segmented by medium category. It is possible that historical rates of appreciation have varied widely over time for different medium categories. Separate hedonic models are estimated for each of the mediums. Figure 2.14 illustrates the indices for five of the medium categories, which together account for 92% of the observations in the sample. Oil paintings are by far the largest category, representing 52% of the volume and 78% of the value of artworks in the sample. The indices indicate that oil was the medium for which there were the largest price increases. Overall, the results seem to indicate that the dramatic price increases occurred especially in the more expensive or higher-end parts of the art market, and for oil paintings.

The sample may be further segmented in order to estimate separate indices for specific artists. Figure 2.15 illustrates indices for four of the leading South African artists, derived from separate hedonic models. Walter Battiss and Gregoire Boonzaier are the top selling artists in the sample in terms of number of sales (most observations). The Battiss index exhibits a sustained quality-adjusted price increase over the period, whereas the Boonzaier index exhibits a more pronounced cyclical pattern.

Irma Stern and JH Pierneef are two of the country's 'masters' and the top-selling artists in terms of turnover. The Pierneef index exhibits a strong increase over the period, but is quite volatile. The Stern index has some missing values towards the beginning of the sample period and is even more volatile. The quality-adjusted prices for Stern do not increase as much as one perhaps would have expected, given the record prices achieved during the sample period. This indicates that the record prices were in line with the model predictions. The large spikes in the Pierneef and Stern

## Art Prices

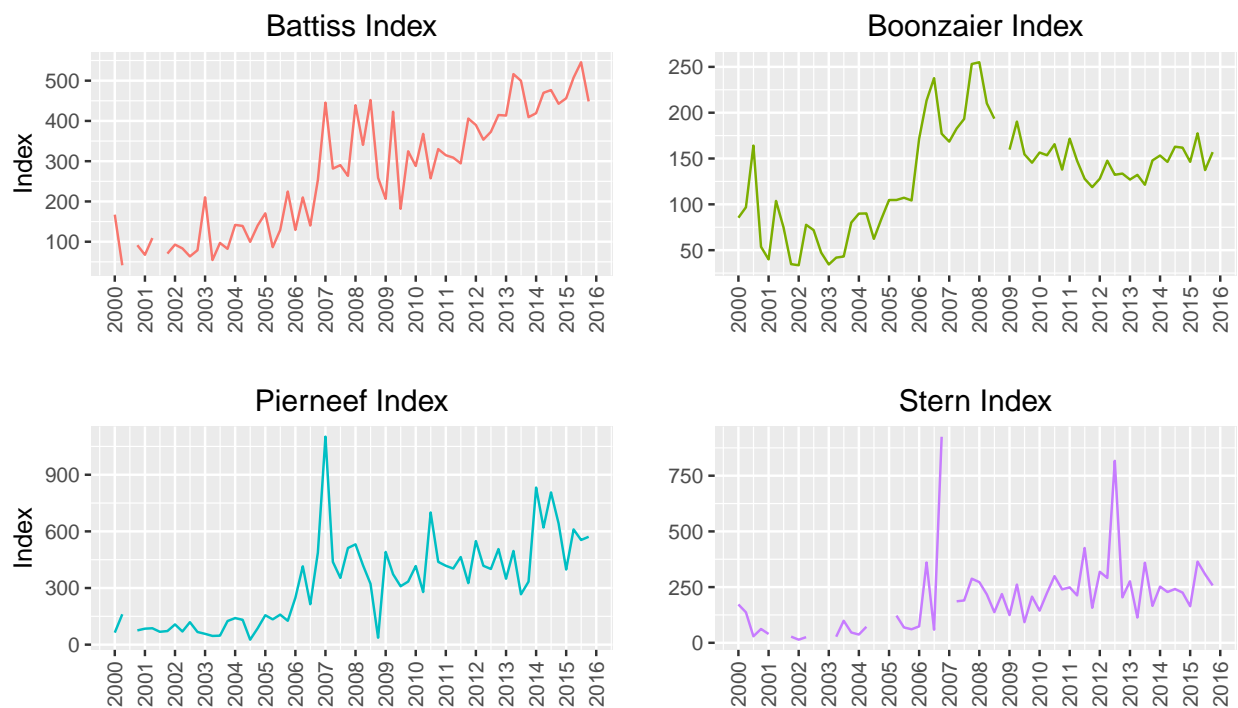


Figure 2.15: Art price indices for individual artists (2000=100)

indices illustrate the small sample problem when slicing the data too thinly. In the quarters with large spikes, only a single artwork by each of the artists was sold. Although these artworks were not particularly expensive, they achieved hammer prices well above the model predictions, which is reflected in the index values.<sup>16</sup> Nonetheless, all four artist indices exhibited large spikes between 2006 and 2007.

## 2.5 Validity Tests and Evaluation

In this section, the internal validity of the indicators is assessed by comparing the indices estimated with the different methodologies, in order to determine whether they provide a consistent picture of price movements in the South African art market. In the absence of an existing South African art price index for an external validity test, the art price indices are compared with traditional South African assets. The indices are then evaluated in terms smoothness, to examine which index provides the most credible gauge of overall price movements in this case.

<sup>16</sup>For instance, the large spike in 2006 in the Stern Index, was due to a single watercolour (gouache) painting *Lady of the Harem*, which sold for R2.2 million at Stephan Welz & Co.

## Art Prices

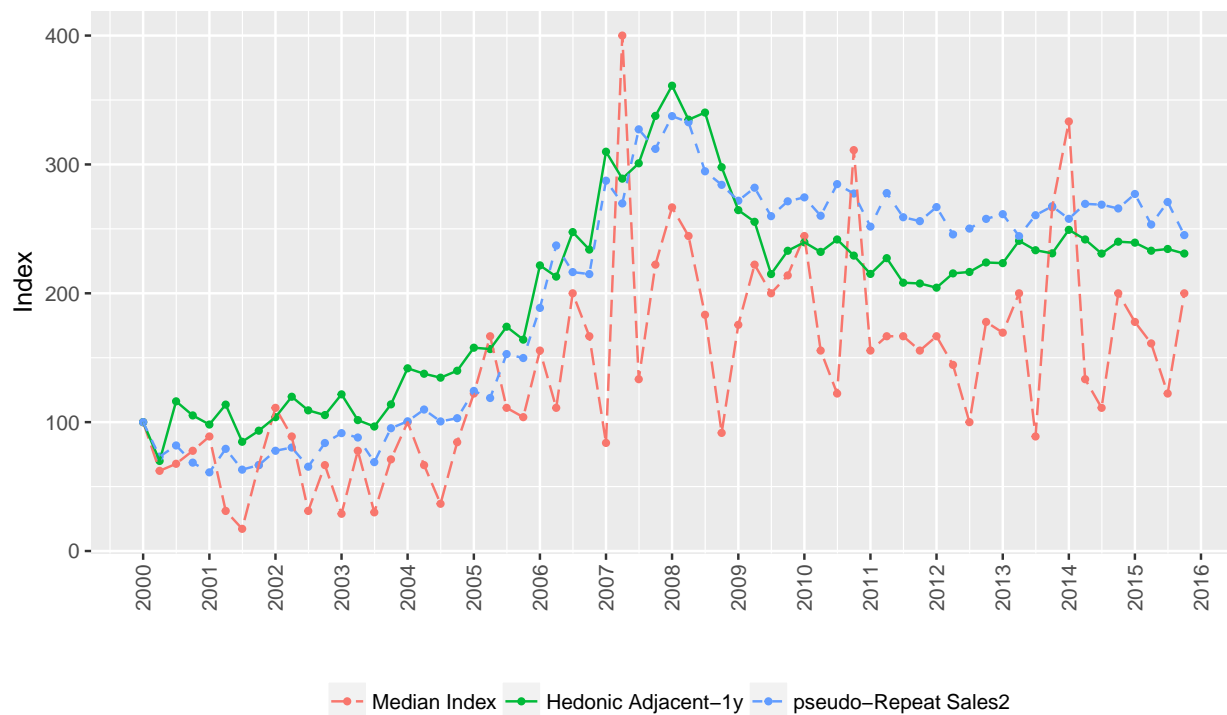


Figure 2.16: Comparing South African art price indices (2000Q1=100)

### 2.5.1 Internal Validity Test

Figure 2.16 illustrates representative indices for the three methodologies: median values, the 1-year adjacent-period hedonic index and the second version (larger sample) of the ps-RS index. The two regression-based indices seem to point to a similar general trend in South African art prices. The simple median index, on the other hand, does not reflect this trend and is much more volatile than the regression-based indices. The regression-based indices are higher in levels than the median index, particularly after the first few years, which implies that lower quality artworks were adjusted upwards from the median levels.

The median index falls outside of the confidence intervals of the regression-based indices for most of the sample period, as illustrated in the Appendix. This implies that it is statistically different from the regression-based indices and that the movements in the regression-based indices are ‘real’ rather than statistical idiosyncrasies. The regression-based methods, which adjust for changes in the composition of artworks sold, therefore seem to provide better estimates of pure price changes for unique assets. The results echo the findings in Els and Von Fintel (2010) for South African real estate prices.

The hedonic and ps-RS indices exhibit similar trends over the sample period, although the hedonic index is at a lower level after 2009. Both measures indicate that the average price of a quality-adjusted artwork increased significantly between 2005 and 2008 and then declined sharply after

## Art Prices

the financial crisis, similar to other asset prices (Shimizu, Nishimura and Watanabe, 2010). Both indices are relatively flat after 2009 in nominal terms, implying that art prices decreased in real terms over latter part of the sample period.

Table 2.2 reports the correlations in the growth rates between the various indices.<sup>17</sup> There is a significant positive correlation between the regression-based indices. This indicates that their general trends are similar, and are different from the simple median. The Fisher central tendency index is also significantly positively correlated with the hedonic indices. This shows that there is some consistency in the estimates from the different methodologies.

Table 2.2: Correlations between art price indices (in growth rates)

	Median	Fisher	Hedonic	Adjacent-1y	Adjacent-2y	Rolling-5y	Repeat Sales	ps-RS1
Fisher	0.13							
Hedonic	0.03	0.38***						
Adjacent-1y	0.09	0.32**	0.90***					
Adjacent-2y	0.10	0.34***	0.95***	0.98***				
Rolling-5y	0.22*	0.35***	0.94***	0.94***	0.95***			
Repeat Sales	0.33**	0.00	-0.12	-0.08	-0.03	-0.09		
ps-RS1	0.06	0.13	0.49***	0.60***	0.59***	0.51***	0.17	
ps-RS2	0.05	0.13	0.49***	0.60***	0.60***	0.50***	0.34***	0.91***

The fact that the regression-based indices are similar, even when the hybrid repeat sales indices are based on smaller subsamples of the data, implies that the potential omitted variable and sample selection bias may not be too pervasive in this case. The ps-RS method acts as an internal validity test, to check that the results are not driven by the inherent biases of a specific method.

The fairly consistent picture offers some confidence that the indices provide a reasonably accurate measure of the price movements in the South African art market. The large increase in art prices between 2005 and 2008 does not seem to be due to a fundamental shift in the types of artworks that were sold over that period. For instance, the top 100 artists in terms of volumes sold, which accounts for 60% of the volume traded and 90% of total turnover, remained remarkably stable over time. Even if the exact same artworks were not being resold, the same artists' work still made up the majority of the market, and the hedonic model controls for the different artists. It is unlikely that the results are driven by sales of systematically better or higher quality artworks by specific artists that appreciated in price before the crisis, and by sales of systematically lower quality artworks by those artists after the crisis. Moreover, paintings are not sold at auction only to profit from price appreciation, or capital gains. A substantial portion of consignments come from the so-called three D's: Debt, Divorce and Death. In other words, many sellers are forced to sell their artworks, even if those artworks have not experienced the largest price appreciation.

<sup>17</sup>The first few periods of repeat sales estimates are often sensitive when the sample size is small, because of the lack of repeat sales in the first few quarters. In this case, very few artworks had been resold in the few quarters, making the index values very volatile. Indeed, there are no true repeat sales in the first three quarters of the sample period. Therefore, the first three index values were excluded from the comparison.

### 2.5.2 External Comparison: Traditional South African Assets

As there are no South African art prices indices available for an external validity check, this section compares the art price indices to traditional South African asset prices. Collectors broadly have three main motives for collecting art: a genuine love of art, social promise and investment possibilities (Findlay, 2012). The first motive relates to the essential (or intrinsic) value of art, sometimes described as ‘aesthetic pleasure’ (Gérard-Varet, 1995). The second relates to the social value or status consumption of art. To the extent that art is a luxury good, collectors may derive utility from the signal of wealth that it conveys (Mandel, 2009). The third motive involves commercial value and relates to the investment possibilities of art as an alternative asset class.

The vast majority of the literature on art price indices focuses on the third motive (e.g. Mei and Moses (2002); Renneboog and Van Houtte (2002); and Kräussl and Van Elsland (2008)). These studies typically analyse the risk-adjusted returns and diversification potential of art as an asset class, to investigate whether art would have formed part of an optimal investment portfolio. Naturally, the results depend on the specific markets and periods under consideration. The majority of studies seems to find at least positive returns to art, although volatility is usually relatively high, and under certain assumptions, some diversification potential (Goetzmann, Mamonova and Spaenjers, 2014).

For instance, Candela and Scorcu (1997) found that Italian art prices increased in line with inflation, but that returns were lower than for traditional assets. Mei and Moses (2002) found that art was less volatile and exhibited a low correlation with traditional assets, which made it attractive for diversification purposes. Renneboog and Van Houtte (2002) found limited diversification potential for German art, and only for investors willing to incur a substantial amount of risk. Kräussl and Van Elsland (2008) concluded German art earned low risk-adjusted returns and was not proportionally rewarded for bearing downside risk. However, art was included in an optimal portfolio under certain assumptions. Campbell (2009) found that the art market was negatively correlated with traditional assets, making art beneficial for portfolio diversification. Kräussl and Logher (2010) analysed the investment characteristics of art in Russia, China, and India. The results varied widely per market, and although the returns were all positive, there was limited diversification potential. Renneboog and Spaenjers (2013) found that the risk-return profile of art was inferior to financial assets, even before transaction costs. However, art did outperform other alternative assets, such as gold, commodities, and real estate. Korteweg (2013) concluded that investors would not include art in a portfolio, after correcting for sample selection bias.

Figure 2.17 provides a comparison of the real South African art price index with real indices for traditional South African assets: the JSE All Share Index, the All Bond Index, and the ABSA House Price Index. The comparison shows that art prices experienced similar rapid price appreciation between 2005 and 2008 to equity prices, and much higher returns than the other assets. This might suggest that art was a relatively good investment over this period of rapid growth. After 2009, however, the art price indices lagged substantially behind the other asset price indices. The failure

## Art Prices

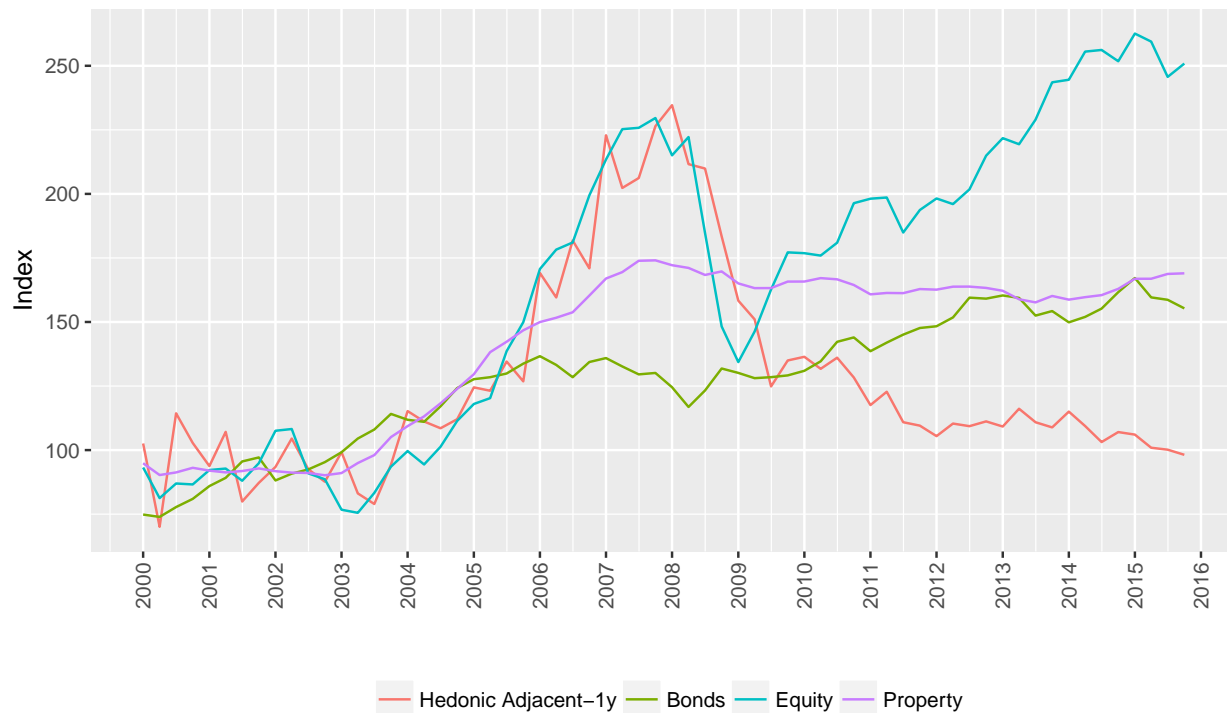


Figure 2.17: Comparing South African real asset price indices (2001=100)

of art prices to show any real growth, or keep pace with other asset prices, after the financial crisis, implies that art was not a good investment over the latter part of the sample period. According to these index estimates, therefore, an investment in a ‘standard artwork’ or the art market in general would not have yielded comparable returns to other investments.

One would not necessarily expect art to provide comparable risk-adjusted returns to traditional assets. One way in which artworks differ from stocks and bonds is that they are also consumer goods that provide direct consumption utility in the form of aesthetic pleasure and social status. Hence, these assets would not necessarily be expected to provide as high investment returns as stocks and bonds, as investors will be further compensated by non-financial ‘aesthetic’ or utility dividends.

The large transaction costs involved in the art market also make it less appealing as an investment. For instance, Strauss & Co charge around 15% buyer’s premium in addition to a seller’s premium, while Bonhams charges premiums of as much as 25% on South African art sales (Olckers, Kannemeyer and Stevenson, 2015). In contrast, stock market transactions for an individual investor may cost around 1.5% for domestic shares and 2.5% for foreign shares. In some cases the insurance premiums associated with artworks can also be relatively large. This implies that one should maximise the consumption utility derived from these assets, so that the non-monetary benefits compensate for the lack of investment returns (Renneboog and Van Houtte, 2002). Indeed, Korteweg (2013) argues that art is primarily an aesthetic investment, not a financial one.



Table 2.3: Correlations between real asset returns

	SA Art: Adjacent-1y	SA Art: pseudo-RS2
SA Bonds	0.02	-0.05
SA Equity	0.38***	0.39***
SA Property	0.40***	0.39***

Table 2.3 reports the correlations in the real returns of the assets. The correlations between the art price indices and the equity and property indices are positive and significant, although the coefficients are relatively low. The significant positive correlations, especially up to the Great Recession, imply that the diversification potential of art as an investment is limited in this case. Botha, Snowball and Scott (2016) investigated the diversification potential of South African art, using VAR models and the CAPI as a measure of South African art prices. They found a significant positive relationship between art prices and lagged share prices, but not bonds or real estate. From this they concluded that the South African art market does not offer the opportunity to diversify investment portfolios. Further research applications might use the new art price indices, and a portfolio optimisation model, to assess the potential diversification benefits of art investments more directly than in Botha, Snowball and Scott (2016).

A caveat of this type of analysis is that it is not currently possible to invest in a representative art price index in the same way as it is possible to invest in an index that replicates the performance of the JSE. In addition, it would be difficult and expensive for an investor to collect a representative portfolio of South African art to replicate the performance of the overall market (as the underlying assets are unique). Indeed, perhaps only a handful of serious collectors would be able to build a representative portfolio of the South African art that is measured in this chapter. It would probably be more practically relevant to evaluate smaller segments of the market (e.g. paintings by specific artists).

Although art might have provided lower returns over the whole sample period, investing in art may offer an interesting alternative when stocks perform below average, i.e. there may be a hedging opportunity (Kräussl and Lee, 2010). For instance, if art prices did not collapse (and only slowed down) during the financial crisis, it is possible that South African art was a relatively ‘safe’ investment. During periods of economic turmoil, investors may prefer to invest in physical assets, which are often considered relatively safe due to their limited supply, in order to protect themselves from crashing markets (Kräussl and Logher, 2010).

A few studies have calculated ‘downside beta’ to analyse the downside risk characteristics of investments in art (e.g. Kräussl and Logher (2010)). Downside beta is calculated by looking at covariance with overall market return, given that the market return is below its average. Assets that have a high downside beta have a high correlation with the market when it declines and do not offer downside protection (Kräussl, 2015). Again, the results depend on the markets and the periods considered. In analysing three large emerging art markets, Kräussl and Logher (2010) found that only Chinese art offered hedging potential during financial market downswings. This was because

Chinese art did not decline as strongly as the stock market, and because of its return characteristics, such as its negative pairwise correlation with many other asset classes. Based on their top 500 Art Market index, Kräussl and Lee (2010) found that art was positively correlated with the global equity index, which meant that it was not a good hedging instrument, although investing in art did lead to diversification benefits under certain assumptions. For the South African art market, Fedderke & Li (2014) found that the JSE was negatively associated with high-end art prices, which they argued might imply that the high-end art market served as a form of risk diversification, increasing in importance as returns on domestic financial markets declined.

In this case, the sample period includes only the single full downswing phase during the Great Recession (i.e. too few to conduct a proper analysis). Figure 2.17 illustrates that the decrease in art prices during the Great Recession was dramatic: as marked as the decrease in equity prices and more so than the declines in the other asset prices. For instance, the 1-year adjacent-period hedonic index decrease by around 40% in real terms between its peak in 2008Q1 and 2010Q1. In the same period, the equity index decreased by around 18% in real terms, while bonds and property prices were relatively flat. Thus, it does not seem as if South African art was a particularly 'safe' store of value during the crisis period, nor that it provided protection during this time of economic turmoil. Moreover, if the wealth effect is an important driver of art prices, as suggested by the positive correlations with equity and property prices (which are typically held by higher parts of the wealth distribution), it suggests collectors did not swap to art as a safe haven during the period of economic turmoil.

Investment in art is generally concentrated in the top of the income distribution. There is some evidence that the buying power of the affluent is an important driver of art prices. For example, Goetzmann, Renneboog and Spaenjers (2011) found that equity markets and top incomes (with income inequality as proxy) had a significant impact on art prices. Renneboog and Spaenjers (2013) found evidence that equity market returns and changes in high-income consumer confidence levels predicted art returns. To the extent that investments in South African art are concentrated in the top wealth brackets, high risk-adjusted returns for this passive investment (even for only short periods) would benefit only those collectors, and would widen income inequality over time. In this case, the large appreciation between 2005 and 2008 may well have increased wealth inequality during this period. However, the large decrease and the flat real returns after 2009 may well have reversed the increase. The art price indices created in this chapter could be used to investigate these potential drivers and developments in future.

### **2.5.3 Evaluation: Smoothness**

In this section, the indices produced with the different methodologies are directly evaluated, in order to assess which index provides the most accurate measure of South African art prices over time. This is not usually attempted for art price indices, given that most papers focus on a specific

method. In other applications, the quality of price indices has often been evaluated based on the diagnostic metrics of the underlying regressions, such as the standard errors of the residuals (see e.g. Hansen (2009)).

However, Guo *et al.* (2014) argued that the regression residuals do not reflect errors in the price index itself, and hence do not directly reflect inaccuracy in the index returns. Even if an index is a perfectly accurate measure of the central tendency of price changes, the regression would still have residuals, and the time dummy coefficients may still exhibit large standard errors, resulting simply from the dispersion of individual art prices around the central tendency. When datasets become large, regression diagnostics can become impressive simply due to the size of the sample. In such cases, tests of economic significance are more valuable than tests of statistical significance. In this case, not all of the indices were generated with regression models. The regressions models that were employed differ in their specifications (in levels or first differences) and the underlying data sets used for estimation.

Guo *et al.* (2014) suggested that signal-to-noise or smoothness metrics are more appropriate for judging the quality of the price index, as they are based directly on the index, as opposed to the underlying model. Signal-to-noise metrics directly reflect the accuracy of the index returns and the economic significance of random error in the indices. Random error in the coefficient estimation leads to ‘noise’ in the index. The volatility and the first-order autocorrelations of the index returns are signal-to-noise metrics that may be useful in comparing the amount of noise in the indices, as is demonstrated in the following simple model:

$$m_t = m_{t-1} + r_t$$

and

$$I_t = m_t + \epsilon_t = \sum_{i=1}^t r_i + \epsilon_t,$$

where  $m_t$  is the true market value level (in logs);  $r_t$  is the true return (i.e. the central tendency) of market prices in period  $t$ ;  $I_t$  is the index in period  $t$ ;  $\epsilon_t$  is the index-level random (white noise) error. This random error causes noise in the index and is therefore important for users of the index. The noise does not accumulate over time.

The index returns can be defined as follows:

$$r_t^* = I_t - I_{t-1} = r_t + (\epsilon_t - \epsilon_{t-1}) = r_t + \eta_t,$$

where  $r_t^*$  is the index return and  $\eta_t$  is the noise in the index return in period  $t$ .

The volatility of the index  $Vol$ , which is the standard deviation of the index return  $\sigma_{r_t^*}$ , and the first-order autocorrelation of the index return  $\rho_{r^*}$ , can be derived as:

$$Vol = \sigma_{r_t^*} = \sqrt{\sigma_r^2 + \sigma_\eta^2}$$

$$AC(1) = \rho_{r^*} = (\rho_r \sigma_r^2 - \sigma_\eta^2/2)/(\sigma_r^2 + \sigma_\eta^2)$$

where  $\sigma_r^2$  and  $\sigma_\eta^2$  are the variances of the true return and the noise, and  $\rho_r$  is the first-order autocorrelation of the true return.

Volatility is a measure of the dispersion in returns over time. There is always true volatility, as the true market prices change over time. An ideal price index will filter the volatility induced by the noise to leave only the true market volatility. In addition to the true volatility, the noise in the index causes excess volatility in the index returns. Excess volatility decreases the first-order autocorrelation of the index returns. Less noise (lower  $\sigma_\eta^2$ ) will lead to lower index volatility and higher autocorrelation. Other things being equal, the lower the volatility and the higher the autocorrelation, the less noisy and more accurate the index. Thus, lower *Vol* or higher *AC*(1) will indicate a more accurate art price index in the sense of less noise or random error.

Guo *et al.* (2014) suggested another test of index quality in terms of minimising random error that is based on the Hodrick-Prescott (HP) filter. The HP filter is a spline-fitting procedure that divides a time series into smoothed trend and cyclical components. The idea is to examine which index has the least deviation from its smoothed HP component, by comparing the sum of squared differences between the index returns and the smoothed returns.

Another option is to compare the smoothness coefficients proposed by Froeb and Koyak (1994). The smoothness coefficient is defined as the average long-run variance of a time series divided by the average short-run variance. The idea is to obtain the spectral density of the time series, which shows the contribution of all frequencies to the data series. The smoothness measure is then taken as the average of the lower half of the frequency range (i.e. the low-frequency or longer-term movements) over the average of the upper half of the frequencies (i.e. the high-frequencies or shorter-term movements). In other words, the smoothness coefficient is the low-frequency portion divided by the high-frequency portion of the periodogram.<sup>18</sup> A higher smoothness coefficient indicates a larger share of variance in the low frequencies and a smoother time series.

Table 2.4 reports these four metrics of index smoothness for the art price indices. The comparison suggests that the regression-based indices are much smoother than the central tendency measures and the classical repeat sales index. The volatilities, autocorrelations and HP filter deviations of the regression-based indices are around the same size. The 1-year adjacent-period hedonic index performs the best in terms of these metrics, with the lowest volatility *Vol* and highest autocorrelation *AC*(1) in returns, the smallest deviation from its smoothed returns, and the highest smoothness coefficient. However, the smoothness coefficients of the regression-based indices are not significantly different.

The following section provides a study of the indices to test for evidence of a bubble in South African art prices. In answering this question, the chapter turns to the literature on bubble detection.

<sup>18</sup>The spectral density is smoothed using the Daniell window, which amounts to a simple moving average transformation of the periodogram values.

Table 2.4: Smoothness indicators

	Volatility	AC(1)	HP-Deviation	Smoothness
Median Index	0.612	-0.416	22.11	-0.02
Fisher Index	0.284	-0.332	4.66	1.00
Hedonic Index	0.114	-0.323	0.75	1.09
Adjacent-1y	0.105	-0.246	0.63	1.39
Adjacent-2y	0.105	-0.303	0.63	1.10
Rolling-5y	0.112	-0.279	0.73	1.27
Repeat Sales	0.549	-0.403	17.76	0.65
pseudo-RS1	0.128	-0.360	0.95	0.94
pseudo-RS2	0.123	-0.342	0.87	1.16

## 2.6 Bubble Detection

Record prices for South African artworks at local and international auctions, especially between 2008 and 2011, prompted many commentators at the time to claim that the market was overheating and suggest the possibility of a ‘bubble’ in the market (e.g. Rabe (2011); Hundt (2010); Curnow (2010)). According to the indices generated above, there was a substantial increase in South African art prices in the run-up to the Great Recession. This section uses the art price indices to investigate whether art prices exhibited bubble-like behaviour over the sample period.

Both emerging and advanced economies suffered severe financial crises around 2008. Yiu, Yu and Jin (2013) argued that these crises were triggered by the collapse of bubbles in asset prices. The adverse effects of bubbles and their related crises have led to a large literature on financial crises and the detection of bubbles in asset prices, including the seminal work by Kindleberger and Aliber (2005) and the modelling approach by Phillips, Wu and Yu (2011).

The starting point is the definition of the term ‘bubble’. Stiglitz (1990) provided the following popular definition: “[I]f the reason the price is high today is only because investors believe that the selling price will be high tomorrow - when ‘fundamental’ factors do not seem to justify such a price - then a bubble exists. At least in the short run, the high price of the asset is merited, because it yields a return (capital gain plus dividend) equal to that on alternative assets.”

Case and Shiller (2003) defined the term as “a situation in which excessive public expectations of future price increases cause prices to be temporarily elevated.” According to the *New Palgrave Dictionary of Economics*, “bubbles refer to asset prices that exceed an asset’s fundamental value because current owners believe they can resell the asset at an even higher price” (Brunnermeier, 2008). These definitions imply that the main features of a bubble are that prices increase above levels that are consistent with underlying fundamentals, and that buyers expect excessive future price increases. In other words, a bubble consists of a sharp rise in a given asset price, beyond a level sustainable by fundamentals, followed by a sudden collapse (Kräussl, Lehnert and Martelin, 2016).

When it comes to the art market, however, it is particularly challenging to determine the fundamental value from which prices potentially deviate. In the case of stocks, dividends have been used to

obtain the expected cash flow as a measure of fundamental value. Rents and convenience yields can potentially be used for real estate prices and commodity prices (Penasse and Renneboog, 2017).

In contrast, artworks do not generate a future income stream (e.g. dividends or rents) that can be discounted to determine the fundamental value. Artworks usually have little intrinsic value, unless the materials used have a high value (Spaenjers, Goetzmann and Mamonova, 2015). Instead, artworks are acquired for a kind of non-monetary utility or aesthetic dividend. This dividend can be interpreted as the rent buyers are willing to pay to own the artwork over a given period. The rent can reflect aesthetic pleasure, but may also provide additional utility as a signal of wealth (Mandel, 2009). The price of an artwork should equal the present value of future private dividends over the holding period, and the expected resale value. The value of the dividend is unobservable and is likely to vary greatly among buyers, based on their motivations and characteristics (Penasse and Renneboog, 2017). Thus, it is almost impossible to determine the fundamental value of art (Kräussl, Lehnert and Martelin, 2016).

To overcome this issue, this section follows Kräussl, Lehnert and Martelin (2016) in using a direct bubble detection method developed by Phillips, Wu and Yu (2011). The method is based on a right-tailed augmented Dickey-Fuller (ADF) unit root test, which is able to detect explosive behaviour in time series. Phillips, Wu and Yu (2011) originally applied the method to stock prices. They showed that there was evidence of explosiveness in stock prices, but not in dividend yields, implying that price explosiveness could not be explained by developments in fundamentals.

Since then, various studies have used the method to investigate bubbles in a number of asset markets, including real estate, commodities and art. Jiang, Phillips and Yu (2015) employed the method to identify explosive periods in real estate prices in Singapore. The results suggested an explosive period from 2006Q4 to 2008Q1. Balcilar, Katzke and Gupta (2015) used the method to detect explosive periods in US real estate prices for the period 1830-2013 and found evidence of several bubble periods. Areal, Balcombe and Rapsomanikis (2013) used the method to test for periods of explosive prices in agricultural markets and found that bubbles occurred for certain commodities, especially around 2007 and 2008. Figuerola-Ferretti, Gilbert and McCorrie (2015) applied the method to examine non-ferrous metals futures prices on the London Metal Exchange. They found that certain commodity futures markets exhibited bubble-like behaviour, with the majority of the bubble periods occurring between August 2007 and July 2008.

In the context of art, Kräussl, Lehnert and Martelin (2016) used the method to detect explosive behaviour in the prices of four art market segments ('Post-war and Contemporary', 'Impressionist and Modern', 'American' and 'Latin American'). They found evidence of explosive behaviour in prices and identified historical bubble episodes in the 'Post-war and Contemporary' and 'American' art market segments, around 2006-2008 and 2005-2008 respectively. The following section sets out the bubble detection framework used to test for the presence of bubble-like behaviour in South African art prices over the sample period.

### 2.6.1 Bubble Detection Framework

The most common bubble detection methods are based on the present value model and a so-called rational bubble assumption (Yiu, Yu and Jin, 2013). According to the present value model, under rational expectations, the price of an asset is equal to the present value of its future income stream, i.e. the expected fundamental value:

$$P_t = \frac{1}{1 + r_f} E_t(P_{t+1} + \gamma_{t+1}),$$

where  $r_f$  is the constant discount rate,  $P_t$  is the asset price at time  $t$ , and  $\gamma_{t+1}$  is the payment received (e.g. dividends, rents or a convenience yield) for owning the asset between period  $t$  and  $t + 1$ . When  $t + n$  is far into the future,  $\frac{1}{1+r_f} E_t(P_{t+n})$  does not affect  $P_t$ , as it tends to zero as  $n$  becomes infinitely large. The present value or market fundamental solution may be written as:

$$F_t = E_t\left[\sum_{i=1}^n \frac{1}{1 + r_f} (\gamma_{t+i})\right]$$

Bubbles occur when buyers are willing to pay more for an asset than the fundamental value, as they expect the future price to be higher than its fundamental value (Yiu, Yu and Jin, 2013). If a bubble occurs, the asset price consists of a fundamental component and a bubble component. In other words, if there is a gap between the fundamental value and the actual price, an additional ‘bubble component’,  $B_t$ , is added to the solution of equation:  $P_t = F_t + B_t$ . In this case  $F_t$  is called the fundamental component of the price, and  $B_t$  is a random variable of the following form:

$$B_t = \frac{1}{1 + r_f} E_t(B_{t+n})$$

Thus, the bubble component is included in the equation, with an expected value in period  $t + 1$  of  $B_t$  multiplied by  $(1 + r_f)$ . The bubble component is called a ‘rational bubble’, as it is in line with the rational expectations framework (Kräussl, Lehnert and Martelin, 2016).

The statistical properties of  $P_t$  are determined by those of  $F_t$  and  $B_t$ . In the absence of a bubble, when  $B_t = 0$ , the degree of non-stationarity in  $P_t$  is determined by the series  $F_t$ , which in turn is determined by  $\gamma_t$ . The current price of the asset is therefore determined by market fundamentals: for example, if  $\gamma_t$  is an I(1) process, then  $P_t$  would be an I(1) process.

When a bubble is present, if  $B_t \neq 0$ , current prices  $P_t$  will exhibit explosive behaviour, as  $B_t$  is a stochastic process for which the expected value in period  $t + 1$  is greater than or equal to the value in period  $t$  (Kräussl, Lehnert and Martelin, 2016). If there is no structural change in the fundamental process or explosiveness in the fundamentals, a period of explosive prices has a non-fundamental explanation. Mildly explosive behaviour in  $P_t$  (i.e. non-stationarity greater than a unit root) provides evidence of bubble-like behaviour. According to this theory, if a bubble exists,



prices should inherit its explosive property (Areal, Balcombe and Rapsomanikis, 2013). Statistical tests can be formulated to detect evidence of explosiveness in the price series (Caspi, 2013).

Early tests were based on unit root and cointegration tests. Campbell and Shiller (1987) suggested a unit root test for explosiveness in prices, based on the idea that during the process of bubble formation, the gap between the asset price and the fundamental value will exhibit explosive behaviour. They identified two scenarios in which the presence of a bubble is implied. In the first case, the asset price is non-stationary while the fundamental value is stationary. In the second, the asset price and fundamental value are both non-stationary (Yiu, Yu and Jin, 2013). In this case, if the asset price and its fundamental value are not cointegrated, their non-stationary behaviour provides evidence of a bubble. Diba and Grossman (1988) showed that explosive behaviour in prices is a sufficient condition for the presence of a bubble, if the fundamental value is not explosive.

However, unit root and cointegration tests are incapable of detecting explosive prices when a series contains periodically collapsing bubbles. Evans (1991) argued that explosive behaviour is only temporary, as bubbles eventually collapse, which means that explosive asset prices may appear more like stationary or  $I(1)$  series. Using simulated data, Evans (1991) showed that bubble detection tests could not differentiate between a stationary process and a periodically collapsing bubble. A series with periodically collapsing bubbles could therefore be interpreted by the standard unit root tests as a stationary series, leading to the incorrect conclusion that the series contained no explosive behaviour (Phillips, Wu and Yu, 2011).

A number of methods have been proposed to deal with this critique (Yiu, Yu and Jin, 2013). The recursive tests proposed by Phillips, Wu and Yu (2011) and Phillips, Shi and Yu (2015) are not subject to this criticism and can effectively distinguish between unit root processes and periodically collapsing bubbles, as well as identify the dates of their origin and collapse. The test proposed by Phillips, Wu and Yu (2011) involves repeatedly implementing a right-tailed unit root test, by estimating an autoregressive model, starting with a minimum sample window size and incrementally expanding the window.

The model typically takes the following form:

$$\Delta y_t = \alpha_w + (\delta_w - 1)y_{t-1} + \sum_{i=1}^k \phi_w^i \Delta y_{t-i} + \epsilon_t,$$

where  $y_t$  is the asset price series,  $\alpha$ ,  $\delta$  and  $\phi$  are the parameters to be estimated,  $w$  is the sample window size,  $k$  is the lag order, and  $\epsilon_t$  is the white noise error term.

Augmented Dickey-Fuller test statistics are calculated from each regression. The null hypothesis of a unit root ( $\delta = 1$ ) is tested against the right-tailed alternative of mildly explosive behaviour ( $\delta > 1$ ). The supremum value of the ADF sequence is then used to test for mildly explosive behaviour. By testing directly for explosive behaviour, the test avoids the risk of misinterpreting a rejection of the null hypothesis due to stationary behaviour.



The method also allows for date-stamping of the origination and termination dates, by comparing the time series of the test statistics with the critical value sequence. In other words, to identify a bubble period, each ADF test statistic is compared with the corresponding right-tailed critical value. The origination point of a bubble is the first observation in which ADF statistic exceeds the corresponding critical value (from below), while the termination point is the first subsequent observation when the ADF statistic falls below the critical value (Caspi, 2013).

A limitation of the method is that it is designed to analyse a single bubble period. Phillips, Shi and Yu (2015) expanded the method to account for the possibility of multiple bubbles, by varying both the starting and ending points of the sample windows. The moving window provides greater flexibility in choosing a subsample that contains a bubble (Yiu, Yu and Jin, 2013). Thus, the method of Phillips, Wu and Yu (2011) is consistent and particularly effective when there is a single bubble period, while the method of Phillips, Shi and Yu (2015) can identify multiple bubble periods. Simulations by Homm and Breitung (2012) indicated that the test worked adequately against other time series bubble detection tests and was particularly effective for real-time bubble detection.

### **2.6.2 Bubble Detection Results**

In this section the South African art market is tested for bubble-like behaviour over the sample period, focusing on a specific aspect of bubbles: explosive prices. This section follows the convention of using the log value of real asset prices, deflated with the CPI (e.g. Kräussl, Lehnert and Martelin (2016), Caspi (2013) and Balcilar, Katzke and Gupta (2015)). In this case, there was only one potential bubble episode, so the Phillips, Wu and Yu (2011) method is sufficient to provide evidence of mildly explosive behaviour.

As explained above, the method involves estimating an autoregressive model, starting with a minimum fraction of the sample and incrementally expanding the sample window. The model starts with 3 years (i.e. 12 observations) and expands the sample by 1 observation each time. Each estimation yields an ADF statistic. In this case, there does not seem to be a deterministic drift present in the log real art price indices, and the intercept is not statistically significant at conventional levels. However, as the results could be sensitive to model formulation, two versions of the autoregressive models are used: one without a constant or drift term and one with a drift term. Lags are included to take possible autocorrelation of the residuals into account, and the number of lags  $k$  is chosen with the Akaike Information Criterion.

In this section, critical values for the tests are derived from Monte Carlo simulations of a random walk process, both including and excluding a drift term, with 2000 replications. In their original study, Phillips, Wu and Yu (2011) used a random walk without drift to estimate the null hypothesis. According to Phillips, Shi and Yu (2014), when the model is estimated with a non-zero drift, it produces a dominating deterministic component that is unrealistic for most economic and financial time series. They argued that a more realistic description of explosive behaviour is given by models

formulated without a constant or deterministic trend. Nevertheless, as a robustness check, the models are formulated with and without a constant, or drift term, included.<sup>19</sup>

The supremum ADF test statistics from both formulations are above the 95% critical values for each of the indices, except for the median index. Therefore, the null hypothesis of a unit root may be rejected in favour of the alternative hypothesis for each of the indices, except the median index. This provides evidence that real art prices experienced periods of explosiveness over the sample period.

The method can be used to date stamp potential bubble periods. Figure 2.18 and Figure 2.19 illustrate the date stamping procedure for three representative series: median values, the 1-year adjacent-period hedonic index and the second version (larger sample) of the ps-RS index. Figure 2.18 illustrates the case of no drift term, while Figure 2.19 illustrates the case with a drift term. The figures compare the ADF test statistic sequence to the 95% and 99% critical value sequences. In both cases the test statistic sequences breach the 95% critical values in the run-up to the financial crisis (2005 and 2006 respectively), before falling below the critical values. The sequence of test statistics for the ps-RS index is higher than for the 1-year adjacent-period hedonic index, and breaches the 99% critical value.

Table 2.5 reports the origination and termination dates for the periods of explosive behaviour, based on 95% critical values. The test statistic sequences for the hedonic indices all indicate a period of explosive prices beginning around 2006/2007 and ending in 2008. The test statistics for the ps-RS indices indicate periods of explosive behaviour that were slightly longer, beginning around 2005/2006 and ending in 2008 or even 2010, depending on the specification. The preferred index in terms of smoothness (i.e. the 1-year adjacent-period index) suggests a period of bubble formation from 2007Q1 to 2008Q3. Phillips, Shi and Yu (2015) recommend that only explosive periods lasting more than  $\log(T)$  units of time should be identified as bubble periods. In this case it implies that the bubble should be at least four quarters in length, and all of the explosive periods identified satisfy this requirement.

The dates identified correspond with many of the explosive periods identified in the literature for a range of assets. In the context of art, Kräussl, Lehnert and Martelin (2016) identified bubble periods for the ‘Post-war and Contemporary’ art segment between 2006 and 2008 and for the ‘American’ art segments between 2005 and 2008, which also corresponds to the pre-financial crisis period. Interestingly, their findings point to evidence in the formation of another bubble in these market segments around the start of 2011. This is not present in the South African art market, which has

<sup>19</sup>Phillips et al. (2014) suggested a random walk process with an asymptotically negligible drift might be useful for allowing for intermediate cases between a model with no drift, and one with a drift term included, i.e. cases where there may be drift in the data but where it may not be the dominant component. Such a model may take the following form:  $y_t = dT^{-\eta} + \theta y_{t-1} + \epsilon_t$ , where  $d$ ,  $\eta$  and  $\theta$  are constant,  $T$  is the sample size, and  $\epsilon_t$  is the white noise error term. The deterministic component depends on the sample size  $T$  and the localising parameter  $\eta$ . When  $\eta > 0$ , the drift term is small relative to the linear trend. The model becomes a model without drift when  $\eta \rightarrow \infty$  and a model with drift when  $\eta \rightarrow 0$ . The results may be sensitive to the value of  $\eta$ , so Phillips et al. (2014) recommend reporting the results for a range of values of  $\eta$ . In this case, different values of  $\eta$  produce qualitatively similar results.

Art Prices

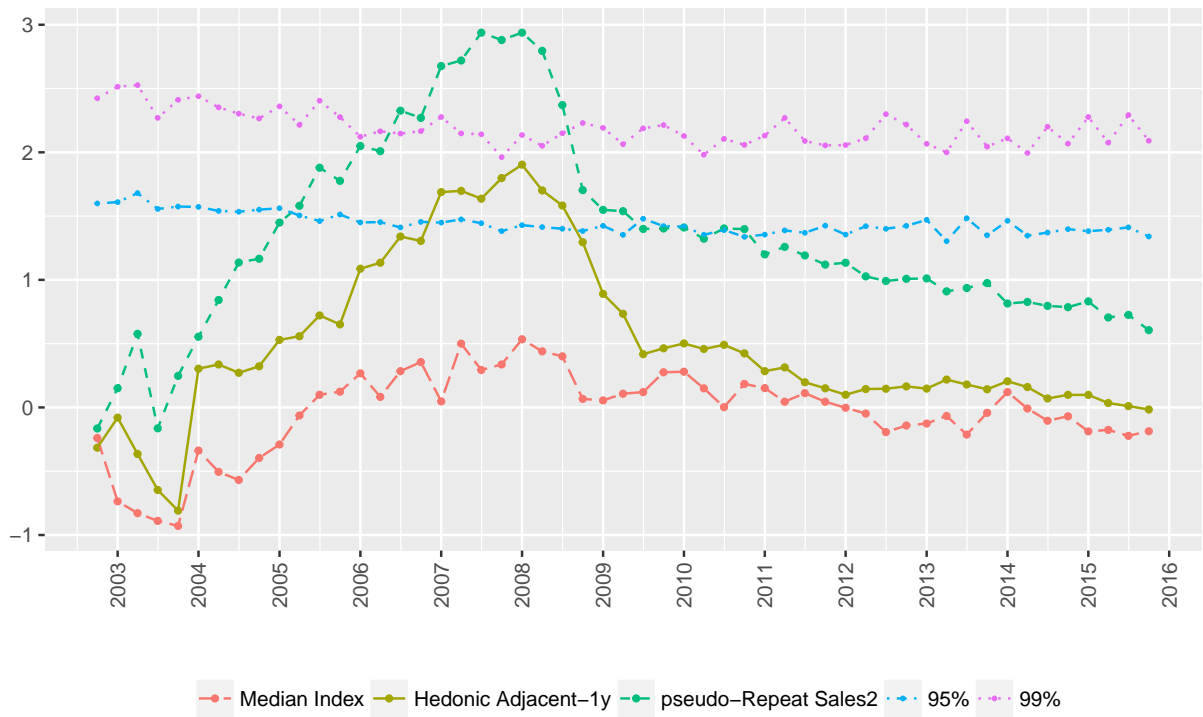


Figure 2.18: Test statistics and critical values for models without drift

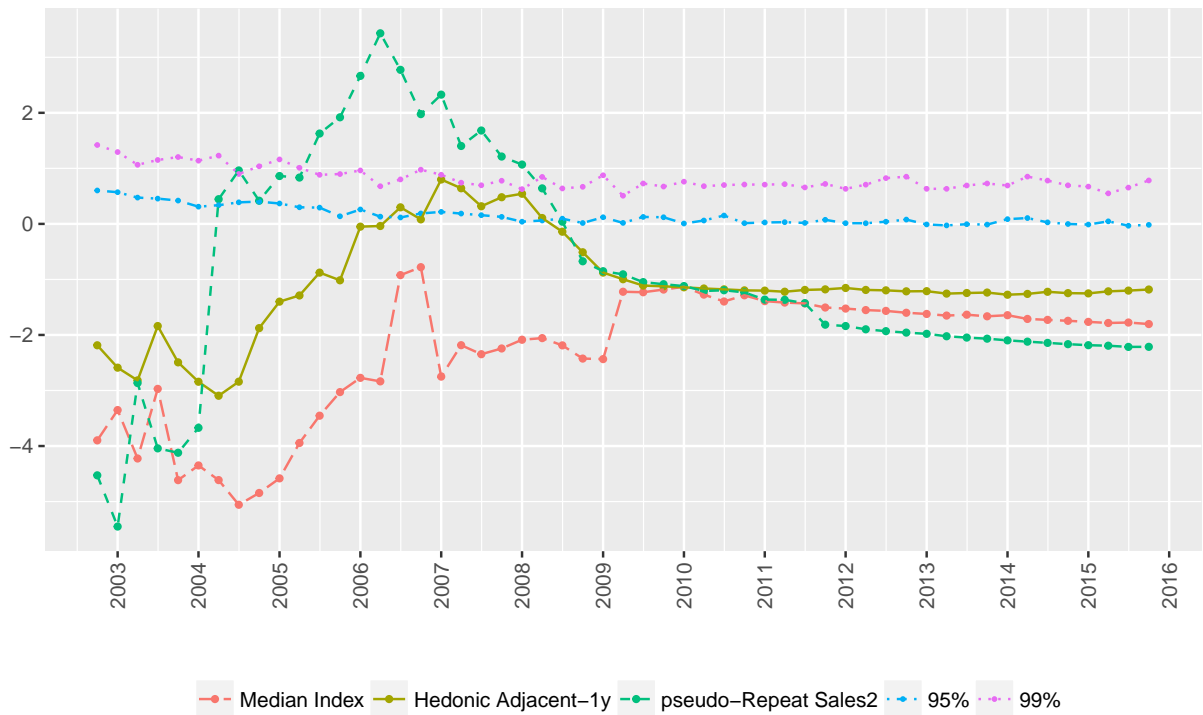


Figure 2.19: Test statistics and critical values for models with drift

Table 2.5: Dates of explosive price behaviour

	No Drift		Drift	
	Start	End	Start	End
Hedonic Full	2007 Q1	2008 Q3	2007 Q1	2008 Q2
Adjacent-1y	2007 Q1	2008 Q3	2006 Q3	2008 Q2
Adjacent-2y	2007 Q1	2008 Q4	2006 Q3	2008 Q2
Rolling-5y	2007 Q2	2008 Q3	2007 Q1	2008 Q2
ps-RS1	2006 Q1	2010 Q1	2006 Q1	2008 Q2
ps-RS2	2005 Q2	2010 Q4	2004 Q2	2008 Q2

remained flat since 2009.

It is also interesting that many of the headline-grabbing auction records for the South African art market occurred in 2011, well after the period of explosive behaviour. This corresponds with findings by Spaenjers, Goetzmann and Mamonova (2015), who observed that the timing of record prices does not always coincide with periods of general price increases. They argued that auction price records often occur in situations characterised by extreme supply constraints, resolution of uncertainty about the potential resale value, social competition among ‘nouveaux riches’, and idiosyncratic shifts in hedonic weights.

Because the art price indices are based on the coefficients from regression models, each index point has a statistical distribution, with standard errors and implied confidence intervals (as illustrated in the Appendix). The standard errors of the coefficients provide an idea of the distribution of the coefficient estimates and the uncertainty surrounding each of the index points. These standard errors may not be constant between index points, which means that a type of heteroskedasticity may be present. The statistical inference in the bubble detection tests should take this potential heteroskedasticity into account.

In the chapter Appendix, the bubble detection tests are estimated after weighting each observation by the inverse of its standard error, similar to a Feasible Generalised Least Squares estimation. The weights are used to indicate that different observations have different standard errors. This places more emphasis on the segments of the data which are estimated more accurately and makes an adjustment for the potential heteroskedasticity between the index points. The bubble detection test results are similar to those reported in Table 2.5, which implies that the inference is acceptable in this case.

### 2.6.3 Market Segments

In this section the price indices for the different segments of the art market are examined, in order to establish how widely dispersed the bubble process was and which segments were responsible for the explosive price increases that occurred. The bubble detection tests were performed on the market segment indices defined above in terms of price, medium and artist value. Again, slicing the data thinly results in small sample sizes (with larger standard errors) and more volatile indices.

## Art Prices

Table 2.6: Dates of explosive price behaviour in the different market segments

	No Drift		Drift	
	Start	End	Start	End
Price Distribution:				
Lower Segment				
Middle Segment	2007 Q2	2008 Q3	2007 Q2	2007 Q4
Upper Segment				
Quantile Regressions:				
Lower (tau=0.25)				
Middle (tau=0.70)	2007 Q3	2008 Q4	2007 Q3	2008 Q2
Upper (tau=0.95)	2007 Q2	2007 Q3	2006 Q4	2008 Q2
Artist Value:				
Lower Segment				
Middle Segment	2007 Q2	2008 Q4	2006 Q3	2008 Q2
Upper Segment	2007 Q1	2008 Q4	2007 Q2	2008 Q2
Mediums:				
Drawing				
Watercolour	2007 Q4	2008 Q3	2007 Q4	2008 Q2
Oil	2006 Q4	2009 Q3	2006 Q3	2008 Q3
Print/Woodcut				
Mixed Media				

This makes it more difficult to discern a pattern and to distinguish the signal from the noise.

The results for the origination and termination dates are reported in Table 2.6. In terms of the price segments, the indices for the middle and upper segments of the price distribution exhibit evidence of explosive behaviour. For the hedonic price indices, the middle segment of the distribution (artworks between R6,500 and R70,000) shows evidence of a brief explosive period. From the quantile regression, the indices for the 70th and 95th percentile show evidence of a similar explosive period.

In terms of average artist value, the indices for the middle (between the median and the 90th percentile) and upper (the top 10%) segments, show evidence of explosive behaviour (although in some versions the lower end of the market also showed some explosive price behaviour). These indices should suffer less from potential endogeneity, as the artists were ranked by the average prices for their artworks in the first five years of the sample period. In terms of the medium segments, the indices for watercolour and oil paintings show evidence of explosive behaviour according to both model formulations.

The explosive behaviour in art market prices therefore seems to have occurred especially in the higher-end of the market (i.e. expensive artworks and artists), and in the oil and watercolour segments. The origination and termination dates are quite consistent in suggesting a bubble formation period between 2006/2007 and 2008.

#### 2.6.4 Discussion

The results assume that the aesthetic or utility dividends associated with South African art did not exhibit explosive behaviour over the period. Aesthetic dividends fluctuate over time, as they depend

on buyers' willingness to pay for art, which in turn depends on preferences and wealth. Preferences regarding art and culture would have had to fluctuate dramatically to explain the movements in art prices over the period. Even if trends can temporarily emerge for specific artists, previous findings in the literature have shown that preferences tend to be stable, even in the long run (Penasse and Renneboog, 2017).

The aesthetic dividend can also fluctuate as wealth fluctuates over time (Spaenjers, Goetzmann and Mamonova, 2015). The literature has provided evidence supporting this idea, with Goetzmann, Renneboog and Spaenjers (2011) finding cointegrating relationships between art prices and top income brackets. Mandel (2009) analysed the satisfaction derived from conspicuous consumption, which increases as the value of art increases. The part of the aesthetic dividend that is a signal of wealth could plausibly lead to price increases, which in turn could lead to another increase in social status consumption. However, it is unlikely that aesthetic dividends, and factors such as collectors' preferences and wealth, exhibited similar explosive behaviour over the period.

Although the bubble detection method provides a consistent basis for identifying periods of explosive behaviour, it does not provide an explanation of the bubble episode. The findings are compatible with several different explanations, including rational bubbles, herd behaviour, and rational responses to fundamentals (Phillips, Wu and Yu, 2011).

The periods of explosive prices could be compatible with a rational bubble, where buyers are willing to pay more for an artwork than their private value, because they expect to resell it at a higher price in the future. Gérard-Varet (1995), for instance, argued that the sharp rise in world art prices in the late 1980s could be explained by a rational bubble. Although buyers believed that prices had reached unsustainable levels in the short run, the prospects for continued increases were sufficient to compensate for the risk of the bubble bursting. Prices may increase at an accelerating rate because the probability of a crash increases, and rational investors require an increasing risk premium to cover this higher probability of a crash (Rosser, Rosser and Gallegati, 2012).

An artwork that would be too expensive under normal conditions might become more attractive during a bubble, because buyers might think they will be compensated by further price increases. Buyers may also be concerned that they will not be able to afford the artwork later. Large expected price increases may have a strong impact on demand if buyers think that prices are unlikely to decrease, so that the perceived risk associated with a purchase is minimal (Case and Shiller, 2003).

Penasse and Renneboog (2017) argued that limits to arbitrage induce a speculative component to art prices. Constraints on short selling and high transaction costs may lead to prices diverging from fundamentals, as they prevent arbitrageurs from pulling prices back to fundamental levels (Balcilar, Katzke and Gupta, 2015). When prices are high, pessimists would like to short-sell, but instead they simply stay out of the market or sell to optimists at inflated prices. Optimists may be willing to pay higher prices than their own valuations, because they expect to resell to even more optimistic buyers in the future. The price of the option to resell the artwork in the future is the difference between

their willingness to pay and their own optimistic valuation. The price of the resale option imparts a bubble component in art prices, and can explain price fluctuations unrelated to fundamentals. These market failures impede the ability to correct price inefficiencies and imply that periods of bubble-like behaviour could exist with little scope for arbitrage. This is especially relevant in art markets, where transaction costs are high, short selling is not possible, and without a rental market, the only possibility to make a profit is by reselling at a higher price (Penasse and Renneboog, 2017).

Penasse and Renneboog (2017) investigated this theory by looking at the behaviour of art prices and volumes. They found that the art market was subject to frequent booms and busts in both prices and volumes. They showed that booms in volume were mainly driven by short-term transactions, which were interpreted as speculative transactions or trading frenzies. Given the high transaction costs in the art market, it is unlikely that these artworks were purchased for the pure aesthetic dividends. The positive correlation between prices and volumes was persistent across art movements, and was larger for the most volatile segments of the art market (i.e. Modern and Contemporary art). When the trading volume was high, they found that buyers tended to overpay, and earned negative returns in subsequent years. This provides evidence for the resale option theory and speculative trading models of bubble formation, which predict that speculative trading can generate significant price bubbles, even if trading costs are large and leverage impossible.

In general, speculative bubbles can act like self-fulfilling prophecies (Rosser, Rosser and Gallegati, 2012). Prices increase because buyers expect them to increase, with this continuing expectation leading to higher demand, which causes further price increases. If some exogenous shock, such as the financial crisis, ends the price increases, the expectation ceases, and the demand suddenly disappears. Prices decline towards their fundamental value, where there is no expectation of price increases.

Kindleberger and Aliber (2005) argued that a boom in one market often spills over into other markets. A famous example in the context of art is the link between the boom in Japanese stock and real estate prices and the Impressionist art market in the second half of the 1980s. Hiraki *et al.* (2009) found a high correlation between Japanese stock prices and the demand for art by Japanese collectors, leading to an increase in the price of Impressionist art during this period. Kräussl, Lehnert and Martelin (2016) found corroborating evidence of a bubble period in the 'Impressionist and Modern' art segment between 1986 and 1991. During this period Japanese credit was freely available, backed by increasing values of stocks and real estate, which led to a consumption and investment spree through the wealth effect. Japanese investors invested heavily in international art and particularly French Impressionist art in the late 1980s. Luxury consumption by Japanese art collectors increased international art prices until the art bubble burst, as a consequence of the collapse of the Japanese real estate market (Penasse and Renneboog, 2017).

Similarly, the run-up to the financial crisis saw large increases in asset prices and credit expansion in South Africa. It is likely that spillovers from credit and asset markets and the wealth effect contributed to the explosive behaviour in South African art prices between 2006 and 2008, by



increasing the demand for luxury goods. Increasing prices might have caused a self-fulfilling prophecy, which was exacerbated by high transaction costs and the absence of short-selling in the art market. The financial crisis caused the bubble to burst and led to a decline in South African art prices. While an in-depth investigation lies outside the scope of the chapter, it does illustrate the usefulness of the art price indices for investigating developments in the South African art market.

## 2.7 Conclusion

This chapter has attempted to make three contributions to the literature. The first was to explore and demonstrate aggregation methods to estimate price indices for unique and infrequently traded items, such as artworks, where the composition or quality-mix is not constant over time. These methods capture the central tendency of distribution of growth rates in prices. In addition, a simple new hybrid repeat sales method was proposed, which addresses the lack of repeat sales in the sample and to some extent the potential omitted variable bias of the hedonic method.

According to the smoothness metrics, the regression-based indices produced better estimates of pure price changes than the simple central tendency indices. This demonstrates the importance of regression-based methods for producing quality-adjusted price indices for unique assets. The regression-based indices seem to point to similar general trends in South African art prices, which offers some confidence that the indices provide an accurate measure of the general price movements in the South African art market.

The hedonic and hybrid repeat sales methods demonstrated in this chapter may be useful in constructing indices for other unique assets, such as real estate, antiques, stamps, coins and wine, where the quality-mix of items differs over time, and there is a lack of repeat sales. These techniques may become more useful as more comprehensive microeconomic datasets become available.

International real estate price indices often employ the repeat sales or hedonic methods to calculate quality-adjusted price indices. The repeat sales method, for instance, is used to calculate the S&P/Case-Shiller Home Price Indices in the US. In South Africa, ABSA, FNB and Standard Bank currently use only stratified central tendency methods to construct their property price indices. These indices often do not correspond closely and often lead to different conclusions on changes in property prices, which implies that the indices have substantial shortcomings. This chapter has argued that the central tendency method does not adequately control for quality-mix changes, as was demonstrated for South African property prices by Els and Von Fintel (2010). Changes in the indices may be due to changes in the quality-mix of properties sold, rather than changes in the general price level of the property market. The use of repeat sales or hedonic methods would substantially improve property price indices in South Africa.

Estimating price indices for the property market is complicated by quality changes over time (i.e. there are improvements and renovations, as well as deteriorations over time). However, these



techniques are standard practice internationally and may be adjusted to take renovations and deteriorations into account. In principle, the hedonic method should take account of quality changes over time. In practice, this will depend on the extent to which the underlying data and specification capture changes in quality. For the repeat sales method, studies have used sub-samples where quality was thought to be relatively constant, included a constant to capture average quality changes, and used a weighted least squares approach to correct for increasing variance (e.g. Case and Shiller (1987)). More sophisticated hybrid models can also exploit all of the sales data by combining repeat sales and hedonic regressions to address not only the quality change problem but also sample selection bias and inefficiency problems. These methods would provide more accurate measures of trends in South African property prices than are currently available.

The second contribution was to produce various quality-adjusted South African art price indices. The art price indices, reported in Table 2.8 in the Appendix below, are useful for investigating and understanding developments in the South African art market. Further research applications might consider the risk-return profile of art as an asset class and evaluate whether art could form part of an optimal investment portfolio. Further research might also evaluate whether specific segments of the South African art market (e.g. artworks by specific artists) would have yielded diversification benefits. Conventional wisdom says that the top artworks by established artists tend to outperform the rest of the market (Mei and Moses, 2002). This so-called ‘Masterpiece effect’ may be investigated by examining the specific collections of artworks in the upper part of the price distribution.

The third contribution was to use the art price indices to test for evidence of episodes of mildly explosive prices, which demonstrates the usefulness of the art price indices. The regression-based indices seem to point to consistent evidence of explosive prices in the run-up to the Great Recession, with the bubble period starting around 2006 and ending around 2008. The explosive behaviour in art prices seems to have occurred mainly in the high-end part of the market, and in the oil and watercolour segments. Future research might further consider the drivers of the bubble formation process, to provide an explanation of the bubble episode. Potential factors that influenced the fluctuations in art prices over time, such as wealth effects, might also be investigated. The quality-adjusted art price indices estimated in this chapter can facilitate these inquiries and enable one to be more concrete about developments in the South African art market.

## **2.8 Appendix**

### **2.8.1 Rolling Coefficient Estimates**

To allow for shifts in the implicit prices, adjacent-period indices were calculated by estimating separate models for adjacent subsamples. The coefficients from these models are similar in magnitude to the full pooled sample model and are significant in virtually all cases. Figure 2.20 and Figure 2.21 illustrate a selection of the coefficient estimates from the 1-year adjacent-period hedonic models over

## Art Prices

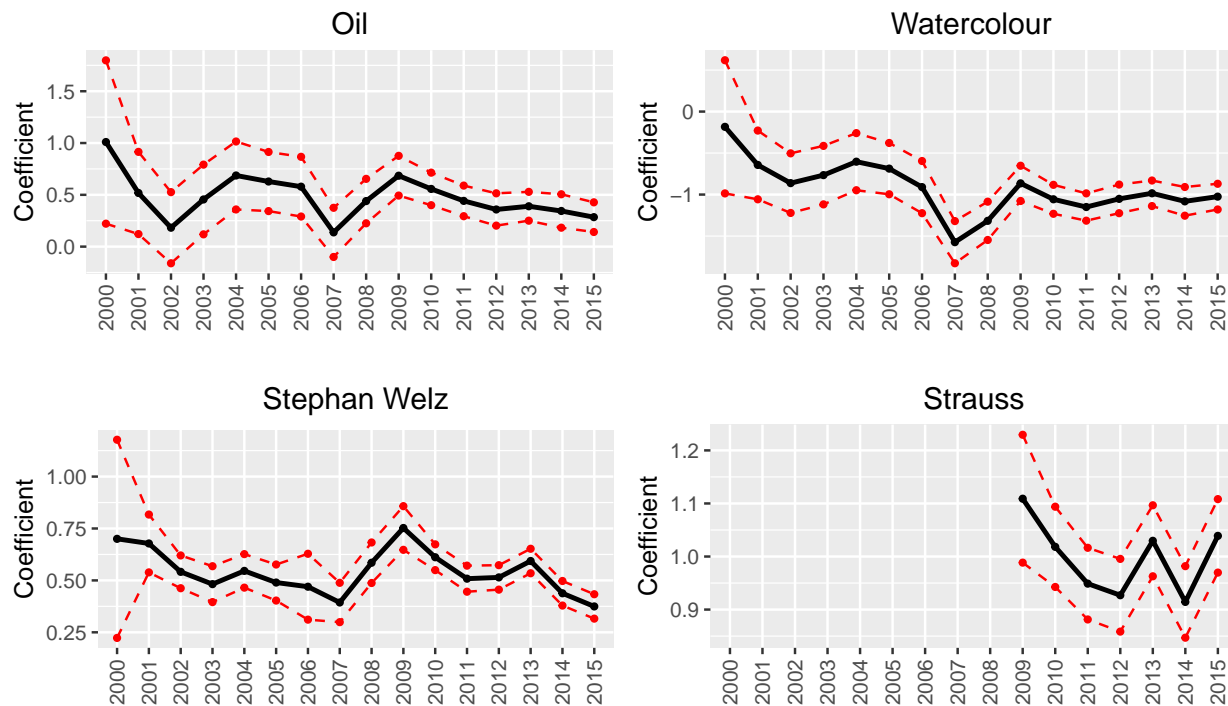


Figure 2.20: Coefficient estimates from the 1-year adjacent-period models

time, along with 95% confidence intervals. The confidence intervals are calculated as two standard errors around the coefficient estimates from each hedonic model (excluding the interaction terms). The coefficient estimates seem to be relatively stable over time, and there do not appear to be any large structural changes over the period. However, there are a few cases in which the estimated parameters fluctuate over time. For example, the coefficient of the ‘signed’ dummy varies between 0 in the model for 2005 and 0.5 in the 2008 model, indicating that it is still useful to allow for shifts in the implicit prices over time.

The adjusted- $R^2$  values for the models over time provides an indication of how influential the hedonic features are at various points in time, and how strong the ‘signal’ of the resulting index is. Figure 2.21 illustrates the adjusted- $R^2$  values for the 1-year adjacent-period hedonic models. The adjusted- $R^2$  is relatively high throughout the sample period, and varies between 0.63 in the 2000 model and 0.83 in the 2013 model. This implies that the hedonic attributes explain quite a large part of the variation in art prices.

### 2.8.2 Index Confidence Intervals

Confidence intervals for the time dummy estimates can be constructed to try to distinguish between real movements in the price indices and statistical idiosyncrasies. The confidence intervals illustrated in this section are calculated as two standard errors around the coefficient estimate.

## Art Prices

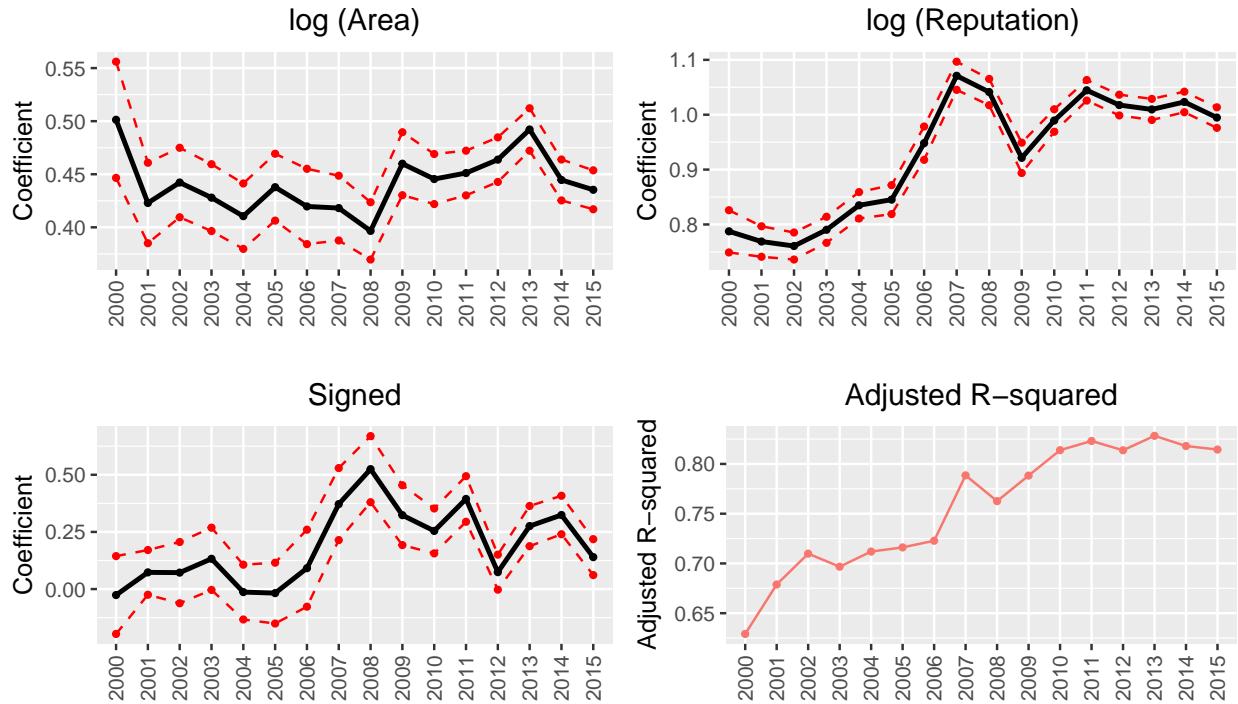


Figure 2.21: Coefficient estimates and adjusted R-squared from the 1-year adjacent-period models

Figure 2.22 illustrates the hedonic indices, along with the 95% confidence interval for the full hedonic model. The adjacent-period indices are within the confidence interval for the entire period, except for the 1-year adjacent-period model, which is below the confidence interval for the latter part of the period. Figure 2.23 illustrates the confidence intervals for the 2-year adjacent-period hedonic model coefficients. The interval becomes narrower over time as there are more observations in the latter part of the sample period.

Figure 2.24 illustrates the pseudo-repeat sales indices, along with the 95% confidence interval for the larger sample model (ps-RS2). The smaller sample ps-RS index is within the confidence interval for the entire period. Because they are based on much smaller samples, the confidence intervals around these indices are wider than those around the hedonic indices.

Figure 2.25 and Figure 2.26 illustrate representative art price indices from each method, with the 1-year adjacent-period index confidence interval, and the larger sample (ps-RS2) confidence interval, respectively. The median falls outside of both confidence intervals for most of the sample, implying that it is statistically different from the regression-based indices. In contrast, the regression-based indices are similar, with few periods where the measures are statistically different. While smaller sample sizes may make the estimation sensitive to outliers, particularly in the earlier periods, the confidence intervals suggest that the movements are ‘real’ rather than statistical idiosyncrasies.

