



Chisha, Keegan (2025) *Integrated perspectives on macroeconomic dynamics: time series disaggregation, financial shock transmission, output clusters and monetary policy communication complexity*. PhD thesis.

<https://theses.gla.ac.uk/85516/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

Integrated Perspectives on Macroeconomic Dynamics: Time Series Disaggregation, Financial Shock Transmission, Output Clusters and Monetary Policy Communication Complexity

Keegan Chisha

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
DOCTOR OF PHILOSOPHY

ADAM SMITH BUSINESS SCHOOL
COLLEGE OF SOCIAL SCIENCES



OCTOBER 2025

*This thesis is dedicated to my late mother,
Maggie Ndawa, 1969 - 2005.*

Abstract

The absence of high frequency economic data and granular information has historically tended to limit the depth of evidence-based research inquiries in the literature aimed at understanding macroeconomic relationships. This thesis takes an empirical approach to address these gaps by, first, advancing an alternative simple data-driven method for generating high frequency and “flash” estimates for macroeconomic analysis. Second, it departs from the literature by determining linguistic complexity per-topic and showing how overall document complexity can obscure the specific topical elements whose linguistic complexity influence financial market movements. The thesis consists of three main parts:

The first chapter addresses the challenge of converting low frequency time series into high frequency time series. Traditionally, when high frequency data is needed but unavailable, researchers have heavily relied on simple regression-based temporal disaggregation methods to transform low frequency data into higher-frequency estimates. However, these methods face significant limitations - they typically utilise a small number of covariates due to multicollinearity issues and are more suited for explaining statistical relationships than predicting economic outcomes. To address these shortcomings, this study introduces a high-dimensional time series disaggregation approach based on Partial Least Squares (PLS) regression. Using U.S. data from the FRED database, comparative analyses of temporal disaggregation and nowcasting demonstrate that the proposed data-driven approach offers an advantage over traditional methods for generating desired high frequency time series and prediction of flash estimates.

The second chapter constructs missing state-level historical quarterly output growth data for the U.S. by leveraging existing annual output growth data by state, state-specific economic indicators, and the PLS method proposed in the preceding chapter. The estimated historical dataset is used in two subsequent analyses: first, the data is used to analyse the differential impact of an aggregate financial shock on state-level real economic activity; and second, it is used to investigate the interconnectedness of economic activity across U.S. states. The study results confirm differentiated response of state-specific out-

put growth to an aggregate financial shock and that the U.S economy forms three main economic clusters with two small ones. These applications offer valuable insights into the differentiated impact of an aggregate financial shock and the dynamics of economic activity connectedness among the individual states. The results highlight weaknesses of one-size-fits-all monetary or fiscal policy in favour of state-specific policy responses.

In the final chapter, I provide a novel approach that determines complexity of specific contents of the monetary policy statements. I examine the response of the 10-year government bond liquidity to the linguistic complexity of the overall central bank communication, and that of the individual topics. In order to generate evidence, I study the market liquidity outcomes when the level of linguistic complexity of monetary policy statements is low and when it is high for the Reserve Bank of South Africa (SARB). Descriptive results show that, first, liquidity is generally low on announcement days relative to the 10-day control day. Density distributional shifts to the right reveal even lower liquidity in the bond market, particularly, when linguistic complexity is high. The sensitivity of liquidity is more when the complexity of forecasting is high. Volatility of liquidity is also found to respond to linguistic complexity of the SARB communication. Main regression results indicate that liquidity is particularly responsive to complexity of overall communication and some specific topics - for example, Forecasting. High communication complexity of forecasting is found to lead to higher quoted spreads suggesting disrupted trading. These findings suggest that market participants respond distinctly to different levels of specific topic communication clarity, providing central banks practical insights to target improvements that promote stable market and economic outcomes.

Therefore, this thesis contributes to the existing literature in three distinctive ways. First, it addresses the limitations of traditional methods of temporal disaggregation of time series by proposing a simple novel, data-driven approach. Second, adds to the existing literature by producing fresh evidence confirming the differential impact of an aggregate financial shock on state-specific output growth. Finally, contributes a novel approach to grade central bank communication complexity by topic, improving financial market understanding and presenting opportunities for pinpointed improvements in central bank communication.

Contents

Abstract	iii
Acknowledgements	xiii
Declaration	xiv
1 Introduction	1
1.1 Motivation	2
1.2 Research Contribution	4
1.3 Thesis Outline	6
2 Enhancing Temporal Disaggregation of Time Series: A Data-Driven Approach Using Partial Least Squares	9
2.1 Introduction	10
2.2 Context of the TS Disaggregation Problem	13
2.3 Review of the Related Literature	17
2.3.1 The Temporal Disaggregation Problem	17
2.3.2 Theoretical Literature on Temporal Disaggregation	21
2.3.3 Empirical Literature on Temporal Disaggregation	34
2.4 Towards Data-Driven Temporal Disaggregation	39
2.4.1 Explanatory versus Predictive Modeling	39
2.4.2 The LASSO Regression Model	40
2.4.3 The Partial Least Squares Regression	42
2.4.4 Cross Validation	46
2.4.5 Forecast Combination	46
2.5 Empirical Results	47

2.5.1	Data Description	47
2.5.2	Performance Evaluation Strategy	49
2.5.3	Temporal Disaggregation of Narrow Indicators	53
2.5.4	Temporal Disaggregation of Broad Indicators	56
2.5.5	Additional Results: GDP Nowcasting	60
2.6	Conclusions	64
	Appendix	66
A	Low Frequency Vs High Frequency Time Series	66
B	Predictor Indicators for the LASSO and PLS	67
C	Application in Data Scarce Environments	69
C.1	Mapping and Characterising Data Gaps	69
C.2	High Frequency Proxies	70
C.3	Tapping Open and Commercial Data Sources	70
D	Strengthening Data Sources in LDCs	70
E	Example of Types of Economic Data	75
F	Delays in GDP Reporting	80
G	Narrow vs Broad Economic Indicators	81
3	Differential Effects of a Financial Shock and the Real Economic Activ-	
	ity Network of the United States	83
1	Introduction	84
2	Context of the States' Data Problem	86
2.1	National Aggregate and State Indicator Selection	86
2.2	Modified Traditional Methods Vs the PLS Method	90
2.3	Applying the Modified Model-Based Methods	92
3	Constructing Historical Data	96
3.1	GDP and the Related Indicators	96
3.2	Accuracy Assessment Strategy and Measures	97
4	Differential State Effects of Financial Disruptions	103
4.1	Background and Motivation	103
4.2	Finance and the Real Economy	104

4.3	Financial Shock and Individual State Macroeconomy	105
4.4	Empirical Approach: State-Level Economic Implications	106
4.5	VAR Empirical Results	111
5	Network of the U.S. Real Economic Activity	115
5.1	Background and Motivation	115
5.2	Network Analysis	116
5.3	Economic Activity Interconnectedness	117
5.4	Empirical Results	119
6	Conclusions	124
	Appendix	126
A	Why 48 Contiguous U.S.States	126
B	Some Econometric Methods for “Backcasting”	129
B.1	Proportional Denton and Benchmarking Methods	129
B.2	Regression Model-Based Backcasting Methods	129
B.3	U-MIDAS and Mixed-Data Sampling (MIDAS)	130
B.4	The Kalman Filter and State-Space Models	130
B.5	Dynamic Factor Models	131
B.6	Machine Learning and Non-Parametric Techniques	131
C	Estimated High Frequency Time Series	154
D	Structural VAR Model	157
4	Topic-Specific Complexity in Monetary Policy Statements: Implications for Liquidity and Volatility in Financial Markets	160
1	Introduction	161
2	Monetary Policy in South Africa	166
2.1	About the SARB	166
2.2	The SARB MPC Policy Communication	168
3	Literature Review	170
4	Data, Methodology and Descriptive Analysis	179
4.1	Measuring Text Complexity	181
5	Empirical Results	186

5.1	Overall and Per-Topic Complexity	186
5.2	Identifying the SARB Communication Effect	189
5.3	Factors Explaining the SARB Communication Effect	197
5.4	Additional Results	207
6	Conclusions	210
	Appendix	213
A	The SARB Key Features	213
B	The SARB MPS (1999 - 2023)	215
C	Monetary Policy Cycles and Key Drivers	218
D	SARB Policy Rate Overtime	221
E	Supplementary Related Literature	222
F	Other Readability Measures	223
G	Liquidity, Volatility, and Complexity Trend	226
G.1	Liquidity of Financial Variables	226
G.2	Volatility of Financial Variables	226
G.3	The Mann-Kendall Test	228
H	Topic-Term Representation	229
5	Concluding Remarks	240

List of Tables

2.1	Explanatory vs. Predictive Modeling	40
2.2	Narrow Economic Indicators in Traditional Procedures	49
2.3	Temporal Disaggregation Results: Narrow Indicators	55
2.4	Temporal Disaggregation Results: GDP	60
2.5	Nowcasting Results During GFC	62
2.6	Nowcasting Results During SP	62
2.7	Nowcasting Results During COVID-19	63
2.8	Predictor Indicators for the LASSO and PLS	67
2.9	Economic Sectors and Key Data Types	75
3.1	Correlations between Output, Employment, and IP Index	89
3.2	Correlations between National and State Output Growth	90
3.3	National and State Low Frequency Models	93
3.4	Framework for Estimating State High Frequency Output	94
3.5	VAR Endogenous Variables	111
3.6	State-Specific Economic Variables	128
3.7	Mean Absolute Error (1991 to 2004)	132
3.8	Network Centrality Parameters	159
4.1	Summary of SARB MPC Statements	180
4.2	The Mann-Kendall Trend Test	186
4.3	Comparison of Mean Liquidity	194
4.4	Comparison of Liquidity Volatility	197
4.5	Results of 10-Minute Regression	201
4.6	Results of 15-Minute Regression	202
4.7	SARB MPC Statements Word Count, 1999 - 2023	217

4.8	Monetary Policy Cycles and Key Drivers	220
4.9	10-Minute Regression Addendum	233
4.10	Results of 10-Minute Interaction Regression	234
4.11	15-Minute Regression Addendum	237
4.12	Results of 15-Minute Interaction Regression	238

List of Figures

2.1	CSPE for Narrow Indicator Estimations	57
2.2	CSPE for Broad Indicator Estimations	59
3.1	National Annual Output, Employment and IP Index	89
3.2	National and State Output Growth Rates	90
3.3	Discrepancy Aggregate Ratio	100
3.4	National Impulse Response Functions	112
3.5	State Impulse Response Functions	113
3.6	Period 1 Real Output Network	121
3.7	Period 2 Real Output Network	122
3.8	U.S Real Economic Activity Clusters	123
3.9	State Discrepancy Aggregates	134
3.10	Conformity of State Output Estimates	138
3.11	Actual LF vs Estimated HF Output Growth	154
4.1	Governor-Specific Word Clouds	169
4.2	Topic Dictionaries	185
4.3	Overall and Per-Topic Communication Features	188
4.4	Liquidity Density Distributions	192
4.5	Volatility Density Distributions	196
4.6	10-Minute Dummy Coefficients	205
4.7	15-Minute Dummy Coefficients	206
4.8	Dummy Coefficients by Complexity Category	209
4.9	SARB Policy Rate Changes	222
4.10	Word Clouds of the Topics	230
4.11	Liquidity Comparisons	231

4.12 Volatility Comparisons	232
---------------------------------------	-----

Acknowledgements

My deepest gratitude goes to my primary supervisor, Professor Dimitris Korobilis, for his invaluable guidance and unwavering support throughout my research journey. His mentorship has been instrumental in my personal growth and the successful completion of this work. I am also profoundly grateful to Dr. Juhee Bae, and the members of my examining committee, Dr. Tuuli Vanhapelto, Professor Jim Malley, and Professor Nicola Viegi for their time and insightful feedback.

Deep gratitude goes to the Bank of Zambia for generously sponsoring my PhD studies in their entirety, making this academic achievement possible. A special thank you to Dr. Luigi Gifuni, Lecturer at Strathclyde University, who co-authored Chapter 3 with me. Thanks to Dr. Santiago Montoya-Blandon, Dr. Anthoulla Phella, and Professor John Tsoukalas for many helpful comments and suggestions at different stages of my studies.

A special thank you to Dr. Jonathan Chipili and Dr. Francis Chipimo of the Bank of Zambia for their support and encouragement before and during my studies. Warm appreciation also to a good friend, Mr. Sydney Chauwa Phiri for the numerous conversations about economics and social mobility wisdom.

Heartfelt thanks to my family, especially my cousin Dr. Kalonde Kasengele and my uncle Dr. Mwango Kasengele, who have always supported me in my education and career. Loving remembrance goes to my late mother, Maggie Ndawa and late cousin Patricia Musonda for believing in me and helping me conceive what was possible.

Most importantly, deepest thanks to my wonderful partner and best friend, Tizyo Mary Bulaka, and my children Lukundo Maggie Chisha, Kundananji Keegan Chisha, and Iwwananji Ezekiel Chisha for their unending love and encouragement. You fulfill my life!

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Keegan Chisha

Chapter 1

Introduction

1.1 Motivation

Accurate and timely economic data serves an essential function for both analysing economic trends and evaluating economic conditions. Timely and representative time series data also establishes the basis for developing economic models and forecasts while allowing experts to understand how different economic systems relate. Therefore, time series data provide essential insights needed to develop evidence-based policies and guide strategic actions while promoting improved economic planning methodologies. In addition, without detailed high frequency data and granular information, analysts and policy makers struggle to generate evidence of complex macroeconomic relationships and trends. The lack of data often become apparent in a variety of forms, such as when there is a need to understand short-term fluctuations or when there is a demand for more localised information on the impact of, for example, an aggregate economic shock. In both examples, having access to more precise and frequent data becomes crucial for effective analysis, decision making, and focused policy.

High frequency data has long been vital for tracking economic short-term shifts as well as evaluating the immediate outcomes of specific events such as monetary policy changes and financial disturbances. Short-term fluctuations in economic indicators lead to material consequences, which makes it crucial to track and address these changes promptly. Economic data related to broad economic indicators frequently remains inaccessible or becomes too costly to acquire when high frequency measurements of especially broad indicators are required. Econometricians and statisticians have created multiple workaround techniques to address this challenge and fill this gap. The methods consist of innovative techniques that disaggregate low frequency time series data and transform them into consistent high frequency data providing detailed insights into economic activity (see, for example, Sax and Steiner (2013)). Although continuous advances in data management practices continue to evolve, traditional estimation methods do not maintain integrity and robustness to handle large information. Put differently, traditional methods of temporal disaggregation fail to offer enough precision and robustness to make accurate predictions with large high frequency time series datasets, thus creating a serious barrier in economic analysis.

For example, in the literature, research on how the aggregate economic shocks affect economic activity of individual states in the U.S. has faced apparent limitations due to the absence of high frequency economic growth data because available quarterly output growth time series data starts from 2005 and, therefore, too short for meaningful econometric analyses of economic systems. The lack of sufficiently long high frequency time series data on regional output means that critical questions about the differentiated responses of individual state economic activity to aggregate disturbances remain largely unanswered (see, for example, Gilchrist and Zakrajšek (2012b)). In the literature, the absence of long enough desirable time series data has, also obstructed researchers from directly examining specific questions such as, how, for example, the U.S. regions interact economically, cluster and influence each other's performance (see, for example, Hamilton and Owyang (2012) and Brownlees et al. (2022)). In response, this thesis, first, introduces a new data-driven approach that can be used to generate high frequency time series based on available low frequency data and high frequency related indicators, and second, apply the new approach to generate missing historical state-specific quarterly output data and conduct nuanced macroeconomic analyses. This novel approach to performing temporal disaggregation of time series aims to produce highly accurate missing high frequency data and therefore, help to enhance macroeconomic analysis capabilities and deliver practical insights for both national and regional policymakers.

Finally, the existing body of research has long addressed how central bank communication affects financial markets, but has focused primarily on the effect of overall policy communication rather than the topical dimensions covered in policy documents. For example, the existing literature typically reports on how financial markets respond to the linguistic complexity of overall monetary policy communication, rather than evaluating the effects of linguistic complexity of specific topics within the policy communication documents. This common approach has significant limitations, for instance, this approach conceals how specific topic clarity or complexity affects financial market instruments. That is because the traditional approach essentially assumes that all key topics in central bank communication have the same level of linguistic complexity, thus preventing researchers

and policy makers from recognising that different topic-specific complexities could impact financial market instruments to varying extents. As a result, monetary authorities that aim to simplify their policy communication miss out on pursuing a more targeted approach to achieving this objective.

1.2 Research Contribution

The structure of this thesis features three main chapters that add distinct perspectives to current academic research. Each chapter addresses a unique research problem and fills an identified gap within the field through a systematic approach driven by clearly defined goals. The primary focus of the first key work concerns the problem of temporal disaggregation of time series data. This chapter introduces a new temporal disaggregation method based on existing research suggestions from the European Statistical System (ESS) as described in Buono et al. (2018). The research presented in this chapter emerges from a substantial foundation of literature which has defined the field of temporal disaggregation of time series in economics. This study builds upon the essential contributions of leading researchers, including but not limited to Chen et al. (2007), Moauro and Savio (2005), Sax and Steiner (2013) and Quilis (2018), especially through their application to economic data. This thesis builds on previous work by implementing a new but simple approach that improves the time series disaggregation process and ensures more precise estimated high frequency time series. This study expands existing methods from previous research by implementing a stronger procedure appropriate for the modern demands of macroeconomic analysis. This chapter establishes its importance in the current debate in the field through its adoption of ESS suggestions combined with previous experts' points of view found in the literature. Moreover, the current proposal has the beauty of simplicity, as it retains the original temporal disaggregation framework.

The thesis transitions into constructing a quarterly historical output growth dataset for U.S. states for the period prior to 2005. The aim is to bridge a data gap and a critical research limitation faced by previous studies that had to rely mainly on proxy indicators such as employment data to study the economic performance of individual states. In the literature, the lack of quarterly output data for individual states has forced researchers

to rely on proxy variables for their analyses. The new data set created in this chapter helps fill this research void and enables me to examine state-level economic performance directly over time. This new development responds to earlier research limitations in which national-level data or proxy variables did not adequately represent the complex economic conditions at the state level. This work presents a strong and direct historical data set on output growth to deepen our understanding of regional economic dynamics. This research facilitates a detailed examination of the unique ways in which an aggregate financial shock affects the economic activity of different states. This research draws connections to earlier investigations such as Gilchrist and Zakrajšek (2012b)'s study of the U.S. economy's responses to financial disruptions. Further, the contributions from Hamilton and Owyang (2012) and Brownlees et al. (2022) substantially enhanced my comprehension of regional economic propagation patterns and economic clusters among states. This chapter builds upon this earlier work and presents a state-specific data generation approach to understand regional economic systems while enabling detailed investigation of economic phenomena unique to each state.

The concluding chapter presents an innovative method for assessing the linguistic complexity present in monetary policy communication discourse. Previous studies have typically measured the linguistic complexity of central bank communications in a broad sense while this thesis analyses linguistic complexity through the lens of different specific key topic dimensions. The primary focus of existing studies is to analyse the complexity of monetary policy statements generally without examining the subtleties of individual topics. This chapter advances foundational research by examining the effects of topic complexity in policy statements on financial market outcomes. The inquiry uses specific topic analysis to achieve precise evaluations of central bank communication elements that influence market reactions. The application of this approach to central bank monetary policy statements enables me to achieve a more precise assessment of how financial markets respond to linguistic intricacies in such communications. The research extends current knowledge by building on the foundational work of Bulíř et al. (2013) who examined central bank communication drivers as well as Chung et al. (2013) who investigated how liquidity interacts with communication strategies. The research conducted by Smales and Apergis (2017b) and Smales and Apergis (2017a) expanded the body of knowledge on

the impact of communication complexity on market reactions during monetary policy disclosures. This chapter presents a novel approach that provides fresh insights on how specific topics within central bank communication affect market dynamics while providing an improved method to evaluate and simplify monetary policy communication.

1.3 Thesis Outline

The structure of the main chapters is as follows:

Chapter 2 is titled *“Enhancing Temporal Disaggregation of Time Series: A Data-Driven Approach Using Partial Least Squares”*. In this study, I introduce a new approach for performing temporal disaggregation of macroeconomic time series data using partial least squares (PLS) regression. In order to appreciate its performance, a horse race is conducted with existing traditional methods that include Chow-Lin, Fernandez, Litterman, Silva-Santos Cardoso (SSC), and a recently proposed data-driven method based on LASSO regression. The PLS and LASSO approaches are inherently predictive methods while the traditional methods are explanatory methods. With regards data, it is downloaded from the FRED database for the period between 1960 and 2020. Key results show that in the majority of temporal disaggregation and nowcasting experiments, the data-driven methods, in particular, the PLS approach is superior in both converting low frequency time series to high frequency time series, and nowcasting. The superior performance is evident through relatively smaller cumulative prediction errors and root mean square errors (RMSE). Strong performance is also evident for extreme periods when accurate predictions are more challenging to achieve. Therefore, this study contributes to the literature, a new data-driven temporal disaggregation approach that improves the accuracy and reliability of temporally disaggregated time series.

Chapter 3 is titled *“Differential Effects of a Financial Shock and the Real Economic Activity Network of the United States”*. This study constructs historical state-specific output growth data and applies it to understanding, first, the transmission of an aggregate financial disruption shock on the economic activity of individual states of the U.S., and second, the output interconnectedness and economic clusters among the states. More

specifically, the study starts by constructing the missing historical quarterly real output growth time series for each of the 48 contiguous states for the period between 1991 and 2004. This is achieved using the PLS approach and state-specific high frequency related indicators. Initial data estimation results show that aggregated estimates of the quarterly time series from the PLS approach closely match the observed annual time series in the absence of fulfilling the temporal constraint - have the smallest error. Using this robust data, I conduct economic analysis using impulse response functions based on a VAR model and a concatenated output dataset (1991-2022). The main results indicate that an aggregate financial shock affects output of specific states differently, with some states' performance showing greater resilience. Finally, the study demonstrates that, over time, the states have increased connectedness among themselves and form three large and two small interconnected communities based on economic activity.

Chapter 4 is titled "*Topic-Specific Complexity in Monetary Policy Statements: Implications for Liquidity and Volatility in Financial Markets*". In this final chapter, I contribute a novel approach that measures complexity of specific topics in the monetary policy communication. Focusing on South Africa, I initially identify the key topics covered by the South Africa Reserve Bank in its monetary policy communication then measure the respective linguistic complexities. This is followed by an examination of the response of the 10-year government bond liquidity to the linguistic complexity of the overall policy statement, and that of the individual topics on the announcement day. The distributional densities of means and standard deviations of the bond on the day of announcement and the 10-day control indicate that liquidity is significantly different when the level of linguistic complexity is low and when it is high. Main regression results indicate that liquidity is particularly responsive to complexity of overall communication and some specific topics - for example, Forecasting. The high complexity of forecasting communication is found to lead to higher quoted spreads, suggesting disrupted trading activity. In this way, this chapter develops a nuanced approach that could be used to assess the linguistic complexity of policy communication based on topical dimensions. The research reveals that traditional methods that measure the complexity of overall communication may overlook key topics that influence the movements of financial market instruments while providing

a more precise understanding of the relationship between communication of individual topics and the South African 10-year government bond. These findings provide central banks with fresh practical evidence to focus on improving the communication of specific topics that promote stable market and economic outcomes.

Chapter 2

Enhancing Temporal Disaggregation of Time Series: A Data-Driven Approach Using Partial Least Squares

2.1 Introduction

Temporal disaggregation of time series is an important technique that is used by many national statistical agencies globally to address gaps in high frequency macroeconomic data. It is also commonly used to ensure data consistency, produce “nowcasts” or timely high frequency “flash” estimates and facilitate evidence-based policy analysis¹. Widely applied across the European Statistical System (ESS) in countries such as France, Italy, Belgium, Portugal, and Spain, temporal disaggregation techniques are essential for producing high frequency estimates of key macroeconomic indicators such as Gross Domestic Product (GDP) (see, for example, Quilis (2018)).

Despite the increasing need for quick, precise “flash” estimates of GDP and other aggregates, current temporal-disaggregation techniques perform poorly when there are a lot of predictors available and relationships deteriorate under pressure - such as economic crises. Is it possible for high-dimensional, shrinkage-based estimators to enhance the accuracy of temporal disaggregation of time series while avoiding the problems of multicollinearity and feature selection? Existing temporal disaggregation approaches suffer at both ends of the data spectrum. First, the regression-based methods are typically known to encounter multicollinearity problems, particularly when dealing with a large number of potential predictor macroeconomic variables (see Moauro and Savio (2005), Chen et al. (2007), and Chamberlin (2010)). As a result, these approaches often utilise limited information by using a small number of related variables, typically one to three related indicators. In addition, they have inherent challenges such as the difficulty of selecting predictor variables and determining the optimal number of predictors. Moreover, these methods are unreliable when macroeconomic relationships among interest variables break down during extreme periods. This may lead to suboptimal predictions for the target high frequency time series².

1. see Buono et al. (2018)

2. Stock and Watson (2006) argues that many predictors provide some resilience against the structural unpredictability that troubles low-dimensional predictions.

Second, while some areas of macroeconometrics, such as forecasting, have already established the prediction corrections, incorporating high-dimensional time series through sophisticated techniques like feature selection and Bayesian shrinkage methods, see, for example, Stock and Watson (2002a), Stock and Watson (2002b), Bai (2003), Boivin and Ng (2006), Li and Chen (2014), Groen and Kapetanios (2016), and Korobilis (2021), temporal disaggregation processes have lagged. Most temporal disaggregation regression-based approaches still rely on single indicators or indeed a small set of related indicators which limits their use of available data in their data conversion process. They are also incompatible with short time series data as they produce unreliable parameters.

However, recent research insights suggest that when high-dimensional data is available, estimation techniques such as the Ridge, Bayesian, Principal Component (PC), Partial Least Squares (PLS), and Least Absolute Shrinkage and Selection Operator (LASSO) regressions among others offer promising solutions to addressing the inherent multicollinearity problems associated with traditional regression-based models (Tibshirani (1996), De Mol et al. (2008), and Ahn and Bae (2022)). Therefore, these relatively large data-intensive techniques also represent a natural pathway for developing temporal disaggregation methods that are best suited for predicting missing high frequency time series³.

Further, in addition to being recommended in the ESS guidelines by Buono et al. (2018), others have also noted, in particular, the potential accuracy gains of predicting time series using such methods as PLS regressions. For example, Groen and Kapetanios (2016) demonstrated that the PLS regression can potentially perform as well as Bayesian regression in forecasting macroeconomic variables. This is an important argument in support of the proposal that this study makes later on. Thus, while traditional regression-based temporal disaggregation methods have their limitations, the availability of large datasets, combined with the opportunities presented by data shrinkage estimation methods, helps identify a clear research gap. These factors not only motivate and position the current study but also underscore its key contribution to the literature.

3. See Appendix A.1 for the description and examples of the difference between low frequency and high frequency time series.

More specifically, the main motivation for this study is to propose a simple novel approach that improves temporal disaggregation outcomes by integrating vast amounts of information to predict high frequency target variables. The study directly addresses the key question of whether the accuracy of temporal disaggregation in time series can be elevated by leveraging high-dimensional data while circumventing issues of multicollinearity among predictors and the challenge of feature selection. In particular, the study proposes a disaggregation approach within the context of “predictive modeling” instead of the traditional “explanatory modeling” which is widely used. In addition, the proposed approach remains theoretically justified, as it effectively uncovers latent factors driving the target variable, which is conceptually understood as identifying underlying economic forces and similar to factor models common in macroeconomics. The inquiry therefore is done mindful of carefully selecting timely and informative related indicators since this is crucial in minimising noise and to achieving quality in-sample temporally disaggregated time series and nowcasting in the large data setting.

To answer the central research question, this study takes an empirical approach that reforms the temporal disaggregation estimation approach by applying the PLS regression to estimate the predicted or systematic component instead of the usual Ordinary Least Squares (OLS) regression in traditional methods. The PLS-based method fits within the standard temporal disaggregation framework. This method may be seen as a direct competitor to the recent LASSO technique proposed by Mosley et al. (2022). As such, this work by Mosley et al. (ibid.) represents the most recent advancement in the literature that is closely related to this study.

In light of the foregoing, current research makes two primary contributions: First, it proposes integration of a PLS regression temporal disaggregation approach in the econometric toolkit for estimating high frequency time series when large data of related variables is available. Second, it highlights the effectiveness of data-driven methodologies compared to traditional approaches in temporal disaggregation and nowcasting, particularly in scenarios characterised by abundant high frequency predictors.

The efficacy of the proposed PLS regression-based procedure is demonstrated through performance evaluations against the traditional and LASSO approaches of temporal disaggregation in three different applications. The comparison strategy adopted here is similar to the one followed in majority of temporal disaggregation literature, for example, Silva and Cardoso (2001), Miralles et al. (2003), Moauro and Savio (2005), Chen et al. (2007), Chamberlin (2010), and Mosley et al. (2022).

Key empirical findings suggest that data-driven approaches generally perform better compared to traditional methods in the majority of cases studied. In particular, the PLS procedure shows relatively superior performance in disaggregating both narrow and broad macroeconomic indicators, outperforming traditional methodologies and the LASSO procedure. The strong performance of the PLS procedure may reflect its ability to capture common trends from many indicators. These findings therefore underscore the potential benefits of using data-driven techniques to improve accuracy in temporal disaggregation, particularly in data-rich environments. To my knowledge, this study is the first attempt to explore and document the potential accuracy gains associated with the PLS approach in the temporal disaggregation of time series within data-rich environments.

The rest of this study is organised as follows: Section 2 provides the context of the temporal disaggregation problem; Section 3 discusses the relevant literature; then, Section 4 discusses different aspects of temporal disaggregation addressed in empirical research and presents the proposed methodology for high dimensional time series temporal disaggregation; next, Section 5 discusses the data collection and experimental results; and lastly, Section 6 offers a conclusion.

2.2 Context of the TS Disaggregation Problem

The modern economy consists of numerous sectors that create extensive high frequency indicators which provide data at regular short intervals. These indicators span multiple economic areas and provide essential knowledge about different aspects of economic performance. Indicators may consist of industry-specific data, for instance, employment data that reveal workforce trends in specific industries or indeed sectoral production levels

that monitor output across different economic sectors, together with retail and wholesale transaction volumes that provide insight into consumer and business activity⁴. For example, studies such as Mosley et al. (2022) have shown that road and port traffic data are an essential high frequency indicator because they measure goods movement and logistics performance (see, for example, Appendix Table 2.9 for an extensive list of examples of sectoral data). Nowcasting uses these narrow economic activity indicators to estimate real-time economic conditions. High frequency indicators often serve as essential input for nowcasting to estimate comprehensive economic indicators which are later used to develop historical high frequency broad growth metrics such as GDP (see, for example, Li et al. (2020)).

Traditional economic growth measures, including GDP, typically experience significant reporting delays as they are released infrequently through quarterly or annual intervals. The delayed reporting intervals of economic data can obstruct effective decision-making processes in rapidly changing economic environments⁵. Even though high frequency data provide potential to deliver timely insights, traditional temporal disaggregation methods continue to serve as an essential tool for generating high frequency time series and ensuring consistency therein. Traditional techniques thus enjoy widespread application, yet rely on basic models with limited inputs when contrasted with modern methods. Traditional temporal disaggregation methods generate less precise estimates due to their inherent simplifications, as noted by Buono et al. (2018).

The limitations of current traditional methods demonstrate an urgent need to change the temporal disaggregation of time series approach. Traditional estimation strategies must be implemented alongside advanced transformed methodologies that can effectively process high-dimensional data composed of many correlated variables. The design of new advanced tools or techniques becomes crucial in overcoming the weaknesses of traditional methods. Fit-for-purpose methods must effectively extract maximum insight from extensive data sets of high frequency indicators with possible interrelationships. Recent academic research shows that forecasting models, for example, using a wider range of predictors enable researchers to access more extensive information than traditional time series explanatory

4. See Appendix A for a brief discussion on low frequency vs. high frequency time series data.

5. A detailed account on why GDP is typically reported with delays is presented in Appendix A.6

models usually provide. This observation demonstrates the necessity of implementing advanced data-driven methods to increase both accuracy and reliability in the temporal disaggregation of time series and nowcasting operations (see, for example, Stock and Watson (2006)).

Furthermore, although advanced techniques offer clear benefits, in the literature, researchers have not yet reached consensus on the standards required for selecting related indicators, particularly in traditional temporal disaggregation methods. The literature contains numerous discussions on this subject, but does not provide a definitive “decision rule” to select the optimal number of related covariates. The complexity of this issue usually stems from the extensive range of potential predictors related to specific target variables and the lack of theoretical frameworks to effectively prioritise the related indicators.

The persistence of the challenge of determining the optimal number of related indicators generates numerous questions that need to be explored in future studies. Researchers must determine whether the use of high-dimensional data can enhance temporal disaggregation through shrinkage methods and other data-driven approaches. The use of advanced methods to manage intricate high-dimensional data sets holds the promise of enhancing time series disaggregation accuracy and reliability by surpassing the limitations inherent in conventional techniques. In addition, researchers must evaluate whether they should perform multiple disaggregation exercises using a complete set of candidate predictor variables and then choose the optimal model based on performance outcomes. When using traditional methods, it is widely accepted that the selection of related variables requires validation through theoretically justified assumptions that clearly outline macroeconomic relationships between the target variable and indicators. However, other questions still emerge, for example, should selection of related indicator series depend only on statistical association strength with the target variable or require supplementary selection criteria?

In practice, any time series disaggregation task faces complex and difficult theoretical and practical questions because researchers can examine many potential related indicators. Therefore, the absence of widely recognised standards for selecting related variables adds complexity to the selection process. In practice, researchers have continuously pursued various methods to arrive at the optimal predictor variables for every analysis. More

specifically, the literature shows that researchers choose related indicator series based on the degree of correlation between each indicator and the target variable. Researchers rely on fundamental information about the relationship between their target and covariates when selecting variables and incorporate their personal preferences into this process (Chamberlin (2010)). Researchers generally address the lack of standardised selection rules by evaluating numerous combinations of variables alongside different methods. Researchers select the variable combination that leads to the most precise high frequency time series by evaluating accuracy through standard measures such as the Mean Squared Error (MSE). They often use this trial-and-error method because it works well, but it is time-consuming, and thus the need for efficient indicator selection methods that follow solid theoretical foundations.

Although significant progress in temporal disaggregation has been made so far, persistent dependence on traditional methods exposes several research gaps that require additional scrutiny. The process of selecting suitable predictors, together with determining the optimal number of related covariates, remains problematic. To address these difficulties, researchers must conduct more studies on high-dimensional data analysis techniques in conjunction with the formulation of better-structured guidelines to choose related indicators. Data-driven approaches operating in high-dimensional spaces have already demonstrated a huge increase in both the accuracy and reliability of estimations. Addressing these open research questions is essential to progress in temporal disaggregation of time series as well.

2.3 Review of the Related Literature

This section is divided into two parts, and its purpose is to present an overview of the theoretical and empirical literature across time. More specifically, the section will have a particular emphasis on shedding light on the difficulties that are connected with the temporal disaggregation of time series using traditional approaches. The current study seeks to clarify and justify the proposed shift from existing approaches to data-driven techniques by highlighting differences in estimation strategies and facilitating a greater appreciation of the placement of the current study. This is accomplished by thoroughly exploring the historical background and evolution of the literature.

As earlier showed, the identified gap in the literature suggests that the unique contribution of this research is the predictive approach of the fitted or predicted component of the high frequency values within the existing general temporal disaggregation framework. In contrast with Ordinary Least Squares (OLS) regression model-based methods, a data-driven approach effectively reduces parameter estimation variance, thereby enhancing in-sample prediction accuracy.

2.3.1 The Temporal Disaggregation Problem

The literature demonstrates that temporal disaggregation techniques have proven to be a useful workaround for data problems such as missing high frequency time series, as evidenced by the widespread use of the methods (Eurostat and Buchow (2013) and Buono et al. (2018)). This is despite the fact that there is no way to completely compensate for missing data in situations where high frequency time series are not available but are required. According to the definition provided by Sax and Steiner (2013), temporal disaggregation is a process that seeks to identify an unidentified high frequency series whose sums, averages, and first or final values are in agreement with a low frequency series that is already known. Estimating a new high frequency time series that is compatible with the low frequency data that has been observed is the fundamental goal of all of the many strategies that are used for temporal disaggregation.

In situations where temporal disaggregation is carried out with the assistance of a related indicator, it is carried out in a manner that maintains the short-term changes that are present in the higher frequency indicator series. Regarding both the problem and the context, this study is also concerned with both of these aspects.

In this section, therefore, the study starts by formally formulating the temporal disaggregation problem according to the definition above and in line with most of the econometrics literature for interpolation, distribution and extrapolation of low frequency series. Below are some general notations to be adopted throughout this study⁶:

Y_t , ($t = 1, \dots, n$) is the observed temporal disaggregates of the missing disaggregated time series, y_t , ($t = 1, \dots, m \times n$) where m is the conversion frequency. So, for temporal disaggregation of yearly series into quarterly series Y_t , ($t = 1, \dots, n$), y_t ($t = 1, \dots, m \times n$), $m = 4$ or quarterly series into monthly, $m = 3$ and lastly annual series into monthly series $m = 12$.

Where temporal disaggregation is done with the help of one or more related series that are known at a higher frequency level, then the information on the related series can be summarised in matrix form by X . Assuming that there are k related series, that is $X = (x_1, \dots, x_k)$, then X will assume the dimensions $mn \times k$ and $x_i = (x_{i,1}, \dots, x_{i,mn})'$, $i = 1, \dots, k$ is a column vector representing data on the i^{th} related time series. In this regard, the temporal disaggregation problem is to derive an estimate y that makes use of information contained in the x_i 's which is consistent with the low frequency, Y .

To ensure consistency between y and Y , the type of data to be disaggregated whether it is a flow or stock, must be clearly specified. Consistency is therefore, achieved through the following equation:

$$Y = Cy, C = I_n \otimes c', c = (w_1, \dots, w_m)'$$

Where c are m -dimensional column vectors of w_i' is representing the weights which must be specified in each specific temporal disaggregation problem. Therefore, for flow data:

6. Notation of the temporal disaggregation presented in this paper follows that of Trabelsi and Hedhili (n.d.) in their OECD working paper in which they develop an alternative temporal disaggregation model and compare its performance to the three popular models, Chowlin, Fernandez and Litterman methods

$$c = (1, \dots, 1)'$$

$$Y_t = \sum_{i=1}^m y_{m \times (t-1) + i} = (1 + B + \dots + B^{m-1})y_{m \times t}$$

$$t = 1, \dots, n$$

B is the backshift operator such that: $B^k y_t = y_{t-k}$. Assuming temporal disaggregation of annual series to quarterly for all cases below, thus $m = 4$, consistency is achieved when

$$Y_t = \sum_{i=1}^4 y_{m \times (t-1) + i}$$

$$t = 1, \dots, n$$

$$c = (1, 1, 1, 1)'$$

.

For index flow data c takes the form:

$$c = \frac{1}{m}(1, \dots, 1)'$$

$$Y_t = \frac{1}{m} \sum_{i=1}^m y_{m \times (t-1) + i} = \frac{1}{m}(1 + B + \dots + B^{m-1})y_{m \times t}$$

$$t = 1, \dots, n$$

Similarly here, consistency is achieved when

$$Y_t = \frac{1}{4} \sum_{i=1}^4 y_{m \times (t-1) + i}$$

$$t = 1, \dots, n$$

$$c = \frac{1}{4}(1, \dots, 1)'$$

.

For stock data observed as the first value of a given period

$$c = (1, \dots, 0)'$$

$$Y_t = y_{m \times (t-1) + 1}$$

$$t = 1, \dots, n$$

And interpolation consistency requires that

$$Y_t = y_{4t-3}, t = 1, \dots, n$$

with

$$c = (1, \dots, 0)'$$

And stock data observed as the last value of a given period

$$c = (0, \dots, 1)'$$

$$Y_t = y_{m \times t}$$

$$t = 1, \dots, n$$

For consistency

$$Y_t = y_{4t}$$

$$t = 1, \dots, n$$

with

$$c = (0, \dots, 1)'$$

2.3.2 Theoretical Literature on Temporal Disaggregation

Many researchers on this subject have noted that theoretical literature has considerably evolved in several directions, from representing the temporal disaggregation problem as simply a mathematical problem predominantly involving preserving variations in the original low frequency data to a model-based problem requiring the formulation of a relationship between the target variable and the predictors. For example, Denton (1971) considers the disaggregation problem via constrained optimisation; Chow and Lin (1971), Fernandez (1981), Rossi et al. (1982), Litterman (1983), Nijman and Palm (1986), Trabelsi and Hillmer (1990), Guerrero and Martínez (1995), and Cabred and Pavía (1999) by regression models using information in covariates together with the observed values of the target variable to determine the high frequency estimates; the state space methods which are estimated by Augmented Kalman Filter. In this section, the study undertakes to review temporal disaggregation methods but with considerable attention and detail paid to the estimation theory that governs traditional methods that this study benchmarks against to justify a data-driven approach as its key point of departure.

From the brief background above, I observe that the technical structure of the temporal disaggregation procedures is generally classified into two main categories. The first and relatively older category is referred to as the optimisation approach which aims to either minimise the within-variable differences from period-to-period changes or indeed minimise the differences between the target variable and the related variable.

The second category refers to the models that assume a linear relationship between an observed low frequency time series and an unobserved high frequency time series also known as regression model-based methods. According to sources such as (Bloem et al. (2001), Eurostat and Buchow (2013), Quilis (2018), and Buono et al. (2018)), model-based procedures play a vital role in compiling quarterly national accounts across many countries.

Many have argued that the reason for their prominence is that, model-based procedures have an inherent ability to provide an objective way of combining the relevance of short-term related indicators and internal consistency of the national accounts. Consequently, the ability to incorporate additional independent information in the data conversion process has made model-based methods the most preferred and popular class of methods for temporal disaggregation of time series globally.

2.3.2.1 Optimisation Methods of Temporal Disaggregation

One can reasonably argue that the early optimisation methods form the very foundation of what has now become the popular temporal disaggregation literature. These early methods simply rely on mathematical principles to derive a smooth path for the unobserved high frequency time series by preserving the movements of the original data series. As alluded to above, within this class, there are two subgroups, the procedures that use a related indicator variable by minimising the difference in the change between the target variable and the related indicator, for example Denton (1971) and another class that is completely a univariate method in nature, minimising the difference in the change between the current period and previous period values subject to the additive constraint holding. Some of the popular methods in this category include the Boot-Feibes-Lisman (BFL) method based on Lisman and Sandee (1964) and Boot et al. (1967), time-domain benchmarking by Denton (1971), Cohen et al. (1971), Al-Osh (1989), Stram-Wei method based on Wei and Stram (1990), and Gómez and Maravall (1994).

2.3.2.2 Boot, Feibes and Lisman Method (BFL)

The BFL method is one of the very early mathematical methods to be applied to temporally disaggregate low frequency series into high frequency time series. This procedure uses a smoothing algorithm to estimate a high frequency time series using first or second differences in an original time series by minimising changes in the level or the change of the resulting high frequency estimates from period to period subject to the additive constraint holding. Formally, the objective function of the BFL method in the first differences sum of squares is specified as follows:

Assume $d = 1$, first difference

$$\min_y p(y) = \sum_{t=2}^T (y_t - y_{t-1})^2 \quad (2.1)$$

Subject to

$$Y_t = \sum_{a=1}^4 y_{t,a} \quad (2.2)$$

Where Y_t is the low frequency time series.

The second differences sum of squares model which is specified by minimising the sum of squares of first differences subject to additive constraint is expressed as follows:

Assume $d = 2$, second difference

$$\min_y p(y) = \sum_{t=2}^T (\Delta y_t - \Delta y_{t-1})^2 \quad (2.3)$$

Subject to the additive constraint, 2.2 above.

The above optimisation problem implies that volatility in the estimated high frequency series is minimised from period to period while ensuring that there is temporal consistency between the low frequency benchmark with its target high frequency series.

In matrix form, the BFL method is formulated as follows:

$$\min_y \phi = u' u = y' D' D y$$

subject to

$$Y = Cy \quad (2.4)$$

Where D : $n \times n$ is the difference operator $\nabla = (1 - B)$ in matrix form:

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 \\ 0 & 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

The solution to the optimisation problem in equation 2.1 yields an estimator of the high frequency, y as follows:

$$\hat{y} = (D'D)^{-1}C'[D'D)^{-1}C']^{-1}Y \quad (2.5)$$

And the estimated high frequency series depends on the level of differencing ($d = 1$ or $d = 2$) and the type of temporal disaggregation, C , depending on whether the series to be disaggregated is to be consistent by sum, average, first or last value.

Denton (1971), proposed an optimisation procedure that incorporates information from both the low frequency series and a related series in a way that is temporally consistent while conditioned on the related series variations. This is the key point of departure between the Denton and BFL method. This procedure is formalised in matrix form as follows:

Assuming an expanded information set to include a related series, x , thus $\Omega = Y, x$ and $d = 1$;

$$\min_y \phi = u'u = (y - x)'D'D(y - x) \quad (2.6)$$

subject to equation 2.4

The solution to the optimisation problem above is

$$\hat{y} = x + (D'D)^{-1}C'[C(D'D)^{-1}C']^{-1}U \quad (2.7)$$

It becomes clear in Denton's solution that the estimated target variable, $\hat{y}$ is a sum of a component linked to the high frequency indicator x and the temporal distribution of a discrepancy

$$(D'D)^{-1}C'[C(D'D)^{-1}C']^{-1}U$$

.

Another variant of the Denton method assumes a relative form of the high frequency residual so that;

$$u_t = \frac{y_t - x_t}{x_t}$$

thus, introducing a diagonal weighting matrix

$$W = \text{diag}(x_1, x_2, \dots, x_n)$$

in the objective function of the optimisation problem:

$$\min_y \phi = u'u = (y - x)'W^{-1}D'DW^{-1}(y - x) \quad (2.8)$$

subject to equation 2.4

The solution to the optimisation problem above is

$$\hat{y} = x + W(D'D)^{-1}WC'[CW(D'D)^{-1}WC']^{-1}U \quad (2.9)$$

The assumption above on the relative form of the high frequency residual leads to a Denton solution that estimates the target variable, $\hat{y}$ as a sum of a component linked to the indicator x and the temporal distribution of a discrepancy $W(D'D)^{-1}WC'[CW(D'D)^{-1}WC']^{-1}U$ that has indicator information through the diagonal weighting matrix, W .

In summary, optimisation temporal disaggregation procedures have been criticised for their failure to fully account for dynamic economic relationships among variables, as such a major shortcoming. Although the current study compares the regression model-based methods to new proposals, here I review the BFL and Denton methods for purposes of offering some initial context, direction and clarity on the way literature regarding estimation in temporal disaggregation has evolved.

2.3.2.3 Regression Model-Based Temporal Disaggregation

The purely mathematical procedures of temporal disaggregation paved the way for the model-based methods which are presented next. At this stage, it is important to begin by describing the universal framework that unifies all model-based temporal disaggregation procedures. This section describes the model-based framework borrowing from the presentation by Sax and Steiner (2013)'s popular paper on temporal disaggregation. Unlike the optimisation procedures reviewed above, regression model-based methods as the class name suggests are set up in a static or dynamic form explicitly assuming a linear relationship between the target variable and the high frequency predictor.

To perform temporal disaggregation of time series, Sax and Steiner (ibid.) showed that the range of the regression model-based temporal disaggregation approaches can be constricted by putting them in a two-step framework: First, a preliminary high frequency series must be estimated; second, the differences between the low frequency values of the preliminary series and the low frequency values of the observed series must be distributed among the preliminary high frequency series. The sum of the preliminary high frequency series, p and the distributed low frequency residuals, Du_l yield the final estimation of the target high frequency series, $\hat{y}$. Formally expressed as,

$$\hat{y} = p + Du_l \tag{2.10}$$

where

D is a $n \times n_l$ distribution matrix, with n and n_l denoting for example, the number of high frequency (quarterly) and low frequency (annual) observations, respectively. This is the general expression and it is easily representative of other conversions such as converting quarterly to monthly series. u_l is a vector of length n_l and contains the differences between the low frequency values of p and the actual low frequency values, y_l :

$$u_l \equiv y_l - Cp \quad (2.11)$$

Multiplying the $n_l \times n$ conversion matrix, C , with a high frequency series performs aggregation. Equation 2.10 constitutes a unifying framework for all disaggregation methods. The model-based methods, therefore, differ in how they determine the preliminary series, p and the assumptions made regarding the behaviour of the error term as shall be shown in the following section with specific procedures.

The formulation of the model-based framework is very important for this study since I develop a data-driven procedure based on it by proposing an alternative way of estimating the preliminary high frequency series or predicted values, p , based on a high dimensional time series estimation method using PLS regression. Next, is to distribute the low frequency residuals or discrepancy, Du_l , equally across conversion frequency, m , to yield the final estimation of the target high frequency series, $\hat{y}$. In what follows, I present and briefly discuss each considered method's unique features.

2.3.2.4 Chow-Lin Method

The Chow-Lin method based on Chow and Lin (1971) is vastly considered among the most popular temporal disaggregation procedures because of its simplicity, transparency and other statistical properties from the point of view of practitioners. The main feature of the Chow-Lin procedure like the Fernandez and Litterman methods among other regression model-based methods is that it assumes a static relationship between the target variable and the indicator variables.

The Chow-Lin method starts with the linearity assumption which leads to the following regression expression:

$$y = x\beta + u \quad (2.12)$$

Where, y is the unobservable target variable, x is the observable high frequency indicators. In the specific case of Chow-Lin, u is assumed to follow a stationary AR(1) process, thereby in essence imposing a restriction that y and x are cointegrated. This assumption also forms the main point of departure between this procedure and other traditional methods reviewed in this study, specifically, Fernandez and Litterman procedures. Furthermore, the Chow-Lin procedure assumes that;

$$u \sim N(0, v)$$

$$u = \sigma^2 [(I_n + \rho \Xi)' (I_n + \rho \Xi)]^{-1}, -1 < \rho \leq 1$$

and the auxiliary matrix Ξ is defined as:

$$\Xi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -1 & 0 \end{pmatrix}$$

ρ determines the initial conditions such that

$$u_0 = N(0, \frac{\sigma^2}{1-\rho^2}), -1 < \rho \leq 1$$

$$u_0 = 0, \rho = 1$$

Although in principle ρ can take values that are between -1 and 1 , i.e., $-1 < \rho \leq 1$. Empirical applications require that ρ is restricted between 0 and 1 . An additional temporal constraint applies with an aggregation defined as;

$$Y = Cy, C = I_N \otimes c | O_{N,n-sN}$$

The temporal constraint ensures consistency!

The Chow-Lin method solves for the best linear unbiased estimator, $\hat{\beta}$ by minimising the mean squared errors subject to satisfying the aggregation constraint thus estimating the target variable according to the linear model specified in equation 2.12 above. Therefore,

$$\hat{y} = x\hat{\beta} + vC'V^{-1}\hat{U} = x\hat{\beta} + L\hat{U} \quad (2.13)$$

where

$$V = CvC'$$

and the

$$\hat{U} = Y - X\hat{\beta}$$

and $X = Cx$. V is covariance matrix and $\hat{\beta}$ is a generalised least squares estimator with the dependent variable Y and aggregated related indicators as independent variables. It is estimated as follows;

$$\hat{\beta} = (X'V^{-1}X)^{-1}(X'V^{-1}Y) \quad (2.14)$$

2.3.2.5 Fernandez Method

In the literature, this procedure is sometimes referred to as the random walk model and was proposed in the study by Fernandez (1981) to correct serial correlation in high frequency estimates. Like the Chow-Lin procedure, this method derives the $\hat{\beta}$ and the linear unbiased estimator for the target variable, y . The key differentiating feature of this disaggregation method is that it assumes that the disturbance term, u_t follows a random walk (as opposed to the AR(1) process assumed in the Chow-Lin procedure). That is,

$$u_t = u_{t-1} + \varepsilon_t$$

where

$$\varepsilon_t \sim N(0, V_{T \times T}),$$

is a vector of random variables with mean zero and covariance matrix, $V_{T \times T}$.

Before estimating the high frequency target variable, the Fernandez procedure requires careful consideration of the behaviour of the series. For the series that are non-stationary and serially correlated, the first difference data should be used to obtain stationary and uncorrelated series. If the first difference is not enough, other transformations may be needed to convert residuals to serially uncorrelated and stationary variables and lastly that generalised least square estimation can be used to test the degree of serial correlation if the series are properly transformed.

2.3.2.6 Litterman Methods

The Litterman method just like the Fernandez method assumes that the residuals follow a non-stationary process:

$$u_t = u_{t-1} + \varepsilon_t$$

where ε_t is an $AR(1)$, thus,

$$\boldsymbol{\varepsilon}_t = \alpha \boldsymbol{\varepsilon}_{t-1} + e_t$$

and e is $WN(0, \boldsymbol{\sigma}_e)$. It can, therefore, be seen from above that the Fernandez method is a special case of Litterman, when $\alpha = 0$.

2.3.2.7 Silva Santos – Cardoso Method (SSC)

Finally, in this section, I review the SSC method. This temporal disaggregation method is also an extension of the classical Chow-Lin method seen above but with a dynamic structure in the relationship between the unobservable target variable y and the high frequency indicators, x . This study follows the formal presentation of the SSC procedure according to Quilis (2018):

$$y_t = \varphi y_{t-1} + x_t \beta + \boldsymbol{\varepsilon}_t, -1 < \varphi < 1, t = 1, 2, \dots, n \quad (2.15)$$

Where x_t is the t^{th} row of x and $\boldsymbol{\varepsilon}_t \sim iidN(0, \boldsymbol{\sigma}_\varepsilon)$. Empirical work assumes φ to take values between 0 and 1.

The model above can be rewritten as a distributed lag as follows:

$$y_t = \sum_{j=1}^p \sum_{h=0}^{\infty} \theta_{j,h} x_{t-h} + \sum_{h=0}^{\infty} \Psi_h \boldsymbol{\varepsilon}_{t-h} \quad (2.16)$$

Where

$$\theta_{j,h} = \beta_j \Phi^h$$

and

$$\Psi_h = \Phi^h$$

are Constraints.

The constraints above ensure that the dynamic model 2.15 is a parsimonious parametrisation of a general dynamic model. The model can be expressed in matrix form as follows:

$$D_{\phi}y = x\beta + q\eta + \varepsilon \quad (2.17)$$

where

$$D_{\phi} : n \times n = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ -\phi & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & -\phi & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -\phi & 1 \end{pmatrix}$$

q and η are referred to as “truncation remainders” holding the initial conditions of the model as follows:

$$q = \begin{pmatrix} \Phi \\ \vdots \\ 0 \end{pmatrix}$$

$$\eta = E(y_0)$$

The matrix representation of the model can then be rearranged to;

$$D_{\Phi}y = z\gamma + \varepsilon$$

where

$$z = [xq]y, \gamma = [\beta\eta]'$$

This means that initial conditions of the model are regarded as additional parameters that must be estimated like the rest of the model parameters. The model solves for y to get:

$$y = D_{\varphi}^{-1}z\gamma + D_{\varphi}^{-1}\varepsilon = z(\varphi)\gamma + u \quad (2.18)$$

Similar to the other procedures reviewed so far, the low frequency model for the equation 2.18 above is achieved by pre-multiplying it with the aggregation matrix, C and it applies only to observable variables of Y and Z : therefore,

$$Y = Z(\varphi)\gamma + U$$

By assuming that the shocks ε_t are Gaussian, it has been shown in the literature that the corresponding log-likelihood function of the low frequency model above can be expressed as:

$$l(\beta, \sigma^2, \varphi) = -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\ln(|Cv(\varphi)|C') - \frac{1}{2\sigma^2}(Y - X\beta)'(Cv(\varphi)C')^{-1}(Y - X\beta)$$

Where the variance-covariance matrix is

$$V(\varphi) = Cv(\varphi)C'$$

and

$$v(\varphi) = \sigma_a(D_{\varphi}D'_{\varphi})^{-1}$$

Since the log-likelihood function is non-linear in Φ , its maximisation is performed using grid search and estimation of all parameters in γ is by means of generalised least squares: therefore,

$$\hat{\gamma} = (Z(\Phi)'(Cv(\Phi)C')^{-1}Z(\Phi))^{-1}(Z(\Phi)'(Cv(\Phi)C')^{-1}Y)$$

And the same best linear unbiased estimator used by the Chow-Lin method is used to find the high frequency time series, y given by:

$$\hat{y} = z(\Phi)\hat{\gamma} + v(\Phi)C'V(\Phi)^{-1}\hat{U} = z(\Phi)\hat{\gamma} + L_{\Phi}\hat{U}$$

2.3.3 Empirical Literature on Temporal Disaggregation

Since the early work by, for example, Lisman and Sandee (1964), and Boot et al. (1967), research on temporal disaggregation of time series has grown considerably, pointing to how long this particular research area has enjoyed attention among researchers, analysts and policymakers alike. However, in terms of specific interests, most of the research on methods of temporal disaggregation that use related indicators has since the seminar paper by Chow and Lin (1971), developed procedures that use a limited amount of information owing in part to the computational constraints and the potential problem of multicollinearity among predictor variables as discussed earlier. Nowadays, however, researchers can handle big data as knowledge, econometric analytical tools and computing power have become available to move in this direction.

To help emphasise and shed more light on the characteristics of the majority of existing methods, below I present some of the popular temporal disaggregation procedures in the empirical literature. I categorise the methods according to their respective estimation strategies, for example, optimisation approaches that do not use related indicators by Lisman and Sandee (1964), Boot et al. (1967), Wei and Stram (1990), optimisation approach that uses related indicators, Denton (1971). Regression model-based approaches that use related indicators according to Chow and Lin (1971), Denton (1971), Fernandez

(1981), Litterman (1983), Cholette and Dagum (1994), Silva and Cardoso (2001), Di Fonzo (2003) and state-space approaches for example, Harvey and Pierse (1984), Durbin and Quenneville (1997), Harvey and Chung (2000), and Proietti (2006). The list of examples highlighted here is far from comprehensive ⁷.

Besides merely noting the various studies performing temporal disaggregation by different traditional procedures, this section also aims to highlight three key inconclusive issues in the empirical literature on temporal disaggregation namely, the problem of selecting related indicators, the problem of multicollinearity among high frequency predictors and lastly, explanatory versus predictive modeling for temporal disaggregation. The literature shows that, indeed over time, various procedures for performing temporal disaggregation of time series have been developed and widely applied to reasonably interpolate, distribute and extrapolate economic data.

On empirical versus theoretical considerations, while in principle, the regression model-based methods can use more than one indicator, most empirical studies have however, tended to use a single predictor but with a few other studies using more than one indicator. For example, among a few notable studies reviewed include, Hall and McDermott (2011), constructs a quarterly GDP series for New Zealand using a Haywood and Campbell diffusion index, Islam (2009) applies the Chow-Lin and Fernandez methods to disaggregate annual GDP to quarterly series using the Index of Industrial Production, Islam (2013) converts annual exports of Bangladesh to quarterly export data using the Chow-Lin, Fernandez and Litterman using Quantum index of Industrial production and compares the quality of the resulting series, Chen et al. (2007) for the US evaluates performance of five regression model-based methods using 45 target variables each using a single indicator. Others are Quilis (2018) for Spain who converts annual employment series using quarterly social security affiliates, etc.

7. see, for example, Chen et al. (2007), Reber and Pack (2014), and Quilis (2018) for extensive conversations on various methods of temporal disaggregation.

The literature shows that most studies that use multiple indicators often run into similar multicollinearity estimation challenges, some examples include; Eurostat (2014), applying the Chow-Lin method to disaggregate the Dutch quarterly GDP to monthly series using three monthly indicators, namely, index of manufacturing, inflation rate and the unemployment rate. Miralles et al. (2003), for Spain conducts an experiment by increasing the number of indicators using the Chow-Lin method to obtain the Spanish Quarterly National Accounts. Like others in later years, the results show that the quality of the estimates decreased when the number of indicators increased possibly due to multicollinearity.

In the literature, the choice of related indicators in many empirical studies has been governed by the observed correlation between the target variable and the indicators and what one would refer to as expert knowledge or common sense. However, correlations between the target and indicator variables may vary at different points in time going from high to low and vice versa, as dynamic relationships change over time. The changes in the strength of association between the target variable and the related indicator over time has implications for the accuracy of the estimated high frequency target variable. Chamberlin (2010) for example, validated this argument and showed that the significance of an indicator largely depends on the relative strength of correlation at the lower frequency level. This underscores the difficulty of finding indicators that demonstrate a stable linear interrelationship over time.

The implications of the “related variable selection problem” are well characterised and summarised by Silva and Cardoso (2001) who note that the ability of the researcher to find a temporal disaggregation model with a good fit depends on the availability of good indicator series and on the characteristics of each problem. In their study, they support their claim by showing that most temporal disaggregation methods tend to perform alike when there is a close relationship between the target variable and the indicator series. This evidence underscores the importance of finding a good and stable related indicator when using the traditional temporal disaggregation methods. Chen et al. (2007) and Quilis (2018) reiterate this view by reflecting that an appropriate combination of benchmark and trackers is critical for a successful application of traditional methods.

Furthermore, even though the popular procedures, for example, Chow and Lin (1971), Fernandez (1981), and Litterman (1983), use unbiased parameter estimates to establish a best linear relationship between the target variable and the high frequency indicator, one can argue that they mischaracterise the temporal disaggregation problem at hand. The temporal disaggregation problem is typically an in-sample prediction problem and therefore, the predictive soundness of a temporal disaggregation model must be emphasised as much as parameter estimation soundness for an explanatory model.

Moreover, given the huge amount of data generated nowadays, it appears that the use of high dimensional time series and big data is an inevitable pathway. As such, temporal disaggregation procedures must primarily embellish much more information and predictive properties rather than causal inferencing characteristics. At this point, the simple regression model-based methods will fail. In critiquing the linear regression, Polson (2017) for example, points out that even in simulated data, we do better than simple linear regression by shrinking the coefficients to reduce the variance. The typical maximum likelihood estimator or the ordinary least squares is designed to have zero bias, which means that it can suffer from high variance. Accordingly, Polson (*ibid.*) notes that the main advantage of linear regression is interpretability (explanatory modeling), not predictability.

An inquiry into the literature on temporal disaggregation using high-dimensional time series data indicates that this strand of literature remains relatively under-researched. However, I note a recent study by Mosley et al. (2022) that proposed a temporal disaggregation approach based on LASSO regression. This work is the closest comparator to my current work. In their study, they applied their new LASSO-based method to UK quarterly gross domestic product data, demonstrating its superiority over traditional methods like the Chowlin technique. While their study focused on the UK dataset and comparisons with the Chowlin and a limited set of indicators, it provides valuable insights applicable to other countries.

In summary, the literature review traces the historical development of temporal disaggregation research, from seminal works to recent advancements. Recent studies, such as the LASSO regression-based approach proposed by Mosley et al. (*ibid.*), are discussed in comparison to traditional methods like the popular Chowlin method. The literature

also shows that there is a prevalent use of regression model-based methods and limited adoption of multiple indicators in empirical studies. It is easy to see that the majority of empirical research has, however, been conducted mindful but tolerant of challenges such as indicator selection and multicollinearity. Although widely used, I also note that various researchers have criticised the linear regression model-based temporal disaggregation framework, particularly because of its parameter estimation strategies which lean heavily on explanatory modeling. Therefore, the question of predictive soundness over explanatory modeling means that since temporal disaggregation is an in-sample prediction problem, there is a clear need to re-balance between explanatory and predictive modeling. Overall, the literature review sets the stage for understanding the evolution, challenges, and recent advancements in temporal disaggregation research, providing a foundation for the study's focus on data-driven approaches.

Based on the literature, I suggest that the short path forward for this study is to introduce a simple yet fundamentally different approach whose estimation approach is based on PLS regression. The robustness of the PLS regression-based procedure is also assessed across a diverse range of countries, including the USA, Japan, Canada, France, and Germany. These two dimensions followed by this study, namely, the temporal disaggregation estimation approach and data coverage, mean that this research expands on existing knowledge by introducing a novel procedure and testing the effectiveness of data-driven methods in multiple data sets, thus complementing the findings of Mosley et al. (2022) and highlighting the competitiveness of such approaches in handling large-scale data.

2.4 Towards Data-Driven Temporal Disaggregation

2.4.1 Explanatory versus Predictive Modeling

In this subsection, the study appeals to the literature to find support for the current proposal to perform temporal disaggregation of time series using data-driven procedures in data rich environments by distinguishing between explanatory modeling and predictive modeling (see, for example, Sainani (2014))⁸.

I provide an overview of the key attributes of specific modeling approaches in Table 2.1 , which helps to lay the foundation for advocating for the alternative temporal disaggregation methods. The proposed PLS temporal disaggregation approach falls well within the space of predictive models as compared to the traditional models like Chowlin, Fernandez, Litterman, and SSC methods that follow an explanatory approach.

Indeed the summary of the features outlined in the table below provide justification that leveraging predictive modeling for temporal disaggregation may represent a natural and logical progression. Therefore, my approach capitalises on the PLS ability to harness a broader range of information with favorable predictive qualities, thereby offering a promising avenue for estimating the high frequency preliminary values, p , as per equation 2.10.

8. High dimensional time series temporal disaggregation techniques can be difficult to apply in settings with limited data because they depend on a wide range of auxiliary indicators to produce reliable and accurate estimates. In order to capture underlying dynamics and lower estimation uncertainty, these techniques usually require rich, high frequency data. The effectiveness of these processes may drastically decline in situations where such data are scarce or untrustworthy. The Appendix C offers a thorough explanation of these restrictions as well as the composition and function of a suggested data warehouse created to overcome this difficulty. And examples of types of economic indicators that may be useful in disaggregating broad indicators are reported in the Appendix E.

Table 2.1: Explanatory vs. Predictive Modeling

	Explanatory Modeling	Predictive Modeling
Main goal	Theory-driven; testing causal hypotheses and relationships between theoretical constructs.	Data-driven; predicting unobserved values through associations among measured variables.
Focus	Minimising model bias to ensure theoretical accuracy.	Balancing bias and variance to achieve optimal predictive performance and generalisability.
Role of Theory	Central—it guides model specification and causal interpretation.	Secondary—models are judged on out-of-sample predictive accuracy.
Role of Statistical Significance	Paramount—parameter estimates must be statistically significant.	None—focus is entirely on prediction error metrics.
Model Evaluation	Fit statistics (for example, R^2 , in-sample RMSE).	Prediction accuracy (for example, out-of-sample RMSE, MAE, MAPE).
Samples Required	One (the estimation sample).	Two (distinct training and testing samples).
Typical Methods	OLS and ML (<i>... as the case may be for traditional temporal disaggregation methods</i>)	Regularised and dimension-reduction techniques (for example, LASSO, PLS).

Directly following the summary of properties of explanatory and predictive models in Table 2.1 above, I discuss the LASSO and PLS regression models. The reason for discussing the LASSO and PLS regression models is that I use these approaches to perform temporal disaggregation of time series and compare the results with those of traditional methods.

2.4.2 The LASSO Regression Model

The LASSO regression short for Least Absolute Shrinkage and Selection Operator was advanced by Tibshirani (1996). The LASSO regression prevents overfitting and ensures the generalisation of a data-driven model. It does so by learning the linear relationship between the features and targets. It implements what is referred to as an $L1$ regularisation term that severely penalises non-important or correlated features by forcing their corresponding coefficients to zero. This means that correlated features in the data are not included during the model development. The LASSO regression seeks to minimise the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant, and it is a well-known sparse regression method which regularises the parameter β under sparse assumption.

The LASSO estimation framework is formalised as follows: Consider a sample consisting of N cases, each of which consists of p covariates and a single outcome. Supposing y_i is the response variable and $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ is the covariate vector for the i^{th} case, $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$, so the objective of the LASSO is to solve the optimisation problem:

$$\operatorname{argmin}_{\beta_o, \beta} = \frac{1}{N} \sum_{i=1}^N (y_i - \beta_o - x_i^T \beta)^2$$

Subject to

$$\sum_{j=1}^p |\beta_j| \leq t$$

where

$t \geq$ is a pre-specified free parameter that determines the amount of regularisation. If t is large, all the coefficients are almost zero. For smaller values of t , the LASSO shrinks some of the estimated coefficients equal to zero.

The LASSO estimator $\hat{\beta}$ (Belloni et al. (2011); Efron et al. (2004); Tibshirani (1996); Zou (2006)) can be expressed as;

$$L(\beta, \alpha) = \min_{\beta} \left(\frac{1}{N} \|y - X\beta\|_2^2 + \alpha \|\beta\|_1 \right)$$

where the tuning parameter $\alpha \geq 0$ is used to balance the empirical error and the sparsity of the model parameter, and where the exact relationship between t and α is data dependent. In my study, we select the final series associated with the value of α that produces the smallest Mean Square Error (MSE) for all experiments.

The main features of the LASSO temporal disaggregation framework are the following; tuning parameter that is associated with the lowest mean square error, recovery of estimated low frequency parameters and selected features and prediction of the final high frequency target using high frequency selected features.

2.4.3 The Partial Least Squares Regression

In this section, I present the general Partial Least Squares (PLS) regression framework and the PLS temporal disaggregation algorithm. Most importantly, the study introduces the PLS temporal disaggregation algorithm which I use to predict the initial high frequency predicted component. The PLS procedure is structured much in the same way as the general temporal disaggregation framework discussed earlier in equation 2.10.

So, to begin with, PLS regression is typically viewed as an extension of the multiple linear regression. It tends to work well in situations where the features are highly colinear and when the number of features is greater than the size of the dataset⁹. It is widely considered a relatively new method for estimating multiple regressions when there is a large, but finite, amount of predictors¹⁰. PLS like the principal component analysis (PC) uses regressors, factors or components which are linear combinations of the original regression variables instead of the original variable. The PLS factors are linear combinations of x_t that give maximum covariance between y_t and factors of x_t while being orthogonal to each other.

However, although PLS is similar to PC, the main difference between the two is that, whereas in PC regressions factors are constructed taking into account only the values of the x_t predictor variables, in PLS, the relationship between y_t and x_t is considered as well in constructing the factors (Groen and Kapetanios (2016)). So, by construction, a PLS regression learns the multidimensional direction in the feature space that explains the maximum multidimensional variance direction in the target space whereby the latent structure corresponding to the most variation in target is first extracted and then explained using a latent structure in the feature space.

PLS regression is also considered to be a better alternative to multiple linear regression and PCA regression methods since it provides for more robust model parameters that do not change with new calibration samples from the population (Geladi and Kowalski (1986); Falk and Miller (1992) and Pirouz (2006)).

9. OLS models are known to overfit linear relationships in such situations whereas PLS model behaves well in terms of developing a generalisable model.

10. see, for example, Wold (1982) and Wold et al. (1987), PLS was developed to address weak theory and weak data.

The number of components cannot be more than the number of features and the final PLS model is constructed with the number of components fewer than the number of features. Unimportant features that are colinear or not strongly correlated to the target variance help to develop an accurate model. In applying the PLS model, the most important hyperparameter is the number of components to be generated - to avoid overfitting. Data pre-processing by scaling the features is essential for developing a robust PLS model.

In summary, PLS regression provides an effective approach for highly collinear predictor sets by projecting the original variables onto a limited number of orthogonal latent factors specifically optimised to maximise covariance with the response variable. This specific dimension reduction not only stabilises coefficient estimates amid multicollinearity but also frequently produces enhanced out-of-sample prediction accuracy relative to unsupervised methods like the principal components. PLS also maintains computing efficiency even when the quantity of predictors surpasses that of observations.

The following is a simple algorithm to construct k PLS factors¹¹:

We assume for simplicity that y_t is demeaned and x_t is normalised to have zero mean and unit variance.

A general PLS algorithm

1. Set $u_t = y_t$, and $v_{i,t} = x_{i,t}$, $i = 1, \dots, N$. Set $j = 1$.
2. Determine the $N \times 1$ vector of indicator variable weights or loadings $w_j = (w_{1j}, \dots, w_{Nj})'$ by computing individual covariances: $w_{ij} = \text{cov}(u_t, v_{it})$, $i = 1, \dots, N$. Construct the j^{th} PLS factor by taking the linear combination given by $w_j' v_t$ and denote this factor by $f_{j,t}$.
3. Regress u_t and $v_{i,t}$, $i = 1, \dots, N$ on $f_{j,t}$. Denote the residuals of these regressions by $\tilde{u}_t$ and $\tilde{v}_{i,t}$.
4. If $j = k$ stop, else set $u_t = \tilde{u}_t$, $v_{i,t} = \tilde{v}_{i,t}$, $i = 1, \dots, N$ and $j = j + 1$ and go to step 2.

11. More detail on constructing PLS factors can be found in Helland (1990) and the current representation is in the lines of Groen and Kapetanios (2016).

Once PLS factors are constructed, y_t can be modeled by regressing y_t on $f_{j,t}$, $j = 1, \dots, k$. A linear regression model using the PLS components as predictors is fitted using least squares. The optimal number of PLS components to keep in the final model is determined by k -fold cross-validation which depends on the size of the sample. This is the approach followed in my determination of the systematic component using the PLS-adapted temporal disaggregation procedure.

Proposed PLS based temporal disaggregation algorithm

A PLS based temporal disaggregation method fits within the standard disaggregation framework (first predicting a high-frequency series, then distributing residuals to match the low-frequency aggregates). It respects the accounting consistency identities (quarterly sums matching annual data) and, therefore, is structured as follows:

1. Step 1: First, we set low frequency data such that $U_t = Y_t$, and $V_{i,t} = X_{i,t}$, $i = 1, \dots, N$. Set $j = 1$.
2. Step 2: For low frequency data, we determine the $N \times 1$ vector of indicator variable weights or loadings $W_j = (W_{1j}, \dots, W_{Nj})'$ by computing individual covariances: $W_{ij} = \text{cov}(U_t, V_{it})$, $i = 1, \dots, N$. Construct the j^{th} low frequency PLS factor by taking the linear combination given by $W_j' V_t$ and denote this factor by $F_{j,t}$.
3. Step 3: For low frequency data, we regress U_t and $V_{i,t}$, $i = 1, \dots, N$ on $f_{j,t}$ and denote the residuals of these regressions by $\tilde{U}_t$ and $\tilde{V}_{i,t}$.
4. Step 4: For low frequency data, if $j = k$ stop, else set $U_t = \tilde{U}_t$, $V_{i,t} = \tilde{V}_{i,t}$, $i = 1, \dots, N$ and $j = j + 1$ and go to step 2.
5. Step 5: We use low frequency PLS factors to estimate the low frequency, y_t s and discrepancy - the difference between actual and predicted values.
6. Step 6: Recover the low frequency weights $W_{ij} = \text{cov}(U_t, V_{it})$, $i = 1, \dots, N$ and construct the j^{th} high frequency PLS factor by taking the linear combination given by $W_j' v_t$ and denote this factor by $f_{j,t}$, where v_t are high frequency. This produces a factor matrix for the high frequency data.

It is important to emphasise that in this approach, the linear combination $W_j'v_t$ implies that I apply the same loading values to the high frequency data to generate the high frequency factors. However, each weight in the weighting matrix, W_j is expanded from the low frequency data using the conversion frequency, m to accommodate the high frequency structure. For example, the annual weights are replicated across the four quarters of each year. I do this in order to ensure that the high frequency data is appropriately accounted for in the estimation of the high frequency factors. Therefore, the expansion aims to maintain the consistency of the relationship between the low frequency and high frequency variables and allows me to estimate factors at the higher temporal level.

Finally, once high frequency time series PLS factors are constructed, the high frequency predicted component of y_t can be modeled by regressing y_t on $f_{j,t}$, $j = 1, \dots, k$. The optimal number of PLS components to keep in the final high frequency model is determined by k -fold cross-validation. The steps that follow are similar to the ones in the temporal disaggregation framework according to Sax and Steiner (2013).

However, there are some situations in which the PLS method may not work well. Typically, the procedure may not perform well when there is little real signal about the target in the selected predictors because its latent factors will capture noise instead of significant patterns. This may result in inaccurate temporally disaggregated time series when using this approach. In addition, because standard PLS only creates linear combinations of inputs, it has trouble with strong non-linearities in covariates. Overfitting can occur when there are too many components kept in relation to the sample size, particularly in high-noise settings, where random fluctuations predominate. Furthermore, there might not be enough data to estimate stable factors when there are very few observations, which could lead to unpredictable loadings and inaccurate predictions.

2.4.4 Cross Validation

Cross-validation plays an important role in validating the final model as the case is in both the LASSO and PLS temporal disaggregation approaches and therefore, warrants some remarks about this method. In principle, cross-validation determines model accuracy by splitting data into two different samples, a training set and a testing set. In k -fold cross-validation, original data is split randomly into k sub-samples which become the training data. Models are estimated using $k - 1$ sub-samples for each fold, with the k^{th} sub-sample serving as the validation sample.

This process repeats until every subsample has served as the validation data and the results aggregated by averaging to produce a single model estimation. In the end, all objects are used for both training and testing, and each object is used for testing only once per fold. This exercise is an important step in the estimation of LASSO and PLS parameter estimates, and therefore, it is closely followed in the implementation of the considered temporal disaggregation procedures as well.

2.4.5 Forecast Combination

In some empirical macroeconomic forecasting studies, there is ample evidence that supports the use of forecast combinations for producing superior forecasts than individual forecasting models. Winkler and Makridakis (1983), Clemen and Winkler (1986), Diebold and Pauly (1990), Stock and Watson (2006), and Huang and Lee (2010) among others, have shown that in many instances even a simple average is a robust enough combination method that outperforms even more sophisticated methods based on model weighting.

In the literature, forecast combination has been used to improve predictions especially where more than one forecasting model has been involved. So, in an additional effort to evaluate the robustness of the models beyond individual methods, this study uses the panel of estimated temporally disaggregated historical time series using different methods and combines the estimated time series. The strategies followed include equal weight (mean) and the Principal Component Analysis (PCA) for estimates of traditional methods. So, in the first stage, the accuracy performance of the individual traditional model estimates

together with the LASSO and PLS estimates are compared to the actual time series. Then combination series based on the simple average and PCA, estimates of the LASSO and PLS frameworks are all compared with the actual observed time series of the target variable.

2.5 Empirical Results

2.5.1 Data Description

Data considerations: To develop and empirically demonstrate temporal disaggregation prediction performance of the PLS procedure and conduct a horse race with five other selected methods, the first consideration is the integrity of the validation process of estimates. This partly hinges on having and using sufficiently long and complete data of both the target and related variables. This is particularly important for the LASSO and PLS procedures which require a long enough dataset to conduct cross-validation exercises.

Data choice: The choice of predictor indicators for the experiments conducted has been carefully selected according to the availability of relevant time series at both low and high frequency levels. The time series spans a reasonably long enough period to enable meaningful analysis and draw sensible conclusions. In this regard, this study uses 67 macroeconomic covariates spanning from 1960 to 2020 for the USA in the first experiment that specifically targets temporal disaggregation of narrow macroeconomic indicators.

To temporally disaggregate GDP, a broader indicator, the study draws on a subset of countries studied by Moauro and Savio (2005), namely, USA, Canada, Germany, France and Japan. The temporal disaggregation and nowcasting experiments are conducted with country-specific and varying numbers of at least 30 covariates in each case. Bai and Ng (2008) for example, argues that principal components that are estimated from at least 30 well-behaved time series data often have reasonably good properties. The time series datasets are downloaded from online resources, the FRED and FRED-MD¹².

12. <https://research.stlouisfed.org/econ/mccracken/sel/> accessed on 14/04/2023

Data pre-processing: In terms of further data considerations, this study uses seasonally adjusted quarterly and monthly time series. The reason for applying seasonally adjusted data series is first and foremost to work with times series that are absent of systematic calendar-related variations that are typically associated with the time of the year or seasonal effects. In addition, using seasonally adjusted time series ensures that the analytical results and conclusions obtained reflect the fundamental movements in a time series. Secondly, the use of seasonally adjusted time series also ensures consistency of the current work with that done by other researchers on this subject, making any comparisons relatively straightforward.

Data used to perform temporal disaggregation using the Chowlin, Fernandez, Litterman and SSC methods in this study is transformed into the first difference of logarithms. The first difference logarithmic data has a nice feature of interpretability as percent change which is very intuitive. To perform temporal disaggregation using the LASSO and PLS frameworks, the target variable is also treated the same way as above (to retain comparability with time series applied in traditional methods) while the predictor variables are scaled to mean 0 and variance 1. Standardising predictor values is a requirement for data-driven procedures to prevent any individual features with high variance in the predictor dataset from dominating the results.

This study uses the downloadable MATLAB library by Quilis (2018) for temporal disaggregation to perform temporal disaggregation of time series according to the Chowlin, Fernandez, Litterman and SSC methods¹³. For the data-driven high-dimension data procedures, new MATLAB algorithms that rely on the inbuilt LASSO and PLS functions are developed and formulated from scratch along the lines of the general temporal disaggregation framework, equation 2.10. The final high frequency estimate is a sum of the predicted high frequency series, p and the distributed low frequency residuals, thus converting the low frequency time series to a high frequency time series.

13. Can be downloaded from: <https://uk.mathworks.com/matlabcentral/fileexchange/69800-temporal-disaggregation/>

Table 2.2 and Appendix Table 2.8, present the list of targets and related indicators used in Experiment 1, that is, the quarterly target and the respective monthly indicator variables used in the traditional and data-driven procedures.

Table 2.2: Narrow Economic Indicators in Traditional Procedures

	Quarterly Time Series	Source	Monthly Indicator Series	Source
1	Industrial Production: Materials, Index	FRED	Industrial Production: Durable Goods Materials, Index	FRED-MD
2	Real Personal Consumption Expenditures	FRED	Real Disposable Personal Income	FRED-MD
3	Real Personal Consumption Expenditures: Nondurable Goods	FRED	Real Disposable Personal Income	FRED-MD
4	Real Personal Consumption Expenditures: Services	FRED	Real Disposable Personal Income	FRED-MD

2.5.2 Performance Evaluation Strategy

The disaggregation exercises are divided into Experiment 1 and Experiment 2, focusing on the performance evaluation of competitor procedures for narrow and broad target variables, respectively. This division aims to demonstrate how the competing procedures perform for different types of macroeconomic time series and serve as a robustness exercises. In Experiment 1, temporal disaggregation is applied to four selected narrow macroeconomic indicators using traditional methods alongside the LASSO and PLS procedures¹⁴. I evaluate each method's performance by tracking the movements and predictions of the original monthly data of narrow macroeconomic indicators. The relevance of this exercise is that analysts, economists and policymakers care about both the direction and level of macroeconomic variables.

In terms of setup, temporal disaggregation using the LASSO and PLS procedures incorporates a large matrix of macroeconomic predictors. In all cases, the procedures are executed to meet the temporal constraint, where $Y = C\hat{y}$ and $\hat{y} = p + Du_l$ ¹⁵. This experiment is carried out under the assumption of *pseudo* absence of monthly time series. So, I

14. This is a subset of the narrow indicators studied by Chen et al. (2007). I use the same target variables in all procedures. Traditional methods use the same predictors while the LASSO and PLS procedures also use the same dataset of related indicators. The selected covariates are downloaded from the FRED-MD, <https://research.stlouisfed.org/econ/mccracken/sel/>

15. As indicated in the motivation of the study, the key difference in the procedures lies in the estimation of the predicted component, p . Key features: Traditional methods rely on OLS estimation, the LASSO reduces data dimensionality through feature selection and finally, the PLS estimates components that explain maximum variations in the target and the covariates.

start by aggregating the observed target monthly time series into quarterly series based on their stock or flow nature. Subsequently, the aggregated low frequency time series undergo temporal disaggregation using all methods, and the resulting high frequency time series are compared against the actual monthly series.

To formally evaluate the performance of competing procedures, I use the MAE and the RMSE same as in the literature - the MAE and RMSE are discussed in detail in Section 2.5.2.1 below. I further assess the competing methods over time with a particular interest in observing how the different procedures perform during extreme economic periods. So, I calculate and plot the cumulative sum of squared prediction errors (CSPE) for each procedure over the full prediction period. The prediction errors, or residuals, are calculated as the difference between the actual value y_i and the predicted value $\hat{y}_i$ at both quarterly and monthly intervals. The CSPE metric emphasises larger errors more than smaller ones, making it especially useful for penalising large errors. Subsequently, the CSPE values are plotted in order to provide a visual representation of how each method's predictions cumulatively differ from the target across the entire sample period. The cumulative sum of squared prediction errors across all n observations is given by the following formula:

$$\text{CSPE} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

This cumulative sum can indicate any systematic bias in the model: a large positive or negative value may suggest consistent underprediction or overprediction, respectively. This value is commonly used in regression analysis to assess the goodness of fit. Lower sum of squared error values indicate that the model's predictions are closer to the actual observations.

Experiment 2 applies all procedures to a broader indicator: the GDP time series for the USA, Japan, Canada, France, and Germany¹⁶. The objective in this experiment shifts to evaluating the in-sample accuracy of the competing methods by predicting the original low frequency series by aggregating the predicted high frequency time series. The original low frequency series prediction is done in the absence of fulfilling the temporal constraint condition. This process highlights the significance of identifying a method that minimises the low frequency discrepancy at each prediction point. To achieve this, quarterly time series data is disaggregated and the resulting monthly series is then aggregated back to quarterly and then compared with the observed values.

This strategy mirrors the approach taken by Armesto et al. (2010) in their forecasting mixed-frequency data study. Another notable study is Clements and Galvão (2008) which used a vector autoregression (VAR) consisting of the monthly indicator variables to provide forecasts of missing monthly values, thereafter aggregated to provide estimates of the quarterly values of the indicators. Others that follow this approach include, for instance, Rünstler and Sédillot (2003) and Zheng and Rossiter (2006).

In summary, Experiment 2 focuses on the GDP of five different countries, converting the observed quarterly time series to monthly time series. I assess how far the resulting model estimates are from the original low frequency target variable. The model estimates are achieved by aggregating the estimated high frequency GDP time series, $\hat{Y} = Cp$ using the predefined aggregation matrix, C . The desire is to have the difference between the observed quarterly series and the estimated series that is as small as possible. The performance of individual approaches is compared using accuracy metrics which I discuss next.