Art Prices

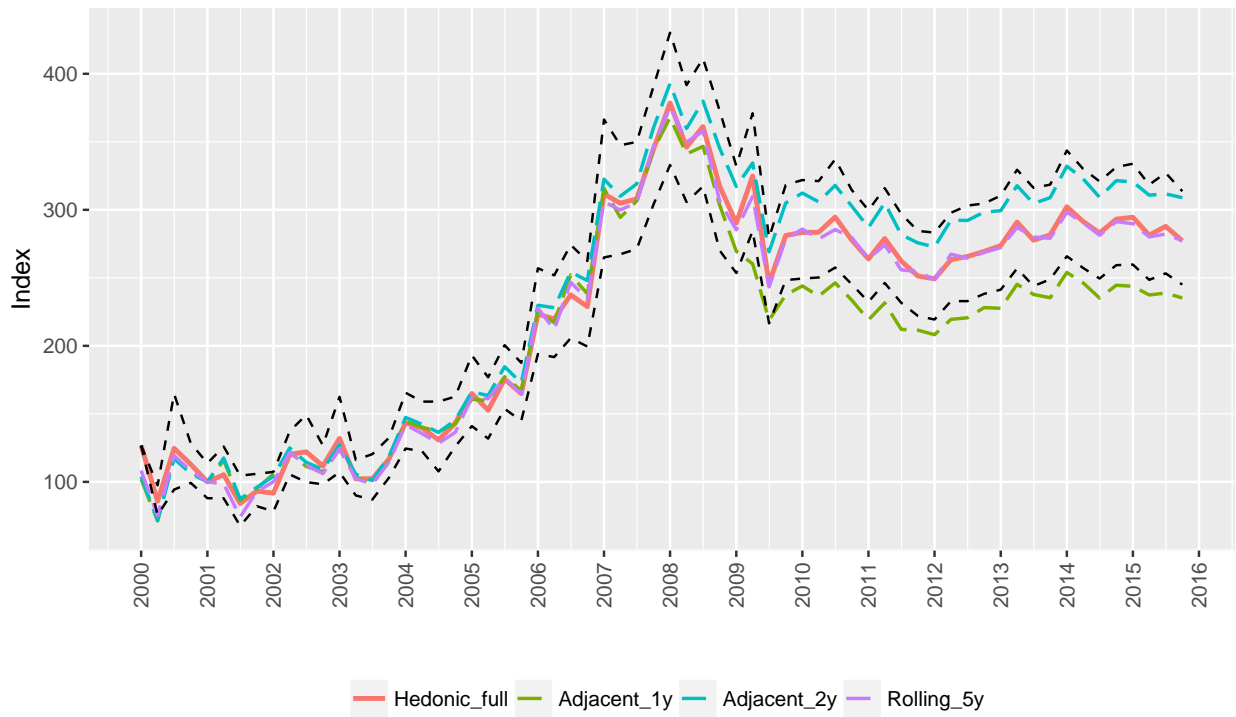


Figure 2.22: Hedonic indices with the full hedonic model confidence intervals

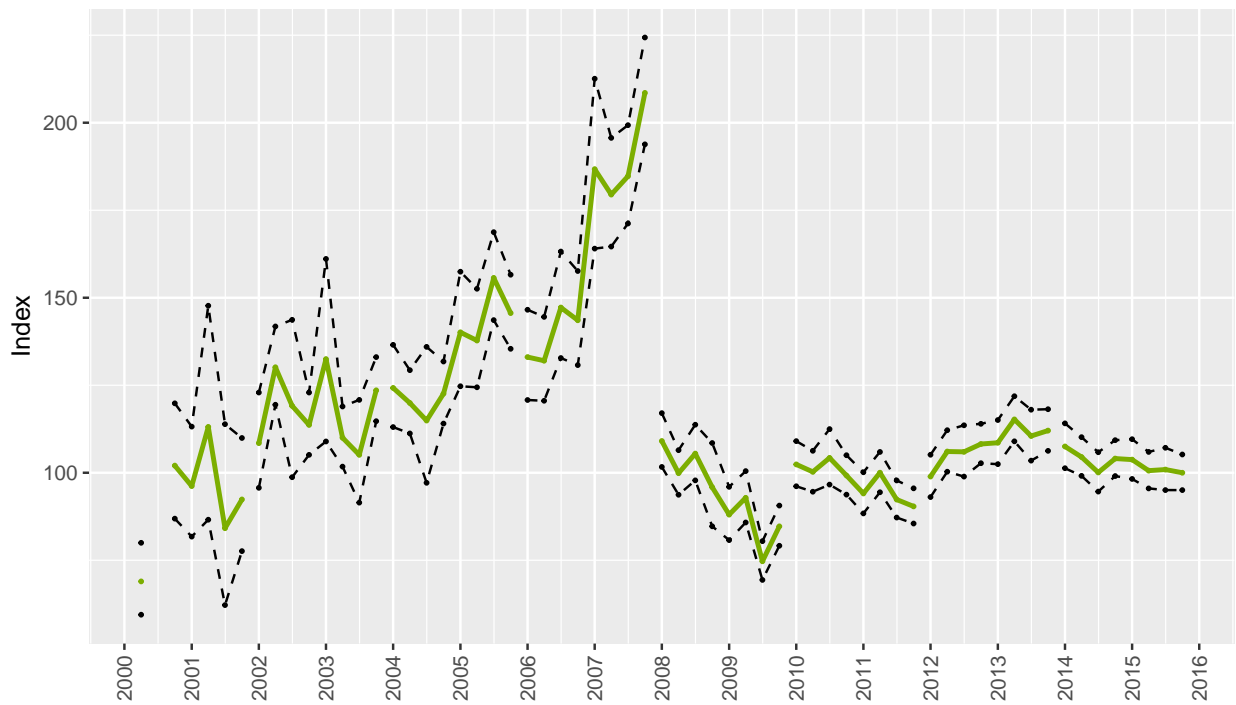


Figure 2.23: 2-year adjacent-period hedonic model coefficients with confidence intervals

Art Prices

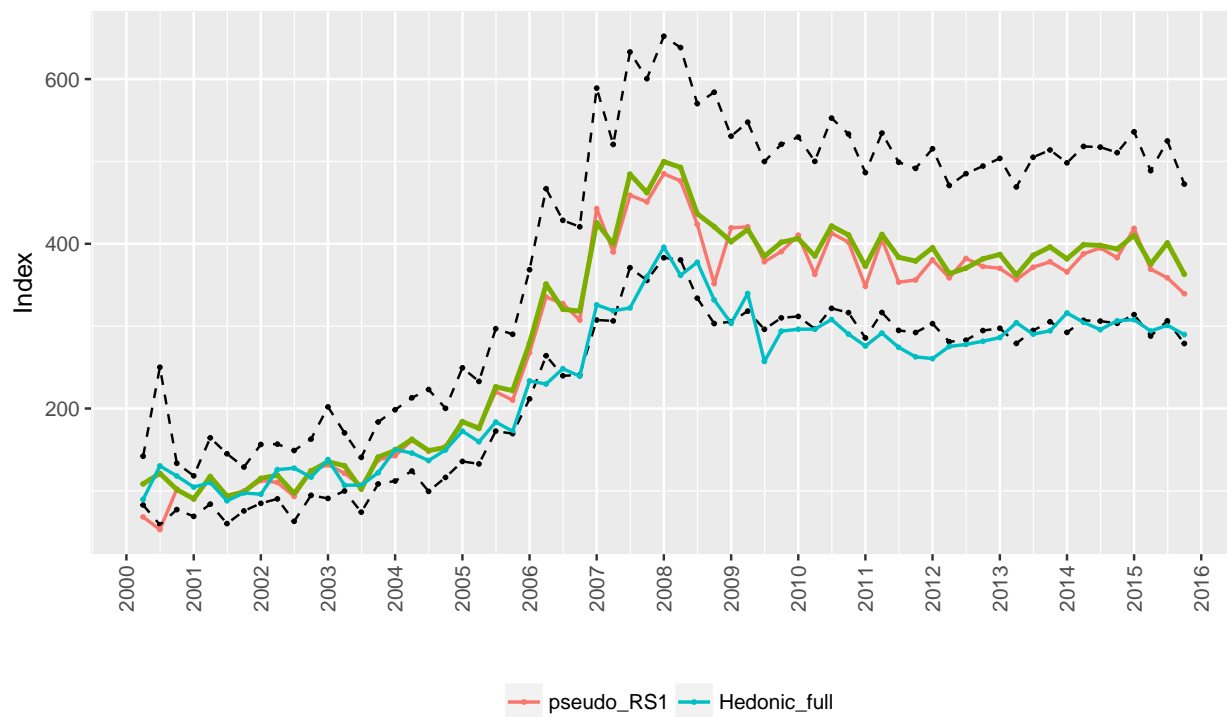


Figure 2.24: Pseudo-repeat sales indices with larger sample (ps-RS2) model confidence intervals

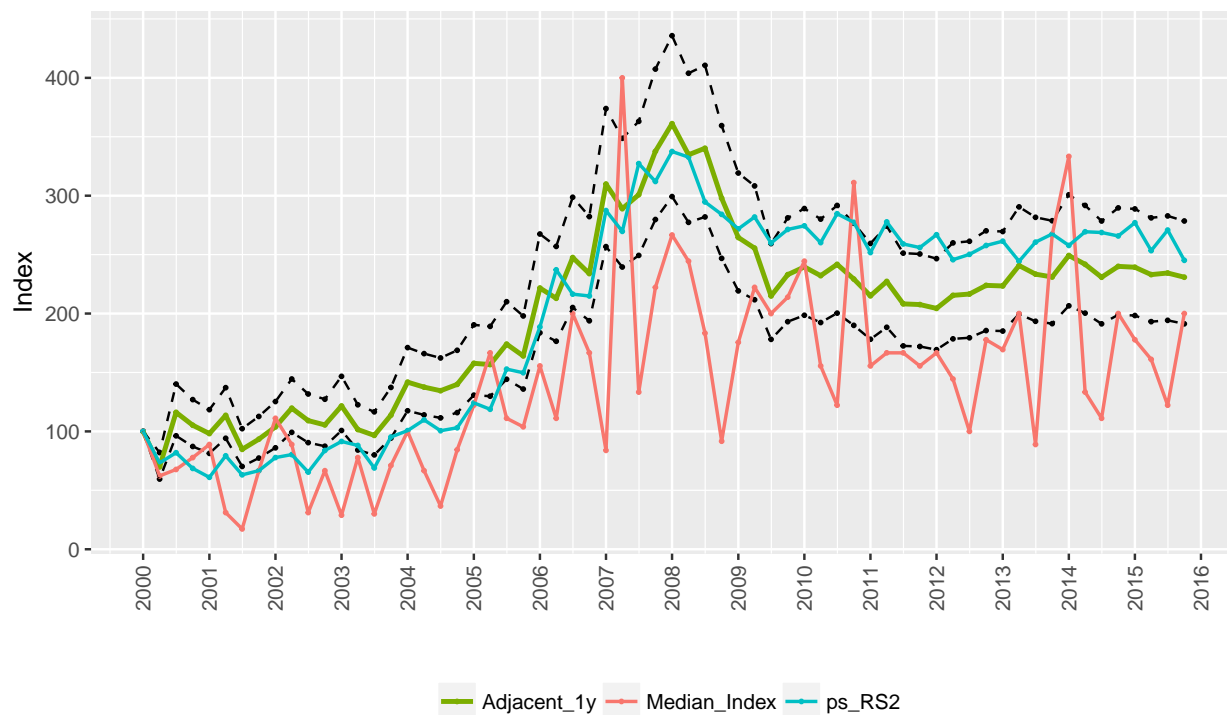


Figure 2.25: Art price indices with the 1-year adjacent-period hedonic model confidence intervals

## Art Prices

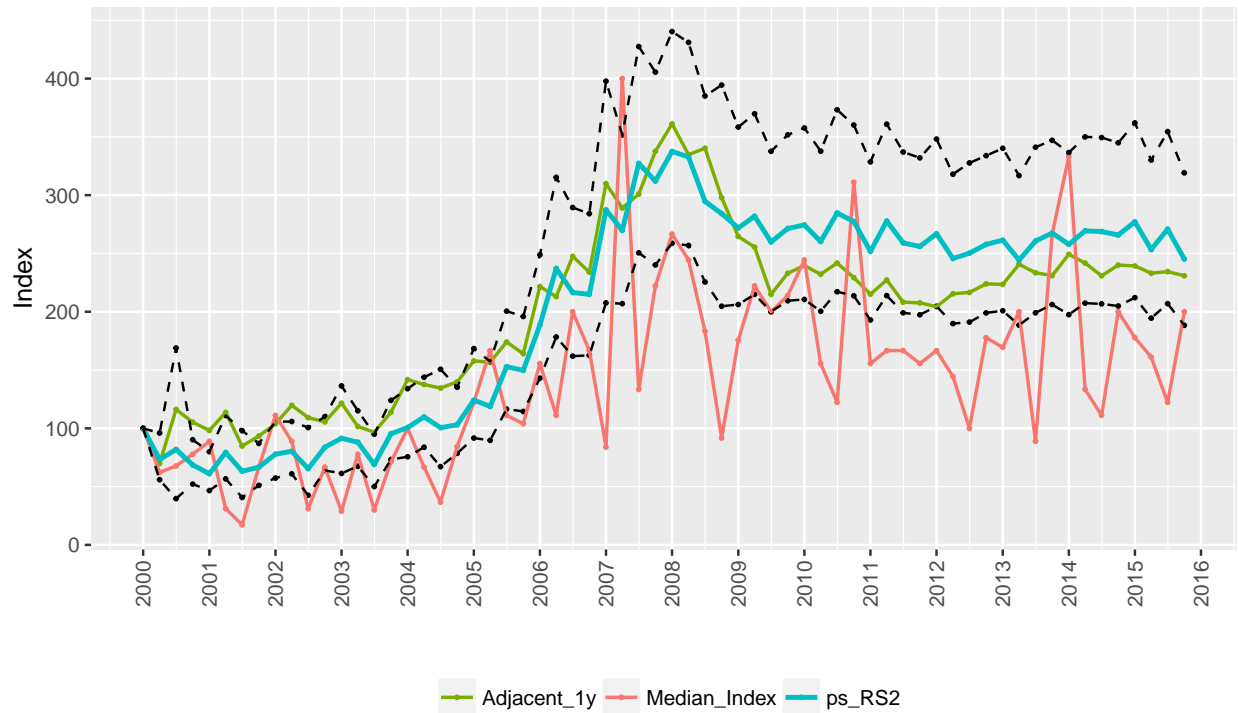


Figure 2.26: Art price indices with ps-RS2 model confidence intervals

### 2.8.3 Weighted Bubble Detection Tests

The standard errors of the time coefficients of the regression provide an idea of the reliability of indices at specific points in time. As illustrated above, the standard errors are not constant between the index estimates (i.e. there is a type of heteroskedasticity present). The statistical inference in the bubble detection tests should take into account the additional variance introduced by the estimation when aggregating the data. Hence, the bubble detection tests are estimated by weighting each observation by the inverse of its standard error, much like Feasible Generalised Least Squares estimation. The weights are used to indicate that different observations have different standard errors. This places more emphasis on the segments of the data which are estimated more accurately from the microeconomic data, and makes an adjustment for the heteroskedasticity between the index points.

Figure 2.27 and Figure 2.28 illustrate the test statistics and critical values for weighted models without and with a drift term. Table 2.7 reports the periods of explosive behaviour derived from these weighted models. The bubble detection test results are similar to those in Table 2.5. The slight exceptions are for the full hedonic and rolling 5-year indices, which only exhibit one period of explosive behaviour. As the results are similar to the results in the main text, it suggests that the inference is acceptable in this case.

Art Prices

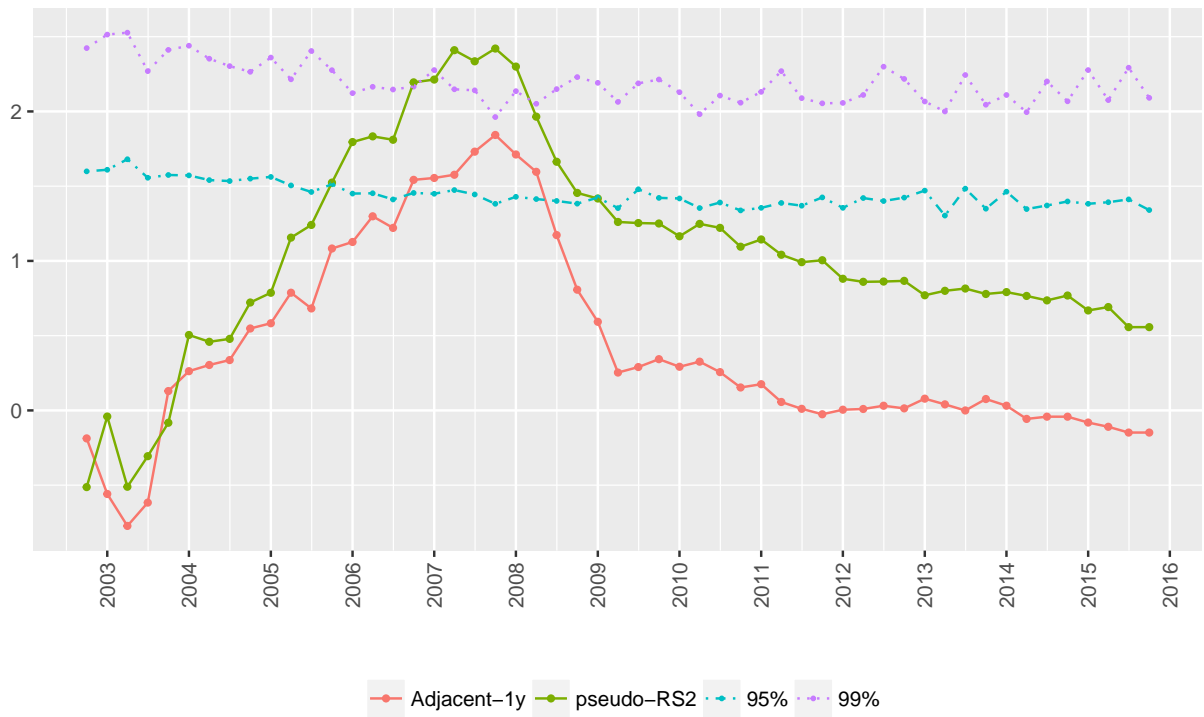


Figure 2.27: Test statistics and critical values for weighted models without drift

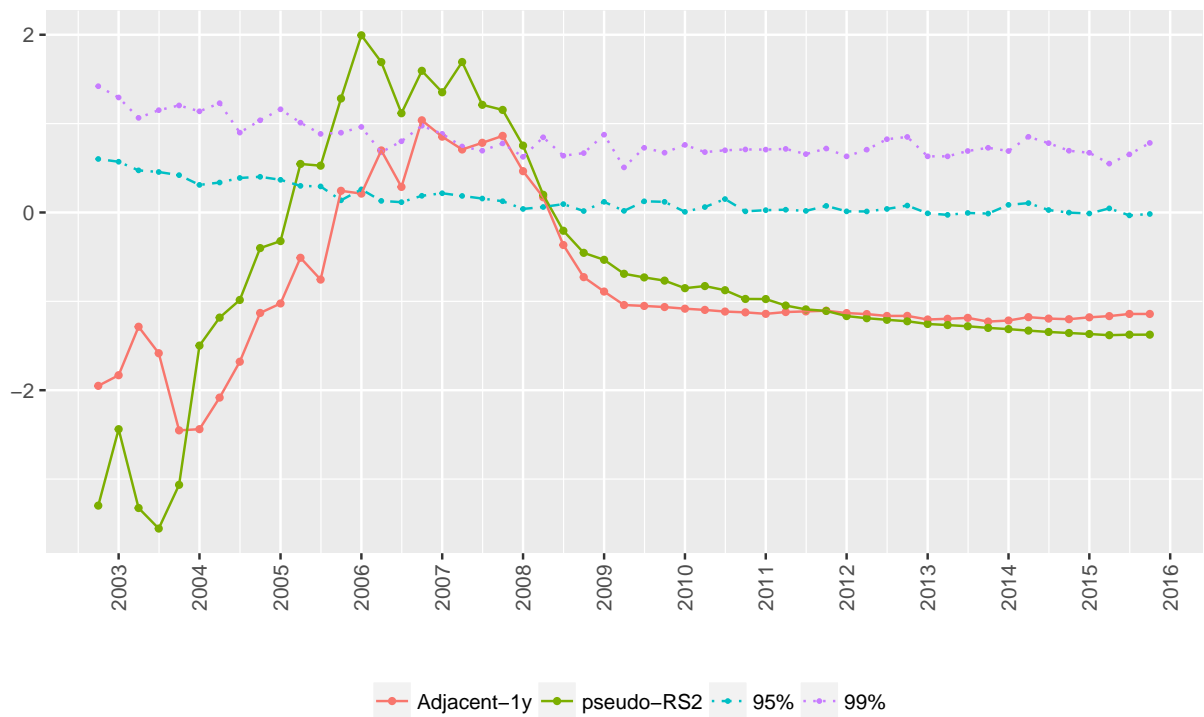


Figure 2.28: Test statistics and critical values for weighted models with drift

## Art Prices

Table 2.7: Dates of explosive price behaviour (weighted models)

	No Drift		Drift	
	Start	End	Start	End
Hedonic Full	2007 Q4	2007 Q4	2006 Q4	2008 Q2
Adjacent-1y	2006 Q4	2008 Q2	2005 Q4	2008 Q2
Adjacent-2y	2006 Q4	2008 Q3	2005 Q4	2008 Q2
Rolling-5y	2007 Q4	2007 Q4	2006 Q4	2008 Q1
ps-RS1	2005 Q3	2009 Q1	2005 Q4	2008 Q1
ps-RS2	2005 Q4	2008 Q4	2005 Q2	2008 Q2

**2.8.4 South African Art Price Indices**



## Art Prices

Table 2.8: Art price indices (2000Q1=100)

Date	Median	Hedonic_full	Adjacent_1y	Adjacent_2y	Rolling_5y	pseudo.RS1	pseudo.RS2
2000 Q1	100.00	100.00	100.00	100.00	100.00	100.00	100.00
2000 Q2	62.22	67.65	69.90	68.96	69.38	69.06	73.24
2000 Q3	67.70	98.31	116.15	112.64	110.18	53.67	81.92
2000 Q4	77.78	89.06	105.22	102.04	99.81	103.29	68.58
2001 Q1	88.89	78.88	98.15	96.19	92.33	91.06	61.03
2001 Q2	31.11	83.04	113.62	113.09	91.03	117.47	79.30
2001 Q3	17.22	66.34	84.78	84.16	68.51	95.07	63.14
2001 Q4	66.67	73.56	93.37	92.37	85.90	100.94	66.67
2002 Q1	111.11	72.33	103.86	100.18	92.39	113.84	77.79
2002 Q2	88.89	94.89	119.71	120.20	112.32	111.45	80.31
2002 Q3	31.11	96.29	109.16	110.01	103.54	94.46	65.43
2002 Q4	66.67	88.04	105.49	104.99	97.92	125.86	83.81
2003 Q1	28.89	104.10	121.58	122.36	114.85	133.22	91.47
2003 Q2	77.78	80.62	101.58	101.61	94.52	122.68	88.08
2003 Q3	30.00	80.70	96.61	97.10	90.86	107.48	68.97
2003 Q4	71.11	92.13	113.81	114.11	105.69	139.18	95.28
2004 Q1	100.00	113.23	141.79	141.78	131.00	144.28	100.66
2004 Q2	66.67	110.09	137.56	136.86	125.09	163.16	109.79
2004 Q3	36.67	103.22	134.52	131.14	118.57	149.85	100.56
2004 Q4	84.44	112.91	139.90	139.90	125.90	155.70	103.07
2005 Q1	122.22	130.13	157.77	159.90	149.63	186.05	124.25
2005 Q2	166.67	120.44	156.72	157.22	148.81	179.16	118.76
2005 Q3	111.11	138.49	174.06	177.64	162.90	222.76	152.84
2005 Q4	103.89	130.04	164.04	166.17	151.83	212.58	149.77
2006 Q1	155.56	176.14	221.66	221.09	210.27	270.93	188.65
2006 Q2	111.11	173.32	212.95	219.35	195.73	339.30	237.13
2006 Q3	200.00	187.35	247.51	244.60	227.78	331.05	216.46
2006 Q4	166.67	180.48	233.91	238.56	216.53	310.89	214.91
2007 Q1	83.89	245.81	309.88	310.29	282.46	447.67	287.39
2007 Q2	400.00	240.42	288.98	298.22	276.92	394.42	269.73
2007 Q3	133.33	243.01	300.87	306.94	281.89	464.20	327.24
2007 Q4	222.22	271.67	337.65	346.50	320.14	455.97	312.07
2008 Q1	266.67	298.57	361.12	377.95	346.87	490.53	337.50
2008 Q2	244.44	272.96	334.73	346.09	323.01	481.75	332.74
2008 Q3	183.33	284.90	340.26	365.47	330.43	428.52	294.68
2008 Q4	91.67	250.35	297.88	332.16	282.83	355.61	284.16
2009 Q1	175.56	228.94	264.54	305.05	263.38	424.09	271.84
2009 Q2	222.22	256.24	255.48	321.69	286.09	425.45	281.92
2009 Q3	200.00	194.02	214.92	258.88	224.69	382.50	259.86
2009 Q4	213.89	221.79	233.00	293.44	257.25	395.02	271.39
2010 Q1	244.44	223.53	239.58	300.43	263.94	415.09	274.43
2010 Q2	155.56	223.60	232.14	294.20	257.25	367.04	260.18
2010 Q3	122.22	232.43	241.76	305.89	263.60	417.49	284.71
2010 Q4	311.11	219.07	229.25	291.14	258.19	406.67	277.41
2011 Q1	155.56	208.16	215.03	276.07	243.75	352.31	251.76
2011 Q2	166.67	220.01	227.26	293.49	253.22	410.11	277.79
2011 Q3	166.67	206.91	208.22	270.99	236.26	357.44	259.03
2011 Q4	155.56	198.25	207.62	265.22	235.01	360.08	256.00
2012 Q1	166.67	196.62	204.35	262.33	229.59	384.85	266.90
2012 Q2	144.44	207.75	215.44	281.29	246.91	362.78	245.68
2012 Q3	100.00	209.57	216.55	281.07	244.19	386.57	250.23
2012 Q4	177.78	212.51	223.91	286.99	248.05	376.75	257.75
2013 Q1	169.44	215.90	223.41	287.95	251.32	374.40	261.40
2013 Q2	200.00	229.49	240.74	305.63	265.07	360.51	244.34
2013 Q3	88.89	219.03	233.41	293.06	258.56	375.71	260.59
2013 Q4	266.67	222.10	231.00	297.16	257.67	382.48	267.50
2014 Q1	333.33	238.42	249.26	319.48	275.38	369.90	257.79
2014 Q2	133.33	229.86	241.77	310.51	268.18	392.16	269.40
2014 Q3	111.11	223.18	230.81	297.41	259.73	399.53	268.77
2014 Q4	200.00	231.24	240.03	309.21	268.97	387.49	265.83
2015 Q1	177.78	232.30	239.31	308.32	267.58	423.43	277.07
2015 Q2	161.11	221.90	233.04	298.95	258.63	373.30	253.39
2015 Q3	122.22	227.06	234.39	299.87	260.57	362.78	270.84
2015 Q4	200.00	218.67	230.83	297.17	256.01	343.20	245.17

### 3 Methods for Aggregating Disparate Qualitative Survey Responses: An Application to Business Sentiment in South Africa

#### 3.1 Introduction

The global financial crisis was associated with unusually low levels of confidence and heightened uncertainty among firms. According to the European Central Bank (2013), weak business sentiment contributed to a large extent to the Great Recession and to the lacklustre subsequent recovery. More recently, there has been increased uncertainty around the implications of the Brexit referendum (Jackson, Tetlow and Kahn, 2017) and the policy direction under President Trump (Shen, 2017). The International Monetary Fund (2017) cited elevated political uncertainty and weak consumer and business confidence when it marked down South Africa's growth forecast for 2018. The idea that weak business sentiment influenced economic activity has inspired a substantial literature examining the impact of changes in sentiment, and especially uncertainty, on investment and output decisions.

Business sentiment covers two distinct concepts: *confidence* and *uncertainty*. For the purposes of this chapter, business confidence is the degree of optimism that firms hold, or their perceptions of, current and future business conditions (Mendicino and Punzi, 2013). Business uncertainty is the inability of firms to forecast the probability of future events occurring (Knight, 1921). It is challenging to measure these concepts (Santero and Westerlund, 1996), as both are not directly observable and their definitions are difficult to operationalise.

The aim in this chapter is to explore aggregation methods for estimating business confidence and uncertainty in South Africa, using the microeconomic data from the Bureau for Economic Research (BER) business tendency surveys. Although measuring business sentiment is challenging, survey-based indicators can be helpful in discovering agents' opinions on future economic developments (Organisation for Economic Co-operation and Development, 2003).

To date, there has been little research on business sentiment in South Africa, in part due to the difficulty of measurement. As far as research on confidence is concerned, only two business confidence indicators are regularly published for South Africa: the South African Chamber of Commerce and Industry Business Confidence Index (SACCI BCI) and the BER Business Confidence Index (BER BCI). The SACCI BCI is a composite measure of economic activity, rather than a confidence indicator in the way used in the literature. The BER BCI is a measure of confidence derived from the BER's business tendency surveys. It is based on a single question on current conditions. The survey responses are weighted in an ad hoc manner, and the services sector survey is excluded from the calculation. As far as research on uncertainty, as opposed to confidence, is concerned, only a few studies have created proxies for uncertainty in South Africa (e.g. Redl (2015) and Hlatshwayo and Saxegaard (2016)). No study has fully exploited the information contained in the BER business

tendency surveys to construct proxies for business uncertainty in South Africa.

The challenge in aggregating the microeconomic data from the BER business tendency surveys is to identify an underlying pattern from the disparate views of individual agents. In the chapter, an attempt is made to capture the full distribution of the qualitative survey responses, by calculating the first and second weighted cross-sectional moments of the distribution. The composite indicators of business confidence and uncertainty are based on these moments.

Two composite confidence indicators are calculated in this chapter: the cross-sectional mean of responses to questions on current business conditions, and the cross-sectional mean of responses to questions on expected future business conditions (Organisation for Economic Co-operation and Development, 2003). Three composite uncertainty indicators are calculated: the scaled cross-sectional standard deviation of forward-looking responses (Girardi and Reuter, 2017); the cross-sectional mean of individual firm forecast errors; and the cross-sectional standard deviation of forecast errors (Bachmann, Elstner and Sims, 2013; Arslan *et al.*, 2015).

The new composite indicators attempt to improve on the existing measures of sentiment for South Africa. The indicators incorporate the survey responses from questions on general business conditions, output, employment, orders placed and profitability. For each question, the responses are weighted by firm size and subsector size to produce sectoral indicators, including the services sector. The sectoral indicators are then weighted by GDP share to produce the overall aggregate composite indicators.

The validity of the indicators is assessed by comparing them with events that were thought to coincide with large changes in confidence and uncertainty, as well as with existing measures for South Africa. The two confidence indicators are compared with the BER BCI and the SACCI BCI. The three composite uncertainty indicators are compared with a measure of financial market uncertainty and the economic policy uncertainty indicator created by Hlatshwayo and Saxegaard (2016). A composite overall measure of uncertainty is constructed, which combines the survey-based uncertainty indicators with the measures of financial market and economic policy uncertainty. Similar to the previous chapter, the indicators are evaluated to determine whether they improve on the existing indicators. In this chapter, the indicators are evaluated according to their comovement with real GDP growth. The leading indicator properties of the confidence indices are also evaluated, in terms of the timing of their turning points and their concordance with the official SARB business cycle.

There has been little analysis of this relationship in the South African context (e.g. Pellissier (2002) and Redl (2015)). The newly constructed sentiment indicators are therefore exploited to further examine the relationship between business sentiment and real economic activity in South Africa. This demonstrates the usefulness of the aggregation methods and provides an additional validity test of the estimated indicators. In particular, the hypothesis is tested that there is significant comovement between the sentiment indicators and real GDP growth, using a standard VAR framework. The

relationship between the indicators and real GDP growth is investigated, including the timing of the relationship and the extent to which correlation is conditional on other economic variables. The following sections provide a review of the literature on confidence and uncertainty.

## **3.2 Confidence**

For the purposes of this chapter, business confidence involves firms' perceptions of, or degree of optimism regarding, current business conditions and the expected future business climate (Mendicino and Punzi, 2013). This section begins with a review of the theoretical links between confidence and macroeconomic outcomes. The section then discusses the measurement challenges and the evidence on the impact of confidence on economic outcomes.

### **3.2.1 Theory on Confidence and Economic Outcomes**

While confidence measures are popular indicators with the media and business players, the stance of the academic literature on the value of these indicators is more ambiguous (Barsky and Sims, 2012). A review of the literature suggests three alternative views, which range from the view that confidence measures play an important causal role in the business cycle, to the view that they contain useful predictive information but play a limited causal role, to the view that they have no value, even in forecasting. In this section these three views are briefly discussed.

According to the so-called 'animal spirits' view, psychological factors have a causal impact on economic fluctuations that is distinct from fundamentals (Carroll, Fuhrer and Wilcox, 1994). This view is most closely associated with Keynes (1936), who argued that: "Our decisions to do something positive, the full consequence of which will be drawn out over many days to come, can only be taken as a result of animal spirits - of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities." The original Keynesian view finds resonance in the more recent literature, with Akerlof and Shiller (2015) arguing that in the face of uncertainty, decisions about the future are based on animal spirits, rather than a weighted average of quantitative benefits and probabilities, as rational theory would dictate.

Thus, psychological factors are drivers of consumption and investment decisions in the face of uncertainty, due to the difficulty of making accurate forecasts (Pagan, 2013). For firms, waves of optimism and pessimism may cause errors in their expectations about future demand and profits. When firms are optimistic about future demand and profits, they decide to accumulate capital. If their expectations are not met, there will be a period of reduced investment, which may cause a recession (Beaudry and Portier, 2004). These errors therefore generate cycles through increases and decreases in investment (Leduc, 2010).

For consumers, optimistic income expectations lead them to increase their discretionary expenditures. These expectations depend not only on economic fundamentals, such as income or prices, but also on psychological factors. Psychological factors therefore influence consumers' perceptions of their economic environment and may become an independent source of economic fluctuations through their impact on their decisions (Mendicino and Punzi, 2013). Blanchard (1993), for instance, argued that a negative consumption shock, which was associated with an exogenous shift in pessimism, was the cause of the 1990/1991 US recession. Akerlof and Shiller (2015) argued that deteriorating confidence, one of the main elements of animal spirits, was an important reason for the global financial crisis and subsequent recession.

Models with multiple equilibria provide a potential causal link between confidence and fluctuations in economic activity, where the equilibria may be determined by sentiment (Taylor and McNabb, 2007). The level of confidence is a potential variable that can determine which equilibrium occurs. For example, if a crisis of confidence causes a banking panic, the economy may settle at a suboptimal equilibrium, in respect of social welfare. If confidence is high, the economy may settle at an optimal equilibrium (Leduc, 2010). In this context, confidence is a prediction of a future outcome, which may become self-fulfilling (Akerlof and Shiller, 2015).

According to the animal spirits view, therefore, confidence has a potentially important causal impact on economic outcomes. In contrast, the so-called 'news' view argues that confidence indicators contain useful predictive information for economic output, but play a limited causal role.

According to the news view, any relationship between confidence measures and subsequent real activity means that confidence measures contain information about current and future fundamentals of the economy (Barsky and Sims, 2012). Confidence can proxy for news that agents receive about future productivity, which is not yet reflected in econometricians' information sets, by aggregating information from various sources (Cochrane, 1994; Barsky and Sims, 2012). Confidence indicators reflect agents' expectations about future fundamentals and economic conditions, which are not summarised in other macroeconomic variables. When agents are optimistic, they give positive responses to surveys. These are confirmed, on average, and real activity eventually increases as predicted by the confidence indicator (Carroll, Fuhrer and Wilcox, 1994).

From the rational expectations point of view, confidence should reflect the expected values of economic fundamentals and should not offer any additional predictive information (Beaudry and Portier, 2004). However, a number of studies (e.g. Beaudry and Portier (2004) and Van Aarle and Kappler (2012)) analyse models where agents receive imperfect signals about future productivity growth and use these signals to make investment decisions. In this context, confidence refers to a state where agents receive an above-average signal, which may generate a wave of optimism. Rational agents then learn gradually about the true state of the economy and adjust their expectations. In this environment, occasional recessions reflect the availability of good quality information on which agents act.

Other factors, such as frictions in capital markets, may explain the predictive information contained in confidence indicators. For instance, an increase in confidence may reflect higher future income, but borrowing constraints can prevent higher current consumption in anticipation of an increase in income. As a result, consumption will increase only when actual income increases, and a rise in consumer confidence will predict the future consumption increase (European Central Bank, 2013).

The literature therefore sets out various theoretical links between confidence and economic activity. Yet, it is not clear whether confidence indicators repackage information already contained in other economic variables, or whether they contain useful independent predictive information about the economy. If they contain predictive information, it is not clear whether they reflect animal spirits, or aggregated information on agents' expectations of future outcomes not yet captured by the macroeconomic data (Mendicino and Punzi, 2013; Akerlof and Shiller, 2015).

### **3.2.2 Empirical Findings**

The empirical literature has tried to establish whether there is predictive information in confidence indicators, over and above economic fundamentals, and if so, whether confidence has a separate causal impact on economic activity. Although the findings have not been conclusive, the majority of studies seems to find that confidence indicators are at least positively related to real economic activity (Taylor and McNabb, 2007). The inconclusive findings may be due to two main challenges: how to construct proxies for confidence and how to establish whether it has a separate causal impact on real economic activity.

#### **3.2.2.1 Measuring Confidence**

As confidence cannot be observed or measured directly (Santero and Westerlund, 1996), analysts typically aggregate responses from business and consumer surveys. These surveys usually contain a small number of qualitative questions, which can be answered quickly by respondents. Indicators are derived from the subjective answers to questions on past, current and future developments. The assumption is that agents form opinions about economic conditions before a specific business activity is implemented (e.g. new production plans, employment, or purchases). These opinions may be called 'confidence'. The most important advantage of these surveys is that they are available long before official statistics become available. Moreover, they are not subject to revisions and avoid problems with trends and seasonality.

The most common and widely used method to aggregate survey responses is to calculate so-called balance statistics. In the context of business tendency surveys, balances are simple averages of survey responses. For most survey questions there are three reply options, e.g. 'up', 'the same', or 'down'. Balances are calculated as the difference between the percentage of positive answers and negative answers. Balances are simple to implement and understand, and are considered both

practical and entirely adequate for cyclical analysis (Organisation for Economic Co-operation and Development, 2003).

Although balances are by far the most common method used by statistical agencies and analysts to aggregate the surveys, a few more sophisticated methods have been discussed in the literature, including a probabilistic approach, a regression approach, and a latent factor approach (Nardo, 2003).<sup>20</sup> However, these approaches usually require actual quantitative reference series for the relevant variables, which is restrictive in the case of business confidence, where quantitative reference series are unavailable. Moreover, these methods can become unreliable when exceptional events have a large impact on the correlation between the survey data and the quantitative reference series (United Nations, 2015).

Nevertheless, the evidence suggests that balance statistics tend to produce indicators that are very similar to those produced by more sophisticated methods. For instance, the Italian National Statistical Agency found a very high correlation between balances and more sophisticated indicators when using three-option responses (Organisation for Economic Co-operation and Development, 2003). Driver and Urga (2004) assessed different ways of aggregating qualitative data from the UK employers' business survey into quantitative indicators for a number of variables. They found that the balance statistic was a satisfactory aggregation method for the survey responses on output, investment, and exports. Weighted balance statistics are therefore used in this chapter to calculate summary statistics for the responses to each survey question.

The balances from multiple questions are typically used to calculate composite confidence indicators, as opposed to using a single question. As no single cause explains all cyclical fluctuations over the long term, it is necessary to have information from many possible sources of change, i.e. to use all potentially important information (Van Aarle and Kappler, 2012). Composite indicators have the capacity to react to various sources of economic fluctuations, while being resilient to fluctuations affecting single components. They often have fewer false alarms and fewer missed turning points than individual components and tend to have more stable lead-times (European Central Bank, 2013). In this chapter, composite confidence indicators are calculated by incorporating the responses to a number of questions.

Composite confidence indicators of this type are available for most countries. The European Commission, for instance, builds composite indicators by aggregating the survey responses from five sectors, using multiple questions on current and expected conditions. For example, the industrial indicator is an average of the balances of questions relating to production expectations, stocks of goods (with an inverted sign), and order books, while the retail trade indicator is an average of balances of questions relating to the present and future business situation and stocks (with an

---

<sup>20</sup>The probabilistic approach assumes a probability distribution for the variable concerned, and the measure is a function of this specific probability distribution. The regression approach uses the relationship between survey responses of the past and actual values to quantify respondents' expectations about the future. The measures are a function of specific regression models (Nardo, 2003). In the latent factor approach, the percentages of each response are a function of a common 'latent measure' observed by respondents, but not by econometricians.



inverted sign) (Organisation for Economic Co-operation and Development, 2003). The aggregate Economic Sentiment Index is a weighted average, using value added shares, of confidence in the manufacturing, construction, retail, and services sectors, as well as for consumers (European Central Bank, 2013). Taylor and McNabb (2007) and Mendicino and Punzi (2013) used these composite confidence indicators for a number of European countries in investigating the impact of confidence on economic activity.

Another prominent example is the German Ifo Business Climate Indicator, which is used as a leading indicator in Germany. It is computed as a geometric mean of the balances referring to the current business situation and the business outlook in the next six months (United Nations, 2015). The results for the manufacturing, construction, wholesale and retail sectors are weighted according to the importance of the industry.

Bachmann, Elstner and Sims (2013) used the Ifo Business Climate Survey, as well as the Philadelphia Fed's Business Outlook Survey, to calculate composite indices of current and expected conditions. Barsky and Sims (2012) used the Consumer Sentiment Index, published by the University of Michigan as a composite forward-looking measure of confidence. Kabundi (2004) used a dynamic factor model to calculate a composite indicator from the French Statistics Institute's business survey data.

Two indicators of confidence are published in South Africa: the BER BCI and the SACCI BCI. The BER BCI, discussed in more detail below, is constructed from the BER's quarterly business tendency surveys, which are similar to the business tendency surveys conducted all over the world. The BER BCI is constructed from a specific question (Q1) that appears in all of the surveys: "Are prevailing business conditions: satisfactory, or unsatisfactory?" The BCI is the weighted percentage of respondents who rated prevailing business conditions as 'satisfactory' and is therefore based on the perceptions of business people (Kershoff, 2002). The survey responses are weighted (except the building survey), and the BER BCI is calculated as the unweighted mean of five sectoral indices (excluding the services sector). The BER BCI is an index of current conditions, as opposed to expected conditions, which is independent of external macroeconomic variables (Pellissier, 2002).

The SACCI BCI, formerly known as the SACOB BCI, is a composite index of 13 quantitative sub-indices thought to have the greatest influence on the business mood. These include the exchange rate, inflation, the prime rate, retail sales volumes, credit extension, commodity prices, import and export volumes, new vehicle sales, utility services, manufacturing production, building plans passed, and the stock market index. The SACCI BCI is an *ex post* measure of actual activity, which is dependent on external macroeconomic variables. The rationale is that recent business activity is indicative of the degree of business confidence (SACCI, 2011). In this sense, the SACCI BCI is a composite measure of economic activity, rather than a confidence indicator in the way it is defined in the literature. This chapter aims to calculate improved composite indicators of business confidence based on the BER business tendency surveys.



### 3.2.2.2 The Impact of Confidence

The majority of studies seems to find that confidence indicators are at least positively related to real economic activity, although this does not necessarily imply a causal relationship (European Central Bank, 2013). Confidence indicators have been found to be useful in some cases as leading indicators, as well as for forecasting, even after controlling for other economic variables. Even in cases where the unique information content is limited, the timeliness of survey indicators may make them useful for monitoring developments and for real-time forecasting (nowcasting).

The empirical literature has often investigated the extent to which confidence indicators contain predictive information over and above economic fundamentals (United Nations, 2015). A number of studies have shown that both consumer and business confidence indicators provided valuable information for forecasting real activity, which was not contained in other economic variables (e.g. Santero and Westerlund, 1996; Ludvigson, 2004; Kabundi, 2004; Parigi and Golinelli, 2004; Taylor and McNabb, 2007; Leduc and Sill, 2013; Mendicino and Punzi, 2013; Martinsen, Ravazzolo and Wulfsberg, 2014; and Kilic and Cankaya, 2016).

In an influential study, Barsky and Sims (2012) found that positive shocks to consumer confidence led to significant, slow-building, and permanent responses in consumption and income. If confidence contained no news about future fundamentals, and reflected only animal spirits, one would expect transitory responses. Barsky and Sims (2012) concluded that their results supported the ‘news’ view of confidence.

The European Central Bank (2013) found that confidence indicators can play a significant role in predicting recessions. They included the European Consumer Sentiment Index, along with the OECD leading indicator for the euro area in a probit model. This model captured business cycle phases relatively well, with probabilities increasing when recessions occurred. The drawback was that probabilities also increased in some periods when there were no recessions, i.e. there were some false positives.

The forecasting ability of confidence indicators might be offset by other indicators during ordinary times, while increasing notably during unusual events (United Nations, 2015). The European Central Bank (2013) found that shocks to confidence played a relatively small role during normal times, but were important during more extreme episodes such as financial crises and recessions. The impact was asymmetric: large decreases in confidence were more important in predicting future changes in consumption than large increases. This pointed to a non-linear and asymmetric relationship between confidence and economic fluctuations.

Even if confidence indicators are just a synthesis of economic variables and do not carry information over and above other economic series, they may still be useful for monitoring economic developments and for real-time forecasting of economic activity. This is because they are available before official quantitative statistics and are subject only to limited revisions (Santero and Westerlund, 1996). In the euro area, for instance, official statistics are released at least 45 days after the reference

month (e.g. data for January is only available by mid-March). Business surveys are usually available before the end of the reference month (e.g. the Italian survey data are released about 45 days before industrial production). Confidence indicators can provide valuable information on the evolution of the economy over this period, which is one of the reasons why they are popular (Parigi and Golinelli, 2004). In this sense, even if the confidence indicators are coincident indicators of real activity, that they are available earlier means that they are quasi-leading indicators.

A number of studies have demonstrated that confidence indicators are useful for nowcasting economic activity. Giannone, Reichlin and Small (2008) examined how key data releases in the US influenced forecasts of GDP and inflation in real time. They used a dynamic factor model together with a Kalman filter to compute real-time forecasts on the basis of unbalanced panels (due to staggered data-release dates). The methodology reduced the dimensionality problem faced by forecasters in real-time, by assuming that comovement in the economy could be described by a few of common factors. They found that survey data were important in determining predictions, particularly for real GDP growth, although this was mainly due to their timeliness.

Girardi and Reuter (2017) evaluated the impact of new releases of financial, real and survey data (using the European Commission surveys) on nowcasting euro-area GDP in real time. They found that survey and real data improved forecast accuracy throughout the sequence of nowcasts. Confidence indicators contained predictive content even after controlling for timeliness, due to their broad sectoral coverage and forward-looking nature. Similarly, Matheson (2010) found that business survey indicators improved real-time forecasting accuracy. This was due not only to their timeliness, but also to the underlying quality of the data. The results were consistent with the literature showing that survey indicators are not only timely proxies for hard data, but also contain complementary information for understanding business cycle developments.

Relatively few studies have analysed confidence indicators in South Africa. Pellissier (2002) examined the ability of the two South African business confidence indicators, the BER BCI and SACCI BCI, as business cycle indicators. He argued that both BCIs seemed to exhibit a coincident rather than a leading relationship with the business cycle, and that the BER BCI seemed to display stable turning point attributes. More recently, Laubscher (2014) found that the BER BCI was one of the closest predictors of the official reference business cycle turning points and was useful as a leading indicator. The BER's BCI is also used by the SARB as one of the component series of its official leading indicator of the business cycle (Venter, 2005).

The BER BCI has occasionally been included in larger datasets in forecasting exercises. For instance, Gupta, Jurgilas and Kabundi (2010) analysed the impact of monetary policy on house price growth in South Africa using a factor augmented vector autoregression. The models were based on 241 quarterly series, including real, nominal, financial and intangible variables, such as confidence indices. Gupta and Kabundi (2011) used similar large factor models, with a large cross-section of macroeconomic time series, to forecast per capita growth, inflation, and the interest rate. Confidence indices were also included in the dataset of 267 quarterly series.

Kabundi, Nel and Ruch (2016) included the BER Consumer Confidence Index (BER CCI) and the SACCI BCI to forecast real GDP growth in South Africa in real time. They argued that the timeliness of the variables was especially important. The BER CCI and the SACCI BCI are published four and two weeks before the end of the reference quarter respectively. This implies that soft data can be useful in forecasting exercises.

In this chapter, the relationship between business confidence and real activity in South Africa is examined. An attempt is made to establish whether there is a significant positive relationship between the indicators and real GDP growth, the timing of this relationship, and whether it remains significant after taking other economic variables into account. The following section turns to the literature on uncertainty.

### 3.3 Uncertainty

Knight (1921) defined uncertainty as agents' "inability to forecast the likelihood of events happening." Uncertainty refers to a lack of knowledge of the set of possible outcomes and their associated probabilities, because the outcome is highly unique or complex, which makes forecasting difficult. According to this definition, uncertainty is distinct from the concept of risk, which refers to a known probability distribution of a set of outcomes (either through calculation *a priori* or from statistics of past experiences). For example, a coin toss entails risk, because there is a known probability distribution (e.g. 50% chance of heads). In contrast, the number of coins ever produced entails uncertainty, because the calculation would require estimating the distribution of all the coins minted in all countries throughout history (Bloom, 2014). While researchers occasionally refer to a mixture of risk and uncertainty (Bloom, 2014), this chapter focuses on uncertainty in the sense of a lack or predictability.

This section begins with a review of the theoretical links between uncertainty and economic outcomes. The section then turns to the empirical literature, by first discussing measurement challenges and the approaches to operationalising the definition of uncertainty, and then examining the evidence on the impact of uncertainty on economic outcomes.

#### 3.3.1 Theory on Uncertainty and Economic Outcomes

The theoretical literature emphasises two negative and two positive channels through which uncertainty can influence economic activity. Most of the focus is on 'real options' theory, based on Bernanke (1983). Uncertainty may have economic consequences when there is a degree of irreversibility to firms' actions. Firms receive new information over time, reducing uncertainty and increasing their ability to undertake the optimal investment. If the value of time, i.e. the benefit of new information, exceeds the costs of committing to a suboptimal project, it is rational to wait before committing to an investment (Binding and Dibiasi, 2017). Because it increases the value of

waiting for new information, uncertainty delays the current rate of investment (Bernanke, 1983). Thus, the option value of waiting increases as uncertainty increases (Bloom, 2014).

This theory has led to the idea of a ‘wait-and-see’ effect (Bloom, 2009). If a firm faces large fixed adjustment costs,<sup>21</sup> higher uncertainty about future demand makes new investments and hiring less attractive. Firms try to minimise the number of times this fixed adjustment cost must be paid. When the future is uncertain, in the sense that demand could be either very high or low, it makes sense to wait until the uncertainty has been resolved (Bachmann, Elstner and Sims, 2013). Facing a more uncertain environment, firms delay investment and hiring, i.e. they ‘wait and see’ how the future will unfold, which leads to a decrease in economic activity. As the future unfolds, there is pent-up demand for capital and labour. Firms are closer to their adjustment triggers in subsequent periods, leading to a rebound and even an overshoot in economic activity. Thus, the initial decrease is followed by a swift recovery and overshoot in economic activity (Bachmann, Elstner and Sims, 2013).

Uncertainty can also negatively affect economic activity through risk aversion and risk premia. If investors are risk averse, higher uncertainty increases risk premia, by increasing the probability of default (Redl, 2015). The accompanying increase in borrowing costs can reduce growth, as highlighted in studies of uncertainty under financial constraints (summarised in Bloom (2014) and Bachmann, Elstner and Sims (2013)). In models where agents have pessimistic beliefs, and uncertainty about the future is too high to form a probability distribution, agents act as though the worst outcomes will occur (so-called ambiguity aversion). As uncertainty increases and the range of possible outcomes increases, the worst possible outcome becomes worse, leading agents to decrease investment and hiring. In contrast, if agents are optimistic (they assume the best case), uncertainty can have a positive impact on activity (Bloom, 2014).

Bloom (2014) also referred to two other channels through which uncertainty can have a positive effect on economic activity. The ‘growth options’ argument is based on the idea that uncertainty can create call option effects, whereby uncertainty may increase investment if the size of the potential prize increases. This is due to the potential for an increase in upside gains, while the downside loss is limited to initial sunk costs, leading to an increase in expected investment returns (Redl, 2015).

The Oi-Hartman-Abel effect highlights the possibility that firms may be risk-loving if they can expand to exploit good outcomes and contract to insure against bad outcomes. For example, if a firm can easily double production if prices increase, and halve production if prices decrease, it should desire a mean-preserving increase in uncertainty. In effect, the firm can partly insure against bad outcomes by contracting and can exploit good outcomes by expanding. For this mechanism to work, firms need to be able to expand or contract easily in response to good or bad outcomes. Bloom (2014) argued that this effect is not very strong in the short run because of adjustment costs,

---

<sup>21</sup>For capital, these costs can be both physical (equipment may have been damaged in installation and removal) and financial (discounts for used goods). For labour, adjustment costs include recruitment, search frictions, training, and severance pay.

but may be more powerful in the medium to long run.

Boncianni and Van Roye (2016) argued that in a general equilibrium framework, these effects may or may not be completely offset. In a New Keynesian Model, for instance, the monetary authority can partially offset the negative effects of uncertainty by reducing the interest rate. They argue that this is the most important reason why many studies do not find a strong effect. However, when the monetary authority is constrained by the zero lower bound, or when there is imperfect pass-through, the effects of uncertainty become more significant, as the central bank cannot perfectly respond to the shock.

The theoretical literature therefore sets out potential channels through which uncertainty may have a positive or negative impact on economic activity. It then becomes an empirical question to determine the direction and significance of the impact. The following section provides a review of the empirical literature on uncertainty.

### **3.3.2 Empirical Findings**

The recent surge in research on uncertainty has been driven by the idea that uncertainty increased during the financial crisis, and its potential role in shaping the Great Recession. In addition, the availability of empirical proxies for uncertainty has increased, along with the ability to include uncertainty in a wide range of models (Bloom, 2014). Although the majority of studies seems to find that uncertainty indicators are at least negatively related to real economic activity, the findings have not been conclusive. The inconclusive findings may be due to the two main challenges when it comes to empirical work on uncertainty: how to construct proxies for uncertainty and how to distinguish a separate causal impact of uncertainty.

#### **3.3.2.1 Measuring uncertainty**

It is unsurprising that there is no perfect measure of uncertainty, given its broad definition and the potential influence of a broad range of factors. A wide range of proxies for uncertainty have been proposed in the literature. These proxies can be grouped into five major categories, depending on the nature of the data used for their construction (Bloom, 2014). All proxies for uncertainty measure a specific type of uncertainty, and have strengths and weaknesses.

The first category uses financial data, with the majority of studies using as proxies the implied or realised volatility in the stock market, GDP, bond yields, and exchange rates. The rationale is that more volatile series are more difficult to forecast, and are associated with a greater degree of uncertainty (Bloom, 2014). Bloom (2009), Boncianni and Van Roye (2016) and Leduc and Liu (2016), for instance, used stock market volatility as a proxy for uncertainty. A popular proxy is the Chicago Board Options Exchange Market Volatility Index (VIX), which focuses on the implied volatility of the S&P 500 Index. It reflects the dispersion of market participants' estimates of future stock

prices, as measured by the implied volatility across all options with a given time to maturity. The most frequent criticism is that developments on stock markets may only partly reflect developments in the real economy (Girardi and Reuter, 2017).

The second category uses new information to construct uncertainty indicators. The most prominent examples are proxies based on references to ‘uncertainty’ in the media. Baker, Bloom and Davis (2015) and Baker, Bloom and Davis (2016), for instance, developed economic policy uncertainty indices based on the frequency of references to policy uncertainty in newspapers. Baker, Bloom and Davis (2015) combined this text mining measure with disagreement among forecasters on future government purchases and inflation, and the number of tax code provisions about to expire to create an overall indicator. One criticism is that the selection of newspapers and search terms entails a certain degree of subjectivity (Girardi and Reuter, 2017).

The third category is derived from the disagreement among professional forecasters. The rationale is that a larger dispersion of opinions about the future indicates a higher degree of uncertainty. Popescu and Smets (2010), for instance, used a proxy for uncertainty based on the dispersion of professional forecasts of consumption, industrial production, investment, output, prices and interest rates in Germany. The downside is that the factors influencing a limited set of professional forecasters might differ from those influencing producers and consumers (Girardi and Reuter, 2017).

The fourth category uses the responses from business and consumer surveys. Bachmann, Elstner and Sims (2013), for instance, used the dispersion of business survey responses, as well as the dispersion in individual forecast errors to construct proxies for the US and Germany. Arslan *et al.* (2015) used squared forecast errors to construct uncertainty indicators for Turkey. Girardi and Reuter (2017) derived indicators of dispersion from the aggregated responses on multiple forward-looking questions. Leduc and Liu (2016) also used a survey-based proxy for uncertainty, measured as the fraction of respondents who listed uncertainty as a factor limiting their spending plans. Survey-based measures have the advantage that they are derived from opinions of key economic agents, as opposed to outside observers (e.g. professional forecasters) or the choices of investors on financial markets (Girardi and Reuter, 2017).

A fifth category was introduced by Jurado, Ludvigson and Ng (2015). They argued that indicators of uncertainty should reflect the common variation across a vast array of variables, and that the forecastable component of each series should be removed when calculating volatility. They constructed new indicators using a large dataset of macroeconomic and financial indicators, as well as firm-level data. They extracted common factors, used them to predict industrial production, and subsequently calculated the forecast errors. Increases in the volatility of forecast errors were interpreted as increases in uncertainty. The disadvantage of their indicator is that it is an *ex post* measure, which requires the actual outcome of the forecasted time series before computing the indicator (Girardi and Reuter, 2017).

A few studies have constructed proxies for uncertainty in the South African context. Redl (2015)

constructed an index of uncertainty for South Africa, based on disagreement among professional forecasters, the number of newspaper articles that mentioned economic uncertainty in South Africa, and references to uncertainty in the SARB's Quarterly Review.

Hlatshwayo and Saxegaard (2016) created a measure for South African economic policy uncertainty, by looking at 'news chatter' in the media, similar to the method used in Baker, Bloom and Davis (2016). They created both economic policy and political uncertainty indices at the sectoral and aggregate level, by counting the number of articles that matched specific search algorithms. Aggregate economic uncertainty, for example, was measured by counting articles containing 3 mentions of words related to policy, economics, and uncertainty (i.e. one mention of each area) within 10 words of 'South Africa'. The absolute counts were normalised and the indices were standardised. McClean (2015) created a similar news-based index for aggregate South African policy uncertainty. He found a moderate correlation between this index, the South African Volatility Index (SAVI) and SA government bond yields.

Pellissier and Fusari (2007) used the BER's manufacturing surveys to construct a measure of uncertainty. 'Volatility' in survey expectations was derived from the (unweighted) percentage of survey respondents changing their expectation between survey periods. 'Realizations' of survey expectations was derived from changes in survey expectations in period  $t - 1$ , compared with survey realisations in period  $t$ . They found a negative relationship between 'Volatility' and 'Realizations' for responses relating to business conditions, production, sales, fixed investment and prices. Hart (2015) also used the BER's manufacturing sector survey to create dispersion measures of uncertainty, similar to the method used in Bachmann, Elstner and Sims (2010).

Recently, North-West University (2016) created a policy uncertainty index for South Africa. The index has three components: the frequency of references to economic policy uncertainty in leading publications, expert opinions drawn from leading private sector economists, and responses from the BER manufacturing survey on whether the political climate is a constraint to doing business. This index is only available from July 2015.

None of these studies has fully exploited the information contained in the BER business tendency surveys. This chapter explores aggregation methods to try to improve on the existing measures of uncertainty for South Africa, using the microeconomic data from BER business tendency surveys.

### **3.3.2.2 The impact of uncertainty**

The majority of studies seems to find at least a negative relationship between uncertainty proxies and economic activity, although this does not necessarily imply causality. In the literature three approaches have been taken to identify the impact of uncertainty on activity (Bloom, 2014). The first approach uses structural models to identify the potential impact of uncertainty shocks. The second approach relies on timing, typically in a VAR framework, by estimating the movements in economic activity that follow changes in uncertainty. The third approach exploits natural experiments such



as exchange rate movements, disasters, and political coups.

In a number of papers, structural models (i.e. DSGE models) have been used to investigate potential mechanisms through which uncertainty may influence economic activity. Empirical VAR models are then used to confirm the theoretical model predictions. In a seminal paper, Bloom (2009) used a structural model to simulate the impact of an uncertainty shock, which produced the rapid decrease and subsequent rebound in aggregate output and employment predicted by the ‘wait-and-see’ effect. This simulated impact was compared with VAR estimations on actual data, using stock market volatility as a proxy for uncertainty. The results matched in both magnitude and timing, with a shock to uncertainty generating a decrease and then an overshoot in employment and production.

Bloom, Bond and Van Reenen (2007) developed a model of firms’ investment decisions to show that, with partial irreversibility, the impact of a firm-level demand shock on investment tends to be weaker for firms that are subject to higher uncertainty. They found evidence of more cautious investment behaviour for firms subject to higher uncertainty. Leduc and Liu (2016) used a structural model with nominal rigidities and search frictions to show the mechanism through which uncertainty could produce large economic effects. Their empirical model found that uncertainty shocks resembled aggregate demand shocks, reducing investment, short-term interest rates and inflation, and increasing credit spreads and unemployment. Bonciani and Van Roye (2016) investigated the impact of uncertainty under financial frictions with a structural model. They found that higher uncertainty reduced activity, and that the impact was potentially larger during a recession.

A number of studies have investigated the timing of the relationship between uncertainty and economic activity in a VAR framework (e.g. Arslan *et al.*, 2015; and Jurado, Ludvigson and Ng, 2015; Baker, Bloom and Davis, 2016; Girardi and Reuter, 2017). The results were generally similar to Bloom (2009), with a positive shock to uncertainty followed by a significant decrease in output, investment and employment.

Bachmann, Elstner and Sims (2010) found that innovations to their survey-based uncertainty indicators had prolonged negative effects on economic activity. The long-run effects of uncertainty shocks were similar to the long-run effects of negative confidence shocks. However, when uncertainty was restricted to have no long-run impact, which is what the ‘wait-and-see’ effect would predict, it did not have a significant impact on activity. They argued that uncertainty could be seen as a symptom of poor economic times rather than a causal mechanism. In a follow-up study, Bachmann, Elstner and Sims (2013) found that positive shocks to uncertainty were associated with a significant decrease in production and employment in both Germany and the US. German production declined and rebounded relatively quickly following an increase in uncertainty, while the response of US output was protracted, with limited evidence of a rebound. The US results suggest that some of the other mechanisms proposed in the literature, such as financial frictions may be important.

A few studies have investigated the interaction of uncertainty and these financial frictions. Popescu and Smets (2010), for instance, argued that once a measure of financial stress is included in the



regressions, the independent role of uncertainty shocks becomes minimal. They found that the real effects of financial risk premia were larger and more persistent than uncertainty effects. Caldara *et al.* (2016) found that uncertainty shocks had a significant negative impact on both financial conditions and real economic activity. Their results suggested that increases in uncertainty associated with tighter financial conditions had a particularly large negative effect on real economic activity.

Other studies have exploited natural experiments such as disasters, political coups, and exchange rate movements. For instance, Baker and Bloom (2013) used natural disasters, terrorist attacks and unexpected political shocks as instruments for the usual stock market proxies of uncertainty. They found that uncertainty shocks accounted for at least half of the variation in GDP growth. Binding and Dibiassi (2017) showed how different uncertainty indicators reacted to an unexpected policy change when the Swiss National Bank decided to return to a floating exchange rate regime in 2015. Firms affected by this exogenous increase in uncertainty decreased their planned investment relative to firms that were unaffected. However, once they controlled for the degree of irreversibility of firm investment, the relationship was no longer significant.

There is relatively little evidence on the impact of uncertainty on economic outcomes in South Africa. Developing countries, such as South Africa, tend to experience higher uncertainty because they tend to have less-diversified economies, which are more exposed to price and output fluctuations of volatile goods such as commodities (Bloom, 2014). Developing countries tend to have more political shocks and often have less effective stabilisation policies. Given that developing countries experience higher levels of uncertainty, it is possible that fluctuations in uncertainty have a more pronounced impact on output.

Redl (2015) argued that analysing uncertainty in developing countries could help to distinguish between the effects of financial and uncertainty shocks. During the Great Recession, many developing countries experienced high uncertainty, while not undergoing the same levels of financial stress as developed countries. He found that an increase in uncertainty in South Africa was associated with subsequent decreases in output, investment, employment, and asset prices. The results were robust to the inclusion of consumer confidence and credit spreads as a measure of financial stress, although the sizes of the responses were moderated.

Hlatshwayo and Saxegaard (2016) explored the role of policy uncertainty in South Africa in reducing the responsiveness of exports to relative price changes, through the wait-and-see effect. They found that increased policy uncertainty reduced the responsiveness of exports to the real effective exchange rate and had short- and long-run effects on export performance. A measure of competitiveness that adjusted for uncertainty and supply-side constraints outperformed the real effective exchange rate in tracking export performance. Similarly, Boshoff (2008) argued that developments in the Rand did not translate into business cycle movements in the South African economy, and that a weaker exchange rate was less likely to boost either foreign investment or export performance in the face of regulatory uncertainty.

Hart (2015) investigated the relationship between sentiment and economic activity in the South African manufacturing sector from 2001Q2 to 2014Q2. The study closely followed Bachmann, Elstner and Sims (2010), which also measured uncertainty in the manufacturing sector using business survey data. A VAR framework was used to estimate the impact of confidence and uncertainty on investment, production and employment in the South African manufacturing sector. None of the uncertainty measures were found to be significant, possibly due to the limited sample period.

In this chapter, the relationship between uncertainty and real activity in South Africa is examined, using standard agnostic econometric methods (VARs). In the following section the BER business tendency surveys used to create the sentiment indicators are discussed.

### **3.4 Data: Business Tendency Surveys**

Business tendency surveys are conducted to obtain qualitative information that is useful in monitoring the current business situation and in forecasting developments in the business cycle. This is reflected in the extensive use of confidence measures as leading indicators of the business cycle (Organisation for Economic Co-operation and Development, 2003). Qualitative surveys often can be completed more easily and quickly than quantitative surveys. The results can be published before official statistics, which are often released with a significant delay. Survey data have the advantage of focusing on the assessments and expectations of economic developments by relevant economic decision makers. Variables related to expectations may reflect cyclical changes earlier than corresponding quantitative statistical series (i.e. expectations lead to plans that are implemented and which will then be picked up in quantitative statistics).

#### **3.4.1 The BER Business Tendency Surveys**

The BER, a research institute attached to Stellenbosch University, has been conducting business tendency surveys in South Africa since March 1954. The BER's quarterly business surveys are similar to the business tendency surveys conducted all over the world, including the European Commission Business Tendency Surveys, the German Ifo Business Climate Survey, the Federal Reserve Bank of Philadelphia's Business Outlook Survey, and the Bank of Japan's Tankan Survey (Organisation for Economic Co-operation and Development, 2003).

During the last month of each quarter, questionnaires are sent to approximately 1,000 firms in each of the manufacturing and services sectors and 1,400 firms in each of the construction and trade sectors (i.e. retail, wholesale and motor vehicles). The questionnaires are completed by senior executives of the firms. The questions have remained largely unchanged since inception, and include those on current and expected future developments regarding, among others, sales, orders, inventories, prices,

employment, and constraints. For the most part, the survey answers fall into three categories: ‘up’, ‘the same’ or ‘down’.

Stratified deliberate sampling is used to design the BER’s survey panels, which is the international norm. Participants are selected to be representative of particular sectors, regions and firm sizes. The respondents are reviewed periodically to ensure reasonable representation of the population universe. The exact number of firms in the universe is unknown to the BER, as censuses of the business sector are not conducted regularly and the BER does not have access to the National Business Register (Kershoff, 2002). Practical experience has shown that non-random samples can give acceptable results in conducting these types of surveys (Organisation for Economic Co-operation and Development, 2003).

The BER makes no provision for firms that were not selected or did not respond during sampling, implicitly assuming that the non-participating or non-responding firms have the same distribution as the responding firms for the period. This corresponds with the ‘missing at random’ assumption, which is typically used internationally (European Commission, 2006). Kershoff (2015) argued that this is a reasonable assumption, given that the responses cannot vary infinitely, and the same factors influence firms in the same sector. He found evidence for this assumption when the inclusion of latecomers had almost no effect on the volatility and tracking record of the results, even at lower levels of aggregation.<sup>22</sup>

The sample of firms remains relatively stable from one survey to the next, effectively creating a panel. The panel is partly fixed and partly rotating, as inactive firms that fail to respond for a period of two years are removed and replaced with new firms. The fixed part reflects the opinions of the same firms over time, which ensures that the results remain comparable between surveys. The results are more likely to reflect changes in the variables under consideration than changes in the sample from one survey to the next (Kershoff, 2002).

Table 3.1 reports the details of the survey data. The sample runs from 1992Q1 to 2016Q3, although the survey of the services sector started only in 2005Q2. Figure 3.1 illustrates the number of respondents over time per sector. Around 1,000 completed questionnaires are received every quarter, leading to an overall sample size of 119,438. All of the surveys have a few missing quarters, when the microeconomic data was lost. The overall panel sizes have remained relatively stable over time.

Figure 3.2 illustrates approximate response rates for the four sectors, assuming that exactly 1,000 questionnaires were sent to firms in the manufacturing and services sectors and exactly 1,400 to the construction and trade sectors.<sup>23</sup> The response rates vary over time and per sector, and are

---

<sup>22</sup>The BER does not adjust individual weights for changes in the response pattern. No calibration or post-stratification is carried out to correct the estimated value. Missing items (specific questions) and missing responses (questionnaires) are not imputed, and the results are not revised to provide for questionnaires received after the results have been processed (Kershoff 2015).

<sup>23</sup>The response rates for the construction sector and the totals are adjusted slightly to take account of the fact that three subsectors (architects, quantity surveyors and civil engineers) were only available from 2001Q2. This adjustment was made by adding the average response rate of these subsectors.

## Business Sentiment

Table 3.1: Sample characteristics

Sector	Sample	Total Obs	Obs/Quarter	Response Rate	Missing Quarters
Manufacturing	1992Q1-2016Q3	36915	384.53	0.38	1997Q4,2000Q1,2005Q4
Construction	1993Q2-2016Q3	28139	312.66	0.26	1993Q4,1998Q3,2000Q2,2005Q4
Trade	1992Q2-2016Q3	40480	426.11	0.30	1992Q4,1993Q3,2005Q4
Services	2005Q2-2016Q3	13904	308.98	0.31	2005Q4
Total	1992Q1-2016Q3	119438	1218.76	0.33	2005Q4

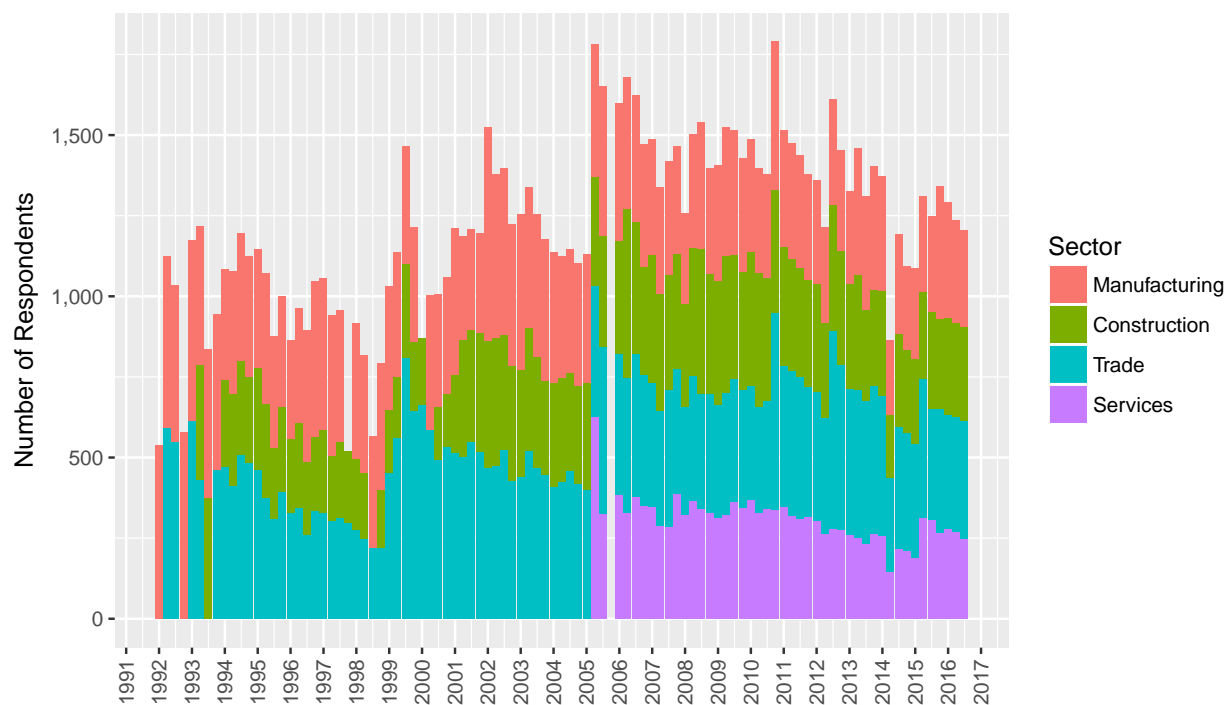


Figure 3.1: Number of respondents over time, by sector (1992Q1-2016Q3)

relatively low by international standards (Kershoff, 2015). The total approximate response rate varied from around 20% to around 60% over the period. Response rates are around 10% higher if the inactive respondents are excluded. The response rates do not seem to follow cycles in real economic output, and exhibit insignificant correlations with real GDP growth.

The spikes correspond to periodic recruitment drives by the BER. Every two to three years the BER removes slightly more than 25% of all respondents from the panel, because they became inactive. The BER tries to ensure that the new recruits are representative of the population, but this does mean that few firms are present throughout the sample period. While the sample of firms remains relatively stable for consecutive surveys, over longer periods the firms respond sporadically and enter and exit the sample often. Since 2005Q2, for instance, when all the data became available, only 6% of respondents replied 75% or more of the time (only 3 firms responded in all periods), 13% replied between 50% and 75% of the time, 19% replied between 25% and 50% of the time and 62% replied less than 25% of the time.

## Business Sentiment

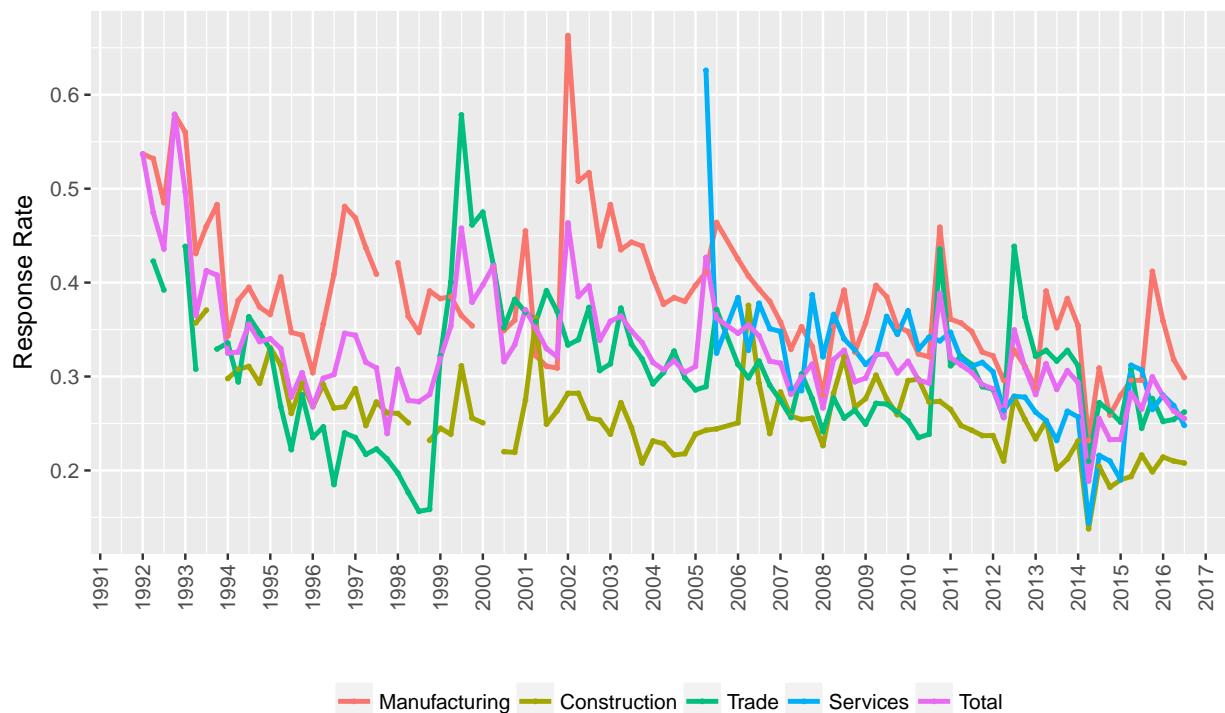


Figure 3.2: Approximate response rates over time per sector (1992Q1-2016Q3)

Panel sizes and response rates determine the representativeness of the sample. In order to be representative, panels have to include a minimum number of participants, which depends on the level of aggregation and the size of the population universe. The results often remain valid even if the sample size is small and the response rate relatively low. According to the Organisation for Economic Co-operation and Development (2003), even as few as 30 respondents might be sufficient to obtain an acceptable level of precision for each stratum. This is because the variance of responses for ordinal-scaled data based on a stable panel is lower than for quantitative data derived from independent surveys. Moreover, certain activities are dominated by a few large firms. Representativeness therefore has a smaller impact on qualitative survey results than on quantitative survey results. A panel that is not fully representative will probably produce similar results to a fully representative one (Kershoff, 2002).

The sample sizes illustrated in Figure 3.1 therefore seem adequate to uncover trends in the data. Kershoff (2002) found that the degree of representation of the BER's construction and trade panels adequately reflects the universes, taking response rates into account and comparing the composition of the survey panels with census and other official data. However, the number of participants per subsector may be too low to consider subsectors or provinces as sub-panels. The survey responses are therefore not disaggregated further into subsectors below.

In order to test whether the attrition rates of firms drive the results, a number of robustness exercises are carried out and reported in the chapter Appendix (section 3.10). The indicators are calculated

by including only firms that form part of smaller, more ‘stable’, samples. The smaller samples include firms that only responded in consecutive surveys, firms that responded to more than half of all the surveys, and firms that responded to more than 75% of all the surveys, respectively. The indicators based on these smaller samples are similar to those for the full sample. This implies that these firms are driving the results, rather than the entry and exit patterns of firms.

### 3.4.2 The BER Business Confidence Indicator

The BER uses these business tendency surveys to construct its business confidence indicator. The BER BCI has proved useful as a leading indicator of the business cycle and economic growth in South Africa. It is used as one of twelve leading indicator series by the SARB to date official turning points in the business cycle. Laubscher (2014) also found that it can improve estimates of cyclical turning points. This is particularly useful in view of the early availability of the index. The BER index results for a particular quarter are available approximately two months before the official GDP estimates (Kershoff, 2000).

In calculating business confidence, the most important issues are which survey questions to use and which weights to apply to the responses. The BER BCI is constructed from a specific question (Q1) that appears in all of the surveys: “Are prevailing business conditions: satisfactory, or unsatisfactory?” The BCI is the weighted percentage of respondents who rated prevailing business conditions as ‘satisfactory’ in a particular sector. The BCI is therefore a rating of business conditions at a specific point in time.

According to Kershoff (2000) there are two reasons for the use of this one question to construct the confidence indicator. Firstly, it is reasonable to assume that respondents who are satisfied with business conditions will have more confidence than those experiencing unsatisfactory conditions. Secondly, respondents take a variety of factors into account when rating prevailing business conditions, which solves the problem of weighting different factors (Kershoff, 2000). The Organisation for Economic Co-operation and Development (2003) argues that responses on general business conditions are usually based on a combination of factors, such as order book appraisals, expectations of interest rates, exchange rates and political developments.

In line with international best practice, all survey responses are weighted (except for the building survey). Each response is multiplied by a factor, which is calculated as the product of a firm size weight and a subsector size weight (except for the motor trade, where there are no subsectors). Each firm receives a weighting in relation to turnover, or the size of workforce in the case of manufacturing.<sup>24</sup> The subsector size weights are based on the composition of production or sales in each subsector, as calculated by StatsSA. The BER does not apply sample weights, as it does

---

<sup>24</sup>Unlike quantitative surveys, where weighting is usually inherent in the variables, weighting is necessary for qualitative surveys, because the variables typically collected do not inherently represent the size of a business. The size weights are necessary because the economic significance of the responses should reflect the size of the firm (UN 2015).

not have access to the National Business Register and cannot calculate selection probabilities.<sup>25</sup> Responses are weighted by firm size and subsector size to obtain five sectoral indices: manufacturing, building contractors (other construction subsectors are omitted), retailers, wholesalers and new vehicle dealers. The BER BCI is calculated as the unweighted mean of the five sectoral indices (services are excluded altogether).

The BER BCI is a measure of current conditions, based on a single question, with survey responses weighted in an ad hoc manner. The business surveys contain a number of questions, all of which potentially have an impact on business confidence. A composite indicator can be calculated by combining the responses to a number of questions, which is often done internationally (European Central Bank, 2013). Moreover, the BER BCI reflects confidence in current conditions, rather than forward-looking confidence. As the surveys contain questions on expectations, forward-looking responses may also provide valuable information.

### 3.5 Methodology

This chapter builds on the BER BCI by calculating composite weighted indicators of confidence on current and expected conditions, as well as composite weighted forward-looking indicators of uncertainty, at a sectoral level and in aggregate. This section describes the methodology for calculating the sentiment indicators based on the microdata from the BER business tendency surveys.

The indicators are based on subjective survey responses, and therefore prone to bias. Tversky and Kahneman (1974) showed that agents rely on a number of heuristics, which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general these heuristics are useful, but sometimes lead to severe and systematic biases. The heuristics and the accompanying reporting biases are important for how agents respond to subjective survey questions.

Anchoring and adjusting is one such heuristic, which entails anchoring with what is well-known, easily recalled from memory, or salient, and then adjusting from that anchor (Tversky and Kahneman, 1974). With anchoring, a respondent's view of the future is anchored in how they feel at present. Moreover, Gehlbach and Barge (2012) showed that survey respondents use anchoring and adjusting, where their response to an initial survey item provides an anchor from which they (insufficiently) adjust in answering the subsequent item, especially when adjacent items on the survey are similar. Thus, over the course of a survey, responses to adjacent item-pairs are likely to be more similar than responses to the same item-pairs in non-adjacent positions. Because the questions in the BER surveys are similar, and the questions on current conditions and those on expected conditions are

---

<sup>25</sup>The BER does not apply sample weights (i.e. the inverse of the probability of selection). This assumes that the probability of selection is the same for all units, which would be the case if firms were selected randomly (OECD 2003).



adjacent, this bias may well be present. The subjective survey responses and the resulting subjective confidence and uncertainty indicators will consequently reflect this bias.

Arguably, there are three types of information contained in these survey responses (Fuhrer, 1988). The first reflects current developments or economic news, not yet reflected in currently available standard macroeconomic time series (e.g. changes in firms' inventory levels). The second type reflects forward-looking information, such as agents' probabilistic assessments of uncertain future policy changes (e.g. impending tax legislation). The third type reflects 'animal spirits', where agents feel optimistic or pessimistic about future prospects for reasons not tied to fundamentals.

The significant correlations between the subjective survey-based indicators and real output in the literature (as well as in this chapter), suggest that the indicators capture at least one of these types of information. The first type of information anticipates data which will be released later. The second type may provide information about events that are either difficult to quantify or predict from the past (e.g. agents assessments of impending policy changes) (Fuhrer, 1988). Thus, they summarise changes in agents' beliefs about the future, i.e. their private information (Acemoglu and Scott, 1994). If agents act on animal spirits, which are reflected in survey data, the third type of information will explain subsequent economic outcomes due to self-fulfilling behaviour (Fuhrer, 1988).

When the agents respond to the questionnaires, they are most likely making an estimate that is partly based on the fundamentals (the first two types of information) and partly based on psychological factors or animal spirits, all of which probably contain biases.

To the extent that the indicators reflect psychological factors, the biases capture the psychological phenomena of confidence and uncertainty - i.e. agents' perceptions or degree of optimism about the future (confidence) and their inability to forecast future outcomes (uncertainty). If animal spirits influence behaviour, over and above fundamentals, these biased measures will determine agents' decisions to some extent and thereby might influence the business cycle. Thus, biased measures are still of interest if they reflect possibly biased psychological factors.

To the extent that the indicators summarise fundamental information, possibly in a biased way, the indicators still provide timely information on current developments, or information on expectations about events that are difficult to quantify. The proof of the usefulness of these potentially biased measures of fundamentals will be in their co-movement with output. Even if the indicators are subject to biases and measurement error, they still seem to contain useful additional information on agents' expectations that is not contained in standard macroeconomic variables (Fuhrer, 1988).

### 3.5.1 Confidence

Formally, one can define a  $k$ -period-ahead expectations measure of confidence ( $C_t^k$ ) at time  $t$  as:  $C_t^k = E_t f(\Delta^h Y_{t+k})$ , where  $Y_{t+k}$  is a measure of real activity (usually output) at time  $t+k$  and



$\Delta^h Y_{t+k} = Y_{t+k} - Y_{t+k-h}$ . A common definition of  $f(\Delta^h Y_{t+k})$  relies on an up, unchanged, or down classification (e.g. Q2A in the BER survey):

$$f(\Delta^h Y_{t+k}) = \begin{cases} -1, & \text{if } \Delta^h Y_{t+k} < 0 \\ 0, & \text{if } \Delta^h Y_{t+k} = 0 \\ 1, & \text{if } \Delta^h Y_{t+k} > 0 \end{cases}$$

An alternative would be to use a binary classification in levels (e.g. Q1 in the BER survey):

$$f(Y_{t+k}) = \begin{cases} -1, & \text{if } Y_{t+k} < a \\ 1, & \text{if } Y_{t+k} \geq a \end{cases}$$

where  $a$  is determined by the preferences of the agent. In this case  $a$  is the subjective benchmark or threshold that determines when conditions are ‘satisfactory’, and the measure of confidence simplifies to:  $C_t^k = E_t f(Y_{t+k})$ .

In this chapter, a distinction is made between indicators of current conditions  $C_t^k$  when  $k = 0$ , and indicators of expected conditions  $C_t^k$  when  $k = 1$ . Both the Conference Board and the University of Michigan make this distinction and report two consumer confidence indices: a current conditions component and an expectations component (Ludvigson, 2004). The confidence measure for current conditions  $C_t^0$  is referred to as ‘current’, as it reflects confidence about the current quarter (in the second month of the quarter). The confidence measure for expected conditions  $C_t^1$  is referred to as ‘expected’, as it reflects confidence about the following quarter.

The BER business tendency surveys make this distinction possible by asking for separate responses relating to current and expected future conditions. The questions on current conditions (e.g. Q2A) all have the following format: “(Estimated development in current quarter) Compared with the same quarter of a year ago, are general business conditions: better, the same, or poorer?” In other words, these questions ask whether the factor under consideration in time  $t$  is better, the same, or poorer, compared with  $t - 4$ .

The forward-looking questions (e.g. Q2P) all have the following format: “(Expected development in next quarter) Compared with the same quarter of a year ago, will general business conditions be: better, the same, or poorer?” As with the questions on current conditions, these questions ask whether the factor under consideration in time  $t + 1$  is expected to be better, the same, or poorer, compared with  $t - 3$ . Responses are relative to the same quarter of the previous year, which corresponds with year-on-year growth rates.

Although the survey questions imply that seasonal adjustment is not required, a common challenge is that respondents may not use the correct reference period when answering the question (Organisation for Economic Co-operation and Development, 2003). For example, answers to the forward-looking

questions may compare expected outcomes in the next quarter  $t + 1$  with period  $t$ , instead of with period  $t - 3$ . In many cases, the time series of balances show some residual seasonality. The indicators are therefore adjusted for seasonality (United Nations, 2015). The results are similar without seasonal adjustments.

As discussed above, confidence indicators are almost always based on balance statistics, which present a single figure summarising the responses of all participants to a particular question (Santero and Westerlund, 1996). It is the cross-sectional mean of survey responses if the standard quantification system is used: ‘better’ is quantified by +1, ‘the same’ by 0, and ‘poorer’ by -1. Confidence in period  $t$  relating to current conditions  $C_t^0$ , and confidence in period  $t$  relating to expected conditions  $C_t^1$ , may be defined as:

$$C_t^0 = \frac{1}{W_t} \sum_{i=1}^N w_{it} E_t f(\Delta^4 Y_{i,t})$$

$$C_t^1 = \frac{1}{W_t} \sum_{i=1}^N w_{it} E_t f(\Delta^4 Y_{i,t+1}),$$

where  $Y_{i,t+k}$  is again a measure of real activity at time  $t + k$  for firm  $i = 1, \dots, N$ ;  $\Delta^h Y_{i,t+k} = Y_{i,t+k} - Y_{i,t+k-h}$  for firm  $i$ ;  $w_{it}$  is the weight that each firm  $i$  receives at time  $t$ ; and  $W_t = \sum_{i=1}^N w_i$  is the sum of the weights.

The weights are calculated as:  $w_{it} = f_{it} s_{jt} / F_{jt}$ , where  $f_{it}$  the firm size weight (i.e. the inner weight reflecting turnover or number of employees) for firm  $i$  at time  $t$ ;  $s_{jt}$  is the subsector weight (i.e. the outer weight reflecting the share of total value added) for subsector  $j$  at time  $t$ ; and  $F_{jt} = \sum_{i=1}^N f_{it}$  is the total firm weight for subsector  $j$  at time  $t$ . These weights are equivalent to an explicit two-step weighting procedure, whereby weighted means are calculated for each subsector separately (using firm size weights), and then aggregated with the subsector weightings (United Nations, 2015). The BER uses similar weights, except that their weights equal the product of firm and subsector weights  $w_{it} = f_{it} s_{jt}$ , without dividing by the total firm weight for the subsector  $F_{jt}$ .

The weighted means are calculated for each question separately. Although the BER uses a single question to calculate its BCI, the business surveys contain a number of questions that may be useful in gauging business sentiment in South Africa. These include questions on general business conditions, production, orders placed, employment, and profitability. Most international institutions calculate composite confidence indicators by combining the responses to a number of questions (European Central Bank, 2013). Composite indicators react to various sources of economic fluctuations, while being resilient to fluctuations affecting single components. They may therefore exhibit fewer false alarms and fewer missed turning points than indicators based on a single question.

This chapter therefore combines the responses to a number of questions in the BER surveys to calculate composite indicators. For consistency, the composite indicators are derived from questions that are present in most of the sectoral business surveys. Table 3.2 reports the questions included in each of the sectoral surveys. These questions cover five types of variables, namely business

Table 3.2: Survey questions used, by sector

Survey Question	Manufacturing	Construction	Trade	Services
Business Conditions	X	X	X	X
Activity	X	X	X	X
Employment	X	X	X	X
Profitability		X	X	X
Orders Placed	X		X	

conditions, activity (production or sales),<sup>26</sup> employment, profitability, and orders placed. Not all of the variables are covered in all the surveys. The measure of confidence about current conditions also include the question (Q1) on business satisfaction used to calculate the BER BCI. The composite sectoral indicators are calculated as the average of the weighted balances for the questions for each sector, as reported in Table 3.2. The results are very similar when the different questions are combined using principal components rather than averages. The sectoral indicators are then weighted by GDP share to form the overall aggregate composite indicators (United Nations, 2015).

### 3.5.2 Uncertainty

Following the literature (e.g. Bachmann, Elstner and Sims (2013), Arslan *et al.* (2015), and Girardi and Reuter (2017)), this section sets out the methodology for calculating three composite forward-looking indicators of uncertainty: (i) the scaled weighted cross-sectional standard deviation of forward-looking responses, (ii) the weighted cross-sectional mean of individual firm forecast errors, and (iii) the weighted cross-sectional standard deviation of firm forecast errors. The BER survey microeconomic data is particularly useful in this case, as it allows individual firm forecast errors to be calculated.

The cross-sectional standard deviation of responses to forward-looking questions (e.g. Q2P)  $D_t^1$  at time  $t$ , is a measure of the dispersion of responses and is often used as a proxy for uncertainty. This measure of dispersion is analogous to the proxy for uncertainty based on forecaster disagreement used by Baker, Bloom and Davis (2015). It may be defined as:

$$D_t^1 = \frac{1}{W_t} \sum_{i=1}^N (w_{it} E_t f(\Delta^4 Y_{i,t+1}) - \mu_{t+1})^2,$$

where the variables are defined in the same way as above, and  $\mu_{t+1} = \frac{1}{W_t} \sum_{i=1}^N w_{it} E_t f(\Delta^4 Y_{i,t+1})$  is the weighted sample mean.

Bachmann, Elstner and Sims (2013) and Girardi and Reuter (2017) noted that the dispersion proxy for uncertainty described above suffers from a major weakness, in that it reflects changes in two factors other than pure uncertainty. First, time variation in the cross-sectional dispersion

<sup>26</sup>The wording of the questions is adapted to the characteristics of each sector (Kershoff 2015). Activity is referred to as the ‘volume of production’ in the manufacturing survey, ‘volume of building activity’ in the construction survey, ‘volume of sales’ in the trade surveys, and ‘volume of business’ in the services survey.

of responses may simply reflect firms reacting differently to aggregate shocks (i.e. heterogeneity), without uncertainty changing over time. Respondents can have legitimately different views on current and future prospects depending on their characteristics, such as their economic sector, export orientation and dependency on external funding. For instance, in an upswing phase with an appreciating currency, export-oriented firms might switch to more negative assessments, while other firms have more positive assessments. This would drive up the dispersion measure, without uncertainty necessarily increasing.

Second, time variation in dispersion may simply reflect time variation in the heterogeneity of expectations (i.e. disagreement), without uncertainty changing over time. Firms might respond differently to the survey questions because they use different information sets. Their assessments might vary widely and translate into high dispersion, without this necessarily indicating that respondents are uncertain of their assessments.

Accordingly, Girardi and Reuter (2017) suggested scaling the forward-looking dispersion measures  $D_t^1$  in period  $t$  by the dispersion of responses to questions on current conditions  $D_{t+1}^0$  in period  $t + 1$ . The idea is that the possible drivers of dispersion differ between these assessments. The dispersion in the responses to forward-looking questions reflects the ‘natural’ degree of dispersion (from heterogeneity and disagreement), as well as uncertainty about the future. The dispersion of responses on current conditions should be less uncertain than assessments of future conditions, and should depend more on the degree to which conditions differ between respondents, i.e. heterogeneity and disagreement. This proxy therefore measures the extent of uncertainty, expressed as a share of the ‘natural’ dispersion. The scaling operation neutralises some of the impact of ‘natural’ dispersion of the responses to forward-looking questions. The scaled uncertainty indicator should at least be closer to actual uncertainty than dispersion based only on forward-looking question.

The first uncertainty indicator  $D_t$ , or ‘dispersion’, is the weighted cross-sectional standard deviation of forward-looking responses  $D_t^1$  at time  $t$ , scaled by the weighted cross-sectional standard deviation of responses on current conditions  $D_{t+1}^0$  at time  $t + 1$ . More formally:

$$D_{t+1}^0 = \frac{1}{W_{t+1}} \sum_{i=1}^N (w_{it+1} E_{t+1} f(\Delta^4 Y_{i,t+1}) - \mu_{t+1})^2$$

$$D_t = \frac{D_t^1}{D_{t+1}^0}$$

One disadvantage of this indicator is that if economic conditions do not remain broadly stable between the two responses, the scaling might neutralise too much or too little of the dispersion and artificially lower or increase the level of uncertainty. Another disadvantage is that it is an *ex post* measure, which requires the outcome at time  $t + 1$  before computing the indicator (Girardi and Reuter, 2017).

Following Bachmann, Elstner and Sims (2013), individual firms’ forecast errors are used to estimate

Table 3.3: Possible forecast errors

		$Q2A_{t+1}$		
		Better	Same	Poorer
$Q2P_t$	E(Better)	0	-1	-2
	E(Same)	1	0	-1
	E(Poorer)	2	1	0

the other two proxies for uncertainty. The panel dimension of the survey is exploited to construct the *ex post* forecast errors. Pairs of questions are used to construct the forecast errors for each respondent, by comparing the expectations in period  $t$  for a specific question with the realisations for that question in period  $t + 1$ . For instance, the survey responses to Q2P in period  $t$  are used to extract the expectations of general business conditions in time  $t + 1$  relative to  $t - 3$ . The errors are then calculated by subtracting these expectations from the realisations of the responses to Q2A at time  $t + 1$  relative to  $t - 3$ . The forecast errors  $\epsilon_{i,t+1}$  in period  $t + 1$  may be defined as the realisations  $E_{t+1}f(\Delta^4 Y_{i,t+1})$  of a specific outcome in period  $t + 1$  minus the expectations  $E_t f(\Delta^4 Y_{i,t+1})$  in period  $t$  of that outcome in period  $t + 1$ :

$$\epsilon_{i,t+1} = E_{t+1}f(\Delta^4 Y_{i,t+1}) - E_t f(\Delta^4 Y_{i,t+1})$$

Table 3.3 illustrates the nine possible forecast errors. For example, for a firm that expected an improvement in (i.e. better) conditions, the realisation of better conditions would be recorded as a 0 forecast error, no change as a -1 forecast error, and poorer conditions as a -2 forecast error.

Arslan *et al.* (2015) argued that firms make forecast errors because of uncertainty and that forecast errors should be treated as uncertainty. Following Arslan *et al.* (2015), the second measure of uncertainty  $A_t$ , or ‘aggregate error’ uncertainty, is the square of the weighted cross-sectional mean of the forecast errors made across firms in each quarter:

$$A_t = \bar{\epsilon}_{it+1}^2,$$

where  $\bar{\epsilon}_{it} = \frac{1}{W_t} \sum_{i=1}^N w_{it} \epsilon_{it}$ .

Aggregate error uncertainty increases if more firms make similar and larger forecast errors. Thus, if more firms make the same forecast errors, aggregate error uncertainty will increase. If the same proportion of firms make positive and negative forecast errors, it implies zero aggregate error uncertainty. This is akin to the measure based on the mean of the absolute forecast errors proposed in Bachmann, Elstner and Sims (2013).

The third measure of uncertainty  $I_t$ , or ‘idiosyncratic error’ uncertainty, is the weighted cross-sectional standard deviation of the forecast errors in each quarter:

$$I_t = \frac{1}{W_{t+1}} \sum_{i=1}^N w_{it+1} (\epsilon_{it+1} - \bar{\epsilon}_{t+1})^2,$$

## Business Sentiment

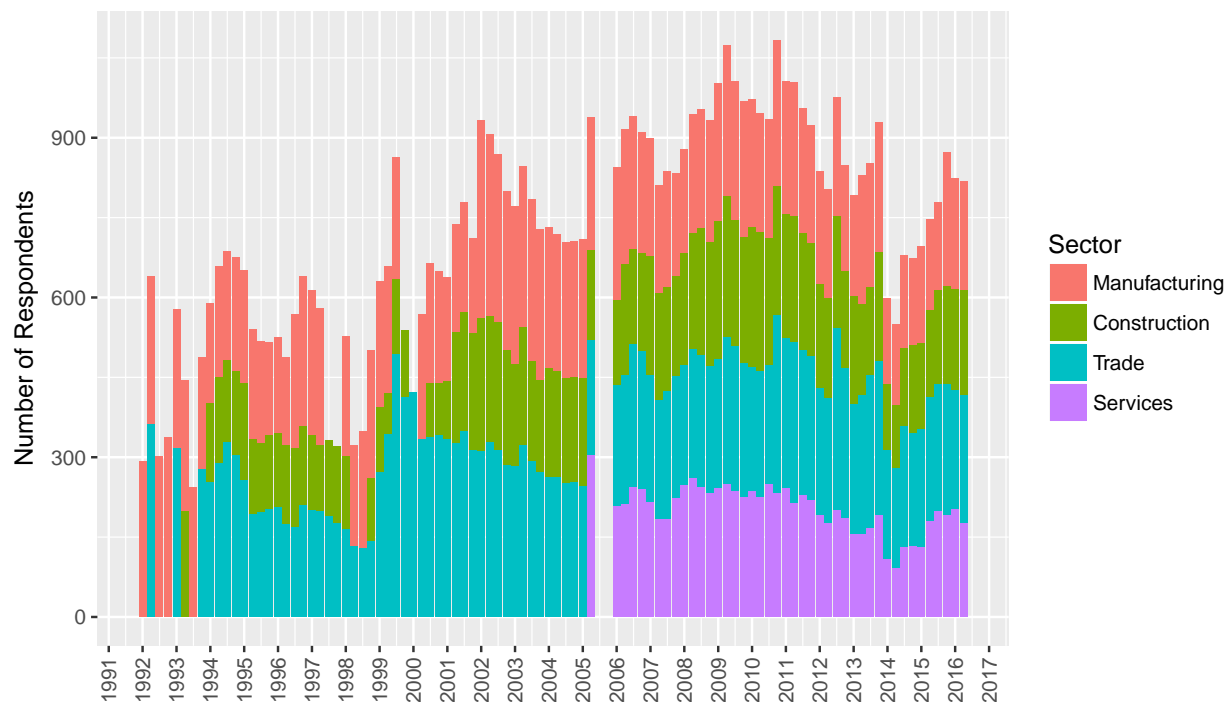


Figure 3.3: Sample sizes of the forecast errors, by sector (1992Q1-2016Q3)

where the variables are defined in the same way as above.

This proxy measures how individual firms depart from the overall mean forecast error. Idiosyncratic error uncertainty increases if firms make more dispersed forecast errors. If all firms make the same forecast error, it implies zero idiosyncratic error uncertainty. This is the measure of uncertainty proposed in Bachmann, Elstner and Sims (2013).

Although these measures are based on the realised forecast errors in the next quarter  $t + 1$ , they depend on the knowledge and level of uncertainty in the current quarter  $t$ . Thus, the mean and standard deviation of realised forecast errors at time  $t + 1$  constitutes uncertainty in  $t$  (Bachmann, Elstner and Sims, 2013).

The composite uncertainty indicators for each sector are then calculated as the average of the same set of survey questions reported in Table 3.2. This should reduce their likelihood of producing ‘false positives’, i.e. signalling high uncertainty where there is none, and ‘false negatives’, i.e. failure to detect mounting uncertainty (Girardi and Reuter, 2017). The sectoral indicators are then aggregated with GDP shares as weights, to form the overall uncertainty indicators.

Figure 3.3 illustrates the sample sizes of the forecast errors by sector. Naturally, these sample size are smaller than the full sample because they require firms to respond in two consecutive quarters. Nevertheless, the sample sizes are still relatively large compared to the full sample, with around 725 forecast errors on average per quarter.

Table 3.4: Comparing sample characteristics in terms of firm size

Firm Size Category	Full Sample		Forecast Error Sample	
	Observations	Percentage of sample	Observations	Percentage of sample
1	25,587	21.43%	14,537	20.88%
2	15,288	12.80%	9,079	13.04%
3	18,554	15.54%	10,936	15.71%
4	13,717	11.49%	8,094	11.63%
5	14,676	12.29%	8,748	12.57%
6	9,140	7.65%	5,331	7.66%
7	6,899	5.78%	3,980	5.72%
8	6,894	5.77%	3,739	5.37%
9	8,667	7.26%	5,166	7.42%

Similar to the case with the non-responses in the full sample, it is assumed that the firms that did not form part of this sample have the same distribution as those included. In other words, the firms that did not respond in consecutive periods have a similar distribution to those that did respond to consecutive surveys. This so-called missing-at-random assumption is common in the international literature when dealing with non-responses in business tendency surveys (EC 2006).

Of course, this need not necessarily be the case. Table 3.4 compares the characteristics, in terms of firm size, between the firms in the full sample and those that only form part of the forecast error sample (i.e. firms that responded in consecutive quarters). The characteristics of the two samples are similar in terms of the size of the firms. As reported in the Appendix, the results are similar when calculating the confidence and uncertainty indicators using only the smaller forecast error sample. The other measures, which do not rely on the panel structure, are therefore robust to calculating them for the more ‘stable’ sample.

Thus, there are three distinct proxies for business uncertainty based on the survey data: dispersion  $D_t$ , aggregate error uncertainty  $A_t$  and idiosyncratic error uncertainty  $I_t$ . Business uncertainty can come from a number of sources and may manifest itself in an array of variables (Jurado, Ludvigson and Ng, 2015). Hence, this chapter also investigates two further proxies for uncertainty, namely economic policy uncertainty and financial market uncertainty.

The economic policy uncertainty indicator is the news-based EPU index created by Hlatshwayo and Saxegaard (2016), discussed above. It is constructed by counting the number of articles that contained 3 mentions of words related to policy, economics, and uncertainty within 10 words of ‘South Africa’. The absolute counts were normalised and the index was standardised.

The financial market uncertainty indicator is a combination of implied and realised stock market volatility. The South African Volatility Index (SAVI) is a forecast of equity market risk on the JSE. It is modelled on the VIX, a popular measure for the volatility of the S&P 500, which has been used in a number of studies (e.g. Bloom (2009)). The SAVI is a forward-looking index that provides a daily prediction of market volatility in three months’ time. It is calculated using implied volatilities obtained daily from specific Top 40 options (JSE, 2014). The SAVI is available only from June 2007. Following the literature (e.g. Bloom (2009), Valencia (2017), Bachmann, Elstner and Sims

(2013) and Redl (2015)), an index of realised stock return volatility was calculated as the standard deviation of the daily JSE All Share index for each quarter. The realised volatility for the period before June 2007 is then chained to the SAVI.

This chapter therefore investigates five proxies for uncertainty. None of them is a perfect measure of an elusive and multidimensional phenomenon, but all of them may contribute to our understanding of uncertainty (Bachmann, Elstner and Sims, 2013). Survey-based measures capture the opinions of key agents in the economy and are driven by changes in firm-level uncertainty. Due to their qualitative nature, however, they are poorly equipped to fully capture heightened uncertainty during extreme events (Bachmann, Elstner and Sims, 2013). Moreover, survey responses are potentially biased and there may be a gap between responses and actual behaviour (Baker and Wurgler, 2007). The survey-based methods do not focus specifically on economic policy uncertainty, which captured in the EPU of Hlathwayo and Saxegaard (2016). The SAVI captures broad uncertainty in financial markets, which is the most popular proxy in the literature, but is derived from a specific segment of firms that are publicly traded.

These five imperfect proxies can be combined to form an overall uncertainty indicator for South Africa. The indicators are combined to attempt to incorporate information from different sources of uncertainty (Leduc and Liu, 2016). This is similar to practices in the literature, where uncertainty indicators are constructed from a range of different proxies (e.g. Baker and Wurgler (2007); Baker, Bloom and Davis (2015), Redl (2015) and North-West University (2016)). The idea is to iron out the remaining idiosyncrasies by averaging the indicators to incorporate information from different sources of uncertainty. This should reduce their likelihood of signalling high uncertainty where there is none, or of failing to detect mounting uncertainty (Girardi and Reuter, 2017). Moreover, incorporating uncertainty from different sources may help to detect exceptionally high uncertainty from specific sources, e.g. from policy changes, which are not captured well by the other indicators.

In constructing their uncertainty measure, Baker and Wurgler (2007) and Baker, Bloom and Davis (2015) used a simple average of their proxies, as well as the first principal component of the series. In this chapter, the first principal component of the five standardised uncertainty proxies is used as an overall combined uncertainty measure ('combined'). A number of papers have used principal component analysis (PCA), or the related factor analysis, to reduce the dimensionality of their data (see Stock and Watson (2002) for a seminal contribution, and Gupta and Kabundi (2011), and Bosch and Ruch (2013) for South African applications). PCA is used to reduce the dimensionality of a dataset consisting of a large number of variables, while retaining as much of the variation as possible (Jolliffe, 2002). The transformation is defined in such a way that the first principal component accounts for as much of the variability in the data as possible (see Jolliffe (2002) for a complete derivation of PCA). The results presented below indicate that the combined indicator exhibits a larger correlation with movements in real output growth than any of the separate components. The results are similar for an equal-weighted overall combined uncertainty index.



## Business Sentiment

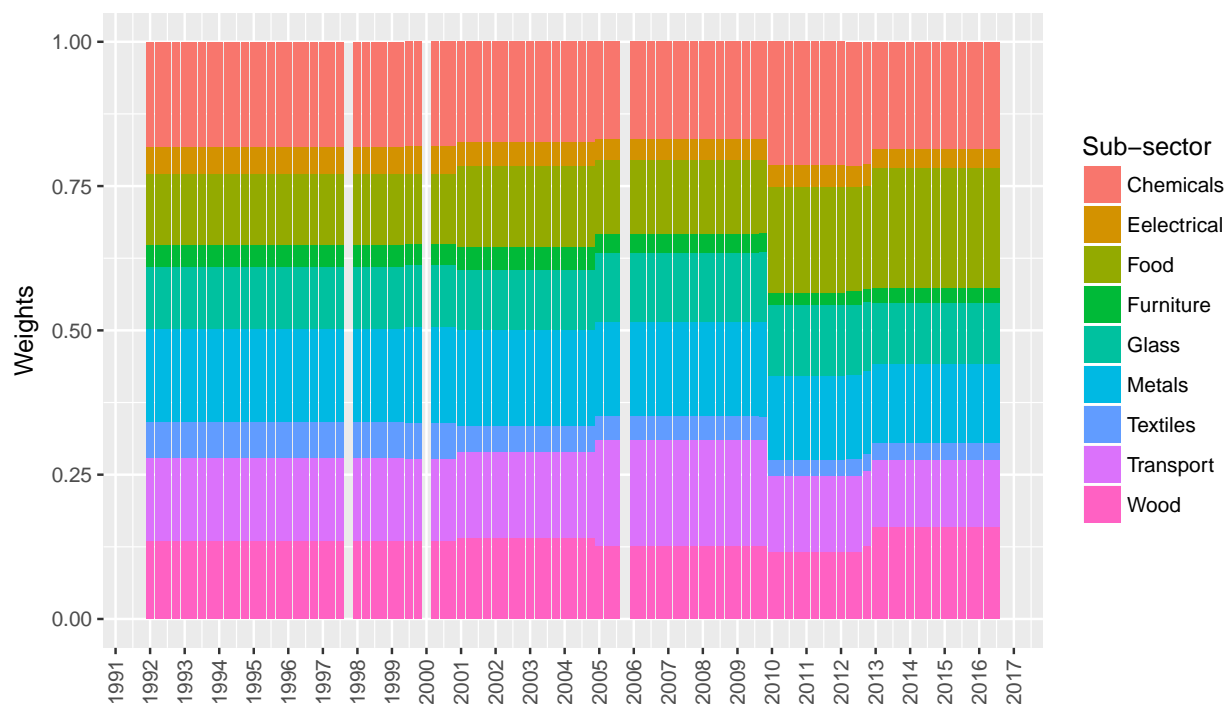


Figure 3.4: Subsector weights applied in the manufacturing sector

### 3.5.3 Weighing the Survey Responses

In this section the weights used to calculate the sentiment indicators are presented. Firm size weights are recorded by the BER for all respondents. The firm size weights are divided into nine categories. In this chapter, the firm size weights are applied to all the responses in all of the subsectors. In contrast, the BER uses exponential firm weights based on the nine categories, except for the building and motor vehicle surveys, where no weights are applied.

Figure 3.4 to Figure 3.7 illustrate the subsector weights for each of the four main sectors: manufacturing, construction, trade and services. The weights for the manufacturing subsectors are updated periodically by the BER, based on the composition of production or sales in each subsector, as calculated by StatsSA. The subsector weights are cleaned versions of those used by the BER in calculating its Manufacturing BCI.

Subsector weights are not recorded by the BER for the construction sector. The BER Building BCI is based on the unweighted responses for contractors only. In this chapter, the relative subsector weights are set equal to the average number of respondents for each subsector over the period. The results are similar when using an equal weighting procedure. The microdata for architects, quantity surveyors and civil engineers are only available from 2001Q2.

The BER also updates the weights for the retail and wholesale subsectors periodically. The weights

## Business Sentiment

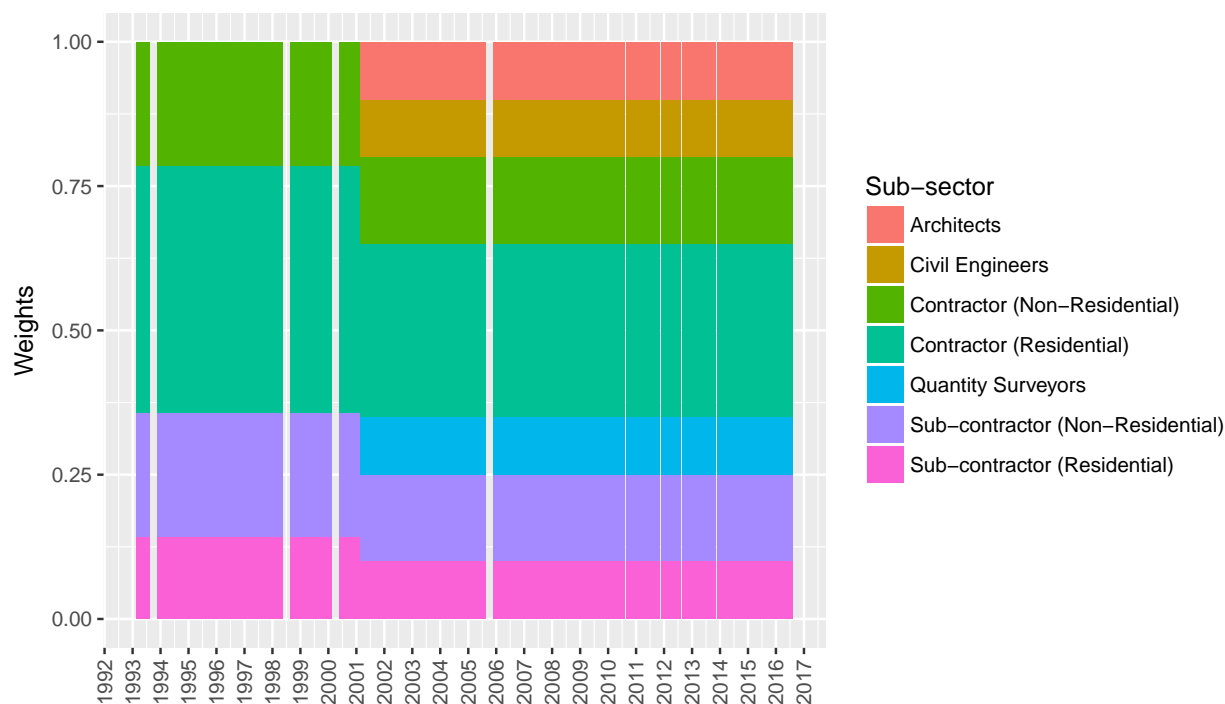


Figure 3.5: Subsector weights applied in the construction sector

for these subsectors in this chapter are the same as those used by the BER in calculating its Retail and Wholesale BCIs. The BER Motor Vehicle BCI does not receive a subsector weighting. The BER assumes an equal weighting for the retail, wholesale and motor vehicle subsectors when calculating its total BCI. In this chapter, the relative weights are set equal to the average number of respondents for each subsector over the sample period. The results are similar when using an equal weighting procedure.

Subsector weights are not recorded by the BER for the services sector and the BER does not publish a Services BCI. In this chapter, the weights are set equal to the average number of respondents for each subsector over the sample period, although the results are similar when using equal weights.

Figure 3.8 illustrates the GDP share weights that are used in aggregating the four sector indicators to calculate the aggregate indicators. The BER BCI, in contrast, is a simple equal weighted average of the sectoral indicators for manufacturing, contractors, retail, wholesale, and motor vehicles.

Naturally, there are other ways to weigh the responses, but experience has shown that the balances are not very sensitive to the choice of weighting procedure (Organisation for Economic Co-operation and Development, 2003). Indeed, in this case the specific weighting procedure turns out to have little impact on the confidence indices. The unweighted versions of the indicators, calculated by stacking all of the available responses from all the surveys (i.e. completely unweighted), are very similar to the weighted versions. The application of the BER weights also provides similar results.

Business Sentiment

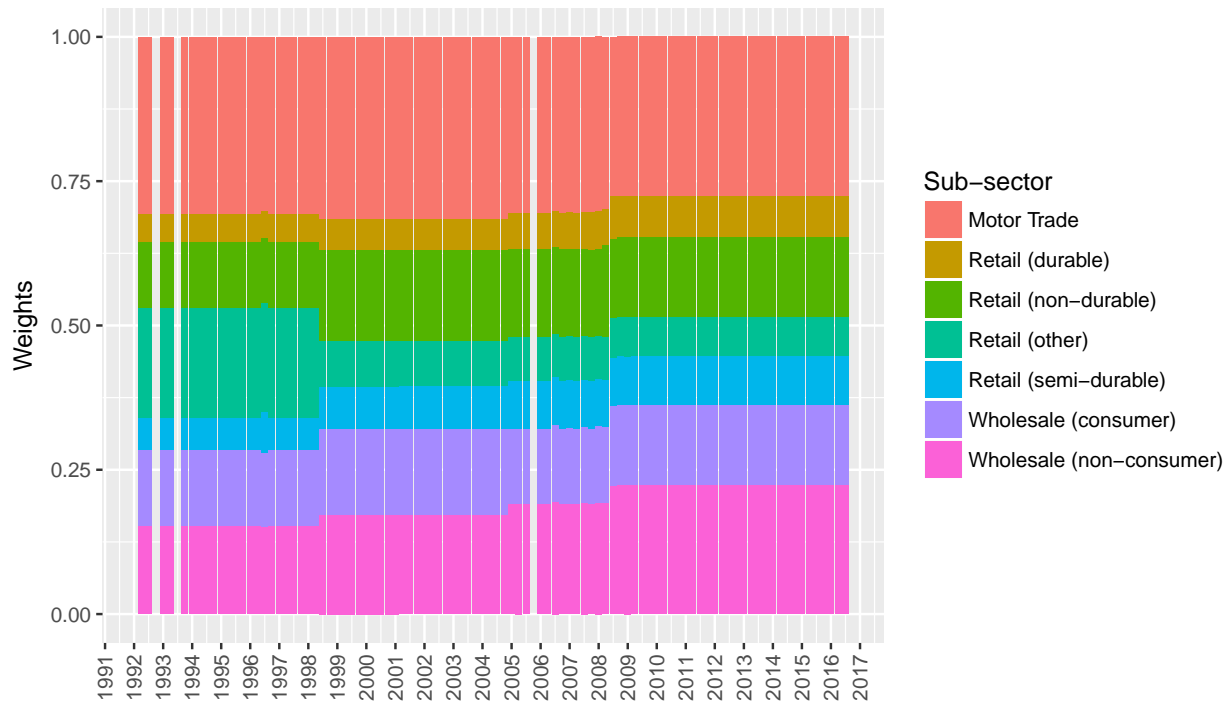


Figure 3.6: Subsector weights applied in the trade sector

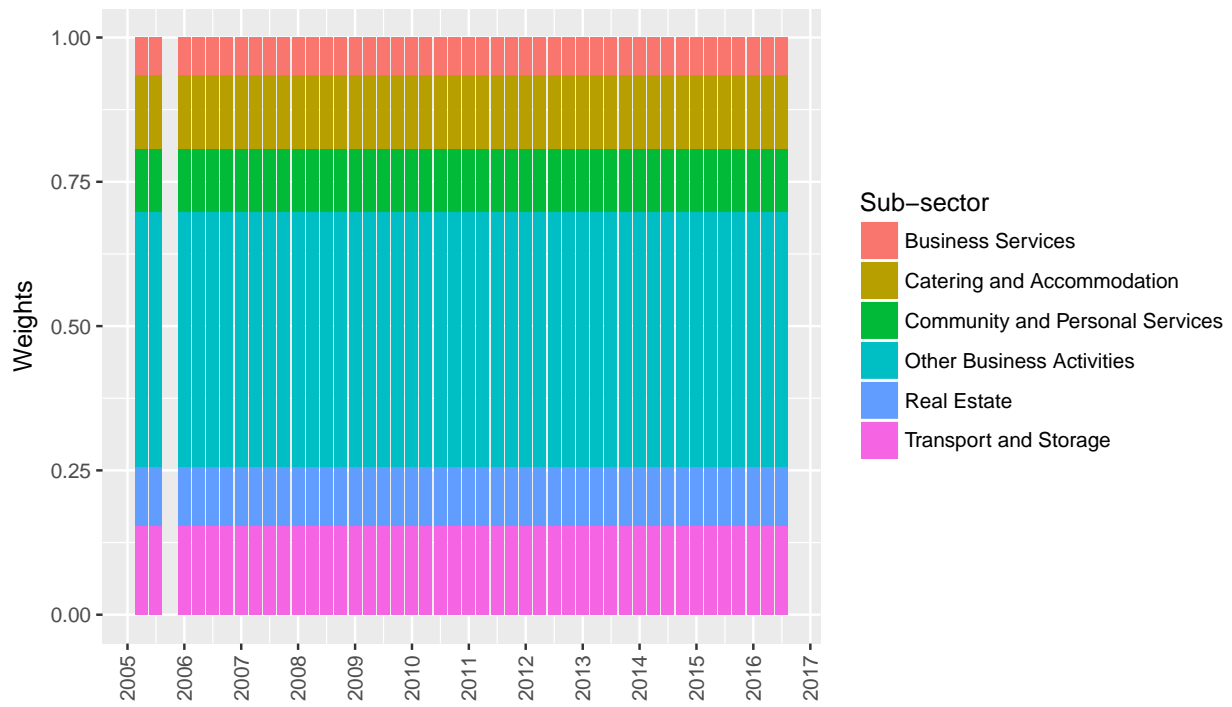


Figure 3.7: Subsector weights applied in the services sector

## Business Sentiment

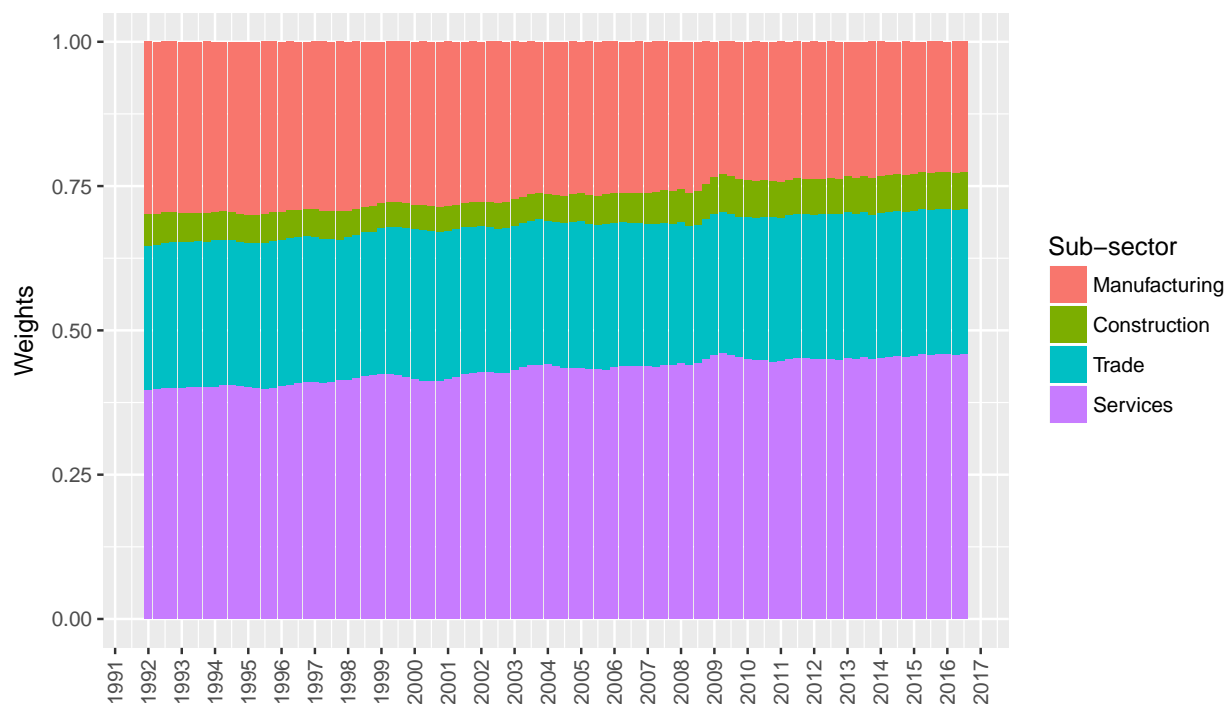


Figure 3.8: GDP share weights applied to the main sectors

The specific weighting procedure adopted therefore does not significantly alter the results.

This confirms the findings by Kershoff (2015), who tested alternative weighting procedures: a different allocation of firm size weights; the introduction of dynamic individual weights (post-stratification), to provide for changes in response patterns (to handle non-responses); the application of a two-step weighting procedure; the inclusion of latecomers to increase the number of responses; and the use of different sector size weights for export variables. The findings showed that the balance statistics were not sensitive to the use of alternative weighting procedures.

### 3.6 Results: Confidence

This section presents the composite sectoral and aggregate business confidence indicators for South Africa. Simple linear interpolation is used for the few missing quarters. The validity of the indicators is assessed by comparing them with events that were thought to coincide with large changes in confidence, as well as with existing measures of confidence for South Africa. The indicators are then evaluated according to their comovement with real GDP growth (i.e. their tracking record), to assess whether they improve on the existing indicators of confidence.

### 3.6.1 Confidence Indicators

Figure 3.9 illustrates the weighted sectoral confidence indicators for current conditions and expected conditions. The indicators appear to capture cyclical movements in the sectors. In general, they display an increase in the early 1990s until just after the first Democratic Elections in 1994Q2. They show a sustained decrease from 1995 into the recession of 1997-1998, associated with the East Asian and Russian crises. After troughs around the start of 1999, the indicators increase up to the global financial crisis at the end of 2007. During this extended upswing phase, the manufacturing and trade sectors reflect the two ambiguous periods in 2001 and 2003, when contractions in the SARB leading and coincident indicators obliged an evaluation of possible reference turning points (Venter, 2005). The construction sector exhibited a particularly strong and sustained increase in confidence during this upswing phase, possibly due to the construction projects associated with hosting the FIFA World Cup in 2010.

The global financial crisis was followed by a large decline in the indicators for all of the sectors, which continued into the subsequent Great Recession. There was a relatively quick recovery in confidence in the manufacturing and trade sectors. Confidence in the construction sector showed a more gradual recovery, especially in confidence on current conditions. In the services sector, confidence on current conditions showed a slight recovery and then continued to decline, whereas confidence on expected conditions was quite erratic. The indicators for the other sectors exhibit a gradual decrease from around 2012, continuing into the downswing phase at the end of the sample period.

Figure 3.10 illustrates the weighted aggregate confidence indicators on current and expected conditions. The shaded areas denote the recessionary periods according to the official turning points of the SARB. The indices follow a similar cyclical trend over the period and are very highly correlated, as is reported in Table 3.5, below.

The indicators appear to match the different phases of the business cycle relatively well. Turning points generally coincide with the official turning points, as is discussed in more detail below. The indicators exhibit an increase following the recession of the early 1990s, with peaks around 1995. There is a prolonged decrease into the recession of 1997-1998, and a strong recovery just before the official trough in 1999. Both ambiguous periods are reflected in moderate decreases in the indicators in 2001 and 2003. Both indicators exhibit a significant decrease following the global financial crisis in 2007, and a relatively mild recovery just before the official trough in 2009. The indicators are relatively flat during the previous upswing phase (2010-2013) and decrease gradually during the downswing phase at the end of the sample period.

As reported in the Appendix, the indicators are robust to calculation based on more stable samples of firms. The Appendix also illustrates confidence intervals around the aggregate and sectoral confidence indices. The confidence intervals show that the distribution of the sample means are relatively narrow, because of the large number observation in each quarter. As a consequence, the

Business Sentiment



Figure 3.9: Weighted sectoral confidence indicators on current and expected conditions

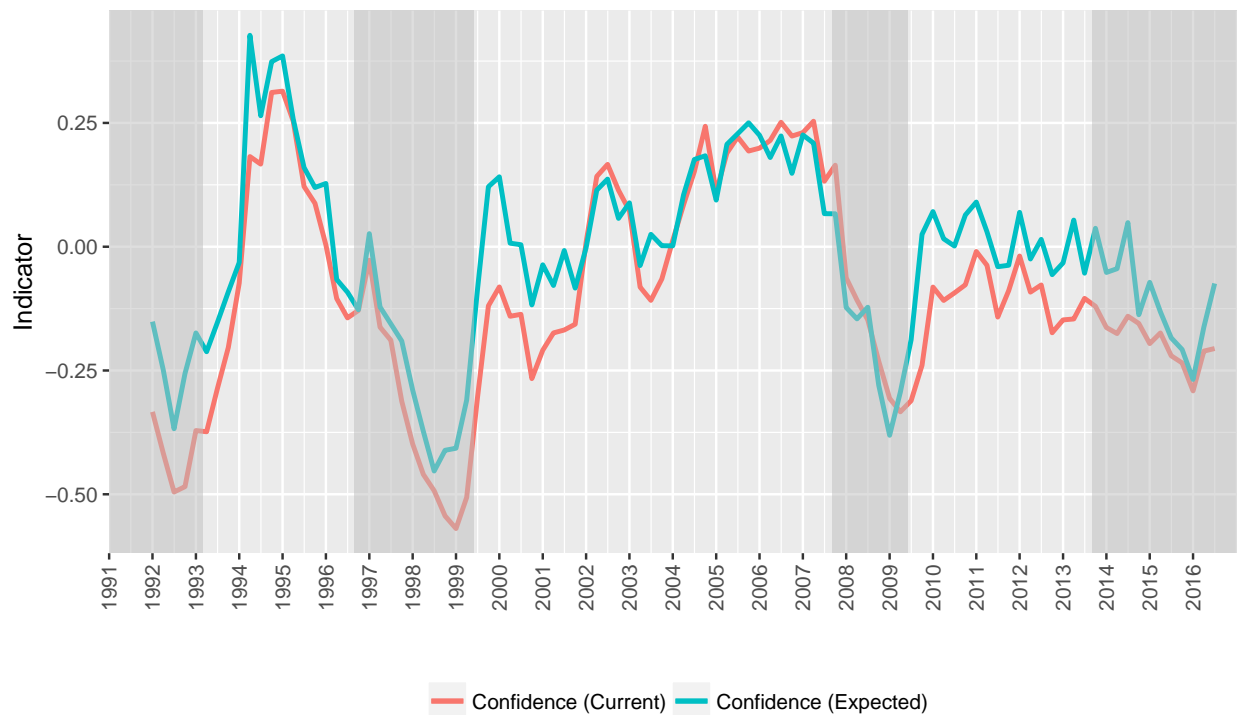


Figure 3.10: Weighted confidence indicators on current and expected conditions

changes in the indices seem ‘real’ rather than statistical idiosyncrasies. The survey-based confidence indicators therefore appear to be plausible and potentially useful indicators of business confidence in South Africa.

### 3.6.2 Validity Tests and Evaluation

This section provides a comparison of the characteristics of the new confidence indicators to the two existing South African business confidence indices, the BER BCI and the SACCI BCI. Correlations are used to analyse the tracking record of the indicators with respect to real GDP growth. The relationships between turning points is reported, to assess their usefulness as leading indicators of the business cycle.

#### 3.6.2.1 Correlations between confidence indicators and real GDP growth

Figure 3.11 compares the confidence indicators on current and expected conditions with the BER BCI,<sup>27</sup> the SACCI BCI, as well as real GDP growth. Real GDP growth is calculated as annual quarter-on-quarter growth rates, e.g. 2015Q1 over 2014Q1, which corresponds to the reference period in the BER surveys. The official recessionary periods are shaded, and the indicators are standardised for plotting. The indicators appear to be strongly pro-cyclical, and follow real GDP growth closely.

Table 3.5 reports the contemporaneous correlations of the indicators and real GDP growth. SACCI BCI growth rates are used to remove unit roots and are calculated as annual quarter-on-quarter growth. All the indicators exhibit a significant positive correlation with one another and with real GDP growth. The current conditions confidence indicator has a marginally higher contemporaneous correlation with real GDP growth than the BER BCI or SACCI BCI, which are also based on current conditions. The correlation between the expected conditions confidence indicator and contemporaneous real GDP growth is also relatively high.

<sup>27</sup>The new confidence indicators differ from the BER BCI in a number of ways. First, the BER BCI is based on a single question related to satisfaction with general business conditions. The new confidence indicators are composite indicators that combine the responses to five types of variables, namely business conditions, activity (production or sales), orders placed, employment, and profitability. Second, the BER BCI excludes the services sector altogether, and excludes the other construction subsectors apart from building contractors (i.e. sub-contractors, architects, quantity surveyors, and civil engineers). The new confidence indicators include all of the available survey responses. Third, the BER BCI weighs each response with a factor, which is calculated as the product of a firm size weight and a subsector size weight,  $w_{it} = f_{it}s_{jt}$ , without dividing by the total firm weight for the subsector  $F_{jt}$ . In contrast, the new confidence indicators use weights,  $w_{it} = f_{it}s_{jt}/F_{jt}$ , which are equivalent to an explicit two-step weighting procedure, whereby weighted means are calculated for each subsector separately, and then aggregated with the subsector weightings. Fourth, the BER BCI uses exponential firm weights, which makes the series particularly volatile. The new confidence indicators use simple linear weights based on the size categories. Fifth, the BER BCI does not weigh the responses from the building contractor and motor vehicle surveys. The new confidence indicators weigh all of the sectors in the same way. Sixth, the BER BCI assumes that the five sectoral indices (manufacturing, building contractors, retailers, wholesalers and new vehicle dealers) have an equal weighting, which increases the importance of motor vehicle dealers substantially. The new confidence indicators combine the sectoral series with weights based on GDP shares to create the aggregate confidence indicators.

## Business Sentiment

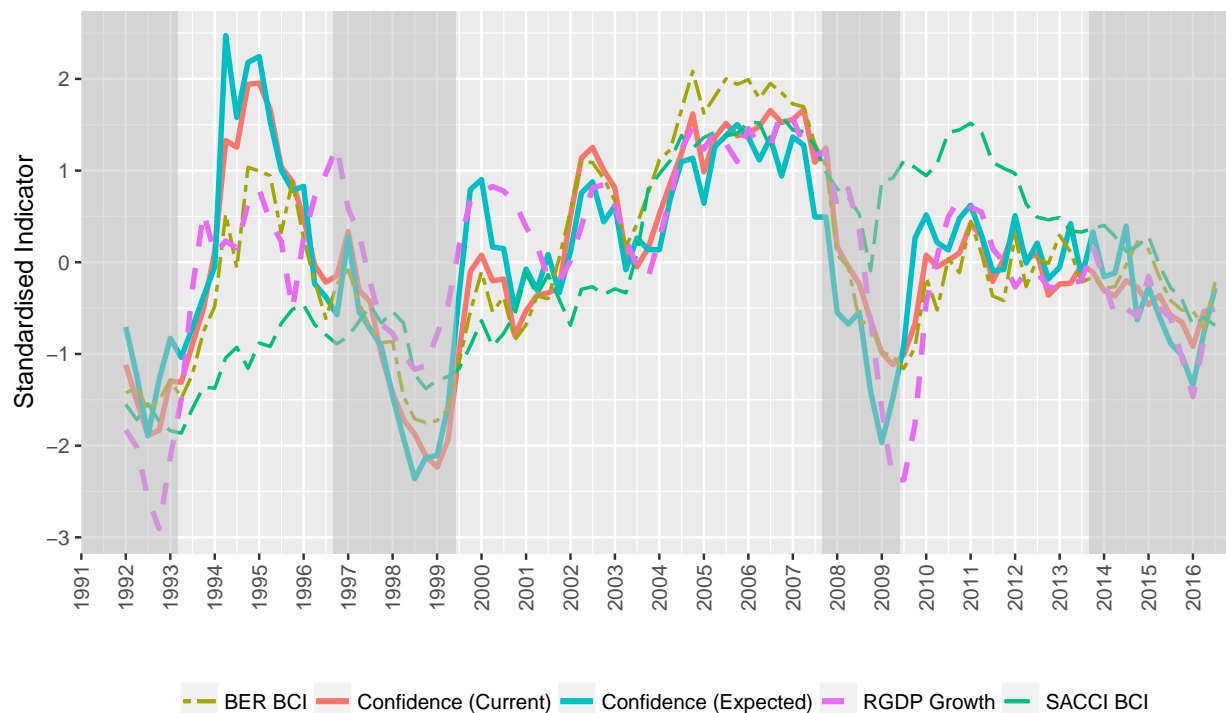


Figure 3.11: Confidence indicators compared to real GDP growth

Table 3.5: Correlations between confidence indicators and real GDP growth

	Confidence (Current)	Confidence (Expected)	BER BCI	SACCI Growth
Confidence (Expected)	0.92***			
BER BCI	0.93***	0.82***		
SACCI BCI Growth	0.35***	0.48***	0.30***	
Real GDP Growth	0.78***	0.70***	0.75***	0.24**

Cross-correlations can be used to illustrate the dynamic relationships between the indicators and real GDP growth. Figure 3.12 illustrates the cross-correlograms for the indicators and real GDP growth. All three survey-based measures exhibit relatively high correlations with contemporaneous and lagged GDP growth. The highest correlation coefficient between the indicators of current conditions and real GDP growth occur contemporaneously. The confidence measure of expected conditions leads GDP growth, and exhibits the highest correlation coefficient when lagged by one period. The results imply that the indicators are all potentially useful leading or quasi-leading indicators of real activity.

Figure 3.13 compares the sectoral current and expected conditions confidence indicators with the BER sectoral indicators, as well as with the corresponding real sectoral GDP growth rates. The indicators capture cyclical movements in real output over the period. Table 3.6 reports the contemporaneous correlations of the sectoral indicators and their respective sectoral real GDP growth rates. All the indicators are highly positively correlated with real sectoral GDP growth rates. For the most part, the current conditions confidence indicators exhibit the highest correlation



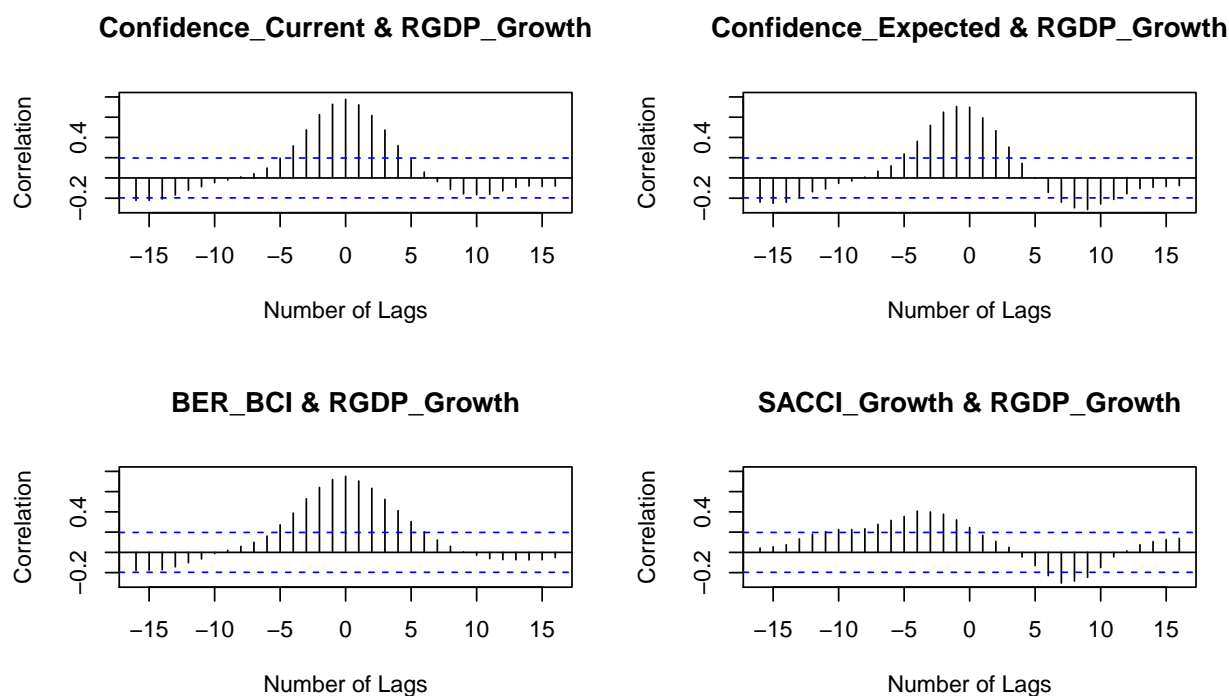


Figure 3.12: Cross-correlograms of the confidence indicators and real GDP growth

with the reference series. In this sense, they are an improvement on existing confidence indicators. The exception is the construction sector, where the BER Building BCI has the highest correlation. This is peculiar, as the BER Building BCI includes only building contractors.

Table 3.6: Correlations between sectoral confidence and real sectoral GDP growth

	Manufacturing			Construction		
	Confidence (Cur)	Confidence (Exp)	BER BCI	Confidence (Cur)	Confidence (Exp)	BER BCI
Confidence (Exp)	0.94***			0.89***		
BER BCI	0.92***	0.85***		0.94***	0.75***	
RGDP Growth	0.68***	0.68***	0.61***	0.74***	0.56***	0.76***
	Trade			Services		
	Confidence (Cur)	Confidence (Exp)	BER BCI	Confidence (Cur)	Confidence (Exp)	BER BCI
Confidence (Exp)	0.87***			0.76***		
BER BCI	0.90***	0.72***				
RGDP Growth	0.61***	0.59***	0.56***	0.76***	0.57***	

Figure 3.14 illustrates the cross-correlograms for the manufacturing indicators and real GDP growth in the manufacturing sector. The results are similar to the aggregate results. Again, all three survey-based measures exhibit relatively high correlations with contemporaneous and lagged GDP growth. The expected conditions confidence measure leads real GDP growth. The cross-correlograms for the other sectors are very similar (not shown), except for the services sector, where the expected conditions confidence measure has an even longer leading relationship with real GDP growth.

Business Sentiment

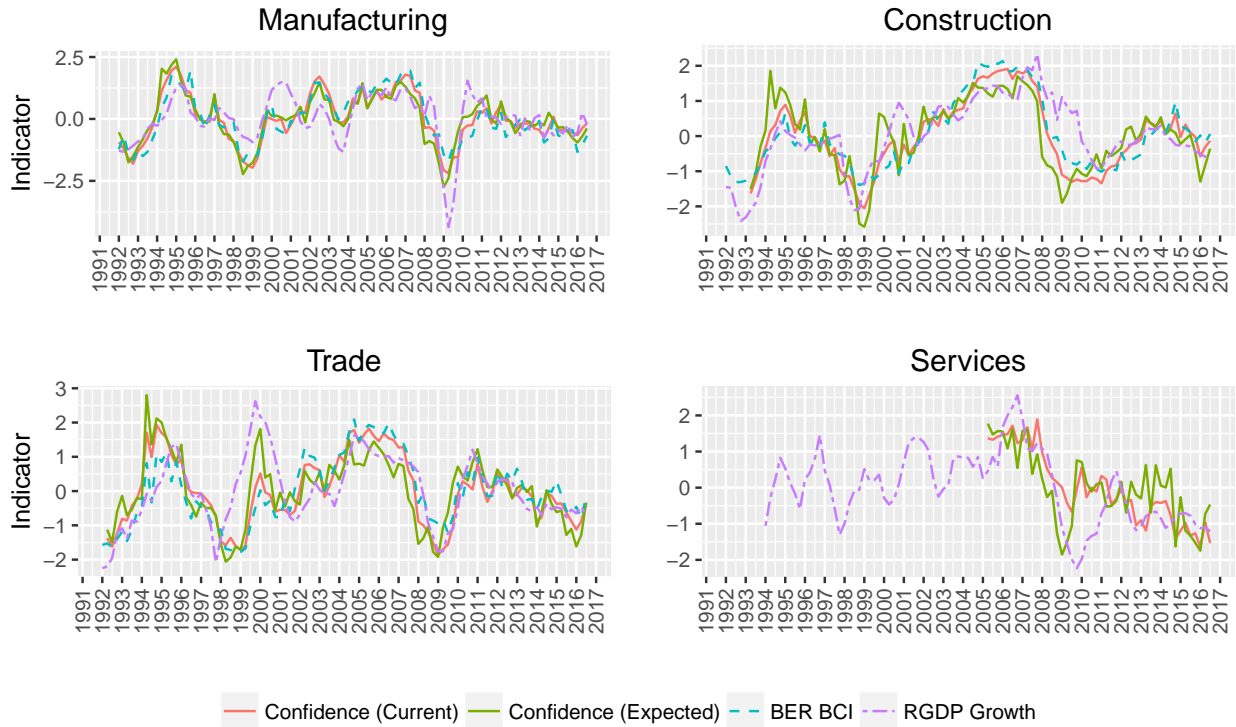


Figure 3.13: Sectoral confidence indicators compared to real sectoral GDP growth

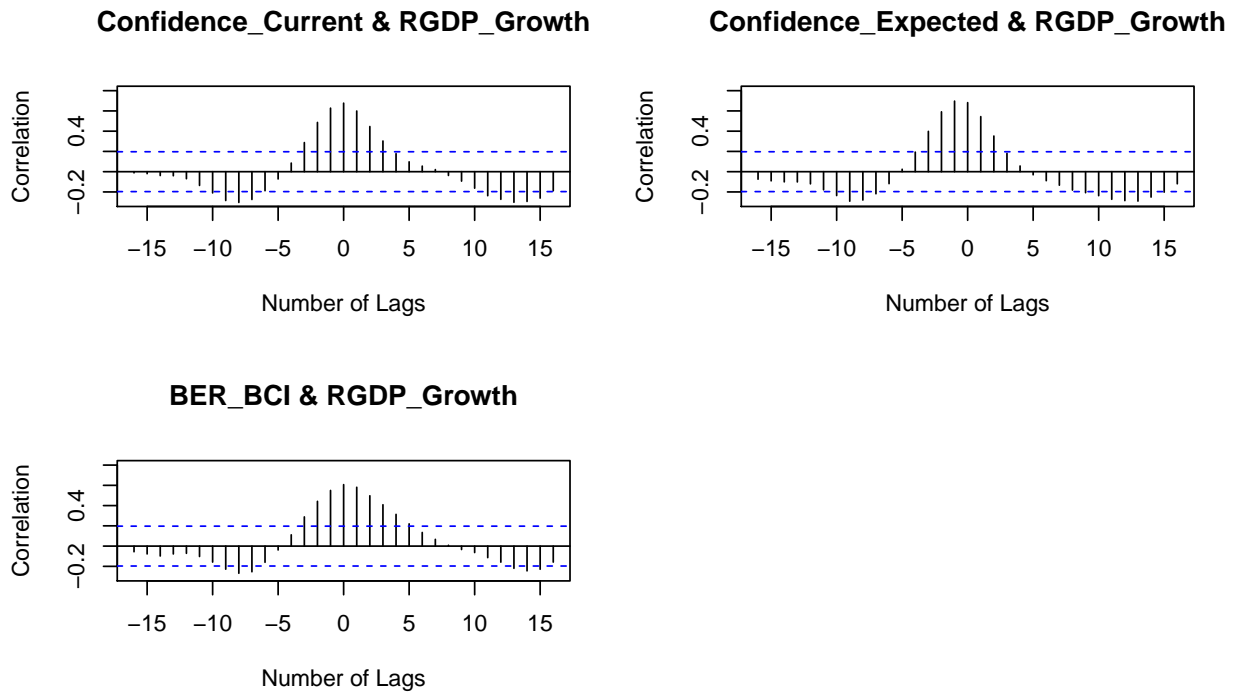


Figure 3.14: Cross-correlograms of the indicators and real GDP growth in the manufacturing sector

### 3.6.2.2 Turning points

An accurate leading indicator should show general conformity to economic activity (i.e. a high correlation), as well as a consistent matching of turning points with the reference cycle. Although there are too few cycles over the sample period to analyse cyclical turning points in full detail, it is still of interest to assess whether the indicators behave in a systematic way around cyclical turning points. In other words, do they systematically lead, coincide with, or lag the peaks and troughs of the business cycle.