16. The study uses GDP for procedure evaluations because it is a commonly used variable for temporal disaggregation due to its significance as an economic indicator, the availability of data at lower frequencies, and the importance of high frequency data for economic analysis and policy-making.

2.5.2.1 Prediction Accuracy Measures

In the literature, performance of temporal disaggregation methods is typically evaluated using the popular loss functions which include the mean absolute error (MAE) and root mean square error (RMSE)¹⁷. Although in practice there are several formulations of loss functions, the choice of these particular functions ensures the consistency of the current study with other studies on temporal disaggregation without detracting from the already existing literature. The basic principle about these loss functions, when applied to temporal disaggregation outcomes, is that the accuracy of the estimated high frequency time series is evaluated against the observed high frequency time series or indeed aggregated and then evaluated against the corresponding observed low frequency time series.

The MAE is defined as a measure of errors between observed values and predicted values by the model. Typically, the difference between the observed values and the values predicted by the model is known as the residuals. The MAE is used to assess the prediction accuracy and smaller values are preferred as they indicate a good fit. MAE is mathematically formulated as follows:

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

The RMSE is another important and commonly used metric for assessing the accuracy of a model. Like the MAE, it measures the difference between the observed values and those predicted by a model. The RMSE is in itself an aggregate of the residuals, and therefore a single measure of predictive power.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

Where, y_i is the observed value of the target variable, $\hat{y}$ is the predicted value of the model, and n is the sample size over which temporal disaggregation is performed.

17. Some researchers have studied the properties and evaluation criteria of different temporal disaggregation methods using the Mean Squared Error (MSE) and Theil's U statistic, for example, Zaier and Abed (2014) and Ajao et al. (2016). Most studies often compare the accuracy and performance of different methods using empirical data.

The resulting value gives us an idea of the average distance between the observed and predicted data. The RMSE essentially puts higher weights on large errors by squaring the error terms. In the literature, it is used mostly as an indicator of comparing the predictive power of different models, especially when there are significant errors.

The MAE on the other hand is somewhat different from the RMSE by construction in that it assigns the same weights to both large and small errors. Therefore, by not putting too much weight on outliers, the MAE only provides a generic measure of how well a model is performing. However, for both MAE and RMSE, smaller values are strictly preferred as they reflect the best-predicting method relative to other methods on historical data.

2.5.3 Temporal Disaggregation of Narrow Indicators

In this section, I present the main results of model performance that compare the estimation results of competing temporal disaggregation procedures applied to relatively narrow macroeconomic indicators. The resulting MAE and RMSE are summarised in Table 2.3. The prediction accuracy results suggest that the performance of the traditional procedures across all four target variables is mixed. As such, for these data, the results suggest that there is no single superior temporal disaggregation procedure for all the cases considered. This observation underscores the notion of complementarity among traditional methods when implemented in real-world scenarios. Moreover, this observation is broadly in line with previous observations by other researchers (see Chen et al. (2007) and Quilis (2018)).

After assessing the predictive performance of each traditional method, I combine the estimated time series of each method by computing the time series of simple averages and Principal Components (PCs). These combined estimates are first compared with the estimates of individual traditional procedures. In particular, the combined time series generated relatively smaller MAE and RMSE compared to the individual method results. The PC in particular generates superior temporally disaggregated time series in three out of the four cases examined in the experiment. Specifically, the PC disaggregated time series generate lower MAE and RMSE for Industrial Production: Materials; Real Personal Consumption Expenditures; and Real Personal Consumption Expenditures: Nondurable

Goods. This observation aligns with existing literature advocating for forecast model combinations. Some studies that generally provide evidence in support of forecast combinations include Diebold and Lopez (1996), Newbold and Harvey (2002), Timmermann (2006) and Clements and Galvão (2009).

However, a comparative analysis of estimates from the LASSO and PLS procedures with the estimates of traditional procedures and combined estimates changes the narrative quite markedly in favour of the data-driven procedures. Evidence from prediction accuracy metrics for the four target variables indicates that the PLS procedure produces better results than traditional and LASSO competitors by significantly reducing MAE and RMSE (see Table 2.3). The LASSO approach is second best in all the cases considered. In general, the study finds that the data-driven methods combined outperform the traditional methods in all the cases considered. Estimates from data-driven procedures even outperform the combined time series of all four traditional methods¹⁸.

Although the number of narrow macroeconomic indicators is small compared to Chen et al. (2007), for example, it is quite obvious that data-driven procedures are competitive against traditional methods in all cases, as seen from the relatively small MAE and RMSE calculated over the entire sample. The results provide evidence for the complementary use and relevance of traditional methods in the temporal disaggregation of the time series literature.

Of course, this result does not by any means suggest that the PLS and data-driven methods in general are strictly better than traditional methods in all instances. The results rather provide some preliminary and useful evidence that data-driven methods can be a sensible alternative to the existing traditional temporal disaggregation methods and have the potential to become more useful methods in performing temporal disaggregation especially when there is an ample number of predictor variables. Table 2.3 summarises the discussion above on the performance of some traditional methods and averages against the LASSO and PLS procedures.

18. A report of the optimal number of features and components for the LASSO and PLS estimations, respectively, used in the final model estimation to generate the disaggregated time series with the lowest mean square error is as follows: Industrial Production: Materials, 5 and 7; Real Personal Consumption Expenditure, 6 and 6; Real Personal Consumption Expenditures: Nondurable Goods, 4 and 18 and finally, Real Personal Consumption Expenditures: Services, 7 and 6.

Table 2.3: Temporal Disaggregation Results: Narrow Indicators

	Industrial Production: Materials		Real Personal Consumption Expenditures		Real Personal Consumption Expenditures: Nondurable Goods		Real Personal Consumption Expenditures: Services	
Method	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Chow-Lin	0.3561	0.5665	0.3865	0.8287	0.5700	0.9909	0.2409	0.6721
Fernandez	0.3622	0.5677	0.3851	0.7950	0.5744	0.9146	0.2416	0.6682
Litterman	0.3613	0.5669	0.4237	0.8703	0.5886	0.9347	0.2581	0.6967
SSC	0.4455	0.7085	0.3842	0.7923	0.5708	0.9346	0.2422	0.6693
Average	0.3575	0.5658	0.3761	0.8082	0.5482	0.9141	0.2398	0.6718
PC	0.3554	0.5613	0.3781	0.8106	0.5455	0.9094	0.2398	0.6719
LASSO	0.3750	0.5101	0.4280	0.7435	0.5694	0.9421	0.2709	0.6223
PLS	0.0766	0.1567	0.1406	0.2844	0.5243	0.8029	0.2906	0.5328

Notes: The entries of the table report the Mean Absolute Errors (MAE) and Root Mean Square Errors (RMSE) corresponding to each target and method. For each target variable, the MAE and RMSE are compared across the methods. The smallest value means that the respective model performs better than the competitor methods. For example, the bold RMSE indicate the lowest RMSE across all models for a given variable. Average and Principal Components (PC) are combinations of predictions based on the simple average and PC of traditional methods.

To appreciate how different models perform during extreme periods, I assess the performance of all procedures against each other by plotting the cumulative sum of prediction errors for the quarterly time series. To determine the quarterly CSPE, I do three things, first, estimate monthly time series of the target variable in the absence of fulfilling the temporal constraint. Second, I aggregate the estimated monthly series into quarterly series using a predefined aggregation matrix, C . Finally, the prediction errors of each model are cumulated over a period of time to demonstrate how well they perform in various economic circumstances. In particular, I calculate a quarterly time series of errors for each method, benchmarked against the actual observations. For each method, I therefore calculate an accumulated series of squared errors. This exercise yields 6 CSPE time series for each target variable. If the cumulative errors of a specific method are on balance smaller than those for other methods at every point, then that procedure is more accurate and as such, to be preferred.

With that said, although I am mindful that a specific procedure may work well on one dataset and not as well on another, evidence from the current study CSPE (see Figure 2.1) indicates that the PLS procedure accumulates relatively smaller errors in 3 out of 4 cases compared to all other methods for the considered target variables. Notably also, is that the PLS procedure does not show sharp increases in errors by the time we arrive into extreme sub-periods like the financial crisis (2008 - 2009) and the COVID-19 crisis (2020 – 2021) as the case may be with traditional methods¹⁹.

This feature of the PLS procedure suggests its superior ability to predict (non-smooth) and thus relatively accurate temporally disaggregated time series even during unstable periods. Further, it is also worth noting that even though the LASSO is outperformed by the PLS in the majority of the cases, its performance is just as good or even better than that of traditional methods. In general, this observation implies that the LASSO and PLS estimates minimise the errors by tracking the actual series relatively better capturing, matching and accurately predicting turning points of target variables.

2.5.4 Temporal Disaggregation of Broad Indicators

This section presents a discussion on Experiment 2, where I continue to focus on demonstrating the efficacy of data-driven temporal disaggregation procedures relative to traditional methods using a broad macroeconomic indicator, GDP. Temporal disaggregation of GDP is quite popular in the literature and in practice. That is because it is available mainly with delays and therefore, at lower frequencies, yet it is an important economic indicator²⁰. However, GDP is often required for regular macroeconomic analyses and policy-making - often every quarter. Examples of previous studies that focus on temporal disaggregation of GDP include among others, Moauro and Savio (2005), Islam (2009), Chamberlin (2010), Hall and McDermott (2011), Islam (2013), and Eurostat (2014).

19. Traditional methods are tending to show sharp increases in errors during unstable periods

20. For a detailed discussion on why GDP is usually reported with delays, please see Appendix F.

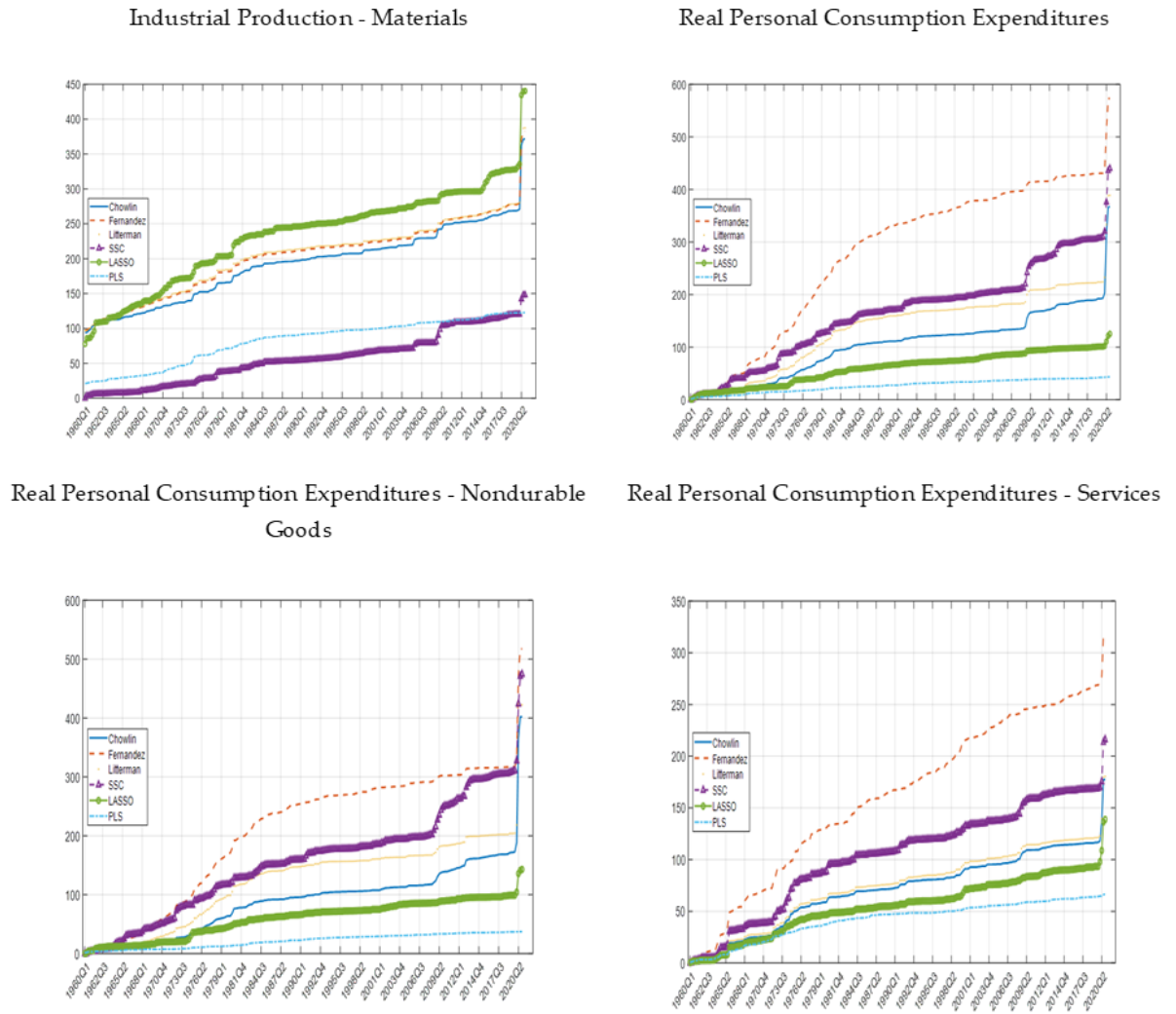


Figure 2.1: CSPE for Narrow Indicator Estimations

Notes: The CSPE for the Chow-Lin, Fernandez, Litterman, SSC, LASSO and the PLS methods are plotted for estimated quarterly Industrial Production: Materials, Real Personal Consumption Expenditures, Real Personal Consumption Expenditures: Nondurable Goods and Real Personal Consumption Expenditures: Services. The method with relatively smaller accumulated errors for the period has superior performance.

In this exercise, all methods are applied to disaggregate GDP growth rates of five selected major economies, including the USA, Japan, Canada, France, and Germany. Quarter-on-quarter changes in real GDP for each country are transformed into monthly series in the absence of fulfilling the temporal constraint. For each country, I use respective quarterly GDP and monthly country-specific related indicators in the traditional methods and the other two methods²¹. Because there are no observed monthly GDP data, I compare the temporal disaggregates at the quarterly level. Therefore, the estimated monthly growth rates are aggregated into quarterly growth estimates. Here, I desire a relatively small aggregate error to be distributed between the monthly preliminary estimates for temporal consistency and enhanced accuracy.

Consistent with the findings of Experiment 1, the PLS procedure, in particular, demonstrates consistently smaller cumulative errors in most of the country cases analysed (see Figure 2.2). Although the LASSO method shows relatively lower accumulated errors for the USA and Canada, the PLS performs better than all other approaches in the remaining countries, as shown by the relatively smaller accumulated model errors. In general, data-driven procedures collectively demonstrate superior performance compared to traditional methods across the countries examined. Moreover, similar to Experiment 1, data-driven methods show a remarkable ability to maintain smaller errors even during the GFC. However, during the COVID-19 crisis, LASSO performed almost as poorly as traditional methods in Japan, France, and Germany, showing a large increase in accumulated errors during this period.

The analysis of the prediction accuracy results, as illustrated in Table 2.4, highlights the competitiveness of the PLS procedure as the leading method for performing temporal disaggregation based on the MAE and RMSE assessments. The relatively smaller MAE and RMSE associated with this method suggest its superior performance in predicting

21. Single indicators in the case of traditional methods and at least 30 macroeconomic indicators are used in the specific case of the LASSO and PLS procedures - I do not report the country-specific covariate names for brevity. More specifically, I use country-specific Total Industrial Production Index (IPI) as a single indicator similar to the majority of previous work in this literature, for example, Moauro and Savio (2005).

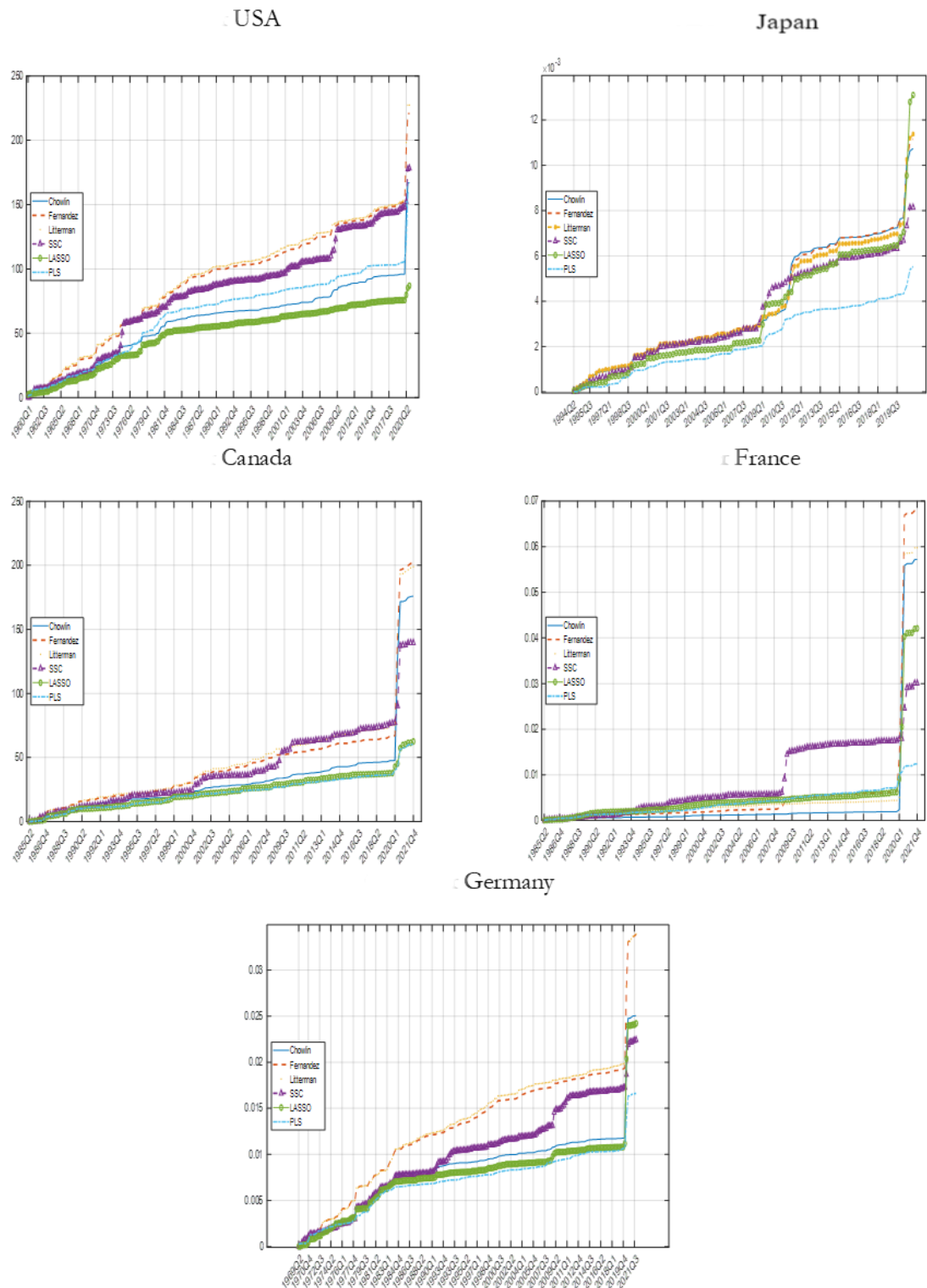


Figure 2.2: CSPE for Broad Indicator Estimations

Notes: The CSPE for the Chow-Lin, Fernandez, Litterman, SSC, LASSO and the PLS methods are plotted for estimated quarterly GDP of respective countries. The method with relatively smaller accumulated errors for the period has superior performance.

missing high frequency time series, even without the imposition of temporal constraints. When contextualised alongside the findings of the earlier temporal disaggregation experiment, this outcome underscores the superior predictive capability of the PLS approach compared to competitor methodologies.

Furthermore, I also observe that the estimates generated by the LASSO and PLS procedures, exhibit a distinct lack of smoothness but demonstrate robust alignment with the actual GDP series. This observation underscores the competitiveness of these methodologies in capturing nuanced variations within the data.

Table 2.4: Temporal Disaggregation Results: GDP

	USA		Japan		Canada		France		Germany	
Method	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Chow-Lin	0.5178	0.8274	0.0055	0.0072	0.5825	1.0941	0.0055	0.0197	0.0062	0.0109
Fernandez	0.6599	0.9514	0.0071	0.0111	0.6804	1.1737	0.0076	0.0215	0.0088	0.0127
Litterman	0.6704	0.9653	0.0070	0.0100	0.6970	1.1631	0.0072	0.0202	0.0089	0.0127
SSC	0.5939	0.8546	0.0071	0.0102	0.6427	0.9741	0.0083	0.0143	0.0073	0.0103
LASSO	0.4815	0.7467	0.0071	0.0103	0.4777	0.6523	0.0078	0.0169	0.0060	0.0107
PLS	0.5559	0.8075	0.0066	0.0087	0.4633	0.6418	0.0065	0.0092	0.0061	0.0089

Notes: The entries of the table report the Mean Absolute Errors (MAE) and Root Mean Square Errors (RMSE) corresponding to each country's GDP and method. For each country, the MAE and RMSE are compared across the methods. The smallest value means that the respective model performs better than the competitor methods. For example, the bold RMSE indicate the lowest RMSE across all models for a given variable.

2.5.5 Additional Results: GDP Nowcasting

In this section, I conduct a sub-sample analysis by nowcasting quarterly GDP and drawing insights from the mixed frequency methodology and approach used by Kuzin et al. (2011). The analysis encompasses *pseudo* nowcasts for GDP growth during different economic phases: stable periods (SP), Global Financial Crisis (GFC) between 2008 and 2009, and the COVID-19 crisis of 2020-2021, focusing on five major economies: the USA, Canada, Japan, Germany, and France.

To ensure fair treatment of competing methods, I employ a recursive prediction strategy where all the models are updated recursively as new data become available over time. This means that the models are re-estimated periodically using the most recent data, allowing them to adapt to changing economic conditions and incorporate new information for nowcasting. In the context of GDP nowcasting, the objective is to predict the most

recent quarterly numbers based on available monthly related economic indicators. The recursive strategy, therefore, allows for the incorporation of the latest information into all the procedures, leading to more accurate and timely predictions of GDP growth. So, similar to Kuzin et al. (2011) the predictions are made at the monthly frequency and then aggregated to derive quarterly estimates, and the prediction errors are computed by contrasting the actual quarterly values with the predicted aggregated time series of GDP growth. In addition, I introduce an auto-regression (AR) model as a reference for each country, a choice supported by previous research, such as the work of Armesto et al. (2010).

The results, summarised in Table 2.5 – Table 2.7, present the relative Mean Absolute Errors (rMAE) and relative Root Mean Squared Errors (rRMSE) to imply comparisons with the AR model for each sub-sample. Notably, nowcasting results indicate that during stable periods, all methods demonstrate lower MAE and RMSE in absolute terms compared to their own performance during crises periods, confirming that all methods generate large errors during periods of high uncertainty. More specifically, it can be seen that the performance of temporal disaggregation methods varies across countries during the financial crisis, and no single method consistently outperforms others. However, these methods generally exhibit favourable performance compared to the benchmark AR model across all countries. However, the most notable feature is that the PLS procedure emerged as the most consistent performer, delivering substantial improvements over the benchmark in all countries during the financial crisis, underscoring its effectiveness in navigating this challenging economic landscape.

During the relatively stable period from 2016 to 2019, the evaluation results in Table 2.6 reveal that the AR model is not particularly easy to outperform even by data-driven methods. In some specific instances, such as Japan and France, the PLS fails to beat the AR model. I observe a similar performance result of the LASSO for France. More specifically, the PLS was the leader in the USA, but LASSO performed better in Germany and Canada. The results also demonstrate that traditional methods tend to be effective in low volatility periods, as demonstrated by the remarkable results in France and Germany.

Table 2.5: Nowcasting Results During GFC

	USA		Japan		Germany		Canada		France	
Method	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE
AR-MIDAS	1.0142	1.1155	1.1180	1.0135	1.0127	1.0099	0.9974	0.9981	1.0897	0.9700
Chowlin	1.0818	1.0737	0.4286	0.4054	0.9873	0.8812	0.5081	0.5002	0.5128	0.5000
Fernandez	1.2446	1.2087	0.6273	0.6577	0.9114	0.8812	0.5675	0.5776	1.6282	1.8100
Litterman	1.1603	1.1057	0.4658	0.4189	0.7342	0.7129	0.7128	0.6937	0.3718	0.3700
SSC	1.7361	1.8121	0.4037	0.3874	1.3038	1.5842	1.1397	1.2057	3.2051	3.4500
LASSO	0.6494	0.6050	0.4224	0.4009	1.2405	1.1485	0.4923	0.5437	0.4872	0.5600
PLS	0.9740	0.8000	0.7391	0.6757	0.8481	0.8020	0.5195	0.4915	0.8333	0.8500

Notes: The entries of the table report the *relative* Mean Absolute Errors (rMAE) and *relative* Root Mean Square Errors (rRMSE) corresponding to each country's GDP and method. The relativity is to the AR(1) model. For each country, the rMAE and rRMSE are compared across the methods. A value smaller than one indicates that the respective model performs better than the AR(1) benchmark model. Sample period is 2008Q1 and 2009Q4.

This observation is quite suggestive of model complementarity in practice, as most methods seem more competitive during the stable period. Although certain methods excel during periods of low volatility, others may prove more effective amid economic instability. However, the overall competitiveness of temporal disaggregation methods against the benchmark AR model underscores their relevance and effectiveness in the nowcasting domain. In addition, the findings shed light on the advantages of utilising data-driven models that are flexible when the macroeconomy is stable.

Table 2.6: Nowcasting Results During SP

	USA		Japan		Germany		Canada		France	
Method	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE
AR-MIDAS	0.4699	0.5238	1.0000	0.9400	0.9828	0.9836	0.9963	0.9966	1.0678	1.1538
Chowlin	0.6128	0.6462	1.0952	1.1200	0.3621	0.4262	0.5984	0.7145	0.3559	0.4000
Fernandez	0.9889	1.0959	1.0000	1.0000	0.8793	0.9016	0.6530	0.7734	0.6441	0.8462
Litterman	1.0302	1.1017	1.0238	1.0200	0.9483	0.9672	0.6653	0.8434	0.7119	0.7385
SSC	1.0087	1.1738	1.0000	1.0000	0.5172	0.5902	0.8314	0.9119	0.7627	0.9385
LASSO	0.6126	0.6727	0.9762	0.9800	0.3793	0.4262	0.4077	0.4566	1.0678	1.1077
PLS	0.4191	0.4625	1.0476	1.0000	0.4655	0.5410	0.5998	0.6459	1.0339	1.1846

Notes: The entries of the table report the *relative* Mean Absolute Errors (rMAE) and *relative* Root Mean Square Errors (rRMSE) corresponding to each country's GDP and method. The relativity is to the AR(1) model. For each country, the rMAE and rRMSE are compared across the methods. A value smaller than one indicates that the respective model performs better than the AR(1) benchmark model. Sample period between 2016Q1 and 2019Q4.

Finally, I highlight the variations in the performance of nowcasting models during the COVID-19 crisis period. The study finds that the performance of the nowcasting models is quite similar to that during the GFC and the results in Table 2.7 show that all data-driven methods surpass the benchmark model. The PLS procedure outperforms all models in

Japan, Germany and Canada and remains among top performers in the USA and France. In general, the findings suggest that the PLS and LASSO, frequently outperform the benchmark in situations where there is a high degree of economic uncertainty. However, the performance of traditional temporal disaggregation approaches is uneven, which highlights the weaknesses of traditional models in terms of their ability to react to sudden economic shocks such as the COVID-19 epidemic. Because I observed similar model performance during the two extreme periods, I refer to the literature for support and find that this competitive edge of data-driven procedures is not unexpected, as others, for example, Lima et al. (2020) have shown that during crisis periods like the COVID-19 crisis, extreme observations in time series data can challenge the assumptions of ordinary least squares estimators.

Table 2.7: Nowcasting Results During COVID-19

	USA		Japan		Germany		Canada		France	
Method	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE	rMAE	rRMSE
AR-MIDAS	0.8671	0.8648	0.7721	0.8232	0.9967	0.9954	1.0000	1.0000	0.4919	0.4300
Chowlin	0.5662	0.6171	0.4182	0.3684	0.8131	0.9379	0.6875	0.7538	0.9052	1.0297
Fernandez	0.5654	0.6075	0.9357	0.8337	0.9672	0.9839	0.7437	0.7744	1.0394	1.0880
Litterman	0.5748	0.6309	0.6032	0.5853	0.9443	0.9632	0.7088	0.7400	0.9302	1.0310
SSC	0.3671	0.4037	0.6649	0.6232	0.5967	0.5862	0.4737	0.5271	0.4884	0.4907
LASSO	0.2704	0.2447	0.7024	0.6653	0.8328	0.9425	0.4118	0.3307	0.8479	0.8302
PLS	0.5593	0.5430	0.4263	0.4105	0.5770	0.6276	0.3748	0.3240	0.3309	0.3172

Notes: The entries of the table report the *relative* Mean Absolute Errors (rMAE) and *relative* Root Mean Square Errors (rRMSE) corresponding to each country's GDP and method. The relativity is to the AR(1) model. For each country, the rMAE and rRMSE are compared across the methods. A value smaller than one indicates that the respective model performs better than the AR(1) benchmark model. Sample period between 2020Q1 and 2021Q4.

Consequently, models relying on OLS estimators may be outperformed by those utilising robust estimation methods like the data-driven LASSO and PLS approaches employed in this study. These findings presented in this section offer the final compelling evidence of the performance of data-driven procedures in nowcasting across diverse economic conditions, including stable and extreme periods.

2.6 Conclusions

In conclusion, this research contributes an original simple and valuable approach to the temporal disaggregation of time series literature through its development and assessment of a new approach, the Partial Least Squares (PLS) method. This method has the appeal of being simple and easily applicable and serves as a reliable competitor to existing traditional methods. The performance of the proposed approach is evaluated through a comparative analysis with the existing traditional Ordinary Least Squares regression-based techniques. The results of the comparative performance analysis between the PLS, LASSO and traditional methods indicates significant relative superiority of the PLS approach across various disaggregation scenarios. In this way, the current study, therefore, demonstrates and highlights the distinct benefits that emerge when data-driven approaches like the LASSO and the proposed PLS method are used in large dataset environments.

The present findings are of particular relevance during this time, when big data analytics plays a growing role in macroeconomic prediction processes and instantaneous data analysis. The PLS procedure demonstrates not only good performance but also superior predictive capabilities in multiple specific situations when compared with traditional methods. The research also demonstrates that individual traditional methods show desirable strengths collectively across different target indicators, indicating that a combination approach could often provide advantages.

The initial experimental investigation, referred to as Experiment 1 centers on temporal disaggregation with a limited number of narrow economic indicators. Under this constraint the study estimates monthly data of the respective quarterly indicators. The research results show that the LASSO and PLS extensions achieve notable reductions in Mean Absolute Error and Root Mean Squared Error when compared to traditional methods under these specific conditions. The relatively small prediction accuracy metrics show strong statistical proof in support of data-driven approaches. Graphical examination of Cumulative Squared Prediction Errors demonstrates the effectiveness of LASSO and PLS techniques in reducing the accumulated error over time. The graphical analysis reveals that data-driven models maintain lower error accumulation, thus demonstrating their outstanding stability and dependability in various testing conditions.

The study then expands its analytical scope to Experiment 2, which explores the temporal disaggregation of quarterly GDP growth for five major developed economies: namely, the United States, Canada, Japan, Germany, and France. This analysis tests the robustness of the models by removing temporal constraints to create a less restrictive testing environment. The main results indicate that the PLS and LASSO procedures jointly maintain superior performance over traditional regression-based methods even when evaluating broad indicators. Therefore, the results confirm the adaptability and effectiveness of data-driven methods through their consistent competitive performance across various countries.

For further robustness checks, the study expands its examination to include nowcasting GDP growth for the five countries. The process of nowcasting which uses estimated current data to predict immediate future conditions plays a vital role in the development of prompt policy decisions. Again, in the nowcasting exercises, the main study results indicate that data-driven methods have relatively superior performance by nowcasting GDP growth with higher precision. This observation provides additional support for the practical relevance and real-world utility of the PLS approach through its ability to deliver precise nowcasts. The proposed method, in particular, shows stability during economic downturns or times of increased volatility like the financial crisis and the pandemic.

Finally, the study, therefore, makes a case and demonstrates through empirical evidence that data-driven methods, in particular, the PLS approach, are indeed a valuable econometric tool for temporal disaggregation and nowcasting of time series. In general, the study demonstrates that the proposed simple method delivers competitive predictive outcomes while maintaining high adaptability and robustness in most applications and data environments. As such, researchers, analysts, and policymakers working in data rich environments can find this method useful because of its capability to manage high-dimensional datasets without overfitting and its stability across various macroeconomic conditions.

Chapter 2: Appendix

A Low Frequency Vs High Frequency Time Series

Here, I attempt to briefly explain the difference between low frequency and high frequency time series. First, low frequency time series are those that have extended, constant gaps between data points, typically occurring on a quarterly or annual basis. As an illustration, when economists report GDP on a quarterly basis, they publish the total output for each of the three months (for example, the GDP for the first quarter of 2025 was USD 1.2 Billion). When all of the economic activity for the year is summed up into a single number, you achieve the annual GDP figure. The low frequency series are typically useful in tracking long-term growth rates, general trends, and structural changes of different countries. Trends are nice for economists because they smooth out short-term variances, which allows them to better capture the economic trajectory. They however, are unable to detect rapid shifts (intra-period changes) in consumer spending or sudden demand shocks within the upcoming quarter or year - this inability to capture intra-period variations is the biggest downside of low frequency time series.

In contrast, a high frequency time series is able to provide researchers with the ability to monitor economic conditions in a relatively short period of time. This is because it collects data at shorter intervals, typically on a monthly, weekly, or even daily basis. Statistical agencies and business forecasters employ high frequency proxies and nowcasting tools to provide monthly “flash” predictions of GDP growth. This is despite the fact that official GDP is rarely reported more frequently than quarterly. A central bank, for instance, might make use of monthly figures on industrial production, retail sales volumes, power consumption, and tax receipts in order to make a prediction about what the quarterly GDP might look like before the official release. During the intervals between the slower official GDP reports, these high frequency indicators assist businesses and governments in identifying economic turning points, reacting rapidly to emerging threats, and improving their projections.

B Predictor Indicators for the LASSO and PLS

Appendix Table 2.8 reports the high frequency variables used in the study experiments for U.S economy.

Table 2.8: Predictor Indicators for the LASSO and PLS

#	Code	Variable Name	Frequency
1	RPI	Real Personal Income	Monthly
2	W875RX1	Real personal income ex transfer receipts	Monthly
3	DPCERA3M086SBEA	Real personal consumption expenditures	Monthly
4	CMRMTSPLx	Real Manu. and Trade Industries Sales	Monthly
5	RETAILx	Retail and Food Services Sales	Monthly
6	INDPRO	IP Index	Monthly
7	IPFPNSS	IP: Final Products and Nonindustrial Supplies	Monthly
8	IPFINAL	IP: Final Products (Market Group)	Monthly
9	IPCONGD	IP: Consumer Goods	Monthly
10	IPDCONGD	IP: Durable Consumer Goods	Monthly
11	IPNCONGD	IP: Nondurable Consumer Goods	Monthly
12	IPBUSEQ	IP: Business Equipment	Monthly
13	IPDMAT	IP: Durable Materials	Monthly
14	IPNMAT	IP: Nondurable Materials	Monthly
15	IPMANSICS	IP: Manufacturing (SIC)	Monthly
16	IPB51222S	IP: Residential Utilities	Monthly
17	IPFUELS	IP: Fuels	Monthly
18	CUMFNS	Capacity Utilisation: Manufacturing	Monthly
19	HWI	Help-Wanted Index for United States	Monthly
20	HWIURATIO	Ratio of Help Wanted/No. Unemployed	Monthly
21	CLF16OV	Civilian Labor Force	Monthly
22	CE16OV	Civilian Employment	Monthly
23	UNRATE	Civilian Unemployment Rate	Monthly
24	UEMPMEAN	Average Duration of Unemployment (Weeks)	Monthly
25	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	Monthly
26	UEMP5TO14	Civilians Unemployed for 5-14 Weeks	Monthly
27	UEMP15OV	Civilians Unemployed - 15 Weeks and over	Monthly
28	UEMP15T26	Civilians Unemployed for 15-26 Weeks	Monthly
29	UEMP27OV	Civilians Unemployed for 27 Weeks and Over	Monthly
30	CLAIMSx	Initial Claims	Monthly
31	PAYEMS	All Employees: Total nonfarm	Monthly
32	USGOOD	All Employees: Goods-Producing Industries	Monthly
33	CES1021000001	All Employees: Mining and Logging: Mining	Monthly
34	USCONS	All Employees: Construction	Monthly
35	MANEMP	All Employees: Manufacturing	Monthly
36	DMANEMP	All Employees: Durable goods	Monthly
37	NDMANEMP	All Employees: Nondurable goods	Monthly

Continued on next page

Table 2.8 – *Continued from previous page*

#	Code	Variable Name	Frequency
38	SRVPRD	All Employees: Service-Providing Industries	Monthly
39	USTPU	All Employees: Trade, Transportation and Utilities	Monthly
40	USWTRADE	All Employees: Wholesale Trade	Monthly
41	USTRADE	All Employees: Retail Trade	Monthly
42	USFIRE	All Employees: Financial Activities	Monthly
43	USGOVT	All Employees: Government	Monthly
44	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing	Monthly
45	AWHMAN	Avg Weekly Hours : Manufacturing	Monthly
46	HOUST	Housing Starts: Total New Privately Owned	Monthly
47	HOUSTNE	Housing Starts, Northeast	Monthly
48	HOUSTMW	Housing Starts, Midwest	Monthly
49	HOUSTS	Housing Starts, South	Monthly
50	HOUSTW	Housing Starts, West	Monthly
51	PERMIT	New Private Housing Permits (SAAR)	Monthly
52	PERMITNE	New Private Housing Permits, Northeast (SAAR)	Monthly
53	PERMITMW	New Private Housing Permits, Midwest (SAAR)	Monthly
54	PERMITW	New Private Housing Permits, West (SAAR)	Monthly
55	AMDMMNOx	New Orders for Durable Goods	Monthly
56	BUSINVx	Total Business Inventories	Monthly
57	M1SL M1	Money Stock	Monthly
58	M2SL M2	Money Stock	Monthly
59	M2REAL	Real M2 Money Stock	Monthly
60	BUSLOANS	Commercial and Industrial Loans	Monthly
61	REALLN	Real Estate Loans at All Commercial Banks	Monthly
62	NONREVSL	Total Nonrevolving Credit	Monthly
63	CONSPI	Nonrevolving consumer credit to Personal Income	Monthly
64	PPICMM	PPI: Metals and metal products	Monthly
65	CPIAUCSL	CPI : All Items	Monthly
66	CPIAPPSL	CPI : Apparel	Monthly
67	CPITRNSL	CPI : Transportation	Monthly

Notes: The table reports monthly economic indicators. All data are downloaded from <https://research.stlouisfed.org/econ/mccracken/fred-databases/>

C Application in Data Scarce Environments

A data-driven temporal disaggregation method, such as LASSO or indeed the PLS is fairly easy to implement when desired related high frequency and target low frequency time series are available. However, in data scarce environments, the application of high-dimensional temporal disaggregation methods as the case may be in most developing countries may not be exactly a straightforward matter because data are generally scarce or fragmented. The procedural requirement for high frequency related indicators demands practical solutions to overcome data constraints.

Most developed countries, on the other hand, have well-developed data collection systems that are able to gather vast economic data regularly in usable format for time series modeling. This makes the use of data-driven procedures relatively easy. So, high-dimensional time series modeling may likely be hampered in many developing country settings due to irregular releases and limited historical observations on various indicators. Therefore, in this section, I briefly discuss suggestions of some well-known practical and non-interpolation solutions that can be followed to reduce data scarcity and support the usability of data-driven approaches in data scarce environments.

C.1 Mapping and Characterising Data Gaps

Given a specific target variable, as a general rule, start with a thorough audit of accessible related high frequency indicator series. For every variable, note the frequency (for example, monthly, quarterly), sample span, and missing intervals. Decide if the gaps are Not Missing At Random (NMAR), Missing At Random (MAR), or Missing Completely At Random (MCAR). This classification helps to select appropriate imputation methods for each specific type of missingness. Complete the time series before initiating the temporal disaggregation approaches.

C.2 High Frequency Proxies

Without directly interpolating high frequency covariates, proxy indicators offer a way to enhance low frequency disaggregation objectives. The use of informative non-traditional proxies, such as satellite nightlights for economic activity, electricity use for industrial production, or web search trends for consumer mood, is becoming more and more popular. Dig deep and find informative related indicators, and trial and error can come in handy.

C.3 Tapping Open and Commercial Data Sources

By looking through both paid and publicly available feeds, one can expand their data universe. The World Bank Development Indicators, FRED, and the IMF International Financial Statistics are a few examples. Other possibilities include scraping labor market dashboards, national statistics offices, central banks, and ministries' websites.

D Strengthening Data Sources in LDCs

Many developing countries lack the resources, infrastructure, knowledge and logistic support to gather fast and accurate economic data. In the following, I review and compile some main strategies that may help to enhance statistical systems and make use of new data sources:

- Build the capability of the institution by increasing the number of professional statisticians, training, and retention. Due to the nature of this process, it is necessary to recruit and keep qualified statisticians, economists, and information technology specialists who have established careers and are paid fairly.
- Provide assistance to statistics agencies and work to improve legal systems that require regular data collecting and reporting. Developing and implementing robust legal frameworks and rules that demand regular data collection, reporting, and disclosure is something that the host LDCs need to do in order to fulfil their obligations. Providing these obligations with legal status often provide organisations with the authority and funds they require.

- Encourage the formation of coalitions by working together with academic institutions and think tanks to provide technical assistance and peer review. The cultivation of a culture of statistical excellence that is long-lasting can be facilitated via collaborations. Through the formation of strategic relationships with academic institutions and research teams, it is possible to acquire the much-required technical support, training programs, and peer evaluations.
- International organisations such as the United Nations Statistics Division, the International Monetary Fund, and the World Bank can provide LDCs with funding and methodological direction, allowing them to strengthen their data systems. It is possible to acquire additional data systems support, which can assist in the incorporation of best practices in survey methodology, data processing, and distribution.
- Processes for conducting surveys and censuses should be revised. Take use of digital data capture: iPads or smartphones should be used in place of paper surveys in order to cut down on errors and speed up the processing time.
- The Procedures for the Census and Survey need to be modified in order to accommodate the LDCs. Traditional paper-based surveys and censuses conducted every 10 years are notoriously expensive, time-consuming, and prone to inaccuracies, especially in areas that are difficult to access or distant. The use of digital data-capture methods, such as Computer-Assisted Personal Interviewing (CAPI) and Computer-Assisted Telephone Interviewing (CATI), will eventually result in a reduction in operational expenses, a speeding up of data processing, and a reduction in transcribing errors.
- Make use of geospatial stratification and stratified sampling in order to guarantee representative coverage, particularly in locations that are geographically isolated. Through stratification according to area, industry, or population group, a sample that has been carefully designed ensures that it is representative of the population and that it allocates resources effectively. Statistical agencies may be able to supply data of a higher quality more frequently and at a cheaper cost if they make use of new survey methodologies.