The turning points in the indicators are determined where the indicators breach the threshold of zero, i.e. they indicate an upswing when they are positive and a recession when they are negative. The two new indicators and the BER BCI are standardised, as their means are below zero over the sample period, and the SACCI BCI enters in growth rates. Censoring rules are used to ensure that phases and cycles have a minimum duration, similar to those used in the so-called Bry-Boschan method (Bry and Boschan, 1971). Following the suggestion of Harding and Pagan (2002), who developed a variant of this method for dealing with quarterly data (the BBQ method), a censoring rule based on a minimum of two quarters for each phase and five quarters for a full cycle is applied.

The resulting phases are illustrated in Figure 3.15, with the recessionary periods shaded. The top panel of each graph illustrates the turning points of the confidence indices, while the bottom panel of each graph shows the official SARB reference turning points. The sample period includes three upswing phases and four downswing phases. In addition, in 2001 and 2003 the SARB indicators pointed to possible reference turning points. Although the SARB dating committee decided at the time that neither of these periods qualified, subsequent data revisions have shown that in hindsight there could have been official peaks, especially in 2003, if the dating procedure had been followed mechanically (Venter, 2005).

The algorithm identifies four recessionary periods in the current conditions confidence indicator and five in the expected conditions confidence indicator. These correspond to the official downswing phases, with the additional downswing phase during the semi-recession in 2001. The turning points in the BER and SACCI BCIs are similar to those for the new confidence indicators. There is some ambiguity towards the latter part of the sample period, as the expected conditions confidence indicator and the BER BCI hover around the zero threshold. On the whole the phases identified with the indicators are longer in duration than the official phases. The indicators mostly exhibit peaks before the official peak dates, by as many as 10 quarters before the official peak. The indicators exhibit troughs concurrent with or after the three official trough dates. The indicators therefore seem to reflect the official business cycle turning points relatively well.

The comovement between these cycle phases can be measured with the concordance statistic suggested by Harding and Pagan (2002). The concordance statistic measures the comovement of two series, by considering the proportion of time the two series are simultaneously in the same phase. This entails testing whether  $I = Pr(S_{xt} = S_{yt})$  is close to 1, where  $S_{xt} = 1$  identifies an expansion

## Business Sentiment

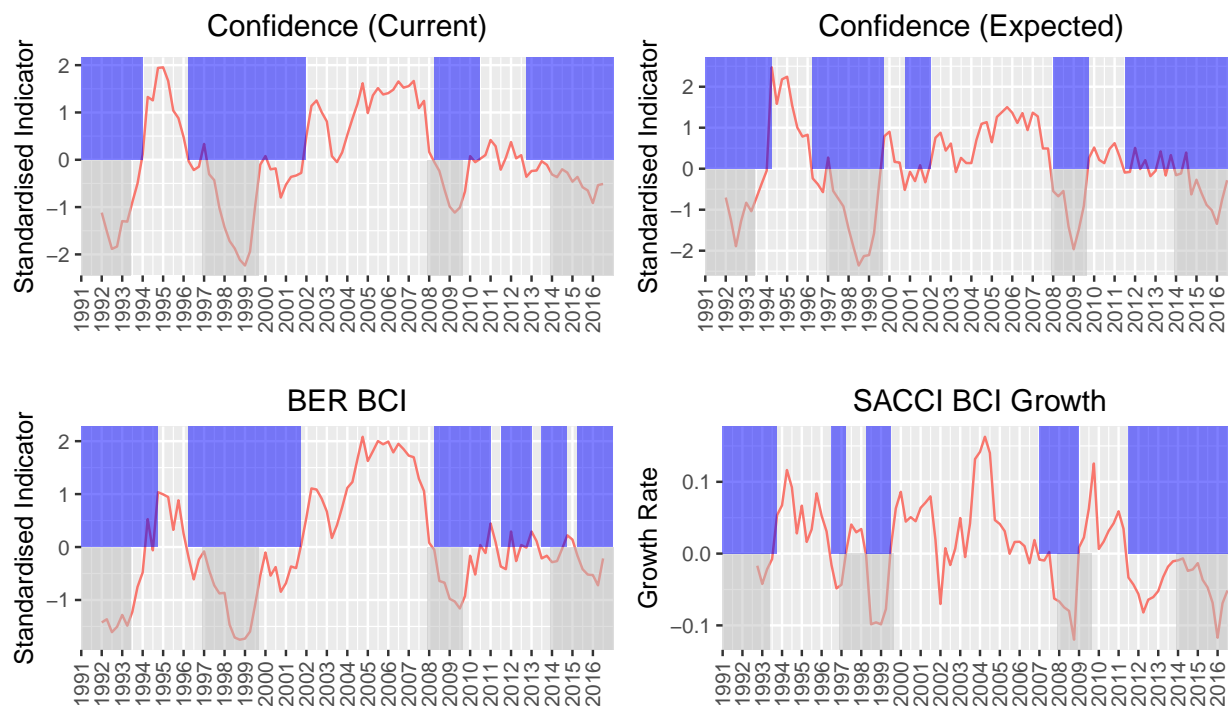


Figure 3.15: Confidence indicator turning points compared to the official SARB turning points

in indicator  $x_t$ , and  $S_{yt} = 1$  identifies a business cycle upswing phase at time  $t$ . The statistic is calculated as follows:  $I = 1/T[\sum_{t=1}^T S_{xt}S_{yt} + \sum_{t=1}^T (1 - S_{xt})(1 - S_{yt})]$ . Following Harding and Pagan (2006), statistical significance is calculated with heteroskedasticity and autocorrelation consistent standard errors.

Table 3.7 reports the concordance statistics for the phases of the indicator variables, compared with the official SARB reference turning points. The indicators all exhibit significant concordance with the official SARB business cycle. The three survey-based indicators have the highest concordance statistic with the official SARB cycle when they are lagged by one or two quarters, but the contemporaneous concordance statistics are all significant.

The indicators therefore seem to reflect the official business cycle turning points relatively well, and provided advance warning especially of the official peaks. The results suggest that the confidence indicators are potentially useful leading indicators of the business cycle. However, the false positives and ambiguous periods imply that the indicators should be used in conjunction with other series when identifying turning points, as in Laubscher (2014). As more microeconomic data from the BER's business tendency surveys become available, the analysis could be expanded by analysing the cyclical properties of the indicators in terms of duration, amplitude and steepness.

Table 3.7: Concordance statistics with the SARB business cycle

	Confidence (Current)	Confidence (Expected)	BER BCI	SACCI BCI Growth
lead=3	0.60	0.62*	0.47	0.72**
lead=2	0.65*	0.67**	0.54	0.75***
lead=1	0.68**	0.70***	0.59*	0.76***
lead/lag=0	0.71***	0.73***	0.62**	0.75***
lag=1	0.72***	0.74***	0.63***	0.70***
lag=2	0.73***	0.69***	0.64***	0.65***
lag=3	0.72***	0.64***	0.63***	0.6**

### 3.7 Results: Uncertainty

This section presents the composite sectoral and aggregate indicators of uncertainty for South Africa. Simple linear interpolation is used for the few missing quarters and all the indicators are standardised. The validity of the indicators is evaluated by assessing whether large changes coincided with events that may have induced uncertainty, as well as by comparing them with existing measures for South Africa. The indicators are then evaluated in terms of their comovement with real GDP growth, to assess whether they improve on the existing indicators of uncertainty. An overall uncertainty indicator is then created, which tries to combine the information in all of the indicators.

#### 3.7.1 Uncertainty Indicators

Figure 3.16 illustrates the weighted sectoral indicators of uncertainty based on dispersion. These indicators are quite volatile by construction (Girardi and Reuter, 2017). The indicators of dispersion for the manufacturing, construction and trade sectors spike during the 1997-1998 recession, associated with the East Asian and Russian crises. In those three sectors the indicators also increased in the recessionary period following the global financial crisis. The dispersion indicator for the manufacturing sector also exhibits a spike at the beginning of the period during the Democratic transition. The dispersion indicator for the services sector is particularly volatile and does not exhibit the large increase during the Great Recession which is present in the indicators for the other sectors.

Figure 3.17 illustrates the weighted sectoral indicators of aggregate error uncertainty. The indicators in the manufacturing, construction and trade sectors spike at similar times as the corresponding indicators of dispersion. There are spikes during the Democratic transition, the 1997-1998 recession, the two semi-recessions (in 2001 and 2003), and the Great Recession. In addition, all four indicators exhibit spikes during the European debt crisis in 2011, and again in 2014, at the start of the downswing phase at the end of the sample period. On the whole, however, the weighted sectoral indicators of dispersion seem to identify periods of uncertainty more accurately than the indicators of aggregate forecast error. This is also reflected in the higher correlations with real GDP growth, presented in Table 3.10 below.

Business Sentiment

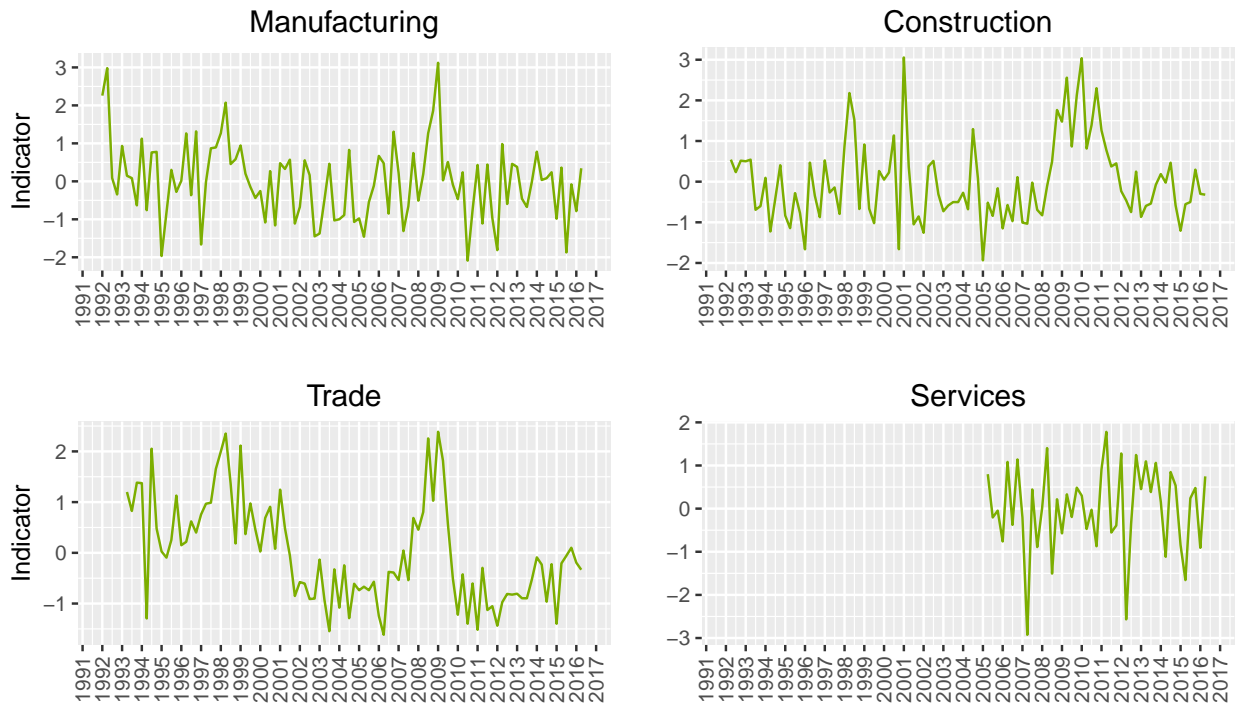


Figure 3.16: Weighted sectoral indicators of dispersion

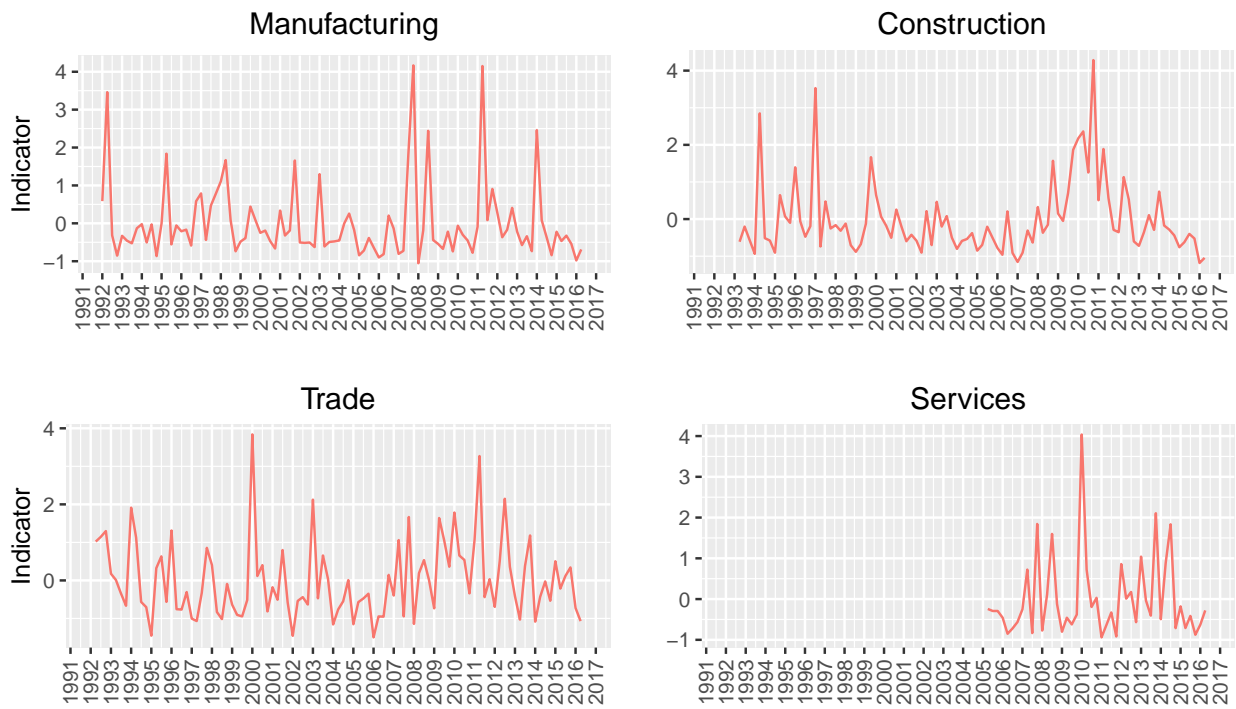


Figure 3.17: Weighted sectoral indicators of aggregate forecast error uncertainty

## Business Sentiment

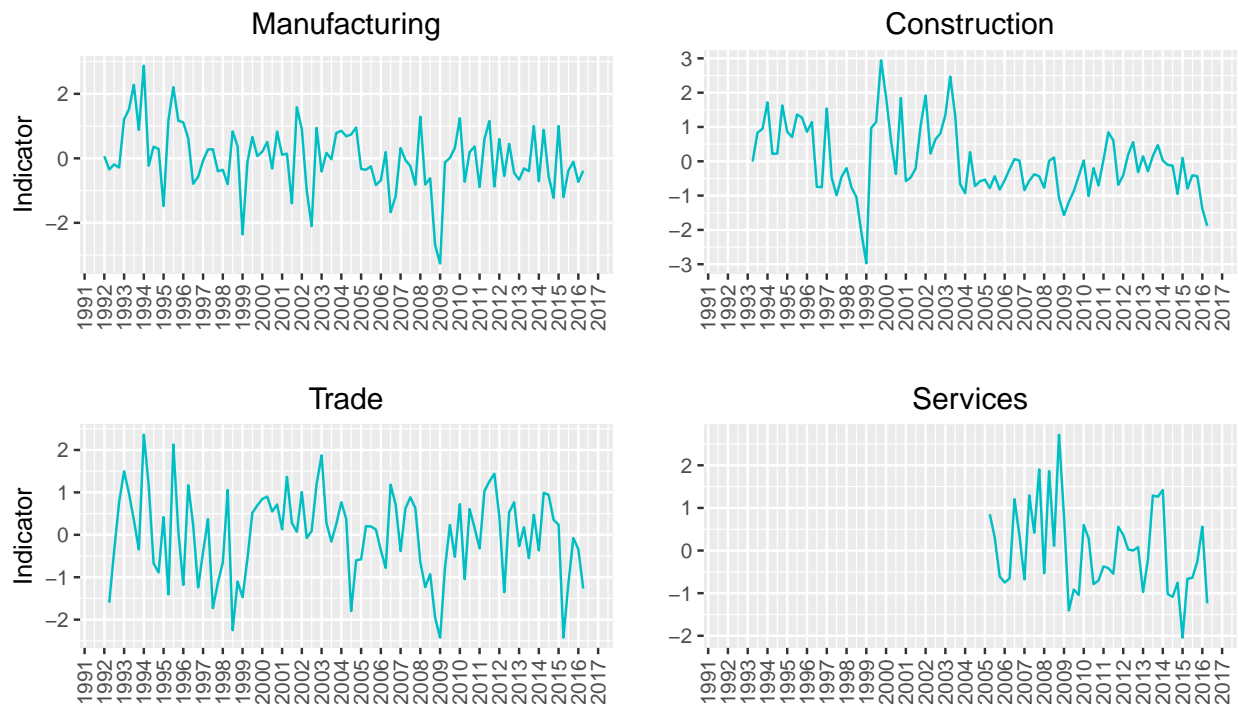


Figure 3.18: Weighted sectoral indicators of idiosyncratic forecast error uncertainty

Figure 3.18 illustrates the weighted sectoral indicators of idiosyncratic error uncertainty. These indicators do not always point to the same periods of heightened uncertainty than the ones highlighted above. The indicator of idiosyncratic error uncertainty in the manufacturing sector exhibits spikes in 1994, with the first Democratic election, and again in 1996, with the crisis in the foreign exchange market and the economic policy uncertainty before the adoption of the Growth, Employment and Redistribution (GEAR) framework (Koma, 2013). This indicator decreases during the Great Recession. The indicator in the construction sector exhibits a marked decrease during the 1997-1998 recession, which is followed by spikes in 2000 and during the two semi-recessions. It is relatively flat for the rest of the period. The indicator for the trade sector is relatively volatile at the beginning of the period, and exhibits substantial decreases during all four recessionary periods. The indicator for the services sector is relatively high and volatile during the Great Recession, and exhibits a spike at the start of the final downswing phase in 2014.

In some cases, therefore, the individual indicators for each sector do not point to the same periods of heightened uncertainty. Table 3.10, below, reports that the indicators are weakly correlated only in a few cases. The lack of correlation is due to the different calculation methods used to construct the proxies. The dispersion indicator measures the disagreement in expectations, expressed as a share of the natural dispersion. The aggregate error and idiosyncratic error uncertainty indicators measure respectively the mean and standard deviation of firm forecast errors. Aggregate error uncertainty will increase if more firms make similar and larger errors, while idiosyncratic error uncertainty will

decrease if more firms make similar errors. Consequently, the indicators do not generally point to the same periods of heightened uncertainty.

This feature is also present for the aggregate indicators. Figure 3.19 illustrates the three weighted uncertainty indicators at the aggregate level, with the recessionary periods shaded. As with the sectoral proxies, the indicators are relatively volatile, and are weakly correlated only in a few cases, as Table 3.9 below reports.

The dispersion indicator seems to follow an anti-cyclical pattern, with spikes during the recessionary periods. In particular, it points to periods of heightened uncertainty during the recessions of the early 1990s, the late 1990s, and the late 2000s. The aggregate error uncertainty indicator also seems to be broadly anti-cyclical. It exhibits large spikes in all four recessionary periods and during the two semi-recessions of the early 2000s. It also exhibits two large spikes in 2010 and 2011, during the period associated with the European debt crisis.

The idiosyncratic error indicator tends to decrease as the economy enters a recessionary period and then to increase towards the end of the recession and into the start of the recovery phase. This is probably because the majority of firms expected poorer general conditions with more certainty, as the recession took hold. Uncertainty about the future then increased around the trough, as expectations became more dispersed. The idiosyncratic error indicator also exhibits the two large spikes in 1994 and 1996, with the first Democratic election and the policy uncertainty before the adoption of the GEAR policy framework. More formal tests of validity are undertaken in the following section.

### **3.7.2 Validity Tests and Evaluation**

This section provides a comparison of the survey-based uncertainty indicators and the two alternative indicators of uncertainty in South Africa, the EPU and the SAVI. The information from all of the indicators is combined to form an overall combined uncertainty indicator, and their correlations with real GDP growth are subsequently evaluated.

#### **3.7.2.1 Correlations between uncertainty indicators and real GDP growth**

Figure 3.20 illustrates the two alternative indicators, as well as the combined overall uncertainty indicator, which was calculated as the first principal component of the five standardised uncertainty indicators. Table 3.8 reports the factor loadings in calculating the combined uncertainty indicator. The dispersion measure and the EPU receive the highest weights in the calculation, while the idiosyncratic error indicator does not enter into the calculation. The results are similar when calculating the combined uncertainty measure as an equal weighted average, except that the idiosyncratic error indicator receives a higher weighting.



## Business Sentiment

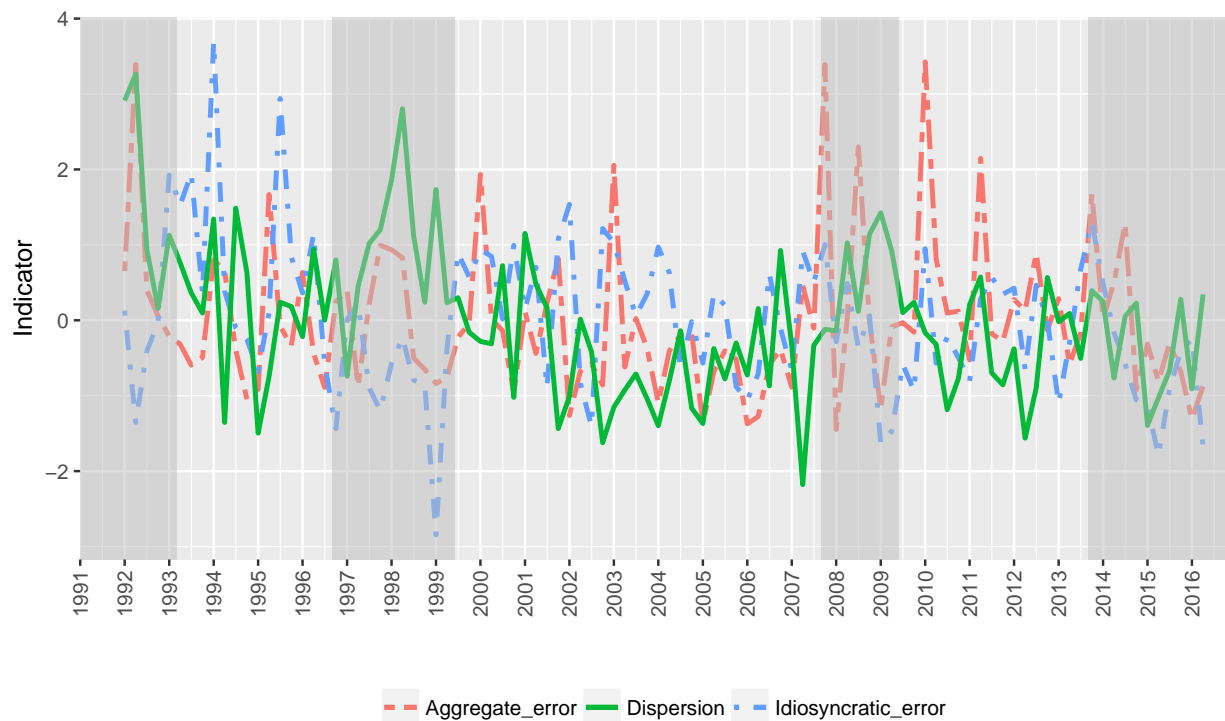


Figure 3.19: Weighted indicators of dispersion, aggregate error and idiosyncratic error uncertainty

Table 3.8: Factor loadings for the first principal component

Indicator	Loadings
Dispersion	0.55
Idiosyncratic error	0.00
Aggregate error	0.46
EPU	0.56
SAVI	0.41

The combined indicator seems particularly plausible as a proxy for uncertainty, as a number of large spikes coincide with periods when uncertainty in South Africa was thought to be relatively high. For instance, uncertainty was relatively high during the Democratic transition up to 1994. There was quite a large spike, mainly in policy and idiosyncratic uncertainty, associated with the foreign currency crisis before the adoption of the GEAR strategy. Other spikes coincide with the East Asian and Russian crises, and the related recessionary period in 1997-1998; the semi-recession in 2003; the global financial crisis in 2008 and the subsequent recession; the European debt crisis in 2011; and the start of the downswing phase in 2014.

Table 3.9 reports the contemporaneous correlations between the indicators and real GDP growth. The dispersion, EPU and combined uncertainty indicators exhibit significant negative correlations with real GDP growth. These indicators are contemporaneously counter-cyclical, as is the case for the majority of the uncertainty indicators in the international literature (e.g. Bloom, 2014). The idiosyncratic error uncertainty indicator does not exhibit the negative correlation with real GDP

## Business Sentiment

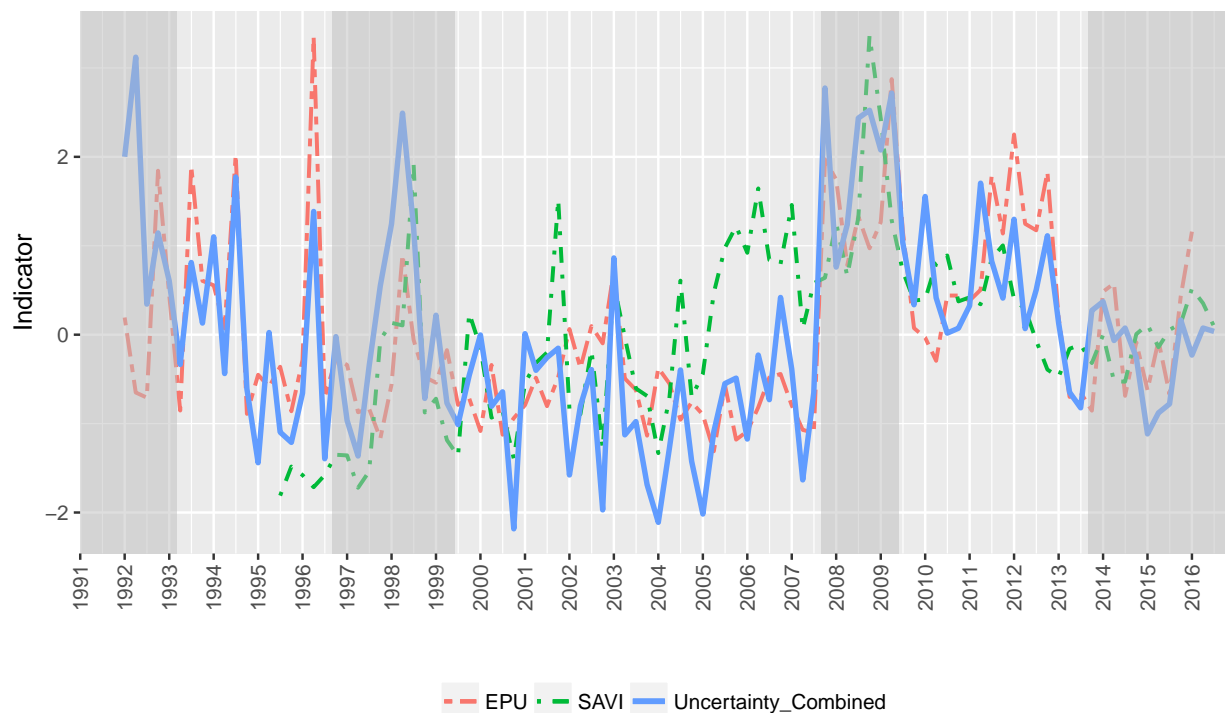


Figure 3.20: Indicators of economic policy, financial market, and combined uncertainty

growth found in Bachmann, Elstner and Sims (2013).<sup>28</sup>

The individual indicators are only weakly correlated with one another in a few cases. This is not too surprising as they attempt to capture different types of uncertainty (Leduc and Liu, 2016). Survey-based measures capture the opinions of key agents in the economy and are driven by changes in firm-level uncertainty. Due to their qualitative nature, however, they are poorly equipped to fully capture large increases in uncertainty during extreme events (Bachmann, Elstner and Sims, 2013). The SAVI captures broad uncertainty in financial markets, but is derived from a specific segment of firms that are publicly traded, while the EPU is focused specifically on policy uncertainty. This is the motivation for using a combined indicator, which captures different types of uncertainty from multiple sources.

The combined uncertainty indicator has a significant positive correlation with all of the indicators, except for idiosyncratic error uncertainty, which reflects the factor loadings used in deriving the first principal component. The dispersion indicator in particular, which in some ways is the simplest measure of uncertainty presented in this chapter, appears to drive the relationship between the combined index and GDP growth. It is therefore the most important measure of uncertainty

<sup>28</sup>It is possible that there is a structural explanation for the different relationship between idiosyncratic error uncertainty and real GDP growth. For instance, South African firms may react later to events such as recessions than US firms (i.e. they may be less forward-looking). Alternatively, it could be that something like growth effect is in operation in South Africa. There may also be problems with the survey data, either in terms of errors or unrepresentativeness.

## Business Sentiment

presented in this chapter.

Table 3.9: Correlations between the uncertainty indicators

	Dispersion	Idiosyncratic_error	Aggregate_error	EPU	SAVI	Combined
Idiosyncratic_error	-0.15					
Aggregate_error	0.20*	0.18*				
EPU	0.14	0.08	0.09			
SAVI	0.06	-0.24**	0.07	0.28**		
Combined	0.64***	-0.10	0.54***	0.65***	0.56***	
RGDP_Growth	-0.44***	0.17*	-0.11	-0.30***	-0.11	-0.43***

Figure 3.21 illustrates the cross-correlograms for the uncertainty indicators and real GDP growth. All of the indicators, except for idiosyncratic error uncertainty, exhibit a significant negative correlation with real GDP growth, albeit at different horizons. All the indicators seem to lead changes in real GDP growth. The combined indicator exhibits the highest negative correlation with real GDP growth at a lag of two quarters.

Table 3.10 reports the contemporaneous correlations for the sectoral indicators and sectoral real GDP growth. The combined uncertainty indicator for each sector is the first principal component of the three survey-based measures. The indicators of dispersion and combined uncertainty for the manufacturing, construction, and trade sectors are significantly negatively, if weakly, correlated with contemporaneous real sectoral GDP growth.

Table 3.10: Correlations between the sectoral uncertainty indicators and real GDP growth

	Manufacturing				Construction			
	Dispersion	Aggregate	Idiosyncratic	Combined	Dispersion	Aggregate	Idiosyncratic	Combined
Aggregate	0.17*				0.39***			
Idiosyncratic	-0.28***	-0.02			-0.26**	0.17		
Combined	0.81***	0.46***	-0.69***		0.89***	0.73***	-0.27***	
RGDP	-0.30***	0.04	0.10	-0.22**	-0.18*	-0.17	-0.05	-0.19*
	Trade				Services			
	Dispersion	Aggregate	Idiosyncratic	Combined	Dispersion	Aggregate	Idiosyncratic	Combined
Aggregate	-0.01				-0.08			
Idiosyncratic	-0.22**	0.18*			-0.01	0.18		
Combined	0.58***	-0.56***	-0.81***		0.35**	-0.77***	-0.70***	
RGDP	-0.28***	-0.09	0.23**	-0.23**	-0.05	-0.19	0.29*	-0.06

Figure 3.22 illustrates the cross-correlograms for the manufacturing indicators and real GDP growth in the manufacturing sector. The correlograms are similar to the aggregated results reported above, where the dispersion and aggregate error measures have a significant negative relationship with real GDP growth. The dispersion indicator again appears to drive the relationship between the combined index and GDP growth. The results for the construction and trade sectors are similar (not shown), while the indicators for the services sectors (not shown) do not exhibit a negative relationship with real services GDP growth.

This section has presented three survey-based indicators of uncertainty, as well as two additional popular proxies from the literature: stock market volatility and the news-based EPU created by Hlathwayo and Saxegaard (2016). All of these imperfect measures may contribute to our

## Business Sentiment

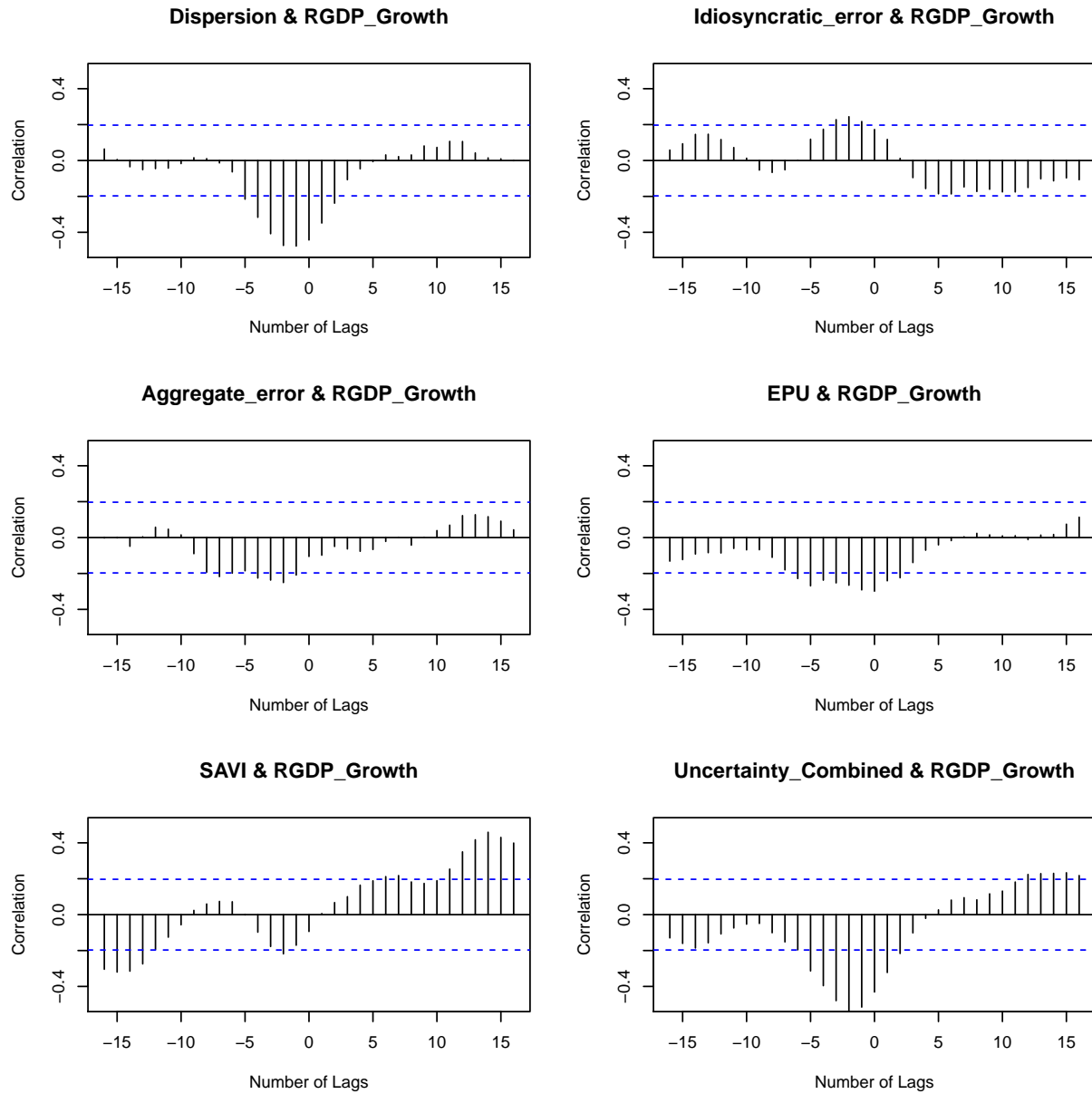


Figure 3.21: Cross-correlograms of uncertainty indicators and real GDP growth

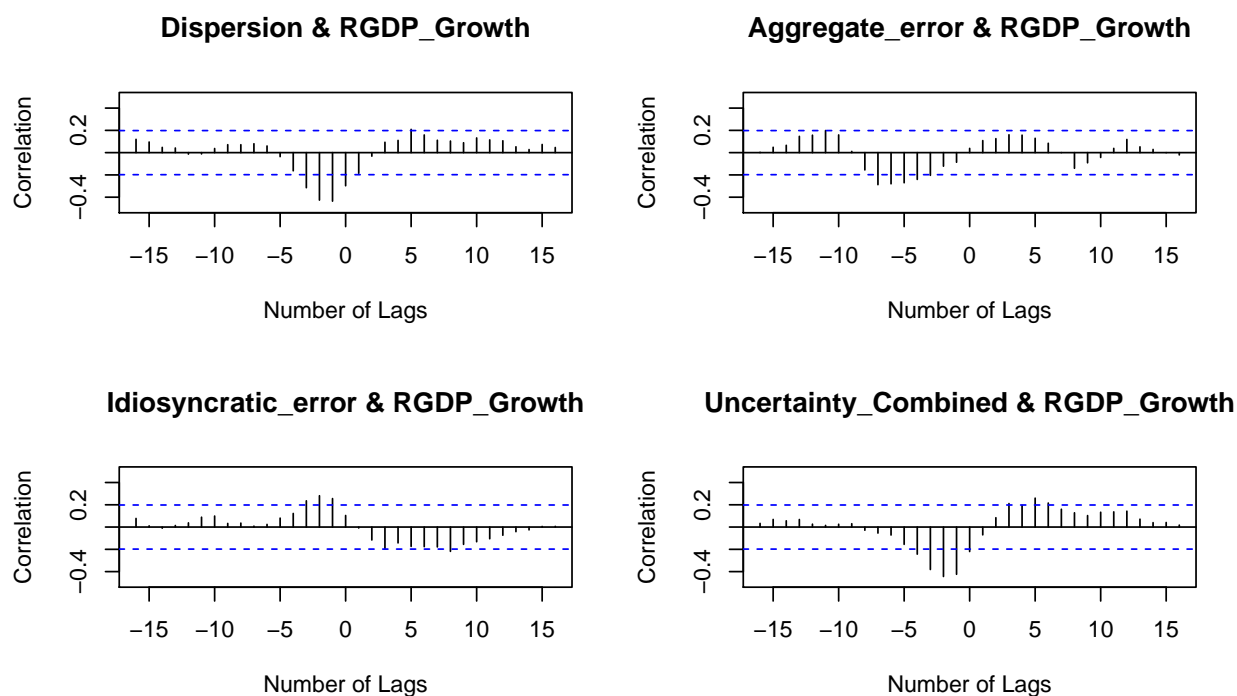


Figure 3.22: Cross-correlograms of manufacturing uncertainty indicators and real manufacturing GDP growth

understanding of uncertainty (Bachmann, Elstner and Sims, 2013). The five indicators were combined to form an overall combined uncertainty indicator, similar to those of Baker, Bloom and Davis (2015) and Redl (2015), to reflect the different sources of uncertainty from the different proxies. This combined indicator appears to be a plausible indicator of uncertainty in South Africa, reflecting key economic events. The composite dispersion and combined uncertainty indicators, in particular, improve on the existing uncertainty indicators in that they exhibit a larger negative correlation with real GDP growth.

### 3.8 The Relationship between Business Sentiment and Real Economic Activity

This section further examines the relationship between business sentiment and real economic activity in South Africa. This demonstrates the usefulness of the aggregation methods and the estimated indicators, and provides an additional validity test of the indicators. The hypothesis is tested that there is significant comovement between the sentiment indicators and real GDP growth. Granger causality tests are used to illuminate the timing of the relationships between the indicators and real output growth. Simple bivariate VARs are then estimated to investigate the dynamic effects of confidence and uncertainty shocks on the economy. A three-variable VAR and an extended VAR are then estimated to examine whether the results hold after the inclusion of additional variables.

## Business Sentiment

Table 3.11: Granger causality tests: confidence

Granger causality H0:	statistic	p-value
Confidence (Current) does not Granger-cause RGDP Growth	2.70*	0.07
RGDP Growth does not Granger-cause Confidence (Current)	1.41	0.25
Confidence (Expected) does not Granger-cause RGDP Growth	3.44**	0.03
RGDP Growth does not Granger-cause Confidence (Expected)	0.58	0.56
BER BCI does not Granger-cause RGDP Growth	4.14**	0.02
RGDP Growth does not Granger-cause BER BCI	1.69	0.19
SACCI Growth does not Granger-cause RGDP Growth	3.23**	0.04
RGDP Growth does not Granger-cause SACCI Growth	0.03	0.97

Table 3.12: Granger causality test statistics: sectoral confidence

Granger causality H0:	Manufacturing	Construction	Trade	Services
Confidence (Current) does not Granger-cause RGDP Growth	4.85***	9.88***	1.04	3.10*
RGDP Growth does not Granger-cause Confidence (Current)	3.23**	1.37	3.86**	0.42
Confidence (Expected) does not Granger-cause RGDP Growth	8.10***	11.19***	1.40	5.90***
RGDP Growth does not Granger-cause Confidence (Expected)	2.45*	0.00	6.01***	0.07
BER BCI does not Granger-cause RGDP Growth	3.79**	5.63**	0.60	
RGDP Growth does not Granger-cause BER BCI	3.01*	0.03	2.84*	

### 3.8.1 Granger Causality Tests

Granger causality tests are often performed when investigating the comovement between variables. This test determines whether one time series is useful in forecasting another, by measuring the ability of lagged values of a time series to predict the future values of another time series. A time series  $Z$  is said to Granger-cause  $Y$  if it can be shown that the  $Z$  values provide statistically significant information of future values of  $Y$ . If the hypothesis that a sentiment indicator does not Granger-cause an economic variable is rejected, it implies that past values of sentiment provide significant information for the economic variable, in addition to its own history.

Table 3.11 reports the results for Granger causality tests for the confidence indicators and real GDP growth. The results suggest that the lagged values of all four confidence indicators significantly predict real GDP growth, with no evidence of Granger-causality in the reverse direction. In other words, the results suggest that all the confidence indicators contain relevant information for the prediction of output growth. This implies that the measures all exhibit a leading relationship with real GDP growth.

Table 3.12 reports the results of the Granger causality tests for the sectoral confidence indicators and their corresponding real sectoral GDP growth rates. The results are similar to those for the aggregate indicators, except for the trade sector, where lagged values of real GDP growth significantly predict all three survey-based confidence indicators. This implies that the confidence indicators for the trade sector are lagging indicators of real GDP growth in that sector.

Table 3.13 reports the results of Granger causality tests for the uncertainty indicators and real GDP growth. The results suggest that the lagged values of three of the uncertainty indicators significantly predict real GDP growth, with no evidence of Granger-causality in the reverse direction. In other

## Business Sentiment

Table 3.13: Granger causality tests: uncertainty

<b>Granger causality H0:</b>	<b>statistic</b>	<b>p-value</b>
Dispersion does not Granger-cause RGDP Growth	3.57**	0.03
RGDP Growth does not Granger-cause Dispersion	1.25	0.29
Aggregate error does not Granger-cause RGDP Growth	7.28***	0.00
RGDP Growth does not Granger-cause Aggregate error	0.13	0.88
Idiosyncratic error does not Granger-cause RGDP Growth	1.20	0.30
RGDP Growth does not Granger-cause Idiosyncratic error	0.98	0.38
EPU does not Granger-cause RGDP_Growth	0.93	0.43
RGDP Growth does not Granger-cause EPU	1.93	0.13
SAVI does not Granger-cause RGDP Growth	1.26	0.29
RGDP Growth does not Granger-cause SAVI	1.01	0.36
Uncertainty (Combined) does not Granger-cause RGDP Growth	5.85***	0.00
RGDP Growth does not Granger-cause Uncertainty (Combined)	0.06	0.94

Table 3.14: Granger causality test statistics: sectoral uncertainty

<b>Granger causality H0:</b>	<b>Manufacturing</b>	<b>Construction</b>	<b>Trade</b>	<b>Services</b>
Dispersion does not Granger-cause RGDP Growth	7.50***	2.69	0.34	0.09
RGDP Growth does not Granger-cause Dispersion	1.76	0.01	0.46	4.54**
Aggregate error does not Granger-cause RGDP Growth	1.52	1.13	2.10	0.44
RGDP Growth does not Granger-cause Aggregate error	1.09	0.28	0.12	0.90
Idiosyncratic error does not Granger-cause RGDP Growth	3.18**	0.42	1.48	1.61
RGDP Growth does not Granger-cause Idiosyncratic error	1.14	0.57	0.73	2.33
Uncertainty (Combined) does not Granger-cause RGDP Growth	9.61***	2.99*	1.60	0.76
RGDP Growth does not Granger-cause Uncertainty (Combined)	1.35	0.02	0.87	1.59

words, the results suggest that the dispersion, aggregate error, and combined uncertainty indicators contain relevant information for the prediction of output growth. This implies that these measures exhibit a leading relationship with real GDP growth.

The results for the sectoral indices, reported in Table 3.14, are not consistent across the sectors. There is some evidence of Granger-causality for a few of the indicators for the manufacturing sector. The tests are not significant at conventional levels in the trade sector, and in the construction sector only the combined uncertainty indicator significantly Granger-causes real trade GDP growth. In the services sector dispersion indicator seems to lag real GDP growth.

### 3.8.2 VAR Analysis

This section provides evidence on the dynamic effects of sentiment shocks on real economic activity. As many economic variables move together over time, without an obvious causal direction, it can be challenging to identify the directions of relationships. In the literature, timing has often been relied on for identification. This section follows the literature (e.g. Taylor and McNabb (2007); Barsky and Sims (2012); Bachmann, Elstner and Sims (2013)) in using standard recursive VARs to trace out the dynamic responses of economic activity to surprise shocks in sentiment. The aim is to investigate whether the indicators have a significant dynamic relationship with real output, whether they contain predictive content for output growth, and whether shocks to sentiment generate responses

that are in line with the theory and the findings in the literature.

The relationships were investigated for the aggregate variables, as well as separately for each sector, using bivariate recursive VARs featuring a measure of sentiment and real GDP growth. A bivariate system is a parsimonious way to model the joint dynamics of sentiment and real economic activity (Bachmann, Elstner and Sims, 2013). In the bivariate case, both variables are treated as endogenous:

$$y_t = \beta_{10} - \beta_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \epsilon_{yt}$$

$$z_t = \beta_{20} - \beta_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \epsilon_{zt},$$

where  $y$  is output,  $z$  is sentiment, and  $\epsilon$  is the residual of each equation.

A range of VARs were estimated for the quarterly data running from 1992Q1 to 2016Q3. The indicators enter in levels, while the real GDP series enter as annual quarter-on-quarter growth rates, which corresponds with the survey reference period. Unit root tests, reported in the Appendix, indicate that virtually all of the aggregate and sectoral indicators, and the corresponding real GDP growth rates are stationary. The exception is confidence on current conditions in the services sector, which may be due to the relatively short sample period. The appropriate number of lags are selected by means of the Akaike information criterion (AIC), the Schwarz criterion (SC) and the Hannan-Quinn criterion (HQ). The most parsimonious model is selected, provided that the diagnostic tests (i.e. no serial correlation, homoskedasticity and normality) are satisfied. In the majority of cases, the information criteria point to two lags. The model fit is best when a constant term is included.

The sentiment indicators are ordered first in a recursive identification strategy, with the Cholesky decomposition used to identify structural shocks. With this ordering, shocks to sentiment are allowed to have a contemporaneous impact on output, but shocks to output have no contemporaneous impact on sentiment ( $\beta_{21} = 0$ ). In other words, innovations to the confidence indicators influence economic output on impact, but not vice versa. This is the identification strategy and ordering used in the literature (e.g. Leduc and Sill (2013), Bachmann, Elstner and Sims (2013), Girardi and Reuter (2017), Baker, Bloom and Davis (2016), and Redl (2015)). It can be motivated by the timing of the surveys before the release of most macroeconomic data (Leduc and Liu, 2016). When the survey is completed in time  $t$ , the respondents do not know the realisations of output growth in time  $t$ , as the response deadline is generally the second month of the quarter.

### 3.8.2.1 Confidence

Impulse response functions (IRFs) can be generated to illustrate the dynamic impact of a shock to sentiment on the system. The shock is an innovation of one standard deviation to the residual in the equation. Figure 3.23 illustrates the IRFs of a bivariate VAR for the confidence indicator on current conditions and real GDP growth. The left panel plots the responses of real GDP growth to



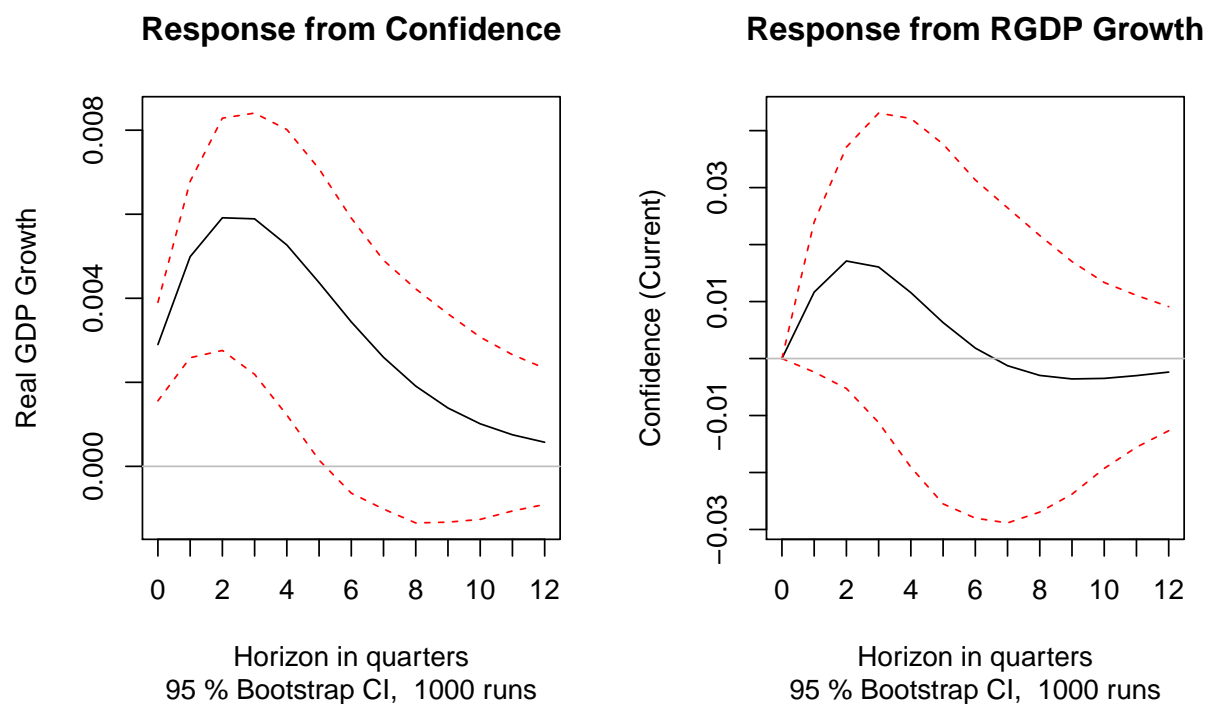


Figure 3.23: IRFs of confidence (current conditions) and real GDP growth

an orthogonal shock in the indicator, with 95% bootstrap confidence intervals. Following an increase in confidence of one standard deviation, real GDP growth increases by around 0.3% on impact, with a peak at two quarters. The impact on the growth rate is transitory, dying out after approximately seven quarters. This is equivalent to a permanent increase in the level of output, which corresponds to the findings in the literature (e.g. Barsky and Sims (2012)). The right panel plots the response of confidence to an orthogonal shock in real GDP growth. Following an increase in real GDP growth, there is an insignificant increase in confidence of around 2% after two quarters. The results are similar for alternative orderings. As reported in the Appendix, the results are remarkably similar for the confidence indicator on expected conditions and the BER BCI, whereas the SACCI growth rate exhibits a smaller significant relationship with real activity after two quarters.

The importance of innovations can also be examined with variance decompositions. While the IRFs describe the reaction of a variable of interest to an exogenous shock, the decomposition of the forecast error variance of a given variable shows how much of the error can be explained by exogenous shocks to the other variables in the system (Girardi and Reuter, 2017). The forecast error variance decomposition (FEVD) shows the proportion of the movements in a sequence due to its own shocks and shocks to the other variable. Figure 3.24 illustrates the FEVDs for the current conditions confidence indicator and real GDP growth. Up to around half (46%) of the movements in real GDP growth are explained by the confidence indicator over the longer term, while real GDP explains up to 2% of the variance in the confidence indicator.

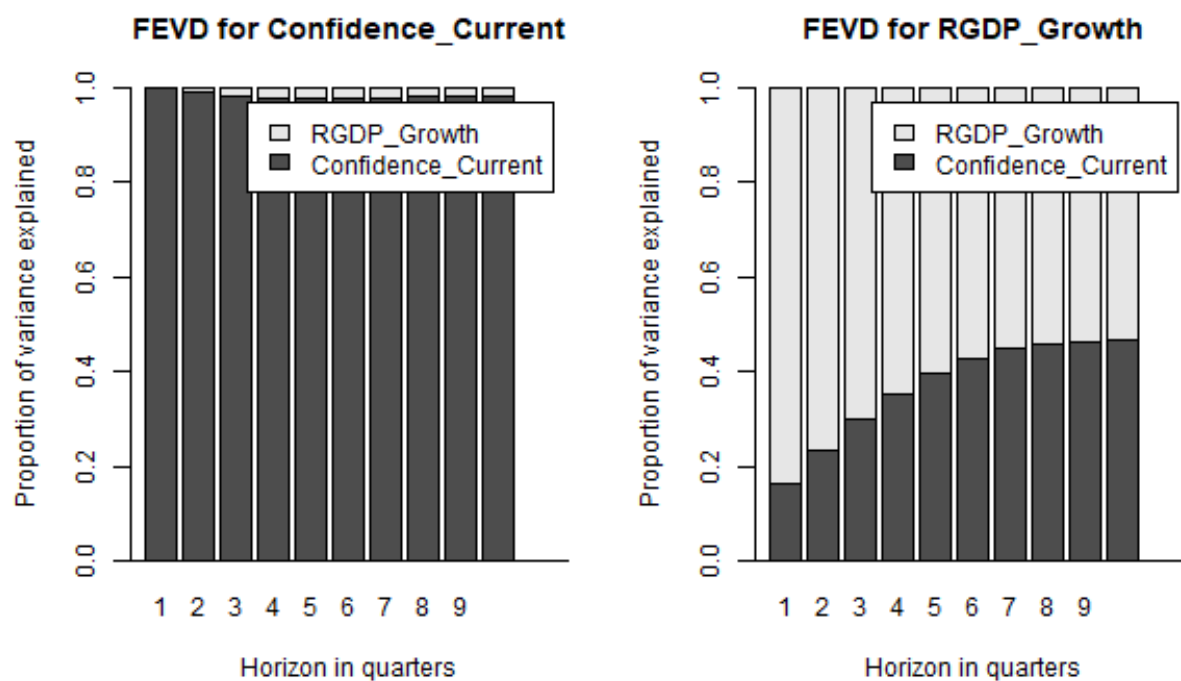


Figure 3.24: FEVDs of confidence (current conditions) and real GDP growth

The results for the sectoral indicators are very similar to the aggregate results. Figure 3.25 illustrates the IRFs of a bivariate VAR for the current conditions confidence indicator in the manufacturing sector and real GDP growth in the manufacturing sector. Following an increase in confidence, real GDP growth increases on impact, with a peak at two quarters. The impact on the growth rate dies out after approximately four quarters. Following an increase in real GDP growth, there is a significant increase in confidence in the following quarter. The results are similar for alternative orderings. The results for the other sectoral indicators (not shown) are very similar to those for the manufacturing sector, with the exception that in the construction sector, the impact of a shock to confidence on GDP growth does not die out within the forecast horizon of 12 quarters.

Figure 3.26 illustrates the FEVDs for the current conditions confidence indicator and real GDP growth in the manufacturing sector. Up to more than a third (37%) of the movements in real GDP growth are explained by confidence over the longer term, while real GDP explains up to 5% of the variance in confidence. Overall, the results suggest that shocks to the confidence indicators account for between 20% and 60% of the forecast error variance of the real GDP growth rate, depending on the level of aggregation and the indicator.

### 3.8.2.2 Uncertainty

Figure 3.27 illustrates the IRFs of a bivariate VAR with the dispersion indicator and real GDP

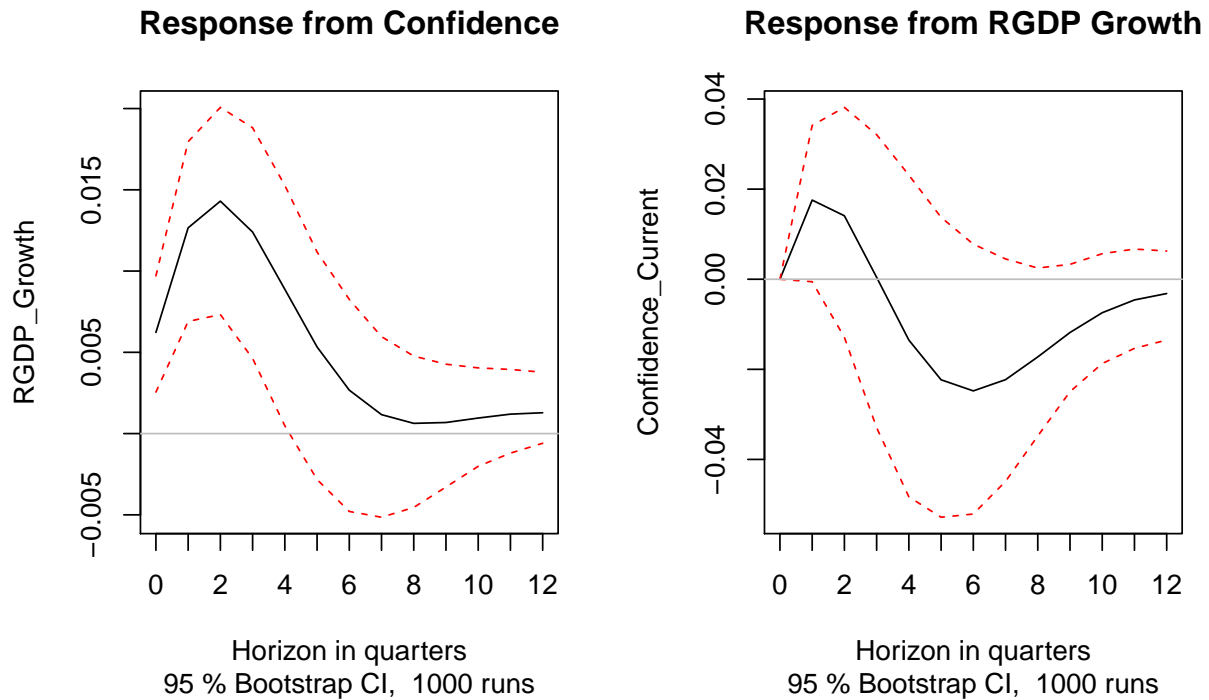


Figure 3.25: IRFs of confidence (current) and real GDP growth in the manufacturing sector

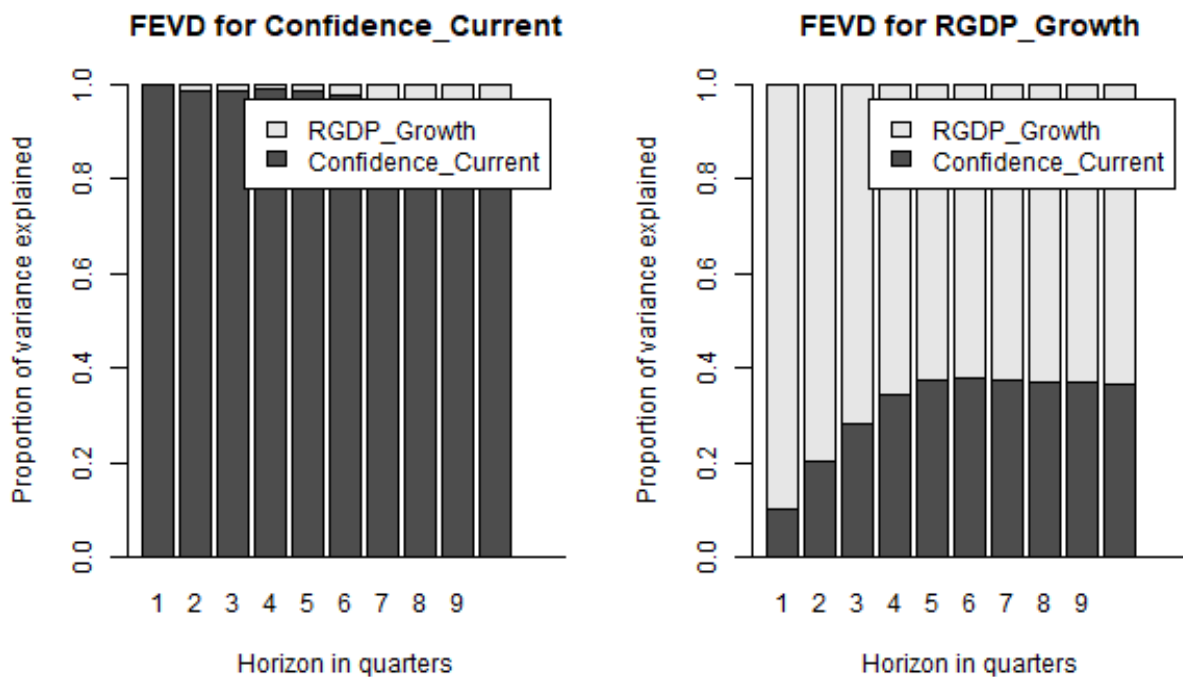


Figure 3.26: FEVDs of confidence (current) and real GDP growth in the manufacturing sector

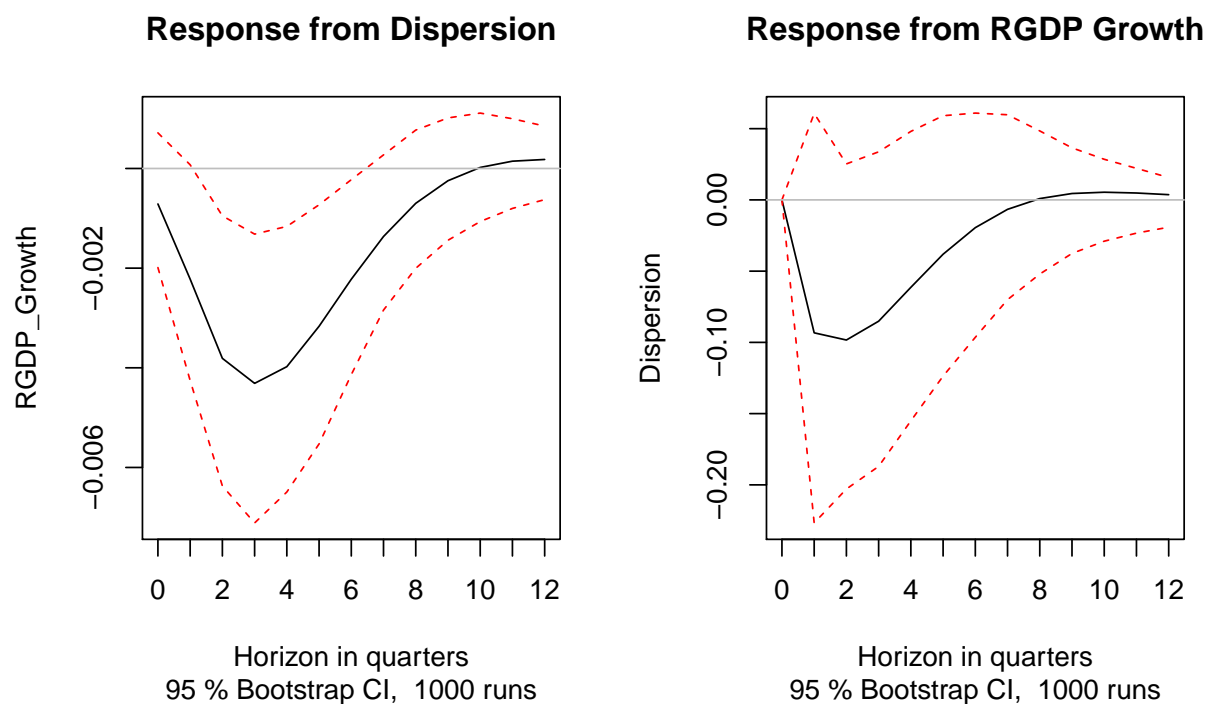


Figure 3.27: IRFs of dispersion and real GDP growth

growth. The left panel plots the responses of real GDP growth to an orthogonal shock in dispersion, with 95% bootstrap confidence intervals. A shock of one standard deviation to dispersion is followed by a significant decrease in real GDP growth in the following quarter. The right panel plots the response of dispersion to an orthogonal shock in real GDP growth. Following a shock to real GDP growth, there is an insignificant response in the dispersion indicator. The IRFs for the aggregate error, EPU, and SAVI indicators of uncertainty are very similar. Shocks to these indicators are associated with a moderately significant decreases in real GDP growth, while shocks to real GDP growth do not lead to a significant changes in the uncertainty indicators. The IRFs for the idiosyncratic error uncertainty indicator are not significant. The results are similar for alternative orderings.

Figure 3.28 illustrates the IRFs of a bivariate VAR with the combined uncertainty indicator and real GDP growth. A shock to uncertainty is followed by a significant decrease in real GDP growth, with a peak at three quarters. The impact is larger and more significant than for any of the component uncertainty indicators separately. The impact on the growth rate is also transitory, dying out after approximately seven quarters. This result corresponds to the findings in much of the literature (e.g. Bachmann, Elstner and Sims (2013) and Redl (2015)), where innovations to uncertainty have protracted negative effects on economic activity.

Figure 3.29 illustrates the FEVDs for the combined uncertainty indicator and real GDP growth. Over 30% of the movements in real GDP growth are explained by the uncertainty indicator over the longer term, while real GDP explains about 1% of the variance in uncertainty. This is in line

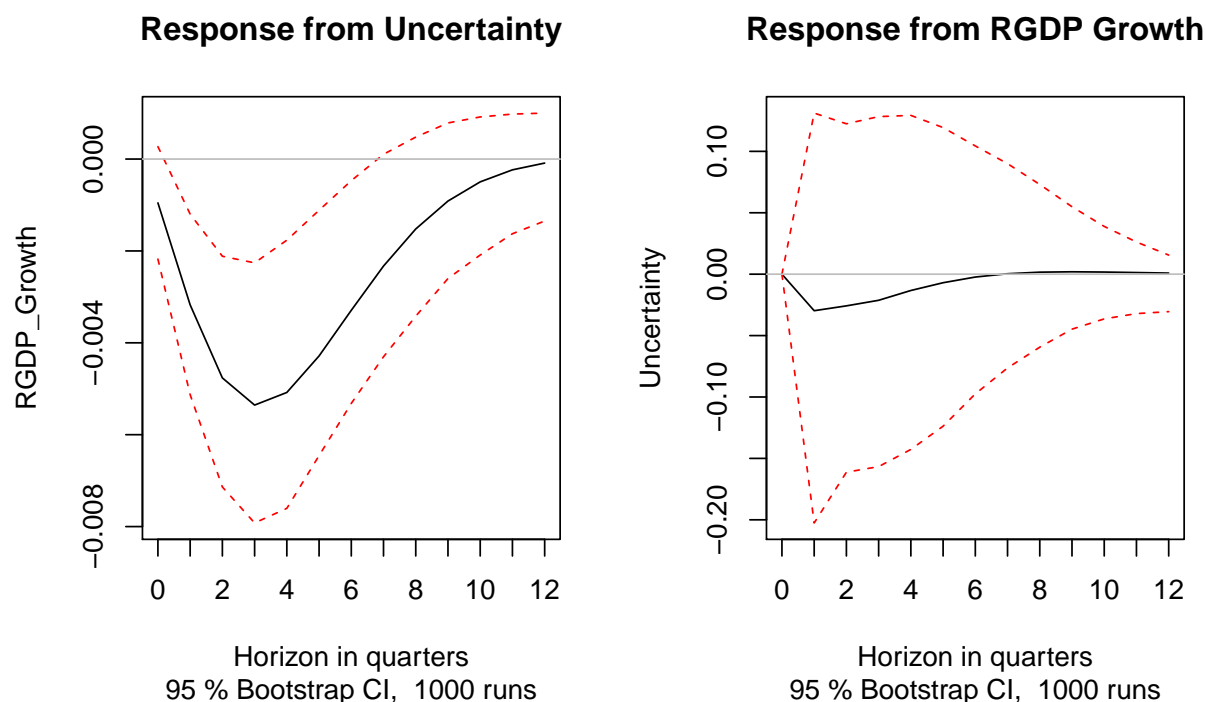


Figure 3.28: IRFs of uncertainty (combined) and real GDP growth

with findings in the literature (e.g. Bachmann, Elstner and Sims (2013)).

Figure 3.30 illustrates the IRFs of a bivariate VAR for the combined manufacturing uncertainty indicator and real GDP growth in the manufacturing sector. A shock to uncertainty is followed by a significant decrease in real GDP growth, with a peak at two quarters. There is even some evidence of a subsequent rebound predicted by the ‘wait-and-see’ effect demonstrated in Bloom (2009). The FEVDs illustrated in Figure 3.31 show that around 30% of the movements in real manufacturing GDP growth are explained by uncertainty over the longer term.

There is no consistent negative relationship for any single indicator and real GDP growth in the other three sectors (not shown). In the construction sector, the dispersion and combined uncertainty indicators have a significant negative impact on real GDP growth. In the trade sector, only the dispersion indicator has a significant negative impact, while in the services sector only the idiosyncratic error indicator has a significant impact on real GDP growth.

### 3.8.2.3 Expanded VAR

Though instructive, the results from a bivariate system are prone to misspecification (Girardi and Reuter, 2017). In order to test the robustness of the negative effect of uncertainty shocks, a number of authors have extended the baseline setup to include measures of confidence (e.g. Girardi and Reuter (2017), Leduc and Liu (2016), Baker, Bloom and Davis (2016) and Bachmann, Elstner

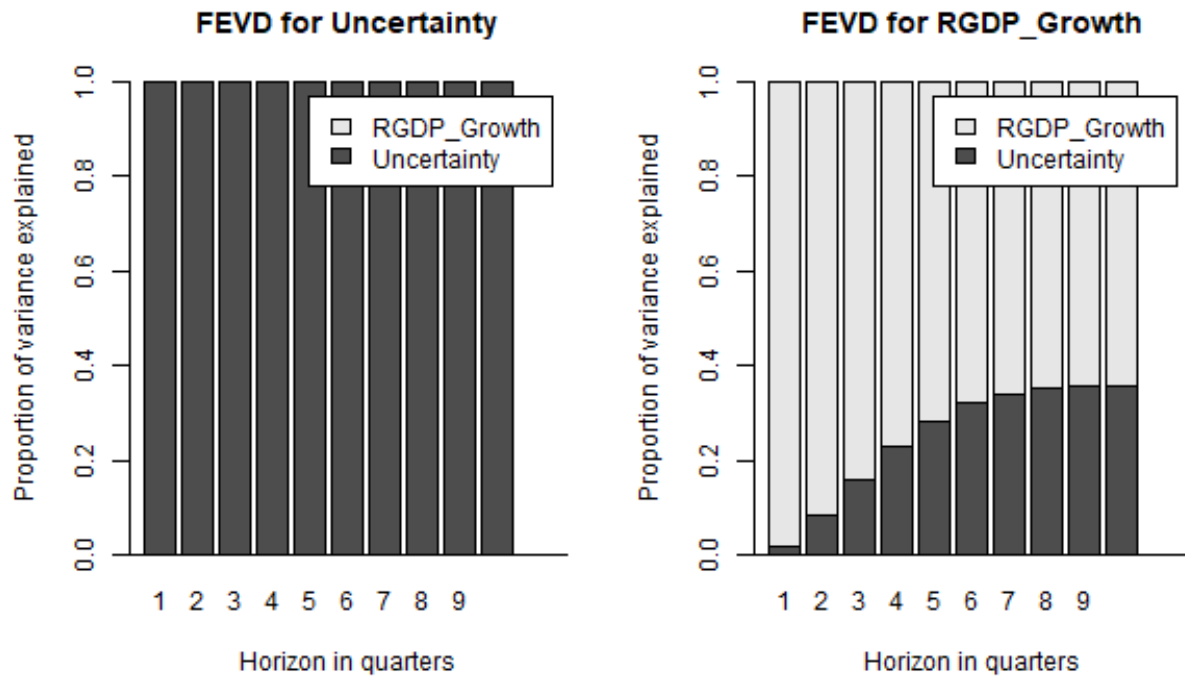


Figure 3.29: FEVDs of uncertainty (combined) and real GDP growth

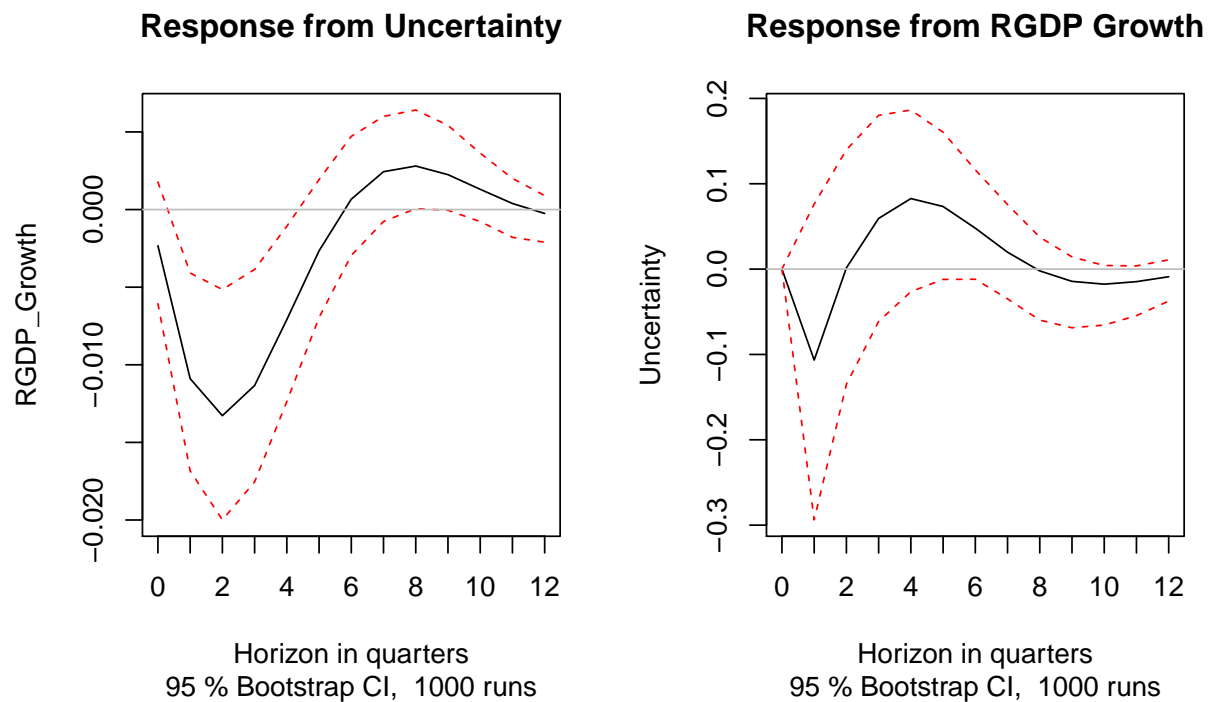


Figure 3.30: IRFs of uncertainty (combined) and real GDP growth in the manufacturing sector

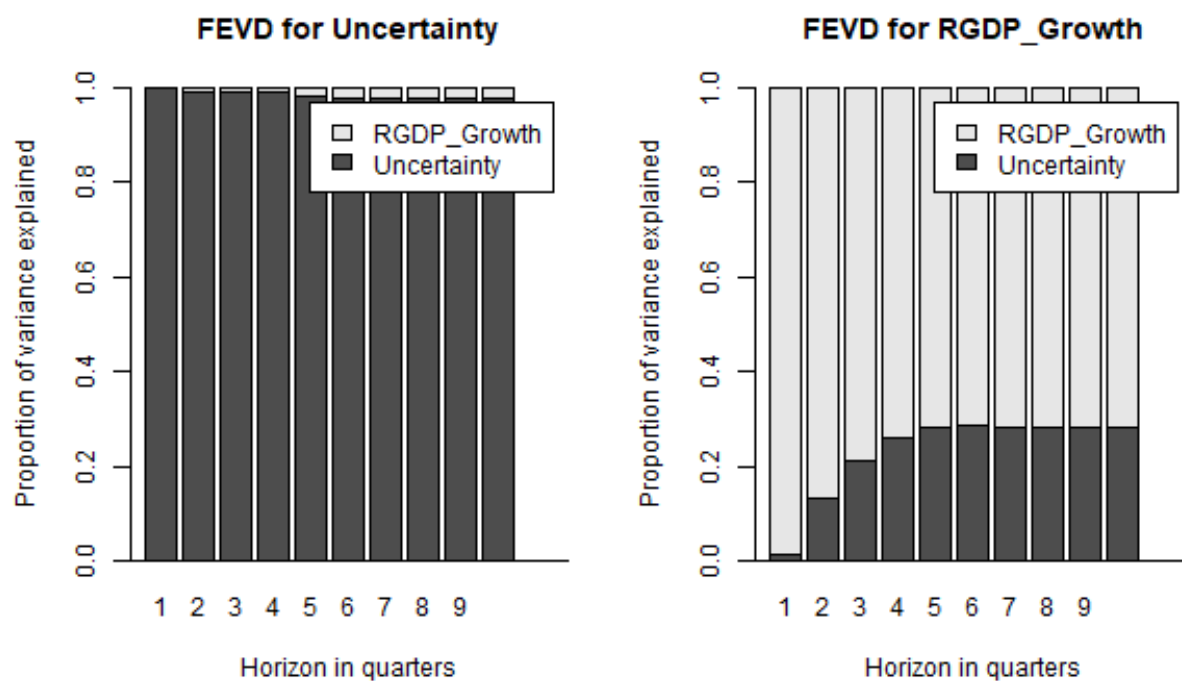


Figure 3.31: FEVDs of uncertainty (combined) and real GDP growth in the manufacturing sector

and Sims (2010)). Periods of increased uncertainty also tend to be periods of bad economic news. Confidence is usually included to control for the possibility that the impact of uncertainty may reflect respondents' perceptions of bad news rather than of an uncertain future (Baker, Bloom and Davis, 2016).

Figure 3.32 illustrates the current conditions measure of confidence and the combined uncertainty indicator with recessionary periods shaded. The two sentiment indicators do not appear to be mirror images of each other, with a correlation of -0.36. Although confidence is pro-cyclical and uncertainty is mostly anti-cyclical, they appear to capture different phenomena.

Figure 3.33 illustrates the IRFs of a three-variable VAR including confidence on current conditions, combined uncertainty and real GDP growth. Following Girardi and Reuter (2017), confidence was ordered first under a recursive identification scheme. The results are very similar to the IRFs for the bivariate VARs reported earlier. A positive shock to confidence is followed by a significant increase in real GDP growth, while a positive shock to uncertainty is followed by a significant decrease in real GDP growth. Figure 3.34 illustrates the FEVDs for this three-variable VAR. Up to around 30% of the variance in real GDP growth is explained by confidence over the longer term, while uncertainty explains around 25% of the variance.

A larger VAR system was also estimated to test the robustness of the relationships. The extended VAR includes the variables suggested by Redl (2015) for South Africa: confidence, uncertainty, the

Business Sentiment

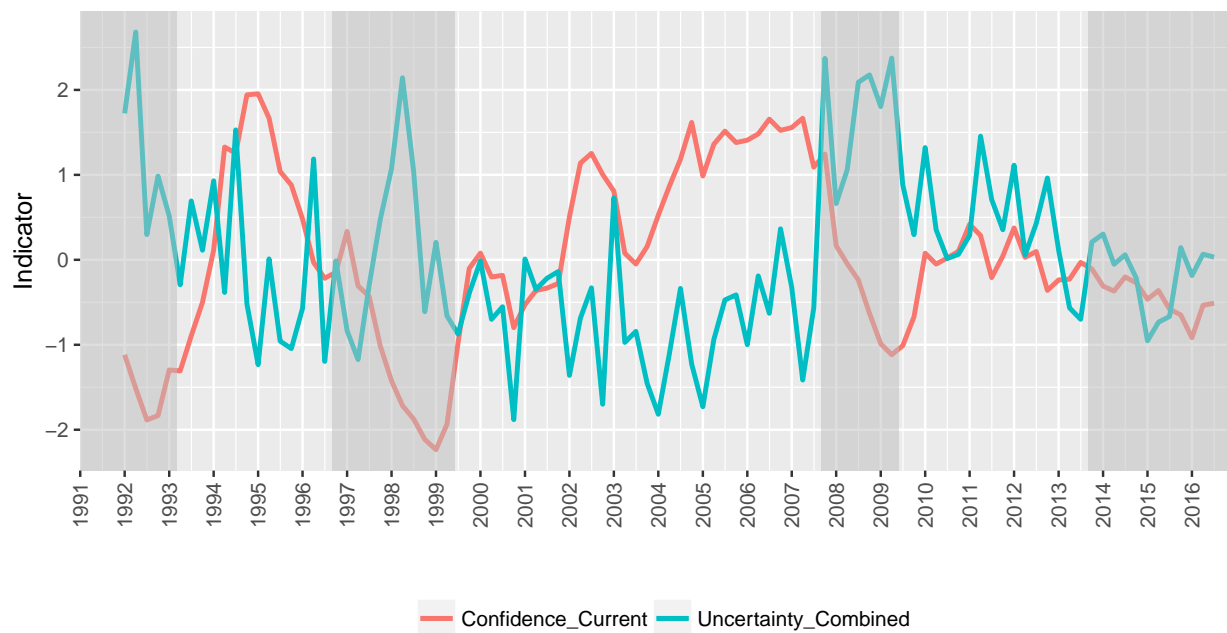


Figure 3.32: Confidence (current) and uncertainty (combined)

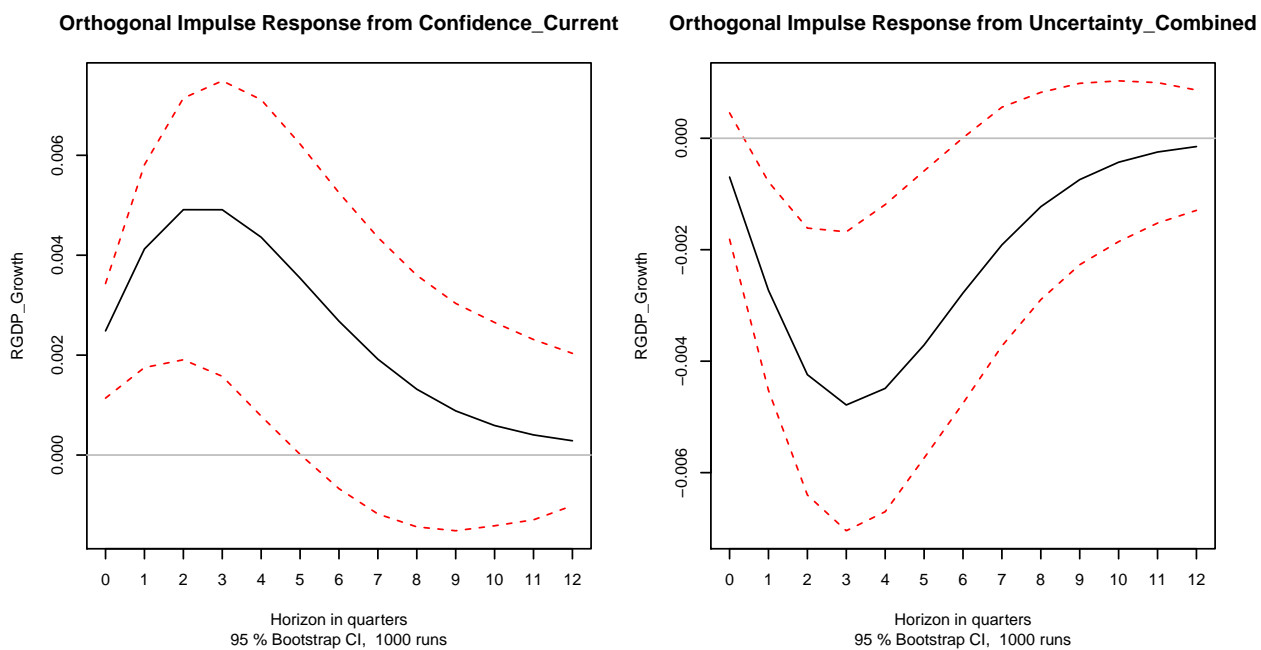


Figure 3.33: IRFs of real GDP growth to confidence and uncertainty in the three-variable VAR



## Business Sentiment

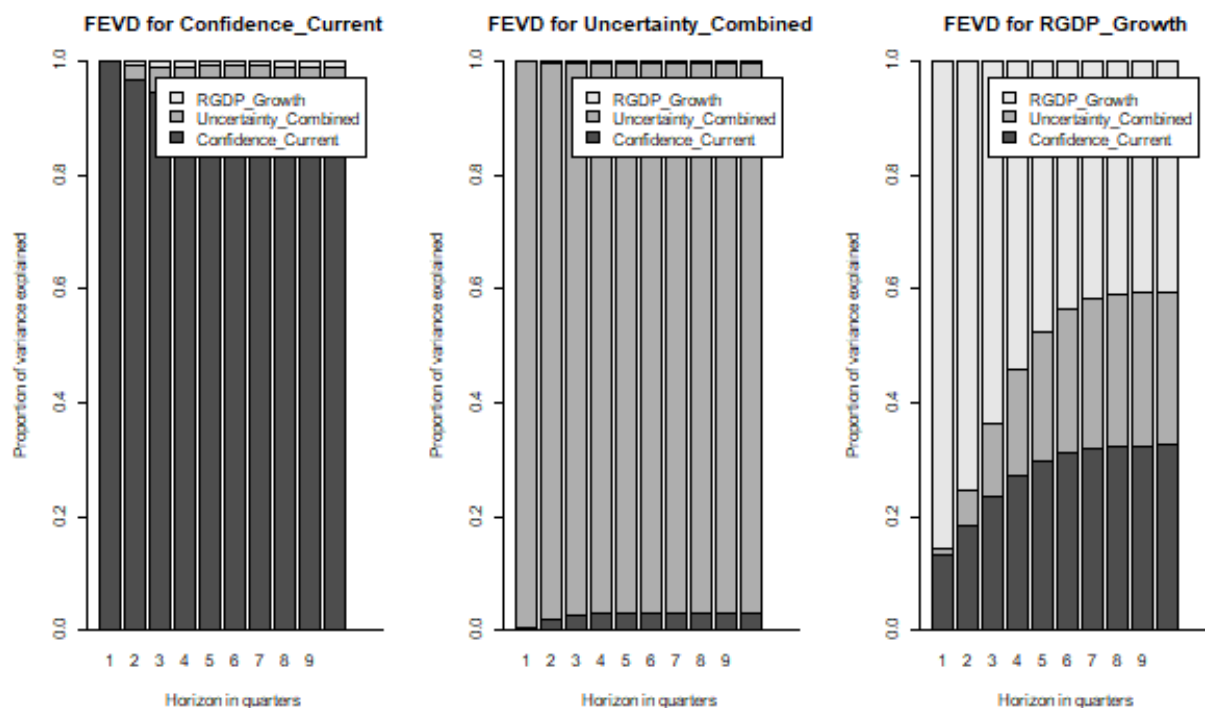


Figure 3.34: FEVDs of the three-variable VAR

JSE All Share Index, the yield spread (i.e. the Government Bond Yield minus the three-month T-Bill rate), GDP, industrial production, investment, and an employment index. These variables are typically included in the literature (e.g. Leduc and Sill (2013), Bachmann, Elstner and Sims (2013), and Baker, Bloom and Davis (2016)).

The variables were ordered with the sentiment variables first, the financial variables next and the real variables last. The financial variables were expected to move faster than the real variables (Redl, 2015). An alternative ordering of placing the sentiment indicators last provides qualitatively similar results. As was the case in the previous VARs, the variables enter as real annual quarter-on-quarter growth rates, except for the sentiment indices and the yield spread. The model was estimated with two lags, with the caveat that the information criteria indicate that more than the maximum number of lags are appropriate. The results with four lags are qualitatively similar.

The IRFs for the impact of confidence (current conditions) and uncertainty (combined) on the growth rate of real GDP, real production and real investment are illustrated in Figure 3.35. The top panels illustrate the responses of the real variables to a shock in confidence, and the bottom panels show the responses to a shock in uncertainty. The larger system seems to provide similar results to the findings in bivariate VARs. The response in real GDP growth are similar to those in the three-variable model. The impacts of the shocks are larger on real production and investment growth than on real GDP growth. This is what the wait-and-see theory would predict. The responses of employment (not illustrated) are very similar to those of real GDP growth. According to the FEVD

## Business Sentiment

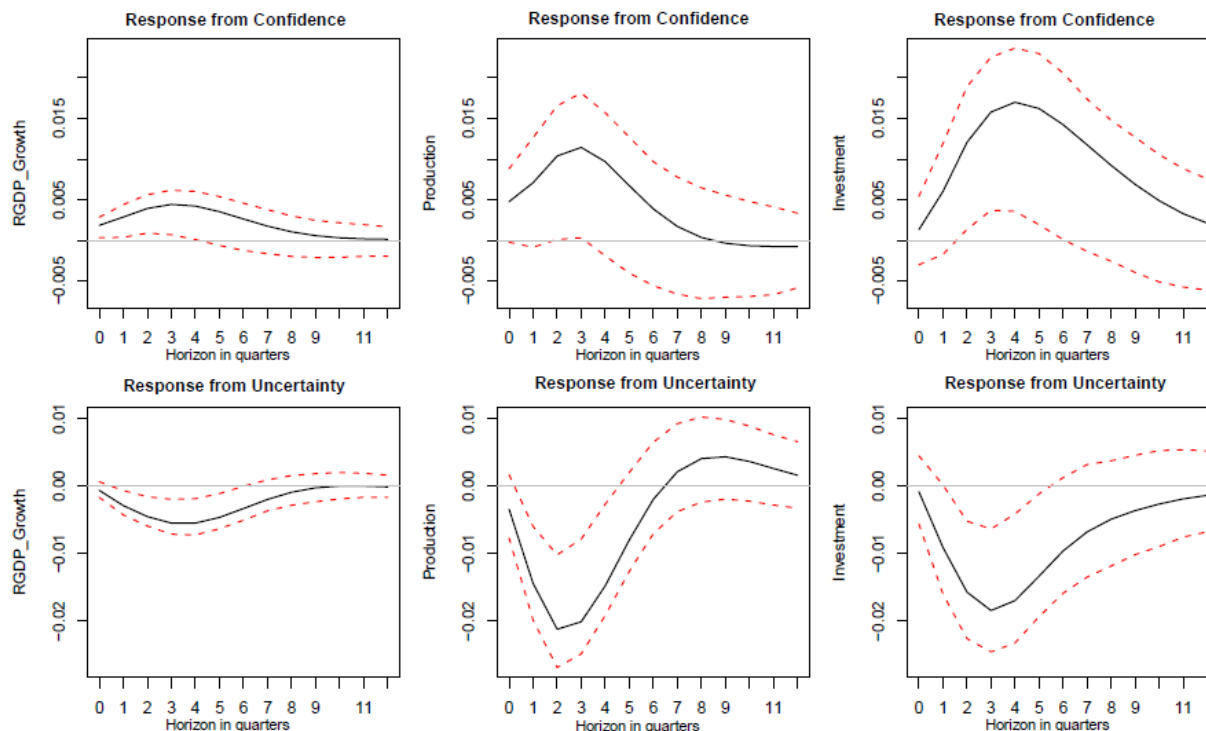


Figure 3.35: IRFs of real GDP, production and investment growth to confidence and uncertainty

(not shown), confidence explains around 35%, 25%, and 40% of the variance in real GDP growth, real production growth and real investment growth respectively, while uncertainty explains around 15%, 25% and 20% of the variance in the three real variables.

### 3.9 Conclusion

This chapter attempted three contributions to the literature. The first was to demonstrate aggregation methods to estimate sentiment indicators from the disparate qualitative responses of individual firms. The chapter used different combinations of the weighted cross-sectional first and second moments of the distribution of the qualitative survey responses to create composite indicators of confidence and uncertainty.

The weighted cross-sectional moments employed in this chapter would be useful in other applications with qualitative survey responses, such as consumer surveys, where there are challenges in capturing the full richness in the data. Consumer confidence measures are popular indicators and are often calculated using qualitative surveys. Examples include the European Commission's Consumer Confidence Indicator and the University of Michigan's Consumer Sentiment Index for the US (United Nations, 2015). The BER calculates consumer confidence for South Africa using their consumer tendency surveys. It would be possible to improve on the existing measures of consumer confidence using the techniques demonstrated in this chapter to identify an underlying pattern

from the disparate views of individual agents. Moreover, it would be possible to create new measures of consumer uncertainty, by calculating the scaled cross-sectional dispersion of responses (which does not require a panel), as long as there are forward-looking questions as well as questions on current conditions. The consumer sentiment indicators may also be combined with the business sentiment indicators to create general sentiment indicators, in the same way in which the European Commission creates its Economic Sentiment Index.

The second contribution was to produce new composite indicators of confidence and uncertainty for South Africa, which are reported in Table 3.19 in the Appendix below. The sectoral indicators are available on request. The new composite indicators outperformed the existing confidence indicators in terms of their correlation with real GDP growth and their concordance with the official SARB business cycle. The BER BCI is often used as a leading indicator of the business cycle, for example, by the SARB and Laubscher (2014). The new confidence indicators may therefore be useful as improved leading indicators of the business cycle. The composite dispersion and combined uncertainty indicators, in particular, exhibited larger negative correlations with real GDP growth than the existing uncertainty indicators.

The third contribution was to use these indicators to contribute to the literature on the relationship between sentiment and real economic activity in the South African setting. The results provide evidence of at least important comovement between the indicators of sentiment and real economic activity. Both sets of sentiment indicators contained significant predictive information for real economic activity, even after controlling for additional variables.

The indicators might be useful for forecasting and nowcasting exercises, especially given that they are available before official statistics are published. The GDP figures used in this chapter are the revised numbers, which are only produced after a significant delay. When the indicators are used for real-time forecasting, however, the issue of data revisions becomes a problem. Econometric forecasts are typically based on revised data but are evaluated against the first release data. However, as Van Walbeek (2006) noted for South Africa, the growth in GDP has been subject to significant upward revisions and bias, especially after 1994. Forecast estimation using the first release versions of the GDP figures would most likely produce markedly different estimates to those using the revised data. In such a case, a ‘poor forecast’ could in fact be a poor first release of the official data (Van Walbeek, 2006). Future research might evaluate the performance of the indicators in real time, as well as the impact of official data revisions.

The uncertainty indicators could be used to further investigate the importance of uncertainty shocks for business cycle fluctuations and credit cycles, and whether this relationship is non-linear or asymmetric. The forecasting ability of the indicators might be offset completely by other variables during ordinary times, while increasing notably in the presence of unusual events. For instance, shocks to sentiment might play an important role only during episodes of economic tension associated with large decreases in confidence and heightened uncertainty. It may also be used to inform other analyses, such as the influence of uncertainty on the responsiveness of exports to relative price

changes, studied in Hlatshwayo and Saxegaard (2016).

The results imply that the current climate of heightened political uncertainty and weak consumer and business confidence are potential determinants of the lower growth that the South African economy is currently experiencing, as argued by the International Monetary Fund (2017). Future research could try to identify the potential causal impact of these changes in sentiment. Moreover, it may be informative to investigate the factors that drive the indicators of sentiment. The new sentiment indicators created in this chapter may facilitate these inquiries.

## **3.10 Appendix**

### **3.10.1 Confidence Intervals for the Confidence Indicators**

This section illustrates confidence intervals around the confidence indices. The intervals are calculated as two times the standard deviation of the weighted sample mean. Figure 3.36 illustrates 95% confidence intervals around confidence indicators on current and expected conditions. Figure 3.37 illustrates 95% confidence intervals around the sectoral confidence indicators on current conditions. The confidence intervals show that the distribution of the sample means are relatively narrow, because of the large number observation in each quarter. As a consequence, the changes in the indices seem to be ‘real’ changes rather than statistical idiosyncrasies. The confidence intervals for the scaled dispersion indicator (not reported) are similarly narrow and imply that the changes are real.

### **3.10.2 Stable Sample Indicators**

In order to test whether the entry and exit or attrition rates of firms drive the results, a number of robustness exercises are carried out. In this section the indicators are calculated by including only firms that form part of smaller, more ‘stable’, samples, which are then compared to the full sample indicators.

#### **3.10.2.1 The sample of firms that responded in consecutive periods (forecast error sample)**

This section presents the confidence and uncertainty indicators based on firms that responded in consecutive quarters (i.e. the forecast error sample). Figure 3.38 and Figure 3.39 compare the full sample indicators to those based only on the more stable forecast error sample. Table 3.15 reports that the correlations between these indicators are high and significant. The table also reports the correlations between the smaller forecast error sample indicators and real GDP growth, which are very similar to those for the full sample. The full sample measures, which do not rely on the panel

Business Sentiment

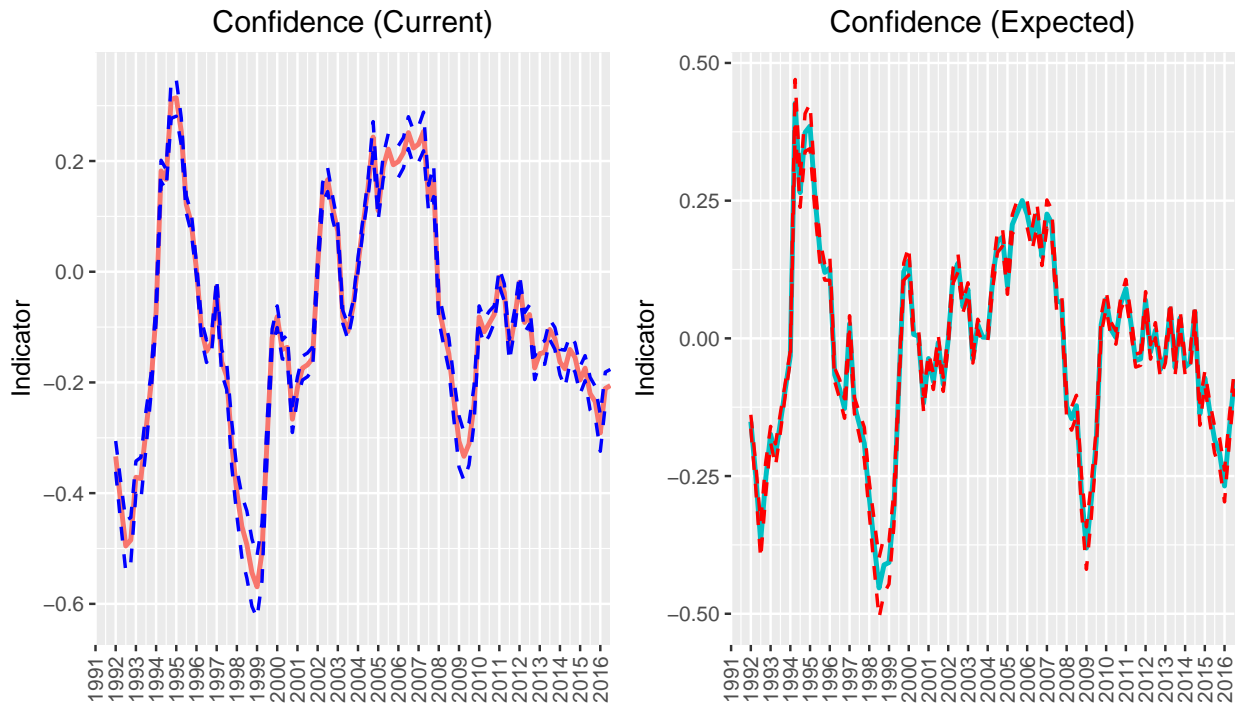


Figure 3.36: Confidence indicators with 95% confidence intervals

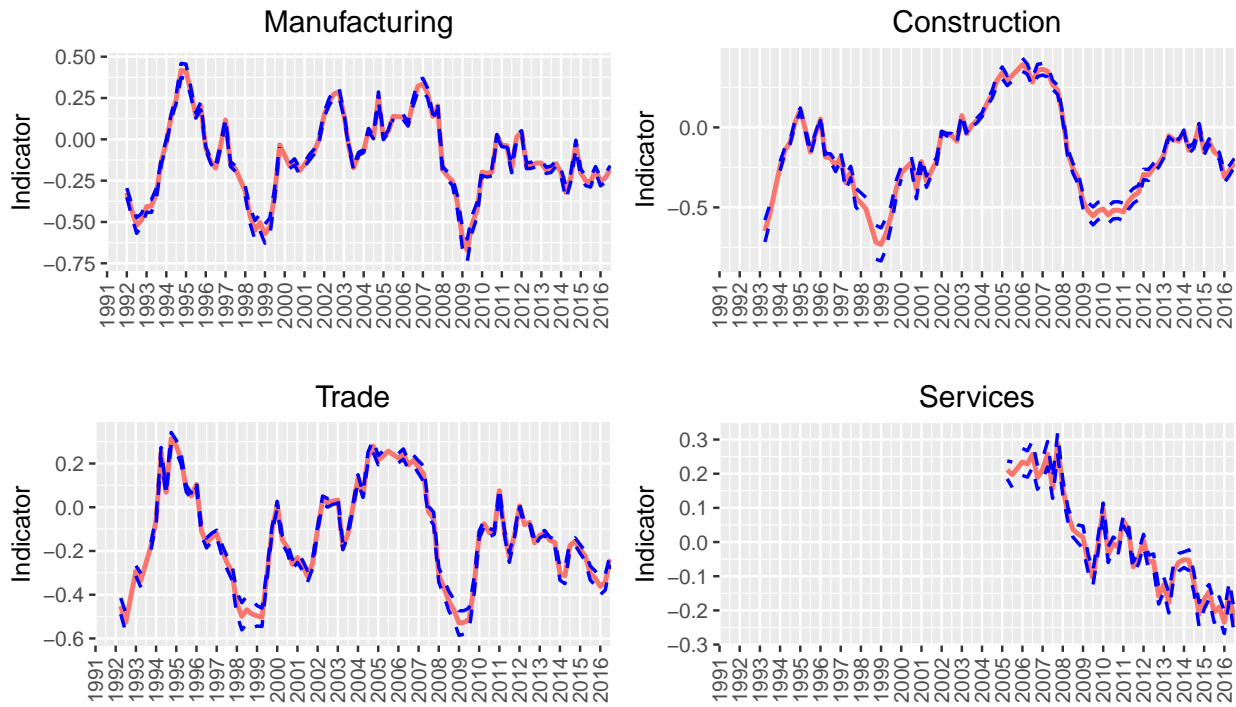


Figure 3.37: Sectoral confidence indicators with 95% confidence intervals

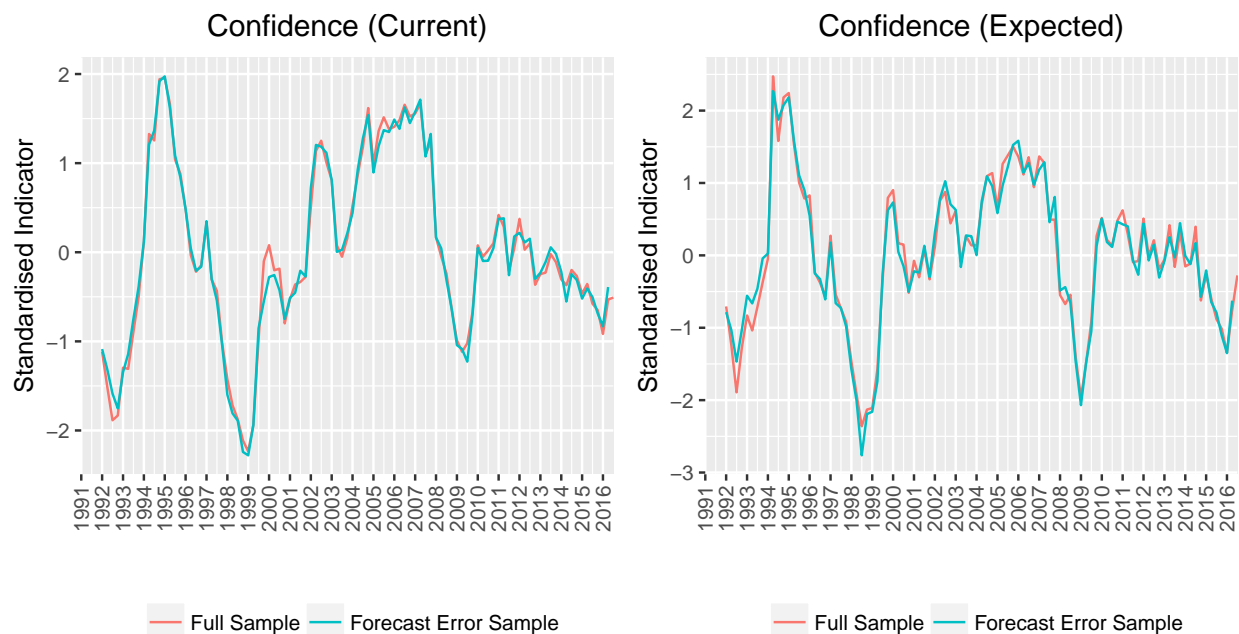


Figure 3.38: Confidence indicators based on the full and forecast error samples

Table 3.15: Correlations between indicators based on the full and forecast error samples

Indicator	Full-Stable Sample	Stable-RGDP Growth
Confidence (Current)	0.99***	0.76***
Confidence (Expected)	0.99***	0.68***
Dispersion	0.91***	-0.44***
Uncertainty (Combined)	0.96***	-0.31***

structure, therefore seem to be robust to only calculating them using this smaller stable sample. This implies that the results are not driven by attrition rates.

### 3.10.2.2 Samples of firms that responded relatively frequently

Figure 3.40 reports the sample sizes for two relatively stable samples of firms, based on how frequently firms responded to the surveys. In the two versions, only firms that responded to more than 50% and to more than 75% of all the surveys over the period are included. Table 3.16 reports that the sample characteristics, in terms of firm size, are similar to those for the full sample, even though these stable samples are substantially smaller.

Figure 3.41 compares the full sample confidence indicators to the two sets of indicators based on the more stable samples. The confidence indicators based on these smaller samples seem to capture very similar trends, even though they are estimated which substantially smaller samples. This provides confidence that the results are not driven by the entry and exit patterns of firms.

Figure 3.42 compares the full sample uncertainty indicators to the two sets of indicators based on

Business Sentiment

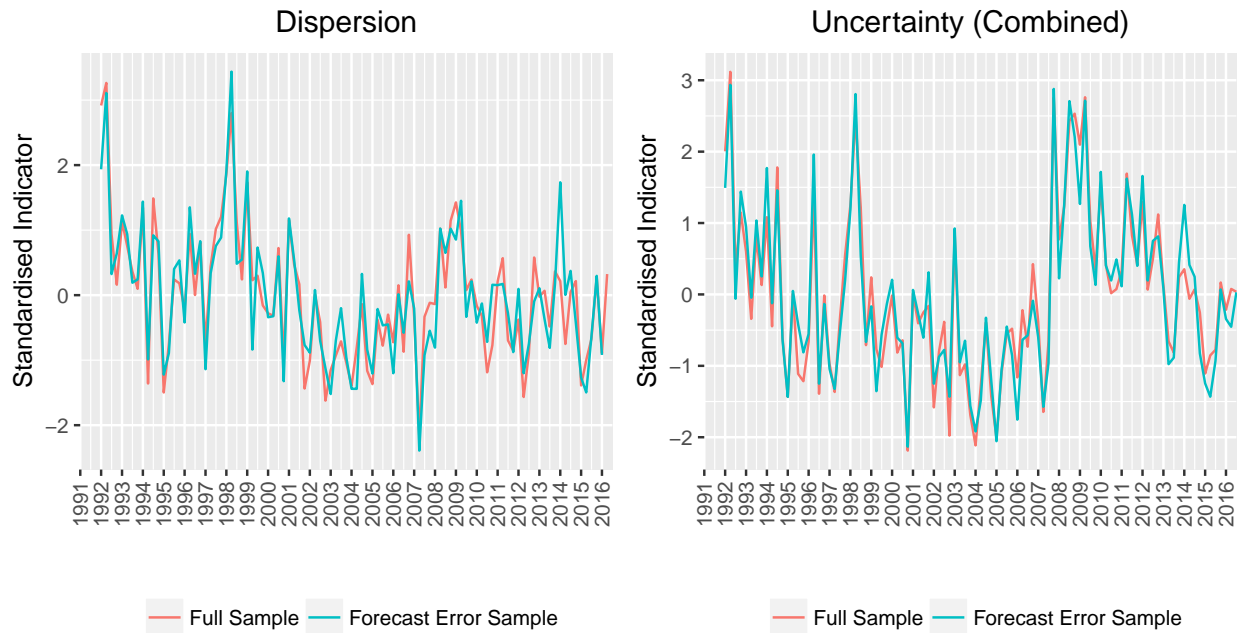


Figure 3.39: Uncertainty indicators based on the full and forecast error samples

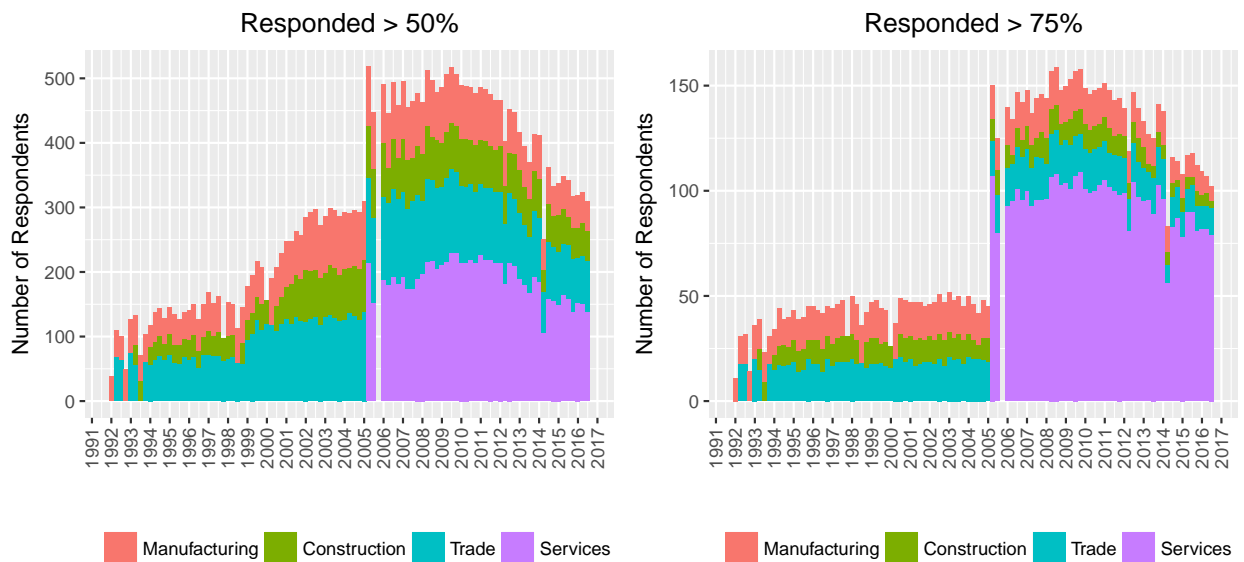


Figure 3.40: Number of respondents in the stable samples

Business Sentiment

Table 3.16: Sample characteristics in terms of firm size (full and stable samples)

Firm Size Category	Full Sample		Greater than 50% sample		Greater than 75% sample	
	Observations	Percentage	Observations	Percentage	Observations	Percentage
1	25,587	21.43%	5,504	18.67%	1,169	14.06%
2	15,288	12.80%	3,843	13.04%	1,232	14.82%
3	18,554	15.54%	4,328	14.68%	917	11.03%
4	13,717	11.49%	2,993	10.15%	724	8.71%
5	14,676	12.29%	4,888	16.58%	1,565	18.82%
6	9,140	7.65%	2,578	8.74%	867	10.43%
7	6,899	5.78%	1,819	6.17%	577	6.94%
8	6,894	5.77%	1,474	5.00%	491	5.90%
9	8,667	7.26%	2,053	6.96%	773	9.30%

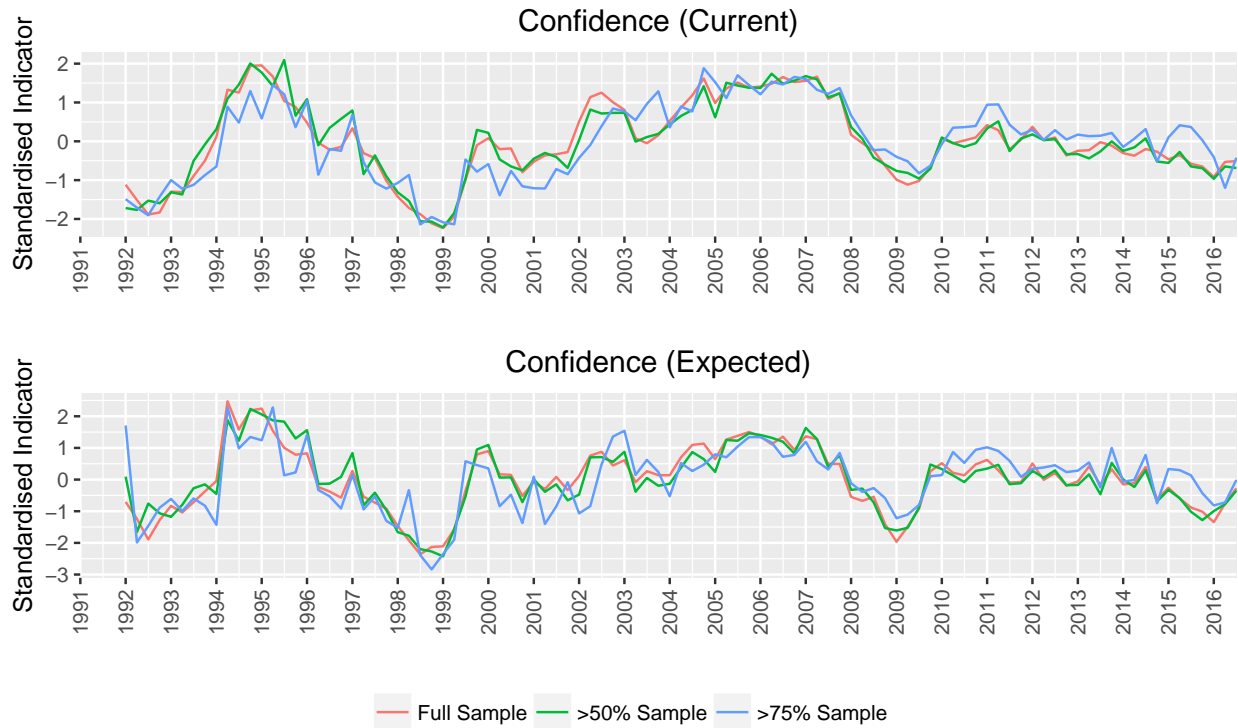


Figure 3.41: Confidence indicators based on the full and stable samples



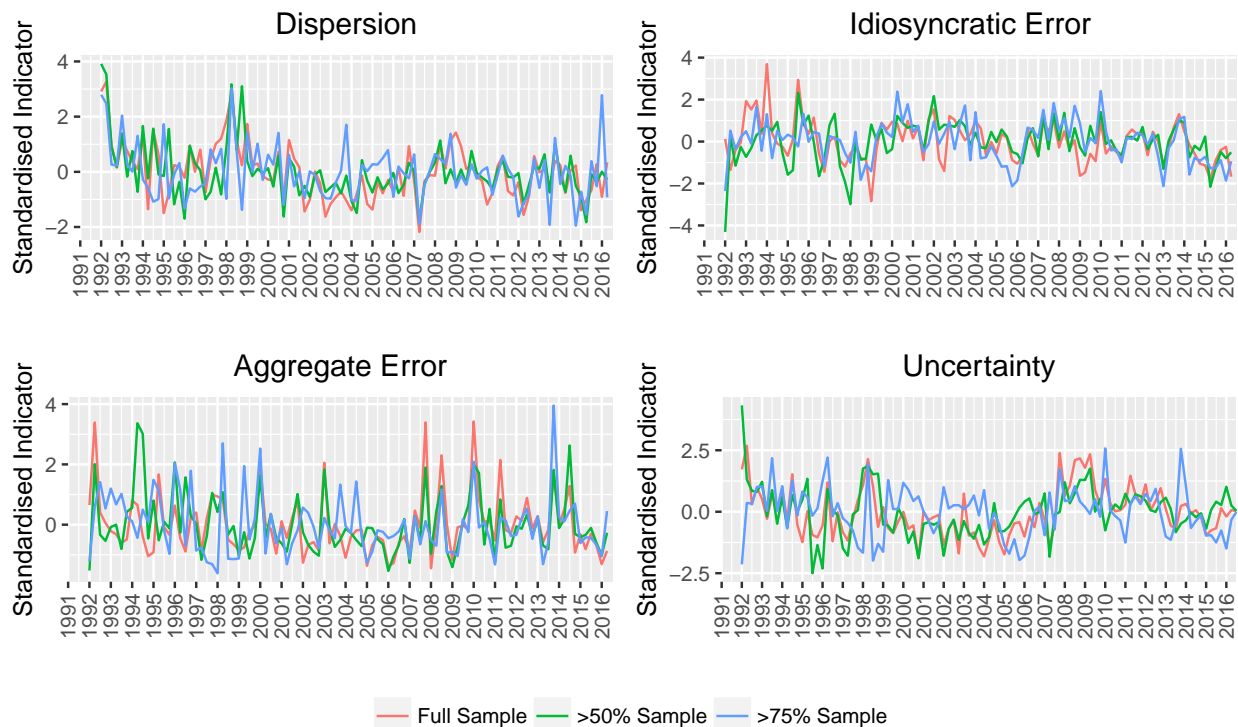


Figure 3.42: Uncertainty indicators based on the full and stable samples

the smaller samples. The indicators still seem to point to similar periods of heightened uncertainty. However, there are a few instances where the indicators seem to differ, which is unsurprising given that the ‘stable’ samples are so much smaller. The combined uncertainty indicators from the smallest sample is the least similar to its full sample counterpart. This is because the PCA procedure exacerbates the differences in the indicators, as its factor loadings are not fixed.

Table 3.17 reports the correlations between the full sample indicators and their smaller sample counterparts, respectively. The correlations for the confidence indicators are high and significant, even for the smallest sample. The correlations are lower but still significant for the uncertainty indicators, even though these indicators are very volatile. Table 3.17 also reports the correlations between real GDP growth and the smaller sample indicators. The confidence indicators still exhibit significant positive correlations with real GDP growth, although the sizes are moderated. The dispersion indicators still exhibit significant negative correlations with real GDP growth. The same holds true for the larger sample (>50%) combined uncertainty indicator, although the correlation is not present for the smallest sample version.

The indicators therefore seem to be relatively robust to only calculating them for more stable samples. This is mostly the case even for the smallest sample, which includes fewer than 50 firms before 2005. This implies that these firms are driving a substantial part of the results, rather than the entry and exit of firms.

## Business Sentiment

Table 3.17: Correlations between indicators based on the full and stable samples

Indicator	>50% Sample		>75% Sample	
	Full Sample	RGDP Growth	Full Sample	RGDP Growth
Confidence (Current)	0.97***	0.77***	0.87***	0.65***
Confidence (Expected)	0.95***	0.66***	0.80***	0.50***
Dispersion	0.70***	-0.35***	0.45***	-0.23**
Aggregate error	0.64***	0.00	0.38***	-0.02
Idiosyncratic error	0.46***	0.26**	0.52***	0.14
Uncertainty (Combined)	0.67***	-0.42***	0.38***	0.00

Table 3.18: Unit root test statistics

	Total	Manufacturing	Construction	Trade	Services
Confidence (Current)	-2.76***	-2.81***	-1.85*	-2.63***	-1.05
Confidence (Expected)	-3.02***	-3.19***	-2.67***	-3.21***	-2.35**
Dispersion	-5.51***	-6.5***	-4.27***	-2.69***	-5.98***
Aggregate error	-5.36***	-8.13***	-4.08***	-6.42***	-4.48***
Idiosyncratic error	-7.28***	-5.73***	-4.53***	-5.93***	-3.54***
Uncertainty (combined)	-4.43***	-6.32***	-3.37***	-4.41***	-3.84***
Real GDP Growth	-2.13**	-4.65***	-1.96**	-2.33**	-3.84***

### 3.10.3 Unit Root Tests

The unit root tests for the series used in the VARs are reported in Table 3.18. The tests indicate that virtually all of the aggregate and sectoral indicators, and the corresponding real GDP growth rates, are stationary. The exception is confidence on current conditions in the services sector, which may be due to the relatively short sample period.

### 3.10.4 IRFs for the Confidence Indicators

This section reports the impulse response functions (IRFs) from bivariate VARs that include each of the other confidence indicators and real GDP growth. As before, the survey-based indicators enter in levels, while the SACCI BCI and real GDP series enter as annual quarter-on-quarter growth rates. The confidence indicators are ordered first in a recursive identification strategy, with the Cholesky decomposition used to identify structural shocks. The appropriate number of lags are selected by means of the AIC, the SC and the HQ.

Figure 3.43, Figure 3.44 and Figure 3.45 illustrate the IRFs from the VARs including confidence on expected conditions, the BER BCI and the SACCI BCI (in growth rates), respectively. The results are remarkably similar for the expected conditions confidence indicator and the BER BCI, whereas the SACCI growth rate exhibits a smaller significant relationship with real activity after two quarters. The IRF of the original BER BCI is slightly moderated compared to the IRF of the current conditions confidence indicator in the main text. This implies that firms are not necessarily responding to the published figures, although there may be some element of self-fulfilling prophecy in firms' decisions. It does suggest that the more robust confidence indicators are capturing some form of true or latent confidence.

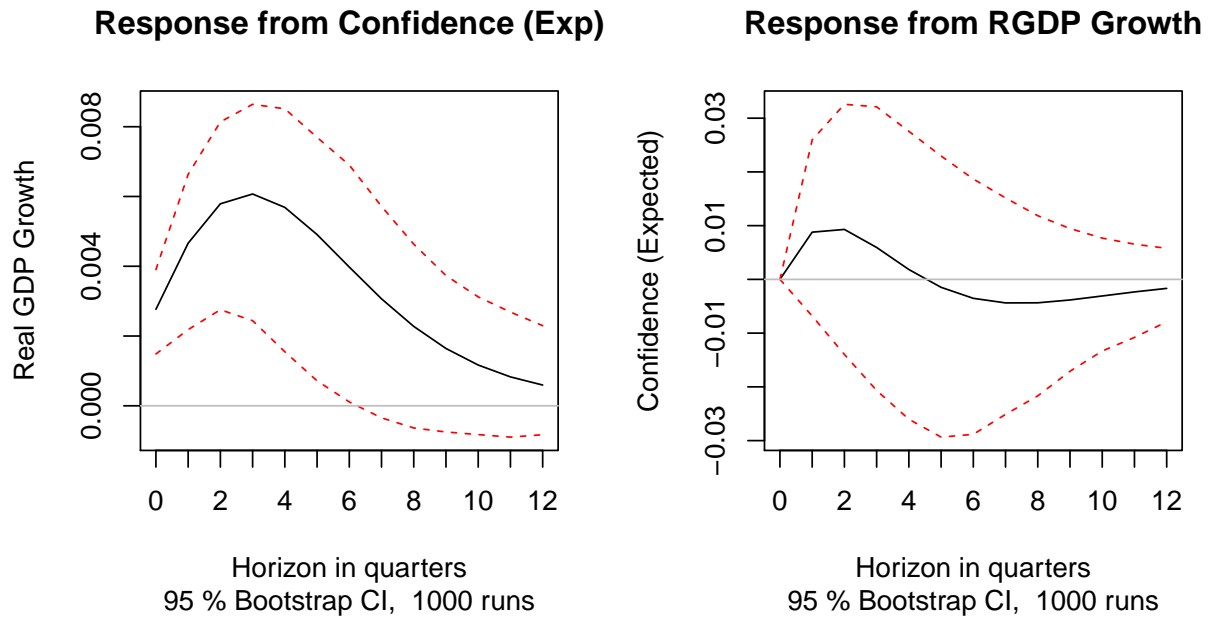


Figure 3.43: IRFs of confidence (expected conditions) and real GDP growth

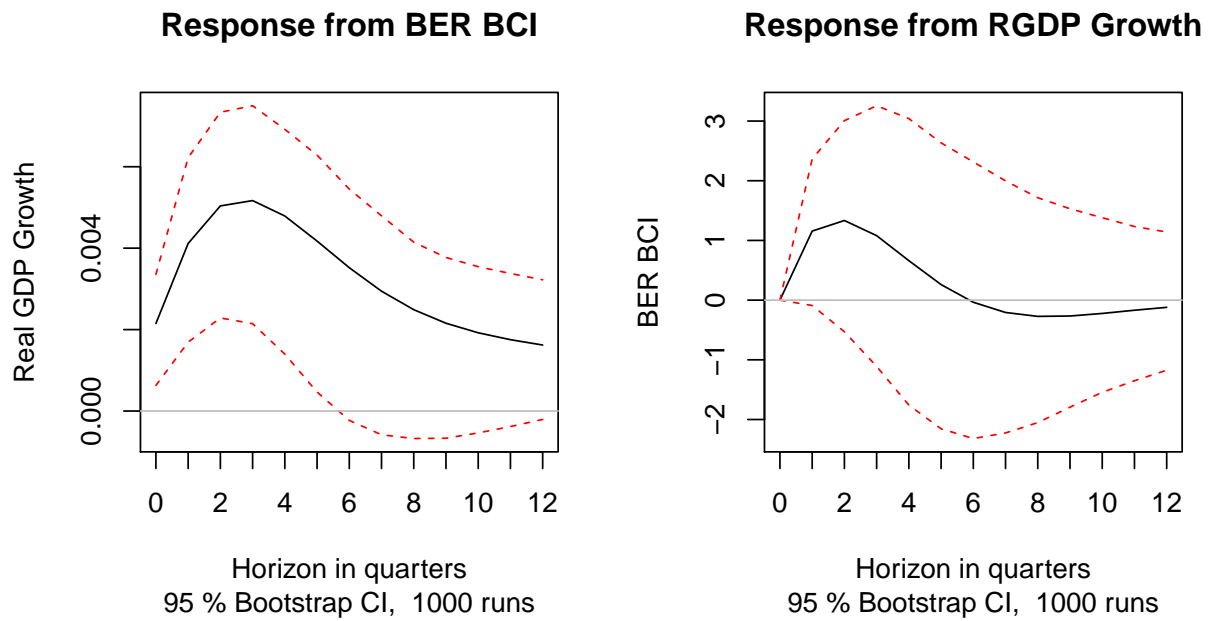


Figure 3.44: IRFs of BER BCI and real GDP growth

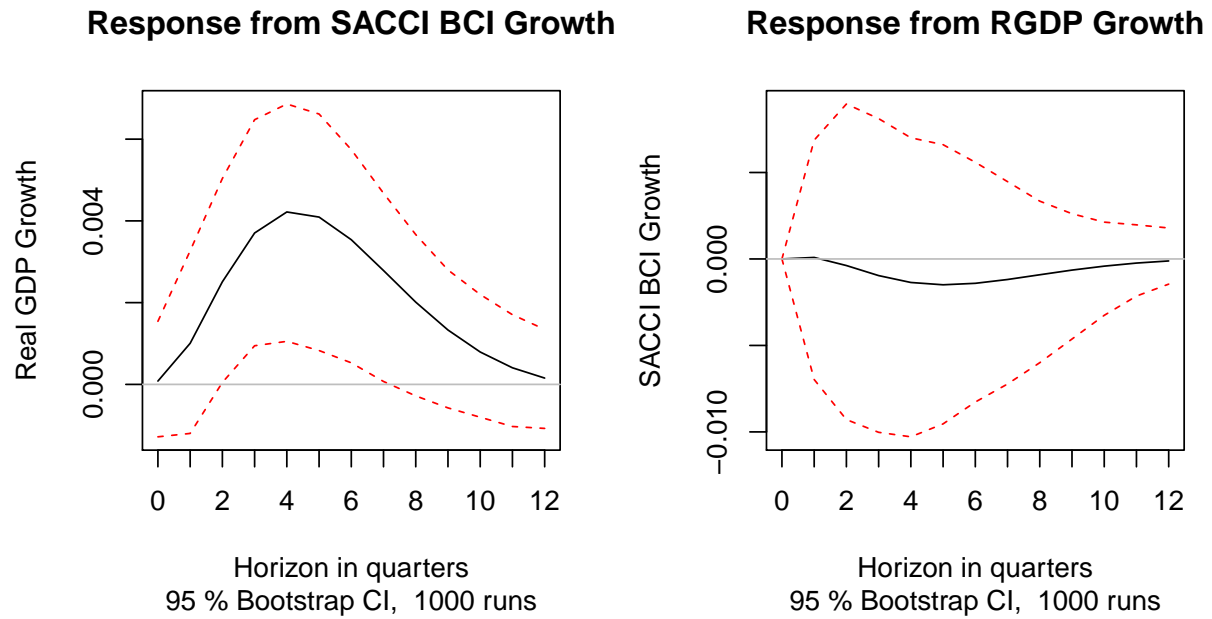


Figure 3.45: IRFs of SACCI BCI growth and real GDP growth

### 3.10.5 Sentiment Indices

## Business Sentiment

Table 3.19: Sentiment indicators

Date	Confidence (Current)	Confidence (Expected)	Dispersion	Idiosyncratic error	Aggregate error	Uncertainty (Combined)
1992-03-31	-0.33	-0.15	2.92	0.13	0.65	2.00
1992-06-30	-0.42	-0.25	3.26	-1.35	3.39	3.12
1992-09-30	-0.50	-0.37	0.94	-0.41	0.40	0.35
1992-12-31	-0.48	-0.26	0.16	-0.04	0.06	1.15
1993-03-31	-0.37	-0.17	1.13	1.93	-0.21	0.60
1993-06-30	-0.37	-0.21	0.76	1.53	-0.32	-0.34
1993-09-30	-0.29	-0.15	0.36	1.95	-0.59	0.81
1993-12-31	-0.20	-0.09	0.10	0.36	-0.49	0.13
1994-03-31	-0.07	-0.03	1.35	3.70	0.80	1.08
1994-06-30	0.18	0.43	-1.36	0.39	0.65	-0.45
1994-09-30	0.17	0.26	1.49	-0.10	-0.39	1.78
1994-12-31	0.31	0.37	0.63	-0.25	-1.05	-0.60
1995-03-31	0.31	0.39	-1.50	-0.68	-0.92	-1.44
1995-06-30	0.25	0.26	-0.75	0.11	1.67	0.01
1995-09-30	0.12	0.16	0.24	2.94	-0.08	-1.11
1995-12-31	0.09	0.12	0.18	0.84	-0.34	-1.22
1996-03-31	0.00	0.13	-0.22	0.36	0.64	-0.66
1996-06-30	-0.10	-0.07	0.94	1.14	-0.44	1.38
1996-09-30	-0.14	-0.09	0.00	-0.58	-0.89	-1.39
1996-12-31	-0.13	-0.13	0.80	-1.44	0.25	-0.01
1997-03-31	-0.03	0.03	-0.74	0.01	0.40	-0.97
1997-06-30	-0.16	-0.12	0.44	0.22	-0.87	-1.37
1997-09-30	-0.19	-0.16	1.02	-0.90	0.22	-0.35
1997-12-31	-0.31	-0.19	1.20	-1.19	0.99	0.56
1998-03-31	-0.40	-0.29	1.86	-0.57	0.93	1.25
1998-06-30	-0.46	-0.37	2.80	-0.23	0.83	2.49
1998-09-30	-0.49	-0.45	1.13	-0.79	-0.50	1.22
1998-12-31	-0.54	-0.41	0.24	-0.85	-0.65	-0.71
1999-03-31	-0.57	-0.41	1.74	-2.84	-0.84	0.24
1999-06-30	-0.51	-0.31	0.23	-0.38	-0.76	-0.77
1999-09-30	-0.30	-0.09	0.30	0.89	-0.23	-1.02
1999-12-31	-0.12	0.12	-0.16	0.57	-0.04	-0.46
2000-03-31	-0.08	0.14	-0.28	0.94	1.93	-0.02
2000-06-30	-0.14	0.01	-0.32	0.85	0.02	-0.81
2000-09-30	-0.14	0.00	0.73	0.02	-0.14	-0.64
2000-12-31	-0.27	-0.12	-1.02	0.99	-0.96	-2.19
2001-03-31	-0.21	-0.04	1.15	0.15	0.11	0.01
2001-06-30	-0.17	-0.08	0.47	0.70	-0.44	-0.40
2001-09-30	-0.17	-0.01	0.17	-0.89	0.23	-0.25
2001-12-31	-0.16	-0.08	-1.44	1.07	0.82	-0.16
2002-03-31	0.01	-0.00	-1.01	1.54	-1.26	-1.58
2002-06-30	0.14	0.11	0.01	-0.84	-0.67	-0.80
2002-09-30	0.17	0.14	-0.42	-1.40	-0.58	-0.38
2002-12-31	0.11	0.06	-1.62	1.22	-0.86	-1.98
2003-03-31	0.07	0.09	-1.15	1.03	2.06	0.85
2003-06-30	-0.08	-0.04	-0.93	0.52	-0.62	-1.13
2003-09-30	-0.11	0.03	-0.71	0.06	0.02	-0.98
2003-12-31	-0.06	0.00	-1.04	0.35	-0.37	-1.69

## Business Sentiment

Date	Confidence (Current)	Confidence (Expected)	Dispersion	Idiosyncratic_error	Aggregate_error	Uncertainty (Combined)
2004-03-31	0.01	0.00	-1.40	0.97	-1.10	-2.11
2004-06-30	0.08	0.10	-0.81	0.63	-0.41	-1.29
2004-09-30	0.15	0.18	-0.14	-0.64	-0.19	-0.39
2004-12-31	0.24	0.18	-1.16	-0.02	-0.16	-1.43
2005-03-31	0.11	0.09	-1.37	-0.57	-1.37	-2.01
2005-06-30	0.19	0.21	-0.38	0.35	-0.68	-1.09
2005-09-30	0.22	0.23	-0.78	0.20	-0.40	-0.55
2005-12-31	0.19	0.25	-0.30	-0.88	-0.51	-0.48
2006-03-31	0.20	0.22	-0.72	-1.06	-1.37	-1.16
2006-06-30	0.21	0.18	0.15	-0.71	-1.28	-0.22
2006-09-30	0.25	0.22	-0.87	0.61	-0.59	-0.73
2006-12-31	0.22	0.15	0.93	-0.11	-0.39	0.42
2007-03-31	0.23	0.23	-0.34	-0.66	-0.89	-0.38
2007-06-30	0.25	0.21	-2.18	0.92	0.47	-1.65
2007-09-30	0.13	0.07	-0.33	0.50	-0.11	-0.65
2007-12-31	0.16	0.07	-0.12	1.00	3.39	2.76
2008-03-31	-0.06	-0.12	-0.14	-0.28	-1.45	0.77
2008-06-30	-0.11	-0.15	1.02	0.49	0.03	1.24
2008-09-30	-0.15	-0.12	0.12	-0.36	2.30	2.43
2008-12-31	-0.23	-0.28	1.14	0.06	-0.02	2.53
2009-03-31	-0.31	-0.38	1.43	-1.63	-1.17	2.10
2009-06-30	-0.33	-0.29	0.98	-1.47	-0.09	2.76
2009-09-30	-0.31	-0.19	0.08	-0.59	-0.03	1.02
2009-12-31	-0.24	0.03	0.24	-0.93	-0.16	0.34
2010-03-31	-0.08	0.07	-0.16	0.95	3.43	1.54
2010-06-30	-0.11	0.02	-0.32	-0.57	0.82	0.41
2010-09-30	-0.09	0.00	-1.19	-0.23	0.10	0.02
2010-12-31	-0.08	0.06	-0.77	-0.50	0.12	0.07
2011-03-31	-0.01	0.09	0.21	-0.80	-0.53	0.33
2011-06-30	-0.04	0.03	0.57	0.28	2.15	1.69
2011-09-30	-0.14	-0.04	-0.70	0.57	-0.18	0.83
2011-12-31	-0.09	-0.04	-0.85	0.35	-0.28	0.41
2012-03-31	-0.02	0.07	-0.38	0.43	0.28	1.30
2012-06-30	-0.09	-0.02	-1.57	-0.65	0.13	0.07
2012-09-30	-0.08	0.01	-0.89	0.47	0.89	0.50
2012-12-31	-0.17	-0.06	0.58	-0.01	-0.13	1.12
2013-03-31	-0.15	-0.03	-0.02	-1.10	0.29	0.13
2013-06-30	-0.15	0.05	0.07	-0.29	-0.58	-0.66
2013-09-30	-0.10	-0.05	-0.48	0.68	-0.17	-0.81
2013-12-31	-0.12	0.04	0.37	1.30	1.67	0.24
2014-03-31	-0.16	-0.05	0.22	0.49	0.02	0.35
2014-06-30	-0.18	-0.04	-0.75	-0.16	0.51	-0.06
2014-09-30	-0.14	0.05	0.05	-0.58	1.29	0.07
2014-12-31	-0.15	-0.14	0.22	-1.06	-0.92	-0.25
2015-03-31	-0.20	-0.07	-1.39	-1.13	-0.31	-1.11
2015-06-30	-0.17	-0.13	-1.00	-1.79	-0.81	-0.85
2015-09-30	-0.22	-0.18	-0.67	-0.91	-0.30	-0.78
2015-12-31	-0.24	-0.21	0.27	-0.41	-0.68	0.17
2016-03-31	-0.29	-0.27	-0.90	-0.25	-1.31	-0.22
2016-06-30	-0.21	-0.16	0.32	-1.67	-0.87	0.08
2016-09-30	-0.21	-0.07				0.04

## 4 Methods for Aggregating Incomplete Time Series from Various Sources: An Application to Commodity Prices in the Cape Colony, 1889-1914

### 4.1 Introduction

Over the past two decades there has been, what Fourie (2016) calls, a data revolution in African economic history. An increase in access to online resources, data-processing software and computing power has enabled scholars to capture and analyse historical statistics on Africa on a much larger scale than before. One example is the Colonial Blue Books, a large collection of reports containing records of Cape government revenue and expenditure, imports and exports, and other statistics. These data sources can now be used to study the incomes and living standards, education and fiscal systems, and transportation networks of African societies (e.g., De Zwart (2011), Greyling and Verhoef (2015), and Fourie, Grundlingh and Mariotti (2017)).

Historical records, and higher-frequency records in particular, are often incomplete (Klovland, 2014). Even if sources are available on an uninterrupted basis, inevitably there will be cases when a particular data series was discontinued or contains substantial gaps. In such cases it can be challenging to aggregate these records in order to estimate complete time series.

The aim in this chapter is to explore aggregation methods for calculating high-frequency commodity price indices, using incomplete historical records. The indices are based on two newly digitised datasets. The first is an expanded version of the historical dataset used in Boshoff and Fourie (2017). The records consist of monthly prices for various commodities in various towns in the Cape Colony from 1889 to 1914, reported in the Agricultural Journals of the Department of Agriculture. The second dataset consists of annual market prices for various commodities in various towns reported in the Cape Colony Blue Books. The aggregation challenge involves combining the incomplete datasets from the two sources that are available at different frequencies. The first dataset is available at a monthly frequency, but is incomplete in terms of the coverage of both products and towns, while the second dataset is only available at an annual frequency and is also incomplete in terms of products and towns.

Klovland (2014) showed that the repeat sales method, which is typically used to create indices for unique and infrequently traded items, such as artworks, may be used to deal with incomplete data in this context. This method provides a consistent way to aggregate the price data from various sources and produces indices with substantially fewer gaps than there are in the individual series. In this chapter the repeat sales method is used to aggregate the incomplete price series for various towns from both sources to estimate relatively complete monthly commodity price indices for the Cape Colony from 1889 to 1914.

To date there has been little research on market prices during this period, in part due to a lack of

records on prices for the Cape Colony before 1910. In this chapter, monthly wholesale commodity price indices are estimated for several individual products. These commodity price indices are intended to shed light on the demand and supply factors that influenced product prices. These indices are then aggregated to form a total commodity price index, with weights based on the production values reported in the 1904 census, and on import values reported in the Blue Books. The total commodity price index is intended to provide a clearer picture of the inflation history of the period. The price indices seem to correspond well with the historical narrative on the economic history of the Cape Colony, with a large increase before and during the Second South African War (1899-1902) and a large decrease prior to unification (1910).

As an internal validity test, the repeat sales indices are compared to simple median indices based on the prices reported in the Agricultural Journals. As in the previous chapters, the external validity of the indices is assessed by testing their conformity to available existing measures. In the absence of an existing monthly commodity price index, the indices are compared with the annual consumer price indices calculated by De Zwart (2011). The indices are then evaluated according to whether they conform to the path implied by available monetary aggregates. The total commodity price index seems to better reflect the monetary history of the Cape Colony than existing price indices.

Higher frequency records, such as these, allow for a more detailed investigation of price histories, business cycles, and market behaviour in general (Mitchell, Solomou and Weale, 2012). In addition to providing a clearer picture of demand and supply conditions and the inflation history of the period, the high-frequency price records may be used to investigate market integration in the Cape Colony around the turn of the 20th century. This chapter therefore explores further methods to aggregate the high-frequency price records. The aggregated time-series indicators are used to investigate internal market integration in the Cape Colony, which demonstrates the usefulness of the aggregation methods and the estimated time series.

Market integration concerns the convergence of prices for the same goods throughout the economy (Andrabi and Kuehlwein, 2010), as well as the efficiency with which price gaps return to equilibrium after a shock (Federico and Sharp, 2013). Higher-frequency data can aid the investigation along both dimensions (Ejrnaes, Persson and Rich, 2008). There is a large international literature on market integration, but very few studies have considered internal market integration in the Cape Colony over this period (e.g. Boshoff and Fourie (2017)). During this period the mineral revolution led to the rapid expansion of the railway network and created a potentially large internal market (Schumann, 1938; Herranz-Loncán and Fourie, 2017). In this part of the chapter, the work of Boshoff and Fourie (2017) is extended, by expanding the period and the number of products, and by giving consideration to both dimensions of market integration.

In examining these two dimensions, this section employs two aggregation methods. The first is to calculate measures of cross-sectional price dispersion among the towns in the Cape Colony, which are used to investigate price convergence. The second is to calculate repeat sales commodity price indices at the regional level, which are used to investigate market efficiency. Time-series techniques



are then used to test for evidence of market integration. The hypotheses that price dispersion between towns was declining over the period and that regional price indices were cointegrated are tested. According to these tests, many of the commodity prices among towns were converging over the period. A number of regional commodity price indices were cointegrated, with the number of cointegrated series increasing in the latter part of the sample period. This implies that there was increasing market integration in the Cape Colony over the sample period.

In Chapter 2, various quality-adjusted measures of the mean of the distribution of growth rates in art prices were estimated with the hedonic and hybrid repeat sales methods. The classical repeat sales method used in this chapter is related to those methods, as it also calculates the mean in the distribution of growth rates. The use of the repeat sales method in the context of estimating complete series from incomplete historical records is a novel application of this method, where each commodity price series (for a particular quality description) in each town is treated as a specific item to track over time.

In Chapter 3, various measures of the mean and standard deviation of survey responses and forecast errors were calculated as proxies for business confidence and uncertainty. The cross-sectional measures of price dispersion calculated in this chapter, such as the coefficient of variation, are related to those measures in that they are measures of the first and second cross-sectional moments of the distribution of commodity prices. In this chapter it is demonstrated that these aggregation methods are useful for calculating macroeconomic time-series indicators in an entirely different context.

## **4.2 Economic History in the Cape Colony: 1886-1914**

The turn of the 20th century was a period of substantial structural change in South Africa. The discovery of gold in 1886 transformed the structure of the economy from an essentially pastoral economy to a mining-intensive one. Within a few decades, the predominantly agricultural economy was transformed into a market and credit economy in which capital-intensive mineral production was of significant importance (Schumann, 1938). The Second South African War (1899-1902) also fundamentally altered the region during this period and led to the formation of the Union of South Africa from the four colonies, in 1910. During this period, therefore, the Cape Colony is an interesting example of an early capitalistic economy undergoing structural change, where globalisation and internal market integration was just beginning to take hold (Boshoff and Fourie, 2017).

This section presents a discussion on the economic history of prices in the Cape Colony between the discovery of gold on the Witwatersrand in 1886 and the start of World War One (WWI) in 1914. A brief description of the Second South African War is then provided, as well as an overview of the existing research on prices and inflation in the Cape Colony over this period. The idea is to provide some context for the subsequent investigation of the price history of the Cape Colony.

Table 4.1: Business cycle turning points

Cape Colony		Great Britain	
Trough	Peak	Trough	Peak
1887/12	1890/12	1886/06	1890/09
1892/12	1893/12		
1894/12	1898/12	1895/02	1900/06
1900/12	1902/12	1901/09	1903/06
1904/12	1906/12	1904/11	1907/06

#### 4.2.1 Overview of Economic Activity in the Cape Colony

The seminal works of De Kock (1924) and Schumann (1938) discussed the economic history of the South African colonies in the late 19th and early 20th centuries. These studies provide similar narratives of the economic development of South Africa around that period. Recently, Greyling and Verhoef (2015) calculated GDP estimates for the Cape Colony, to facilitate the analysis of its long-term growth and development. This section provides a brief review of the economic history of the period, by relating the historical narratives to the quantitative evidence from the GDP estimates. The GDP estimates are necessarily imprecise and are intended to illustrate trends for the Cape Colony.

Figure 4.1 illustrates the real GDP estimates for the Cape Colony reported in Greyling and Verhoef (2015), with recessionary periods shaded. The turning points of the indices are determined with the non-parametric dating algorithm suggested by Bry and Boschan (1971). This method automates the Burns and Mitchell approach to determining classical turning points. The algorithm identifies local minima (troughs) and maxima (peaks) in a logged time series. A censoring rule based on a minimum of one year for each phase and two years for a full cycle is used to ensure that phases and cycles have a minimum duration, as suggested by Harding and Pagan (2002). The algorithm has the advantage that it is simple, transparent, easily replicable, and provides intuitively appealing results (Du Plessis, 2006). Following Du Plessis (2006), the algorithm is applied to the log of real GDP, after subtracting a simple deterministic linear trend. The resulting turning points differ slightly from those derived by Greyling and Verhoef (2015) using the Baxter-King band-pass filter.