- Mix “Big” Data with administrative records to achieve the following: The ability to track GDP components, employment, and commerce is made possible through the linking of data from social protection, business registrations, tax filings, and customs declarations. Administrative records, which include tax returns, business registrations, customs declarations, and social protection databases, are a hidden gold mine of economic data produced by government operations. This is no longer a secret; it is now common knowledge that these records contain a wealth of information. Through their incorporation into official statistics, these sources have the potential to alleviate the burden placed on respondents and to offer continuous and nearly real-time insights into the activity of the economy.
- Utilise the resources that are available in the private sector: Work together with telecommunications companies to provide anonymised call-detail records for real-time mobility and activity proxies. It is possible that partnerships with the business sector will result in anonymous streams of “big data”. An example of this would be the records of credit and point-of-sale transactions, which can provide insight into retail trends. Additionally, mobile phone call logs can serve as a proxy for changes in the population and patterns in consumption.
- Nowadays, it is possible to estimate production in the manufacturing, mining, and agricultural sectors by using satellite pictures, such as night-light data. For example, in the District of Columbia, satellite photography has supplied an extra and supplementary perspective on economic output. This perspective includes statistics on agricultural land-use classification and nocturnal light intensity, particularly for enterprises that are informal or located in rural areas. In order to provide economic indicators that are more comprehensive, up-to-date, and comprehensive, LDCs have the ability to utilise this approach, which allows them to combine data from administrative and big-data sources with data from traditional surveys.
- The International Monetary Fund’s Government Finance Statistics Manual, the System of National Accounts (SNA), and the International Standard Industrial Classification (ISIC) should all be utilised. ensuring that the data maintains its quality and standards. For the purpose of producing high-quality statistics, the

foundation consists of categories that are accepted internationally and methodological criteria that are stringent. Countries that are still in the process of developing should use the most recent editions. The use of standardised classifications makes it considerably simpler to facilitate methodological modifications, cross-national comparisons, and cross-temporal relationships.

- In order to improve data systems, it is essential to develop metadata and papers. Open documentation, which contains full metadata, sample frames, data-processing techniques, and modification procedures that inform users of the benefits, limitations, and potential biases of the data, is an essential component.
- It is important to establish quality assurance processes, which include checking and verifying data on a regular basis, resolving source discrepancies, and monitoring non-response or undercoverage. Methods of quality assurance that are performed on a regular basis, such as inter-agency reconciliations, external peer reviews, and internal audits, are helpful in identifying and rectifying errors, validating sampling plans, and monitoring undercoverage or non-response. This continuous dedication to quality helps to cultivate trust among users and serves as a guide for the formulation of policy that is evidence-based.
- Encourage the use of open data and transparency. Through the process of making government data more accessible to the general public, open-data projects encourage accountability and innovation. In order to enable programmatic access through application programming interfaces (APIs) and publish data under open licenses, governments have the ability to construct centralised web-based portals. Businesses, academic institutions, and organisations that are part of civil society are all encouraged to develop innovative apps, carry out their own independent research, and monitor the performance of the government as a result of this. When distributing data, it is important to ensure that priority gaps and user demands are taken into consideration. This can be accomplished by the holding of frequent stakeholder meetings, which incorporate members from the private, academic, and public sectors. This transparency in data governance, which includes clear disclos-

ures of the data source and revision deadlines, contributes to an even greater gain in public trust. It is possible for developing countries to boost the impact of their statistics investments and achieve greater economic and social advantages if they adopt the principles of open data.

- Improve both the frequency and the timeliness of your communications. Instantaneous forecasting and short-term indicators: It was said before that timely data is essential for responsive policy interventions, particularly during times of emergency or periods of fast structural change. Such interventions are especially important. Nowcasting techniques are vital and can be used to complement slower, more traditional signals such as quarterly GDP or monthly CPI by utilising high frequency proxies such as credit card spending, industrial power consumption, or online search trends. These proxies can be used to supplement slower signals. The rapid ingestion, validation, and publication of essential metrics on a weekly or even daily basis can be made possible with the assistance of automated data pipelines that are based on cloud architecture infrastructure. In a similar vein, interactive dashboards enhance situational awareness and decision-making by providing the general public and politicians with access to trends in real time. Statistical agencies are able to deliver the actionable data that modern economies require and do so by placing an emphasis on timely release and increasing the frequency of their publications.
- It is important to make investments in both creativity and technology. New technologies have made it feasible to develop powerful tools that can improve data collection, processing, and analysis. Switching to cloud systems that are both secure and scalable can have a number of benefits, including the centralisation of data storage and the encouragement of agency collaboration. Consequently, the classification of company registry documents, customs filings, or online job postings can be accomplished through the utilisation of machine-learning algorithms and natural language processing technology. This provides early indicators of changes in the labour market or growth in specific sectors. When it comes to monitoring environmental changes and agricultural productivity, applications for geospatial analytics and remote sensing offer a level of granularity that is unmatched. Stat-

istics agencies have the ability to cultivate an innovative environment by means of pilot projects and partnerships with other stakeholders. This environment enables the statistical agencies to test new methods, evaluate the efficiency of those strategies, and progressively incorporate winning concepts into their normal operations. This dedication to technological progress has the potential to guarantee that statistical systems will continue to be adaptive and relevant.

E Example of Types of Economic Data

Nowadays, there are huge amounts of high frequency indicators whose data are generated frequently across various sectors, the table below is just an example of types of economic data that may be useful in applying a data-driven method in both data rich and scarce environments in order to disaggregate broad economic indicators.

Table 2.9: Economic Sectors and Key Data Types

Sector	Types of Economic Data
Agriculture, Forestry and Fishing	• Production and Output: crop yields; livestock counts; fish catch volumes • Prices and Indices: farm-gate prices; commodity price indices; input-cost indices • Trade and Stocks: export/import volumes and values; on-farm inventories • Investment and Inputs: land-use area; machinery and irrigation investment • Employment and Wages: seasonal and permanent employment; average farm wages; rural unemployment rate

Table 2.9 (continued)

Sector	Types of Economic Data
Mining and Quarrying	• Extraction Volumes: ore; coal; metallic minerals; aggregates • Sales and Exports: domestic sales; export volumes and values • Prices and Margins: international commodity prices; domestic sale prices • Capacity and Utilisation: operating mines; capacity-utilisation rates • Employment and Wages: workforce numbers; average mine wages; occupational categories • Safety and Incidents: accident and injury statistics
Manufacturing	• Output and Indices: production indices by sub-sector • Capacity and Utilisation: installed capacity; utilisation rates • Orders and Inventories: new orders; raw-materials and finished-goods stocks • Trade Flows: export/import by HS code; trade balances • Prices and Costs: PPI; unit-labour costs; input-cost indices • Investment and R and D: capital expenditure; R and D spending; equipment orders • Employment and Wages: headcount; average wages; hours worked

Table 2.9 (continued)

Sector	Types of Economic Data
Construction	• Activity and Output: construction-put-in-place; housing starts/completions • Permits and Backlogs: building permits issued; contract backlogs • Prices and Costs: construction cost index; material price indices • Investment: public and private construction spending • Employment and Wages: labour force size; average hours worked; average wages
Utilities (Electricity, Gas, Water)	• Production and Consumption: generation (MWh); throughput (m³); water supply (m³) • Prices and Tariffs: retail/wholesale tariffs; peak/off-peak rates • Capacity and Reliability: installed capacity; outages; reserve margins • Investment and Infrastructure: grid; pipeline • Employment and Wages: workforce size; average wages
Wholesale and Retail Trade	• Sales and Turnover: retail sales (value/volume); wholesale turnover • Price Indices: retail price index; margin spreads • Inventories: stock levels; days-of-stock • Profitability: gross/net margins • Employment and Wages: number of employees; average hours; average wages

Table 2.9 (continued)

Sector	Types of Economic Data
Transportation and Logistics	• Freight and Passenger Volumes: ton-km; passenger-km by mode • Throughput: airport passenger and cargo; ports • Prices and Fares: freight rates; passenger fares • Fleet and Utilisation: vehicle/ship counts; capacity utilisation • Investment: infrastructure spending on roads, ports, airports • Employment and Wages: workforce size; average hours; average wages
Information and Communication Technology (ICT)	• Subscriptions and Usage: mobile/broadband subscriptions; data volumes • Revenues and Expenditure: telecom revenues; IT services spending • ARPU and Prices: average revenue per user; equipment price indices • Innovation Metrics: patents filed; R and D intensity • Employment and Wages: ICT workforce size; average wages; skill distribution
Financial and Insurance Services	• Balance Sheets: assets; liabilities; capital ratios • Credit and Lending: loan volumes; NPLs; interest spreads • Market Activity: equity market cap; trading volumes; bond issuances • Rates and Premiums: deposit/loan rates; insurance premiums • Employment and Wages: number of employees; average compensation; vacancy rates

Table 2.9 (continued)

Sector	Types of Economic Data
Real Estate and Rentals	• Property Prices: residential/commercial indices; rental yields • Transactions: sales volumes and values; building permits • Vacancy and Utilisation: vacancy rates; occupancy levels • Investment Flows: Foreign Direct Investment in real estate • Employment and Wages: sector employment; average wages
Professional, Scientific and Technical Services	• Service Revenues: consulting; legal; accounting fees • Prices and Rates: average hourly/daily rates • Innovation Output: reports published; patents supported • Employment and Wages: headcount; billable hours; average rates
Health and Education	• Expenditure and Funding: public/private spending; tuition fees • Utilisation: hospital admissions; student enrollments • Prices and Wages: medical/tuition price indices; staff wages • Employment and Wages: number of practitioners/teachers; turnover rates
Hospitality, Arts and Recreation	• Activity and Revenue: hotel occupancy; revenues; ticket sales • Price Indices: accommodation and admission-fee indices • Capacity and Utilisation: room and event-space usage • Employment and Wages: sector headcount; average wages; employment type breakdown

F Delays in GDP Reporting

Countries' through their respective national statistics organisations utilise GDP as a comprehensive measure of economic activity. However, compiling it generally requires gathering data from many different sources. Output numbers must all be calculated using industrial production surveys, retail sales records, trade statistics, business investment questionnaires, government finance accounts, agricultural yields, services sector revenues, and more. Because every one of these component series arrives on a different timetable, for example, monthly, some quarterly, and others even less frequent, the statistical offices have to wait for the most current data from the slowest contributors before compiling a thorough GDP figure.

Apart from the simple speed of data arrival, GDP compilation grapples with discrepancies in reporting frequency and timing. It is argued that data comes in at different frequencies, for instance, annual for some administrative registers, quarterly reports for business-investment surveys, and monthly reports for retail sales. As such, to produce a consistent quarterly GDP figure, statisticians have to interpolate or allocate lower-frequency observations into the suitable quarter, usually using benchmark ratios or temporal-disaggregation methods. Creating these estimates calls for greater time and detailed methodological judgement since they must accurately reflect underlying economic changes.

After acquiring the raw data, agencies have to do thorough validation, balancing, and seasonal adjustment techniques. Raw survey responses and administrative data could contain discrepancies, reporting mistakes, or outliers. Statistical analysts then have to clean and reconcile these data to guarantee that the accounting identity, that is, output equals income equals expenditure must hold exactly. They then apply seasonal adjustment filters to remove weather-related patterns or holiday-related recurrent calendar events. Although all these activities add to the general delay, every one of these quality control techniques is required to provide reliable and internally consistent GDP figures.

Many agencies publish a preliminary or “flash” GDP estimate based on known proxies and partial data just after the end of the quarter, sometimes within 30 to 45 days, therefore providing users rapid information. As more full data becomes available such as finalised trade figures or detailed tax returns—the agency publishes second and final estimates, typically two to three months following the first release. Periodic annual or multi-year thorough revisions including further methodological reconsiderations help to further hone past GDP numbers and ensure the long-run series stays as accurate as possible.

Finally, national accounts, are also known to follow strict international standards such as those described in the UN System of National Accounts. These systems provide precise definitions of economic activities, valuation rules, and industry classifications helping to ensure comparability between countries and over time. Official numbers can only follow from the careful, deliberate effort supporting high quality economic statistics. Preparing GDP in line with these conventions requires for expert knowledge and precise application of the suggested techniques.

G Narrow vs Broad Economic Indicators

Narrow indicators are economic indicators that give a limited perspective owing to the nature of what they measure. They are normally compiled regularly in real time and at high frequency. They are single series measures that track a particular aspect of economic activity. For example, among them are industrial output indices, building permits, new car registrations, retail sales volumes, and the purchasing managers’ index (PMI). Every one of these provides data on a different aspect of the economy including manufacturing, trade, housing, mood, and so forth. Narrow indicators all have the possibility to vary considerably in reaction to shocks unique to a particular sector, such as a strike in car plants or an increase in the number of home starts. Narrow indicators have the possibility of being very timely and sensitive to turning points since they only concentrate on one sector.

In contrast, broad economic indicators provide a more complete view of the general performance of the economy by combining several various data sources or complete aggregates. Among the examples deemed classic are the consumer-price index, a weighted average of prices across a basket of goods and services; the gross domestic product, an aggregate of output across all sectors; and composite leading indicators, which include data on employment, production, finances, and sentiment. Although they are usually reported less often (quarterly or annually) and have greater delays than narrower measures, broad measures are usually more stable and better reflect the underlying trend. Since they filter out sectoral noise, broad indicators are crucial for assessing the underlying momentum of the economy and for modifying policy. Nevertheless, to do this, they give up timeliness and granularity.

Chapter 3

Differential Effects of a Financial Shock and the Real Economic Activity Network of the United States

1 Introduction

A major obstacle to understanding regional economic dynamics in the United States is the absence of high frequency real GDP growth time series at the state level. There are no official quarterly real GDP growth data for the period prior to 2005 for individual states, and previous research in the literature has had to rely on proxies or aggregate national data to infer state level economic outcomes. This lack of data not only makes it more difficult to evaluate how regional economies react to macroeconomic shocks, but it also makes it difficult to see how recovery and growth patterns vary from state to state.

Therefore, in this chapter, I aim to achieve three main objectives. First, I aim to create the missing historical quarterly real GDP growth data spanning from 1991 to 2004 for the 48 contiguous U.S. states (excluding Alaska and Hawaii). This is achieved using quarterly state-specific multiple predictor indicators, annual state-specific GDP and the PLS temporal disaggregation procedure. Second, the study utilises the newly constructed quarterly dataset together with the official time series from 2005 onward to identify the differential effects of a financial shock on the real economy of individual states and, lastly, identify the economic clusters among states¹.

More specifically, I use the excess bond premium developed by Gilchrist and Zakrajšek (2012b) to identify an aggregate financial disruption and analyse its impact using a structural VAR model for each state. Further, the study uses both the separate and the merged time series to examine the unobserved characteristics of economic activity interdependence among U.S. states using Network Analysis, and the economic clusters or communities among the states. The objective here is to assess whether there is connectivity progression in economic activity of individual states, and clustering among them which has important implications for development planning and economic risk management.

1. Official state-level GDP data from the U.S Bureau of Economic Analysis can be downloaded from: <https://www.bea.gov/>

Because data is the foundation of the investigation that is now being conducted, a significant amount of attention is devoted to discussing the many factors to take into account and the complexities involved in producing the high frequency historical data that is missing. This is because a data set that has been carefully constructed is essential for gaining relevant economic insights into the many dynamics of particular states, which in turn helps to improve the validity and relevance of the conclusions of the study.

The data estimation results reveal that the PLS technique using industry employment indicators that are specific to individual states produces superior time series compared to other approaches that make use of national indicators and state-specific indicators. The final resultant time series is used to contribute to two strands of literature by making use of relatively accurate estimations of high frequency time series. First, the estimated state level output growth data are used to provide fresh insights into the heterogeneous effect of a financial shock on regional real economies. This is in contrast to previous studies that have utilised employment statistics as proxy variables for economic activity, such as Carlino and DeFina (1998). The analytical approach in the current study follows the conventional structural Vector Autoregression (VAR) model which is in line with the methodology described by Gilchrist and Zakrajšek (2012b) and Gilchrist and Zakrajšek (2012a). This methodology incorporates a number of economic indicators that have a multi-state dimension. So, following Gilchrist and Zakrajšek (2012b) I investigate the effects that a financial shock proxied by the excess bond premium has on the macroeconomic system using impulse response functions. The impulse response functions trace the responses of individual state output growth and other macroeconomic variables to an aggregate financial shock. One of the most important findings suggests that a financial shock causes a decline in outputs in the majority states, although to varying degrees. Certain states demonstrate resilience in the face of financial disturbances by having insignificant impulse response functions. The heterogeneity of states across the United States is brought into focus by this observation. The findings of this study fill a vacuum in the existing body of research on the influence of financial shocks on the real economy at the state level. Furthermore, they highlight the relevance of approximated high frequency time series in the process of understanding the dynamics of the macro economy when desired high frequency time series is missing.

Further, I employ both the estimated high frequency time series (1991Q1 - 2004Q4) and official data (2005Q1 - 2022Q4), both separately and in conjunction with one another, to discover aspects of economic activity interconnectedness across the states of the United States through the utilisation of network analysis. The findings indicate that there has been a rise in the connection of economic activity among states, as well as the clear formation of communities that are interrelated. By gaining an understanding of the output network, individuals, organisations, and policymakers can be assisted in the development of strategies that can better utilise network effects and improve economic performance.

To my knowledge, this study is an early attempt to construct the historical high frequency state-specific output growth using the PLS procedure and state-specific high frequency indicators and analyse the differential effects of a national financial shock on state-specific output and the economic activity network of the U.S. Studies in the literature use available employment data as a proxy for economic activity, for example, Hamilton and Owyang (2012) and Brownlees et al. (2022).

The rest of this study is organised as follows: Section 2 discusses the context of the states' data problem; then; Section 3 construction of historical state data; Section 4 discusses the differential state effects of an aggregate financial shock; Section 5 provides a network analysis of the United States output; and lastly, Section 6 offers a conclusion.

2 Context of the States' Data Problem

2.1 National Aggregate and State Indicator Selection

It is crucial to acknowledge that there are no official state-level quarterly output growth estimates for the individual U.S. states for the period prior to 2005. It is generally agreed that relying solely on available annual state data for economic analyses hugely obscures critical intra-year fluctuations, which are essential for understanding macroeconomic dynamics, for example, how an aggregate economic shock propagates across states. The same applies to appreciating the interconnectedness of economic activity of different states. Using low frequency data, therefore, has the disadvantage of potential misguidance of short-term to medium-term policy responses. This is why many studies in the literature

have tended to use proxy quarterly variables for state level research inquiries despite of annual times series being available. These considerations underscore the importance of estimating the missing state-level high frequency time series data. In this study, I employ the PLS procedure and various temporal disaggregation techniques discussed in the previous chapter and construct historical individual state output time series. I select the most robust time series from competing methods for utilisation in the subsequent analyses.

An initial assessment reveals that crucial requirements for conducting temporal disaggregation of annual state data are not met immediately, particularly regarding access to state-specific high frequency related indicators. Therefore, the missing data problem at the state level can be characterised in two ways. First, traditional indicators commonly used for converting low frequency output to higher frequency data are either unavailable or are too short to facilitate credible parameter estimation in state level models, and later on, credible estimates by traditional methods. Second, while high frequency industry-specific statistics, such as employment, are accessible at the state level, their availability only extends back to 1990. However, these time series can serve as informants in data-driven procedures and, therefore, help extend the available output growth time series. These data availability challenges have costs that significantly influence my analytical approach and constrain the period for which I can generate missing data. I note that successful navigation of these constraints requires careful consideration to ensure the reliability and validity of my applied research.

Given the absence of suitable state indicators, I initially propose performing temporal disaggregation of state annual output time series by using national predictors as covariates in individual state models. In particular, this means that addressing the primary missing data issue at the state level requires adjustments to the temporal disaggregation framework employed by traditional methods. In this formulation, national-level indicators act as surrogates for capturing short-term fluctuations in the target individual state output growth. Therefore, the estimated high frequency series of the target mirrors the short-term dynamics observed in the national-level indicator series while ensuring temporal coherence with the low frequency observations of the state target. This step aims at creating a benchmark time series to compare against.

A secondary missing data problem is that there is also no sufficiently long data on preferred related indicators. More specifically, there is a short time series for the state's economic indicators such as total number of non-farm employees with no state-specific total industry production index. This is predominantly the case for all the states for the early reporting years leading up to the 2000s. Naturally, this means that conducting temporal disaggregation using traditional related macroeconomic indicators at the state level remains a daunting task. To address this issue, as indicated above, this study adopts an adjusted framework that relies on national indicators exhibiting strong correlations with output at both national and state levels and uses these readily available national-level indicators, which show robust relationships with state-level output².

Preliminary evaluations reveal strong correlations between national and state-level output growth rates, as well as with national-level related indicators³. These observations are good in so far as supporting the approach of utilising national-level indicators in this case where sub-regional data are unavailable⁴. In addition, the selected national indicators, namely non-farm employees (PAYEMS) and total industry production index (IPI), are widely used in the literature as they often carry strong signals of short-term movements in GDP and are thus deemed suitable for my analysis⁵.

The national output growth time series strongly correlates with de-trended non-farm total employees data and the total industry production index (see Figure 3.1 and Table 3.1).

2. Based on the ESS guidelines (Buono et al. (2018)), candidate high frequency indicators should closely approximate the expected short-term movement of the target variable and show a good correlation with the original target variable when converted to the low frequency.

3. See correlation analysis for annual data between 1963 and 2004. This is also true for quarterly data for the official national and state gross domestic product between 2005 and 2022. Alaska has fairly distinctively low correlation with all regions of the USA

4. This is one of the most surpassing requirements in performing temporal disaggregation using a related indicator as others have argued, for example, Miralles et al. (2003); Chen et al. (2007); Chamberlin (2010); Armesto et al. (2010); Quilis (2018); Cuartas et al. (2020) and, more importantly, the European Statistical System (ESS) guidelines according to Buono et al. (2018) among others.

5. In literature, temporal disaggregation of gross domestic product in various countries using traditional methods has been performed with the help of some popular related indicators, namely, index of industrial production, exports data or employment data, for example, Islam (2009); Chamberlin (2010); Hall and McDermott (2011); Islam (2013), and Eurostat (2014) among others.

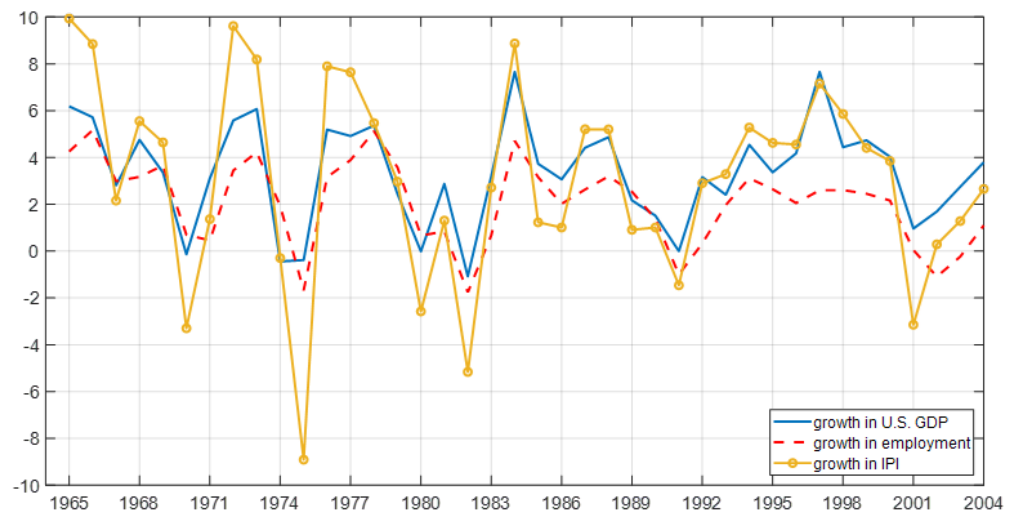


Figure 3.1: National Annual Output, Employment and IP Index

Notes: For the period between 1965 and 2004, this chart plots the historical national annual output, employment and IP index growth rates of the U.S. The variables show strong co-movements with synchronised turning points. See correlation coefficients in Table 3.2.

Table 3.1: Correlations between Output, Employment, and IP Index

	Growth in US GDP	Growth in Employment	Growth in IP Index
Growth in US GDP	1		
Growth in Employment	0.769	1	
Growth in IP Index	0.915	0.830	1

National output growth is also observed to be very strongly correlated with state-level output growth (see Figure 3.2). For the period 1965 to 2004, sub-trends in the national and state-level real GDP growth of annual time series show very strong and synchronised dynamics or cycles (see Table 3.2). Further assessments indicate that there is a strong positive association among the various state output dynamics, as reflected in the observed high correlation between the annual growth rates of the states and the states and the national output. This study exploits these strong associations observed and applies these variables in traditional procedures⁶.

6. The closest observed correlation is between national US output growth and growth in the total production index.

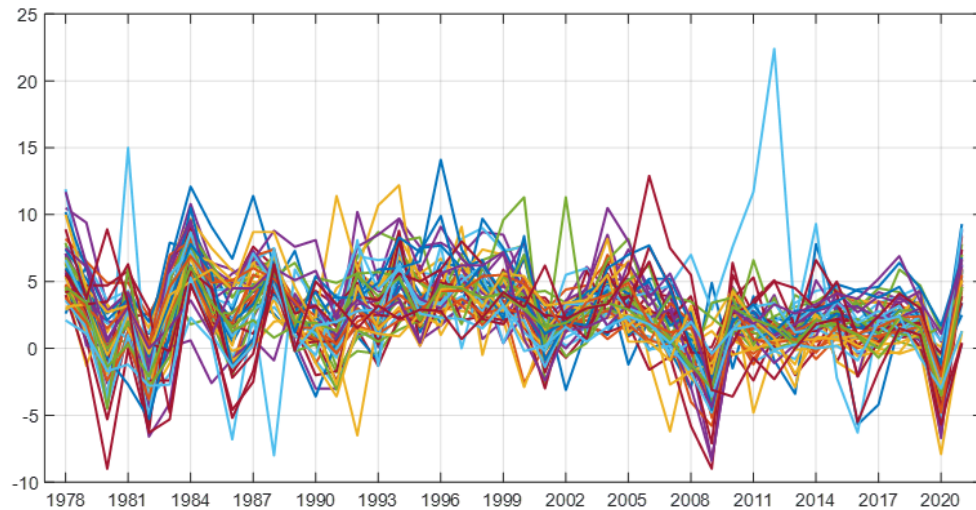


Figure 3.2: National and State Output Growth Rates

The historical national and 48 contiguous states' annual output growth rates show similar trends with strong co-movements.

Table 3.2: Correlations between National and State Output Growth

Region	United States	Alabama	Alaska	Arizona	Arkansas	California	...	Wyoming
United States	1							
Alabama	0.997	1						
Alaska	0.760	0.777	1					
Arizona	0.994	0.988	0.727	1				
Arkansas	0.992	0.996	0.775	0.986	1			
California	0.993	0.988	0.777	0.980	0.975	1		
...	...	...	...	...	...	...	...	
Wyoming	0.731	0.757	0.959	0.704	0.771	0.731	...	1

Notes: The table reports correlations between national output and individual state output. The correlations are presented for a few states because the results are very similar across all the states, thus making the rest of the table redundant - for brevity. The essence, therefore, is to reflect my key observation that there is a strong correlation between national and individual state output growth rates for the period between 1965 and 2004.

2.2 Modified Traditional Methods Vs the PLS Method

Initially, the study constructs historical estimates of quarterly state-level GDP growth time series for the U.S. using available state annual data⁷. To achieve this with traditional methods given the data problems identified earlier, I need to utilise a constrained temporal disaggregation framework for each state along with either national or state-specific related indicators.

7. Sax and Steiner (2013), Buono et al. (2018) and Jarmin (2019) all argue that although there is no substitute for official data, temporal disaggregation still is a useful workaround to provide missing high frequency time series data.

With regard to possible methods, Cuartas et al. (2020) highlight a diverse range of temporal disaggregation methods that are available for disaggregation of regional annual accounts and argue that the choice of method depends on factors such as available information, procedure, and required assumptions⁸. For instance, I know that traditional methods perform well when applied with a limited set of indicators, stable periods and sufficiently long time series. This is also supported in the temporal disaggregation model presentations outlined in the ESS guidelines, as discussed by Buono et al. (2018) among others. So, given the scarcity of extensive and relevant macroeconomic data at the state level, I explore the national indicators to inform state output dynamics when using traditional methods.

So, the study constructs historical output growth for individual state output growth using both national indicators and an extensive set of industry employment data specific to each state from 1991 to 2004 in the case of applying the PLS procedure. In particular, the state estimates are made using three models whose results are compared, namely, first, I estimate a traditional model for each state with national parameter estimates and a national-related indicator (indirect approach). Second, a traditional model for each state using a state-specific related indicator such as employment (direct approach), and finally, the PLS procedure using state-specific covariates (direct approach)⁹. In addition, the study illustrates the limitations associated with using proxy national indicators in informing the regional temporal disaggregation process in traditional methods¹⁰.

Related to the first modified framework, I compute the low frequency discrepancy by comparing a specific state's observed low frequency output growth rate with the low frequency predicted value - based on both national elasticities and related indicators. This strategy maintains the observed relationship between output growth and specific macroeconomic indicators at the national level while ensuring that estimates of high frequency state-level output growth rate series remain coherent with both the observed

8. For an extensive review of the temporal disaggregation methods, an interested reader may read Mazzi et al. (2005), Chen et al. (2007), Pavia-Miralles et al. (2010) and Quilis (2018) among others.

9. We have extensive state-specific industry employment data from 1990, for example, All employees: government, manufacturing, construction, durable goods, non-durable goods, wholesale trade, retail trade, financial activities, pharmaceuticals, etc.

10. For traditional methods, since information on relevant related high frequency indicators is not sufficiently long at the state level, this study uses U.S. national high frequency economic indicators to construct the missing state-level output growth time series.

annual state time series and short-term fluctuations in selected national indicators. This assumption is indeed strong and thus a key downside to estimating the missing high frequency state output growth. The approach imposes strict constraints on the temporal disaggregation process, limiting its applicability¹¹. However, recognising this limitation, the study then attempts to use state-specific single indicators. Furthermore, I also relax this constraint by taking an even more flexible approach, particularly when employing the PLS method with state-specific indicators.

Given the fundamental differences between the PLS procedure and traditional methods in their utilisation of information and underlying assumptions, this study embraces both frameworks as competitors to generate a wider range of high frequency state output growth time series. This approach enables me to harness the benefits of multiple methodologies and assess the resulting high frequency time series thus allowing me to select the most robust series for my final analyses.

To determine the best estimated time series, I evaluate the statistical properties of competing quarterly unconstrained temporally aggregated time series using the MAE and the adapted Discrepancy Aggregate Ratio (DAR), ensuring a thorough evaluation and selection process. In general, the use of national indicators in sub-regional analyses resonates with the evolving interest and literature in this field, exemplified by recent contributions from scholars like Pfeiffermann (2015), Bean (2016), Ferrara and Mazzi (2017), and Koop et al. (2020), who have explored the collection of official statistics and proposed timely methods to bridge data gaps and complement traditional survey-centric approaches¹².

2.3 Applying the Modified Model-Based Methods

In order to adapt the various traditional procedures, the temporal disaggregation problem and estimation strategy is now formulated as follows: Suppose that we have annual state-level gross domestic product growth over n years. Let y^i be a $(4n \times 1)$ vector of the unknown i^{th} state quarterly gross domestic product growth and x be a $(4n \times k)$ vector of

11. The main downside of this approach is the assumption that the slope between the indicator and target variables is the same for the national model and all the state models.

12. However, Koop et al. (2020) notes that official data is to be strictly preferred over predicted data whenever it is available.

national level quarterly covariates, for example, employment, IPI, etc. The implication of the assumption made above that governs the first experiment is that both national and state-level models are estimated using the same parameter estimates and predictor variables. This in literal terms is synonymous with using national-level information to estimate state-level output growth. From a practical standpoint, this particular specification is based on the assumption that all states react in the same manner to national economic signals. This means that they all share a same set of elasticities, but they differ in the amount and level of their unobserved residuals. This is the biggest downside of this approach.

For the current application $k = 2$, which means state models include a constant and one indicator series. The mathematical and compact formulation of the low frequency national and state regression models may be represented as taking the form in Table 3.3.

Table 3.3: National and State Low Frequency Models

Low Frequency National Model
$Y_t = X\beta + U$
Low Frequency State Model
$Y_t^i = X\beta + \epsilon_t^i$
where
$t = 1, \dots, T$ runs at annual frequency.
$i = 1, 2, \dots, 48$ denotes the i^{th} state.
Y_t is USA national output growth in period t .
Y_t^i is output growth for the i^{th} state in period t .
X is a matrix of national low frequency indicators, employment or IPI or both.
U and ϵ_t^i are vectors of disturbances of the national and state-specific models, respectively.
In the state models, β represents elasticities of the national model.

In this regard, this study modifies the two-step regression model-based framework in two ways: First, for each state, a preliminary low frequency series estimated using national-level indicators must be determined. Second, using the low frequency state model estimates, calculate the differences between the observed low frequency series and the corresponding low frequency values of the preliminary series for each state and distribute the discrepancy among the preliminary high-frequency series in order to determine the final high frequency state time series.

Table 3.4: Framework for Estimating State High Frequency Output

Step	Equation & Definitions
(i) National low frequency model	$Y_t = X_t \hat{\beta} + U_t$ • Y_t: observed U.S. annual output growth in year t. • X_t: $1 \times K$ vector of national annual indicators. • $\hat{\beta}$: estimated elasticity vector. • U_t: residual capturing unexplained annual variation.
(ii) State preliminary high frequency estimates	$p_q^i = x_q \hat{\beta}, \quad \hat{Y}_t^i = \sum_{q=1}^4 p_q^i = C p^i$ • x_q: $1 \times K$ vector of quarterly national indicators for quarter q. • p_q^i: preliminary quarterly output for state i. • $p^i = (p_1^i, \dots, p_4^i)'$. • C: 1×4 aggregation row vector of ones (i.e. sums four quarters). • $\hat{Y}_t^i$: aggregated annual estimate for state i.
(iii) Annual state discrepancy	$U_t^i = Y_t^i - \hat{Y}_t^i$ • Y_t^i: observed annual output growth for state i. • U_t^i: annual residual for state i, to be re-distributed.
(iv) Final state high frequency model	$y_q^i = p_q^i + D_q U_t^i$ • y_q^i: final quarterly output growth for state i. • D is a $n \times n_l$ distribution matrix, with n and n_l denoting, the number of high frequency and low frequency observations, respectively.

In general, from Table 3.4, it becomes clear that the sum of the preliminary state high frequency series, p and the distributed state-specific low frequency discrepancy, DU^i yield the final estimate for the i^{th} state of the target high frequency series, $\hat{y}^i$. Formally expressed as,

$$\hat{y}_q^i = p_q^i + D_q U^i \quad (3.1)$$

Where

p denotes the predicted or preliminary series, $x_q \hat{\beta}$.

D is a $n \times n_l$ distribution matrix, with n and n_l denoting for example, the number of high frequency (quarterly) and low frequency (annual) observations, respectively.

This is a general expression, and it is easily representative of other conversions such as converting quarterly to monthly series. U^i is a vector of length n_l and contains the differences between the aggregate values of p and the actual low frequency state values, Y^i or discrepancy:

$$U^i \equiv Y^i - \hat{Y}^i \quad (3.2)$$

Where

$$\hat{Y}^i = C p^i$$

Multiplying the $n_l \times n$ conversion matrix, C , with a high frequency series performs aggregation. Equation 3.1 constitutes a unifying framework for all disaggregation methods. The models used in this study technically differ in how they estimate the preliminary series, p .

3 Constructing Historical Data

This section discusses the challenging task of constructing and selecting meaningful estimates of historical quarterly state-level output growth series for the U.S by state. Whether it is using state-specific covariates with the PLS approach or traditional methods with national and a state-specific high frequency indicator, I am aware that each approach has its own set of limitations or costs. As a result, I use the approach that is the most effective among them in order to generate the final data set for subsequent analyses¹³.

3.1 GDP and the Related Indicators

For the inputs into the modified traditional models, the study relies on data obtained from multiple sources, primarily the Bureau of Economic Analysis (BEA) Interactive Tables. Here, both national and state-level All Industry Total in Chained Dollars Real GDP data is obtained. The time series by state spans the period from 1990 to 2004 and corresponds with the available high dimensional state-specific high frequency indicators¹⁴. Among the available key high frequency indicators is employment data, which is accessible starting from 1939.

Specifically, I download two covariates. First, the PAYEMS series represents All employees, Total Non-farm, Thousands of Persons, and Seasonally Adjusted. Second, since the study incorporates an alternative covariate, I also download the industrial production index, denoted INDPRO, which is also available since 1939. This index reflects the Industrial Production: Total Index, adjusted to the base year of 2017 = 100 and is also Seasonally Adjusted¹⁵.

13. See Appendix B for a brief review of alternative econometric methods for backcasting.

14. Annual series, SAGDP9 is the percentage change from the preceding period while quarterly series, SQGDP9 is the percentage change from the preceding period (annualised)

15. The study also considers state-specific unemployment rate, Percent, Monthly, Seasonally Adjusted, which is available since 1976. This variable is also used as an indicator in the temporal disaggregation procedures, but the results do not improve, so the results are not reported in this report because the tables were too large with little value addition.

In order to ensure consistency and comparability across the datasets, all utilised data series are transformed into growth rates. This process involves synchronising the data by removing the initial observation from each series. Furthermore, for the PLS framework, supplementary data is sourced from the Economic Research Federal Reserve Bank of St. Louis website¹⁶. This state-specific data set comprises a range of employment and other economic variables, including occupational titles and the corresponding employment figures for each occupation (see Appendix Table 3.6). An assessment using state-specific unemployment rates as indicators in the traditional temporal disaggregation procedures did not yield meaningful results.

3.2 Accuracy Assessment Strategy and Measures

3.2.1 Strategy of Accuracy Assessment

In this subsection, the study discusses the strategy for assessing the quality of estimates based on the method and related variables utilised in estimations. I want to understand how close the aggregated estimates of each model are to the observed low frequency data. Ultimately, this helps to ensure that the estimated historical high frequency state-specific economic growth estimates have the smallest discrepancy and are obtained using the best approach.

I follow a two-stage quality assessment which is focused initially, on the procedures and on the estimated historical data. The first stage evaluates the information content in national indicators and the methods used to disaggregate state output growth. The objective is to determine whether a given procedure-variable combination used for predicting a time series delivers reliable results for the subsequent analyses¹⁷. In this way I am able to assess the viability of using the indirect methods, leveraging national level information to disaggregate state-level annual values before committing to its implementation. This assessment utilises the “Discrepancy Aggregate Ratio (DAR)”. In the second stage, the

16. See <https://www.stlouisfed.org/>

17. Procedure-variable combination means, for example, Chow-Lin method with employment, Fernandez method with IPI, SSC method with employment, etc.

study assesses the accuracy of the resultant time series for each procedure-variable combination using two metrics: Mean Absolute Error (MAE) and Normalised Disaggregation Error (NDE). These metrics provide an evaluation of the precision and reliability of the disaggregation methods applied to state output growth data.

3.2.2 Discrepancy Aggregate Ratio

So, how much information content do I retain for each state arising from using both national parameter estimates and indicators to predict the state output growth rates? To directly answer this question, this paper introduces a metric called the DAR¹⁸. Indeed, I observed strong correlations between the target national and state outputs and related indicator variables. However, this is not enough as it is only good for deciding related variables to work with. In order to determine how well the “national-to-state” models are working, the study assesses the extent of information loss incurred when utilising national indicators to estimate quarterly output growth for individual states.

This evaluation is important because my strict preference is to have a high percentage of the state target explained by the national features and therefore, at least maintain the same integrity of information captured at the national level while deriving state-level estimates. In this approach, this is expected to ensure reliability of the resulting high frequency state-level time series. Calculating the DAR helps to quantify the information loss as follows; the proportion of the observed that is predicted by respective models. This helps to gauge the fidelity of the estimation process. The evaluation is fundamental to ascertain the trustworthiness of the derived high frequency state-level time series, which is vital for subsequent analytical exercises.

The DAR is an adaptation of the principle construct of the Noise Aggregate Ratio (NAR) to my current and specific temporal disaggregation problem. Therefore, for this purpose, the study formulates the DAR as follows:

18. The DAR is constructed based on the noise aggregate ratio, NAR because the logic is similar. In Klemenjak et al. (2021), NAR was used as a noise measure utilised to determine the noise levels of real-world aggregate signals found in datasets. A direct reference states as follows: The ratio can be computed for any type of power signal, provided that readings of the aggregate and individual appliances are available. In this way, the NAR quantifies the residual term’s share in an aggregate signal. It was first utilised in 2015 by Makonin et al. (2015).

$$DAR = \frac{\sum_{t=1}^T U^i}{\sum_{t=1}^T Y_t^i} = \frac{\sum_{t=1}^T |Y_t^i - Cp_t^i|}{\sum_{t=1}^T Y_t^i} \quad (3.3)$$

or

$$\frac{\sum_{t=1}^T |Y_t^i - \hat{Y}_t^i|}{\sum_{t=1}^T Y_t^i}$$

where

$$\hat{Y}_t^i = Cp_t^i$$

and

$$p_t^i = x\hat{\beta}$$

$$C = (1, \dots, 1)$$

is the aggregation matrix for the predicted quarterly time series.

The study reports $(1 - DAR) \times 100$ so that I have a direct and positive interpretation of the metric. The reported results indicate the percent of the observed value that can be attributed to the predicted component of the model while the rest is attributed to the residual term.

When I compare traditional methods, the results of the DAR for the “benchmark” national models indicate that Chowlin using PAYEMS has the smallest aggregate discrepancy error with 73.3% of the observed annual value attributable to the aggregate predicted component of the model. The Litterman method with the IPI is the least-performing procedure-variable combination at the national level. The DAR average for all methods for the national models is about 64.6%. Further, between the national models and the state models, it is apparent that there is a huge loss in performance of the procedure-variable combinations across all states with averages of 50% in the majority of cases.