Taken as a whole, the Cape Colony experienced relatively rapid expansion over this so-called ‘gold-mining’ period, during which the structure of the economy was transformed from an agricultural economy to an agricultural-mineral one (Schumann, 1938). According to these estimates, the Cape Colony experienced six upswing and six downswing phases between 1886 and 1909. Despite the imprecise GDP estimates for the Cape Colony, Table 4.1 reports that these cycles corresponded quite closely to the business cycles in Britain, as reported in Moore and Zarnowitz (1984).

The discovery of gold was followed by an upswing phase, lasting roughly from 1888 to the end of 1890. During this period, there were large investments in the gold mining industry, a substantial proportion of which was directed through the Cape Colony banks. This prosperity was accompanied by speculation in gold shares and property, and easy credit extension by the banks (Greyling and

## Historical Commodity Prices

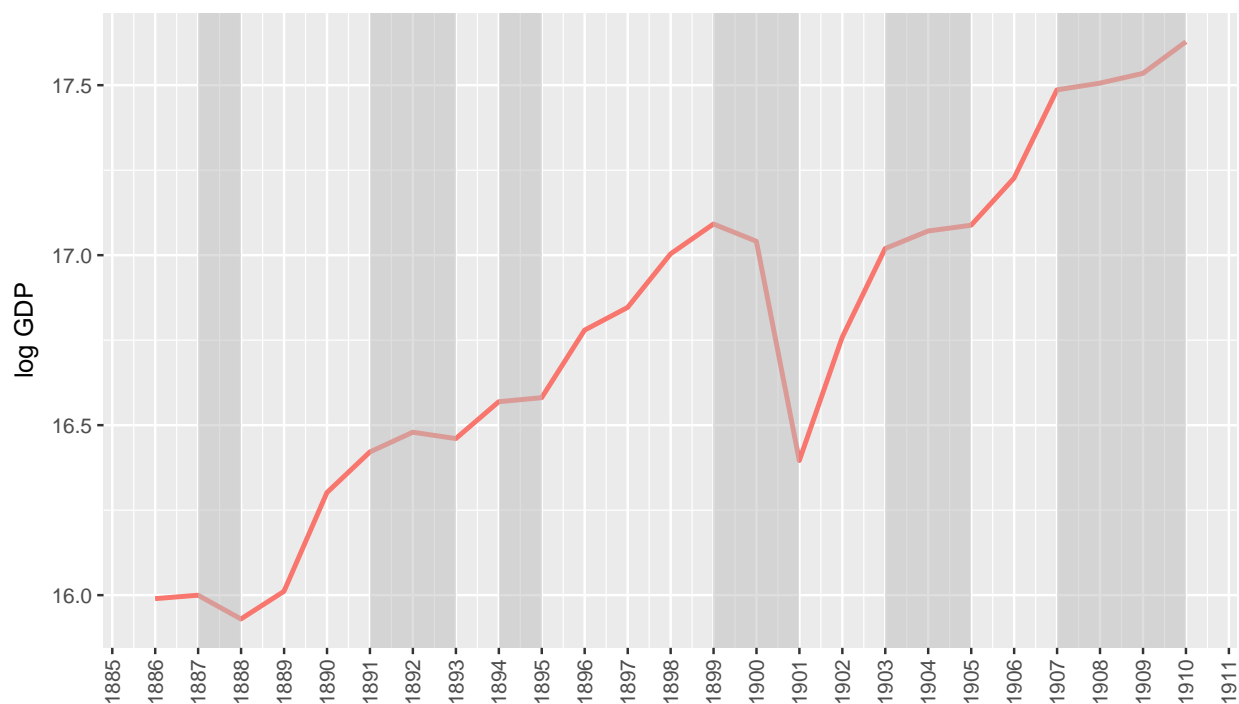


Figure 4.1: Real GDP in the Cape Colony (1886-1909)

Verhoef, 2015).

This upswing phase was followed by a depression, as a result of the financial and banking crisis in 1889-1890. This crisis was due to the overextension of credit and massive speculation in property and gold shares (Schumann, 1938). According to Gilbert (1933), the speculative bubble was burst by the realisation that many mines located on outside reefs would not be profitable and that deep-level mining would have to be attempted by the mines on the reef. Most of the District Banks in the Cape got into financial difficulties and were absorbed by the larger banking institutions (De Kock, 1924). The price of gold shares, land and prospecting firms plummeted, while railway traffic, government revenue and imports declined marginally (Greyling and Verhoef, 2015). Schumann (1938) reported that the banks of the Cape were more severely affected than the economy as a whole. Agriculture in the Cape Colony did not suffer immediately from the recession, as the harvest prospects for 1890 were good. A relatively stagnant period followed, with a short upswing phase in 1893.

There was another speculative boom in 1895, with large investments in the gold mining industry from abroad. According to Schumann (1938), this was followed by a relatively mild recession in 1896-97, although the GDP figures do not indicate a turning point. Political and natural factors contributed to the stagnant conditions in 1896 (Gilbert, 1933). The Jameson Raid at the start of 1896 had an unsettling effect on business conditions. Tensions between the Boer Republics and Britain led to the closure of the drifts in September and October 1895, disrupting road transportation of goods into the interior and outward to the coastal ports. A severe drought in 1896 and the rinderpest cattle

disease caused widespread livestock losses, especially in the interior (De Kock, 1924; Schumann, 1938).

Economic growth had just started to accelerate when the outbreak of the Second South African War (1899-1902) caused a recession (De Kock, 1924). According to the GDP estimates, the Cape Colony experienced a short but severe recession, at least during this initial shock.<sup>29</sup> The outbreak of the war paralysed trade between the Cape Colony and the two Boer Republics and hampered trade with the northern districts of the Cape Colony. Gold mining in the Transvaal came to an almost complete standstill until 1902, with the value of gold production declining from £16m in 1898 to £1m in 1901 (Schumann, 1938). The Cape Colony therefore lost what had become two important markets for their produce (driven by railway integration, as discussed below), as well as a major export commodity through its ports. The war had a particularly devastating effect on the economies of the Boer Republics. Agricultural prospects were severely hampered following the ‘scorched earth’ policy of the British forces, which caused the destruction of farm buildings, crops and livestock in the Transvaal and the Orange Free State, and to a lesser extent, in parts of the Cape Colony and Natal.

According to the GDP estimates, the short recession was followed by a relatively swift recovery in the Cape Colony. The large increase in expenditure by the British Government in connection with military operations created prosperity in parts of the Cape Colony and Natal, especially among the farming and trading communities. The war cost Great Britain around £250 million, a large proportion of which was spent in the Cape Colony and Natal to purchase agricultural and pastoral produce for the troops. Within a few months of the outbreak of the war, the number of British soldiers in South Africa increased to 130,000, and subsequently to around 250,000 (nearly 25% of the white population of the four territories combined) (Schumann, 1938).

Figure 4.2 illustrates the imports and exports over the period. The Report on the Customs Transactions of the Colony for 1900 reported that the marked increase in imports into the Cape Colony could be attributed to the large number of immigrant residents, expenditure on troops en route, and the restricted production of foodstuffs in the districts affected by the war. The increase in imports from 1899 to 1903 was due in large part to the increased imports of military supplies for the British troops (Gilbert, 1933). Expenditures by the British military therefore provided an injection for consumption and production in the Cape Colony and Natal (Greyling and Verhoef, 2015).

At the same time, exports from the Cape and Natal ports declined substantially, from almost £25 million in 1898 to £7.5 million in 1900, due to the increased local demand for produce and the temporary cessation of gold mining. The decline in exports did not seriously affect purchasing power abroad, as the large remittances of funds by the British Government from London to South Africa served to more than offset the reduction in value of the exports. The increase in imports

---

<sup>29</sup>While the size of the decrease may seem implausibly large (almost 50%), the estimates are only intended to be indicative of trends over the period.

## Historical Commodity Prices

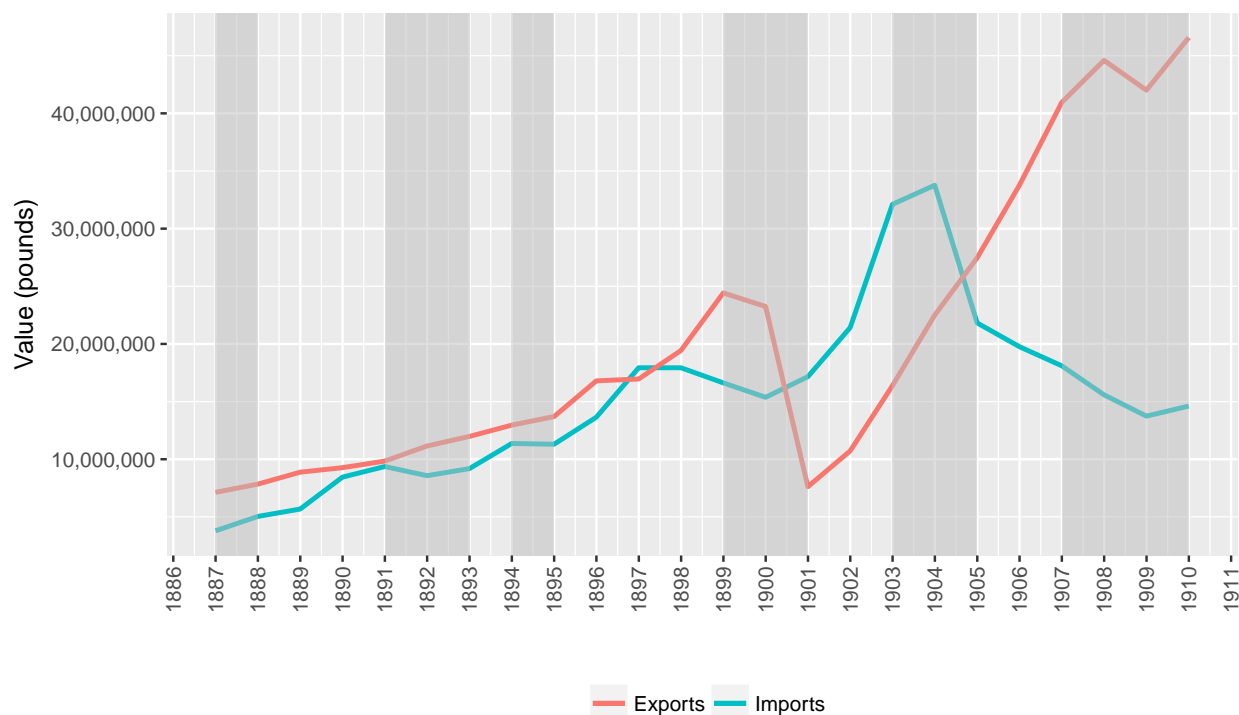


Figure 4.2: Total imports and exports in the Cape Colony (1886-1909)

and railway revenues from the transportation of troops and war materials also enhanced the public revenues (Schumann, 1938). The war demand for goods and services therefore stimulated increases in prices and production in the Cape Colony (De Kock, 1924). The large negative shock initially therefore seems to have been short-lived. According to the Cape Colony GDP estimates, the trough had already been reached by the end of 1900.

The short recession was followed by a relatively swift recovery. According to Schumann (1938), the post-war boom was the result mainly of a general feeling of optimism after the conclusion of the war, both in England and in South Africa. This optimism resulted in the extension of bank credit and speculation in property, as well as an increase in the importation of capital and goods to aid reconstruction. The reconstruction schemes, including a development loan of £3 million granted by the British Government, stimulated economic activity and a speculative environment (De Kock, 1924). Immigration also increased significantly after the war, by those presumably in search of mineral wealth (Gilbert, 1933).

The swift recovery was short-lived. According to the GDP estimates, there was a mild recession from 1903 to the end of 1904. The recession was caused by a cyclical reaction to the large economic disequilibrium (Schumann, 1938). During the war, production and trade had increased due to the greater demand for foodstuffs and other commodities. After the war, production and trade increased further, based on the expected expansion and prosperity of the gold and other industries. Despite the relatively early resumption of gold production, progress was hampered by a lack of unskilled

labour. The inflow of purchasing power due to the war ended, and large stocks of military goods were sold. When the expectations of growth were not fully realised, a recession followed (De Kock, 1924). The recession may also in part be explained by the extensive destruction of the interior during the war (Schumann, 1938).

After a brief upswing phase from 1905 to the end of 1906, another recession set in around 1907. The financial panic in Europe and America in 1907 and its effects on South African trade, as well as the drought of 1908 may have contributed to the recession (De Kock, 1924). The recession seems to have lasted up to 1910, the year of the unification of the four colonies, which was in line with reviving conditions in England (Schumann, 1938).

The subsequent recovery period seems to have taken place up to the outbreak of WWI in 1914 (the end of the sample period). De Kock (1924) referred to this as the period of economic reorganisation, with an expansion in the agricultural, mining and manufacturing industries. The benefits of the construction of branch railway lines in many parts of the country promoted the development of farming enterprises. The gold industry was growing rapidly, and the diamond industry had recovered quickly and resumed its growth. Although this was not a period of exceptional prosperity, conditions were generally fairly stable (De Kock, 1924).

Overall, the turn of the 20th century was a turbulent period in South African history. Globalisation and the discovery of minerals transformed the Cape Colony into a prosperous exporting region with a steadily increasing GDP (Herranz-Loncán and Fourie, 2017). The Second South African War (1899-1902) was a particularly important event during the sample period and involved large demand and supply shocks, which one would expect to have had an impact on prices. The following section provides a brief description of the war period.

#### **4.2.2 The Second South African War (1899-1902)**

The Second South African War (1899-1902), or Anglo-Boer War, was fought between Britain, with its two colonies (the Cape Colony and Natal), and the two independent Boer Republics (the South African Republic, or Transvaal, and the Orange Free State). It proved to be the longest (almost 3 years), costliest (over £200 million), bloodiest (estimates vary between 59,000 and 68,000 lives), and most humiliating colonial war for Britain between 1815 and 1914 (Packenham, 1979). Marks (2011) argues that the war was as important in shaping modern South Africa as the American Civil War was for the United States.

After the discovery of gold on the Witwatersrand, immigrants (labelled *Uitlanders*) poured into the Transvaal. The tensions around the franchise of the *Uitlanders*, many of whom were of British origin, became the main source of conflict between the Transvaal and British Governments, and ostensibly became the cause of the war. However, there were also wider considerations relating to the powerful financial interests concerning the control of the gold mines (Marks, 2011). Factors such as *Uitlander*

---

## Historical Commodity Prices

---

grievances, tariffs, and monopolies were directly connected with gold mining interests. Any attempt by the British Government to gain concessions from the Transvaal Government made it the agent of the mining magnates (De Kiewiet, 1941). Moreover, Britain was pursuing its lengthy quest for a united South Africa (De Kiewiet, 1941) and its need to assert itself in the face of economic competition from America and Germany (Marks, 2011).

War was declared by the Boer Republics on 11 October 1899, after the expiry of President Kruger's ultimatum to the British government to withdraw all its troops from the border of the Transvaal. Kruger had hoped for a short war, like the clash with Britain in 1881, when the British quickly gave up the fight. This time round a long, drawn-out war ensued, with numerous unforeseen consequences (Fourie, Grundlingh and Mariotti, 2017).

Within a year the British captured the capitals of the two Republics, Bloemfontein (on 13 March 1900) and Pretoria (on 5 June 1900), without meeting much resistance (Fourie, Inwood and Mariotti, 2014). Contrary to expectations, however, this did not put an end to the war. There were too many areas still not under British control, and the line of communication through the Orange Free State was poorly guarded and vulnerable (Pretorius, 1998). Hostilities lasted for another two years. Giving up on the impossible objective of trying to halt the British advance, Boer commando forces switched to guerrilla tactics (Pretorius, 1998). Moving in mobile commandos, Boer soldiers launched running attacks on British columns and supply lines (Marks, 2011) and strategically intercepted British outposts over the Highveld terrain (Fourie, Inwood and Mariotti, 2014).

The British responded with a threefold strategy to end the war, involving block houses, intensifying the scorched earth policy, and the use of concentration camps. Block houses were erected to protect strategic points such as bridges and railways, and as barriers to limit the free movement of the Boers across the countryside. The scorched earth policy involved the destruction of homesteads, burning of foodstuffs and crops, and the slaughtering of livestock. Boer women and children were then sent to concentration camps (Fourie, Inwood and Mariotti, 2014). These soon became notorious, as insanitary conditions, poor rations and overcrowding led to major epidemics and high mortality rates. Thousands of displaced black South Africans, too, were rounded up and put into segregated concentration camps, many of which also served as labour depots for the British army (Marks, 2011).

By 1902 the Republics had been ground down by the large resources of the British army, the destruction of farms, the high mortality rates among women and children in the concentration camps, and the increasingly tight British control of the countryside through blockhouses and barbed wire (Marks, 2011). The Boer leadership was forced to sue for peace, and on 31 May 1902 the Peace Treaty of Vereeniging was signed (Pretorius, 1998). At the signing of the peace, the British Government promised £3 million towards rehabilitating the Transvaal and the Orange Free State.

Estimates of the mortality rates vary, but it is estimated that at least 22,000 British and 7,000 Boer soldiers died. More than 27,000 Boer women and children died in the concentration camps (out of



## Historical Commodity Prices

Table 4.2: Census commodity production numbers for the Cape Colony

Year	Wheat (200lbs)	Maize (200lbs)	Cattle	Woollen Sheep	Horses
1891	909,000	965,000	2,211,000	13,631,000	444,000
1904	567,000	1,131,000	1,954,000	8,465,000	255,000
1911	1,305,000	1,728,000	2,715,000	13,239,000	334,000

Source: *De Kock (1924)*

the 118,000 inmates). An estimated 12,000 black South Africans, who fought in large numbers on both sides, lost their lives, although these estimates are less precise (Fourie, Inwood and Mariotti, 2014). Recent estimates suggest that African camp mortality rates may have been even higher than that of the Boer camps (Marks, 2011).

For the purposes of movements in commodity prices, two factors were especially important over this period. From the demand side, the war involved an influx of British troops. In 1897, there were at most 10,000 British troops within the Cape Colony (Evans, 2000). By 1900, the British had shipped in massive reinforcements for a counteroffensive against the Republics. By the end of the war, almost 450,000 British regular and colonial forces had been involved in the war (although all were not necessarily present at any given time), compared with an estimated 88,000 Boer and volunteer forces (Fourie, Inwood and Mariotti, 2014). The size of the invading army was as large as the combined white citizenry of the Republics (Marks, 2011).

From the supply side, the scorched earth policy led to the systematic devastation of the Republics (Pretorius, 1998). The policy started as specific reprisals for Boer attacks, whereby farms in the vicinity of attacks would be burnt. In September 1900 the policy was intensified, to involve the destruction of all farms and food supplies within 16 km of an incident, or if a farm had been used as a commando base. In December 1900 the land clearance policy was intensified again to deny support to Boers in the field. Over the course of the war, some 30,000 houses, including farms, were burned down or extensively damaged (Evans, 2000). The policy involved the destruction of all food supplies, with fields of grain and maize burned and livestock killed in enormous numbers. The devastation in the two Republics led to subsequent food shortages (Pretorius, 1998).

Table 4.2 reports the 1891, 1904 and 1911 census production numbers for five commodities in the Cape Colony. Although the scorched earth policy caused the destruction of crops and livestock mostly in the interior of the country, its effects were still visible in the census numbers for the Cape Colony in 1904, two years after the war. De Kock (1924) showed that the substantial decreases in livestock were visible in the official statistics for all four of the regions during the war. Cattle herds were depleted to such an extent that for years, large quantities of frozen meat had to be imported to satisfy the local demand. The war depleted more than half of the stock of horses in South Africa and reduced the number of woollen sheep significantly, causing a great setback in the wool industry (Fourie, Inwood and Mariotti, 2014).

Moreover, trade between the territories had all but collapsed with the outbreak of the war. According to the Report on the Customs Transactions of the Colony for 1899 (in the Blue Books), the war



paralysed trade between the Cape Colony and the two Boer Republics and was seriously interfering with trade with the northern districts of the Cape Colony. Indeed, all trade was suspended between the Cape Colony and the Boer Republics on the 12th October 1899, and a Proclamation (No. 277) was issued, prohibiting such trade.

One would expect these developments to have led to an increase in agricultural commodity prices around the war period. The two new datasets provide a way of investigating whether these developments in demand and supply were reflected in the commodity prices. The following section provides a brief overview of the existing research on prices and inflation in the Cape Colony over this period.

### 4.2.3 Overview of Existing Research on Prices and Inflation

Price indices are useful in studying inflation, business cycles, and market behaviour in general. The first generation of price indices, such as Jevons (1865) and Giffen (1879), focused on the effect of the gold supply on the general price level. Gilbert (1933) followed this tradition in his analysis of the economic effects of gold discoveries on South Africa, using the annual weighted aggregated index of the prices of 13 foods for 1883-1907, reported below. Another strand of the literature, which includes Persons and Coyle (1921) and Silberling (1923), focused on the measurement of business cycles. The seminal work of Schumann (1938) on business cycles in South Africa follows this tradition. He used a quarterly wholesale price index from 1910-1936 in his analysis.

No wholesale price index exists for the Cape Colony before 1910, although a few studies have calculated consumer price indices (CPIs). Gilbert (1933) computed a weighted aggregate index of the retail prices of 13 foods in nine towns in the Cape Colony for the years 1883 to 1907. The indices are available only graphically, and are reproduced in Figure 4.3. According to the index, prices in the Cape Colony increased significantly during the two speculative booms in 1889 and 1895, mentioned above, and declined slightly during the recessionary or stagnant periods that followed. There was a rapid increase in prices during the war period and the post-war recovery. This was followed by a rapid decline that lasted at least up to 1908. Similarly, Schumann (1938) reported that retail prices increased markedly during the war, with a peak in 1902, before decreasing up to 1909.

Figure 4.4 illustrates the annual CPI indices for the Cape Colony calculated by De Zwart (2011), in order to estimate the cost of living in the Cape Colony in the 19th century. De Zwart (2011) calculated CPI based on a ‘bare bones’ and a ‘respectable’ basket of consumption goods. This involved creating consumption baskets that reflected the consumption patterns of workers, as well as price series for the products in those baskets. The indices included 8 and 10 products<sup>30</sup> respectively, based on the annual ‘average market prices of consumables’, ‘average prices of agricultural produce’

---

<sup>30</sup>These were beef, bread or wheat, butter, candles, cheese, coal, cotton, peas and beans, fish, soap and wine. The bare bones consumption basket reflected the cost of bare subsistence, while the respectable consumption basket included higher quantities of the products, as well as beef, cheese, and wine instead of fish.

Historical Commodity Prices

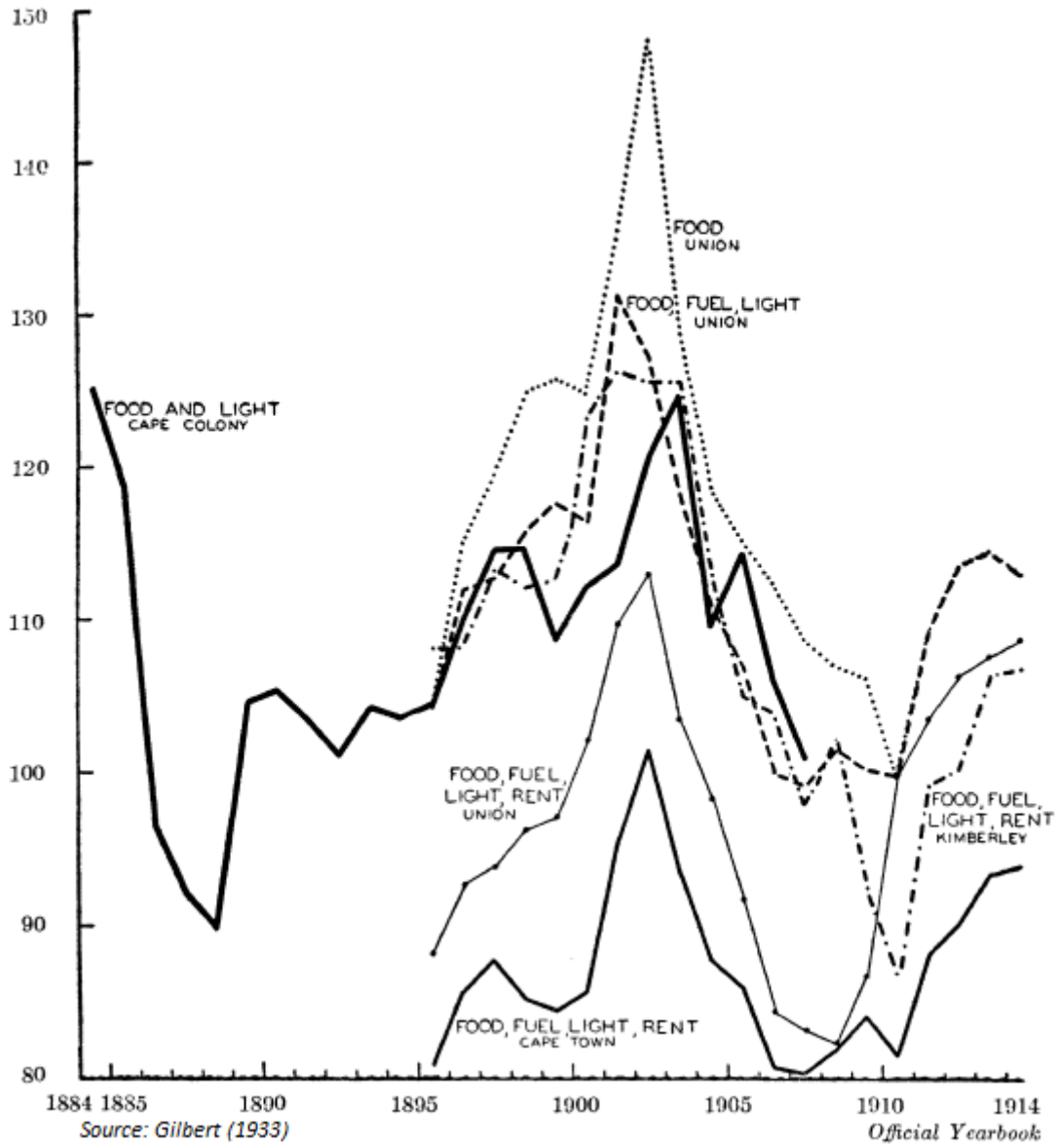


Figure 4.3: Cape Colony domestic retail price indices (1910=100)

## Historical Commodity Prices

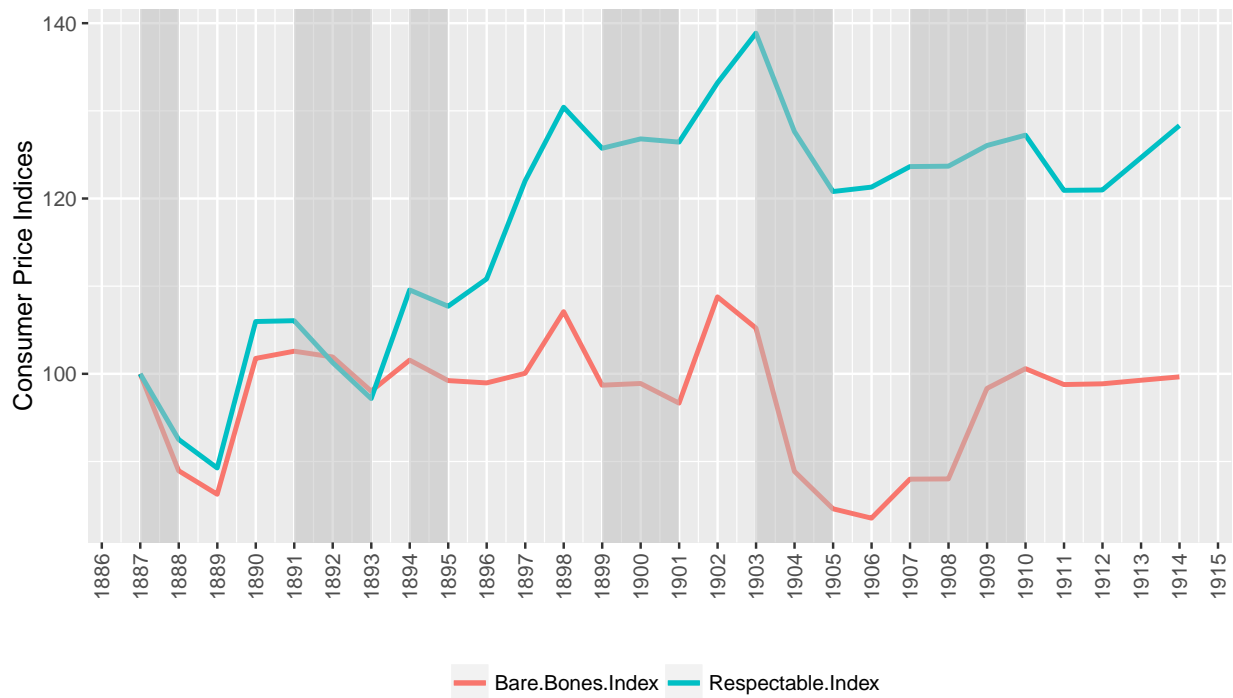


Figure 4.4: Consumer price indices for the Cape Colony

and import and export prices, as reported in the Blue Books. Because of missing price points, it was necessary to interpolate and extrapolate some of the series.

The CPIs illustrate two of the main challenges with the historical statistics. First, they are often available only at an annual frequency, which means that high-frequency events cannot be analysed. Second, because the records are often incomplete, many series have to be interpolated and extrapolated to form continuous series.

### 4.3 Data: Commodity Prices

This chapter aims to demonstrate aggregation methods for calculating a high-frequency (i.e. monthly) commodity price indices from 1889 to 1914, using incomplete historical records for various towns from two sources. The commodity price indices are based on market prices<sup>31</sup> from two newly digitised datasets. This section presents these two datasets and then investigates the pricing dynamics in the monthly price records.

<sup>31</sup>All prices were reported in pounds sterling (£), shillings (s) and pennies (d) and converted to pennies. There were 20 shillings (s) in a pound (£), and 12 pennies (d) in a shilling (s).

### 4.3.1 Commodity Prices in the Cape Colony

The first dataset is an expanded version of the monthly prices used in Boshoff and Fourie (2017). It reports the monthly ‘Current Market Rates (Wholesale) of Agricultural Produce’, as telegraphed by the Civil Commissioners and reported in the Agricultural Journals of the Department of Agriculture of the Cape Colony. The Journals are available from September 1889 to July 1914, which dictates the period that the indices cover. Monthly data are available for 24 commodities (e.g. wheat, eggs and tobacco)<sup>32</sup> in 19 towns across the Cape Colony (e.g. Cape Town, Kimberley and Port Elizabeth).<sup>33</sup>

Figure 4.5 illustrates the number of monthly observations by commodity. The sample suffers from substantial missing data points. The Journal volumes do not record prices for all of the towns or for all of the products over the sample period, and for some months, no prices were recorded. The increase in observations around the turn of the century is because prices were recorded weekly over that period.

Figure 4.6 illustrates the monthly prices for one of the commodities, wheat, by town. Wheat prices varied widely over the period, especially around the turn of the century. In this chapter, these time series form the basis of the monthly index of wheat prices for the Cape Colony.

The second dataset reports annual market prices in the Colonial Blue Books of the Cape Colony, available from 1889 to 1907. Annual market prices in the Blue Books were collected by the colonial administrators in South Africa and sent to the Colonial Office in London (De Zwart, 2011). This information on ‘Average Market Prices of Agricultural Produce’,<sup>34</sup> ‘Provisions’,<sup>35</sup> and ‘Stock and Animal Productions’<sup>36</sup> was used to create a database that includes prices of 50 agricultural products in 48 towns.<sup>37</sup> The annual data is also incomplete in terms of the coverage of products and towns. The annual dataset does include a number of important products and towns that are omitted from the monthly dataset. Wool is one example of an important export product in the Cape Colony,

<sup>32</sup>The 24 products are wheat, wheat flour, boer meal, mealies, mealie meal, barley, oats, oathay, lucerne hay, potatoes, tobacco (boer roll), beef, mutton, fresh butter, eggs, cattle (slaughter), sheep (slaughter), pigs (slaughter), bread, oranges, saddle-horses, transport oxen, milk cows, woollen sheep.

<sup>33</sup>The 19 Cape Colony towns are: Aliwal North, Beaufort West, Burghersdorp, Cape Town, Clanwilliam, Cradock, Dordrecht, East London, Graaff-Reinet, Grahamstown, Kimberley, King William’s Town, Malmesbury, Mossel Bay, Port Alfred, Port Elizabeth, Queenstown, Tarkastad, and Worcester. The data also includes a few prices for towns in the other territories (e.g. Johannesburg) in the final few months of the sample. Unfortunately, these are too few to include in the indices.

<sup>34</sup>Wheat, barley, rye, oats, mealies, peas and beans, potatoes, pumpkins, aloes, argol, wine and brandy.

<sup>35</sup>Oatmeal, flour, bread, mutton, beef, pork, bacon, butter, cheese, tea, coffee, sugar, rice, tobacco, dried fruit, salt, wine, brandy, beer, milk, candles, and lamp oil.

<sup>36</sup>Saddle horses, draught mules, asses, draught oxen, milk cows, woolled sheep, Cape sheep, swine, goats, fowls, ducks, washed wool, fat and tallow, soap, hides, sheep skins, and goat skins.

<sup>37</sup>The 48 Cape Colony towns are: Albany, Albert, Aliwal North, Beaufort West, Bredasdorp, Caledon, Cape Town, Ceres, Clanwilliam, Colesberg, Cradock, East London, Fort Beaufort, George, Glen Grey, Graaff-Reinet, Humansdorp, Kimberley, King William’s Town, Knysna, Ladismith, Malmesbury, Middelburg, Mossel Bay, Oudtshoorn, Paarl, Philipstown, Piquetberg, Port Elizabeth, Prince Albert, Queenstown, Richmond, Riversdale, Robertson, Somerset East, Stellenbosch, Swellendam, Tulbagh, Uitenhage, Uniondale, Willowmore, Worcester, Vanrhynsdorp, Wynberg, Mount Currie, Kokstad, Umtata.

Historical Commodity Prices

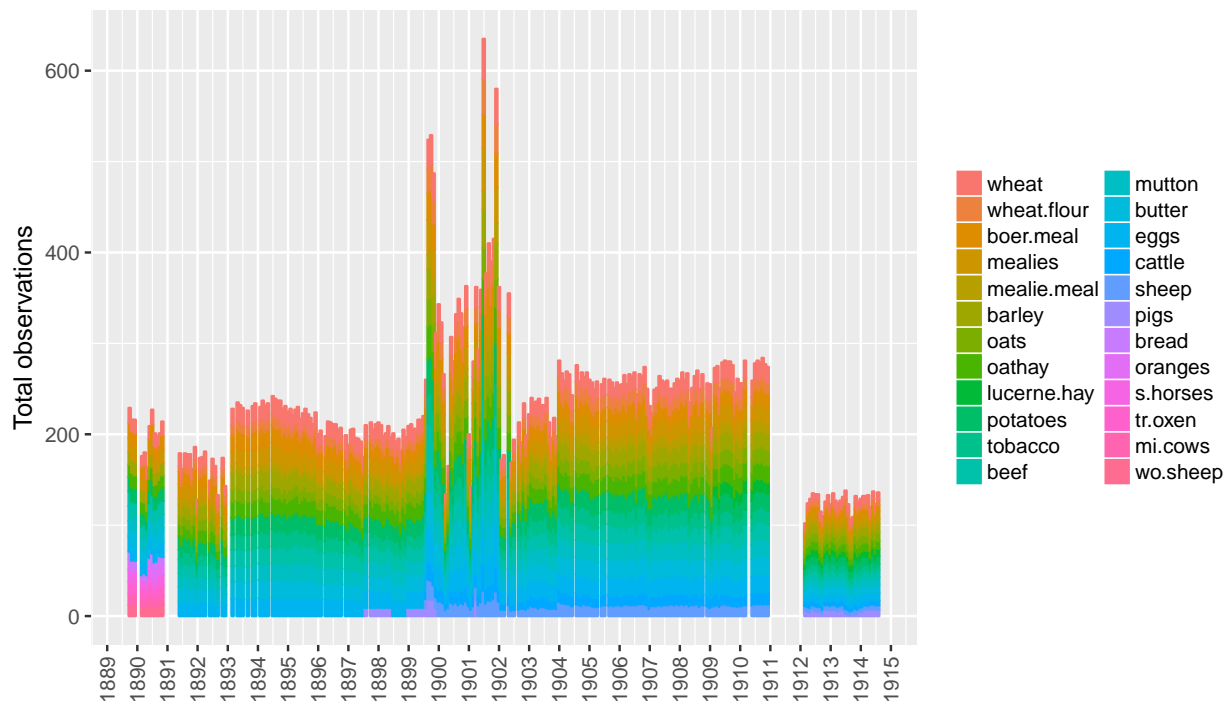


Figure 4.5: Total number of monthly observations by commodity

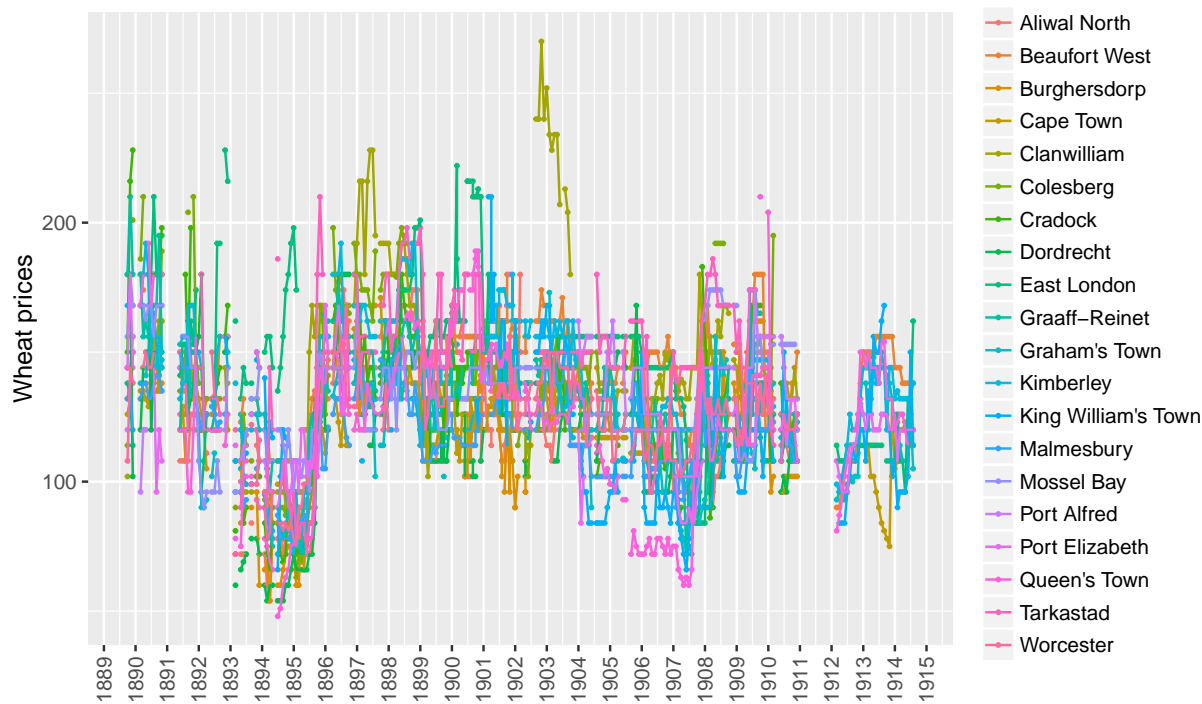


Figure 4.6: Monthly wheat prices by town in the Agricultural Journals

which is not present in the monthly records.

In some cases the two datasets provide prices for the same products in the same towns, although they are mostly reported in different units. For instance, in the monthly dataset, wheat prices are reported in pounds (lbs), whereas in the annual data, the wheat prices are reported per bushel. This means that the trends can be compared, even though the levels are different. For the most part, the average prices reported in the two datasets seem to capture a similar trend over time. In a few cases there were multiple quality descriptions for a specific commodity, such as three series for wine (ordinary quality, better quality and wine without a description) and two series for beer (English and Colonial). The challenge is to combine and aggregate these two sets of prices in a consistent way in order to create complete price indices for commodities in the Cape Colony.

### 4.3.2 Pricing Dynamics in the Monthly Records

The monthly price records can be used to gain some insight into the pricing behaviour over this period of the Cape Colony's history. General changes in the price level are the result of many adjustments. Only a fraction of prices change in a particular month, and price changes may vary in size. This section describes an investigation, following Ruch, Rankin and Du Plessis (2016) and Ruch (2016), into the frequency and size of price changes, to examine the underlying dynamics of pricing behaviour.

#### 4.3.2.1 The size and frequency of price changes

Inflation is the outcome of many different price increases and decreases as well as changing and unchanging prices all occurring in a particular month. In this section the size of price changes and the fraction of prices changing in a specific month are investigated.

A price change is defined as  $dp_{ikt} = (p_{ikt} - p_{ikt-1}) * 100$ , where  $p_{ikt}$  is the log price of a specific product  $i$  in town  $k$  in time  $t$ ; for example, the change in the log price of wheat in Cape Town from one month to the next.

Figure 4.7 plots the distribution of the price changes, in bins of 5 percentage points, as well as the distribution of the standardised price changes. The monthly changes are dominated by no change or small changes, with around 86% of changes being smaller than 10%, in either direction. Three very large price changes occurred for tobacco, cattle and sheep in 1900 during the war. For comparison, Ruch (2016) recently analysed the prices of the goods included in the current CPI and found that monthly price changes were dominated by no change, which accounted for 75% of the changes, while 80% were smaller than 5% in either direction.

## Historical Commodity Prices

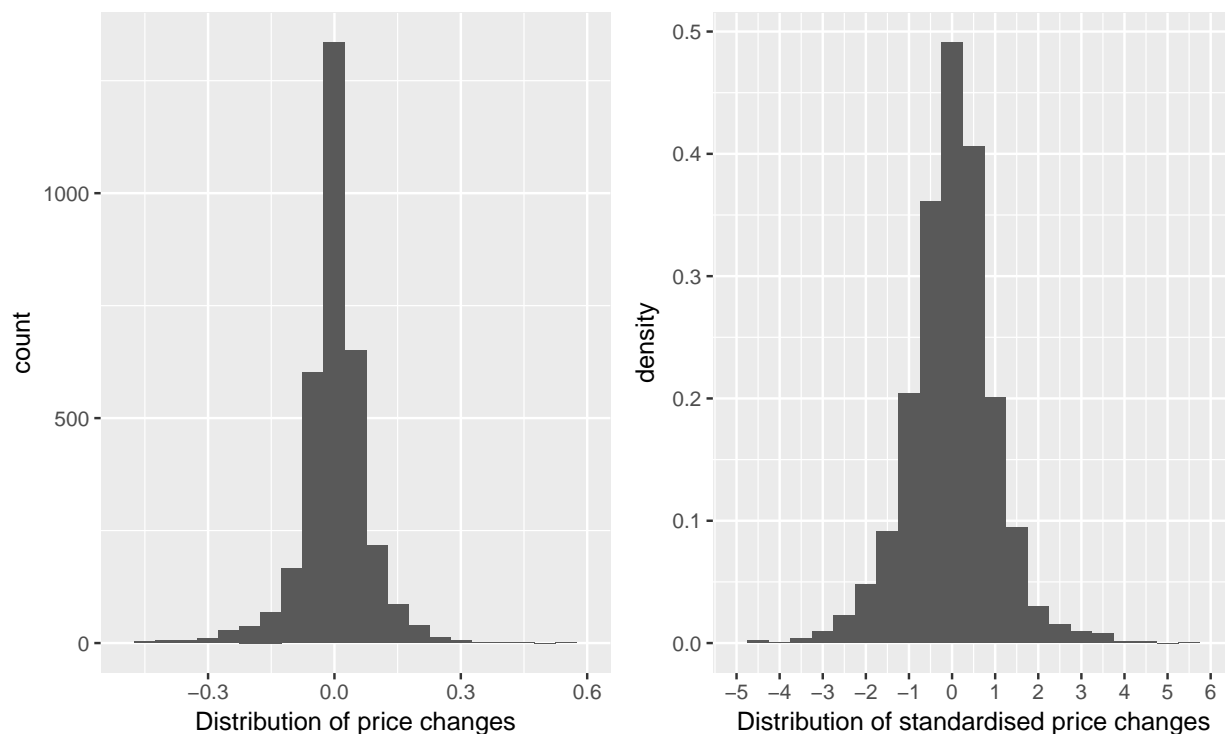


Figure 4.7: The size of price changes

An indicator variable is created to calculate the frequency of price changes:

$$I_{ikt} = \begin{cases} 1, & \text{if } dp_{ikt} \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

The mean of the indicator variable  $I_{it}$  is then calculated for each commodity over all the towns. The indicator variable  $I_{it}$  is therefore equal to the mean frequency of price changes across towns.

Figure 4.8 illustrates the mean and the 12-month moving average of the frequency of monthly price changes over the period. The sample mean, reported in Table 4.3 below, is 56.3%, with a standard deviation of 9.18%. The frequency of price changes decreases in the recession of the early 1890s, increases during the war and then decreases thereafter. The inverse of the frequency measure provides a rough approximation of the duration in months between price changes. The mean frequency implies that the prices of goods changed on average every 1.9 months, or 57 days. For comparison, Ruch (2016) found that the prices of the goods included in the current CPI changed every 3.6 months (according to the mean frequency).

The underlying data allows the decomposition of inflation into the Extensive Margin (EM), i.e. the fraction of price changes, and Intensive Margin (IM), i.e. the size of the price changes. In this instance, inflation is calculated simply as the weighted average size of price changes over all the towns  $k = 1, \dots, K$  and products  $i = 1, \dots, N$ . The commodity weights  $w_i$ , discussed in more detail

## Historical Commodity Prices

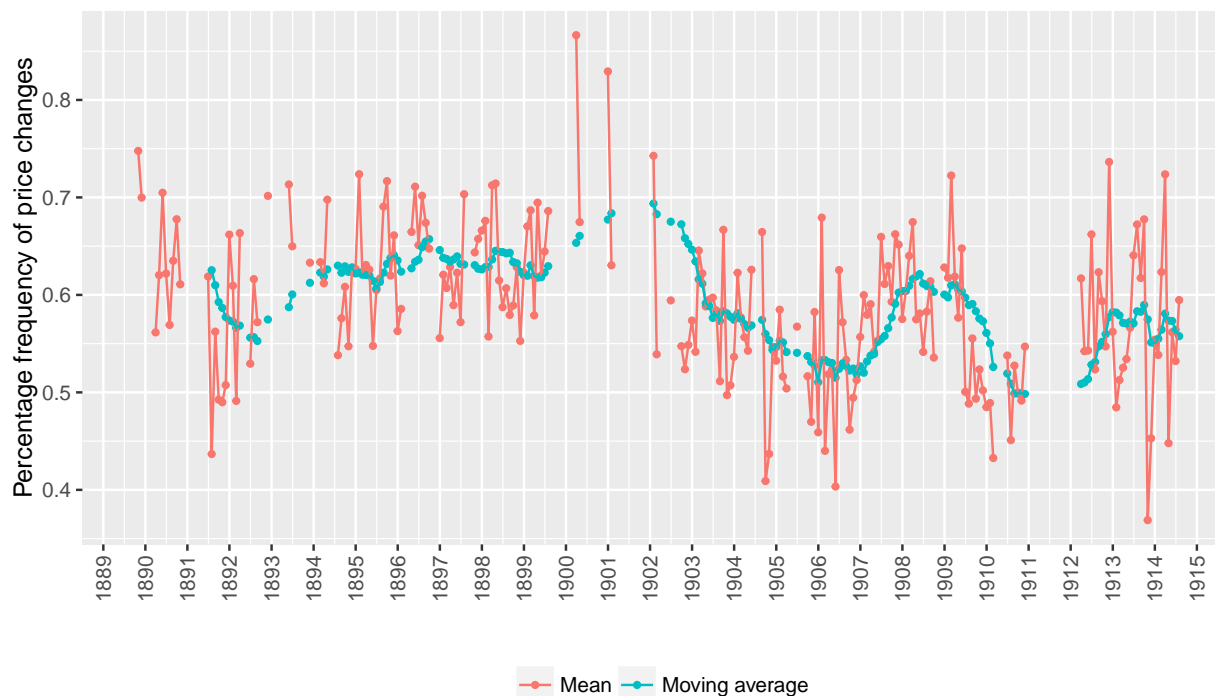


Figure 4.8: The frequency of price changes

below, are the domestic production value shares reported in the 1904 census. The decomposition can aid in the understanding the pricing dynamics over time, providing insight into whether inflation is driven by the average size of changes or how often prices change (Klenow and Kryvtsov, 2008; Ruch, 2016). Inflation is the product of the fraction of items with price changes (EM) and the average size of price changes (IM):

$$\pi_t = \frac{1}{NKW} \sum_{i=1}^N \sum_{k=1}^K w_i (p_{ikt} - p_{ikt-1})$$

$$\pi_t = EM_t * IM_t$$

$$\pi_t = \frac{1}{NKW} \sum_{i=1}^N \sum_{k=1}^K w_i I_{ikt} \frac{\frac{1}{NKW} \sum_{i=1}^N \sum_{k=1}^K w_i (p_{ikt} - p_{ikt-1})}{\frac{1}{NKW} \sum_{i=1}^N \sum_{k=1}^K w_i I_{ikt}},$$

where the variables are defined in the same way as above.

Figure 4.9 illustrates the 12-month moving average of the monthly inflation rate  $\pi_t$ , the intensive margin  $IM_t$ , and the extensive margin  $EM_t$  over the period. The extensive margin is scaled for ease of reference. Monthly inflation averages 0.03%, as reported in Table 4.3, with a standard deviation of 3.38%, although there seem to be some pronounced cycles over the period. The average inflation rates is due to an average fraction of 56.3% of prices changing every month (EM) and -0.15% average monthly price increase (IM). The intensive margin is highly correlated with this



## Historical Commodity Prices

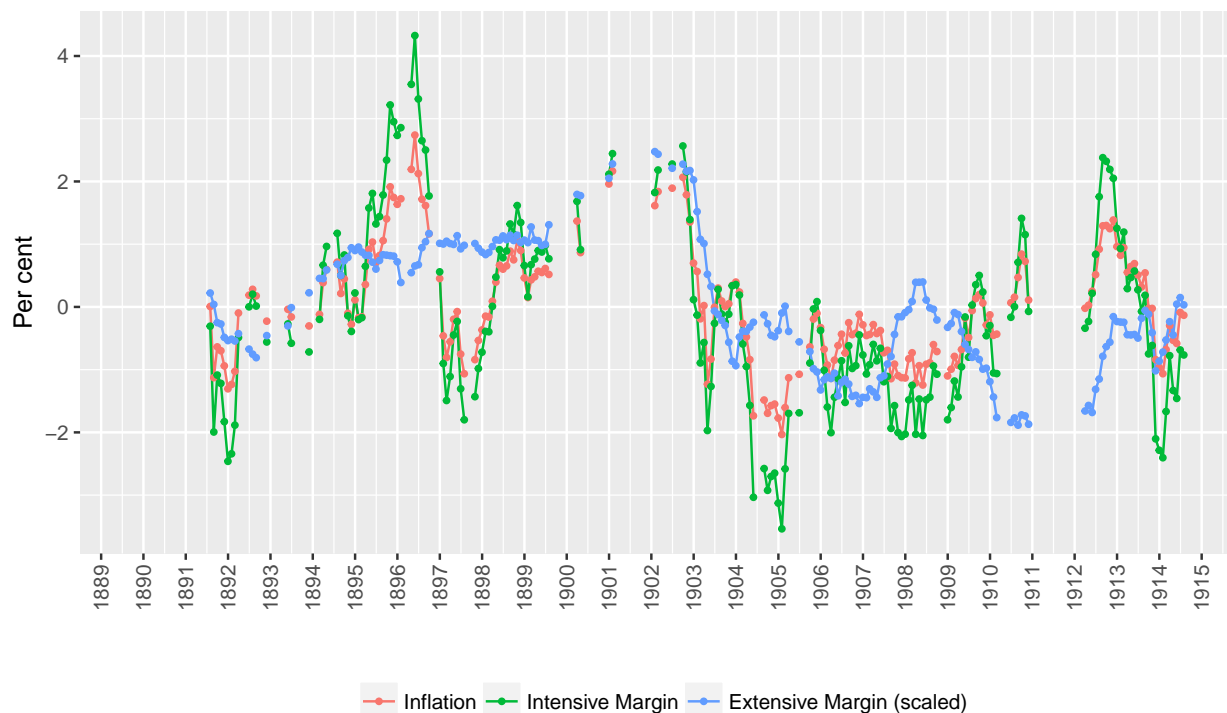


Figure 4.9: Decomposition of inflation (12-month moving averages)

measure of inflation, as reported in Table 4.3, as seems to be the more important driver of inflation.

Table 4.3: Pricing dynamics

	Mean %	Stan Dev %	Cor with Inflation
Inflation	0.03	3.38	1.00
Intensive Margin	-0.15	5.50	0.98
Extensive Margin	56.30	9.18	0.24

Clearly there are substantial gaps in the records, which prevent more informative decompositions of inflation, as in Ruch (2016) and Klenow and Kryvtsov (2008). The missing data problem is amplified when inflation is calculated as simple period-on-period growth rates.

#### 4.4 Methodology

This section first discusses the limitations of the existing methods for aggregating these datasets. It then presents the specific methodology used to create the historical commodity price indices in this chapter.

## Historical Commodity Prices

Table 4.4: Repeat sales example with wheat prices

Date	Beaufort West	Cape Town	Worcester	Median	Repeat Sales
May 1891	150.00		150.00	150.00	100.00
Jun 1891	135.00	138.00		136.50	93.75
Jul 1891	135.00	150.00		142.50	99.75
Aug 1891			144.00	144.00	92.16
Sep 1891		144.00		144.00	94.90
Oct 1891		144.00		144.00	94.04
Nov 1891	120.00	144.00		132.00	93.20
Dec 1891			144.00	144.00	88.48
Jan 1892		126.00	144.00	135.00	84.94
Feb 1892		126.00		126.00	84.94

#### 4.4.1 Limitations of Traditional Techniques

A simple example may be used to illustrate different methods to aggregate these incomplete series. Table 4.4 reports the wheat prices per 100 lbs for 3 towns, over a period of 10 months, recorded in the Agricultural Journals.

As discussed in Chapter 2, the simplest method for constructing a price index is to calculate a measure of the central tendency of the price distribution. In this instance, a price index for a specific product could be calculated as the central tendency (i.e. the mean or median) of the growth rates across towns. However, when inflation is calculated as period-on-period growth rates, as above, every missing price observation would be associated with two missing inflation observations. In the example, for instance, there would be no period-on-period growth rates for August and September 1891, because there are no consecutive price observations. The inflation calculation would produce missing values for those months, even though market prices were recorded in those periods. Thus, in cases where consecutive prices were not recorded, those single observations would not be incorporated into the index calculation. For instance, the Beaufort West price for November 1891 and the Worcester price for August 1891 would be discarded. Moreover, the annual prices recorded in the Blue Books would be discarded with this method, as month-on-month growth rates could not be calculated.

This central tendency method would therefore require substantial interpolation and in some cases substantial extrapolation to create complete series. For instance, interpolation methods (e.g. time-series models and state space models such as the Kalman filter) could be implemented to complete the individual monthly series, and temporal distribution methods (such as the Chow-Lin or Litterman methods) could be implemented to convert the annual prices from the Blue Books to monthly series. However, interpolation methods can become inaccurate when there are consecutive missing values (Fung, 2006). In these historical records, every commodity price series for a particular town contains numerous missing observations. In some cases, there are long periods where no prices were recorded. In this case, the individual series are too patchy to produce sensible results and interpolation and extrapolation methods are not feasible.

The traditional procedure for dealing with the problem of gaps in historical data is to splice the series when there is overlapping information (as in Solar and Klovland (2011)). In the example, for instance, it would be natural to combine the price increases from Beaufort West and Cape Town for July 1891. For February 1892, the Cape Town price series might be chained to the Worcester price series. Yet, the problem of incomplete information remains in a number of the periods. Additional series, with more overlapping information, might aid the calculation. However, in practice, with many time series, this procedure is difficult and time consuming to handle consistently and requires considerable work and great care (Klovland, 2014). In many cases, when there is no unique way of splicing the time series, the procedure may become quite subjective. Thus, if there are many time series and numerous gaps in the data, manual splicing can become difficult to implement consistently.

An alternative is to calculate the central tendency of the prices in levels. The median may be preferred to the mean to diminish the impact of outliers. The problem is that these indices are more dependent on which prices were recorded than on changes in the underlying market prices. In the example, for instance, the median index decreases in November 1891 and February 1892, simply because the lower prices for Beaufort West and Cape Town were included in those periods. Thus, the inclusion or exclusion of a specific town may have a large impact on the median index.

A further difficulty with this method in the current context is that only prices that were recorded in the same measurement units and for the same quality descriptions are included in the simple median calculation. This means that the prices recorded in the Blue Books, which refer to different measurement units than in the Agricultural Journals, would be excluded from the calculation. Different quality descriptions for the same product would also be excluded from the calculation.

Nevertheless, in this chapter simple median indices are calculated to provide an internal validity test for the price indices calculated with the repeat sales method. The median indices are calculated as the median price for an individual product across towns, using the Agricultural Journal records, which were converted to the same measurement units for each product.

Klovland (2014) suggested that the repeat sales method could be useful for aggregating underlying series characterised by incomplete observations. The repeat sales method was specifically developed for markets in which the price of an item is recorded infrequently and at irregular intervals, which is typical in the real estate market. The problems in aggregating the high-frequency historical data are similar in principle, although the gaps between the observations are usually shorter. The repeat sales approach compares the price of that same item over time, with a sale occurring whenever a data point is observed. One can think of each commodity price series (for a particular quality description) in each town as a specific item (e.g. an artwork) to track over time. The following section presents the repeat sales method in more detail.

#### 4.4.2 The Repeat Sales Method

As discussed in Chapter 2, Bailey, Muth and Nourse (1963) initially proposed the repeat sales method to calculate house price changes. Their procedure was a generalisation of the chained-matched model method previously used to construct real estate price indices. While the repeat sales model was derived as the differenced hedonic model in Chapter 2, following Wang and Zorn (1997), it can also be derived in the following manner. Where  $t_1$  and  $t_2$  are the times of the first and second transactions, a price ratio  $\frac{P_{i,t_2}}{P_{i,t_1}}$  can be modelled as:

$$\frac{P_{i,t_2}}{P_{i,t_1}} = \frac{I_{t_2}}{I_{t_1}} \times u_{i,t},$$

where  $P_{it}$  is the price of commodity  $i$  at time  $t$ ;  $I_t$  is the true but unknown index for period  $t$ ; and  $u_{i,t}$  is the error term. Taking logs

$$\ln \frac{P_{i,t_2}}{P_{i,t_1}} = -\ln(I_{t_1}) + \ln(I_{t_2}) + \ln(u_{i,t})$$

This relationship can be expressed in vector notation as  $y = X\beta + \epsilon$ , where  $y$  is the vector of logged price ratios;  $\beta$  is a  $T$ -dimensional column vector of unknown logarithms of the index numbers to be estimated, such that the  $t$ -th component of the  $\beta$  vector is  $\beta_t = \ln(I_t)$ ;  $X$  is a matrix of  $(n \times T)$  dimensions such that the  $t$ -th component of each row is  $-1$  if  $t = t_1$ ,  $+1$  if  $t = t_2$ , and  $0$  otherwise; and  $\epsilon$  is the vector of  $\log(u_{i,t})$  values.

The standard specification of the repeat sales method usually takes the following form:

$$\ln \frac{P_{it}}{P_{is}} = \sum_{t=1}^T \beta_t D_{it} + u_{it},$$

where  $P_{it}$  is the price of a specific commodity  $i$  (e.g. wheat in Cape Town) in time  $t$ ;  $P_{is}$  is the price of the same commodity  $i$  at time  $s$ ;  $\beta_t$  is the parameter to be estimated for time  $t$ ;  $D_{it}$  represents a time dummy equal to  $1$  in period  $t$  when the resale occurred,  $-1$  in period  $s$  when the previous sale occurred, and  $0$  otherwise; and  $u_{it}$  is a white noise residual.

Thus, in the standard repeat sales model, the dependent variable is regressed on a set of dummy variables reflecting time periods. The coefficients are estimated on the basis of changes in item prices over time. The index numbers are derived from the dummy variable coefficients. The price index is simply the antilog of the series of estimated coefficients:  $\hat{\beta}_1, \dots, \hat{\beta}_T$ .

At each time  $t$ , there exists a distribution of growth rates in the sample. The repeat sales estimator is a measure of the central tendency of this distribution, in the form of the geometric mean of the growth rates of the items that sold more than once. Bailey, Muth and Nourse (1963) showed that the  $\beta$  estimates (i.e. the mean logged price indices) consist of the period-by-period weighted

---

## Historical Commodity Prices

---

averages of the logged price ratios, with weights being proportional to the sample sizes. In other words, the regression solutions are complicated weighted averages of the average logged price ratios. Wang and Zorn (1997) showed this clearly for a relatively simple and intuitively appealing example with two periods. For example, the first index value in  $t_1$  is the weighted average of two quantities: the average price appreciation of items observed in period  $t_0$  and  $t_1$ ; and the average appreciation from  $t_0$  to  $t_2$  minus the average appreciation from  $t_1$  to  $t_2$ . The weights given to the two quantities are proportional to their sample sizes.

The repeat sales method therefore uses a simple least squares regression on time dummies to produce the estimated index values. In this case, repeat sales pairs are formed from the price ratios of the same commodity for each town. In the example above, 12 sales pairs and 10 time dummies can be created for the 10 time periods. For instance, three sales pairs can be formed for Beaufort West (e.g. 135/150). Running a least squares regression produces a set of coefficient estimates for the 10 dummy variables. Taking the antilog to these estimates produces the index in Table 4.4. In this case the additional prices from Worcester help to make the index complete, specifically by adding price observations in August and December 1891, where there were only missing observations. Thus, including more price series will lead to fewer gaps in the index. Even if the prices are at different levels, due to factors such as transport costs, for the purposes of the index, the trends (or growth rates) are compared over time.

The repeat sales method utilises all of the information in the dataset, compared to the cumbersome traditional approach of manually splicing individual time series (Klovland, 2014). As Wang and Zorn (1997) noted, if the number of observations does not vary across time periods (i.e. if there are no gaps in the time series), the repeat sales estimator simplifies to an ordinary chain index.

This method has the advantage of being able to handle gaps of any length in the data series and can incorporate prices for different measurement units and quality descriptions. In this case, new quality descriptions are occasionally introduced for a product (e.g. Colonial beer per gallon) and old ones disappear (e.g. English beer per bottle). The repeat sales approach can handle this instance by tracking prices for the same quality description over time. Moreover, the annual prices from the Blue Books may be easily incorporated, by forming price ratios whenever these prices were recorded, and tracking the same measurement units over time.

As discussed in Chapter 2, the disadvantages usually associated with the repeat sales method are that single-sale data are discarded and that there is sample selection bias in the types of commodities that are sold more than once (Hansen, 2009). In this case, however, these disadvantages are not big concerns. All of the commodities were sold numerous times and could be included in the index estimates.

The repeat sales approach treats all observations as equal, which means that towns that were observed more frequently than others will exert a stronger influence on the index, simply because there will be more observations in the dataset originating from this item. For example, if Cape Town

wheat prices are recorded and included in the index calculations more frequently than Worcester prices, the Cape Town prices will have a larger implicit weighting in the index. In other words, the prices for the same commodity in different towns are treated equally, and towns with more observations have a higher implicit weighting. This might be a sensible approach, as regularly quoted prices are often the most frequently traded (Klovland, 2014).

It is possible that the smaller towns have more price observations in some periods that do not properly reflect price movements in the larger towns, which might then bias the index to some degree. For instance, there might be an idiosyncratic spike in Worcester wheat prices in a period when only Worcester prices were reported. However, idiosyncratic price movements in specific towns are less likely to bias the overall indices because the estimates include a large number of towns from two separate datasets.

The repeat sales method therefore provides a consistent way to aggregate the data from the different towns for a specific commodity. Price series from a number of towns help to make the index continuous, by adding observations where individual series suffer from missing observations. Thus, including more price series will lead to fewer gaps in the index. The repeat sales method is particularly useful when prices from different sources, in different units, need to be aggregated. In this case, it provides a simple way to include the prices from the Blue Books, by forming sales pairs between the annual data points. Even if the prices are at different levels due to factors such as transport costs, for the purposes of the index, the growth rates are compared over time. Hence, this method produces an index with substantially fewer gaps than there are in the individual series, without the need for substantial interpolation or extrapolation.

#### **4.4.3 Creating Indices for Specific Commodities**

Klovland (2014) suggested that the simple unweighted version of the repeat sales model is most applicable for specific commodities, i.e. at the lowest level of aggregation, when the explicit weighting of different price observations is less crucial. The towns are weighted implicitly according to coverage, i.e. the number of observations in each dataset. Following this suggestion, the repeat sales method is used to aggregate the different time series for each commodity from all of the towns in the Cape Colony.

The annual prices from the Blue Books are also incorporated in the monthly indices. This is useful because the monthly commodity indices still contain substantial gaps where no data is available. It is also useful because the monthly dataset omits a number of important products and towns.

The Blue Books report the annual market prices as averages for November of each year. The simplest way to incorporate the annual prices from the Blue Books is to treat them as the prices for November of each year, and then to simply form sales pairs between the annual data points whenever they were recorded. The repeat sales method is then applied to all of the sales pairs for a

## Historical Commodity Prices

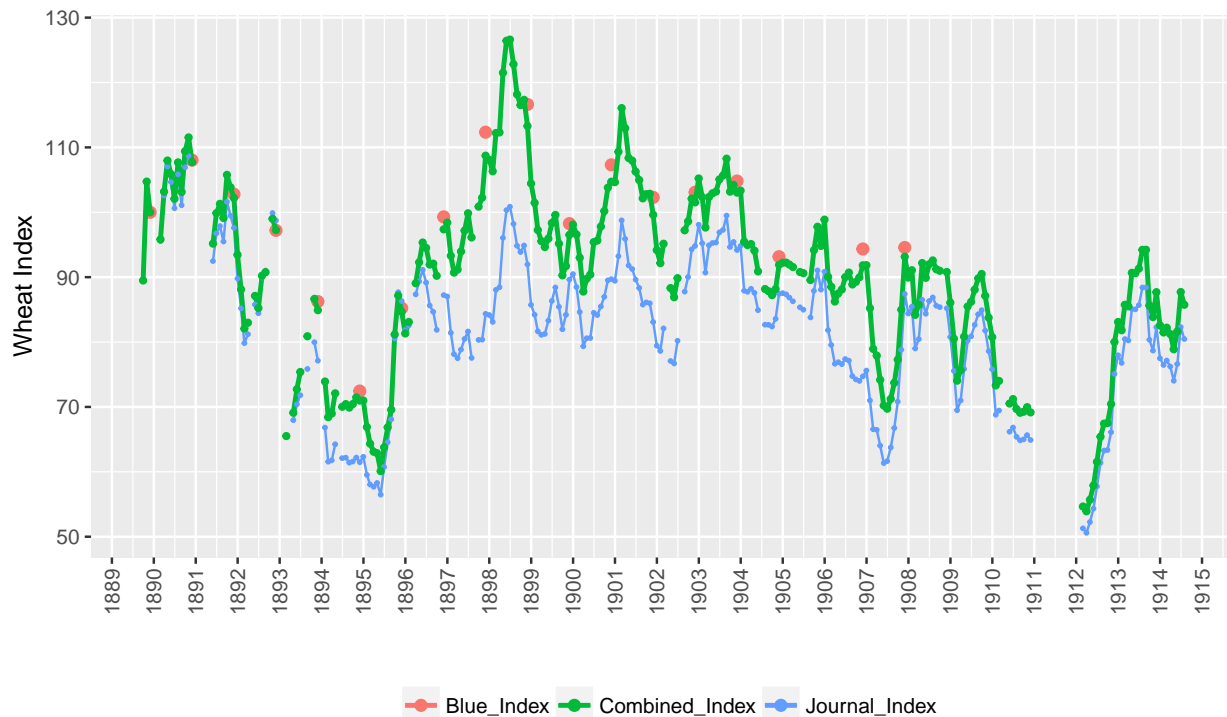


Figure 4.10: Wheat price indices for the Cape Colony (Nov-1889=100)

specific commodity. This procedure differs slightly from the one used in Klovland (2014), in that his price records were extracted from more underlying sources (4 or 5 primary sources), and referred to fewer towns (12).

Figure 4.10 illustrates the index for wheat prices in the Cape Colony, as well as two indices estimated separately from the two datasets.<sup>38</sup> For the separate indices, the monthly wheat prices for 19 towns in the Cape Colony were combined to form a monthly wheat index, while the annual wheat prices for the 48 towns were combined to form an annual wheat price index. The total index combines the prices from both sources and closely reflects the separate indices. The repeat sales method allows for the estimation of an index with more complete coverage, although there are still a few gaps for the months when no price information was recorded. The index is interpolated at this stage to obtain a complete index. This involves significantly less interpolation than would have been necessary if each individual series was interpolated from the start.

A central assumption of the methods used to combine the prices from these two sources is that the growth rates were similar across sources. It is possible that changes in factors such as mark-ups

<sup>38</sup>As a robustness check, the repeat sales method was also estimated on the two datasets separately, to obtain two indices for each commodity, one monthly and one annual. After interpolating (temporally distributing) the annual index, the two monthly indices for each commodity were combined again with the repeat sales method to obtain one reasonably complete index for each commodity that covers all the towns in the Cape Colony. With this method, the indices from the two data sources implicitly received an equal weighting. The results were similar to the results reported below.

---

 Historical Commodity Prices
 

---

or data coverage (regionally and over time) may call this assumption into question. To test the assumption, the average prices from the two sources can be compared where there is overlapping information. Table 4.5 reports the correlations (in levels and in growth rates) between the average annual prices for a number of commodities for which there is sufficient overlapping information from the two sources. The overlapping information includes the prices for November of each year from the Agricultural Journals in those towns for which the Blue Books also reported prices. The correlations are mostly high and significant, which implies that average annual prices were at least similar and that the assumption seems to hold reasonably well.

Table 4.5: Correlations in overlapping average annual commodity prices

Commodity	Levels	Growth
wheat	0.80***	0.63**
mealies	0.92***	0.87***
barley	0.82***	0.72**
oats	0.84***	0.47*
potatoes	0.85***	0.74***
tobacco	0.87***	0.58**
beef	0.98***	0.87***
mutton	0.99***	0.96***
butter	0.52**	0.30

#### 4.4.4 Weighing the Commodity Indices

The construction of the indices involve three stages of aggregation. The first stage involves aggregating the price information from the various towns from the two sources into indices for 43 individual commodities. This is accomplished by using the repeat sales method described above. It involves combining the individual data series on each commodity in an efficient way to form an index, using all available price information. For many commodities, there are prices from both sources, as well as prices for different product qualities (e.g. Colonial and English Beer, or high and ordinary quality wine). This stage involves only an implicit weighting, where the towns with the greatest number of observations are the most influential in determining the coefficient estimates (Klovland, 2014).

The next stage involves aggregating the individual commodity price indices into seven broader commodity group indices, which are weighted together in the conventional manner (e.g. with the Laspeyres price index). For instance, Klovland (2014) aggregated his 110 commodities into 16 commodity groups (grain, meat, etc.) using chained Laspeyres indices, with weights based on output or trade values in 1835, 1870, 1890 and 1910. This raises the question of what weights to use to aggregate the different commodities.

The Agricultural Journals, particularly in the earlier volumes, refer to these prices as ‘wholesale’ rates. ‘Wholesale’ prices typically refer to sales in large lots, often the first commercial transaction in major trading centres (Klovland, 2014). The records for Provisions in the Blue Books possibly reflect average retail prices, rather than the wholesale prices of the other series (De Zwart, 2011).



---

 Historical Commodity Prices
 

---

Table 4.6: Commodity classification

Crops	Agricultural Produce	Pastoral Products	Livestock	Pastoral Provisions	Agricultural Provisions	Other Provisions
Wheat	Tobacco	Wool	Cattle	Beef	Bread	Tea
Mealies	Dried Fruit	Hides	Horses, Mules & Asses	Mutton	Flour	Coffee
Barley	Wine	Skins	Sheep	Pork	Mealie Meal	Sugar
Oats	Brandy	Cheese	Pigs	Eggs	Boer Meal	Beer
Oathay		Fat & Tallow	Goats	Butter	Oatmeal	Rice
Rye		Soap	Fowls & Ducks	Milk		Salt
Peas & Beans						Candles
Potatoes						

One would expect the retail prices to follow a similar trend to the wholesale prices and this chapter follows Klovland (2014) in including these prices in the index calculations.

The first price indices were often called wholesale price indices and typically included both imported and domestically produced goods, although weights were not always applied consistently (Klovland, 2014). Nowadays producer price indices (PPIs) are more popular. PPIs focus on prices obtained by domestic producers, and therefore include domestic goods sold at home, and in some cases, exported goods.

In this chapter, the focus is on domestic production value shares. The produce returns reported in the 1904 census are used as the benchmark weights. The value shares were calculated as the average market price multiplied by the volume of production of each commodity. The difficulty is that the census did not cover all of the commodities in the sample, and a number of important products lack weights, e.g. wheat flour and beef. The weights for these products were based on the import values reported in the Blue Books. This approach follows Greyling and Verhoef (2015), where agricultural output was based on the volume of production reported in the censuses, and was also supplemented with information on export and import values. Thus, the different commodity indices for the Cape Colony are aggregated by applying weights based on the value shares of each commodity in domestic production, and supplemented by import value weights. This weighting scheme makes the indices akin to the wholesale price index used in Klovland (2014).