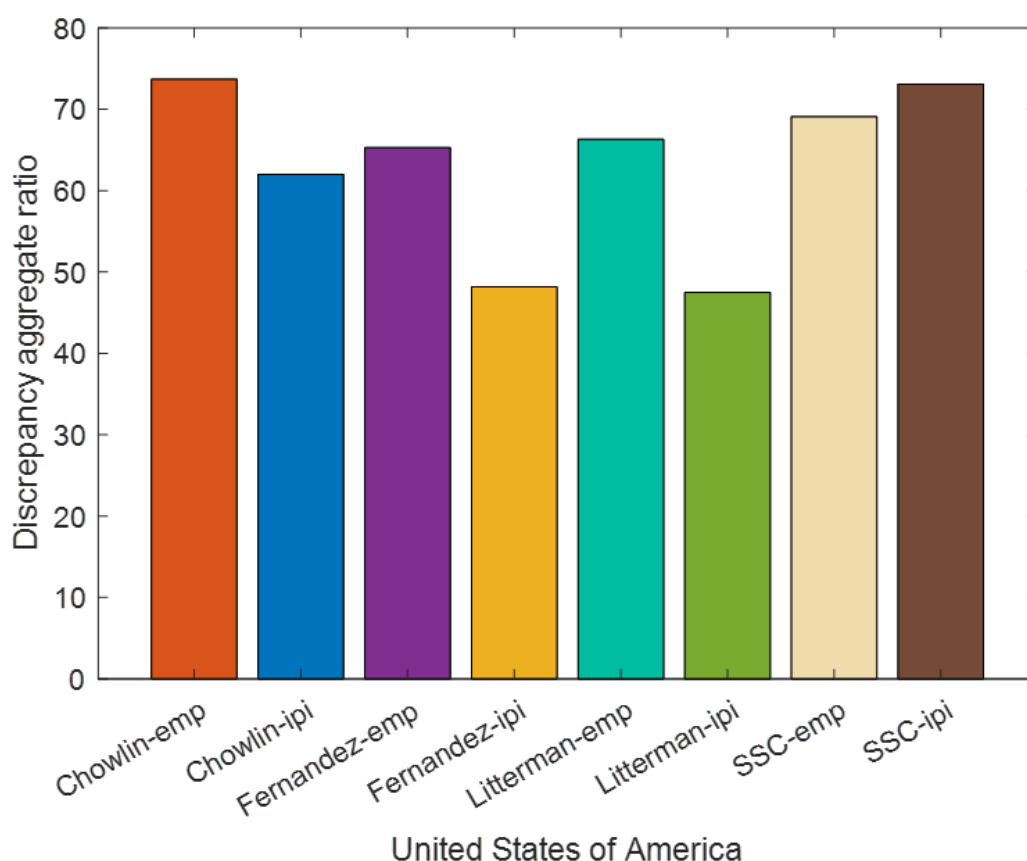


Figure 3.3: Discrepancy Aggregate Ratio

Notes: This chart shows the discrepancy aggregate ratio represented as $(1 - DAR) \times 100$ for the national models of each traditional method. The chart shows that the best performing models explain just over 70% of the observed value. In addition, a method pairwise comparison by indicators shows that national employment has stronger signals relative to the IPI. An exception is SSC. Similar analyses for individual states indicate massive loss of information with the majority of the ratios falling below 50%.

However, the information loss is not surprising, as parameters in the state models are only proxies, and the indicators too, are not state-specific but national. This observation also underscores the inherent risk of drawing conclusions at the national level and extrapolating the same to individual states. This is because aggregate national predictors jointly with national parameters fail to significantly explain state-specific economic outcomes. The losses are so great particularly in some states, indicating considerable failure of this indirect approach, for example, Alaska, District of Columbia, Hawaii, Louisiana and West Virginia. Surprisingly, I get even worse results in most states when I employ state-specific unemployment rates as single related indicators in traditional methods. The DAR for the national models is shown in Figure 3.3 while the individual state DARs are reported in Appendix Figure 3.9.

3.2.3 Estimated Time Series Accuracy Evaluation

Having determined the information content of the procedure-indicator above, I adopt two distance measures to assess the performance of each procedure-indicator combination based on the resulting time series. The first is a well-known and popular metric used in the econometric literature, the MAE, where smaller values are preferred as they indicate a good fit, as discussed in Chapter 2.

The second accuracy measure is defined by Non-Intrusive Load Monitoring (NILM) or load disaggregation scholars, for example, Kolter and Jaakkola (2012) and Klemenjak et al. (2021), popular in effective energy management studies, as the Normalised Disaggregation Error (NDE), defined as:

$$NDE = \sqrt{\frac{\sum_{t=1}^T (Y_t^i - \hat{Y}_t^i)^2}{\sum_{t=1}^T (Y_t^i)^2}}$$

Unlike MAE, NDE belongs to the class of normalised metrics and is therefore a dimensionless metric. As such, it will allow me to make fair comparisons of procedure-indicator performance between different states. Similarly to the MAE, smaller values of the NDE are preferred.

The results of both the MAE and the NDE indicate that the best method and indicator combination varies across states. However, consistent with the DAR above from these accuracy metrics, national employment as an indicator has relatively strong signals across all methods. This observation along with data availability forms part of my key considerations to use a wide range of state-specific industry employment indicators in a high dimensional time series model to perform temporal disaggregation in the data-driven predictive approach. Furthermore, the accuracy measures indicate that the PLS achieves the smallest MAE and NDE compared to all other model outcomes (see the MAE results in Appendix Table 3.7, while the NDE table has been omitted for brevity).

3.2.4 Conformity of Estimates to Actual Data

In this section, I evaluate how well the predicted state-level output growth figures conform to the observed data. This evaluation is carried out by method and indicator variable in the absence of satisfying the temporal aggregation constraint. In particular, annual output growth is first disaggregated into quarterly series. The estimated quarterly values are aggregated and then compared with the actual annual data. The estimates indicate that aggregated time series from the PLS procedure strongly align with the annual state-level output growth rates in all cases. The conformity charts for each state are reported in Appendix Figure 3.10.

Based on the conformity evaluation, I also show the evolution of the PLS estimated quarterly output growth by state. It is easy to see that the quarterly time series results from the PLS method also co-move with the actual annual data fairly well. Appendix Figure 3.11 reports quarterly estimates from the PLS method superimposed on actual annual data for the years 1991 to 2004. The features confirm how closely the PLS-derived high-frequency estimates co-move with the official annual data.

In general, the PLS approach produces the most reliable and practical high-frequency output growth series for the years 1991 to 2004 when combined with state-specific predictor variables, according to accuracy metrics. As a result, quarterly output series produced by the PLS technique are used in the application sections that follow.

In summary, results of all accuracy measures suggest that a relatively more accurate and usable time series is the one estimated by the PLS procedure with state-specific predictor variables for the period between 1991 and 2004. Accordingly, I use high frequency output time series estimated using the PLS procedure in the application sections that follow.

4 Differential State Effects of Financial Disruptions

4.1 Background and Motivation

In this section, I would like to understand the effect of an aggregate financial disruption on individual state real economic activity. This is because the financial accelerator effects might be stronger in some regions than others, meaning that shocks to asset values tighten borrowing restrictions and amplify downturns differently because of the different underlying financial and economic structures. I study this problem along the lines of the framework developed and applied by Gilchrist and Zakrajšek (2012a) (GZ hereafter) in which a similar problem is studied at an aggregate level of the United States¹⁹. Although their study is at the national level, it provides a lot of useful insights for macroeconomic policy. However, a key critique of the approach is that it essentially assumes that the effect of a financial shock is identical in all the states. As such, this analytical approach fails to recognise the differences in regional business cycles that can be attributed to the differences in regional characteristics, such as the industrial composition.

In macroeconomic management, relying solely on an aggregate approach can be misleading when assessing the impact of financial shocks on individual states. This can potentially lead to sub-optimal policy responses from local authorities. Moreover, earlier research such as Carlino and DeFina (1998) validates the individualised approach by showing that monetary policy changes affect core regions, such as New England, the Mideast, Plains, Southeast, and the Far West—differently than the three non-core regions: the Great Lakes, Southwest, and the Rocky Mountains. Therefore, this analysis seeks to adopt a more nuanced approach by examining the effects of a common aggregate financial shock

19. An analysis and approach similar to mine is one by Carlino and DeFina (1998) who examined how the actions of monetary policy differentially affect regions and economic activity of each state. They estimated an eleven-variable structural VAR for each state with economic activity being measured by the state's real personal income growth and employment growth. Personal income focuses on the income received by individuals and households whereas real output growth focuses on the total economic output and production. So, although similar, there are three distinct points of departure which include measure of economic activity, type of shock and economic model.

on the economies of individual states by utilising a state-specific target broad indicator, namely, real output growth by state. This approach will arguably provide granular insight into the varied regional responses to a financial shock and facilitate more tailored policy implications.

4.2 Finance and the Real Economy

To start with, the heterogeneous nature of the states means that I have to modify and adapt the main arguments of the transmission channels through which disruptions in financial markets influence individual state output outcomes. This is done along the lines of GZ as follows:

- i. A pullback in spending owing to reductions in wealth: State variation means that States with relatively higher consumption patterns may feel these effects more acutely.*
- ii. Balance sheet mechanisms lead to a widening of credit spreads, which curtails the ability of households and businesses to obtain credit. State variation means that States with higher levels of debt or more interest rate-sensitive sectors (like housing and durable goods) may experience more significant impacts.*
- iii. The direct effect of impairments in the ability of financial institutions to intermediate credit. State variations means that states with significant financial sectors or large banking industries may see more pronounced impacts.*

A detailed exploration of the theoretical benchmark macroeconomic model and empirical literature on the relationship between finance and the real economy can be found in the Gilchrist and Zakrajšek (2012b). Additional empirical support for this type of mechanism is provided in Gilchrist and Zakrajšek (2012a) in which they show that increases in risk aversion due to financial disruptions lead to contraction in credit supply, both through the corporate bond market and broader commercial banking sector and ultimately, economic activity.

The related literature on the differentiated regional effects of monetary shocks identifies three primary reasons for these variations. First, regional differences in sectoral composition play a significant role. Certain sectors, such as manufacturing and construction, are highly sensitive to interest rate changes. When interest rates rise due to a monetary shock, these sectors experience significant slowdowns as higher borrowing costs deter investment and spending. Regions with a higher concentration of these sectors are more adversely affected by monetary shocks, as demonstrated by Carlino and DeFina (1998).

Second, the impact of monetary shocks also depends on the size distribution of firms within a region. Smaller firms typically face higher costs of external finance compared to larger firms. During periods of tight monetary policy, borrowing costs increase more sharply for small firms, making it harder for them to access credit. Regions with a higher concentration of small firms are therefore more vulnerable to monetary shocks, as argued by Gertler and Gilchrist (1993) and Oliner, Rudebusch et al. (1995).

Lastly, the ability of banks to adapt their balance sheets in response to monetary policy changes affects regional resilience. Banks in some regions may be better equipped to find alternative sources of funding for deposits and loans when monetary policy tightens. This ability allows them to continue lending even in a restrictive monetary environment, thereby mitigating the adverse effects of monetary shocks on the regional economy, as highlighted by Kashyap and Stein (1995) and Carlino and DeFina (1998).

4.3 Financial Shock and Individual State Macroeconomy

Addressing the differential effects of a financial disruption on state output is a challenge especially due to the lack of data at the preferred quarterly level. The official state quarterly real output time series only begins in 2005 Q1²⁰. The relatively short time series may not produce reliable and accurate estimates in such an econometric analysis. In principle, relying on such limited data introduces several risks, including overfitting, imprecise estimates, and reduced degrees of freedom due to the high number of parameter estimates required. Additional challenges include selecting appropriate lag lengths

20. See <https://www.bea.gov/itable/> for quarterly GDP by state time series.

and ensuring model stability. Overfitting, in particular, can result in models that capture noise rather than the underlying relationships, leading to erratic and unreliable impulse response functions, which are not suitable for policy analysis or decision-making (Lütkepohl (2005)).

So, to abate the short time series problem, I use the state specific historical quarterly data constructed in the previous section and the official state quarterly output data, concatenate the two time series datasets for each state (i.e estimated 1991Q1 – 2004Q4 + Official 2005Q1 – 2019Q4), full sample, hereafter and use a vector autoregression (VAR) framework to trace the dynamic effect of an unexpected financial shock on key macroeconomic and financial variables.

In terms of identifying a common aggregate financial distress variable, I use GZ's indicator of financial market distress whose construction is based on secondary market prices of bonds issued by U.S financial institutions (Gilchrist and Zakrajšek (2012b)). From the flexible empirical credit-spread pricing framework, GZ decompose credit spreads into a component identifying counter-cyclical movements in expected defaults and a component that represents the cyclical changes in the relationship between default risk and credit spreads. The latter component is what they refer to as the financial bond premium (EBP) which they show that it represents the shifts in the risk attitudes of financial intermediaries. An increase in the EBP (positive financial shock) means an increase in financial market distress.

Based on the evidence provided by the base study, for the purposes of this study (see Gilchrist and Zakrajšek (ibid.)), I also assume that the three transmission channels identified above link changes in financial conditions and the individual state output outcomes.

4.4 Empirical Approach: State-Level Economic Implications

Having discussed the state output data, the link between finance and the real economy (transmission channels) and the current research blind spot on the implications of a financial shock on individual state real economy, I now turn my attention to the empirical examination of the effect of a financial shock (as measured by the EBP) on the individual state output. To do this I estimate the following standard VAR with quarterly variables:

The Economic Model

We assume an economic model with the following eight endogenous variables for each state as follows:

$$Y_t = \begin{bmatrix} PCE_t \\ INV_t \\ GDP_t \\ INFL_t \\ EBP_t \\ SP500_t \\ GS10_t \\ FFR_t \end{bmatrix}$$

Each state's model includes state-specific output (GDP_t^{State}) alongside aggregate economic variables:

$$\text{Alabama: } Y_t^{\text{Alabama}} = \begin{bmatrix} PCE_t \\ INV_t \\ GDP_t^{\text{Alabama}} \\ INFL_t \\ EBP_t \\ SP500_t \\ GS10_t \\ FFR_t \end{bmatrix}$$

$$\text{Arizona: } Y_t^{\text{Arizona}} = \begin{bmatrix} PCE_t \\ INV_t \\ GDP_t^{\text{Arizona}} \\ INFL_t \\ EBP_t \\ SP500_t \\ GS10_t \\ FFR_t \end{bmatrix}$$

$$\vdots$$

$$\text{Wyoming: } Y_t^{\text{Wyoming}} = \begin{bmatrix} PCE_t \\ INV_t \\ GDP_t^{\text{Wyoming}} \\ INFL_t \\ EBP_t \\ SP500_t \\ GS10_t \\ FFR_t \end{bmatrix}$$

The VAR representation is given by

$$Y_t = AY_{t-1} + \varepsilon_t$$

Where Y_t is a vector of endogenous variables, A is an $N \times N$ matrix holding the autoregressive coefficients of the model and ε_t is a vector white noise error term. Therefore, for each state, I estimate an econometric model VAR with 2 lags:

$$Y_t^i = A_1^i Y_{t-1}^i + A_2^i Y_{t-2}^i + \varepsilon_t^i, \quad i \in \{\text{Alabama, Arizona, ... , Wyoming}\}$$

So for each state i :

$$Y_t^i = \begin{bmatrix} PCE_t^i \\ INV_t^i \\ GDP_t^i \\ INFL_t^i \\ EBP_t^i \\ SP500_t^i \\ GS10_t^i \\ FFR_t^i \end{bmatrix}$$

$$A_1^i = \begin{bmatrix} a_{11}^{(1),i} & a_{12}^{(1),i} & \cdots & a_{18}^{(1),i} \\ a_{21}^{(1),i} & a_{22}^{(1),i} & \cdots & a_{28}^{(1),i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{81}^{(1),i} & a_{82}^{(1),i} & \cdots & a_{88}^{(1),i} \end{bmatrix}, \quad A_2^i = \begin{bmatrix} a_{11}^{(2),i} & a_{12}^{(2),i} & \cdots & a_{18}^{(2),i} \\ a_{21}^{(2),i} & a_{22}^{(2),i} & \cdots & a_{28}^{(2),i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{81}^{(2),i} & a_{82}^{(2),i} & \cdots & a_{88}^{(2),i} \end{bmatrix}$$

$$\boldsymbol{\varepsilon}_t^i = \begin{bmatrix} \varepsilon_{PCE,t}^i \\ \varepsilon_{INV,t}^i \\ \varepsilon_{GDP,t}^i \\ \varepsilon_{INFL,t}^i \\ \varepsilon_{EBP,t}^i \\ \varepsilon_{SP500,t}^i \\ \varepsilon_{GS10,t}^i \\ \varepsilon_{FFR,t}^i \end{bmatrix}$$

Where:

1. Y_t^i is the vector of endogenous variables for state i .
2. A_1^i and A_2^i are the coefficient matrices for lag 1 and lag 2 for state i .
3. $\boldsymbol{\varepsilon}_t^i$ is the vector of error terms for state i .

EBP time series are available on a monthly basis and are therefore, converted to quarterly while the SP500 is converted to quarterly and first difference. Others are log-difference of consumer price index (CPI), log-difference of real business fixed investment (INV), log-difference of real personal consumption expenditure (PCE), state output growth (GDP) is full sample (no transformation required). No transformation is required for the effective (nominal) federal funds rate (FFR) and the 10-year (nominal) Treasury yield (GS10).

In order to identify the structural shocks in the VAR model above, I assume a Cholesky identification strategy in which the variance-covariance matrix Σ of the reduced-form residuals $\boldsymbol{\varepsilon}_t$ results in a lower triangular matrix L such that $\Sigma = LL'$. This recursive ordering matters to ensure a sensible structure so that $PCE_t, INV_t, GDP_t, INFL_t$ reacts to the financial shock with a lag while it contemporaneously affects $SP500_t, GS10_t, FFR_t$. Table 3.5 reports the data and its sources.

Table 3.5: VAR Endogenous Variables

Full Variable [Model Name][Frequency]	Source
State GDP [GDP][Q]	<i>Estimated</i> + BEA
Personal Consumption Expenditures [PCE][Q]	FRED Economic Data
Consumer Price Index [INFL][Q]	FRED Economic Data
Investment [INV][Q]	FRED Economic Data
Treasury Securities at 10-Year Const. Maturity [GS10][Q]	FRED Economic Data
Federal Funds Rate [FFR][Q]	FRED Economic Data
S&P 500 [S&P 500][M]	http://www.econ.yale.edu/~shiller/data.htm
Excess Bond Premium [EBP][M]	http://www.federalreserve.gov/econres

Notes: The table reports variables and data sources for the VAR models.

4.5 VAR Empirical Results

I estimate a VAR with two lags, similar to GZ, and find that the shape and behaviour of impulse response functions of all endogenous variables are consistent with those observed of the U.S. average response in the original study depicted in Figure 3.4. The state output IRFs are presented in Figure 3.5. This observation is broadly similar to Carlino and DeFina (1998) findings that the estimated state responses all took the same general profile found using the regional data.

Results indicate that a positive shock to the EBP leads to a reduction in real economic activity (for the majority of states), consumption, investment and prices²¹. However, the response of individual state outputs to a financial shock is somewhat different across the different states. In addition, this empirical result despite being at a micro level is consistent with the work of Gilchrist and Zakrajšek (2012a) and Gilchrist and Zakrajšek (2012b) that found that an increase in the excess bond premium causes a drop in asset prices and a contraction in economic activity through the financial accelerator mechanisms for the U.S.²². The amount of variation in the endogenous variables explained by the orthogonalised shocks to the excess bond premium using aggregate data are shown in Figure 3.4.

21. A price puzzle emerges when I conduct this analysis on shorter samples. Sensible and consistent results were only observed using the full sample.

22. The financial accelerator mechanism highlights the interplay between financial markets and the real economy, demonstrating how credit conditions can amplify economic fluctuations.

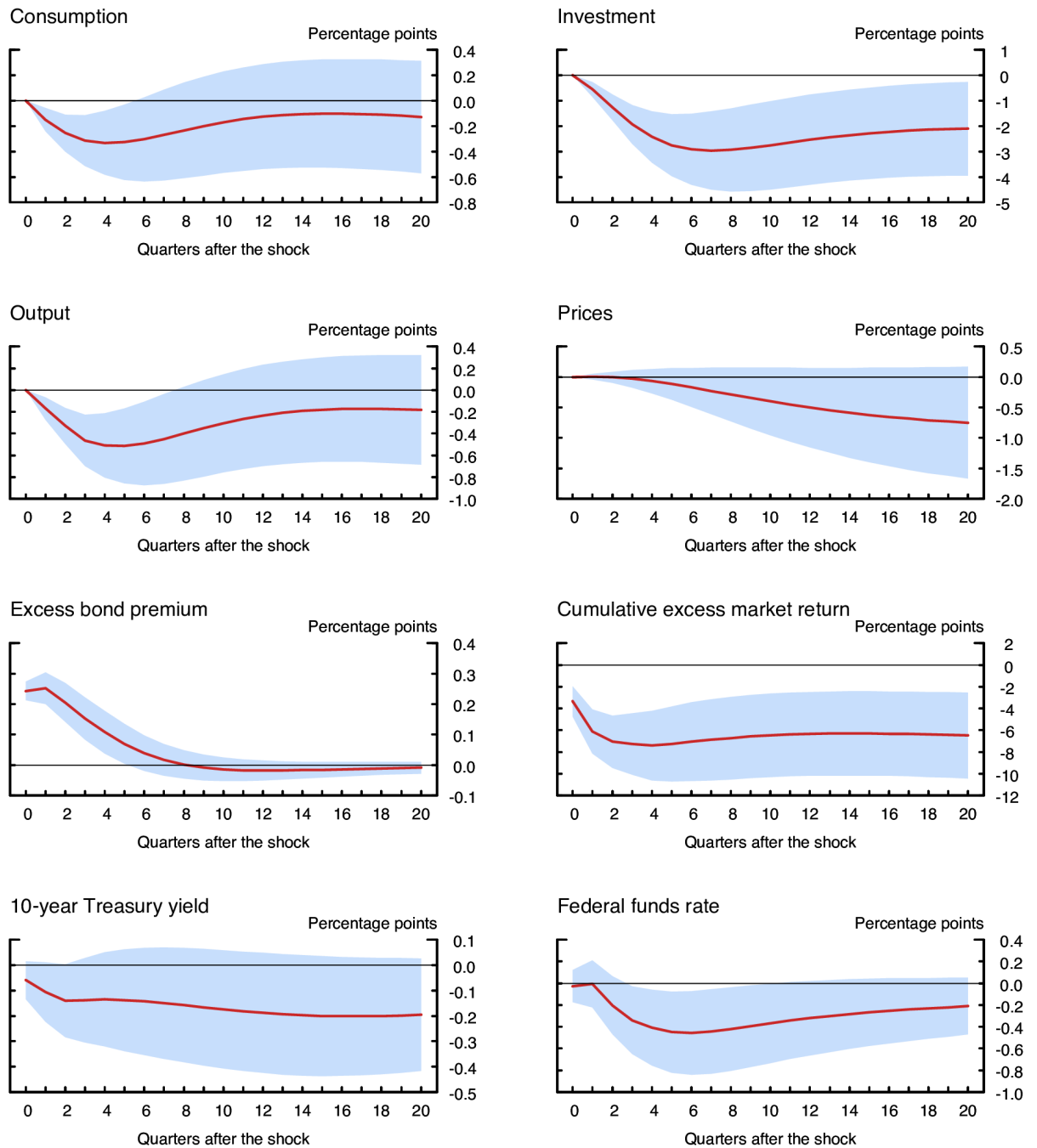


Figure 3.4: National Impulse Response Functions

Notes: Macroeconomic implications of a shock to the excess bond premium. I have estimated a VAR model for each of the 48 contiguous states which has resulted in a large set of impulse response functions (IRFs) and because of space considerations, I refer to the results of the original study because the observed results of the aggregate macroeconomic variables for each state are broadly similar. The difference, however, is notable in the response of state-specific output to a financial shock whose output IRFs I have reported in the Heat Map.

Source: Gilchrist and Zakrajšek (2012a)

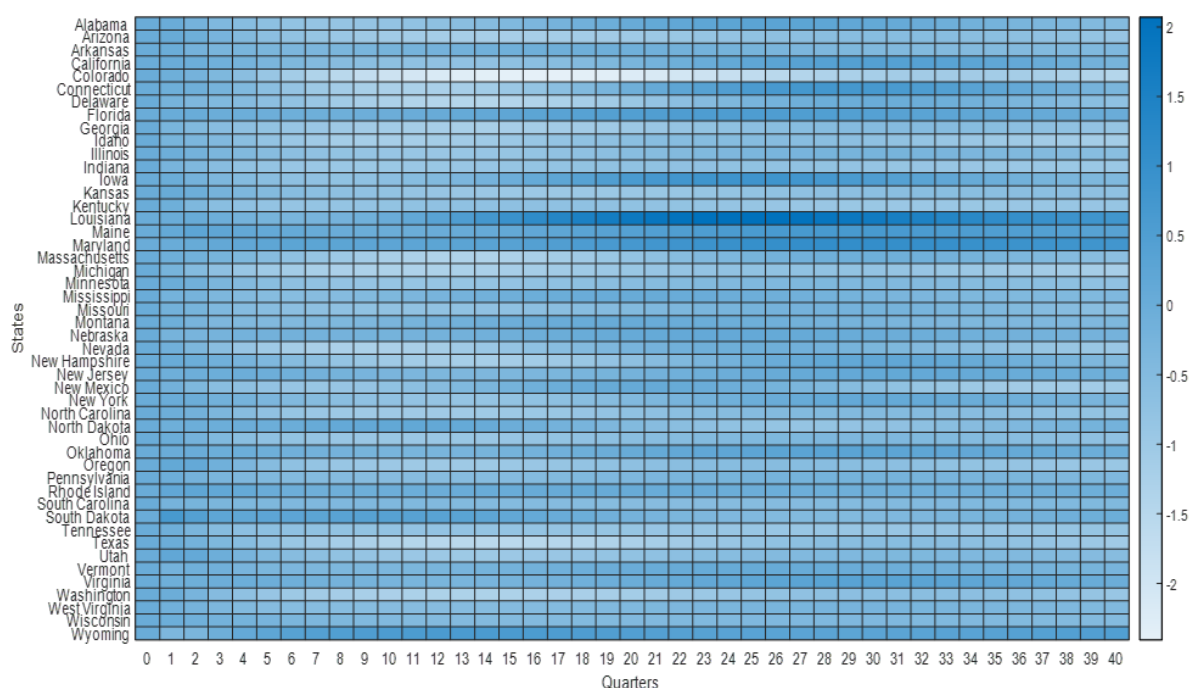


Figure 3.5: State Impulse Response Functions

Source: Author's Computations. Notes: This heatmap shows the implications of a financial shock on state output. This chart represents only the response of individual state output to a positive shock to the Excess Bond Premium (EBP). The response of the majority of the states is a reduction in real economic activity while some remain more or less unchanged. In general terms, I note that the states' economic fabric differs quite a lot also, for example, Arkansas, South Dakota, Rhode Island and New Jersey have diversified economies in agriculture, retail and services. They are further supported by natural resources. Florida has a diverse economy with strong tourism, agriculture, and aerospace sectors. Maine is diversified through tourism, agriculture, fishing, forestry and paper. Maryland benefits from proximity to Washington, D.C., with a strong presence of federal government jobs, healthcare, and biotechnology sectors. Virginia also has a significant number of federal government and defense-related jobs, along with a growing tech sector. Louisiana and North Dakota have an oil and gas industry, along with significant agricultural and shipping sectors. Finally, Wyoming is rich in natural resources such as coal, natural gas, and oil.

The heatmap in Figure 3.5 illustrates the impulse responses of quarterly state-level GDP growth for each of the 48 contiguous states (along the vertical axis) to a common financial disruption shock over a 40-quarter horizon. While a pale (nearly white) colour suggests a negative reaction, a deep blue colour indicates a positive deviation in the opposite direction from the trend growth.

During the first few quarters after the financial shock, majority of the states suffer a contraction in output growth that is visible with nearly white shade. This contraction typically occurs in the early quarters. After that, the majority of states begin to recover by the eighteenth quarter, the colours change to medium blues (around zero), which indicates a return to the growth trend. The effects finally fade after the twenty-first quarter; by the thirty-first quarter, the shading in virtually every state returns to a level that is close to zero.

With regards heterogeneity, despite the fact that the overall pattern of a mid-horizon recovery following an initial drop is usual, the amplitude of that rebound can vary quite a bit. Louisiana, Maryland, Massachusetts, and Maine are among the states that have some of the most unresponsive economic activity to the financial disruption, which suggests that these states are not susceptible to this particular financial shock. The output IRFs for these states are not statistically different from zero following a financial shock. The investigation of the reasons why these states are resilient is, however, beyond the scope of the current study. On the other hand, states that are large in terms of manufacturing, such as Indiana and Michigan, as well as those that export energy, such as Texas, have responses that are more pronounced.

The geographical trends that emerge suggest that states in the Midwest and Plains, such as Iowa, Nebraska, and North Dakota, have near zero response, states in the Northeastern region, such as Massachusetts and Maine, and some economies in the Mid-Atlantic region, such as Maryland and New Jersey, have similar outcomes. The southern states are tending to be more vulnerable.

The observations may suggest differences in the underlying general characteristics of individual states, such as unique economic structure, demographic characteristics, and policies that contribute to their ability to withstand financial disruptions as insinuated by Carlino and DeFina (1998). The implications of the heterogeneity of the state responses means that it is not possible to have a policy response that is uniform to a national financial shock. While it is possible that some states would fall behind and require targeted assistance, it is also possible that other states will recover strongly and be able to endure a firmer monetary stance. Furthermore, the shock's period of around two to four years highlights the need of medium-term planning in fiscal and regulatory policy.

5 Network of the U.S. Real Economic Activity

5.1 Background and Motivation

In this section, the study builds on the previous analysis by examining the real output growth network of the United States. The main aim is to identify states whose business cycles are interdependent and co-move closely over time. This application, albeit with differing research questions, is in the lines of Hamilton and Owyang (2012), Leiva-Leon (2017) and Brownlees et al. (2022) studies on business cycle synchronisation in the United States²³. Therefore, although the previous studies in the literature provide useful insights into the overall synchronisation pattern in a given sample period, the main critique is that none use real output growth time series due to its unavailability at high frequency by state. In particular, they use proxy measures of economic activity, for example, both for Hamilton and Owyang (2012) and Brownlees et al. (2022) use the same seasonally adjusted, annualised quarter-on-quarter growth rates of payroll employment data and Leiva-Leon (2017) uses the monthly coincidence indicator of overall economic activity by state²⁴.

23. For example, Hamilton and Owyang (2012) proposed a cross-validation procedure that estimated the number of communities to be three. The states with similar business cycle timing were clustered into communities using their Bayesian methodology based on a Markov-switching model.

24. Hamilton and Owyang (ibid.) noted that the measure most synonymous with GDP at the state level is gross state product (GSP) but is only available at an annual frequency and at a two-year lag, making it nonviable for a study of the business cycle.

Research by Acemoglu et al. (2012) introduced economic models where aggregate economic fluctuations are determined by the most interconnected sectors, underscoring the importance of networking for macroeconomic outcomes. In addition, the literature indicates lessons from the Great Financial Crisis that highlighted that a high degree of integration and interconnectedness among financial firms can make the entire financial system vulnerable (Goldin and Vogel (2010)). Further, the business cycle interconnectedness tends to have implications for the movement of human and physical capital, technology transfer, macroeconomic policies and shocks, and ultimately, regional incomes.

5.2 Network Analysis

In general, network analysis provides conceptual and methodological tools to investigate empirically relational structures between interacting units as shown by Wasserman and Faust (1994). In the literature, interconnectedness concepts such as globalisation have been examined using network analysis, for example, Snyder and Kick (1979), Kim and Shin (2002) and De Lombaerde et al. (2018), the connectedness of financial firms by Diebold and Yilmaz (2014), etc. The literature further shows that network analysis has also been used to understand interactions of economic entities before and after certain key events. For example, some studies examine the impact of natural disasters before and after a specific disaster (Martin et al. (2006)). Other contributions investigate topical subjects such as the impact of financial crises on trade typically before and after the crisis (Ando and Kimura (2012) and Korniyenko et al. (2017)).

In this section, however, I attempt to have an understanding of the U.S. real output interdependence by addressing the following related questions based on the two datasets used in the preceding section: (i) Has there been an increase or decrease in the real output interconnectedness among states over time? (ii) What are the systematically important (central) states in the business cycle system by region, i.e. Midwest, Northeast, South and West? and finally, (iii) Do states form clusters or communities within the U.S.?

And, the Chicago Fed National Activity Index(CFNAI) is used as a monthly measure of the US national business cycle.

Based on the externalities of networking, particularly over time, it is expected that the economic activity network of the U.S. forms clusters consistent with findings of other studies using proxy measures of economic activity. Literature, for instance, shows that clustering can be due to economic and political factors such as efforts of synchronised regional capital investments, technology and federal policy responses over the years (for example, De Lombaerde et al. (2018)). To establish the changes in the interconnectedness of the state-specific business cycles, the study examines changes in the network centrality parameters for each state. The study further identifies clusters using the Blockbuster Algorithm proposed by Brownlees et al. (2022) in their recent work.

5.3 Economic Activity Interconnectedness

5.3.1 Definitions of Centrality Measures

Generally, centrality determines the central actors in a network's structure. Several such measures of network centrality are used to identify structurally important actors in a network. The most common centrality measures are the following: degree, betweenness, closeness and eigenvector. Their analytical usefulness depends heavily on the context of the network, the type of relation being studied and the underlying network structure. As such, for this study, I will define and apply two measures, namely degree and eigenvector centrality.

5.3.2 Degree Centrality

Degree centrality simply counts how many edges each node has - the most degree central actor is the one with the most ties.

5.3.3 Eigenvector Centrality

Eigenvector centrality though related to degree centrality has a fundamentally different idea behind it. The idea behind it is that we might think that power comes from being tied to powerful entities. So, if X and Y have the same degree of centrality, but X is tied to all high-degree entities and Y is tied to all low-degree entities, then reasonably X has a higher score than Y.

5.3.4 Community Detection

Community detection is a process used in network analysis to identify groups or clusters of nodes that are more densely connected internally than with the rest of the network. These groups, or communities, represent subsets of the network where nodes have stronger interactions or similarities with each other compared to those outside the community. Community detection is crucial in understanding the structure and function of complex networks, such as social and economic networks, biological networks, and technological networks.

In the literature, one of the most popular models for studying community structure is the stochastic block model (SBM, Holland et al. (1983)) whose basic intuition is that vertices are partitioned into groups and the probability of the edge between two vertices is p if they belong to the same group and q otherwise. It is assumed that vertices are more likely to be linked within the same group if $p > q$. Furthermore, in the literature, there are other prominent measures of connectedness, for instance, the Diebold–Yilmaz dynamic connectedness measurement framework based on variance decompositions which define weighted, directed networks (see Diebold and Yilmaz (2014)).

This study, however, employs a recent community detection approach (The Blockbuster Algorithm) developed by Brownlees et al. (2022) to study the community structure of the U.S. This is a partial correlation network model with a community structure for a large panel of time series. The series in the network are partitioned into groups such that correlation is higher within groups than between them. The communities are detected using the eigenvalues of the sample covariance matrix. They argue that this approach is ideal for modeling economic and financial time series since the series typically exhibit a factor structure or influenced by a set of common factors. The methodology is as follows:

The Blockbuster Algorithm

Inputs:

- Sample data Y_t for $t = 1, \dots, T$.
- Number of communities k .

Procedure:

1. Compute the sample covariance matrix $\hat{\Sigma}$:

$$\hat{\Sigma} = \frac{1}{T-1} \sum_{t=1}^T (Y_t - \bar{Y})(Y_t - \bar{Y})^T,$$

where $\bar{Y} = \frac{1}{T} \sum_{t=1}^T Y_t$.

2. Construct the eigenvector matrix $\hat{U}$:

- Compute the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ of $\hat{\Sigma}$.
- Select the k largest eigenvalues and form the $[n \times k]$ matrix $\hat{U}$ using their corresponding eigenvectors.

$$\hat{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k],$$

where each $\mathbf{u}_i$ is an eigenvector corresponding to the eigenvalue λ_i .

3. Standardise each row of $\hat{U}$:

- Normalise each row $\mathbf{u}_i$ of $\hat{U}$ by its Euclidean norm.

$$\hat{X}_{ij} = \frac{\hat{U}_{ij}}{\|\hat{U}_i\|},$$

where $\|\cdot\|$ denotes the Euclidean norm.

4. Apply the k-means algorithm to the rows of $\hat{X}$:

- Partition the rows of $\hat{X}$ into k clusters $\hat{V}_k = \{\hat{V}_1, \dots, \hat{V}_k\}$ using the k-means algorithm.

Output:

$$\hat{V}_k = \{\hat{V}_1, \dots, \hat{V}_k\},$$

where each $\hat{V}_i$ is a cluster or community.

5.4 Empirical Results

The results indicate that there have been changes in network centrality parameters that are reported in Appendix Table 3.8. In particular, the results show that the degree and eigenvectors for the majority of the states have increased for the period 2005 - 2022 when compared to the period 1991 - 2004. This observation means that there is greater economic activity connectedness among the U.S. real output. States such as Georgia (8

to 19), Indiana (5 to 13), Oklahoma (2 to 14) and Rhode Island (7 to 14) have shown the most considerable increase in their degree of centrality between the two periods considered. This observation resonates with Leiva-Leon (2017) who finds that contractionary shocks were more synchronously and uniformly spread toward most of the U.S. since the early 1990s meaning that the U.S. regions had become more interdependent.

However, I also observe that some states have had their activity connectedness decrease over the two periods, for example, Alabama, Massachusetts, Minnesota, Nevada, New Jersey, New York and Virginia. Although the eigenvector centrality is conceptually different from degree centrality, I observe similar trends in both measures, thus, leading to an overall conclusion that economic activity connectedness has generally increased over time.

Regarding the centrality of states, the results indicate that the important states in each region have changed between the two time periods. For the period 1991 - 2004, the centres were New Jersey, Alabama, Wyoming and Ohio for the Northeast, South, West and Midwest, respectively. However, in the period between 2005 - 2022, the centres change to Rhode Island, Georgia and Michigan for the Northeast, South and Midwest, respectively. Wyoming remained the centre of the West region. All the states exhibit interconnectedness within the block region they belong to (see Figure 3.6 and Figure 3.7).

The community detection results indicate that the states are partitioned into three large clusters and two small clusters. Although these results are not directly comparable to previous prominent studies, disregarding the two small clusters yields results that broadly converge with the Hamilton and Owyang (2012) and Brownlees et al. (2022) observations. Indeed, in their study, Hamilton and Owyang (2012) use more information with other explanatory variables and show that some states have a low posterior probability of being a member of any cluster. However, Brownlees et al. (2022) argues that the Blockbuster algorithm assigns every state to a community. This could explain the extra two small clusters I observe when using my augmented historical data. From Figure 3.8, another notable observation is that the largest cluster has state membership from all four regions whereas the other two are predominantly Midwest (including California and Nevada) and West (including Texas, Pennsylvania and Vermont).

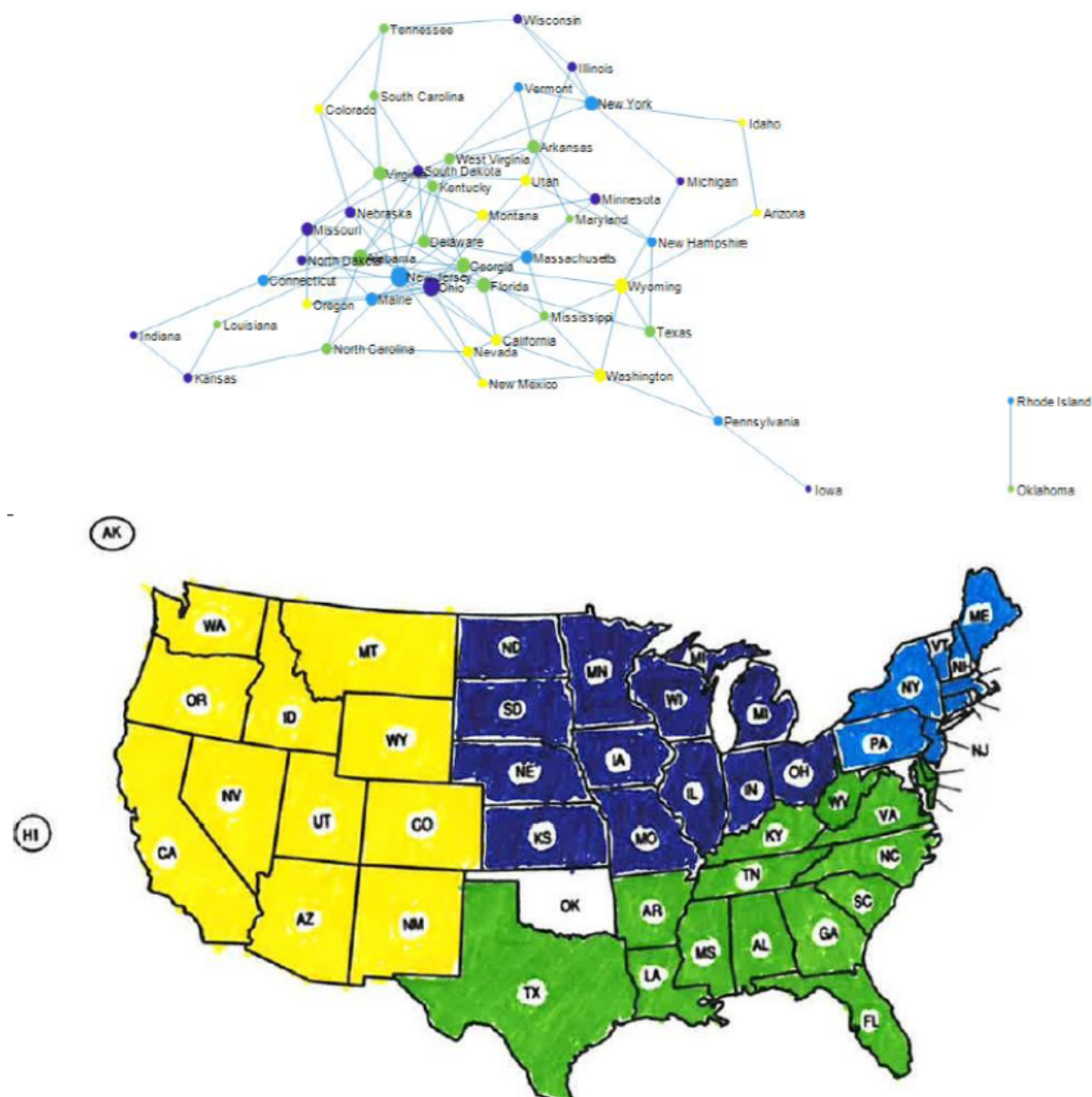


Figure 3.6: Period 1 Real Output Network

Notes: The top panel shows the connectivity network of the U.S. Real output between 1991 and 2004 whereas the bottom panel shows the main regional divisions. The size of the dots in the top panel indicates the extent of connectivity (central actor status) of a given state so that a bigger dot represents the importance of a given state at the national level and within a given region.