The individual commodities included in each of the seven commodity group indices are reported in Table 4.6. The categorisation is informed by the classification of the commodities in the data sources. For instance, in the Blue Books, the prices are divided into Agricultural Produce, Stock and Animal Productions, and Provisions.<sup>39</sup>

The weights are required only to be relative value shares, i.e. weights relative to the other commodities in the larger commodity groups, which makes it easier to compare the commodities in each group. For instance, the livestock production numbers from the census may be compared with one another

---

<sup>39</sup>A few commodities had to be excluded from the analysis because of a lack of observations. These include lucerne hay and oranges from the Agricultural Journals, as well pumpkins, aloes, argol, condensed milk and lamp oil from the Blue Books.

to form the livestock index. However, only a portion of livestock produced would have been sold in any given period, which makes it difficult to compare these numbers with the other commodities, e.g. the gallons of milk. This is also the reason that the commodities in the indices for provisions are kept apart from the other commodities. For example, wheat (Crops) and wheat flour (Agricultural Provisions) are included in separate commodity group indices, because the weights for Crops are based on production value shares from the 1904 census, whereas the weights for Agricultural Provisions is based on import value shares.

A number of products are reported only in the Blue Books, which means that the indices reflect only values for November of each year. As in Klovland (2014), these indices are interpolated to form monthly indices, which makes them smoother than the actual monthly price would have been. The series are seasonally adjusted and the weighted average of the price ratios are calculated, as recommended by the International Monetary Fund (2004) for the construction of the PPI.

The final stage combines the seven indices for the commodity groups to form an overall aggregate wholesale commodity price index. The difficulty is that the weights are not all comparable, given that some of the commodity groups are weighted with reference to production values in the 1904 census, while the three provision indices are weighted with reference to import values. There will always be some difficulties in finding the best weights with historical data and different assumptions may be used to create relative value shares. Klovland (2014), for instance, estimated the relative weights for each commodity within the group as follows: the commodity with the greatest market value was given a weight of 10, and the other commodities were scaled proportionately, using rounded weights, subject to the constraint that all the time series for which data was available would receive a weight of at least one.

In order to combine the seven commodity groups into a total wholesale commodity price index, a similar assumption is made about the relative weights of the commodity groups. The following weights are based on the available information on production and import value shares: Crops (5), Agricultural Produce (5), Pastoral Products (3), Livestock (5), Pastoral Provisions (5), Agricultural Provisions (3), and Other Provisions (3). The intention is merely to use reasonable weights to aggregate the indices (Klovland, 2014).<sup>40</sup> The following section presents the results for the commodity price indices, for each level of aggregation.

## 4.5 Index Results

This section presents the results for the repeat sales commodity price indices at each level of aggregation. The individual price indices are intended to shed light on the demand and supply

---

<sup>40</sup>Klovland (2014) makes two further points. First, if a raw material is used in the production of a more finished commodity, some reduction in the weights based on gross output measures may be necessary. Second, a large portion of agricultural products may have been consumed on the farms, rather than sold on the market. In practice, it is difficult to find evidence of how large these adjustments should be.

factors that influenced product prices. The more comprehensive total wholesale commodity price index is intended to provide a clearer picture of the inflation history of the period.

#### 4.5.1 Repeat Sales Price Indices

Figure 4.11 illustrates the seasonally adjusted commodity price indices for four individual commodities: wheat, tobacco, cattle and beef. There was a large variation in price movements when comparing individual commodity prices. The wheat price index decreased significantly after the recessionary period of the early 1890s. The index reached a trough in May 1895 and subsequently recovered during the upswing phase up to mid-1888. Subsequently, wheat prices were on a general downward trajectory, with two exceptions occurring during the war (early 1901) and the post-war recovery (mid-1903). Wheat prices reached a trough in November 1911, before recovering somewhat towards the end of the sample period.

In contrast, the tobacco price index remained relatively stable up to the outbreak of the war, before increasing substantially during the war period, from October 1899 to May 1901. Tobacco prices returned to pre-war levels during the post-war recession. De Kock (1924) reports that there was abnormal demand for products like tobacco and wine during the war, followed by a substantial decline in prices during the subsequent recession.

The indices for cattle and beef follow similar paths, which is reassuring for such closely related products (cattle prices were derived from separate series for cattle, transport oxen and milk cows). Both price indices increased significantly in the run-up to the war, with subsequent peaks occurring in the post-war recovery period from 1902 to around 1906. Cattle prices exhibited spikes during the war in the middle of 1900 and during the post-war recovery phases. There was a relatively sharp decrease into the recession in 1907, and the subsequent cycle peaked towards the end of 1911. The effect of the rinderpest outbreak in 1896 is barely visible, probably because cattle herd losses occurred mainly in the Transvaal, Natal, the Orange Free State and the eastern districts of the Cape, as opposed to the entire Cape Colony (De Kock, 1924).

Beef prices increased significantly in the run-up to the war, with a large peak in 1903 during the recovery phase. This was followed by a significant decrease during the post-war recessionary period, and a moderate increase towards the end of the period. De Kock (1924) reports that large quantities of frozen meat had to be imported for years after the war to satisfy local requirements.

The indices for the commodity groups are illustrated in Figures 4.12 and 4.13. Figure 4.14 illustrates the total wholesale commodity price index, with recessionary periods shaded. The trends in the wholesale commodity price indices conform to the literature on the economic history of the Cape Colony. Commodity prices decreased during the recessionary periods of 1891-1893 and 1894-1895, and increased during the upswing phase from 1895. Prices decreased briefly in 1897, reflecting the recession that Schumann (1938) identified in that year. Virtually all of the indices increased

## Historical Commodity Prices

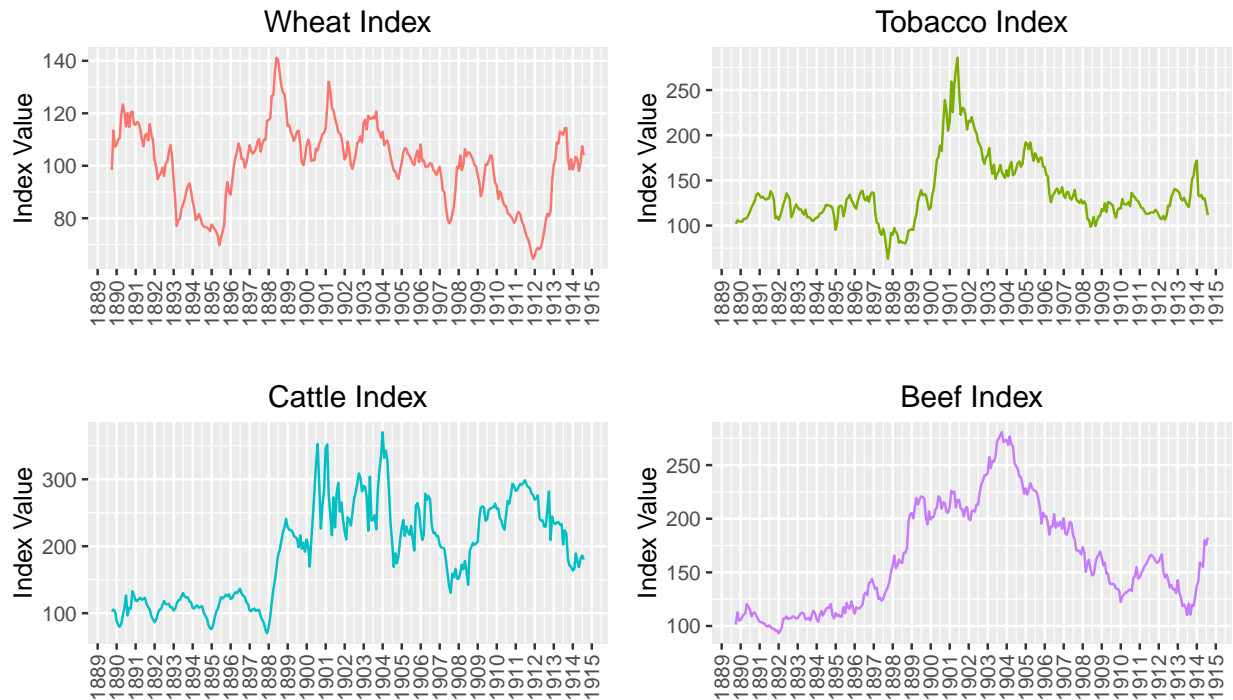


Figure 4.11: Individual commodity price indices

markedly during the war, with the total commodity price index reaching a peak in 1901. The exception was the general decline in the index for ‘Other Provisions’ (e.g. tea and coffee). Prices remained at a high level and even increased in some cases during the post-war recovery up to 1903. Most of the price indices decreased markedly during the post-war recession of 1903-1905 and again during the recession of 1907-1910, to their levels before the war. In most cases prices were stable or increased slightly towards the end of the sample period.

The price indices therefore exhibited clear cyclical trends over the period, with a large increase before and during the war and a large decrease after the war up to unification. A number of studies have found similar large increases during the war-time periods (e.g. Schumann, 1938; Friedman, 1952; and Klovland, 2014). The inflationary episodes in the total commodity price index seem to correspond with the expansions and contractions in the money supply, as discussed in more detail below.

In summary, the wholesale commodity price indices estimated with the repeat sales method seem to conform to the economic history of the Cape Colony. The price indices provide some insights into the supply and demand developments over the period, especially around the time of the Second South African War. In general, prices exhibited a pronounced cyclical trend over the period, with a large increase before and during the war and a large decrease after the war and up to unification.

Historical Commodity Prices

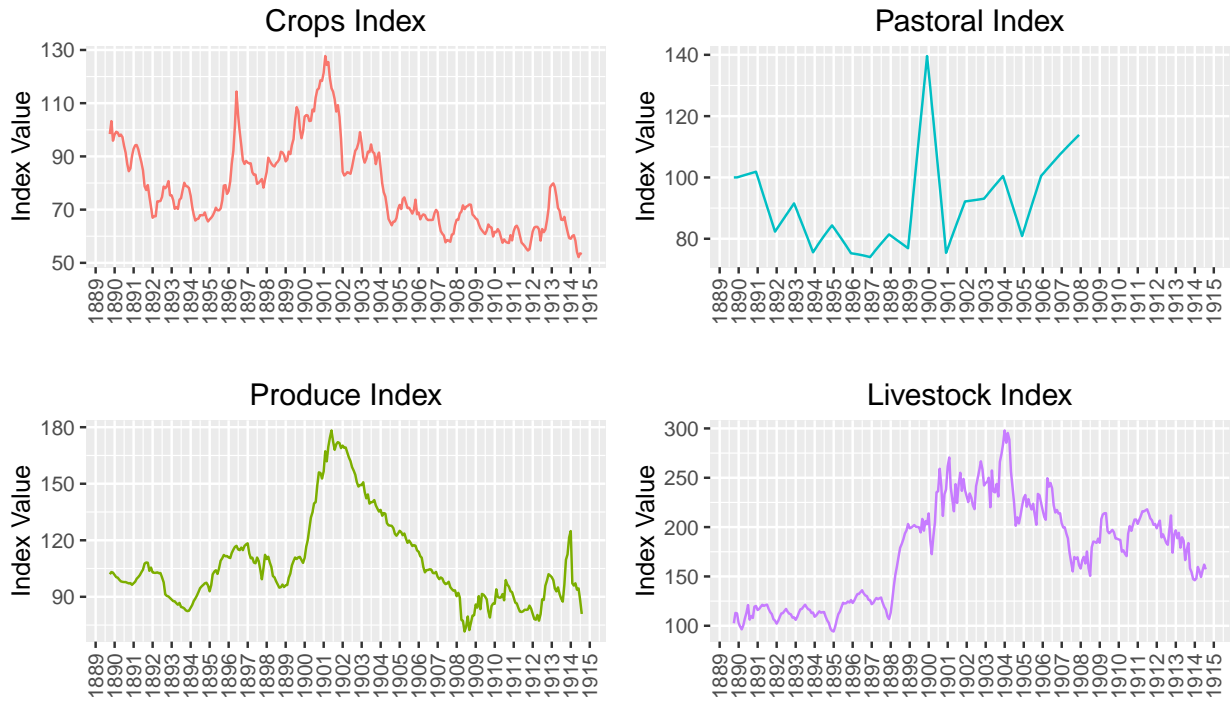


Figure 4.12: Commodity group indices (1)

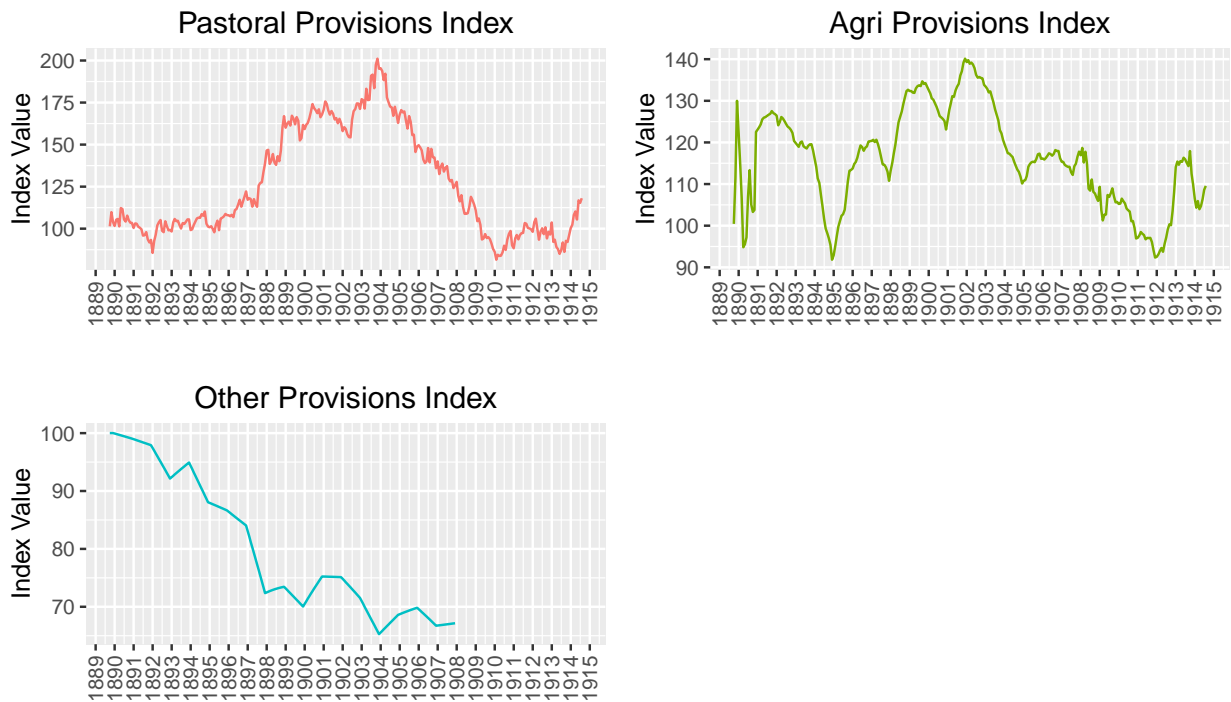


Figure 4.13: Commodity group indices (2)

---

 Historical Commodity Prices
 

---

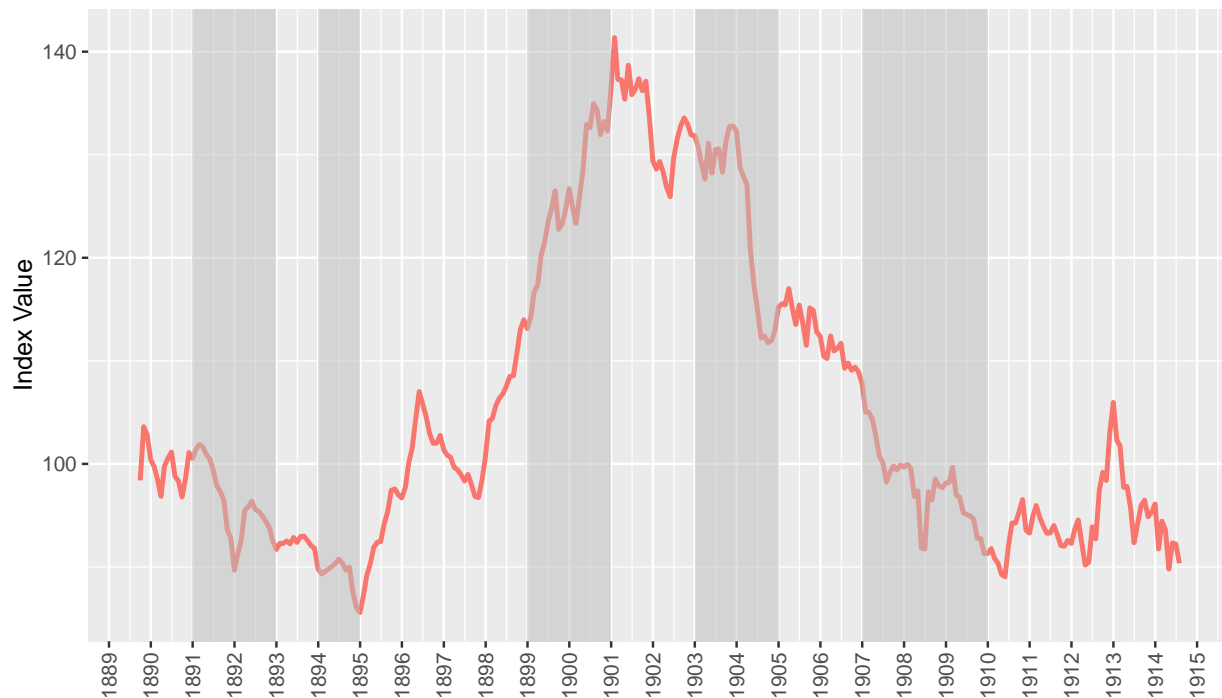


Figure 4.14: The total commodity price index for the Cape Colony

## 4.6 Validity Tests

In this section, the price indices calculated with the repeat sales method are compared with simple median indices to provide an internal validity check. The median indices are calculated as the median prices for each individual product across towns, using the *Agricultural Journal* records, which were converted to the same measurement units for each product. The external validity of the indices is assessed in this section by testing their conformity to existing price indices for the Cape Colony. In the absence of an existing monthly commodity price index, the total commodity price indices are compared with the annual consumer price indices reported in De Zwart (2011).

### 4.6.1 Internal Validity

The price indices for four individual commodities are compared in Figure 4.15 and the total commodity price indices are compared in Figure 4.16. The repeat sales and median indices exhibit similar trends over the sample period. The repeat sales indices tend to be less volatile than the median indices. The median index mirrors the developments in the repeat sales index, but is more volatile and shows a larger increase during the war and post-war recovery period. The difference is due to the fact that the repeat sales index includes the price indices for numerous additional products recorded only in the *Blue Books*.

Historical Commodity Prices

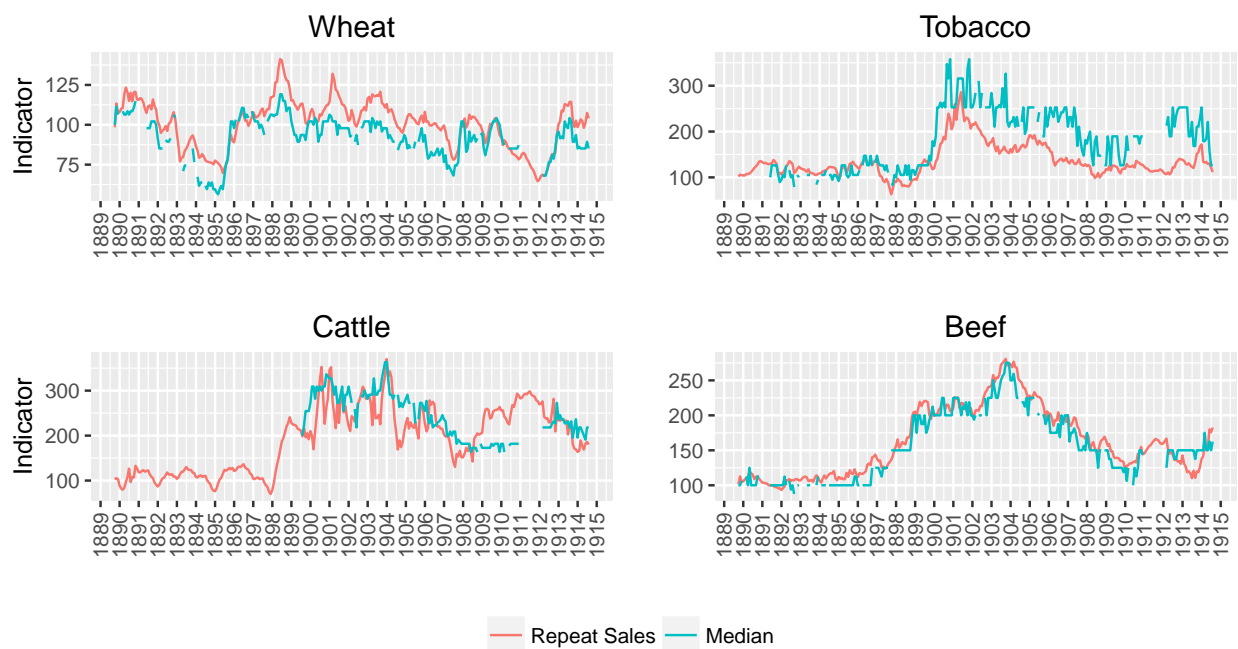


Figure 4.15: Comparing individual commodity price indices

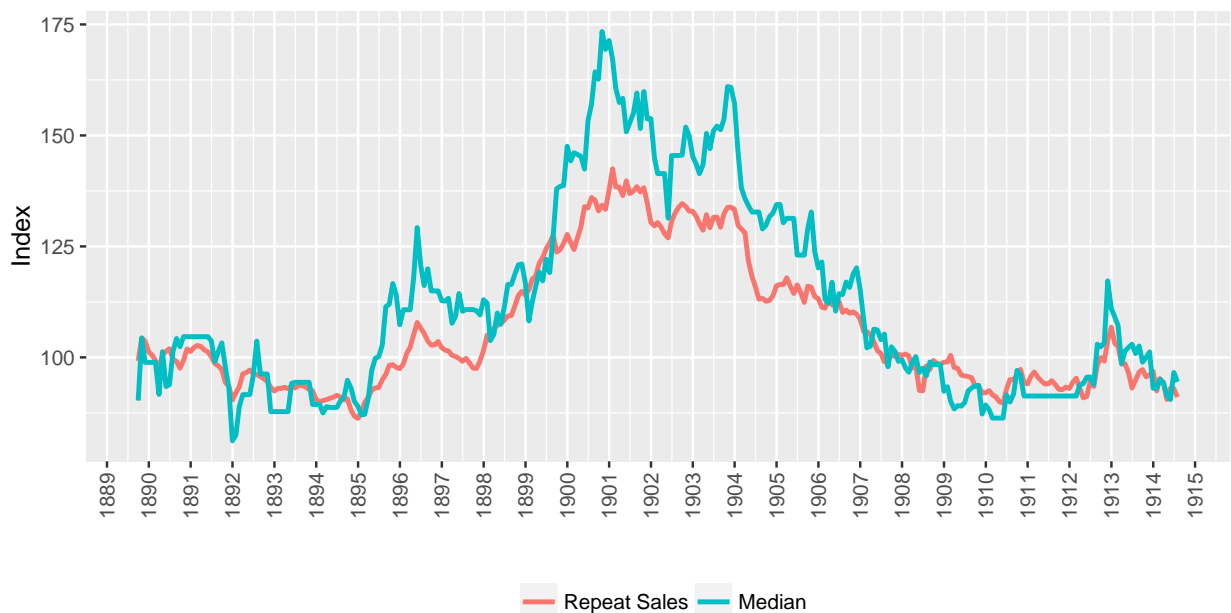


Figure 4.16: Comparing total commodity price indices for the Cape Colony (1890=100)

---

 Historical Commodity Prices
 

---

Table 4.7 reports the correlations among the repeat sales and median indices for each commodity, both in levels and in growth rates. There is a significant positive correlation among virtually all of the commodity price indices calculated with the two methodologies. The correlations between the total repeat sales index and the total median index are also high and significant. The repeat sales indices therefore seem to capture the central tendency in the underlying Agricultural Journal records, while being able to incorporate additional information from the Blue Books.

Table 4.7: Correlations between repeat sales and median indices

	Levels	Growth
wheat	0.84***	0.56***
wheat.flour	0.62***	0.45***
boer.meal	0.73***	0.48***
mealies	0.94***	0.70***
mealie.meal	0.69***	0.60***
barley	0.81***	0.56***
oats	0.88***	0.54***
oathay	0.77***	0.57***
potatoes	0.68***	0.59***
tobacco	0.80***	0.37***
beef	0.96***	0.33***
mutton	0.84***	0.36***
butter	0.49***	0.27***
eggs	-0.09	-0.13**
cattle	0.49***	0.15*
sheep	0.90***	0.31***
Total	0.96***	0.36***

#### 4.6.2 External Validity: Consumer Price Indices

As discussed above, De Zwart (2011) calculated consumer price indices (CPIs) based on a ‘bare bones’ and a ‘respectable’ basket of goods for consumption. The two CPIs are available only at an annual frequency, and differ from the total repeat sales commodity price index in terms of the underlying prices, composition, weighting and construction method (Klovland, 2014). Nevertheless, it is of some interest to compare the indices to see whether they provide the same basic picture of price trends in the Cape Colony.

Figure 4.17 compares the annualised total repeat sales and median indices to the bare bones and respectable basket CPIs. There is a large deviation between the two new total price indices and the two CPIs. The CPI based on the respectable consumption basket is the most similar to the total repeat sales index in that it is derived from a subset of the prices reported in the Blue Books.

The correlations in levels and in growth rates among the indices are reported in Table 4.8. Although there is a significant positive correlation between the repeat sales index and the respectable CPI in levels, the correlation in growth rates is not significant. For comparison, the correlation between growth rates in the annual CPI and PPI in South Africa between 1993 and 2015 was around 0.6. The correlation between the total repeat sales index and the bare bones CPI is not significant, perhaps due to the differences in the composition, weighting and construction method.



Historical Commodity Prices

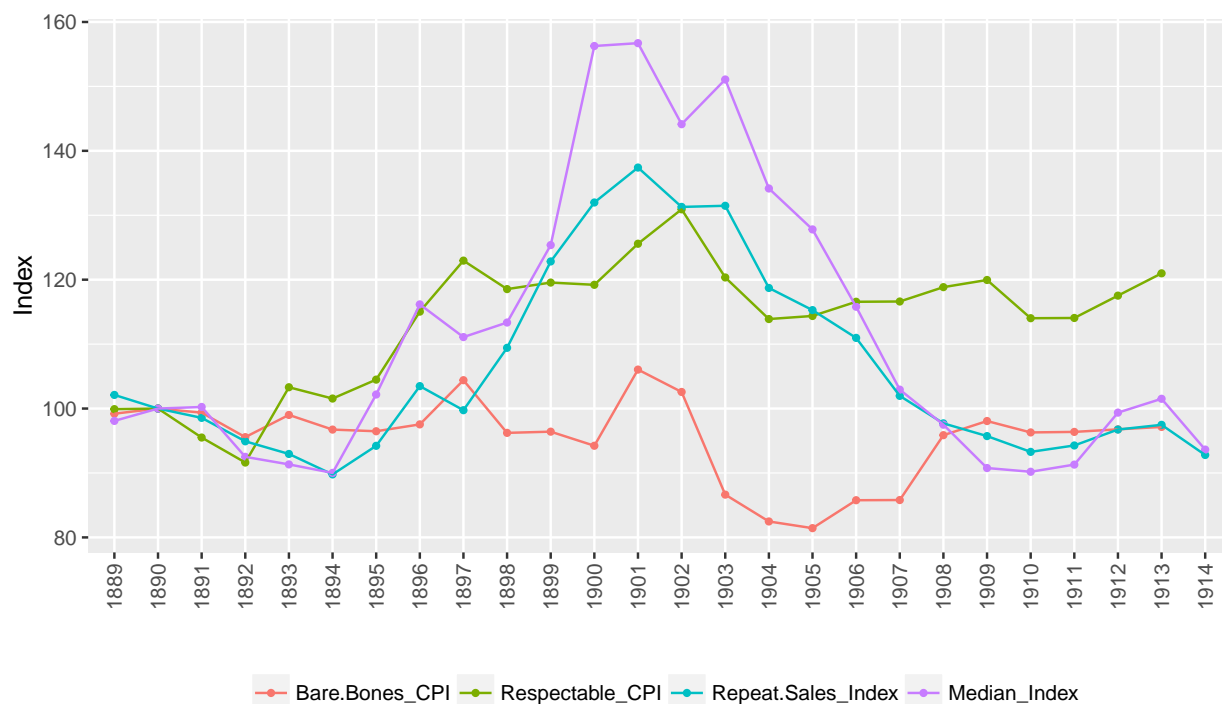


Figure 4.17: Comparing the total price index to the CPIs for the Cape Colony

Table 4.8: Correlations between the price indices

Levels	Bare Bones CPI	Respectable CPI	Repeat Sales Index
Respectable CPI	0.00		
Repeat Sales Index	-0.15	0.57***	
Median Index	-0.16	0.56***	0.97***
Growth Rates	Bare Bones CPI	Respectable CPI	Repeat Sales Index
Respectable CPI	0.65***		
Repeat Sales Index	-0.06	0.20	
Median Index	-0.14	0.14	0.84***

In general, the total repeat sales index exhibits larger movements than the CPIs. Short-term cycles in wholesale price indices (WPIs) often exhibit more pronounced amplitudes than cycles in CPIs, because WPIs contain a larger share of producer goods (e.g. raw materials), which are often more sensitive to business cycles than consumer goods, due to the inelasticity of agricultural production (Klovland, 2014). Movements in retail price indices often exhibit a smaller amplitude or a greater stickiness than movements in WPIs, as they contain more constant cost elements, such as the costs of distribution and fuel, lighting and rent (Schumann, 1938).

### 4.7 Evaluation: The Money Supply

The total commodity price index is intended to provide a clearer picture of the inflation history of the Cape Colony over the sample period. According to Lucas (1980), an implication of the quantity

theory of money is that a change in the growth rate of the quantity of money induces an equal change in the rate of price inflation. Overall price inflation should therefore reflect developments in the money supply.

This section first reviews the developments in the monetary and banking system in the Cape Colony, which developed rapidly under the stimulus of the mining industry during this period (De Kock, 1924). The development of a more modern monetary and banking system, points to the rise of a more modern capitalistic economy in the Cape Colony, and the beginning of business cycles in the modern sense (Schumann, 1938).

In order to evaluate whether the new commodity price index is useful as a measure of inflation, it is compared to the path implied by available monetary aggregates. The caveat is that the data series are imperfect measures of these concepts. The total commodity index is not a perfect measure of commodity prices, let alone a broader measure of the price level, as it places disproportionate emphasis on agricultural products and raw materials. The available monetary aggregate (M1) is an imperfect measure of the broader concept of the money supply. As the GDP estimates are imprecise, the estimates of the money supply per unit of output are only indicative. Moreover, the effects of changes in the money supply may be moderated by prevailing fiscal policy and the government's balance sheet (Smith, 1988).

#### **4.7.1 The Monetary and Banking System in the Cape Colony**

Figure 4.18 illustrates the measures of the M1 money supply reported in Greyling and Verhoef (2017), calculated as currency (coins and paper money in circulation) plus deposits (private bank fixed deposits). According to this measure of M1, there were large changes in the money supply over the period, reflecting episodes of instability. The most important of these were the banking crisis of 1889-90, the smaller crisis of 1895-96, and the war period around 1899-1904. The movements in the money supply were related to developments in the monetary and banking system during this period. According to Schumann (1938), these developments closely influenced the nature of the business cycle, with the banks instrumental in amplifying the expansionary phases.

The rapid increase in currency and bank deposits in 1886-1889, followed by the exceptional decrease in 1890-1892, reflects the overextension and drastic contraction of credit before and after the banking crisis of 1889-90. A speculative boom resulted from the opening of the gold mines, which led to the overexpansion of credit and inflation (Schumann, 1938). These investments led to an increase in bank deposits and currency imports. Between 1887 and 1889, more than £4 million in specie was imported, nearly two-thirds of the total capital of the gold mines (Gilbert, 1933).

During the crisis most of the District Banks in the Cape got into financial difficulties. There were hardly any restrictions on banks and their note issues. In the aftermath of the crisis, seven of the eight District Banks were liquidated or amalgamated into the three Imperial Banks: Standard Bank,

## Historical Commodity Prices

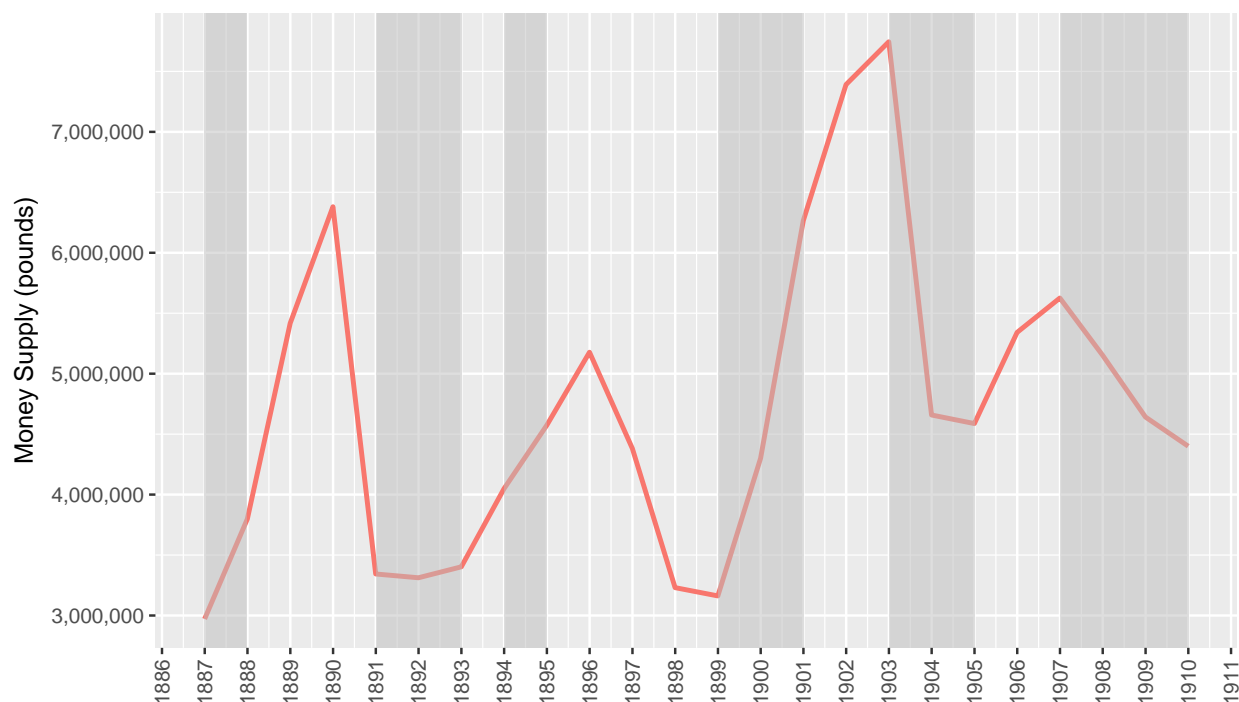


Figure 4.18: Money supply (M1) in the Cape Colony (1886-1909)

African Banking Corporation, and the Bank of Africa (De Kock, 1924). In this way the banking business of the Cape Colony was concentrated, with the exception of the Stellenbosch District Bank, in the hands of the three Imperial banks. The Colonial Banking Act was passed in 1891 to prevent similar banking crises and, at the same time, to secure a market for Government stock. According to the Act, the banks were to deposit government securities with the Treasury to the full amount of their note issue, the notes being legal tender and redeemable in gold. The total sum issued was limited to the amount of paid-up capital and reserves (De Kock, 1924).

After the banking crisis of 1889-90, there was again a period of considerable inflow of capital from abroad, in the form of large investments in the gold mining industry. A substantial proportion was directed through the banks in the Cape Colony, and in 1895, over £5 million in specie was imported. This was followed by a similar, albeit less severe, crisis in the latter part of 1895, during which large losses were sustained on gold shares (Gilbert, 1933).

There was an exceptional increase in both deposits and loans during the war period, especially in the Cape Colony, as well as vast specie imports (Schumann, 1938). The British Government required local labour and supplies to support their troops. Capital and merchandise imports continued for two years after the war, partly from borrowing for reconstruction and partly from the resumption of mining activities. There was a sharp decrease in the money supply after 1903, as accumulating debt servicing came to exceed new capital. The result was a large outflow of specie after 1903 (Gilbert, 1933). Imports of specie again occurred around 1906, but was curtailed by the local and

## Historical Commodity Prices

Table 4.9: Banking statistics in South Africa (in £'000s)

Year	Reserves	Currency	Deposits	Loans	Dep/Res	Dep/Cur
1891		4,473	12,295	9,820		2.75
1899	6,162	8,077	25,153	19,684	4.08	3.11
1902	8,814	11,472	47,222	34,985	5.36	4.12
1905	8,439	8,907	44,039	37,020	5.22	4.94
1907	8,217	7,968	37,791	30,801	4.60	4.74
1910		7,766	42,474	37,133		5.47

Sources: *De Kock (1924)*; *Schumann (1938)*

international recession in 1907 (Gilbert, 1933). The branch-banking system was firmly established and the concentration of banking resources increased further after the establishment of the Union in 1910 (Schumann, 1938). According to De Kock (1924), banking activities developed rapidly from 1910 to 1914, with large imports of specie.

Friedman and Schwartz (1963) discussed three major channels through which any change in the stock of money occurs. The first channel is high-powered money, which is the total of currency held by the public plus bank reserves. An increase in high-powered money leads to an equal percentage increase in the money supply. High-powered money in this case was constrained by the international gold standard. The amount of high-powered money changed through imports and exports of specie in order to produce a balance with other countries on the same standard (Friedman and Schwartz, 1963; Aghevli, 1975).

Changes in high-powered money were determined by the flow of international gold and an excess of gold production over consumption (Aghevli, 1975). The flow of international gold was determined by the balance on the current account and the flow of capital. Generally, gold flowed into the country through export surpluses, as illustrated in Figure 4.2, and large capital investments in the mines. Gilbert (1933) provided some evidence on the timing of capital inflows, by looking at capital applications in London from South African mines, new railroad miles completed annually, and imports of machinery into the Cape Colony. These measures reflect the cyclical movements in capital imports, with peaks between 1902 and 1904.

The large increases in the money supply generally corresponded to export surpluses, as well as an inflow of investment in the mines and remittances during the war. The large decreases in the money supply generally corresponded to increases in imports and outflows of investment funds. Gilbert (1933) argued that after 1905 the rapid growth of debt servicing and other invisible imports decreased merchandise imports and produced an export surplus that lasted until WWI. Thus, the large export surplus after 1905 did not translate into a large increase in the money supply.

Taken as a whole, this was a period of rapid expansion of the commercial banking system. During this period the Imperial banks opened a large number of new branches, not only in the Cape Colony, but throughout South Africa. Table 4.9 reports the banking statistics for all the banks in South Africa (in £'000s). Reserves and currency peaked in 1902, towards the end of the war.

The second channel through which a change in the stock of money may occur is the ratio of commercial bank deposits to reserves (Friedman and Schwartz, 1963). For a given amount of reserves, the higher this ratio, the larger the amount of deposits outstanding and the larger the money supply. The ratio of deposits to reserves depends on the willingness of the banks to create deposits relative to their reserves, on regulations, and on the amount of high-powered money available. According to Table 4.9, the ratio of deposits to reserves reached a peak in 1902 and declined towards 1907. The rise in the deposit-reserve ratio reflects the rise and maturation of the Imperial Banks, which found it profitable to operate with lower reserves relative to deposits.

The third channel is the ratio of commercial bank deposits to currency held by the public. The higher this ratio, the larger the portion of high-powered money that is used as bank reserves, and the larger the money supply. The ratio of deposits to currency is determined by the public, government (the legal conditions under which currency and deposits may be issued) and the banks (the services and interest offered to depositors) (Friedman and Schwartz, 1963).

According to Table 4.9, the ratio of deposits to currency increased throughout the period. At the outset, the public held £2.75 of deposits for each pound of currency. This increased to almost £5.5 by 1910. The low initial value probably helped to produce and intensify the banking crisis of 1889-90. In general, banking difficulties are followed by a decrease in the ratio of deposits to currency, as the public seeks to convert the one into the other (Friedman and Schwartz, 1963). The rise in the ratio after 1890 probably reflected the spread of the branch-banking system, the more widespread availability of checking facilities, and the greater usefulness of bank deposits. Another important factor was probably the growing real income per capita over the period (De Zwart, 2011).

#### **4.7.2 Money Supply and Inflation**

In this section, these changes in the money supply are compared to the total commodity price index, in order to gauge if it provides a plausible history of inflation for the Cape Colony over this period. Figure 4.19 illustrates the standardised versions of the annual total commodity price index compared to the measure of M1 money supply, as well as money supply per unit of output. The money supply per unit of output provides a simple gauge of the implied inflation history, based on the quantity theory of money, given the tenuous assumption that the velocity of money was constant over the period (Schumann, 1938; Friedman and Schwartz, 1963). According to Friedman (1952), inflation occurs when the money supply increases more rapidly than output. The annual total commodity price indices seem to reflect the developments implied by fluctuations in the money supply.

Table 4.10 reports that the correlations between the growth rates in the measures of money supply and the measures of inflation. The correlations between the growth rates in the money supply per unit of output and the repeat sales and median indices of total commodity prices are moderately significant. It should be borne in mind that these are rough measures, and one would not expect a one-to-one relationship. Although the correlations with the money supply are not significant,

## Historical Commodity Prices

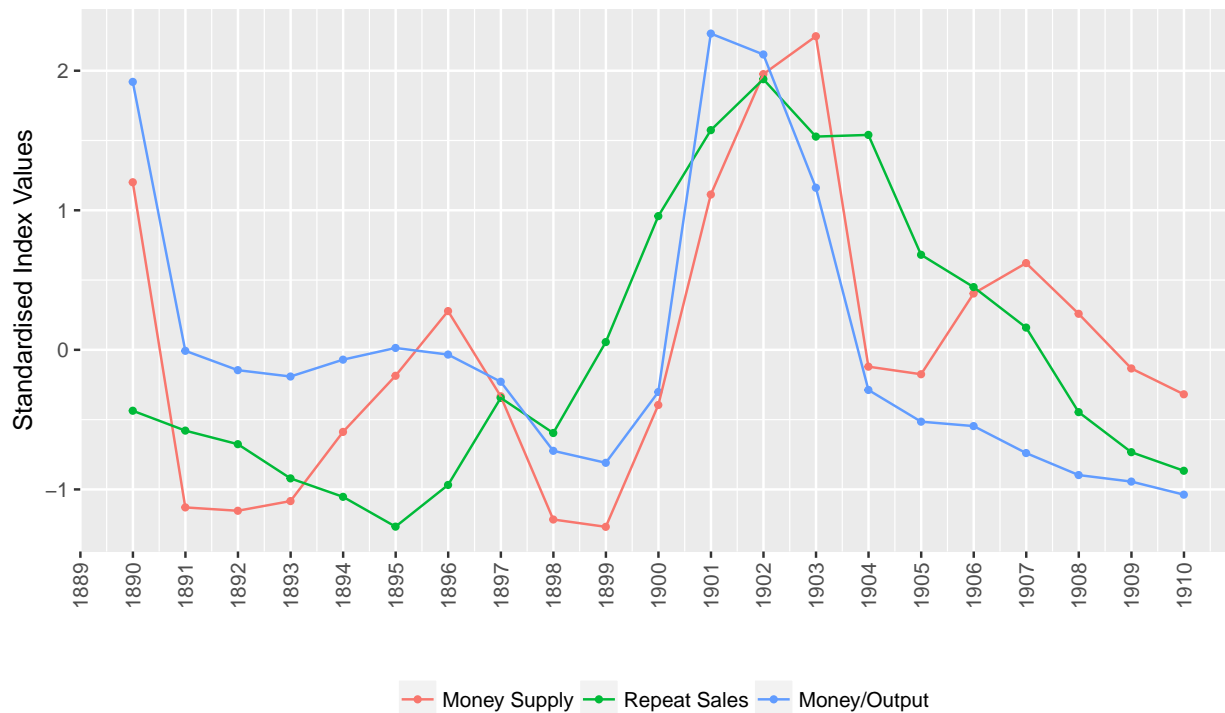


Figure 4.19: Comparing the total commodity price index to the monetary aggregates

they at least illustrate a similar picture and a positive relationship. The total commodity price indices seem to follow developments in the money supply more closely than the existing measures of CPI (which were calculated to estimate the cost of living), and may provide a more accurate and comprehensive picture of the inflation history of the period.

Table 4.10: Correlations between the price indices and monetary aggregates (growth rates)

	Bare Bones CPI	Respectable CPI	Repeat Sales	Median Index
Money Supply	0.20	0.18	0.27	0.19
Money/Output	0.14	0.12	0.44*	0.49**

The caveat in using the total commodity price index as a gauge of inflation is that the products included are primarily agricultural goods and raw materials. As such, they omit some important non-agricultural products and services. Nevertheless, the total commodity price index does include many more products than the CPIs, which are based on the annual prices of 8 and 10 products respectively. The total commodity price index may therefore be useful for future studies that require a measure of inflation for the Cape Colony.

In addition to providing a clearer picture of demand and supply conditions and the inflation history of the period, the high-frequency price records may be used to investigate market integration in the Cape Colony. This chapter therefore explores further methods to aggregate the price records, in order to create indicators that may be used to investigate internal market integration in the Cape

Colony around the turn of the 20th century. The following section turns to the literature on market integration.

## 4.8 Market Integration

There is a large international literature on market integration, but very few studies have considered internal market integration in the Cape Colony around the turn of the 20th century (e.g. Boshoff and Fourie (2017)). During this period the mineral revolution led to the rapid expansion of the railway network and created a potentially large internal market (Schumann, 1938; Herranz-Loncán and Fourie, 2017). The following section comprises a brief discussion on the history of railway development in the Cape Colony as a potential driver of internal market integration.

### 4.8.1 Development of the Railways

For most of the 19th century, transport between towns in South Africa was basic and expensive, due to the large distances between towns and the rugged terrain between the coast and the interior (Fourie, Grundlingh and Mariotti, 2017). The ox wagon was the main form of freight transport between towns. The mineral revolution created an urgent need for an efficient means of transporting goods and passengers between the coastal ports and the mining towns in the interior. The result was a rapid expansion of the railway network and a decline in transportation costs (De Kock, 1924).

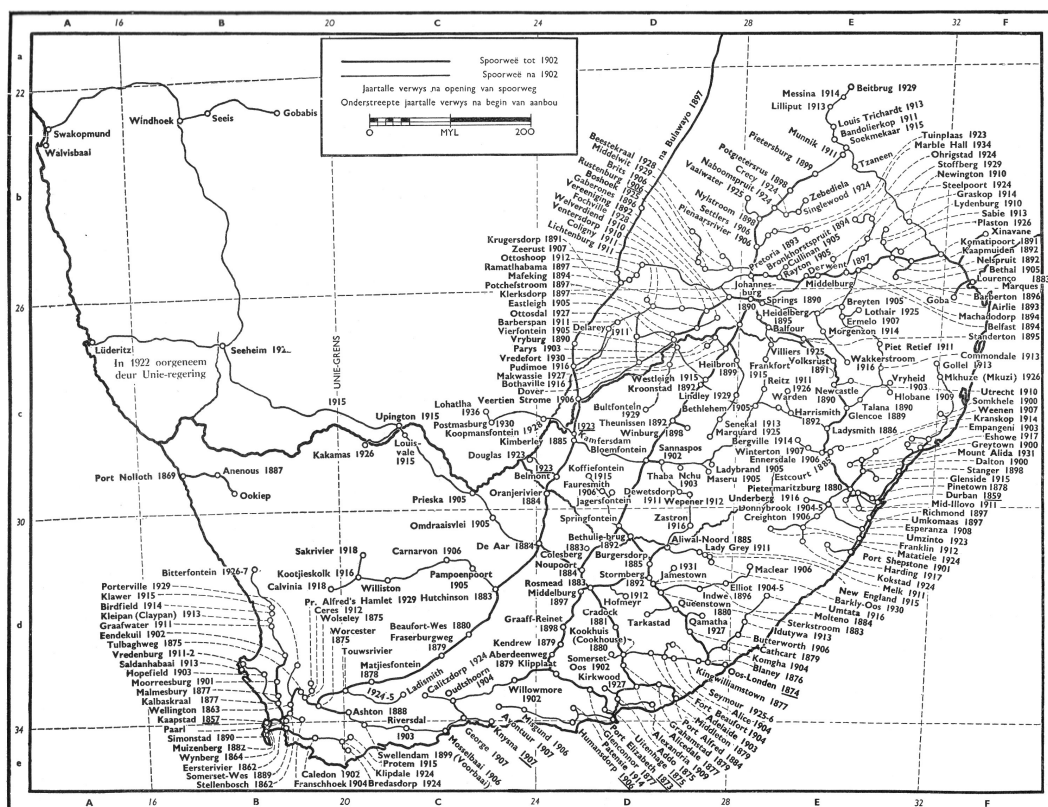
The first railway lines were built mainly to connect the ports of Cape Town, Port Elizabeth and East London with the diamond fields in Kimberley and later with the gold mines in the Witwatersrand (Herranz-Loncán and Fourie, 2017). The railway line from Cape Town to Kimberley was completed in 1885. Railway lines from Kimberley to Port Elizabeth, and to the gold mines were completed in 1892. The railway network was used to transport diamonds and gold to the coast, as well as to transport produce, provisions and large mining machinery to the growing towns in the interior (Fourie, Grundlingh and Mariotti, 2017). By reducing the cost of transport to the interior, the railway eased the movement of foodstuffs, labour, and capital goods between regions. The completion of these main railway lines between the coastal towns and the mining towns served to ‘integrate transaction flows’ in an unprecedented way (Greyling and Verhoef, 2015).

Prior to the war, the Cape Government had been concerned primarily with constructing main lines and developing main-line traffic. The Cape Colony’s railway network initially consisted of three main lines connecting Cape Town, Port Elizabeth and East London with the diamond-producing region in Kimberley. Access was limited to towns through which the main railway lines passed. Towns further from the main lines had to rely on transport riders, who were expensive over long distances (Herranz-Loncán and Fourie, 2017).

A second period of rapid expansion of the railway lines occurred just after the war. Investment focused on branch lines to act as feeders to the main lines or as connections between them (e.g. the



## Historical Commodity Prices



78. ONTWIKKELING VAN SPOORWEE  
 Source: *Geskiedenis Atlas* (1948)

Figure 4.20: Development of the railways in South Africa

route between Cape Town and Port Elizabeth) (Fourie, Grundlingh and Mariotti, 2017; Herranz-Loncán and Fourie, 2017). A number of branch and subsidiary main lines were built to develop the agricultural prospects of the Colony (Schumann, 1938).

Figure 4.20 illustrates the development of the railways in South Africa. The dates indicate the opening of the railway line at each town (underlined dates indicate the start of construction). The thicker lines denote railway lines completed before 1902, while the thinner lines denote lines completed after 1902. Between 1902 and 1910, 1,194 miles of new railway lines were opened by the Cape Government, compared with 244 miles during the previous ten years, from 1892 to 1901 (De Kock, 1924). By 1910, the railway network in the Union was by far the largest and densest network in Africa (Herranz-Loncán and Fourie, 2017).

The railway was seen as an instrument for the economic development of the different areas of the Colony. Herranz-Loncán and Fourie (2017) found that the development of the railway made a large contribution to economic growth in the Cape Colony. According to their estimates, 22%-25% of the growth in labour productivity in the Cape Colony between 1873 and 1905 derived directly from the railway. The railway was instrumental for the expansion of the mining districts, as well



as the development of the Western region of the Cape Colony, as its production could expand in response to increased domestic demand and market integration. The railways were used mainly to carry freight, although passenger traffic was also sizeable. Railway freight was dominated by the transport of domestic produce and imports to the internal markets. Herranz-Loncán and Fourie (2017) reported that in 1908 the biggest categories in terms of total freight revenues were ‘general products’ (36%), i.e. manufacturing goods, and agricultural goods (25%).

Presumably, internal market integration was aided by the expansion of the railway network during this period. Theory suggests that integrated markets generate important economic benefits, such as a comparative advantage from greater regional specialisation, economies of scale, increased investment opportunities, and greater incentives for technical progress and innovation (Andrabi and Kuehlwein, 2010).

While the development of the railways was no doubt important, it was not the only factor that could have aided market integration (Andrabi and Kuehlwein, 2010). There were also a number of significant changes in technology and infrastructure, such as telegraphs and roads, occurring before and during this period which could have contributed to price convergence. Better information sharing throughout the country may have alerted producers to large price differences of which they could take advantage. The unification of the four colonies probably also contributed, through institutional changes such as the elimination of internal tariffs and the use of a common currency.

In all likelihood, therefore, market integration was aided by the development of the railway network, as well as technological developments and unification. There was a large expansion in railway lines just before the sample period, and again after the war, with significant investment in branch lines to act as connections between the main lines.

#### **4.8.2 Market Integration in the Cape Colony**

Cournot characterised an integrated market as “an entire territory of which the parts are so united by the relations of unrestricted commerce that prices take the same level throughout with ease and rapidity” (cited in Federico (2012)). In other words, an integrated market must fulfil two conditions: equilibrium prices should be equal after trading costs (i.e. the law of one price should hold), and prices should return quickly to this equilibrium after a shock (i.e. there should be a degree of efficiency) (Chilosi *et al.*, 2013). In the parlance of Federico (2012), there should be price convergence and market efficiency.

The first dimension concerns the convergence of prices for the same goods throughout the economy. Market integration along this dimension features a decline in price gaps or price dispersion. Decreasing transportation costs should lead to lower price dispersion and more integrated markets (Andrabi and Kuehlwein, 2010). The second dimension concerns the efficiency with which price gaps return to equilibrium after a shock. This reflects growing market efficiency, especially through

better circulation of information about prices and fundamentals (Federico and Sharp, 2013). These two conditions are tested using different techniques. Trends in price gaps or dispersion are tested by looking at pairwise trends in relative prices or, for a large number of markets, at trends in the coefficient of variation. Tests of efficiency typically involve time-series comovement and cointegration methods (Federico and Sharp, 2013).

There is a growing literature on market integration in economic history (see Federico (2012) for a review of the literature). Overall, the international literature finds substantial convergence in commodity prices in several economies in the late 1800s and early 1900s, and that railway development was one of the main causes of this integration (Andrabi and Kuehlwein, 2010). Railways improved market integration in many regions, including Europe (Federico, 2012), the US (Federico and Sharp, 2013) and India (Donaldson, 2012). Andrabi and Kuehlwein (2010), however, showed that not all integration was necessarily a consequence of railway development and that commodity prices in India were converging almost as rapidly among districts without railways as they were among districts linked by railway lines.

To date there has been little research on regional market integration in the Cape Colony around the turn of the 20th century. Boshoff and Fourie (2017) showed that product markets in the Cape Colony were fairly well integrated by the turn of the century. They found that Western Cape agricultural markets were quite well integrated internally, as evidenced by the significant correlations among prices in Malmesbury, Clanwilliam, Worcester and Cape Town. The wheat and maize markets of major towns located on the main railway lines were integrated with those of Cape Town, particularly those of Beaufort West and Colesberg and the key market of Kimberley. They also found that Eastern Cape wheat and maize markets were on the whole less well integrated with Cape Town markets, although the towns with good railway links, such as Aliwal North, Cradock, Dordrecht and King William's Town, did have significant connections with the market in Cape Town. They estimated a long-term equilibrium relationship, using the bounds test method suggested by Pesaran, Shin and Smith (2001). The results confirmed that Eastern Cape wheat markets were not integrated with Cape Town's but that there was a long-term relationship between Kimberley's wheat and maize prices and those of Cape Town.

Their study of market integration was complicated by the limited availability of data. For instance, the correlations were based on monthly data for the period 1897-1906 for just four products, while the bounds tests could be applied only to the prices in five towns. Moreover, correlations and cointegration tests concern only the second dimension of market integration, i.e. market efficiency. The work in this part of the chapter builds on that of Boshoff and Fourie (2017), by considering the first dimension of market integration, i.e. price convergence, and expanding the period and number of products analysed along the second dimension.

In examining these two dimensions, this part of the chapter employs two aggregation methods. The first is to calculate measures of cross-sectional price dispersion among the towns in the Cape Colony, which are used to investigate price convergence. The second is to calculate repeat sales commodity

price indices at the regional level, segmenting the data by region, which are used to investigate market efficiency. Time-series techniques are then used to test for evidence of market integration. The hypotheses that price dispersion between towns was declining over the period and that regional price indices were cointegrated are tested.

### 4.8.3 Methodology: Tests of Market Integration

The literature divides the tests of market integration into those for price convergence and those for market efficiency. Tests of price convergence typically involve trends in price gaps or, for a large number of markets, trends in the coefficient of variation (e.g. Klovland (2005) and Chilosi *et al.* (2013)). In order to avoid the problem of choosing a satisfactory threshold for price differences, trends in price differences over time are evaluated (Federico, 2012). Tests of market efficiency typically involve time-series correlation and cointegration methods (e.g. Federico and Sharp (2013) and Boshoff and Fourie (2017)). In some cases the tests of the two dimensions might provide different results. This is unsurprising, as they refer to different aspects of integration (Federico, 2012). For example, if the absolute prices for a specific product converge quickly between two towns, the tests of efficiency may not exhibit significant comovement over time.

According to Razzaque, Osafa-Kwaako and Grynberg (2007), the most appropriate way of evaluating trends in commodity prices is to examine price behaviour at the individual commodity level. In this chapter, tests are carried out for a number of commodities, as opposed to focusing on wheat, which is the most common commodity analysed in the literature (Klovland, 2005). Federico (2012) found substantial differences between patterns of convergence for different commodities in the first half of the 19th century in Europe, and concluded that one should be cautious in accepting the results for wheat as being representative for all commodities. This section presents the aggregation methods and the time-series tests used to investigate market integration in the Cape Colony.

#### 4.8.3.1 Tests of price convergence

Most of the literature on market integration is based on the idea that when transaction costs are small and information is shared, price differences are quickly arbitrated away, and prices converge (Chilosi *et al.*, 2013). In a competitive equilibrium, price differences are equal to trading costs. If trading costs are smaller than price differences, traders will arbitrage away the differential to make a profit. Thus, arbitrage should ensure that homogeneous commodities sell for one price, after taking trading costs into account (Chilosi *et al.*, 2013). One way to test for market integration is to measure the extent to which prices of the same commodities converge over time (O'Rourke and Williamson, 2004). Price convergence among locations can be examined graphically, or estimated more formally with a regression of price trends (Federico, 2012).

This chapter undertakes direct tests for this type of market integration by testing for price convergence among the towns in the Cape Colony. Two aggregate measures of cross-sectional price dispersion

among the towns in the Cape Colony are calculated, both of which are related to the measures of the first and second cross-sectional moments of the distribution used in Chapter 3.

The first is based on the measure of absolute deviations suggested by Klovland (2005), which examines the development of the annual absolute deviations over time. For a group consisting of  $K$  towns, the absolute deviation (AD) for commodity  $i$  is defined as:

$$AD_{it} = \frac{1}{K} \sum_{k=1}^K Abs(1 - P_{kit}/\mu_{it}),$$

where  $P_{kit}$  is the price of commodity  $i$  in town  $k$  month  $t$ , and  $\mu_{it}$  is the cross-sectional average price of commodity  $i$  in month  $t$  across all the towns  $\mu_{it} = \frac{1}{K} \sum_{k=1}^K P_{kit}$ .

$AD_{it}$  is therefore a measure of the average absolute deviation for a specific commodity across all of the towns over time. The absolute value is used to avoid positive and negative price gaps averaging out. This measure reflects the extent of deviations from the law of one price over time. If the price gaps for a specific product between towns are large, the index will show relatively high values. Consider two extreme cases. If the law of one price held exactly for all the towns,  $AD_{it}$  would equal zero, whereas if price gaps were 100%, the measure would equal 1. A decreasing absolute deviation over time implies price convergence and increasing market integration (Klovland, 2005).

The second measure looks at dispersion by computing the cross-sectional coefficient of variation (CV), i.e. the standard deviation normalised by the mean. The CV is a popular measure for investigating price convergence in the literature (e.g. see Federico (2012), Federico and Sharp (2013), Chilosì *et al.* (2013)). The coefficient of variation (CV) for commodity  $i$  is defined as:

$$CV_{it} = \sigma_{it}/\mu_{it},$$

where  $\mu_{it}$  again is the average price across towns, and  $\sigma_{it}$  is the cross-sectional standard deviation of prices across towns  $\sigma_{it} = \sqrt{\frac{1}{n-1} \sum_{k=1}^K (P_{kit} - \mu_{it})^2}$ . A higher CV implies that prices are more dispersed among the towns. A decreasing CV over time implies price convergence and increasing market integration.

To test for price convergence, it is therefore necessary to test the AD and CV series for a decreasing trend. To test this formally, the trend in these two measures can be explored by running the following log-linear regression on either the AD or CV series:

$$\ln CV_t = \alpha + \beta TIME + \epsilon_t$$

A negative and significant coefficient for the trend variable  $TIME$  is evidence of convergence and of integration (Federico, 2012). This model assumes that the process is trend-stationary. Unit root tests indicate that this is the case for most of the AD and CV series.

A number of authors (e.g. Bleaney and Greenaway (1993), Razzaque, Osafa-Kwaako and Grynberg

(2007), Federico (2012)) have estimated error-correction models to take into account the possibility that a series follows a difference-stationary process:

$$\Delta \ln CV_t = \alpha + \beta TIME + \psi \ln CV_{t-1} + \phi \Delta \ln CV_{t-1} + \epsilon_t$$

This model is in the form of the Augmented Dickey-Fuller unit root test, which encompasses both the trend and difference stationary models (Razzaque, Osafa-Kwaako and Grynberg, 2007). The lagged dependent variable is included to take account of potential serial correlation.

There are four possibilities (Bleaney and Greenaway, 1993). If  $\beta = 0$  and  $\psi = 0$ , the series has a random walk with zero mean. If  $\beta = 0$  and  $-1 < \psi < 0$ , the series has no long term but tends to be pulled back towards its historic mean. If  $\beta \neq 0$  and  $\psi = 0$ , the series has a random walk with drift, so that if  $\beta$  is negative, it is more probable that it will be less than its current value. If  $\beta \neq 0$  and  $-1 < \psi < 0$ , the series has a non-zero deterministic trend to which it reverts after a shock. The long-run trend rate can be computed as  $t = -(\beta/\psi)$ . A negative trend implies long-run convergence and integration (Federico, 2012).

#### 4.8.3.2 Tests of market efficiency

Prices in markets that are integrated should move together over time. The intuition is that there is a limit to how far prices can deviate before either demand-side or supply-side substitution forces them back in line. Time-series methods, such as correlations and cointegration, can be used to investigate how closely prices move together over time (Klovland, 2005). According to Federico (2012), the measures of efficiency can be classified into three types. The first type is comovement or correlation tests, whereby arbitrage forces prices to move together over time. The second is cointegration tests, whereby arbitrage forces price differentials to return to their equilibrium level after a shock. The third is variance tests, whereby arbitrage reduces the effect of local shocks and thus the volatility of prices. Following Boshoff and Fourie (2017), this chapter provides a study of correlation and cointegration to test for short- and long-term relationships among the commodity prices in the towns of the Cape Colony.

The difficulty in applying these tests in this case is that the price series for the individual towns have large gaps. This chapter therefore splits the data into geographical regions, and uses the repeat sales method to calculate indices for these regions. By combining the price records for each product in a number of towns, more complete price indices can be estimated. The time-series tests therefore focus on regional commodity price indices. In making a regional classification, there is the trade-off between including enough towns to create a reasonably complete and accurate index for each commodity, and in making the regions small enough to be reasonably well integrated internally. As the records are incomplete, it is useful to include the information from all of the towns in mutually exclusive regions that make economic sense.

As discussed above, Cape Town, Port Elizabeth, and East London, were the three main international

ports, with three main lines running to the diamond fields in Kimberley. The diamond fields played a central role in the railway system. In 1905, for instance, apart from Cape Town, Port Elizabeth, and East London, Kimberley was the station that generated by far the highest revenue (Herranz-Loncán and Fourie, 2017).

The sample is therefore split into four geographic regions: towns in the present day Western Cape along the Cape Town main line<sup>41</sup>; towns in the present day Eastern Cape along the Port Elizabeth main line<sup>42</sup>; towns in the Eastern Cape along the East London main line<sup>43</sup>; and towns in the present day Northern Cape<sup>44</sup> (chiefly Kimberley).<sup>45</sup> The correlation and cointegration tests are carried out on pairwise combinations of the regional price indices.

An implicit assumption is that the towns in each regional grouping are already reasonably well integrated. To get a rough idea of whether this assumption is valid, the average ADs and CVs for the different regions are calculated, for all periods and commodities. Table 4.11 reports the average ADs and CVs for different regions. The towns in the Western Cape exhibited lower dispersion than the Colony as a whole. An alternative grouping would be to divide the Western Cape region into smaller regional groupings (e.g. only towns around Cape Town), which exhibit lower price dispersion. However, towns further away with relatively complete information (e.g. Beaufort West) are included in the Western Cape region in order to improve coverage, even if this increases internal price dispersion somewhat. The towns on the Port Elizabeth and East London trunk lines exhibit lower variation than the towns in the Eastern Cape as a whole. The East London values are relatively high because of the inclusion of Queenstown. However, because Queenstown is a close neighbour to East London and to improve coverage, its information is included in this regional grouping. The smaller group of towns in the Northern Cape exhibited the lowest variation.

There is lower price dispersion, on average, for the towns within each regional grouping than for the Cape Colony as a whole. The assumption that the regions were relatively well integrated seems to hold reasonably well, albeit to varying degrees for the regions. Alternative regional grouping do not change the results qualitatively. The main trunk railway lines, which determine this regional grouping, were completed before or right at the beginning of the sample period. Nevertheless, there is the caveat that within-region integration probably increased over the period.

If two non-stationary time series have a linear combination that is stationary, they are said to be

<sup>41</sup>The Western Cape towns are Beaufort West, Bredasdorp, Caledon, Cape Town, Ceres, Clanwilliam, George, Knysna, Ladismith, Malmesbury, Mossel Bay, Oudtshoorn, Paarl, Piquetberg, Prince Albert, Riversdale, Robertson, Stellenbosch, Swellendam, Tulbagh, Uniondale, Worcester, Vanrhynsdorp, and Wynberg.

<sup>42</sup>The Eastern Cape towns along the Port Elizabeth main line are Albany, Cradock, Graaff-Reinet, Grahamstown, Humansdorp, Middelburg, Port Alfred, Port Elizabeth, Somerset East, Uitenhage, and Willowmore.

<sup>43</sup>The Eastern Cape towns along the East London main line are Albert, Aliwal North, Burghersdorp, Cradock, Dordrecht, East London, King William's Town, Middelburg, Queen's Town, Tarkastad, Mount Currie, Kokstad, and Umtata.

<sup>44</sup>The Northern Cape towns are Colesberg, Kimberley, Philipstown, and Richmond.

<sup>45</sup>An analysis that distinguishes between the towns along the main and branch railway lines was prevented by the lack of price observations for the towns on the branch lines. Such an analysis may be undertaken as new price records are added in future.

## Historical Commodity Prices

Table 4.11: Price dispersion in each region

Region	Average ADs	Average CVs
Cape Colony	0.17	0.24
Western Cape towns	0.10	0.20
Eastern Cape towns	0.15	0.23
East London main line	0.13	0.22
Port Elizabeth main line	0.08	0.19
Northern Cape towns	0.03	0.16

cointegrated (Engle and Granger, 1987). The concept of cointegration is useful in studies of market integration (Klovland, 2005). This chapter employs the Johansen test for cointegrating relationships, which studies the properties of the long-run matrix of a VAR to test whether two or more price series exhibit stable long-term relationships (Johansen, 1991).

Assume that we know that all of the  $n$  elements of vector  $y_t$  are  $I(1)$ . The Johansen test is based on a vector error-correction model (VECM) of the following form:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^k \alpha_i \Delta y_{t-i} + \epsilon_t,$$

where  $y_t$  is a vector of price indices in this case;  $k$  is the number of lags, selected by means of the AIC, SC and HQ; and  $\epsilon_t$  is a white noise error term.

If the variables are cointegrated, the matrix  $\Pi$  can be written in terms of a matrix of adjustment parameters  $\alpha$  and a matrix of cointegrating vectors  $\beta$ :  $\Pi = \alpha\beta'$ . If  $\Pi$  is a matrix of zeros, the variables are not cointegrated and the relationship reduces to the VAR in first differences:  $\Delta y_t = \sum_{i=1}^k \alpha_i \Delta y_{t-i} + \epsilon_t$ .

One way to test for this is to test the rank of the  $\Pi$  matrix. If the variables are cointegrated, the rank will equal the number of cointegrating vectors. The number of cointegrating vectors is less than or equal to the number of variables  $n$  and strictly less than  $n$  if the variables have unit roots. If the rank is less than the number of rows and columns in the  $\Pi$  matrix, then one or more eigenvalues is zero.

The Johansen procedure consist of a maximum eigenvalue test and a trace test. Both of these test the null hypothesis of no cointegration against the alternative of cointegration. The maximum eigenvalue test examines whether the largest eigenvalue is zero relative to the alternative that the next largest eigenvalue is zero. The eigenvalues are ordered by size,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . The first test is a test whether the rank of the matrix  $\Pi$  is zero. The null hypothesis is that  $rank(\Pi) = 0$ , while the alternative hypothesis is that  $rank(\Pi) = 1$ . If the largest eigenvalue  $\lambda_1 = 0$ , then the rank of  $\Pi$  is zero and there are no cointegrating vectors. If  $\lambda_1 \neq 0$ , then the rank of  $\Pi$  is greater than or equal to one and there is at least one cointegrating vector. If  $\lambda_{n-1} \neq 0$ , then the test considers whether  $\lambda_n = 0$ . If  $\lambda_n = 0$ , then there are  $n - 1$  cointegrating vectors. If  $\lambda_n \neq 0$ , the variables do not have unit roots.



The trace test examines whether the rank of the  $\Pi$  matrix is  $r$ , relative to the alternative that there are more than  $r$  cointegrating relationships. The null hypothesis is that  $\text{rank}(\Pi) = r$ , i.e. that there are at most  $r$  cointegrating relationships. The alternative hypothesis is that  $r < \text{rank}(\Pi) \leq n$ , where  $n$  is the maximum number of possible cointegrating vectors, i.e. there are more than  $r$  cointegrating relationships. It therefore tests whether the sum of the smallest  $n - r$  eigenvalues is close enough to zero to not reject the null. If this null hypothesis is rejected, the next null hypothesis is that  $\text{rank}(\Pi) = r + 1$  and the alternative is that  $r + 1 < \text{rank}(\Pi) \leq n$ . Testing proceeds as for the maximum eigenvalue test.

In this chapter the trace tests are applied to pairwise regional price indices. Thus, the maximum number of cointegrating vectors is two. If  $\lambda_1 = 0$  then there are no cointegrating vectors. If  $\lambda_1 \neq 0$  and  $\lambda_2 = 0$  then there is one cointegrating vector. If  $\lambda_1 \neq 0$  and  $\lambda_2 \neq 0$  then the variables do not have unit roots.

To determine whether market efficiency increased over the period, the sample is divided into two periods: the period before the war, i.e. from September 1889 to September 1899, and the post-war period, i.e. from June 1902 to July 1914. The second period corresponds with the second rapid expansion in railway lines. The war-time period is excluded to avoid the possibility of overstating integration, as the regions were subject to the same exogenous shock (Federico, 2012). The efficiency test are applied for the two periods separately, as well as for the entire sample period. One would expect efficiency between the regions to improve in the post-war period.

#### 4.8.4 Market Integration Results

This section presents the aggregate measures and the results of the tests for price convergence and market efficiency. The tests for convergence look at price gaps and dispersion across all the towns in the Cape Colony over time. The tests of efficiency examine correlations and cointegrating relationships between regional price indices, and distinguish between the periods before and after the war.

##### 4.8.4.1 Price convergence

Figure 4.21 and Figure 4.22 illustrate the monthly AD and CV measures, respectively, for selected commodities. The two measures provide very similar evidence, which is reflected in their high correlations, reported in Table 4.12. The absolute deviations and coefficients of variation are relatively low for these commodities. For comparison, Federico and Sharp (2013) found that the coefficient of variation for wheat prices in European markets declined from a peak of slightly below 0.4 during the Napoleonic wars to about 0.10-0.15 in the heyday of free trade in the 1870s. In this sample, the CV for wheat was around 0.15 over the period.

Both measures suggest a declining trend over the sample period. To test this more formally, the



### Historical Commodity Prices



Figure 4.21: Absolute deviations over time for selected commodities



Figure 4.22: Coefficient of variation over time for selected commodities

## Historical Commodity Prices

Table 4.12: Correlations between ADs and CVs

	Levels	Growth
wheat	0.93***	0.86***
wheat.flour	0.84***	0.86***
boer.meal	0.93***	0.81***
mealies	0.88***	0.82***
mealie.meal	0.94***	0.88***
barley	0.82***	0.83***
oats	0.77***	0.72***
oathay	0.94***	0.80***
potatoes	0.94***	0.89***
tobacco	0.89***	0.68***
beef	0.93***	0.86***
mutton	0.89***	0.80***
butter	0.91***	0.84***
eggs	0.91***	0.85***

simple model (Model 1), as well as the extended error-correction model (Model 2) were estimated using both measures for 14 commodities.<sup>46</sup> For the most part, the diagnostic tests for normality and homoskedasticity, and no serial correlation are satisfied. The lagged dependent variable in levels  $\psi$  was negative and significant in all cases (using the Dickey-Fuller distribution), except for that of boer meal.