Source: Author's Computations

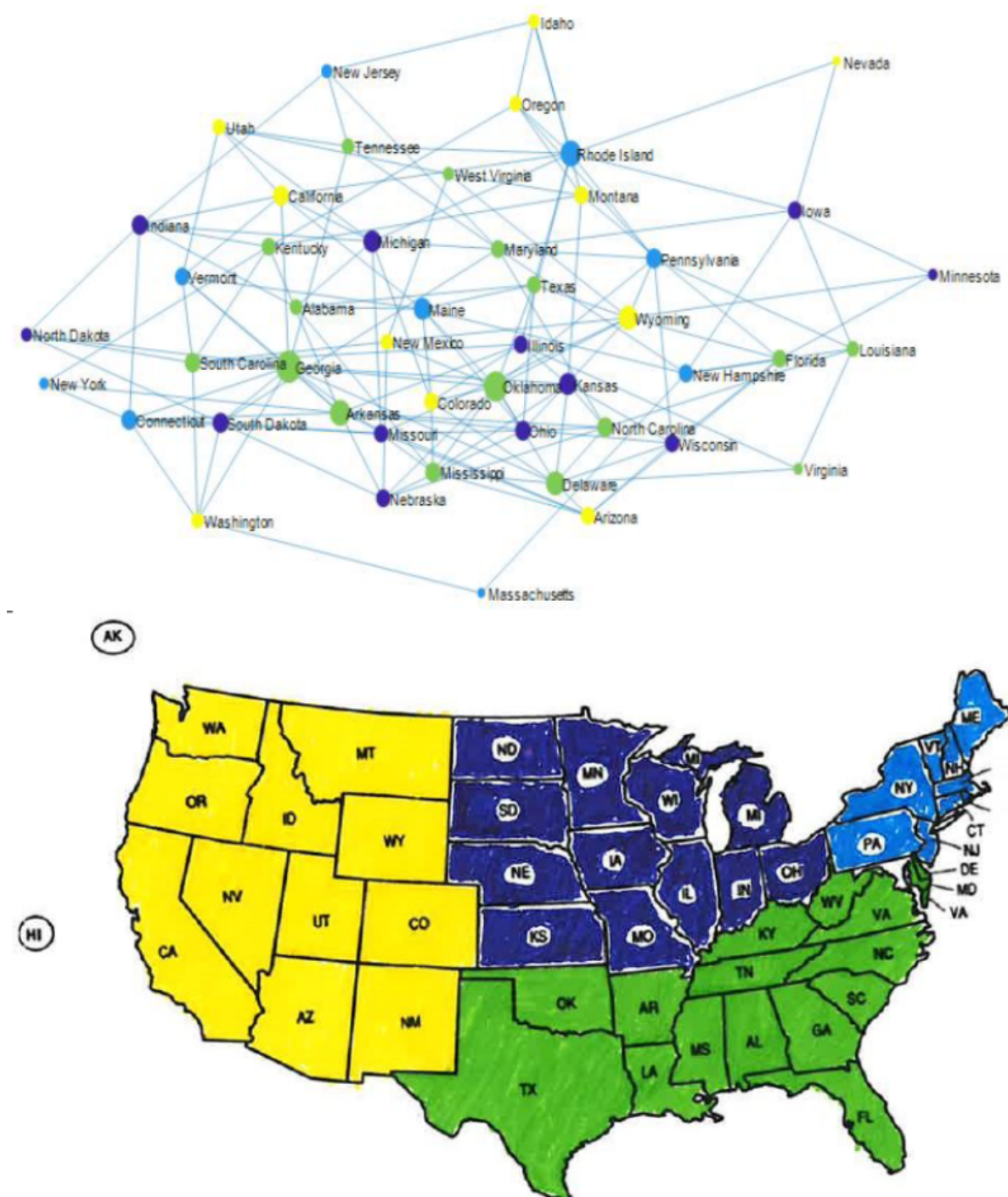


Figure 3.7: Period 2 Real Output Network

Notes: The top panel shows the connectivity network of the U.S. Real output between 2005 and 2022 whereas the bottom panel shows the main regional divisions. The size of the dots in the top panel indicates the extent of connectivity (central actor status) of a given state so that a bigger dot represents the importance of a given state at the national level and within a given region.

Source: Author's Computations

6 Conclusions

The current study estimates a historical state-level output data set and performs an analysis on how an aggregate financial shock affects economic activity across different states in the United States. The analysis further explores the complex interconnectedness and regional clustering trends of real economic activity in the United States. The research primarily extended the official state-level economic datasets in order to produce an extended historical quarterly real GDP growth estimates for each of the 48 contiguous states. The data set now encompasses output growth data from 1991 through to 2004. The estimation of state-level quarterly GDP growth was achieved using the Partial Least Squares temporal disaggregation method that integrated annual state output growth figures with quarterly state-specific economic indicators. The novel method enabled me to derive high frequency quarterly data from annual state-level economic statistics, marking an essential advancement for states previously missing quarterly economic data.

By extending state-level output time series, the study is in a position to add valuable insights to two separate research areas. The first strand of research to which the current study contribute is - the examination of how financial shocks or disruptions affect real economic performance across individual states. In this chapter, the study investigates how an aggregate financial disturbance produce different economic outcomes across states due to varying levels of regional economic strength and susceptibility. The second strand of economic literature to which I contribute is the literature that examines the relationships and groupings of economic activity across individual states. The analysis of these two dimensions enhances our understanding of regional economic dynamics in the United States.

The primary results of the study indicate multiple essential findings. The main findings show that an aggregate financial shock lead to negative macroeconomic impacts on the economic activities of individual states. The excess bond premium, which serves as an indicator of financial stress, reveals measurable impacts on the economic performance of individual states. Impulse response functions demonstrate that a financial shock leads to widespread reductions in economic activity across most states throughout the country. Each state experiences different levels of economic deterioration following financial dis-

ruptions. The financial shock impact some states harshly, while others maintain resilience which indicates varied regional capacities across the U.S. to manage financial disruptions. This research provides simple yet valuable information about economic variation throughout the U.S. while documenting the effects of a specific financial shock.

Further, the research utilised network analysis methods to examine how economic activities are interlinked across different U.S. states. The analysis identified important interconnectedness patterns which demonstrated strong economic connections between states' economic activities. The interconnectedness of states across two data sets confirmed the increased interconnectedness among states. Other results from the "BlockBuster Algorithm" revealed that there are three main groups of economically linked states alongside two minor groups displaying robust economic connections. The formation of these clusters demonstrates shared economic characteristics across U.S. regions. Identifying these clusters helps reveal the economic interdependencies that cross state boundaries. These results underscore that financial shock assessments must include both state-specific impacts and larger regional interactions which define the U.S. economy.

The findings provide valuable insights for policymakers responsible for economic stability who must tackle financial disruption challenges. Policymakers must navigate the complex task of balancing each state's unique economic features and necessities against the wider context of business cycle and economic linkages between states. The identification of connected clusters of states shows that policy makers need to evaluate the level of interdependence between states because economic disturbances can spread rapidly to nearby or linked states. In addition, policymakers need to comprehend regional heterogeneity because uniform national policies are unlikely to meet the needs of every state efficiently. Different states require customised and region-specific strategies to navigate their unique economic challenges during periods of disruption. Although national policies address broad economic conditions effectively, this study indicates the necessity of focusing more on how different regions perform economically and their mutual economic connections. Designing precise and effective policy responses to economic disruptions requires an intricate understanding of how each state's economic activity connects with other regions.

Chapter 3: Appendix

A Why 48 Contiguous U.S.States

Applied economists typically focus their attention on the 48 states that are adjacent to one another in the United States because they represent the vast majority of the nation's population, economic activity, and institutional diversity while maintaining a constant geographic and policy framework. This is because the surrounding states are located in close proximity to one another. The states that are considered to be part of the "lower-48" collectively contribute more than 80% of the gross domestic product of the United States (Jacobson et al. (2012)). These states also have employment, industrial, and consumer percentages that are proportionally large. In order to evaluate a dataset that is representative of the core of the American economy, most studies in studies in the literature eliminate Alaska and Hawaii - which I also follow. These two states are geographically isolated and rely on distinct supply networks. The subtraction of these states generally allows researchers to evaluate the dataset without introducing distortions that are caused by significant geographic peculiarities. As a consequence of this choice, the outcomes of empirical research are more pertinent to stakeholders and policymakers who are operating within the contiguous mainland. This is achieved by ensuring that tendencies at the national level in areas such as employment markets, housing prices, tax legislation, and interstate trade are appropriately portrayed.

In terms of methodology, the contiguous states constitute a natural unit for the purposes of panel-data econometrics and geographical analysis. This is because the states are geographically close to one another. There are a number of contemporary methods that, in order to function properly, require exact definitions of adjacency and distance. Although this is not the case for my current methodology, below are some examples of these methods; gravity-type trade regressions, difference-in-differences with surrounding controls, and spatial autoregressive models.

The consistency of the data and the availability of the data are just some of the criteria that play a role in determining which states are considered to be contiguous. The lower 48 states are the default coverage for state-level series that are released by federal statistics agencies on topics ranging from employment and income to the formation of firms and health outcomes. These series cover a wide range of topics. There is a chance that the frequent participation of Alaska and Hawaii in microdata releases, maybe in the form of footnotes, aggregated categories, or with cell sizes hidden for privacy reasons, could make it more challenging to compare microdata between states and to guarantee that time series are complete. In my current study, like other researchers before, I avoid the problems of patchy or irregular series, missing data, and the need for intensive imputation or special-case coding by employing the technique of limiting analysis to the 48. This allows me to dodge these concerns. I am able to keep the sample size and econometric quality of their study intact as a result of this.

As alluded to earlier, numerous regional and sectoral studies have increased their external validity by concentrating on the states in the United States that are within close proximity to one another. Markets for shared labour and products have an effect on economic policies, which usually lead to differences across jurisdictions that are located in close proximity to one another physically. The minimum wage standards, state income taxes, and environmental regulations are all examples of policies that are included in this area when it comes to policy examples. Being able to travel or obtain inputs across nearby borders makes it easier to separate the causal impacts of changes when these policies are evaluated in a setting that is geographically integrated. This is because citizens and businesses are able to move across neighbouring boundaries. In addition, because the lower 48 states share infrastructure corridors that connect their economies, many researchers in the literature have been able to explore transmission processes in a realistic context. This is because the corridors connect the economies of the states. The electricity grids, rail networks, and roadways that make up these corridors are important. Coherence, breadth, and analytical tractability are all important aspects of American economic research, and this method strikes the ideal balance between these three aspects. This approach to problem solving has proven successful when all of the relevant aspects are taken into consideration and these are the reasons I have adopted to justify the approach taken in my study.

Table 3.6: State-Specific Economic Variables

Task ID	Variable Name
1	Unemployment Rate, Percent, Monthly, Seasonally Adjusted
2	Labor Force Participation Rate, Percent, Monthly, Seasonally Adjusted
3	All Employees: Total Nonfarm
4	All Employees: Manufacturing
5	Coincident Economic Activity Index, Index 2007=100, Monthly, Seasonally Adjusted
6	Civilian Labor Force
7	All Employees: Construction
8	All Employees: Government
9	All Employees: Non-Durable Goods: Printing and Related Support Activities
10	All Employees: Trade, Transportation, and Utilities
11	All Employees: Financial Activities
12	All Employees: Wholesale Trade: Merchant Wholesalers, Durable Goods
13	All Employees: Financial Activities: Agencies, Brokerages, and Other Insurance Related
14	All Employees: Professional and Business Services
15	All Employees: Education and Health Services
16	All Employees: Retail Trade
17	All Employees: Retail Trade: Department Stores
18	All Employees: Manufacturing: Non-Durable Goods
19	All Employees: Manufacturing: Durable Goods: Navigational, Measuring, and Electromedical
20	All Employees: Manufacturing: Durable Goods: Machine Shops; Turned Product
21	All Employees: Construction: Heavy and Civil Engineering Construction
22	All Employees: Government: State Government Excluding Education
23	All Employees: Government: State Government
24	All Employees: Government: Federal Government
25	All Employees: Other Services: Repair and Maintenance
26	All Employees: Leisure and Hospitality: Food Services and Drinking Places
27	All Employees: Professional and Business Services: Employment Services
28	All Employees: Financial Activities: Central Bank
29	All Employees: Financial Activities: Finance and Insurance
30	All Employees: Transportation and Utilities: Truck Transportation
31	All Employees: Retail Trade: Miscellaneous Store Retailers
32	All Employees: Manufacturing: Durable Goods: Medical Equipment and Supplies
33	All Employees: Manufacturing: Durable Goods: Transportation Equipment
34	All Employees: Manufacturing: Durable Goods: Machinery Manufacturing
35	All Employees: Manufacturing: Durable Goods
36	All Employees: Service-Providing
37	All Employees: Government: Local Government Non-Educational Services
38	All Employees: Government: Local Government Educational Services
39	All Employees: Government: Local Government
40	All Employees: Other Services: Personal and Laundry Services
41	All Employees: Government: U.S. Postal Service
42	All Employees: Other Services: Religious, Grantmaking, Civic, and Similar Organisations
43	All Employees: Professional and Business Services: Services to Buildings and Dwellings

Notes: The table reports all employees in Thousands of Persons, Monthly, Seasonally Adjusted for the period 1990 to 2004. All data are downloaded from <https://fred.stlouisfed.org/>.

B Some Econometric Methods for “Backcasting”

A thorough tool set has been developed by econometricians for the purpose of “backcasting”, which refers to the process of filling in previous data at higher frequencies or more granular levels than was first observed. The following is a list of some popular substitution strategies that go beyond simple bridge equations and competing procedures such as the missing data predictions based on PLS:

B.1 Proportional Denton and Benchmarking Methods

The Denton proportional-dent approach is one of the strategies that has been around the longest and is the most often used. Using a low frequency series (such as annual GDP) and a corresponding high frequency indicator (such as quarterly industrial production), it generates a quarterly path that (a) precisely adds up to the annual totals and (b) follows the indicator’s quarterly profile in the least-squares sense as closely as possible. A quarterly path is created. On the other hand, the proportional variation reduces the number of percentage modifications, and the additive Denton reduces the number of absolute adjustments. Despite the fact that these methods are extremely reliable in practice and do not call for any distributional assumptions to be made regarding the residuals, they are unable to incorporate more than one indication at the same time without requiring extensions.

B.2 Regression Model-Based Backcasting Methods

The first time that regression-based frameworks were presented was by Chow and Lin (1971). These frameworks regard low frequency series as an aggregate of a latent high frequency process that is driven by contemporaneous high frequency indicators. Fernandez (1981) and Litterman (1983) brought about improvements to the approach by proposing various error-structure assumptions and efficient estimate methods. This led to the method’s further improvement. Estimation of the parameters of the AR(1) or ARMA process, which the high frequency residual in these models follows, is performed in conjunc-

tion with the estimation of the regression coefficients. Despite the fact that these “bridge equation” systems are able to handle many indicators, they are dependent on parametric assumptions concerning the dynamics of the disturbance. If these assumptions are broken, it is possible that the disaggregated series will be skewed.

B.3 U-MIDAS and Mixed-Data Sampling (MIDAS)

For the purpose of directly embedding high frequency covariates and low frequency dependent variables within a single regression, mixed-frequency (MIDAS) regressions make use of polynomial lag-weight functions. By establishing a weighted polynomial that encapsulates the impact of each lag in a parsimonious manner, for example, it is possible to regress quarterly GDP on a number of monthly indices. U-MIDAS, which allows for an unlimited number of lag coefficients while still allowing for control over the total number of parameters, is a further development of this. These models are particularly useful for predicting and nowcasting since they make use of raw high frequency data without aggregating the data based on time. The inclusion of a large number of lags or indicators, on the other hand, may cause them to be sensitive to the weighting system that is selected, and they will require rigorous regularisation.

B.4 The Kalman Filter and State-Space Models

The evolution of the unobserved high frequency series over time is modeled as a latent state vector by state-space techniques. These approaches are used in conjunction with measurement equations that connect it to low frequency data and indicators that have been observed. The Kalman filter then produces the most accurate estimations of the state throughout each period, provided that all of the data that is available is considered. This flexible architecture is able to support a wide variety of data sources, including those with varied publication lags, stochastic volatility, and factors that change over time. Even though state-space models require careful specification and a significant amount of computational effort, they usually create historical series that are smoother and more coherent. This is because they frame the problem of temporal disaggregation as a statistical filtering solution.

B.5 Dynamic Factor Models

When there is a wide panel of high frequency indicators available, dynamic factor models are able to extract a few common latent factors that describe the co-movements that occur across series. Once this is accomplished, these variables can be immediately incorporated into mixed-frequency VARs (MF-VARs) or connected to low frequency objectives through the utilisation of bridge equations. The resulting factor-augmented framework enhances accuracy and resilience in comparison to single-indicator approaches by utilising information from hundreds of series. This is accomplished while taking into consideration idiosyncratic noise. Estimation is often accomplished through the use of principal components or Bayesian approaches, and the models can be extended to nowcast in real time with jagged-edge datasets.

B.6 Machine Learning and Non-Parametric Techniques

In more recent times, researchers have begun employing machine learning techniques, including as neural networks, boosting, and random forests, in order to nowcast or interpolate missing historical values. The management of interactions and nonlinearities is possible with these methods since they do not need imposing rigorous distributional assumptions. For example, a gradient-boosted tree model can automatically select important variables and lag structures while simultaneously learning detailed correlations between a low frequency series and dozens of high frequency predictors. This can accomplish both of these goals simultaneously. Post-processing is required to ensure that these procedures are consistent with the low frequency totals that have been seen, despite the fact that they are capable of producing good out-of-sample results. In general, these techniques do not have the precise aggregation limitations that are associated with traditional disaggregation.

Table 3.7: Mean Absolute Error (1991 to 2004)

State	CL ipi	CL emp	Fern ipi	Fern emp	Lit ipi	Lit emp	SSC ipi	SSC emp	PLS
AL	0.912	1.202	1.570	1.508	1.605	1.471	1.701	1.489	0.177
AZ	3.274	3.516	2.418	4.296	2.387	4.245	2.736	3.501	0.637
AR	1.298	1.767	1.214	2.109	1.219	2.073	1.801	1.956	0.591
CA	2.105	2.173	2.346	2.424	2.355	2.402	1.681	2.115	0.109
CO	2.743	2.681	2.356	3.212	2.338	3.169	2.022	2.478	0.308
CT	2.365	2.031	2.789	1.944	2.804	1.937	2.146	1.896	0.391
DE	2.809	2.486	2.599	2.727	2.603	2.712	2.413	2.280	0.559
FL	1.381	1.516	0.857	2.310	0.861	2.259	1.261	1.528	0.653
GA	1.787	1.870	1.393	2.689	1.388	2.638	1.315	1.901	0.327
ID	3.045	3.311	2.833	3.966	2.824	3.923	2.931	3.432	0.374
IL	0.813	0.691	1.582	1.087	1.616	1.043	1.054	0.842	0.174
IN	1.190	1.411	1.455	1.859	1.471	1.815	1.335	1.563	0.306
IA	2.001	2.321	2.159	2.602	2.163	2.580	2.257	2.524	0.578
KS	0.930	0.937	1.868	1.237	1.909	1.215	1.183	1.165	0.572
KY	1.046	1.311	1.706	1.703	1.725	1.659	1.467	1.571	0.605
LA	2.602	2.434	2.928	2.505	2.945	2.490	2.859	2.445	1.011
ME	1.489	1.447	1.946	1.508	1.974	1.495	1.718	1.435	0.254
MD	1.501	1.573	1.777	1.703	1.805	1.695	1.743	1.506	0.248
MA	1.807	1.528	2.065	1.838	2.078	1.817	1.349	1.339	0.224
MI	1.672	1.784	2.415	1.775	2.442	1.774	2.094	1.972	0.405
MN	1.576	1.884	1.525	2.373	1.520	2.328	1.467	2.021	0.229
MS	1.451	1.472	1.792	1.566	1.803	1.551	1.836	1.690	0.364
MO	1.020	1.229	1.831	1.545	1.858	1.523	1.233	1.429	0.165
MT	1.412	1.661	1.559	1.908	1.581	1.885	2.097	1.928	0.442
NE	1.946	2.024	1.926	2.260	1.940	2.244	2.256	2.104	0.596
NV	2.982	3.427	2.239	4.246	2.210	4.195	2.641	3.483	0.474
NH	1.806	2.127	1.590	2.715	1.593	2.671	1.543	2.074	0.332
NJ	1.522	1.259	2.299	1.161	2.335	1.153	1.610	1.235	0.171
NM	3.295	3.855	3.148	4.275	3.139	4.246	3.750	4.185	2.095
NY	1.872	1.572	2.437	1.447	2.471	1.454	1.855	1.402	0.217
NC	1.213	1.290	1.331	1.946	1.337	1.895	1.035	1.351	0.295
ND	2.849	2.954	2.795	3.084	2.798	3.077	2.902	3.009	1.173
OH	1.122	1.105	1.881	1.253	1.908	1.232	1.490	1.329	0.307
OK	1.294	1.165	1.861	1.413	1.901	1.397	1.443	1.246	0.242
OR	2.515	2.814	2.208	3.391	2.191	3.355	2.265	2.950	0.328
PA	0.968	0.749	1.882	0.913	1.923	0.890	1.370	0.921	0.118
RI	1.680	1.794	2.053	1.847	2.080	1.832	1.825	1.751	0.383
SC	0.686	0.814	1.100	1.427	1.128	1.383	0.971	0.992	0.331
SD	2.274	2.448	2.028	2.888	2.028	2.851	2.352	2.599	1.274

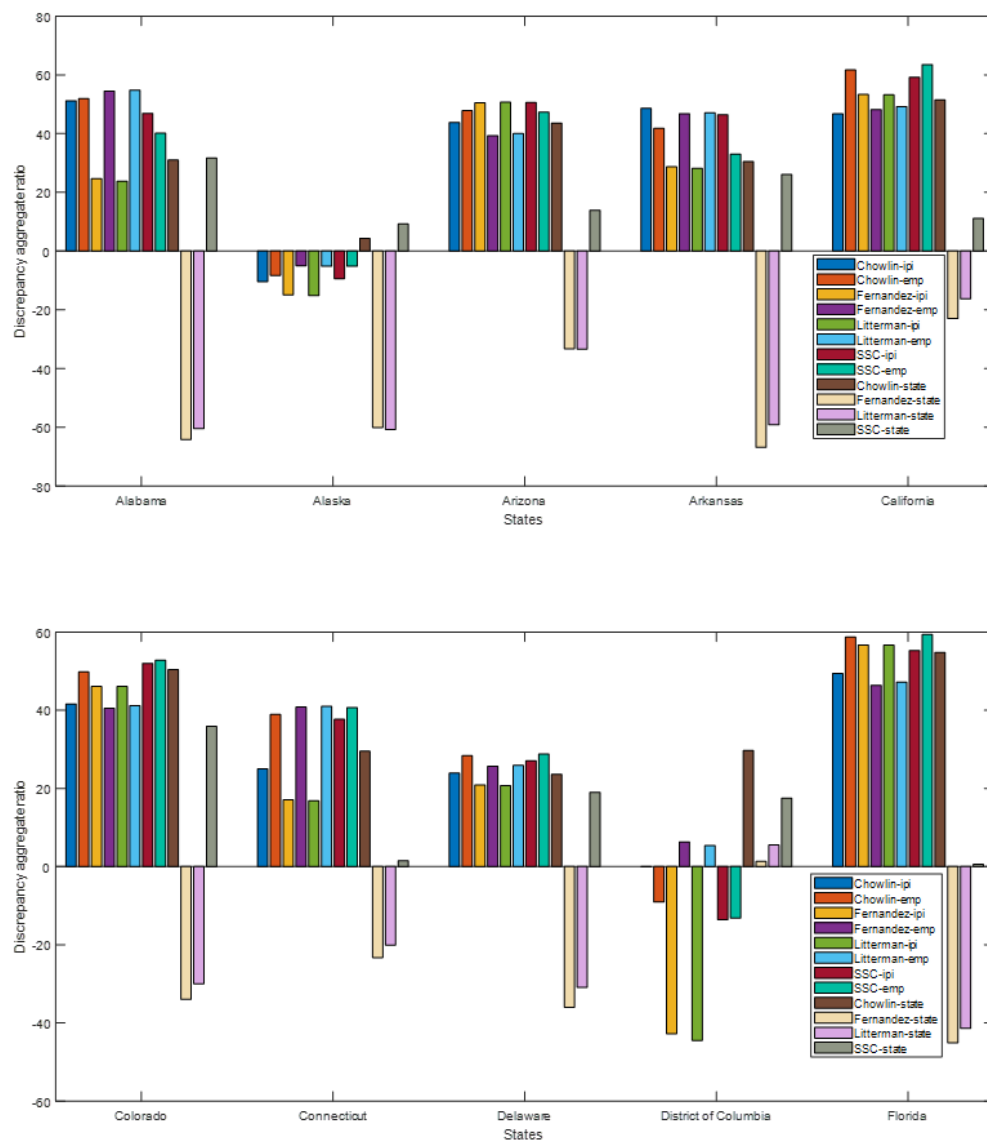
Continued on next page

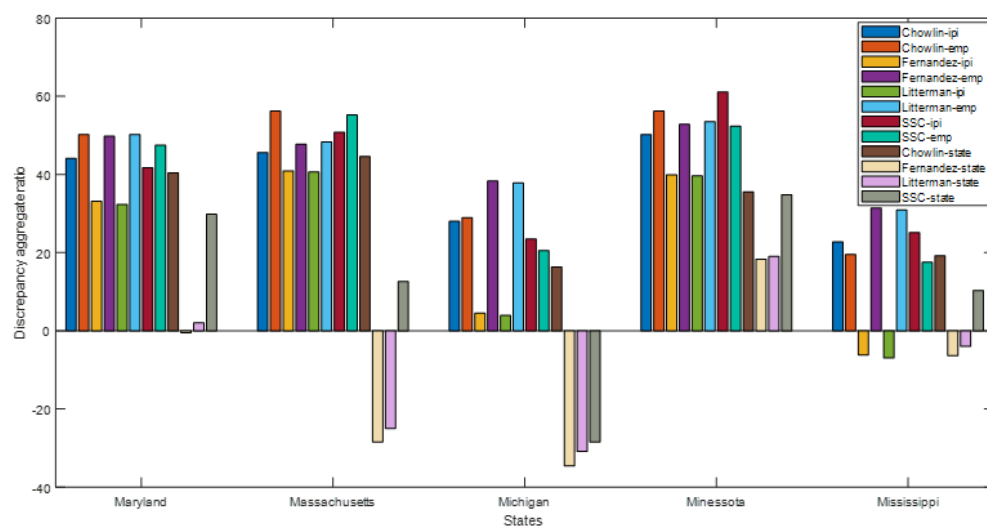
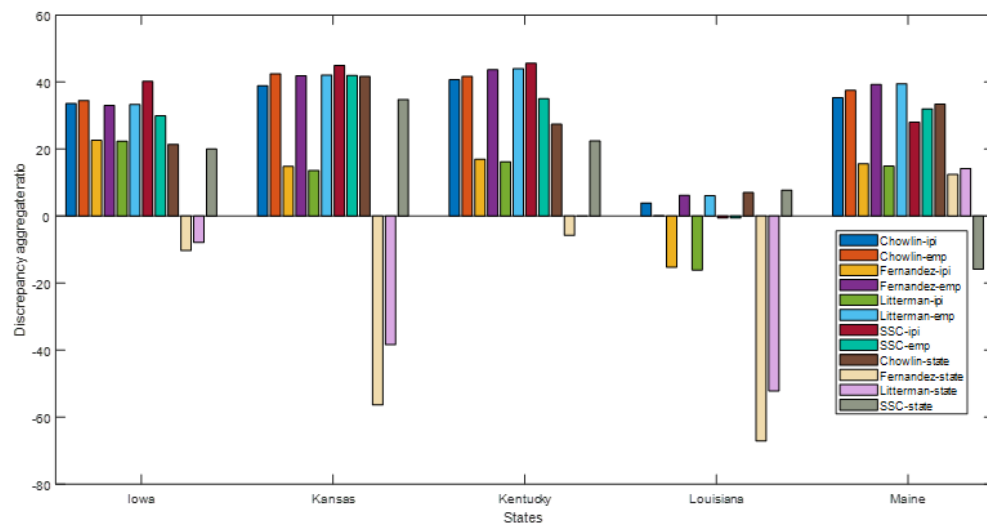
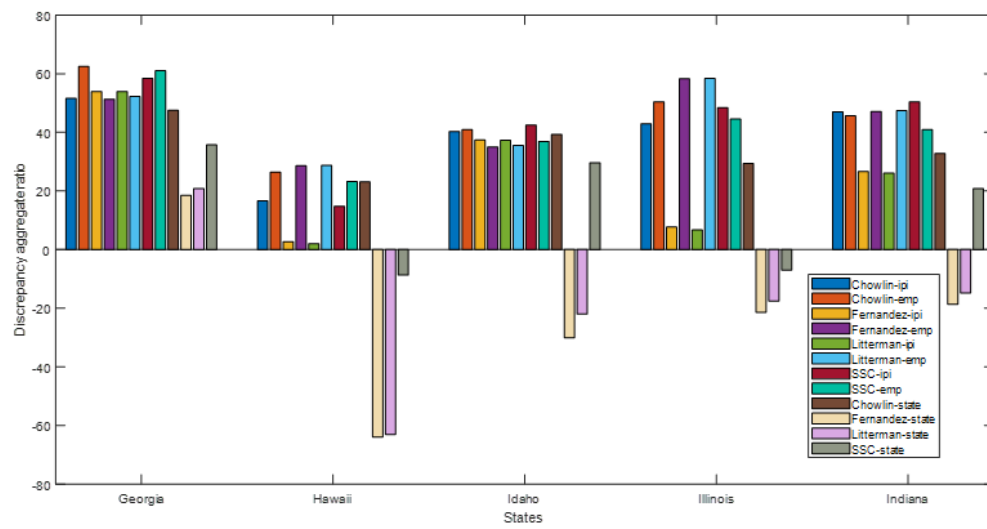
Continued from previous page

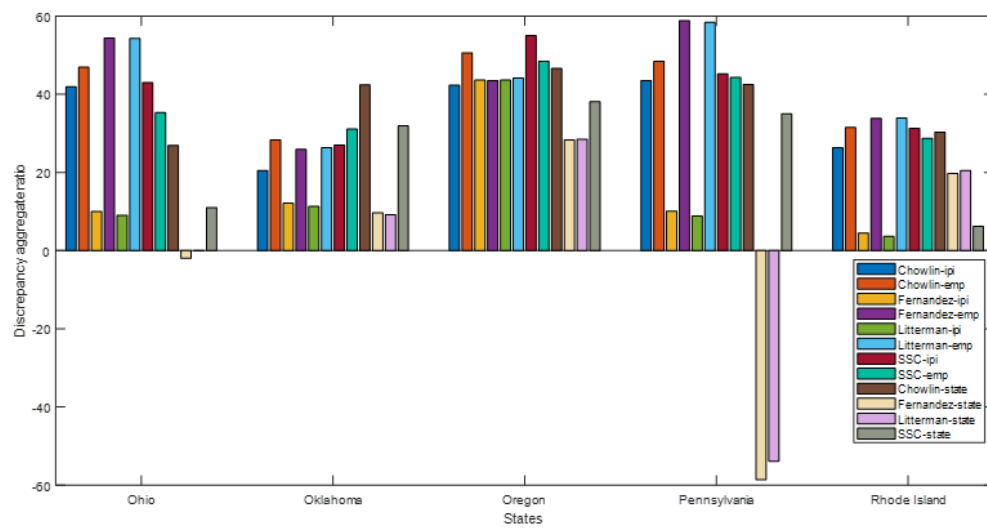
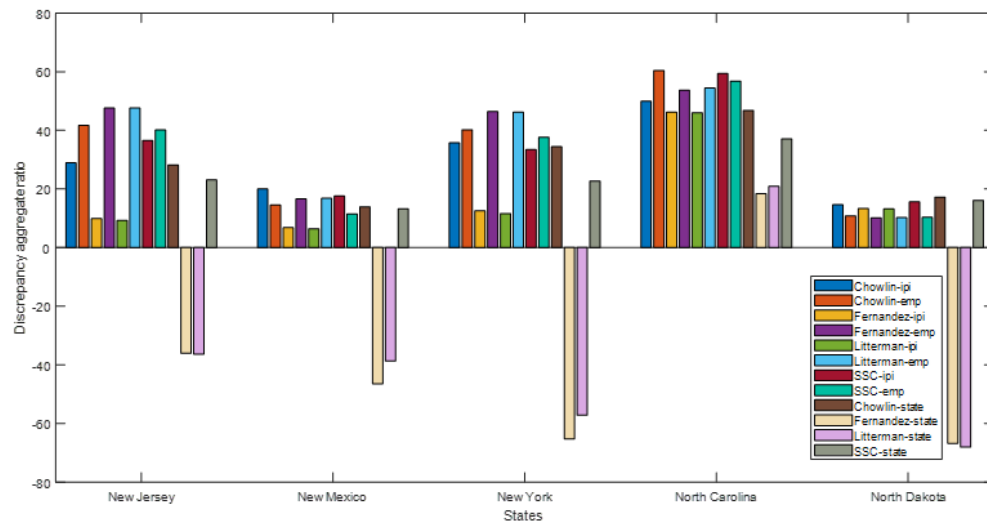
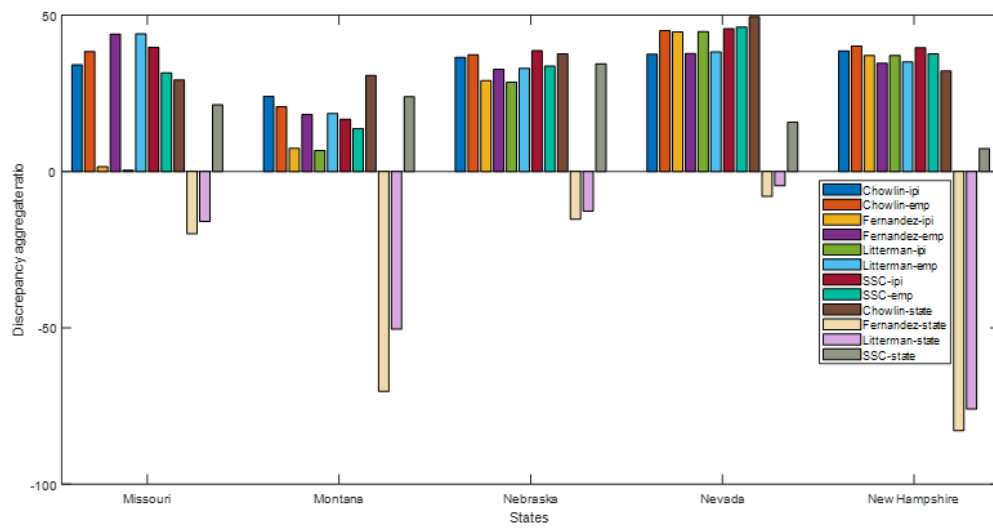
State	CL ipi	CL emp	Fern ipi	Fern emp	Lit ipi	Lit emp	SSC ipi	SSC emp	PLS
TN	1.235	1.683	1.258	2.182	1.263	2.138	1.750	1.994	0.158
TX	1.485	1.598	1.221	2.310	1.219	2.259	1.088	1.528	0.156
UT	2.036	2.091	1.680	2.910	1.677	2.859	1.785	2.119	0.287
VT	1.770	2.073	1.760	2.360	1.776	2.338	2.044	2.129	0.589
VA	1.335	1.163	1.460	1.573	1.481	1.536	1.061	1.137	0.237
WA	1.532	1.739	2.090	1.908	2.102	1.893	1.409	1.888	0.710
WV	1.322	1.193	2.332	1.008	2.366	1.005	1.780	1.381	0.515
WI	0.801	0.906	1.024	1.612	1.052	1.568	1.080	1.092	0.180
WY	1.480	1.527	1.982	1.700	2.022	1.669	2.019	1.578	0.294

Notes: AL=Alabama, AZ=Arizona, AR=Arkansas, CA=California, CO=Colorado, CT=Connecticut, DE=Delaware, FL=Florida, GA=Georgia, ID=Idaho, IL=Illinois, IN=Indiana, IA=Iowa, KS=Kansas, KY=Kentucky, LA=Louisiana, ME=Maine, MD=Maryland, MA=Massachusetts, MI=Michigan, MN=Minnesota, MS=Mississippi, MO=Missouri, MT=Montana, NE=Nebraska, NV=Nevada, NH=New Hampshire, NJ=New Jersey, NM=New Mexico, NY=New York, NC=North Carolina, ND=North Dakota, OH=Ohio, OK=Oklahoma, OR=Oregon, PA=Pennsylvania, RI=Rhode Island, SC=South Carolina, SD=South Dakota, TN=Tennessee, TX=Texas, UT=Utah, VT=Vermont, VA=Virginia, WA=Washington, WV=West Virginia, WI=Wisconsin, WY=Wyoming.

Figure 3.9: State Discrepancy Aggregates







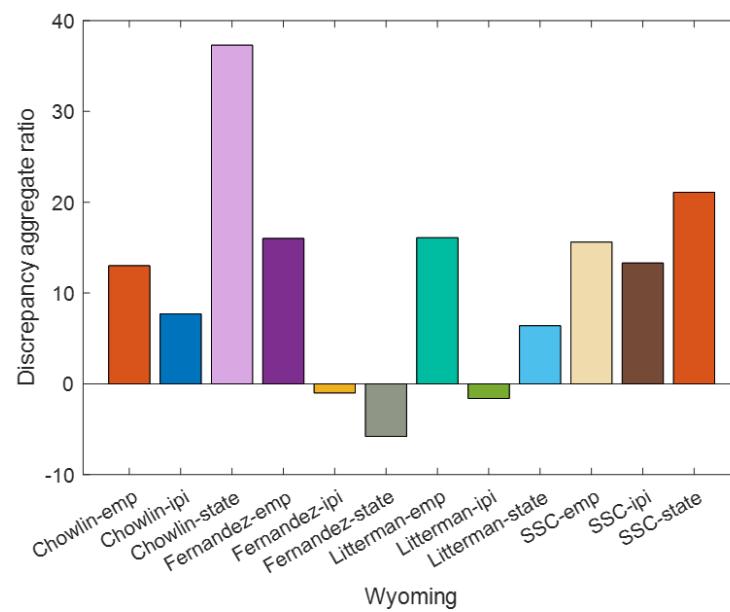
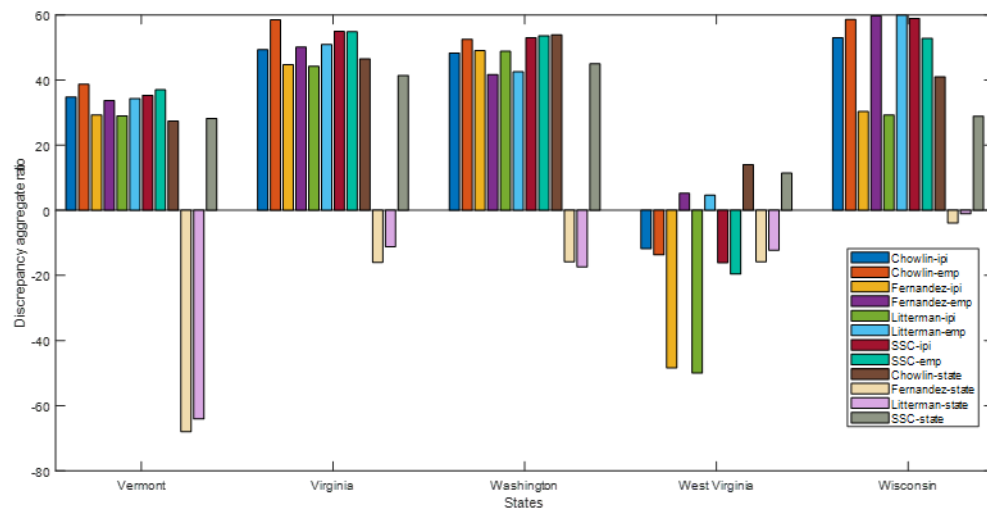
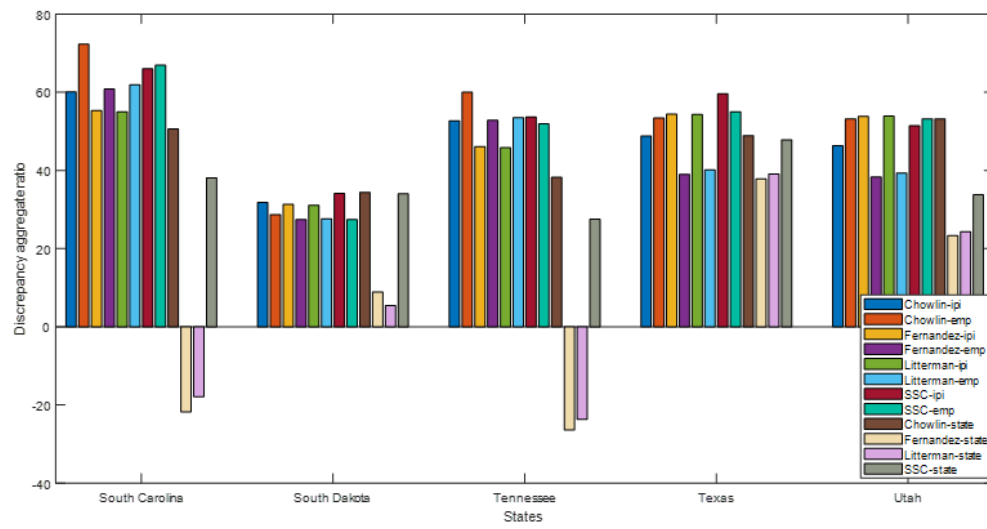
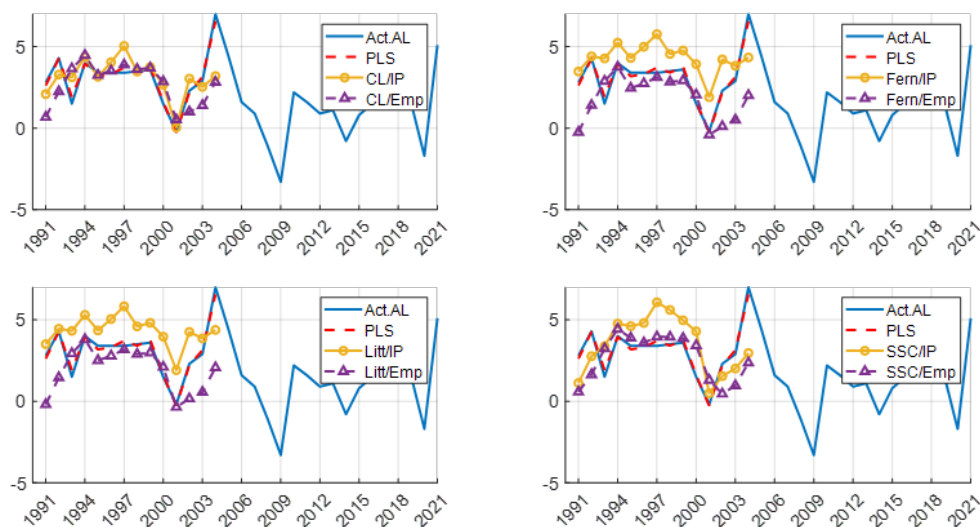
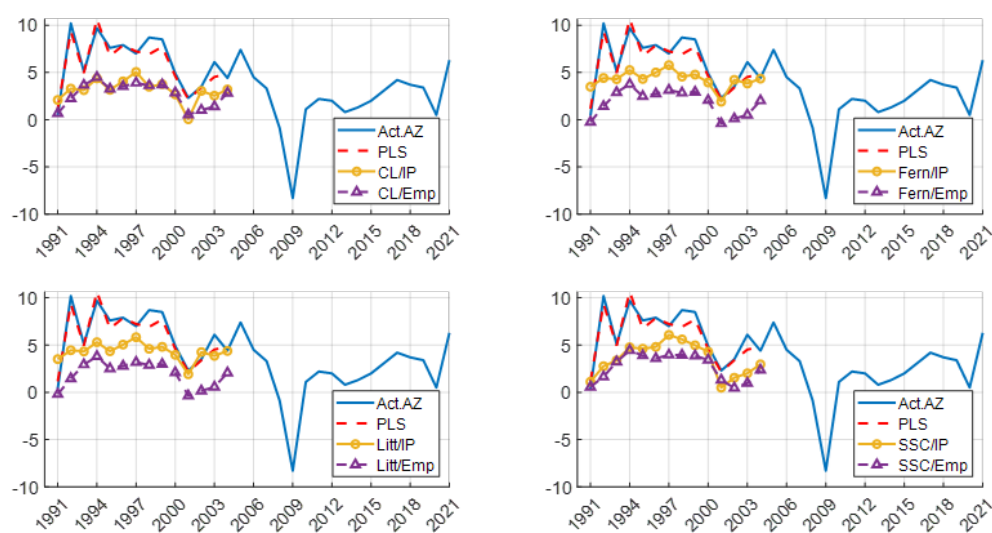


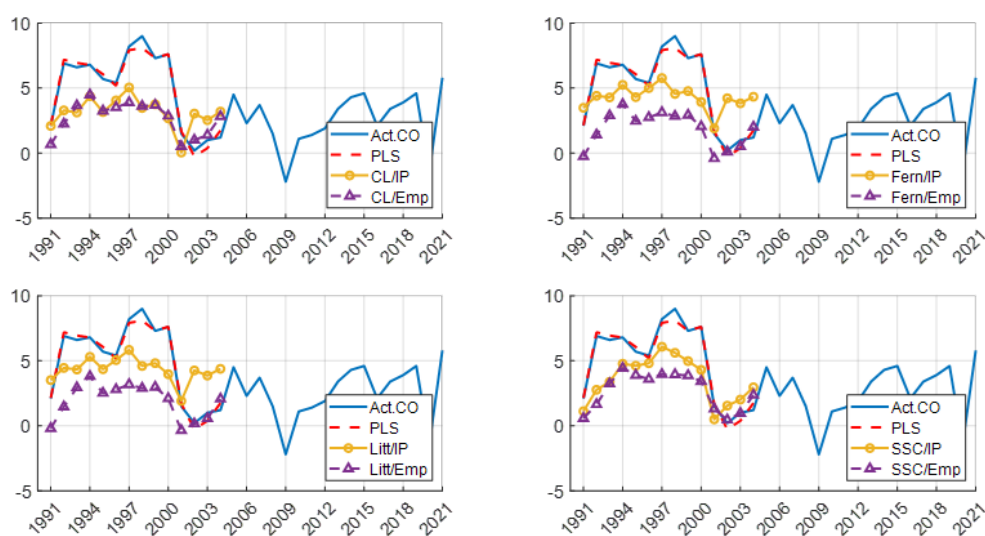
Figure 3.10: Conformity of State Output Estimates



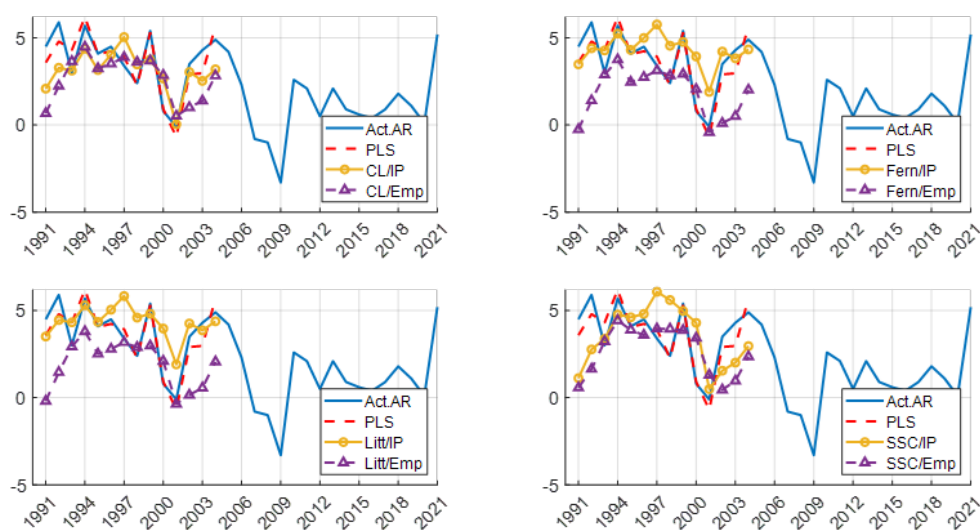
Alabama



Arizona



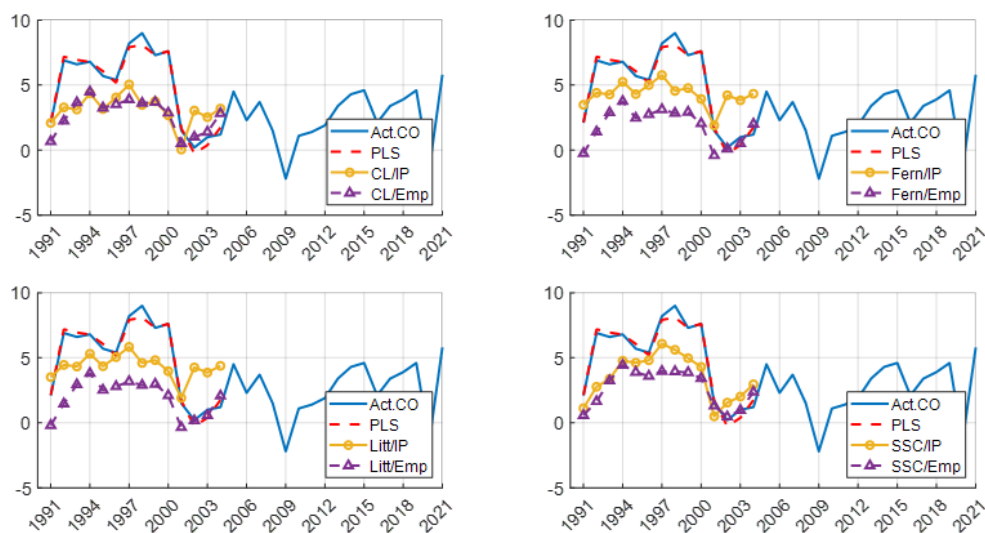
Arkansas



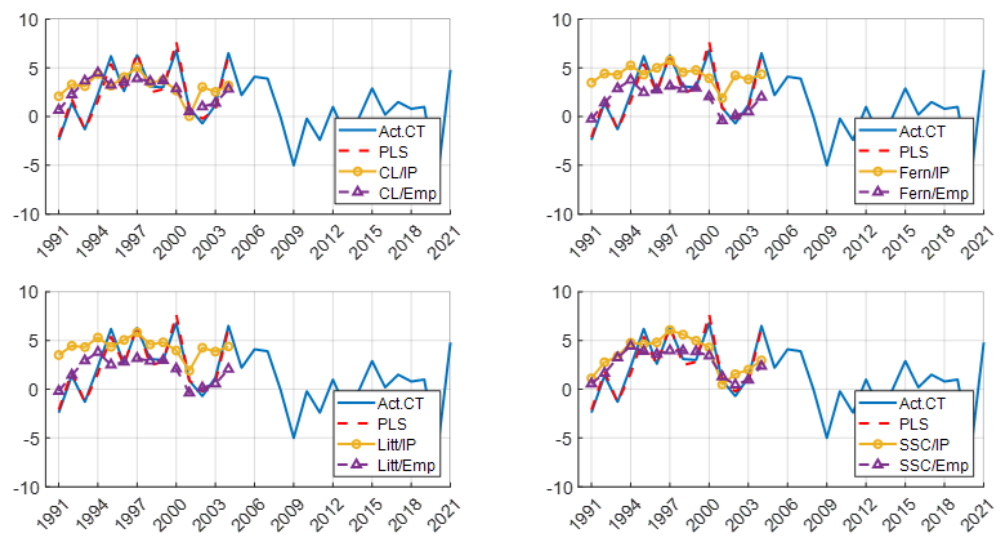
California



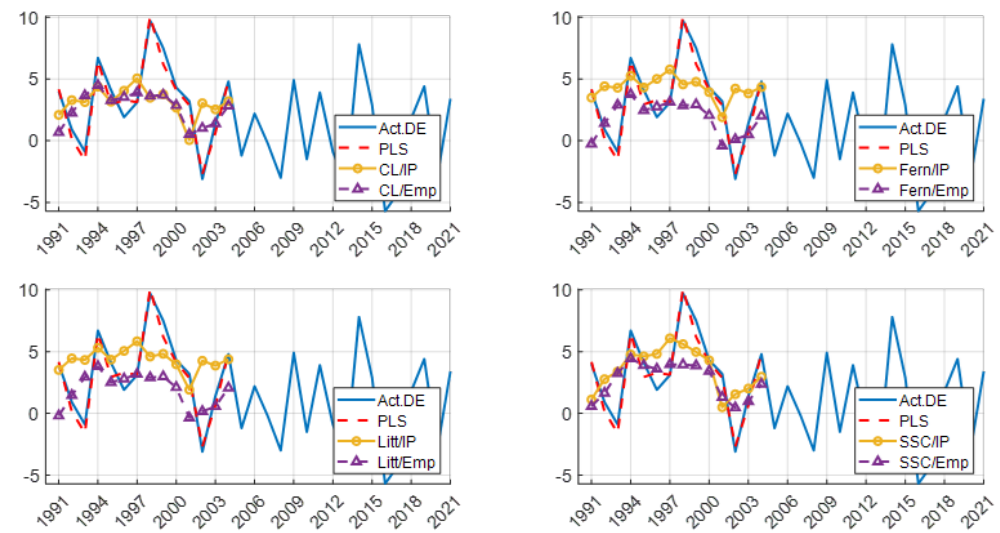
Colorado



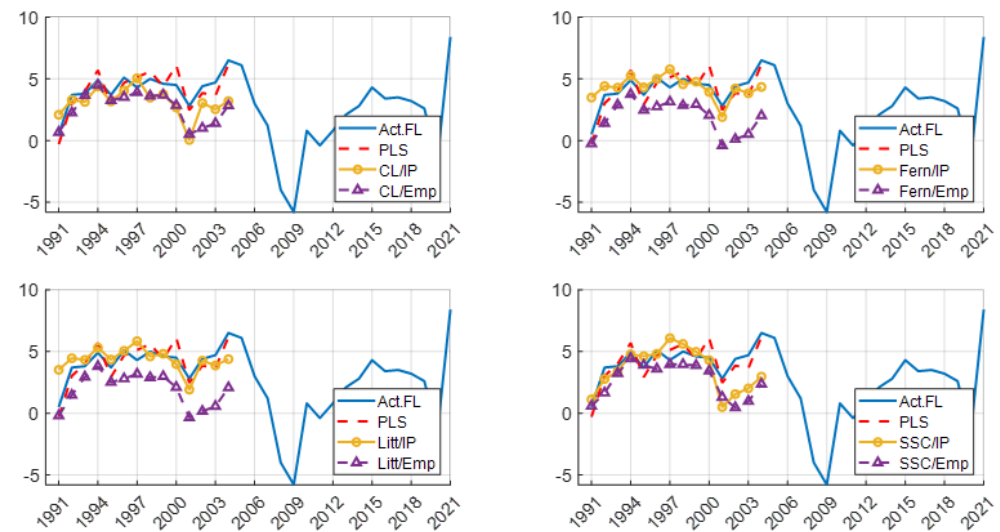
Connecticut



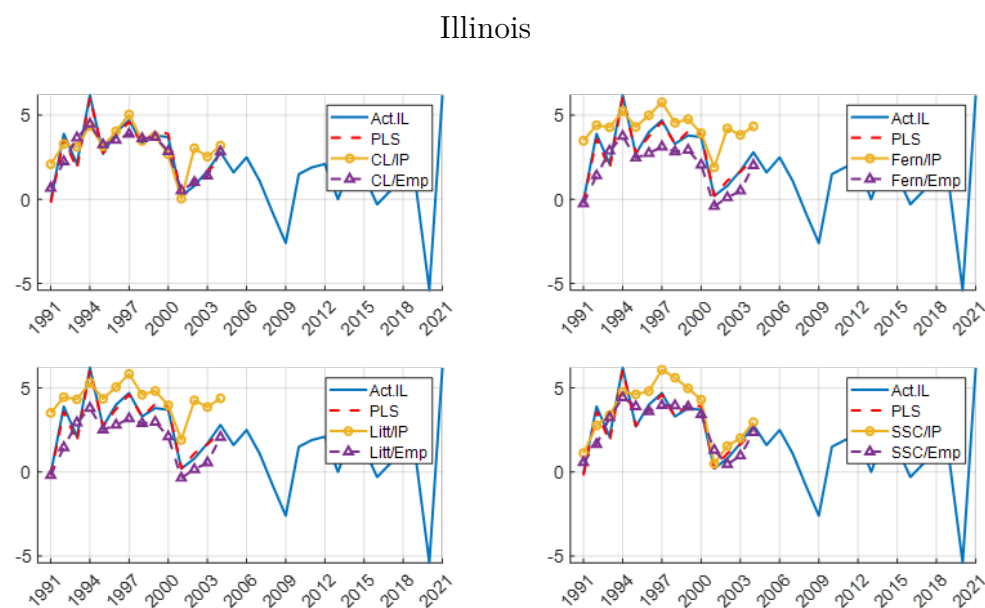
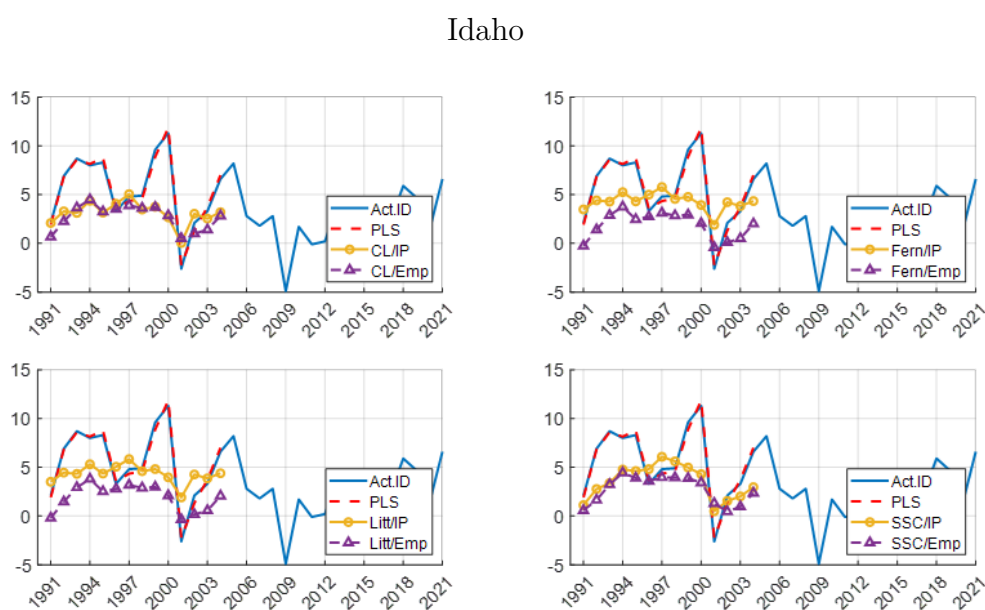
Delaware

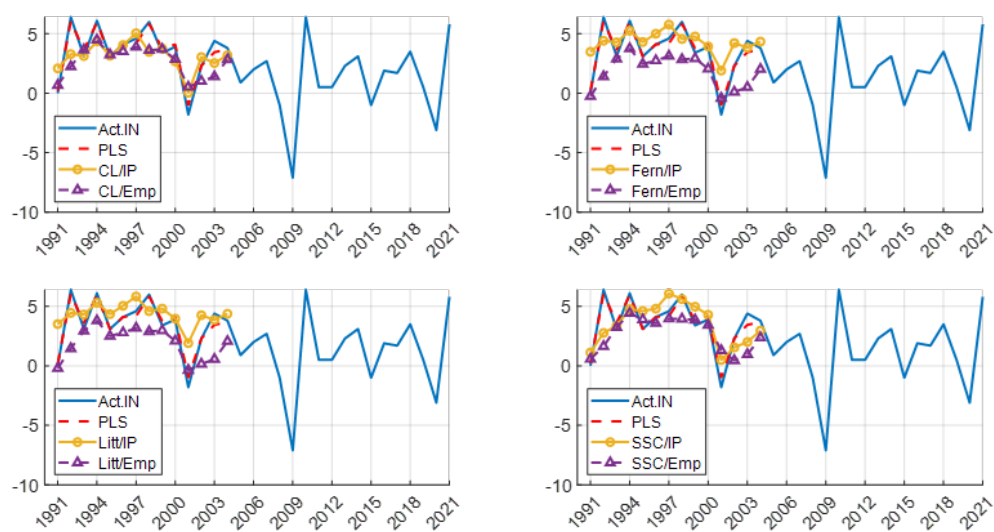


Florida



Georgia





Iowa



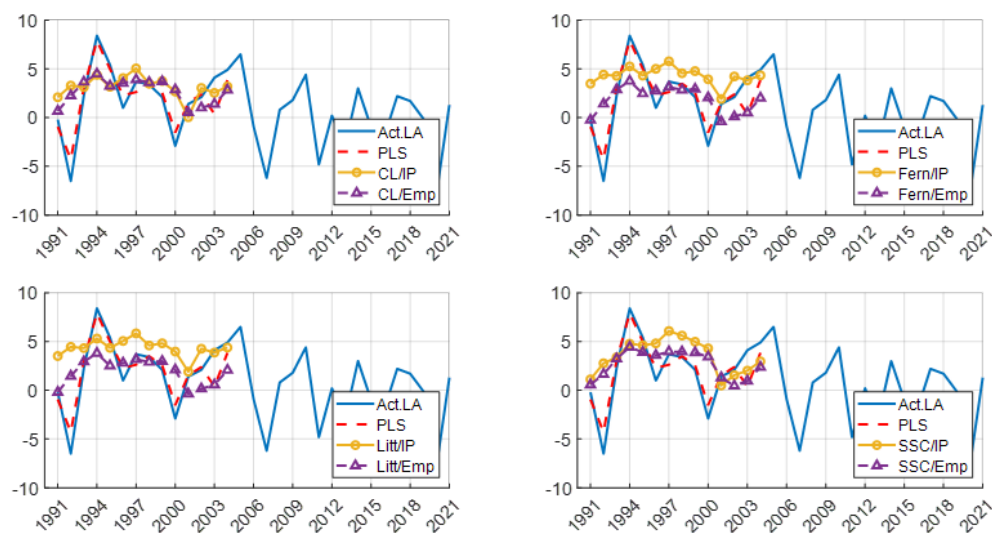
Kansas



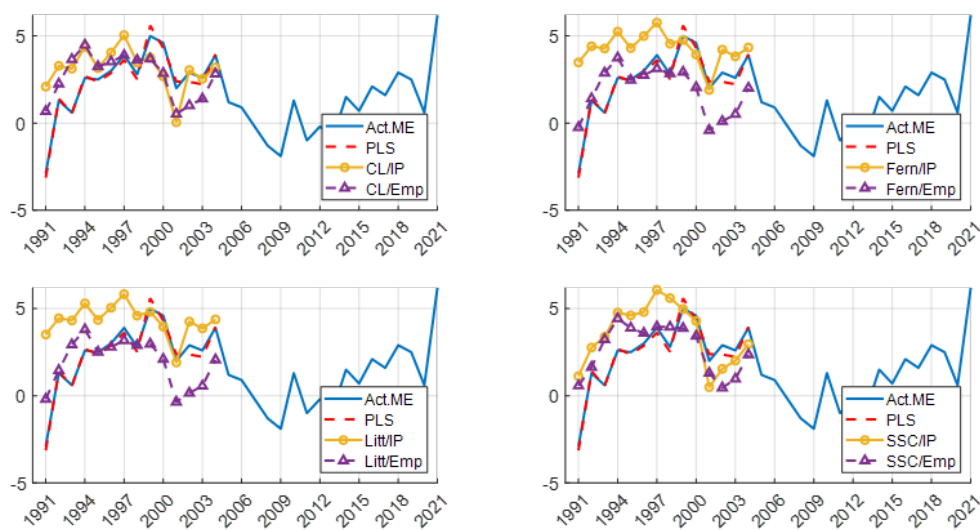
Kentucky



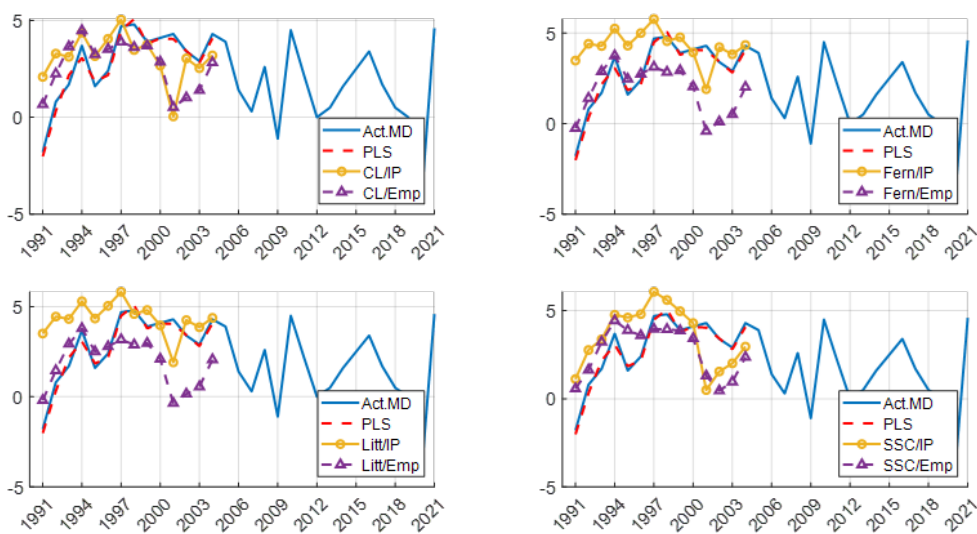
Louisiana



Maine



Maryland



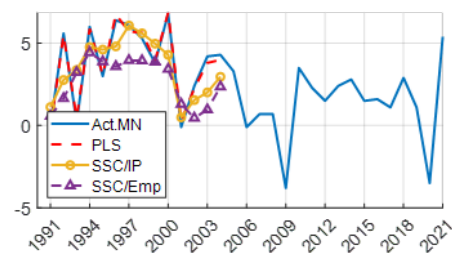
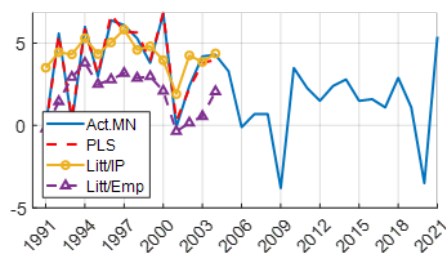
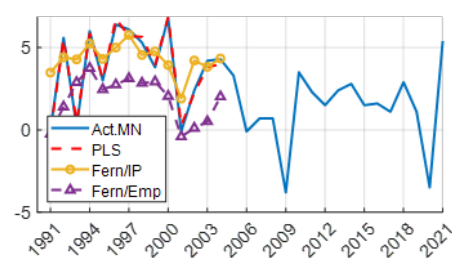
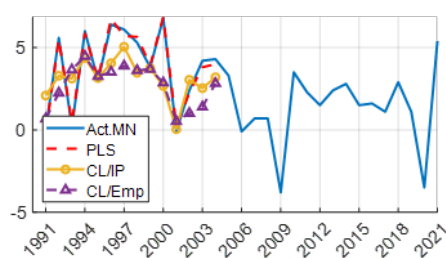
Massachusetts



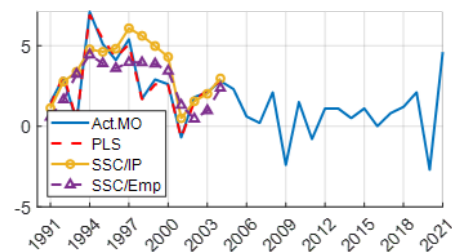
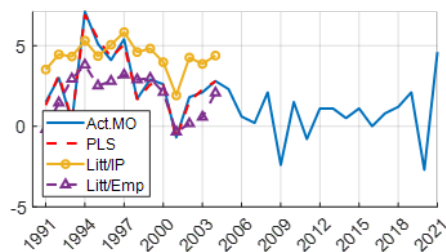
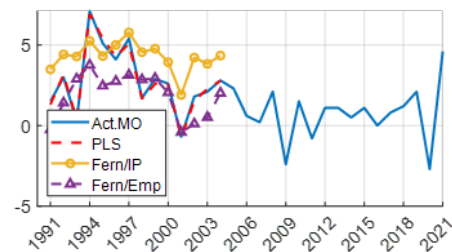
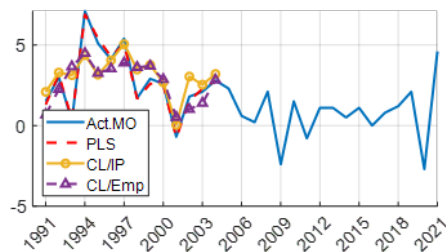
Michigan



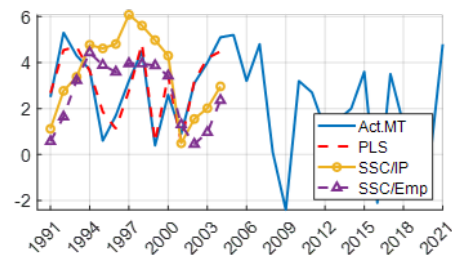
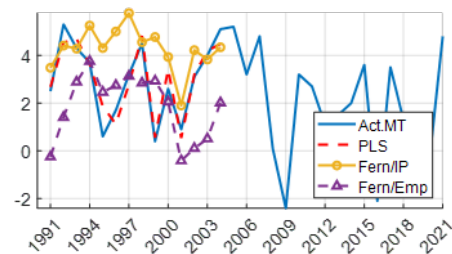
Minnesota



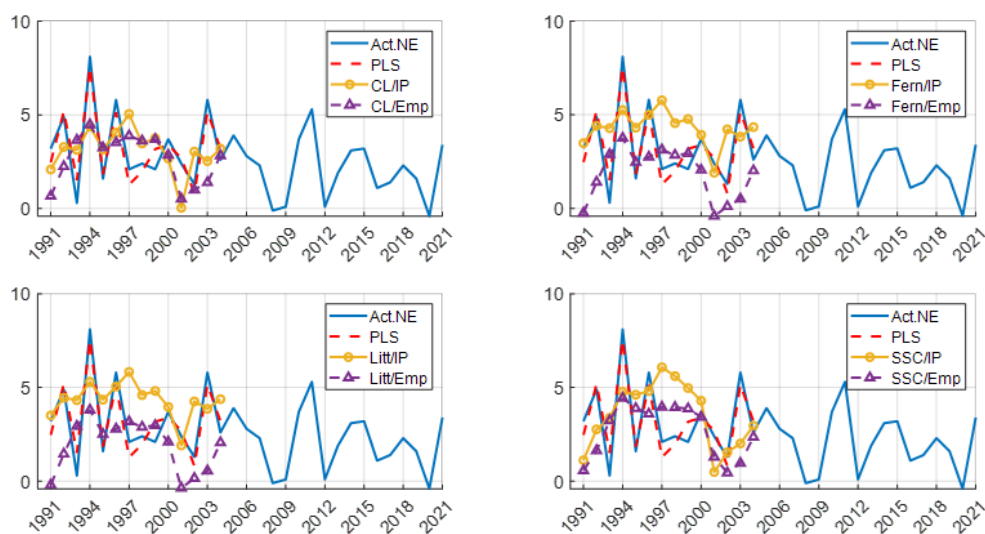
Mississippi



Missouri



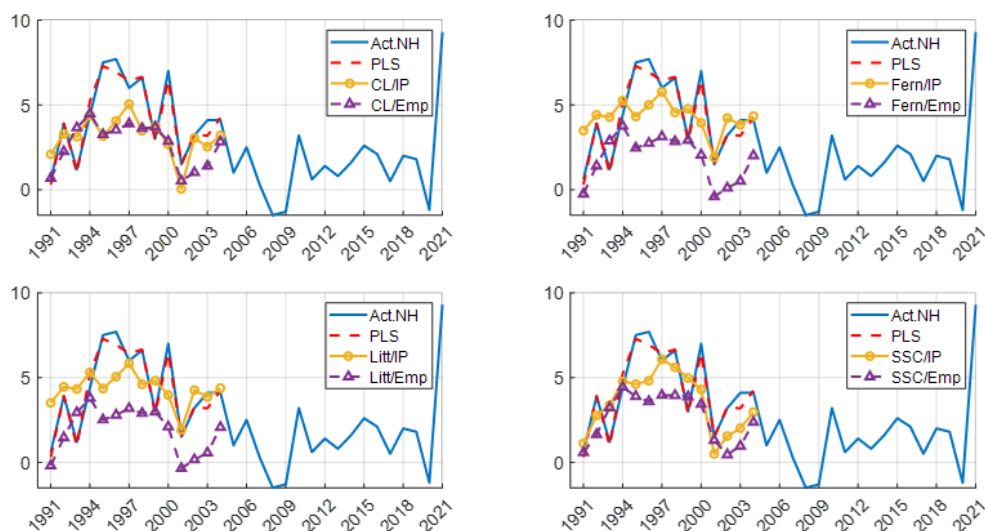
Montana



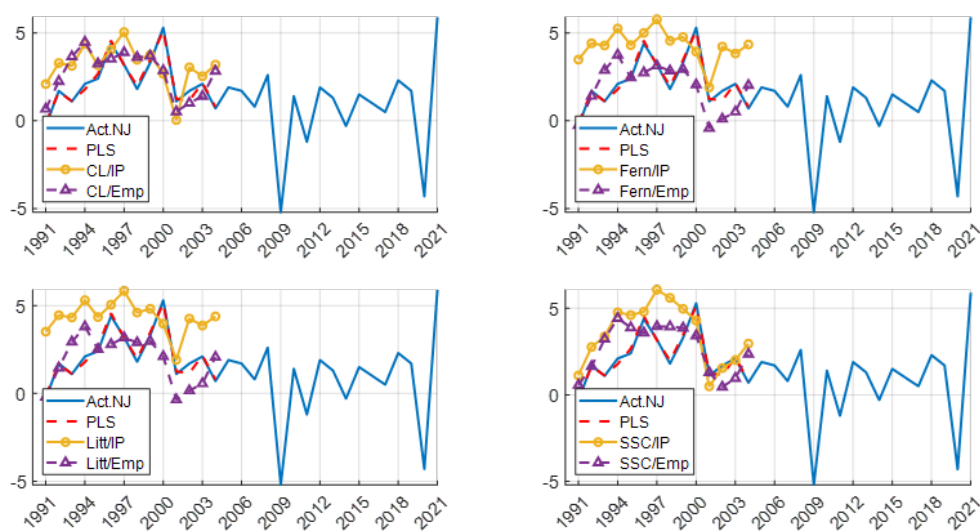
Nebraska



Nevada



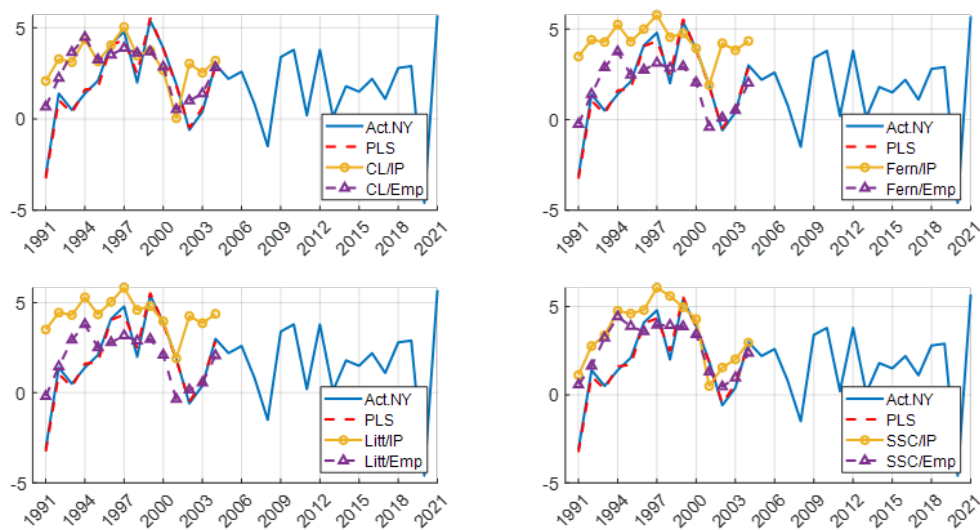
New Hampshire



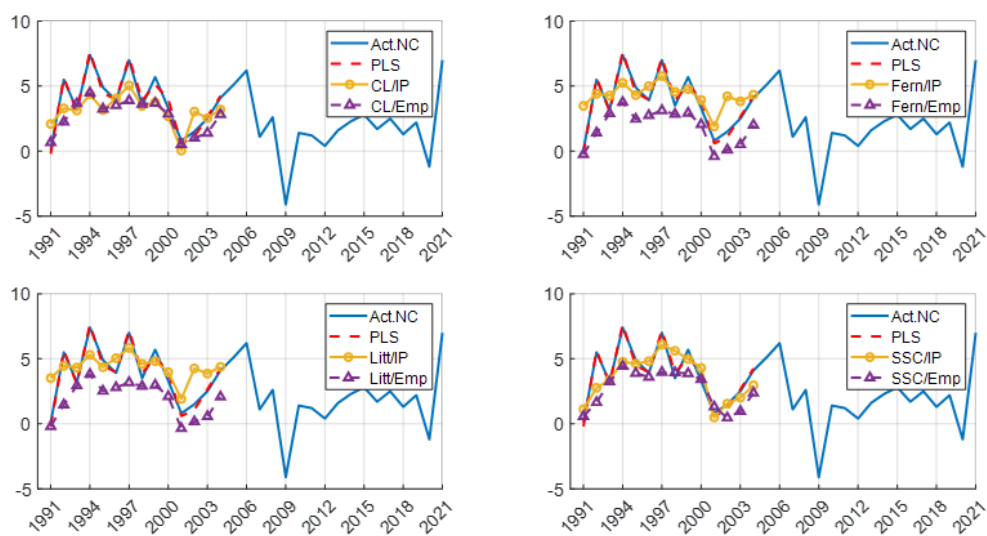
New Jersey



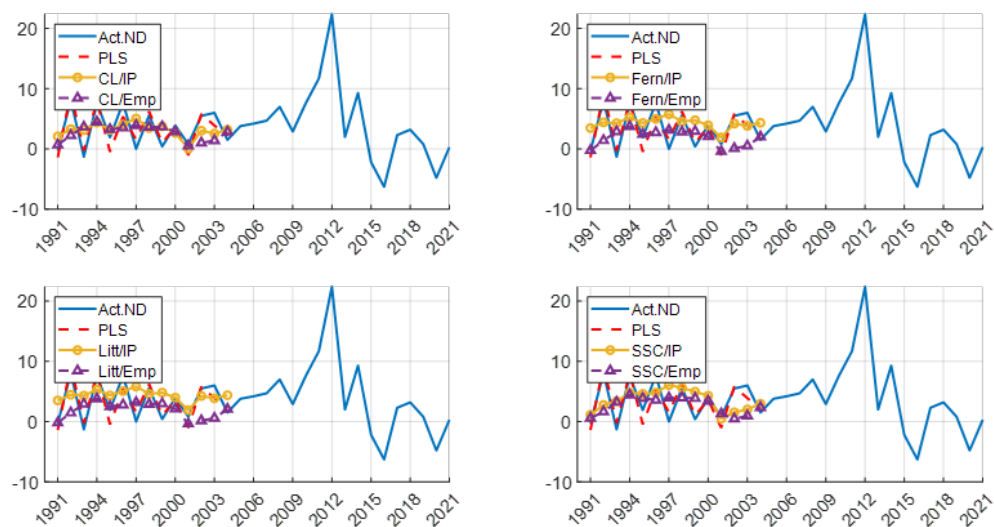
New Mexico



New York



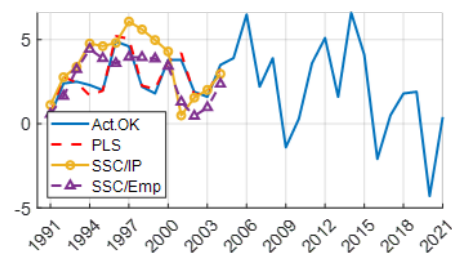
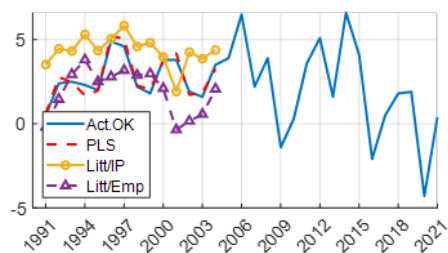
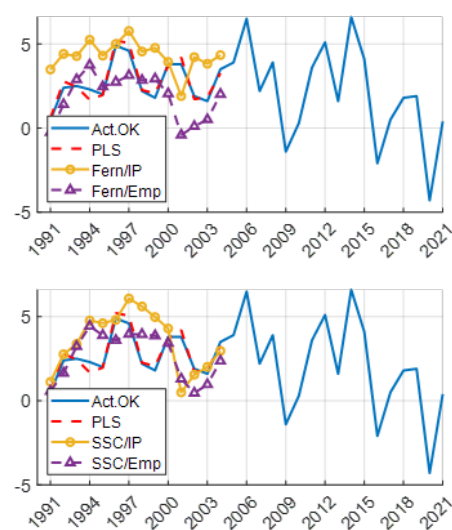
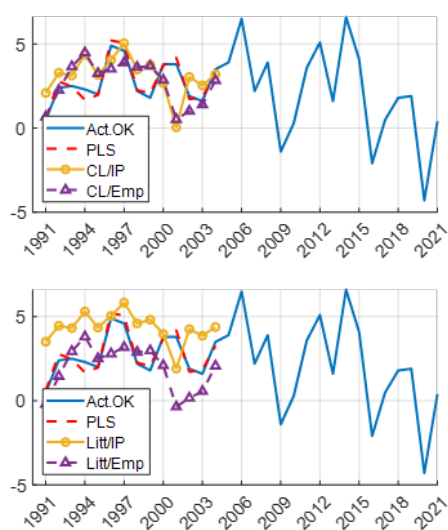
North Carolina



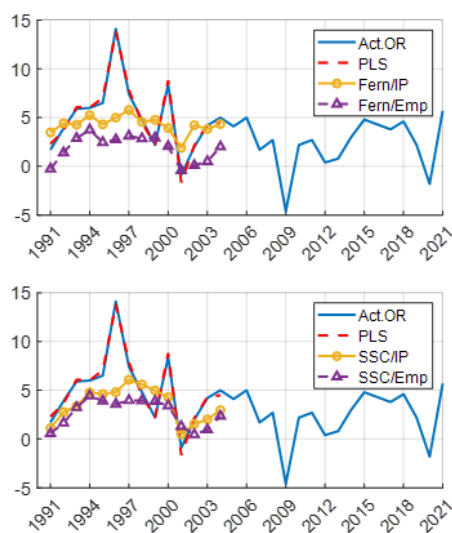
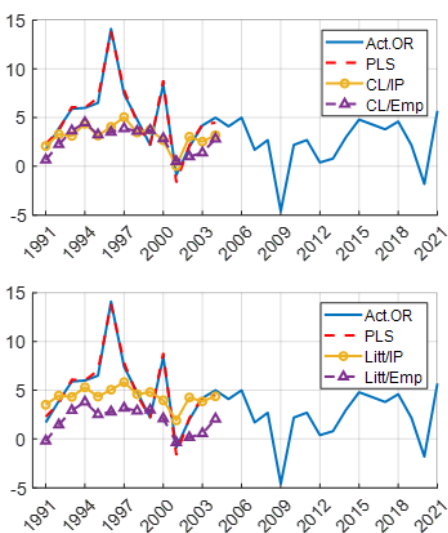
North Dakota



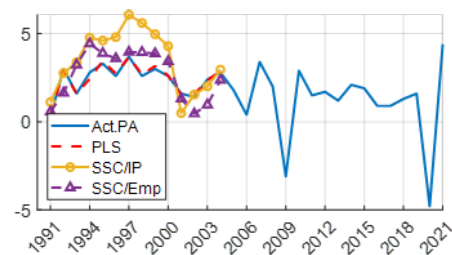
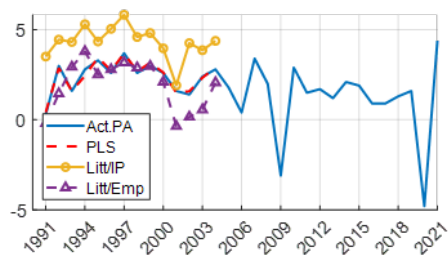
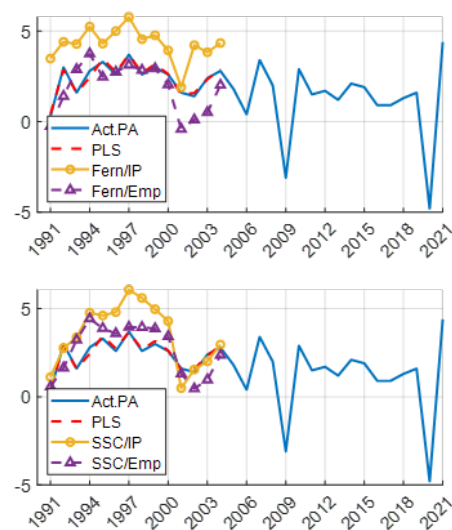
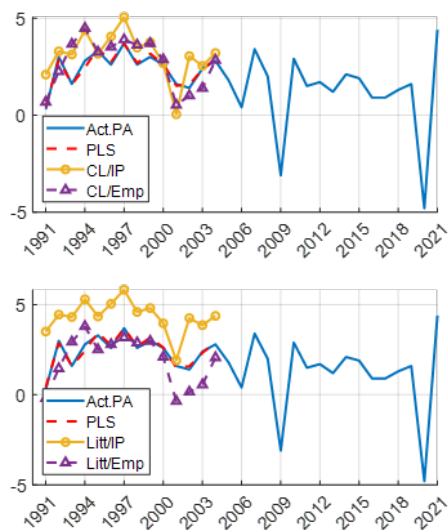
Ohio