Table 4.13 reports the coefficients for the time trend variable  $\beta$  for Model 1 and the long-run trend rate  $t = -(\beta/\psi)$  for Model 2. There is strong evidence of convergence for wheat flour, mealies, oats, oat-hay, tobacco, butter, and eggs, as both models point to a significant negative trend. There is weaker evidence of convergence for wheat, potatoes, beef and mutton, in the sense that the simpler models point to a decreasing trend. The commodities for which there is no convergence over the period, mealie meal and barley, already exhibited low values of price dispersion, and the results imply that their prices revert towards the historic mean.

The significant monthly trend rates imply annual growth rates that range between -0.48% and -3.67% per annum for Model 1 and between -0.53% and -1.12% per annum for Model 2. In similar exercises, the estimates in Federico and Sharp (2013) range between -0.5% and -1.75% per annum, and those in Razzaque, Osafa-Kwaako and Grynberg (2007) range between -0.8% and -1.4% for one dataset and -0.92% and -3.27% for another. Boshoff and Fourie (2017) did not analyse this dimension of market integration.

Overall, there is relatively strong evidence of price convergence among the towns in the Cape Colony. Table 4.14 reports that 12 of the 14 commodities exhibited a significant declining trend in price dispersion over the sample period when using the first specification (Model 1). When using the stricter model specification (Model 2), 7 (according to the ADs) and 6 (according to the CVs) of the 14 commodities exhibited a significant declining trend in price dispersion. This implies that market integration along this dimension increased over the sample period.

<sup>46</sup>The following products were excluded because of limited data: lucerne hay, bread, pigs, horses, cattle and sheep.

## Historical Commodity Prices

Table 4.13: Coefficients of time trends

	Absolute Deviation		Coefficient of Variation	
	Model1	Model2	Model1	Model2
wheat	-0.00075***	-0.00064	-0.00079***	-0.00067
wheat.flour	-0.00201***	-0.00209**	-0.0004**	-0.00033
boer.meal	-0.00194***	-0.00233	-0.00127***	-0.00159
mealies	-0.00254***	-0.00217***	-0.00167***	-0.00123**
mealie.meal	-0.00015	0.00063	-0.00031	0.00065
barley	-0.00028	0.00036	-0.00004	0.00054
oats	-0.00185***	-0.00186**	-0.00149***	-0.00147**
oathay	-0.00268***	-0.00217***	-0.00217***	-0.00173***
potatoes	-0.00124***	-0.00133	-0.00063***	-0.00087
tobacco	-0.00182***	-0.00177**	-0.00202***	-0.00204***
beef	-0.00132***	-0.00113	-0.00101***	-0.00062
mutton	-0.00068***	-0.00056	-0.00041***	-0.00028
butter	-0.00192***	-0.00174***	-0.00175***	-0.00174***
eggs	-0.00202***	-0.00161***	-0.00162***	-0.00131***

Table 4.14: Number of significant negative time trends

	Model1	Model2
AD	12	7
CV	12	6
Total	14	14

#### 4.8.4.2 Market efficiency

In order to investigate market efficiency, regional commodity price indices are estimated for towns in the present day Western Cape (WC), towns along the Port Elizabeth (PE) main line, towns along the East London (EL) main line, and towns in the present day Northern Cape (NC). The repeat sales method is used to calculate regional wholesale price indices for the 14 individual commodities for these four regions. Figure 4.23 illustrates the regional indices for four selected commodities.

Table 4.15 reports the pairwise contemporaneous correlations for the growth rates in regional commodity prices for the full period. On the whole, the regional markets seem to have been relatively well integrated over the short term. A number of commodities show significant positive correlations between two regions (especially wheat, mealies, oat-hay, potatoes and eggs). The commodities that exhibit less significant correlations include barley, beef and butter.

The number of commodities with significant positive correlations between the pairwise regional price indices is reported in Table 4.16, distinguishing between the pre-war and post-war periods. The number of positive significant correlations increased in the post-war period for most of the pairs, and particularly between the Western Cape and the other three regions. In many cases the correlations also increased in size in the post-war period (e.g. wheat flour, potatoes and eggs). This implies increasing market integration along this dimension over the sample period.

Table 4.17 reports the results for Johansen trace test for cointegrating relationships between regional price indices for the full sample period. The trace test statistics are reported for the null hypothesis that there are zero cointegrating relationships  $r = 0$ . The null can be rejected for a large proportion

Historical Commodity Prices

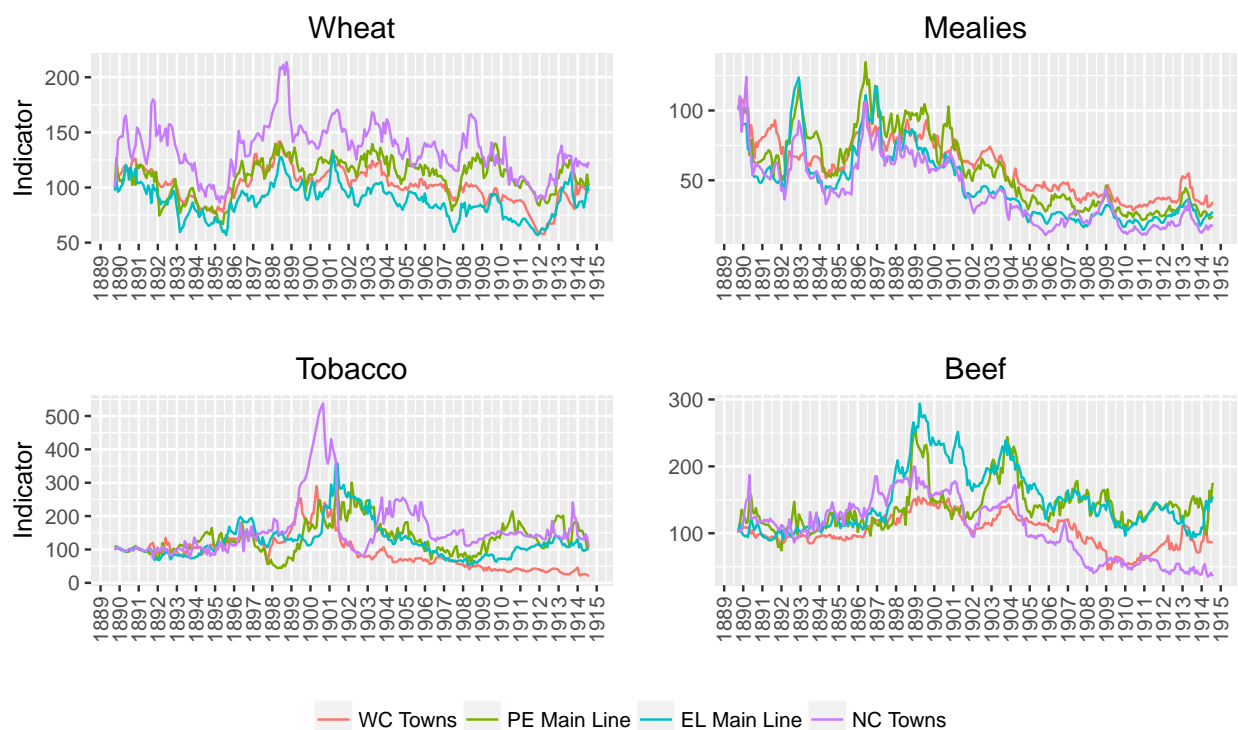


Figure 4.23: Regional commodity price indices

Table 4.15: Correlations between regional price indices (growth rates)

	WC-PE	WC-EL	WC-NC	PE-EL	PE-NC	EL-NC
wheat	0.25***	0.21***	0.21***	0.39***	0.06	0.25***
wheat.flour	0.23***	0.02	0.04	0.00	0.11*	0.22***
boer.meal	0.08	0.30***	0.12**	0.44***	0.15**	0.32***
mealies	0.26***	0.10*	0.31***	0.52***	0.32***	0.39***
mealie.meal	-0.03	0.09	0.14**	0.16**	0.32***	0.38***
barley	0.03	0.03	0.06	0.27***	0.17**	0.02
oats	0.04	0.15**	0.13**	0.07	-0.04	0.11*
oathay	0.15**	0.20***	0.18**	0.22***	0.18**	0.33***
potatoes	0.58***	0.23***	0.20***	0.35***	0.30***	0.67***
tobacco	0.04	0.21***	0.02	0.15**	0.04	0.12**
beef	-0.07	0.15**	0.06	0.13**	0.03	0.05
mutton	-0.15**	-0.05	-0.04	0.25***	0.05	0.2***
butter	0.08	-0.05	0.03	-0.05	-0.07	0.26***
eggs	0.62***	0.62***	0.56***	0.62***	0.6***	0.51***

Table 4.16: Number of significant positive correlations

	WC-PE	WC-EL	WC-NC	PE-EL	PE-NC	EL-NC
Pre-war	3	4	5	9	7	9
Post-war	6	8	6	9	5	10
Full period	6	8	8	11	7	11
Total commodities	14	14	14	14	14	14

## Historical Commodity Prices

Table 4.17: Cointegration trace test statistics

	WC-PE	WC-EL	WC-NC	PE-EL	PE-NC	EL-NC
wheat	20.34**	38.72**	48.44**	25.75**	34.59**	44.02**
wheat.flour	15.71	7.74	11.52	9.04	16.20	7.15
boer.meal	18.17*	17.19	27.21**	26.52**	25.44**	27.68**
mealies	30.70**	23.44**	31.37**	33.71**	51.15**	45.51**
mealie.meal	18.65*	11.73	31.23**	19.52*	33.35**	31.48**
barley	27.11**	18.26*	22.54**	21.72**	21.07**	15.58
oats	24.64**	9.64	14.48	30.66**	22.26**	9.71
oathay	35.89**	46.00**	29.54**	46.73**	30.32**	43.32**
potatoes	32.41**	37.22**	25.80**	40.81**	53.69**	43.49**
tobacco	22.58**	17.01	12.31	21.40**	20.17**	26.61**
beef	18.24*	11.53	13.49	21.28**	13.01	9.42
mutton	8.60	18.97*	10.11	12.33	11.21	18.15*
butter	40.12**	72.76**	36.93**	61.31**	57.75**	59.78**
eggs	34.73**	23.2**	27.57**	34.37**	26.63**	31.57**

Table 4.18: Number of commodities with one significant cointegrating relationship

	WC-PE	WC-EL	WC-NC	PE-EL	PE-NC	EL-NC
Pre-war	4	2	4	2	5	6
Post-war	4	3	5	4	7	4
Full period	8	4	7	9	9	7
Total commodities	14	14	14	14	14	14

of the commodities, i.e. there are more than zero cointegrating relationships. The regional markets were relatively well integrated over the longer term, as evidenced by the cointegrating relationships for commodities such as wheat, mealies, oat-hay, potatoes, butter and eggs. The exceptions are wheat flour, beef and mutton. The next step is to test the null hypothesis that there is at most one cointegrating relationship, relative to the alternative that there are two cointegrating relationships. This hypothesis is rejected in few cases, especially for the butter price indices, implying that these series do not contain unit roots.

The number of commodities for which there was one significant cointegrating relationship between the pairwise regional price indices is reported in Table 4.18, again distinguishing between the pre-war and post-war periods. The number of commodity price index pairs with one significant cointegrating relationship was relatively high, especially considering that a few of the series did not contain unit roots over the different sample periods. The number of commodities with one significant cointegrating relationship increased in the post-war period among almost all of the regions. Specifically, the number increased between the Western Cape region and the East London and Northern Cape regions, as well as between the Port Elizabeth region and the East London and Northern Cape regions. This implies increasing market integration among these regions over the sample period.

The majority of regional commodity price indices therefore exhibited significant short-term and longer-term relationships. This implies that these markets were relatively well integrated along the efficiency dimension, which corresponds with the findings in Boshoff and Fourie (2017). In general, the regional price indices showed strong evidence of significant pairwise correlations and

cointegrating relationships for eggs and crops, especially wheat, mealies, oat-hay and potatoes. Only a few commodities, such as wheat flour, beef and mutton, did not show much evidence of significant relationships between regions. This difference in results for various commodities shows the importance of analysing other commodities in addition to wheat.

The results imply that all three coastal regions were relatively well integrated with the interior, and Kimberley in particular. Integration was relatively high between the Eastern Cape regions and the Northern Cape region. Integration was highest between the two Eastern Cape regions, as one would expect from their close proximity. Integration was lowest between the Western Cape region and the two regions of the Eastern Cape, although it increased in the post-war period. This concurs with the findings in Boshoff and Fourie (2017) that Eastern Cape wheat and maize markets were less well integrated with Cape Town markets.

This section has shown that internal market integration was increasing in the Cape Colony along both dimensions. The difference in results for various commodities shows the importance of analysing other commodities in addition to wheat. On the whole, both sets of tests pointed to integration in the markets for wheat, mealies, oat-hay, potatoes and eggs. In a few cases, the tests provided different results. This is unsurprising, as they refer to different aspects of integration (Federico, 2012). For example, wheat flour prices seemed to have converged quickly between two towns, but did not show significant comovement over time.

In all likelihood, the increasing market integration among these regions was aided by the expansion of branch railway lines between the towns. In addition, during this period there were a number of significant changes in technology and infrastructure, as well as institutional changes, which may have contributed to market integration. Future studies may investigate the factors that contributed to this increasing internal market integration, especially if transaction volumes become available for analysis.

## 4.9 Conclusion

This chapter has attempted three contributions to the literature. The first was to demonstrate the use of the repeat sales method to aggregate incomplete information from historical records. The repeat sales method employed in this chapter may prove useful in a diverse range of applications where incomplete information from various sources can be combined. This is likely to be a more common problem in future as more statistical records are digitised and become available for analysis. The repeat sales method may be useful in these cases to estimate more continuous indices, for example, for incomes, wages and trade volumes.

Further methods to aggregate the high-frequency price records were explored, in order to investigate market integration in the Cape Colony around the turn of the 20th century. To investigate price convergence, two measures of cross-sectional price dispersion among the towns in the Cape Colony

were calculated: absolute deviation and the coefficient of variation. The measures of price dispersion used in this chapter to investigate price convergence are closely related to the measures used in Chapter 3. In that chapter, measures of cross-sectional dispersion of survey responses were used to create proxies for uncertainty. In this chapter, measures of cross-sectional price dispersion were used as a gauge of market integration in terms of price convergence. This demonstrates that these methods may be useful in a diverse range of applications.

The second contribution was to produce numerous monthly wholesale commodity price indices for the Cape Colony. The price indices can shed new light on selected episodes in the economic and financial history. In wartime periods, prices may be among the few available time series that can be accurately measured, and may convey information on the state of the economy (Klovland, 2014). The total commodity price index, reported in Table 4.19 in the Appendix below, may also be useful as an improved measure of inflation for the Cape Colony over this period. Further research might supplement the indices by including prices from other sources, and potentially import and export prices. Another extension would be to add prices from the other territories in South Africa, i.e. Natal, the Transvaal, and the Orange Free State, to investigate and compare the price histories of the four territories.

In addition, two new sets of historical records on market prices in the Cape Colony were digitised. The first is an expanded version of the Agricultural Journal records used in Boshoff and Fourie (2017), which captures high-frequency market prices, and the second is a set of records on annual prices from the Cape Colony Blue Books. These new datasets will be of use to economic historians doing research on this period of the Cape Colony's history.

The third contribution was to investigate market integration in the Cape Colony over the period concerned. Although the data is patchy, and the number of commodities limited, the results indicated that there was evidence of increasing market integration among towns in the Cape Colony over the sample period. Potential further research could investigate convergence among towns along the main and branch railway lines, if more records for specific towns are added. The factors that aided market integration could also be investigated, as well as the impact of market integration on welfare and long-term growth (Federico, 2012). The high-frequency price indices can facilitate these inquiries.

## 4.10 Appendix

## Historical Commodity Prices

Table 4.19: Total commodity price index

Date	WPI	Date	WPI	Date	WPI	Date	WPI	Date	WPI	Date	WPI
1889-09-30	98.40	1893-11-30	91.78	1898-01-31	104.17	1902-03-31	128.30	1906-05-31	111.24	1910-07-31	94.26
1889-10-31	103.61	1893-12-31	89.80	1898-02-28	104.38	1902-04-30	126.82	1906-06-30	111.70	1910-08-31	94.25
1889-11-30	102.83	1894-01-31	89.32	1898-03-31	105.68	1902-05-31	125.92	1906-07-31	109.27	1910-09-30	95.32
1889-12-31	100.42	1894-02-28	89.54	1898-04-30	106.37	1902-06-30	129.64	1906-08-31	109.79	1910-10-31	96.54
1890-01-31	99.69	1894-03-31	89.78	1898-05-31	106.79	1902-07-31	131.51	1906-09-30	109.08	1910-11-30	93.55
1890-02-28	98.49	1894-04-30	90.06	1898-06-30	107.52	1902-08-31	132.80	1906-10-31	109.39	1910-12-31	93.29
1890-03-31	96.85	1894-05-31	90.34	1898-07-31	108.49	1902-09-30	133.58	1906-11-30	108.89	1911-01-31	95.02
1890-04-30	99.77	1894-06-30	90.76	1898-08-31	108.55	1902-10-31	132.95	1906-12-31	107.80	1911-02-28	95.96
1890-05-31	100.58	1894-07-31	90.35	1898-09-30	110.70	1902-11-30	131.91	1907-01-31	105.02	1911-03-31	94.79
1890-06-30	101.15	1894-08-31	89.70	1898-10-31	113.07	1902-12-31	131.87	1907-02-28	105.00	1911-04-30	93.98
1890-07-31	98.79	1894-09-30	89.97	1898-11-30	113.98	1903-01-31	130.73	1907-03-31	104.30	1911-05-31	93.27
1890-08-31	98.31	1894-10-31	87.48	1898-12-31	113.13	1903-02-28	129.12	1907-04-30	102.69	1911-06-30	93.30
1890-09-30	96.79	1894-11-30	86.05	1899-01-31	114.35	1903-03-31	127.64	1907-05-31	100.71	1911-07-31	94.02
1890-10-31	98.56	1894-12-31	85.57	1899-02-28	116.67	1903-04-30	131.11	1907-06-30	100.13	1911-08-31	93.12
1890-11-30	101.10	1895-01-31	87.24	1899-03-31	117.37	1903-05-31	128.22	1907-07-31	98.23	1911-09-30	92.07
1890-12-31	100.53	1895-02-28	89.14	1899-04-30	120.23	1903-06-30	130.50	1907-08-31	99.17	1911-10-31	92.01
1891-01-31	101.43	1895-03-31	90.30	1899-05-31	121.53	1903-07-31	130.57	1907-09-30	99.80	1911-11-30	92.57
1891-02-28	101.88	1895-04-30	91.90	1899-06-30	123.48	1903-08-31	128.30	1907-10-31	99.41	1911-12-31	92.27
1891-03-31	101.64	1895-05-31	92.41	1899-07-31	124.83	1903-09-30	131.33	1907-11-30	99.89	1912-01-31	93.71
1891-04-30	100.91	1895-06-30	92.46	1899-08-31	126.50	1903-10-31	132.75	1907-12-31	99.68	1912-02-29	94.57
1891-05-31	100.46	1895-07-31	94.20	1899-09-30	122.76	1903-11-30	132.79	1908-01-31	99.95	1912-03-31	92.30
1891-06-30	99.34	1895-08-31	95.41	1899-10-31	123.33	1903-12-31	132.30	1908-02-29	99.56	1912-04-30	90.18
1891-07-31	97.86	1895-09-30	97.44	1899-11-30	124.69	1904-01-31	128.75	1908-03-31	96.83	1912-05-31	90.47
1891-08-31	97.30	1895-10-31	97.58	1899-12-31	126.71	1904-02-29	127.89	1908-04-30	97.38	1912-06-30	93.90
1891-09-30	96.49	1895-11-30	97.00	1900-01-31	124.85	1904-03-31	127.10	1908-05-31	91.83	1912-07-31	92.72
1891-10-31	93.58	1895-12-31	96.69	1900-02-28	123.33	1904-04-30	120.80	1908-06-30	91.73	1912-08-31	97.46
1891-11-30	92.83	1896-01-31	97.72	1900-03-31	125.95	1904-05-31	117.37	1908-07-31	97.30	1912-09-30	99.18
1891-12-31	89.68	1896-02-29	100.06	1900-04-30	128.44	1904-06-30	115.03	1908-08-31	96.50	1912-10-31	98.39
1892-01-31	91.32	1896-03-31	101.54	1900-05-31	132.93	1904-07-31	112.22	1908-09-30	98.53	1912-11-30	102.98
1892-02-29	92.62	1896-04-30	104.29	1900-06-30	132.60	1904-08-31	112.41	1908-10-31	97.85	1912-12-31	105.96
1892-03-31	95.52	1896-05-31	107.02	1900-07-31	134.96	1904-09-30	111.76	1908-11-30	97.68	1913-01-31	102.28
1892-04-30	95.83	1896-06-30	105.82	1900-08-31	134.36	1904-10-31	111.96	1908-12-31	98.13	1913-02-28	101.75
1892-05-31	96.38	1896-07-31	104.56	1900-09-30	131.95	1904-11-30	112.95	1909-01-31	98.22	1913-03-31	97.72
1892-06-30	95.58	1896-08-31	102.90	1900-10-31	133.22	1904-12-31	115.23	1909-02-28	99.67	1913-04-30	97.81
1892-07-31	95.40	1896-09-30	101.98	1900-11-30	132.31	1905-01-31	115.54	1909-03-31	96.97	1913-05-31	95.72
1892-08-31	94.94	1896-10-31	102.03	1900-12-31	136.52	1905-02-28	115.44	1909-04-30	96.78	1913-06-30	92.35
1892-09-30	94.43	1896-11-30	102.77	1901-01-31	141.36	1905-03-31	117.02	1909-05-31	95.27	1913-07-31	94.16
1892-10-31	93.82	1896-12-31	101.38	1901-02-28	137.30	1905-04-30	115.16	1909-06-30	95.09	1913-08-31	95.96
1892-11-30	92.53	1897-01-31	100.83	1901-03-31	137.27	1905-05-31	113.52	1909-07-31	94.94	1913-09-30	96.48
1892-12-31	91.69	1897-02-28	100.66	1901-04-30	135.38	1905-06-30	115.41	1909-08-31	94.64	1913-10-31	94.86
1893-01-31	92.29	1897-03-31	99.67	1901-05-31	138.69	1905-07-31	113.72	1909-09-30	92.79	1913-11-30	95.30
1893-02-28	92.28	1897-04-30	99.43	1901-06-30	135.83	1905-08-31	111.50	1909-10-31	92.78	1913-12-31	96.11
1893-03-31	92.53	1897-05-31	98.94	1901-07-31	136.34	1905-09-30	115.13	1909-11-30	91.28	1914-01-31	91.74
1893-04-30	92.23	1897-06-30	98.33	1901-08-31	137.39	1905-10-31	114.90	1909-12-31	91.31	1914-02-28	94.45
1893-05-31	92.87	1897-07-31	98.98	1901-09-30	136.20	1905-11-30	112.80	1910-01-31	91.79	1914-03-31	93.59
1893-06-30	92.38	1897-08-31	97.96	1901-10-31	137.13	1905-12-31	112.37	1910-02-28	90.84	1914-04-30	89.79
1893-07-31	92.96	1897-09-30	96.83	1901-11-30	133.70	1906-01-31	110.43	1910-03-31	90.36	1914-05-31	92.33
1893-08-31	92.98	1897-10-31	96.73	1901-12-31	129.39	1906-02-28	110.22	1910-04-30	89.27	1914-06-30	92.22
1893-09-30	92.55	1897-11-30	98.42	1902-01-31	128.59	1906-03-31	112.42	1910-05-31	89.05	1914-07-31	90.35
1893-10-31	92.06	1897-12-31	100.78	1902-02-28	129.34	1906-04-30	110.95	1910-06-30	92.05		



## 5 Conclusion

Large microeconomic datasets are becoming increasingly available, due to technological developments, and will become more commonplace in the future. These datasets provide an opportunity to improve on the measurement of a range of economic phenomena and to bolster economic research (Schutt and O’Neil, 2013). One way in which these large amounts of data can aid economic analysis is to enable the construction of better macroeconomic indicators. The construction of such indicators requires the aggregation of the microeconomic data and this dissertation has focused on particular challenges in aggregating information.

This dissertation attempted three main contributions, which involved answering the set of research questions posed in Chapter 1. The first contribution was to explore suitable aggregation methods to overcome some of the challenges in aggregating the relatively large microeconomic datasets. In Chapter 2, the challenge involved compositional changes over time. The hedonic and hybrid repeat sales methods were used to estimate quality-adjusted price indices for South African art. A simple new pseudo-repeat sales method was proposed that, to some extent, addressed the lack of repeat sales in the sample and the potential omitted variable bias of the hedonic method. In Chapter 3, the challenge was the identification of aggregate time series from disparate qualitative survey responses. Various measures of the weighted first and second moments of the distribution of responses were used to create indicators of business confidence and uncertainty for South Africa. In Chapter 4, the challenge involved incomplete coverage over time and across units. The repeat sales method was used to estimate high-frequency commodity price indices for the Cape Colony, as well as regional commodity price indices. In addition, two measures of cross-sectional price dispersion between the towns in the Cape Colony were estimated to investigate price convergence.

The second contribution was to produce useful macroeconomic time-series indicators: art price indices, sentiment indices, and historical commodity price indices. The regression-based indices, estimated in Chapter 2, pointed to the same general trend in South African art prices, with a large increase in the run-up to the Great Recession and a flat trend after 2009. The composite confidence and uncertainty indicators, calculated in Chapter 3, appeared to be plausible and useful proxies for business sentiment in South Africa, reflecting key developments in real economic activity. The commodity price indices, estimated in Chapter 4, seemed to correspond well with the economic history of the Cape Colony, with a large increase before and during the Second South African War and a large decrease prior to unification. Chapter 4 also presented two newly digitised sets of historical records on market prices in the Cape Colony. These records will be of use to economic historians doing quantitative research on this period of the Cape Colony’s history.

To demonstrate how the indicators might be used to aid economic analysis in each field, and to provide a further assessment of the validity of the proposed aggregation methods, the third contribution was to use the indicators to test a specific hypothesis in each case: explosive South African art prices, significant predictive comovement between sentiment and real activity in South

Africa, and increasing market integration in the Cape Colony.

Chapter 2 demonstrated the consistent results provided by the regression-based indices in terms of the explosive periods in South African art prices, with a potential bubble most likely beginning in 2006 and ending in 2008. Further research could investigate the potential drivers of art price and the bubble formation process. In Chapter 3, the VAR models indicated that positive shocks to the confidence and uncertainty indicators were followed by a significant increase and decrease in real GDP growth respectively. Future research may investigate the usefulness of these indicators in out-of-sample and real-time forecasts, as well as the potential causal impact of sentiment on real activity. Chapter 4 provided evidence of increasing market integration in the Cape Colony around the turn of the 20th century. Future research could investigate the factors that aided the market integration, such as railway expansion, as well as the impact of market integration on welfare and long-term growth.

The aggregation techniques demonstrated in this dissertation may become ever more useful as additional microeconomic datasets become available for economic analysis. The time-series indicators may prove useful for further research in each of the relevant fields and may be easily updated over time as additional data become available. The rest of Chapter 5 presents brief chapter summaries.

## 5.1 Art Prices

Chapter 2 explored methods for constructing quality-adjusted South African art price indices, using a relatively large dataset of auction prices, which is subject to compositional changes in the sample. Three methodologies for addressing this challenge were demonstrated in the chapter: central tendency methods, hedonic methods and hybrid repeat sales methods.

It was argued that central tendency measures do not adequately control for compositional changes over time. Various indices were estimated with the hedonic regression method, which is able to control more adequately for quality-mix changes by attributing implicit prices to a set of asset characteristics. A shortcoming of indices based on the hedonic method is that they may suffer from potential omitted variable bias. The scarcity of repeat sales observations in the dataset limited the usefulness of the classical repeated sales approach in this case. Proposed, therefore, in this chapter was a simple new pseudo-repeat sales methodology, which addresses to some extent the problem of the scarcity of repeat sales and the potential omitted variable bias inherent in the hedonic method (Guo *et al.*, 2014).

The indices estimated with the regression-based methods point to the same general trend in South African art prices, which provides some confidence that omitted variable bias and sample selection bias are not pervasive in this case (Calomiris and Pritchett, 2016). The indices were evaluated in terms of smoothness metrics, which suggested that the 1-year adjacent period hedonic index performed the best, although the regression-based indices performed similarly.

The hedonic and hybrid repeat sales methods demonstrated in this chapter produced indices that improve upon the more common central tendency indices, in that they reflected a more plausible cyclical trend and exhibited better signal-to-noise metrics. These methods may be useful in constructing indices for other unique assets, such as real estate, antiques, coins and wine, where the quality-mix of items differs over time, and there is a lack of repeat sales.

In order to demonstrate the usefulness of the estimated price indices for investigating developments in the South African art market, the indices were studied for evidence of a price bubble. The direct bubble detection tests proposed by Phillips, Wu and Yu (2011) were used to investigate whether South African art prices exhibited mildly explosive behaviour over the period. The regression-based indices provided relatively consistent results in terms of the explosive periods in the South African art market, with a potential bubble most likely beginning in 2006 and ending in 2008. The process seemed to be particularly prevalent in the higher-end, and oil and watercolour segments of the market.

## 5.2 Business Sentiment

Chapter 3 presented an exploration of aggregation methods to create proxies for business sentiment in South Africa, using the microeconomic data from the BER's business tendency surveys. Although measuring business sentiment is not straightforward, survey-based indicators can be helpful in discovering agents' opinions on future economic developments (Organisation for Economic Co-operation and Development, 2003). Survey-based measures have the advantage that they are derived from the opinions of key economic agents (Girardi and Reuter, 2017), are available with a shorter lag than official statistics, and are usually not subject to revisions (European Central Bank, 2013). The challenge in aggregating the qualitative survey responses is to fully exploit the disparate views of individual agents.

The chapter demonstrated aggregation methods to capture the full distribution of the qualitative survey responses, by calculating different combinations of the weighted cross-sectional first and second moments of the distribution of the qualitative survey responses. The composite indicators incorporated the survey responses from questions on general business conditions, output, employment, orders placed and profitability. For each question, the responses were weighted by firm size and subsector size to produce sectoral indicators. The sectoral indicators were then weighted by GDP share to produce the overall aggregate composite indicators.

Two composite confidence indicators were calculated: the cross-sectional mean of responses to questions on current and expected future business conditions (Organisation for Economic Co-operation and Development, 2003). These confidence indicators exhibited a significant positive correlation with real GDP growth and significant concordance with the official SARB business cycle. Three composite uncertainty indicators were calculated: the scaled cross-sectional standard deviation of forward-looking responses (Girardi and Reuter, 2017); and the cross-sectional mean and

standard deviation of individual firm forecast errors (Bachmann, Elstner and Sims, 2013; Arslan *et al.*, 2015). An overall uncertainty indicator was created, by combining the information from the survey-based indicators, the SAVI and the EPU created by Hlathwayo and Saxegaard (2016). The combined indicator appeared to be a plausible proxy of macroeconomic uncertainty in South Africa. The composite dispersion and combined uncertainty indicators, in particular, exhibited significant negative correlations with real GDP growth.

The weighted cross-sectional moments employed in this chapter would be useful in other applications with qualitative survey responses, such as consumer surveys, where there are challenges in capturing the full richness in the data. It would be possible to improve on the existing measures of consumer confidence and create new measures of consumer uncertainty using these methods.

To demonstrate the usefulness of the aggregation methods and the estimated indicators, as well as provide an additional validity test, the relationship between business sentiment and real economic activity in South Africa was further examined. The hypothesis that there was significant comovement between the sentiment indicators and real GDP growth was tested, using the standard VAR framework.

In the VAR models, positive shocks to the confidence indicators were followed by significant increases in real GDP growth. This was the case for the aggregate indicators as well as the sectoral indicators. Positive shocks to the uncertainty indicators were generally followed by a decrease in real GDP growth. The significant relationships between the sentiment indicators and real economic activity held after the inclusion of additional variables. The impact of shocks on real production and investment growth were larger and more significant than on real GDP growth. The results provided evidence of at least important comovement between the sentiment indicators and real economic activity. The indicators may therefore be useful for real time monitoring of economic developments and contain potentially useful predictive content.

### 5.3 Historical Commodity Prices

Chapter 4 explored aggregation methods for developing relatively complete monthly wholesale commodity price indices for the Cape Colony around the turn of the 20th century, using incomplete historical records. The indices were based on an expanded version of the historical dataset used in Boshoff and Fourie (2017), which captures high-frequency market prices for various commodities in various towns from the Agricultural Journals, as well as newly digitised annual prices for various commodities in various towns from the Cape Colony Blue Books. The challenge in aggregating these two datasets was that the data was incomplete in terms of the coverage of both products and towns.

The chapter demonstrated that the repeat sales method provides a consistent way to aggregate the incomplete price data and produces indices with substantially fewer gaps than in the individual series (Klovland, 2014). Monthly wholesale commodity price indices were estimated for several

individual products. These indices were then aggregated to form a total commodity price index, with weights based on the production values reported in the 1904 census, and on import values reported in the Blue Books.

The commodity price indices estimated with the repeat sales method corresponded well with the economic history of the Cape Colony. In general, prices exhibited a pronounced cyclical trend over the period, with a large increase before and during the Second South African War (1899-1902) and a large decrease prior to unification (1910). The total commodity price index provides a clearer picture of the inflation history over the sample period than the existing measures, in the sense that it exhibited a higher correlation with available monetary aggregates.

The chapter subsequently explored further aggregation methods for the high-frequency price records, to investigate internal market integration in the Cape Colony around the turn of the 20th century. This demonstrated the usefulness of the aggregation methods and the estimated time series. In this part of the chapter the work of Boshoff and Fourie (2017) was extended, by expanding the period and the number of products, and investigating both dimensions of market integration, i.e. price convergence and market efficiency.

In examining these two dimensions, two aggregation methods were used. The first was to calculate two measures of cross-sectional price dispersion between the towns in the Cape Colony, absolute deviation and the coefficient of variation, which were used to investigate price convergence. These measures are related to the measures used in Chapter 3, in that they are measures of the first and second cross-sectional moments of the distribution of prices. This demonstrates that these aggregation methods are useful for calculating macroeconomic time-series indicators in a variety of settings. These measures showed evidence of convergence in prices across towns, as price dispersion exhibited a declining trend over the sample period for the majority of the commodities. This implies increasing market integration between towns in the Cape Colony along the dimension of price convergence.

The second method was to calculate repeat sales commodity price indices at the regional level, which were used to investigate market efficiency. The classical repeat sales method used in this chapter is related to the methods used in Chapter 2, as it also calculates the mean in the distribution of growth rates. The use of the repeat sales method in the context of estimating complete series from incomplete historical records is a novel application of this method. The majority of regional repeat sales price indices exhibited significant short-term and longer-term relationships. There was an increase in the number of regional commodity prices with significant positive correlations and with significant cointegrating relationships in the post-war period, relative to the pre-war period. This implies increasing market integration between regions in the Cape Colony along the dimension of market efficiency.

---

## References

- Acemoglu, D. and Scott, A. (1994) ‘Consumer Confidence and Rational Expectations: Are Agents’ Beliefs Consistent with the Theory?’, *The Economic Journal*, 104(January), pp. 1–19.
- Aghevli, B. B. (1975) ‘The Balance of Payments and Money Supply Under the Gold Standard Regime: U.S. 1879-1914’, *American Economic Review*, 65(1), pp. 40–58.
- Akerlof, G. and Shiller, R. J. (2015) *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton University Press.
- Anderson, R. C. (1974) ‘Paintings as an Investment’, *Economic Inquiry*, 12(1), pp. 13–26. doi: 10.1111/j.1465-7295.1974.tb00223.x.
- Andrabi, T. and Kuehlwein, M. (2010) ‘Railways and Price Convergence in British India’, *Journal of Economic History*, 70(2), pp. 351–377.
- Areal, F. J., Balcombe, K. and Rapsomanikis, G. (2013) ‘Testing for Bubbles in Agriculture Commodity Markets’, *Munich Personal RePEc Archive*, 48015.
- Arslan, Y., Atabek, A., Hulagu, T. and Sahinöz, S. (2015) ‘Expectation Errors, Uncertainty, and Economic Activity’, *Oxford Economic Papers*, 67(3), pp. 634–660. doi: 10.1093/oep/gpv003.
- Aye, G. C., Balcilar, M., Bosch, A. and Gupta, R. (2014) ‘Housing and the Business Cycle in South Africa’, *Journal of Policy Modeling*, 36(3), pp. 471–491. doi: 10.1016/j.jpolmod.2014.03.001.
- Bachmann, R., Elstner, S. and Sims, E. R. (2010) ‘Uncertainty and Economic Activity: Evidence From Business Survey Data’, *NBER Working Paper*, 16143.
- Bachmann, R., Elstner, S. and Sims, E. R. (2013) ‘Uncertainty and Economic Activity: Evidence From Business Survey Data’, *American Economic Journal: Macroeconomics*, 5(2), pp. 217–249.
- Bailey, M. J., Muth, R. F. and Nourse, H. O. (1963) ‘A Regression Method for Real Estate Price Index Construction’, *Journal of the American Statistical Association*, 58(304), pp. 933–942. doi: 10.1080/01621459.1963.10480679.
- Baker, M. and Wurgler, J. (2007) ‘Investor Sentiment in the Stock Market’, *Journal of Economic Perspectives*, 21(2), pp. 129–151. doi: 10.3386/w13189.
- Baker, S. R. and Bloom, N. (2013) ‘Does Uncertainty Reduce Growth? Using Disasters as Natural Experiments’, *NBER Working Paper*, 19475. doi: 10.3386/w19475.
- Baker, S. R., Bloom, N. and Davis, S. J. (2015) ‘Measuring Economic Policy Uncertainty’, *NBER Working Paper*, 21633.
- Baker, S. R., Bloom, N. and Davis, S. J. (2016) ‘Measuring Economic Policy Uncertainty’, *The*

## References

- Quarterly Journal of Economics*, 131(4), pp. 1593–1636. doi: 10.1093/qje/qjw024.Advance.
- Balcilar, M., Katzke, N. and Gupta, R. (2015) ‘Identifying Periods of US Housing Market Explosivity’, *University of Pretoria: Department of Economics Working Paper Series*, 2015-44.
- Barsky, R. B. and Sims, E. R. (2012) ‘Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence’, *American Economic Review*, 102(4), pp. 1343–1377. doi: 10.1257/aer.102.4.1343.
- Beaudry, P. and Portier, F. (2004) ‘An Exploration into Pigou’s Theory of Cycles’, *Journal of Monetary Economics*, 51(6), pp. 1183–1216. doi: 10.1016/j.jmoneco.2003.10.003.
- Bernanke, B. S. (1983) ‘Irreversibility, Uncertainty, and Cyclical Investment’, *Quarterly Journal of Economics*, 98, pp. 85–106.
- Binding, G. and Dibiasi, A. (2017) ‘Exchange Rate Uncertainty and Firm Investment Plans: Evidence from Swiss Survey Data’, *Journal of Macroeconomics*. Elsevier Inc., 51, pp. 1–27. doi: 10.1016/j.jmacro.2016.11.004.
- Blanchard, O. J. (1993) ‘Consumption and the Recession of 1990-1991’, *American Economic Review*, 83(2), pp. 270–274.
- Bleaney, M. and Greenaway, D. (1993) ‘Long-Run Trends in the Relative Price of Primary Commodities and in the Terms of Trade of Developing Countries’, *Oxford Economic Papers*, 45(3), pp. 349–363.
- Bloom, N. (2009) ‘The Impact of Uncertainty Shocks’, *Econometrica*, 77(3), pp. 623–685. doi: 10.3982/ECTA6248.
- Bloom, N. (2014) ‘Fluctuations in Uncertainty’, *Journal of Economic Perspectives*, 28(2), pp. 153–176. doi: 10.1257/jep.28.2.153.
- Bloom, N., Bond, S. and Van Reenen, J. (2007) ‘Uncertainty and Investment Dynamics’, *Review of Economic Studies*, 74, pp. 391–415.
- Bonciani, D. and Van Roye, B. (2016) ‘Uncertainty Shocks, Banking Frictions and Economic Activity’, *Journal of Economic Dynamics and Control*. Elsevier, 73, pp. 200–219. doi: 10.1016/j.jedc.2016.09.008.
- Bosch, A. and Ruch, F. (2013) ‘An Alternative Business Cycle Dating Procedure for South Africa’, *South African Journal of Economics*, 81(267), pp. 491–516.
- Boshoff, W. H. (2008) ‘Rethinking ASGISA and the Rand Exchange Rate’, *South African Journal of Management Sciences*, 11(1), pp. 112–120.
- Boshoff, W. H. and Fourie, J. (2017) ‘When Did South African Markets Integrate Into The Global



## References

- Economy?', *Journal of Studies in Economics and Econometrics*, 41(1).
- Botha, F., Snowball, J. and Scott, B. (2016) 'Art Investment as a Portfolio Diversification Strategy in South Africa', *South African Journal of Economic and Management Sciences*, 19(3), pp. 358–368.
- Brunnermeier, M. K. (2008) 'Bubbles', *New Palgrave Dictionary of Economics*. 2nd edn. Edited by S. N. Durlauf and L. E. Blume. Basingstoke: Palgrave Macmillan.
- Bry, G. and Boschan, C. (1971) 'Programmed Selection Of Cyclical Turning Points', *National Bureau of Economic Research*, I, pp. 7–63.
- Caldara, D., Fuentes-Albero, C., Gilchrist, S. and Zakrajek, E. (2016) 'The Macroeconomic Impact of Financial and Uncertainty Shocks', *European Economic Review*, 88(1166), pp. 185–207. doi: 10.1016/j.euroecorev.2016.02.020.
- Calomiris, B. C. W. and Pritchett, J. (2016) 'Betting on Secession: Quantifying Political Events Surrounding Slavery and the Civil War', *American Economic Review*, 106(1), pp. 1–23.
- Campbell, J. Y. and Shiller, R. J. (1987) 'Cointegration and Tests of Present Value Models', *Journal of Political Economy*, 95(5), pp. 1062–1088. doi: 10.1086/261502.
- Campbell, R. J. (2009) 'Art as a Financial Investment', *Collectible Investments for the High Net Worth Investor*, pp. 119–150. doi: 10.1016/B978-0-12-374522-4.00006-8.
- Candela, G. and Scorcu, A. E. (1997) 'A Price Index for Art Market Auctions', *Journal of Cultural Economics*, 21, pp. 175–196.
- Candela, G. and Scorcu, A. E. (2001) 'In Search of Stylized Facts on Art Market Prices: Evidence from the Secondary Market for Prints and Drawings in Italy', *Journal of Cultural Economics*, 25(3), pp. 219–231.
- Capgemini (2010) *World Wealth Report 2010*. Capgemini & Merrill Lynch Global Wealth Management.
- Capgemini (2013) *World Wealth Report 2013*. Capgemini & RBC Wealth Management.
- Carroll, C. D., Fuhrer, J. C. and Wilcox, D. W. (1994) 'Does Consumer Sentiment Forecast Household Spending? If So, Why?', *American Economic Review*, 84(5), pp. 1397–1408.
- Case, B. and Quigley, J. M. (1991) 'The Dynamics of Real Estate Prices', *Review of Economics and Statistics*, 73(1), pp. 50–58.
- Case, K. E. and Shiller, R. J. (1987) 'Prices of Single-Family Homes Since 1970: New Indexes for Four Cities', *New England Economic Review*, (September), pp. 45–56. doi: 10.3386/w2393.
- Case, K. E. and Shiller, R. J. (2003) 'Is There a Bubble in the Housing Market?', *Brookings Papers*



## References

- on *Economic Activity*, 34(2), pp. 299–362.
- Caspi, I. (2013) ‘Rtadf: Testing for Bubbles with EViews’, *Munich Personal RePEc Archive*, 58791.
- Chilosi, D., Murphy, T. E., Studer, R. and Coşkun Tunçer, A. (2013) ‘Europe’s Many Integrations: Geography and Grain Markets, 1620-1913’, *Explorations in Economic History*, 50, pp. 46–68. doi: 10.1016/j.eeh.2012.09.002.
- Cochrane, J. H. (1994) ‘Shocks’, *Carnegie-Rochester Conference Series on Public Policy*, 41(December 1994), pp. 295–364.
- Collins, A., Scorcu, A. and Zanola, R. (2009) ‘Reconsidering Hedonic Art Price Indexes’, *Economics Letters*, 104(2), pp. 57–60. doi: 10.1016/j.econlet.2009.03.025.
- Curnow, R. (2010) ‘South Africa’s Booming Art Market’, *CNN World*, 18 June.
- Davis, P. and Garcés, E. (2010) *Quantitative Techniques for Competition and Antitrust Analysis*. Princeton University Press.
- De Kiewiet, C. W. (1941) *A History of South Africa, Social & Economic*. Oxford: Clarendon Press.
- De Kock, M. H. (1924) *Selected Subjects in the Economic History of South Africa*. Cape Town, South Africa: Juta.
- De Zwart, P. (2011) ‘South African Living Standards in Global Perspective, 1835-1910.’, *Economic History of Developing Regions*, 26(1), pp. 49–74. doi: 10.1080/20780389.2011.583003.
- Deng, Y., McMillen, D. and Sing, T. (2012) ‘Private Residential Price Indices in Singapore: A Matching Approach’, *Regional Science and Urban Economics*, 42(3), pp. 485–494.
- Diba, B. T. and Grossman, H. I. (1988) ‘The Theory of Rational Bubbles in Stock Prices’, *Economic Journal*, 98(392), pp. 746–754.
- Donaldson, D. (2012) ‘Railroads of the Raj: Estimating the Impact of Transportation Infrastructure’, *NBER Working Paper*, 16487.
- Dorsey, R. E., Hu, H., Mayer, W. J. and Wang, H. C. (2010) ‘Hedonic versus Repeat-Sales Housing Price Indexes for Measuring the Recent Boom-Bust Cycle’, *Journal of Housing Economics*, 19(2), pp. 87–105. doi: 10.1016/j.jhe.2010.04.001.
- Driver, C. and Urga, G. (2004) ‘Transforming Qualitative Survey Data: Performance Comparisons for the UK’, *Oxford Bulletin of Economics and Statistics*, 66(1), pp. 71–89. doi: 10.1111/j.1440-1754.2007.01273.x.
- Du Plessis, S. (2006) ‘Reconsidering the Business Cycle and Stabilisation Policies in South Africa’,

## References

- Economic Modelling*, 23(5), pp. 761–774. doi: 10.1016/j.econmod.2005.10.006.
- Econex (2012) *Citadel Art Price Index*. March, pp. 1–20.
- Ejrnaes, M., Persson, K. G. and Rich, S. (2008) ‘Feeding the British: Convergence and Market Efficiency in the Nineteenth-Century Grain Trade’, *Economic History Review*, 61, pp. 140–171. doi: 10.1111/j.1468-0289.2007.00386.x.
- Els, M. and Von Fintel, D. (2010) ‘Residential Property Prices in a Sub-Market of South Africa: Separating Real Growth from Attribute Growth’, *South African Journal of Economics*, 78(4), pp. 418–436. doi: 10.1111/j.1813-6982.2010.01244.x.
- Engle, R. F. and Granger, C. W. J. (1987) ‘Co-Integration and Error Correction: Representation, Estimation, and Testing’, *Econometrica*, 55(2), p. 251. doi: 10.2307/1913236.
- Epple, D. (1987) ‘Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products’, *Journal of Political Economy*, 95(1), pp. 59–80.
- European Central Bank (2013) ‘Confidence Indicators and Economic Developments’, *ECB Monthly Bulletin*, (January), pp. 45–58.
- European Commission (2006) *European Economies: The Joint Harmonised EU Programme of Business and Consumer Surveys*. Directorate-General for Economic; Financial Affairs (Special Report No. 5).
- Eurostat (2013) *Handbook on Residential Property Price Indices (RPPIs)*. European Commission.
- Evans, G. W. (1991) ‘Pitfalls in Testing for Explosive Bubbles in Asset Prices’, *American Economic Review*, 81(4), pp. 922–930. doi: 10.2307/2006651.
- Evans, M. F. M. (2000) *Encyclopaedia of the Boer War*. Santa Barbara, California: ABC-CLIO.
- Fedderke, J. W. and Li, K. (2014) ‘Art in Africa: Market Structure and Pricing Behavior in the South African Fine Art Auction Market, 2009-2013’, *ERSA Working Paper*, 466.
- Federico, G. (2012) ‘How much do we know about Market Integration in Europe?’, *Economic History Review*, 2, pp. 470–497.
- Federico, G. and Sharp, P. (2013) ‘The Cost of Railroad Regulation: The Disintegration of American Agricultural Markets in the Interwar Period’, *Economic History Review*, 66(4), pp. 1017–1038. doi: 10.1111/1468-0289.12009.
- Figuerola-Ferretti, I., Gilbert, C. L. and Mccrorie, R. (2015) ‘Testing for Bubbles in LME Non-Ferrous Metals Prices’, *Journal of Time Series Analysis*, 36, pp. 763–782.
- Findlay, M. (2012) *The Value of Art: Money, Power, Beauty*. Munich: Prestel Publishing.
- Fourie, J. (2016) ‘The Data Revolution in African Economic History’, *Journal of Interdisciplinary*

## References

- History*, XLVII(2), pp. 193–212. doi: 10.1162/JINH.
- Fourie, J., Grundlingh, A. and Mariotti, M. (2017) ‘Poor South Africa! Will No Nice English People Ever Come Out Here?’ The South African Constabulary of the Anglo-Boer War’, *Journal of Imperial and Commonwealth History*, 45(4), pp. 580–606.
- Fourie, J., Inwood, K. and Mariotti, M. (2014) ‘Can Historical Changes in Military Technology Explain the Industrial Growth Puzzle?’ London School of Economics (Mimeo).
- Friedman, M. (1952) *Prices in Wartime Periods*. Chicago: Aldine Publishing Company.
- Friedman, M. and Schwartz, A. J. (1963) *A Monetary History of the United States, 1867-1960*. Princeton: Princeton University Press.
- Froeb, L. and Koyak, R. (1994) ‘Measuring and Comparing Smoothness in Time Series: The Production Smoothing Hypothesis’, *Journal of Econometrics*, 64, pp. 97–122.
- Fuhrer, J. C. (1988) ‘On the Information Content of Consumer Survey Expectations’, *The Review of Economics and Statistics*, pp. 140–144.
- Fung, D. S. (2006) *Methods for the Estimation of Missing Values in Time Series*. PhD thesis. Edith Cowan University.
- Gehlbach, H. and Barge, S. (2012) ‘Anchoring and Adjusting in Questionnaire Responses’, *Basic and Applied Social Psychology*, 34(5 (September)), pp. 417–433. doi: 10.1080/01973533.2012.711691.
- Gérard-Varet, L. V. (1995) ‘On Pricing the Priceless: Comments on the Economics of the Visual Art Market’, *European Economic Review*, 39, pp. 509–518.
- Giannone, D., Reichlin, L. and Small, D. (2008) ‘Nowcasting: The Real-Time Informational Content of Macroeconomic Data’, *Journal of Monetary Economics*, 55(June), pp. 665–676. doi: 10.1016/j.jmoneco.2008.05.010.
- Giffen, R. (1879) ‘On the Fall of Prices of Commodities in Recent Years Source’, *Journal of the Statistical Society of London*, 42(1), pp. 36–78.
- Gilbert, D. W. (1933) ‘The Economic Effects of the Gold Discoveries Upon South Africa: 1886-1910’, *Quarterly Journal of Economics*, 47(4), pp. 553–597.
- Girardi, A. and Reuter, A. (2017) ‘New Uncertainty Measures for the Euro Area Using Survey Data’, *Oxford Economic Papers*, 69(1), pp. 278–300. doi: 10.1093/oep/gpw058.
- Goetzmann, W., Mamonova, E. and Spaenjers, C. (2014) ‘The Economics of Aesthetics and Three Centuries of Art Price Records’, *NBER Working Paper*, No. 20440.
- Goetzmann, W., Renneboog, L. and Spaenjers, C. (2011) ‘Art and Money’, *American Economic Review*, 101(3), pp. 222–246.
- Greyling, L. and Verhoef, G. (2015) ‘Slow Growth, Supply Shocks and Structural Change: The

## References

- GDP of the Cape Colony in the Late Nineteenth Century’, *Economic History of Developing Regions*, 30(1), pp. 23–43. doi: 10.1080/20780389.2015.1012711.
- Greyling, L. and Verhoef, G. (2017) ‘Savings and Economic Growth: A Historical Analysis of the Cape Colony Economy, 1850–1909’, *Economic History of Developing Regions*. Taylor & Francis, 32(2), pp. 127–176. doi: 10.1080/20780389.2017.1327808.
- Griliches, Z. (1961) *Hedonic Price Indexes for Automobiles: An Econometric of Quality Change*. Edited by Price Statistics Review Committee. National Bureau of Economic Research, pp. 173–196. doi: 10.1017/CBO9781107415324.004.
- Guo, X., Zheng, S., Geltner, D. and Liu, H. (2014) ‘A New Approach for Constructing Home Price Indices: The Pseudo Repeat Sales Model and its Application in China’, *Journal of Housing Economics*, 25, pp. 20–38. doi: 10.1016/j.jhe.2014.01.005.
- Gupta, R. and Kabundi, A. (2011) ‘A Large Factor Model for Forecasting Macroeconomic Variables in South Africa’, *International Journal of Forecasting*, 27(4), pp. 1076–1088. doi: 10.1016/j.ijforecast.2010.10.001.
- Gupta, R., Jurgilas, M. and Kabundi, A. (2010) ‘The Effect of Monetary Policy on Real House Price Growth in South Africa: A Factor-Augmented Vector Autoregression (FAVAR) Approach’, *Economic Modelling*, 27(1), pp. 315–323. doi: 10.1016/j.econmod.2009.09.011.
- Hansen, J. (2009) ‘Australian House Prices: A Comparison of Hedonic and Repeat-Sales Measures’, *Economic Record*, 85, pp. 132–145. doi: 10.1111/j.1475-4932.2009.00544.x.
- Harding, D. and Pagan, A. (2002) ‘Dissecting the Cycle: A Methodological Investigation’, *Journal of Monetary Economics*, 49(2), pp. 365–381. doi: 10.1016/S0304-3932(01)00108-8.
- Harding, D. and Pagan, A. (2006) ‘Synchronization of Cycles’, *Journal of Econometrics*, 132, pp. 59–79. doi: 10.1016/j.jeconom.2005.01.023.
- Hart, C. (2015) ‘The Effects of Uncertainty: Evidence from the South African Manufacturing Sector’, *MComm Thesis*. Stellenbosch University.
- Herranz-Loncán, A. and Fourie, J. (2017) ‘For the public benefit?’ Railways in the British Cape Colony’, *European Review of Economic History*, pp. 1–28.
- Hiraki, T., Ito, A., Spieth, D. and Takezawa, N. (2009) ‘How Did Japanese Investments Influence International Art Prices?’, *Journal of Financial and Quantitative Analysis*, 44(06), p. 1489. doi: 10.1017/S0022109009990366.
- Hlatshwayo, S. and Saxegaard, M. (2016) ‘The Consequences of Policy Uncertainty: Disconnects and Dilutions in the South African Real Effective Exchange’, *IMF Working Paper*, WP/16/113.
- Hoehn, T., Langenfeld, J., Meschi, M. and Waverman, L. (1999) ‘Quantitative Techniques in

## References

- Competition Analysis', *Office of Fair Trading: Research Paper*, 17.
- Homm, J. and Breitung, U. (2012) 'Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods', *Journal of Financial Econometrics*, 12(1), pp. 198–231.
- Hundt, S. (2010) 'Art Auction Round-Up', *SANLAM Private Investments Art Advisory Service*.
- International Monetary Fund (2004) 'Producer Price Index Manual', *IMF Publications*, September.
- International Monetary Fund (2017) 'World Economic Outlook Update', *IMF Publications*, July 23.
- Jackson, G., Tetlow, G. and Kahn, M. (2017) 'Brexit Uncertainty is Holding Back Investment, Mark Carney Warns', *Financial Times*, 3 August.
- Jevons, W. S. (1865) 'On the Variation of Prices and the Value of the Currency Since 1782', *Journal of the Statistical Society of London*, 28(2), pp. 294–320.
- Jiang, L., Phillips, P. C. and Yu, J. (2015) 'New Methodology for Constructing Real Estate Price Indices Applied to the Singapore Residential Market', *Journal of Banking and Finance*. Elsevier B.V., 61, pp. S121–S131. doi: 10.1016/j.jbankfin.2015.08.026.
- Johansen, S. (1991) 'Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models', *Econometrica*, 59(6), p. 1551. doi: 10.2307/2938278.
- Jolliffe, I. (2002) *Principal Component Analysis*. New York: Springer. doi: 10.2307/1270093.
- JSE (2014) *A New Way of Measuring Expected Volatility*. JSE Limited.
- Jurado, K., Ludvigson, S. C. and Ng, S. (2015) 'Measuring Uncertainty', *American Economic Review*, 105(3), pp. 1177–1216. doi: 10.1162/REST\_a\_00207.
- Kabundi, A. (2004) 'Estimation of Economic Growth in France Using Business Survey Data', *IMF Working Paper*, WP/04/69.
- Kabundi, A., Nel, E. and Ruch, F. (2016) 'Nowcasting Real GDP Growth in South Africa', *ERSA Working Paper*, 581.
- Kershoff, G. (2000) 'Measuring Business and Consumer Confidence in South Africa', *Bureau for Economic Research*, (December).
- Kershoff, G. (2002) 'An Analysis of The BER's Trade and Building Survey Panels', *Journal for Studies in Economics and Econometrics*, 26(1), pp. 1–17.
- Kershoff, G. (2015) 'South Africa: The BER's Business Tendency Surveys', *Bureau for Economic Research*, pp. 1–19.
- Keynes, J. M. (1936) *General Theory of Employment, Interest and Money*. London: Macmillan. doi: 10.2307/2143949.
- Kilic, E. and Cankaya, S. (2016) 'Consumer Confidence and Economic Activity: A Factor Augmented

## References

- VAR Approach', *Applied Economics*, 48(32), pp. 3062–3080. doi: 10.1080/00036846.2015.1133902.
- Kindleberger, C. P. and Aliber, R. Z. (2005) *Manias, Panics, and Crashes*. 5th edn. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Klenow, P. J. and Kryvtsov, O. (2008) 'State Dependant or Time Dependent Pricing: Does it Matter For Recent U.S Inflation?', *Quarterly Journal of Economics*, CXIII(August), pp. 1–55. doi: 10.1017/CBO9781107415324.004.
- Klovland, J. T. (2005) 'Commodity Market Integration 1850-1913: Evidence from Britain and Germany', *European Review of Economic History*, 9(2), pp. 163–197.
- Klovland, J. T. (2014) 'New Methods for Construction of Historical Price Indices, with an Illustration from Norway, 1777-1920', *European Review of Economic History*, 18(3), pp. 277–305. doi: 10.1093/ereh/heu012.
- Knight, F. H. (1921) *Risk, Uncertainty, and Profit*. New York: Sentry Press.
- Koma, S. (2013) 'The Trajectory of Economic Development Policies in South Africa: The Case for Public Policy Analysis', *Administratio Publica*, 21(1), pp. 142–160.
- Korteweg, A. G. (2013) 'Research: Is Art A Good Investment?', *Stanford Business*, 21 October.
- Kräussl, R. (2015) 'Art as an Alternative Asset Class: Risk and Return Characteristics of the Middle Eastern & Northern African Art Markets', in Velthuis, O. and Curioni, S. B. (eds) *Cosmopolitan canvases*. Oxford University Press: Oxford. doi: 10.1093/acprof:oso/9780198717744.001.0001.
- Kräussl, R. and Lee, J. (2010) 'Art as an Investment: the Top 500 Artists', *Business*, 31(February), pp. 1–26.
- Kräussl, R. and Logher, R. (2010) 'Emerging Art Markets', *Emerging Markets Review*, 11, pp. 301–318. doi: 10.1016/j.ememar.2010.07.002.
- Kräussl, R. and Van Elsland, N. (2008) 'Constructing the True Art Market Index: A Novel 2-Step Hedonic Approach and its Application to the German Art Market', *Center for Financial Studies*, 2008/11.
- Kräussl, R., Lehnert, T. and Martelin, N. (2016) 'Is there a Bubble in the Art Market?', *Journal of Empirical Finance*, 35, pp. 99–109. doi: 10.1016/j.jempfin.2015.10.010.
- Lancaster, K. J. (1966) 'A New Approach to Consumer Theory', *Journal of Political Economy*, 74(2), pp. 132–157.
- Laubscher, P. (2014) 'A New Recession-Dating Algorithm for South Africa', *Stellenbosch Economic Working Papers*, 06/14.
- Leduc, S. (2010) 'Confidence and the Business Cycle', *FRBSF Economic Letter*, 2010-35(November),



## References

pp. 1–4.

Leduc, S. and Liu, Z. (2016) ‘Uncertainty Shocks are Aggregate Demand Shocks’, *Journal of Monetary Economics*. Elsevier, 82, pp. 20–35. doi: 10.1016/j.jmoneco.2016.07.002.

Leduc, S. and Sill, K. (2013) ‘Expectations and Economic Fluctuations: An Analysis Using Survey Data’, *Review of Economics and Statistics*, 95(October), pp. 1352–1367.

Lucas, R. E. (1980) ‘Two Illustrations of the Quantity Theory of Money’, *American Economic Review*, 70(5), pp. 1005–1014.

Ludvigson, S. C. (2004) ‘Consumer Confidence and Consumer Spending’, *Journal of Economic Perspectives*, 18(2), pp. 29–50.

Mandel, B. R. (2009) ‘Art as an Investment and Conspicuous Consumption Good’, *American Economic Review*, 99(4), pp. 1653–1663. doi: 10.1257/aer.99.4.1653.

Marks, S. (2011) ‘War and Union, 1899–1910’, *Cambridge History of South Africa*. Cambridge University Press.

Martinsen, K., Ravazzolo, F. and Wulfsberg, F. (2014) ‘Forecasting Macroeconomic Variables using Disaggregate Survey Data’, *International Journal of Forecasting*, 30(1), pp. 65–77. doi: 10.1016/j.ijforecast.2013.02.003.

Matheson, T. D. (2010) ‘An Analysis of the Informational Content of New Zealand Data Releases: The Importance of Business Opinion Surveys’, *Economic Modelling*. Elsevier B.V., 27(1), pp. 304–314. doi: 10.1016/j.econmod.2009.09.010.

McClean, L. (2015) ‘Policy Uncertainty in Post-Apartheid South Africa’, *MComm Thesis*. Stellenbosch University.

McMillen, D. P. (2012) ‘Repeat Sales as a Matching Estimator’, *Real Estate Economics*, 40(4), pp. 743–771. doi: 10.1111/j.1540-6229.2012.00343.x.

Mei, J. and Moses, M. (2002) ‘Art as an Investment and the Underperformance of Masterpieces’, *American Economic Review*, 92(February), pp. 1656–1668. doi: 10.1257/000282802762024719.

Mendicino, C. and Punzi, M. T. (2013) ‘Confidence and Economic Activity: The Case of Portugal’, *Banco de Portugal*, (Winter), pp. 39–49.

Mitchell, J., Solomou, S. and Weale, M. (2012) ‘Monthly GDP Estimates for Inter-War Britain’, *Explorations in Economic History*, 49(4), pp. 543–556. doi: 10.1016/j.eeh.2012.06.001.

Moore, G. H. and Zarnowitz, V. (1984) ‘The Development and Role of the National Bureau’s Business Cycle Chronologies’, *NBER Working Paper*, 1349. doi: 10.3386/w1394.

Nagaraja, C. H., Brown, L. D. and Zhao, L. H. (2011) ‘An Autoregressive Approach to House Price

## References

- Modeling', *Annals of Applied Statistics*, 5(1), pp. 124–149. doi: 10.1214/10-AOAS380.
- Naidoo, P. (2013) 'Art Market: Auction Houses Reflect SA', *Financial Mail*, 22 August.
- Nardo, M. (2003) 'The Quantification of Qualitative Survey Data: A Critical Assessment', *Journal of Economic Surveys*, 17(5), pp. 645–668.
- North-West University (2016) 'Policy Uncertainty Index', *NWU School of Business and Governance*, (February).
- Olckers, M., Kannemeyer, C. and Stevenson, M. (2015) 'Art Critic Index: A Proxy for Cultural Value in the Context of the South Africa Art Market', *ERSA Working Paper*, 500(February).
- Organisation for Economic Co-operation and Development (2003) *Business Tendency Surveys: A Handbook*. Edited by E. Giovannini and E. Burgeat. Paris, France: OECD Publications.
- O'Rourke, K. H. and Williamson, J. G. (2004) 'Once More: When did Globalisation Begin?', *European Review of Economic History*, 8(1), pp. 109–117.
- Packenham, T. (1979) *The Boer War*. New York: Random House.
- Pagan, A. (2013) 'Patterns and their Uses', *NCEER Working Paper Series*. School of Economics, University of Sydney, 96(October 2013 October 2013), pp. 1–31.
- Parigi, G. and Golinelli, R. (2004) 'Consumer Sentiment and Economic Activity', *Journal of Business Cycle Measurement and Analysis*, 2004(2), pp. 147–170. doi: 10.1787/jbcma-v2004-art10-en.
- Pellissier, M. (2002) 'Business Confidence and the South African Business Cycle', *Journal for Studies in Economics and Econometrics*, 26(2), pp. 51–67.
- Pellissier, M. and Fusari, A. (2007) 'Volatility as an Indicator of Uncertainty and its Impact on the Realizations of Industrial Business Expectations', *Third Joint EC-OECD Workshop on International Development of Business and Consumer Tendency Surveys*. Brussels.
- Penasse, J. and Renneboog, L. (2017) 'Speculative Trading and Bubbles: Evidence from the Art Market', *CentER Discussion Paper*, 2014-068.
- Persons, W. M. and Coyle, E. S. (1921) 'A Commodity Price Index of Business Cycles', *Review of Economics and Statistics*, 3(11), pp. 353–369.
- Pesaran, M. H., Shin, Y. and Smith, R. J. (2001) 'Bounds Testing Approaches to the Analysis of Level Relationships', *Journal of Applied Econometrics*, 16(3), pp. 289–326. doi: 10.1002/jae.616.
- Phillips, P. C. B., Shi, S. and Yu, J. (2014) 'Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behaviour', *Oxford Bulletin of Economics and Statistics*, 76(3), pp. 315–333. doi: 10.1111/obes.12026.
- Phillips, P. C. B., Shi, S. and Yu, J. (2015) 'Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500', *International Economic Review*, 56(4), pp. 1043–1078.



## References

doi: 10.1111/iere.12132.

Phillips, P. C. B., Wu, Y. and Yu, J. (2011) ‘Explosive Behavior In The 1990S Nasdaq: When Did Exuberance Escalate Asset Values?’, *International Economic Review*, 52(1), pp. 201–226. doi: 10.1111/j.1468-2354.2010.00625.x.

Popescu, A. and Smets, F. R. (2010) ‘Uncertainty, Risk-Taking, and the Business Cycle in Germany’, *CEifo Economic Studies*, 56(4), pp. 596–626. doi: 10.1093/cesifo/ifq013.

Pretorius, F. (1998) *The Anglo-Boer War, 1899-1902*. Cape Town: Don Nelson.

Rabe, J.-M. (2011) ‘Beautiful Bubbles Burst’, *Personal Finance Magazine*, 25 October.

Razzaque, M. A., Osafa-Kwaako, P. and Grynberg, R. (2007) ‘Long-Run Trend in the Relative Price: Empirical Estimation for Individual Commodities’, in Grynberg, R. and Newton, S. (eds) *Commodity prices and development*. Oxford University Press: Oxford, pp. 35–67.

Redl, C. (2015) ‘Macroeconomic Uncertainty in South Africa’, *ERSA Working Paper*, 509.

Renneboog, L. and Spaenjers, C. (2013) ‘Buying Beauty: On Prices and Returns in the Art Market’, *Management Science*, 59(1), pp. 36–53. doi: 10.1287/mnsc.1120.1580.

Renneboog, L. and Spaenjers, C. (2015) ‘Investment Returns and Economic Fundamentals in International Art Markets’, in Velthuis, O. and Curioni, S. B. (eds) *Cosmopolitan canvases*. Oxford University Press: Oxford. doi: 10.1093/acprof:oso/9780198717744.001.0001.

Renneboog, L. and Van Houtte, T. (2002) ‘The Monetary Appreciation of Paintings: From Realism to Magritte’, *Cambridge Journal of Economics*, 26, pp. 331–357. doi: 10.1093/cje/26.3.331.

Rosen, S. (1974) ‘Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition Authors’, *Journal of Political Economy*, 82(1), pp. 34–55.

Rosser, J. B., Rosser, M. V. and Gallegati, M. (2012) ‘A Minsky-Kindleberger Perspective on the Financial Crisis’, *Journal of Economic Issues*, 46(2), pp. 449–458. doi: 10.2753/JEI0021-3624460220.

Ruch, F. (2016) *Second-Round Effects on Inflation, and Underlying Inflation*. PhD thesis. Stellenbosch University.

Ruch, F., Rankin, N. and Du Plessis, S. (2016) ‘Decomposing Inflation Using Micro-Price-Level Data: Sticky-Price Inflation’, *SARB Working Paper Series*, WP/16/07.

SACCI (2011) ‘Updated and Revised SACCI Business Confidence Index’, *South African Chamber of Commerce and Industry*.

Santero, T. and Westerlund, N. (1996) ‘Confidence Indicators and Their Relationship to Changes in Economic Activity’, *OECD Economics Department Working Papers*, 170. doi: 10.1787/537052766455.

Schumann, C. G. W. (1938) *Structural Changes and Business Cycles in South Africa, 1806-1936*.

## References

New York: Staples Press.

Schutt, R. and O’Neil, C. (2013) *Doing Data Science: Straight Talk from the Front Line*. Sebastopol, CA: O’Reilly Media.

Shen, L. (2017) ‘President Trump Is Causing More Economic Uncertainty Than the 2008 Financial Crisis’, *Fortune Magazine*, 19 May.

Shimizu, C., Nishimura, K. G. and Watanabe, T. (2010) ‘Housing Prices in Tokyo: A Comparison of Hedonic and Repeat Sales Measures’, *Jahrbucher fur Nationalokonomie und Statistik*, 230(6), pp. 792–813.

Silberling, N. J. (1923) ‘British Prices and Business Cycles, 1779-1850’, *Review of Economic Statistics*, V(2), pp. 223–235.

Smith, B. D. (1988) ‘The Relationship Between Money and Prices: Some Historical Evidence Reconsidered’, *Federal Reserve Bank of Minneapolis Quarterly Review*, 1232(Summer).

Solar, P. M. and Klovland, J. T. (2011) ‘New Series for Agricultural Prices in London, 1770-1914’, *Economic History Review*, 64(1), pp. 72–87. doi: 10.1111/j.1468-0289.2009.00527.x.

Spaenjers, C., Goetzmann, W. N. and Mamonova, E. (2015) ‘The Economics of Aesthetics and Record Prices for Art since 1701’, *Explorations in Economic History*, 57, pp. 79–94. doi: 10.1016/j.eeh.2015.03.003.

Stigler, G. J. and Sherwin, R. A. (1985) ‘The Extent of the Market’, *Journal of Law and Economics*, 28(3), pp. 555–585.

Stiglitz, J. E. (1990) ‘Symposium on Bubbles’, *Journal of Economic Perspectives*, 4(2), pp. 13–18. doi: 10.1257/jep.4.2.13.

Stock, J. H. and Watson, M. W. (2002) ‘Forecasting Using Principal Components From a Large Number of Predictors’, *Journal of the American Statistical Association*, 97(460).

Taylor, K. and McNabb, R. (2007) ‘Business Cycles and the Role of Confidence: Evidence for Europe’, *Oxford Bulletin of Economics and Statistics*, 69(2), pp. 185–208. doi: 10.1111/j.1468-0084.2007.00472.x.

Tissot, B. (2016) ‘Closing Information Gaps at the Global Level – What Micro Data Can Bring’, *Combining Micro and Macro Statistical Data for Financial Stability Analysis*. Proceedings of the IFC Warsaw Workshop, 14-15 December 2015.

Triplett, J. (2004) *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application To Information Technology Products*. Paris: OECD. doi: 10.1787/643587187107.

Tversky, A. and Kahneman, D. (1974) ‘Judgment under Uncertainty: Heuristics and Biases’, *Science*,

References

---

185(4157), pp. 1124–1131.

United Nations (2015) *Handbook on Economic Tendency Surveys*. New York: United Nations Publications.

Valencia, F. (2017) ‘Aggregate Uncertainty and the Supply of Credit’, *Journal of Banking and Finance*, 81, pp. 150–165. doi: 10.1016/j.jbankfin.2017.05.001.

Van Aarle, B. and Kappler, M. (2012) ‘Economic Sentiment Shocks and Fluctuations in Economic Activity in the Euro Area and the USA’, *Intereconomics*, 47(1), pp. 44–51. doi: 10.1007/s10272-012-0405-z.

Van Walbeek, C. (2006) ‘Official Revisions to South African National Accounts Data: Magnitudes and Implications’, *South African Journal of Economics*, 74(4), pp. 745–765.

Venter, J. C. (2005) ‘Reference Turning Points in the South African Business Cycle: Recent Developments’, *SARB Quarterly Bulletin*, (September), pp. 61–70.

Wang, F. T. and Zorn, P. M. (1997) ‘Estimating House Price Growth with Repeat Sales Data: What’s the Aim of the Game?’, *Journal of Housing Economics*, 6(1997), pp. 93–118. doi: 10.1006/jhec.1997.0209.

Yiu, M. S., Yu, J. and Jin, L. (2013) ‘Detecting Bubbles in Hong Kong Residential Property Market’, *Journal of Asian Economics*, 28(October), pp. 115–124.