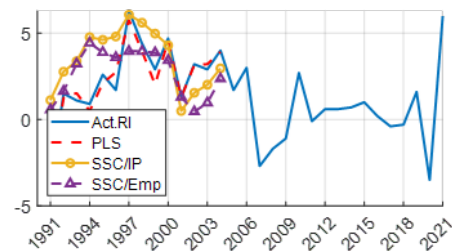
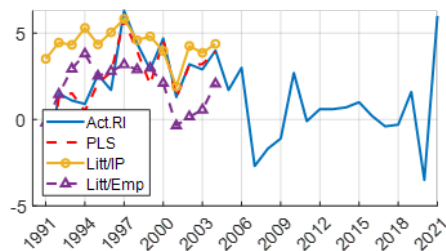
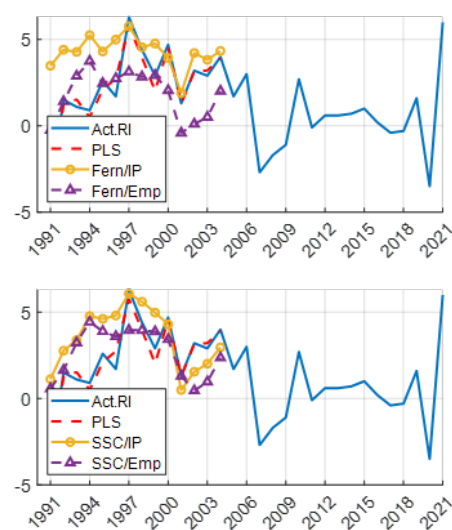
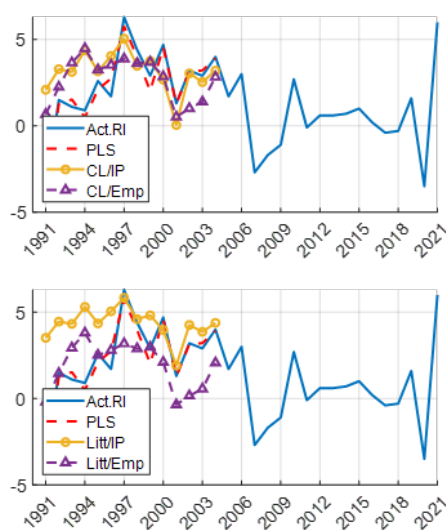
Oklahoma



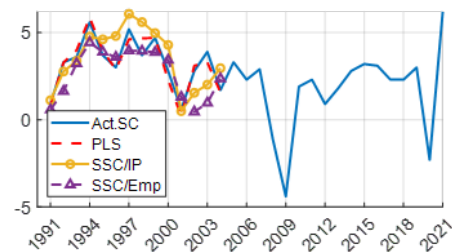
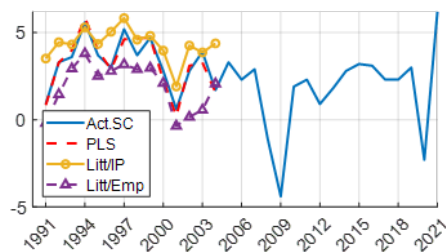
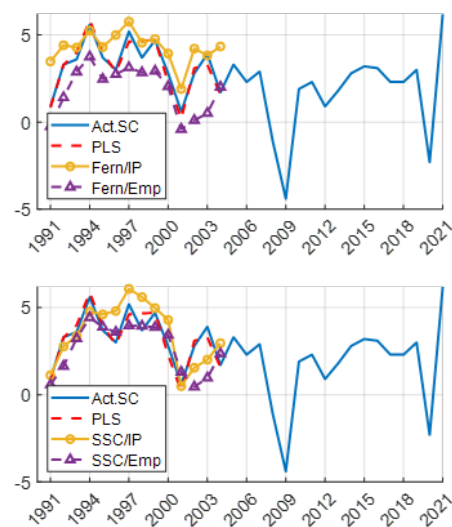
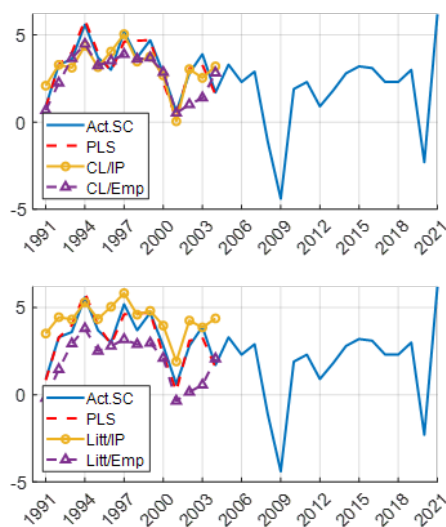
Oregon



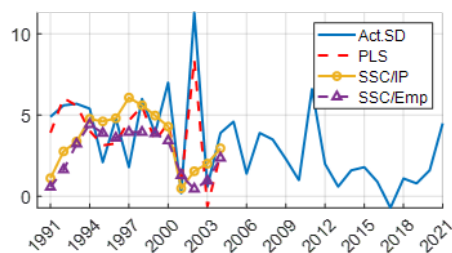
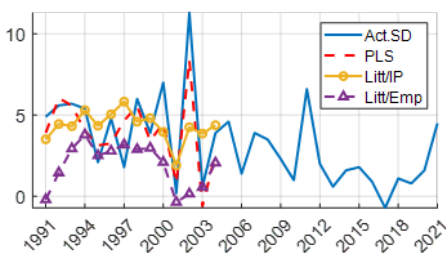
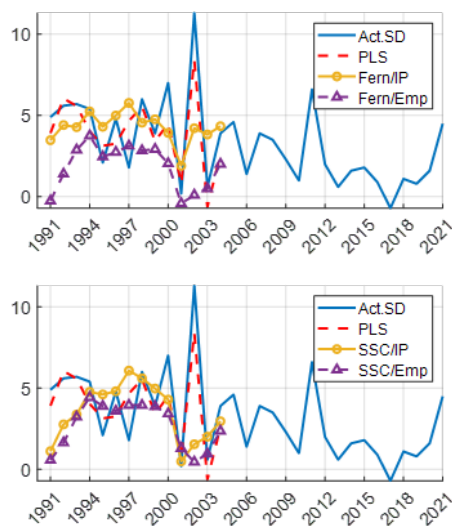
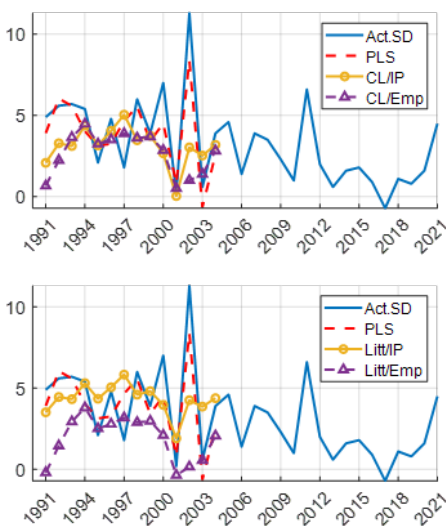
Pennsylvania



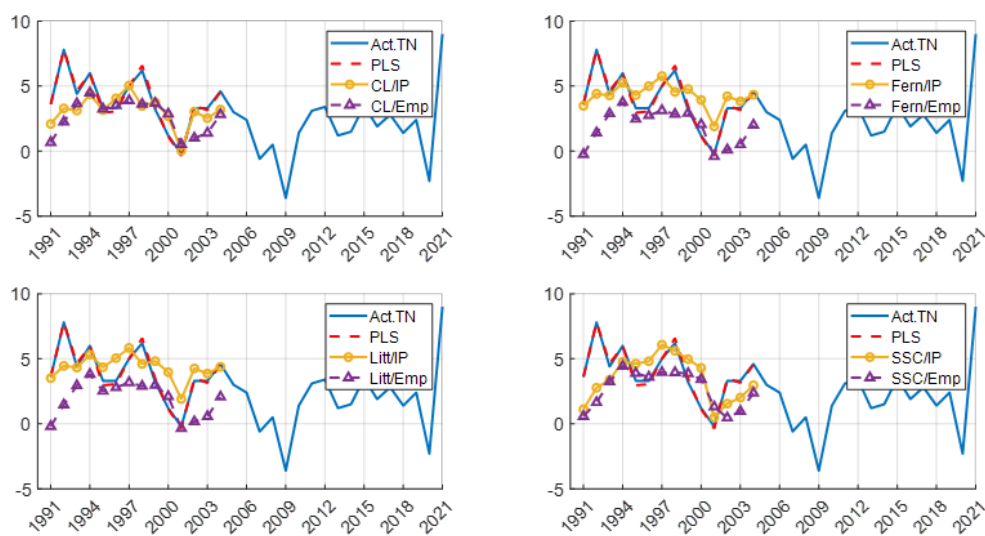
Rhode Island



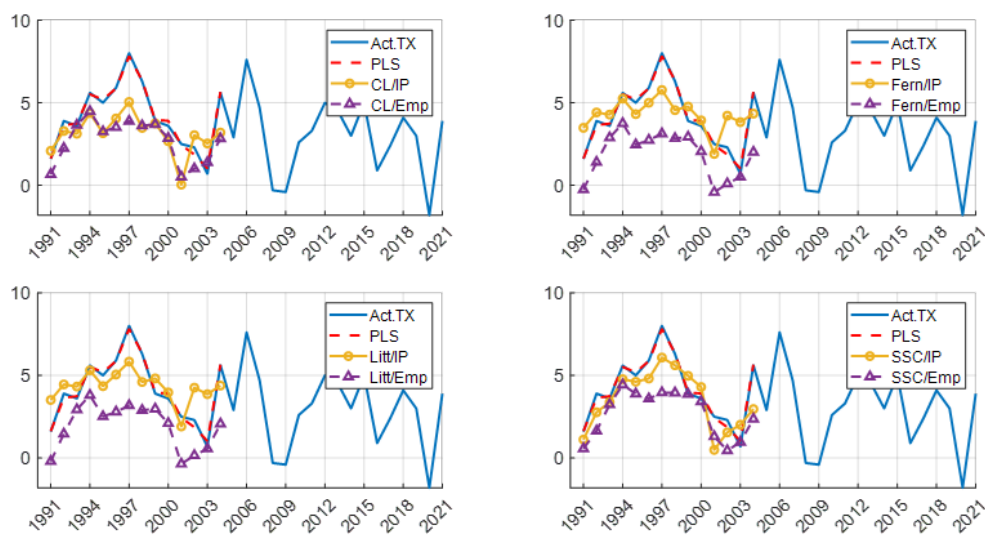
South Carolina



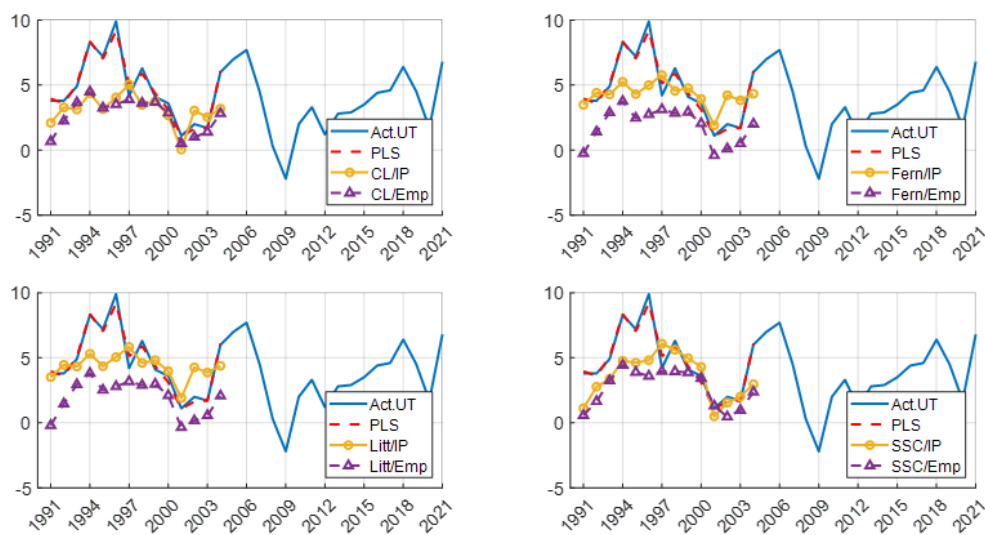
South Dakota



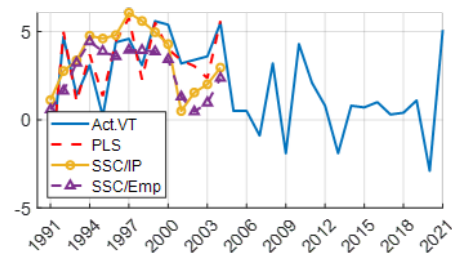
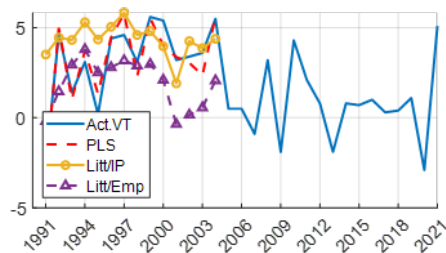
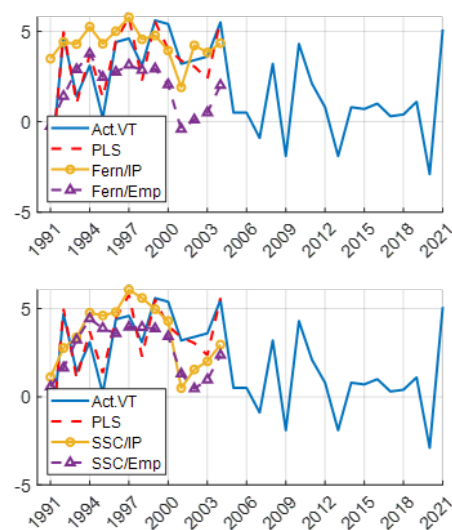
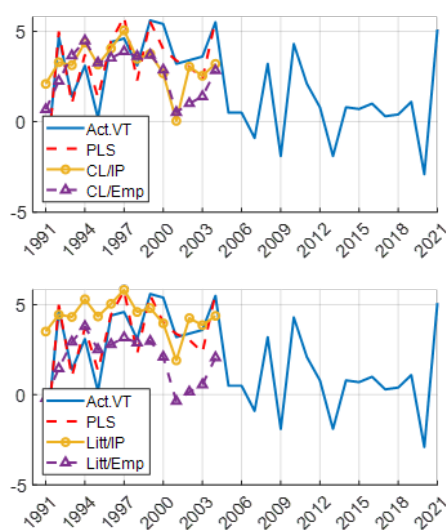
Tennessee



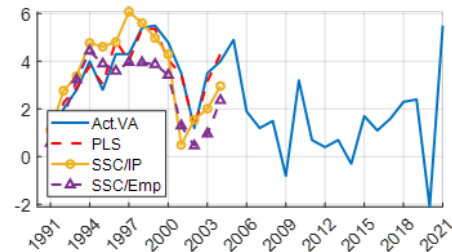
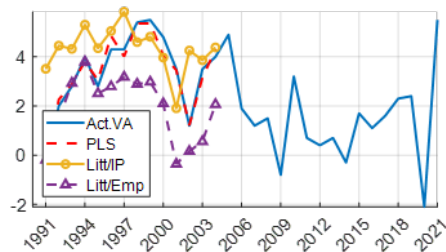
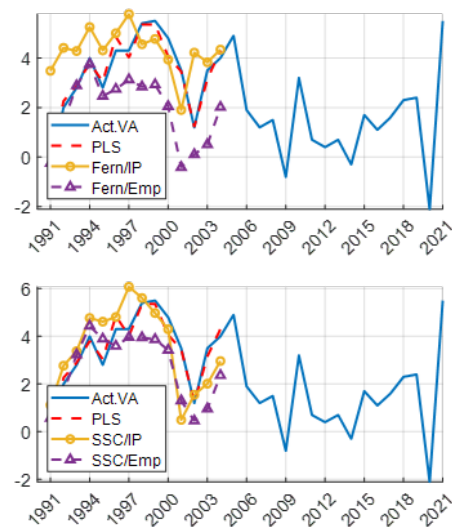
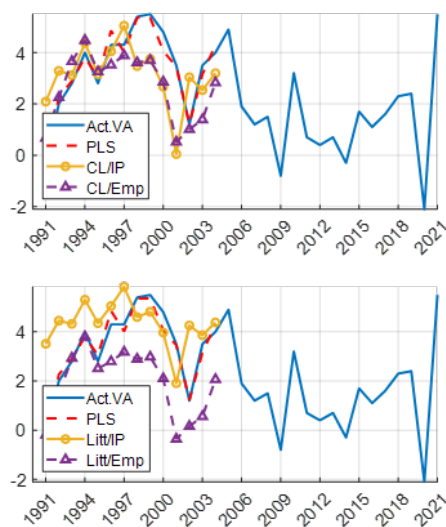
Texas



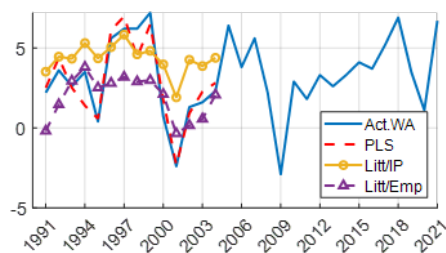
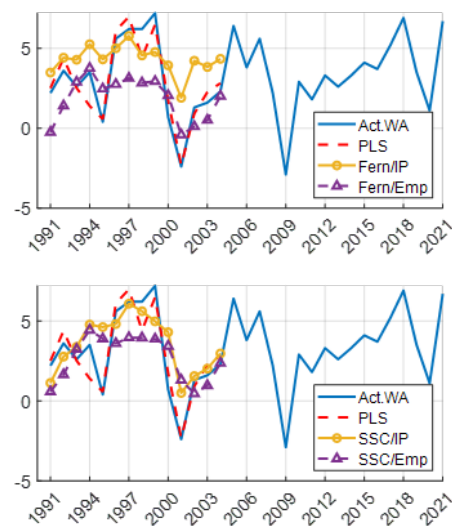
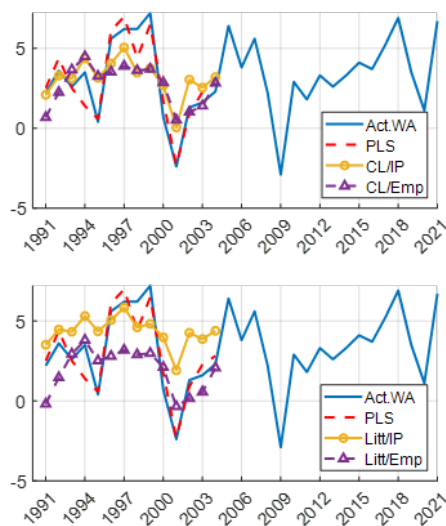
Utah



Vermont



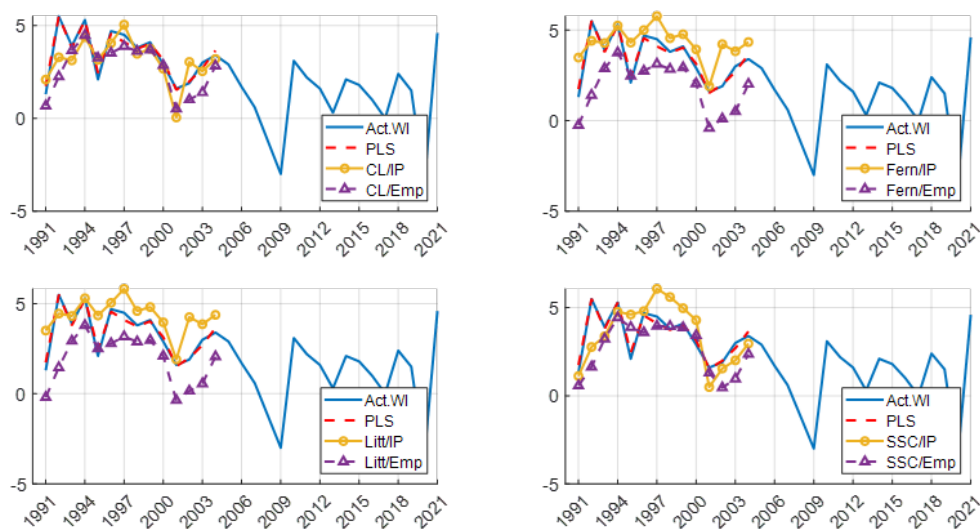
Virginia



Washington



West Virginia



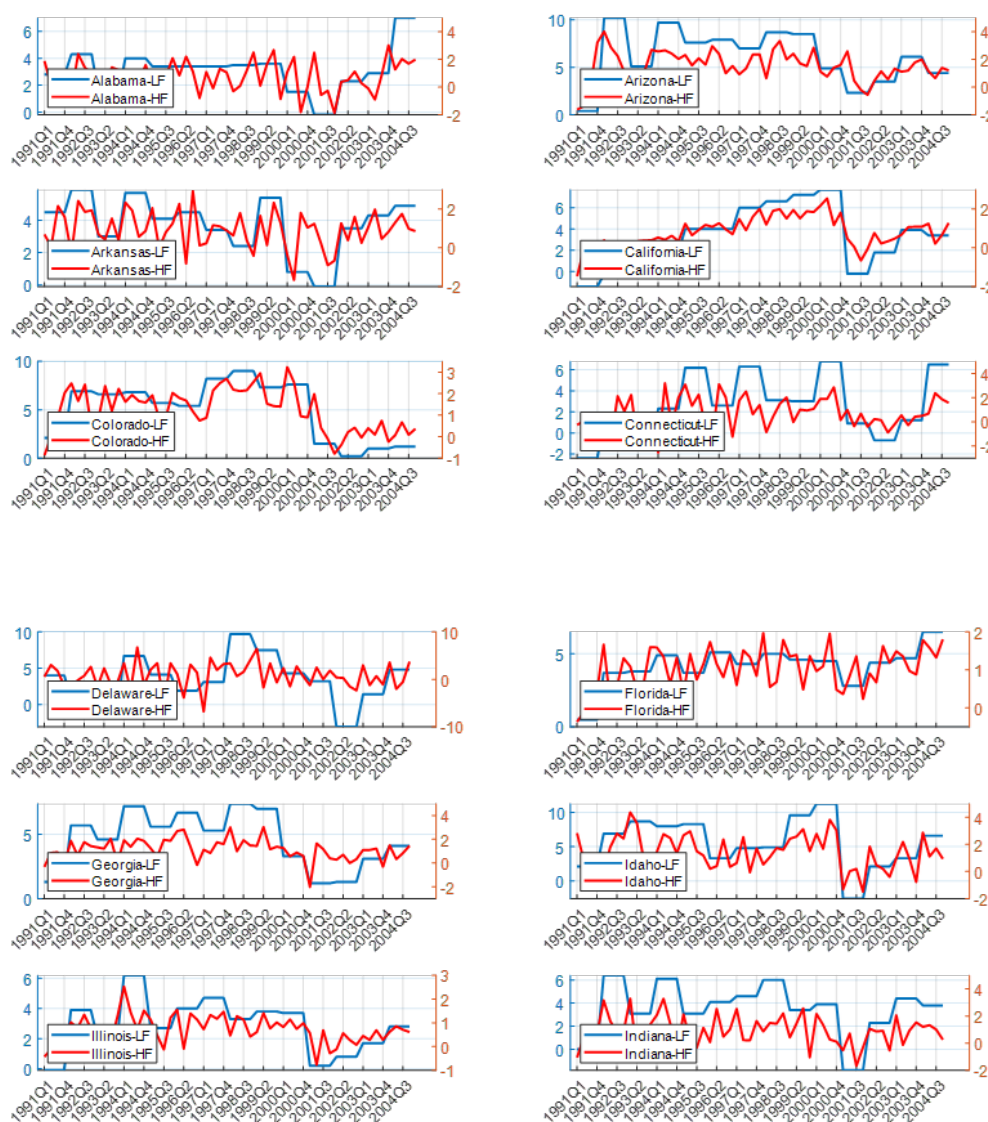
Wisconsin

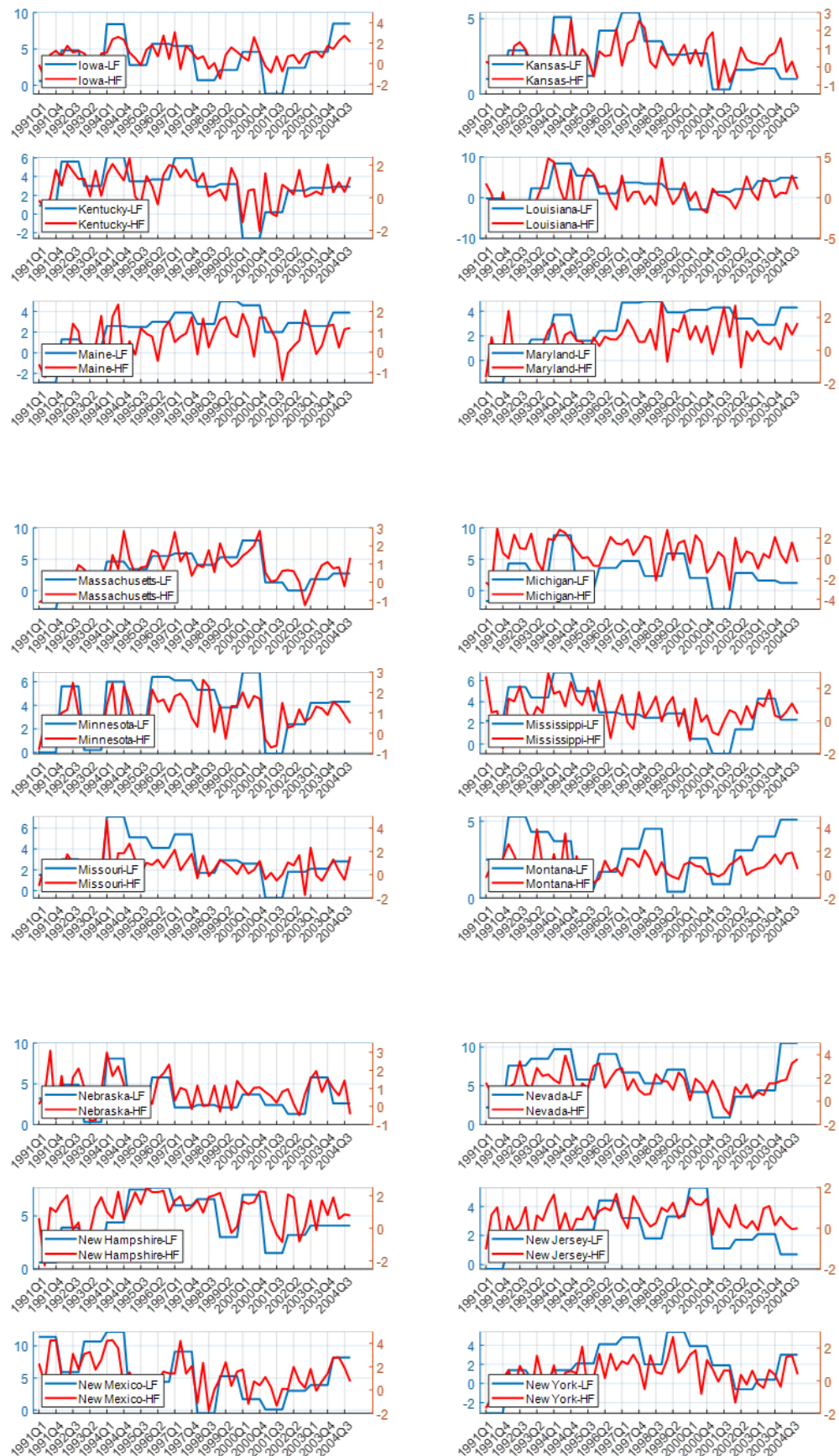


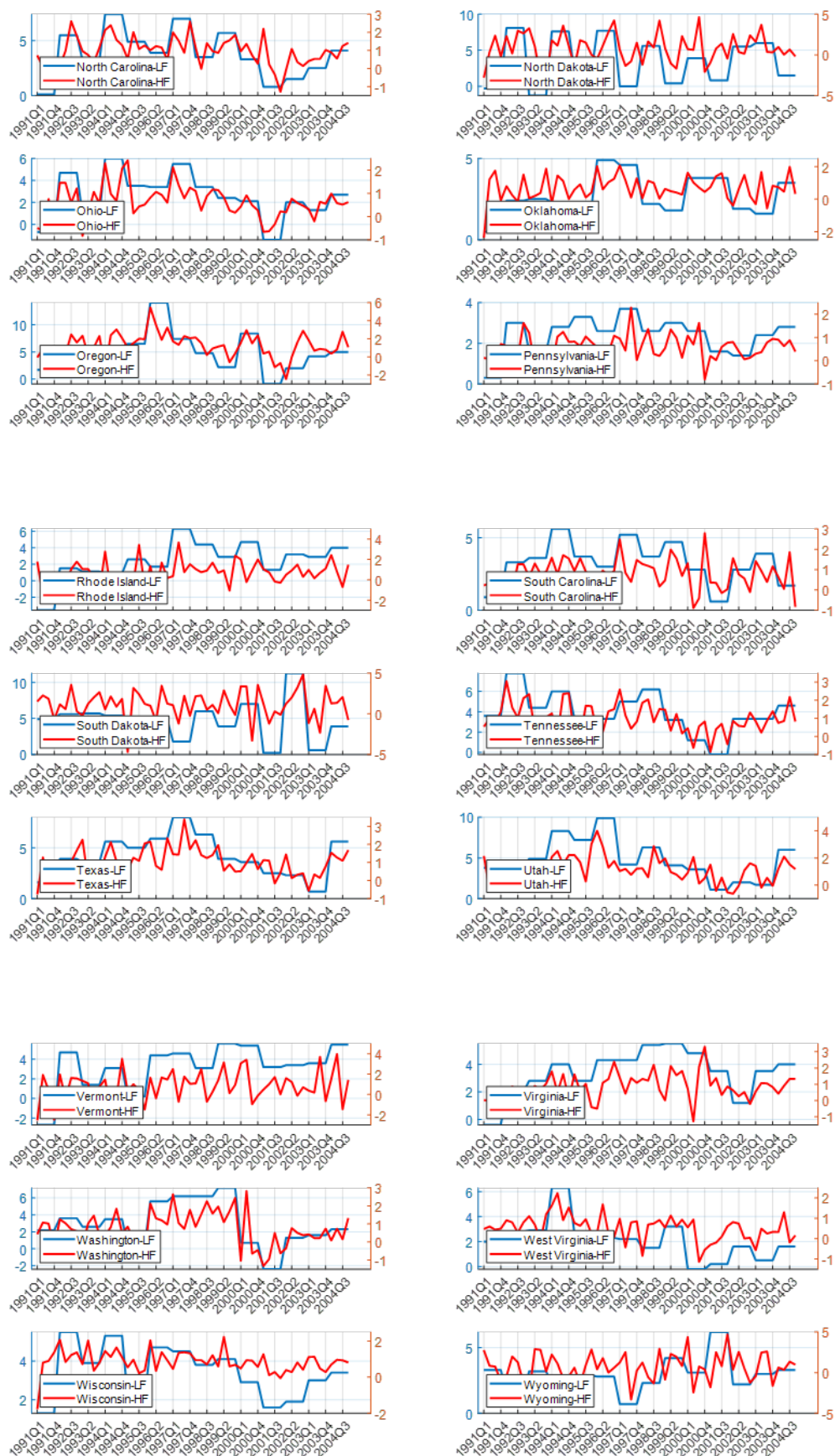
Wyoming

C Estimated High Frequency Time Series

Figure 3.11: Actual LF vs Estimated HF Output Growth







D Structural VAR Model

A Structural Vector Autoregression (SVAR) augments the standard reduced-form VAR by imposing economically motivated restrictions that allow for the identification of contemporaneous relationships and disentangling of “structural” shocks. While a p -lag reduced-form VAR in K endogenous variables can be written as

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u),$$

the SVAR expresses the same dynamics in structural form:

$$A_0 y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I_K).$$

Here A_0 captures the contemporaneous interactions among the variables, and ε_t are the orthogonal structural shocks. Because $u_t = A_0^{-1} \varepsilon_t$ and $\Sigma_u = A_0^{-1} (A_0^{-1})'$, identifying A_0 requires at least $K(K-1)/2$ independent restrictions drawn from economic theory.

Key Econometric Steps

1. Stationarity and Pre-testing. Test each series for unit roots (e.g. ADF or KPSS tests). If non-stationary, determine whether to difference or to model cointegration via a Vector Error Correction Model (VECM).
2. Lag-Length Selection. Estimate the reduced-form VAR for a range of lags and choose p by information criteria (AIC, BIC) or sequential likelihood-ratio tests.
3. Estimate Reduced Form. Fit

$$\hat{y}_t = \sum_{i=1}^p \hat{A}_i y_{t-i} + \hat{u}_t,$$

obtain residuals $\hat{u}_t$ and their covariance matrix $\hat{\Sigma}_u$.

4. Identification of Structural Shocks. Impose sufficient restrictions on A_0 (and/or on long-run impact matrix $\Theta(1) = \sum_{i=0}^{\infty} \Psi_i$) so that A_0 is uniquely determined. Common schemes include:

- *Short-run (Cholesky) zero-restrictions*: assume a recursive ordering so that A_0 is lower-triangular.
 - *Long-run restrictions*: impose restrictions on the long-run effects of certain shocks (e.g. $\Theta_{ij}(1) = 0$).
5. Solve for Structural Parameters. Given $\hat{\Sigma}_u$ and the restriction matrix R , recover $\hat{A}_0$ (and thus $\hat{\epsilon}_t = \hat{A}_0 \hat{u}_t$) by solving

$$\hat{\Sigma}_u = \hat{A}_0^{-1} (\hat{A}_0^{-1})'.$$

6. Impulse Response and Variance Decomposition. Compute the moving-average representation

$$y_t = \sum_{j=0}^{\infty} \Psi_j \epsilon_{t-j}, \quad \Psi_j = A_0^{-1} \sum_{i=1}^j B_i \Psi_{j-i},$$

then derive impulse-response functions (IRFs) and forecast-error variance decompositions (FEVD).

7. Diagnostic Checks. Verify the stability condition (all eigenvalues of the companion matrix lie within the unit circle).

Table 3.8: Network Centrality Parameters

State	Degree		Eigenvector	
	1991-2004	2005-2022	1991-2004	2005-2022
Alabama	13	10	0.0424	0.0255
Arizona	5	11	0.0133	0.0236
Arkansas	7	11	0.0172	0.0230
California	6	11	0.0216	0.0269
Colorado	5	8	0.0139	0.0204
Connecticut	8	7	0.0219	0.0179
Delaware	9	13	0.0261	0.0301
Florida	11	9	0.0321	0.0179
Georgia	8	19	0.0265	0.0441
Idaho	6	6	0.0191	0.0128
Illinois	5	9	0.0113	0.0181
Indiana	5	8	0.0118	0.0293
Iowa	3	8	0.0068	0.0123
Kansas	4	12	0.0074	0.0205
Kentucky	7	9	0.0200	0.0203
Louisiana	6	6	0.0167	0.0113
Maine	7	11	0.0229	0.0239
Maryland	5	8	0.0154	0.0151
Massachusetts	9	8	0.0303	0.0162
Michigan	7	13	0.0219	0.0257
Minnesota	4	6	0.0108	0.0116
Mississippi	6	10	0.0159	0.0264
Missouri	6	9	0.0188	0.0216
Montana	8	11	0.0222	0.0228
Nebraska	5	8	0.0169	0.0180
Nevada	6	2	0.0201	0.0039
New Hampshire	6	6	0.0166	0.0094
New Jersey	10	6	0.0302	0.0112
New Mexico	12	10	0.0384	0.0240
New York	10	7	0.0281	0.0151
North Carolina	7	7	0.0191	0.0135
North Dakota	12	9	0.0377	0.0245
Ohio	12	9	0.0379	0.0209
Oklahoma	2	14	0.0045	0.0342
Oregon	6	9	0.0163	0.0206
Pennsylvania	4	13	0.0080	0.0327
Rhode Island	7	14	0.0200	0.0275
South Carolina	5	11	0.0167	0.0253
South Dakota	8	14	0.0288	0.0326
Tennessee	6	6	0.0151	0.0144
Texas	8	9	0.0227	0.0221
Utah	6	8	0.0190	0.0167
Vermont	8	7	0.0263	0.0173
Virginia	10	9	0.0286	0.0206
Washington	6	8	0.0182	0.0194
West Virginia	9	4	0.0272	0.0089
Wisconsin	5	10	0.0136	0.0221
Wyoming	8	13	0.0237	0.0280

Notes: Degree and eigenvector are centrality parameters from the network analysis indicating interconnectedness and network connectivity, respectively. Sample periods: 1991 to 2004 and 2005 to 2022.

Chapter 4

Topic-Specific Complexity in Monetary Policy Statements: Implications for Liquidity and Volatility in Financial Markets

1 Introduction

Central bank communication plays a critical role in shaping market expectations and influencing economic behaviour. Speeches, inflation reports, and minutes are not merely channels for transmitting data; they are meticulously crafted documents replete with technical jargon, subtle nuances, and at times, intentional ambiguity, which may result in misinterpretations or delayed market responses (Andersson et al. (2006)). Enhancing the clarity and readability of these monetary policy statements, therefore, increases the ratio of meaningful information to background noise, reduces uncertainty among market participants, and enables a more precise interpretation of policy signals.

Nonetheless, despite the recognised benefits of clear policy messages, a substantial body of research has examined central bank communication complexity in advanced economies, with limited attention paid to its manifestations in emerging economies. Moreover, the common practice is to assess the clarity of central bank communication using objective readability metrics, such as the Flesch-Kincaid Grade Level (FK) proposed by Flesch (1948) and refined by Kincaid (1975), on the policy statement as a whole. However, aggregating readability into a single metric for an entire Monetary Policy Committee (MPC) statement obscures the distinct linguistic factors contributing to complexity, thereby implying that all topics addressed by policymakers carry equal importance.

More specifically, research on central bank communication has remained an active area with the literature showing ample documentation of the communication strategies of central banks in advanced economies (AEs) as seen from works of, for example, Andersson et al. (2006), Ehrmann and Fratzscher (2007), Cieslak et al. (2019), Ehrmann and Talmi (2020), Ter Ellen et al. (2022) among others. Over the years, the literature has consistently shown that communication remains an integral component of modern monetary policy, serving as a critical tool along with other traditional central bank tools. Geraats (2009) and Casiraghi and Perez (2022) confirm the rise in central banks' communication regardless of the monetary framework. The latter does so by adding that the most important message of central bank communication is its policy regime which is typically enough to reveal the logic behind policy decisions.

Furthermore, (Binder (2017)) confirms that various aspects of central bank communication have so far been studied, for example, the effect of monetary policy communication on the bank's own ability to meet its own objectives (Woodford (2005) and Ehrmann and Fratzscher (2007)), impact of central bank communication on: market expectations and understanding (Rosa and Verga (2007) and Hayo and Neuenkirch (2010)), interest rates (Kohn and Sack (2003) and Demiralp and Jorda (2004)), yields (Gürkaynak et al. (2004)), bonds liquidity and stock markets (Kohn and Sack (2003), Ehrmann and Fratzscher (2005), Anderson (2010), Chung et al. (2013), Smales and Apergis (2017a) and Hayo et al. (2022)), exchange rate volatility (Evans and Speight (2010), Rosa (2011a), and Vyshnevskiy et al. (2024)). From above, it can be seen that the literature has ample evidence that central bank communication influences financial markets, and recent studies, for example, Smales and Apergis (2017a) also find evidence that the linguistic complexity of monetary policy communication, in particular, is among the factors that significantly affect the liquidity of financial instruments on the day of policy communication.

So, although the literature on monetary policy communication is vast, an important omission that may possibly explain the scarcity of research in the EMEs and DCs is that, most central banks in these regions only started modernising and communicating regularly in the last two decades. So, a large part of the communication irregularities may be attributed to their inexperience as they are still just navigating the structural frictions of transitioning from opaque to transparent policy frameworks. For this reason, I believe that there are many nuanced adverse characteristics of monetary policy communication in these new adopters of modern monetary policy frameworks. Therefore, in this study, I attempt to build on the existing literature and explore an issue that is closely related to the effect of communication complexity on financial markets. That is, the existing research to date has only concerned itself with the understanding complexity of monetary policy communication at an aggregate level, for example, the level of the whole policy statement. So, I want to investigate the linguistic complexity of the specific topics in the policy statements and their relationship with financial markets in South Africa within a

tight window, in particular, on announcement day¹. This current research is closely related to the previous work of, Anderson (2010), Jansen (2011), Rosa (2011a), Bulíř et al. (2013), Chung et al. (2013), and Smales and Apergis (2017b), and aims to make a direct contribution to this literature. The study also connects to a similar methodology proposed in Barbaglia et al. (2023) and Barbaglia et al. (2024), which introduce FiGAS for text extraction and sentiment analysis using unigrams and bigrams. Therefore, I start from the premise that the linguistic complexity of individual topics within monetary policy documents have differentiated influences on financial market variables.

I am especially motivated by two existing blind spots in the literature. First, until now, the literature analyses complexity of the central bank communication and its impact on the economy based on the overall policy document rather than the linguistic complexity of the specific topics covered. While this approach has provided valuable insights, it captures only the overall complexity of a document, which implicitly assumes that all topics addressed in central bank communications are understood equally, hold equal importance for market players and that their clarity has uniform significance. Second, as indicated in the opening parts above, I have observed the lack of research on emerging and less developed financial markets which creates a knowledge gap on understanding the effectiveness on monetary policy communication, in particular, due to complex communication of specific topics. My view is that this gap requires a dedicated, general but granular inquiry that is representative of the overall policy communication.

Therefore, in response I present this study that distinguishes itself in the following specific and significant ways: first, the study develops an innovative dictionary-based algorithm that isolates topic-specific sections from each document by employing designated keywords and text segments as inputs. This method enables me to apply the readability index at a granular, per-topic level and construct a distinct readability series for each topic. My approach demonstrates that linguistic complexity not only evolves over time across different statements but also varies significantly within individual documents that address multiple topics. Second, unlike previous studies that focus on the complexity of the entire

1. Within the region, the South African financial markets that are considered the most developed and liquid in Africa. South African financial markets are also considered developed by global standards reflecting credible and independent monetary policy, diversified productive sectors and strong financial institutions.

policy document, the current study examines whether and how the complexity of each specific topic covered in the policy document correlate with financial market variables; and lastly, by focusing on South Africa, a leading economy in Sub-Sahara Africa (SSA), this study offers a unique and valuable perspective to this active research subject. I, therefore, initialise with the following hypotheses:

H1: *Central bank communication complexity can be divided into two components: the overall linguistic complexity of the entire document and the complexity of the specific topics it covers, which may vary across different policy documents.*

I argue that central bank policy documents cover various aspects of the four blocks of the economy, namely, fiscal, monetary, external and real sectors. For this reason, I expect that the language and accordingly, that complexity of the policy document will vary depending on the complexities of the individual topics covered and over time.

H2: *The greater the complexity of specific subjects covered in the monetary policy statements, the lower the expected liquidity of the 10-year government bond on the day of the announcements.*

H3: *The greater the complexity of specific subjects covered in the monetary policy communication, the higher the volatility of liquidity of the 10-year government bond on the day of announcements.*

My dataset comprises 151 Monetary Policy Committee (MPC) statements issued by the SARB between October 1999 and July 2023 and total word count of 300,323. Latent topics are extracted from central bank statements using Blei et al. 2003's Latent Dirichlet Allocation (LDA). My topic identification findings reveal that SARB communications generally address five primary topics: *Inflation*, *Forecasting*, *Monetary Policy*, *Financial Markets*, and *Economic Activity*. Each topic occupies a different proportion of a statement and is articulated in a distinct tone, such that the overall readability is largely determined by the more heavily emphasised topics. Notably, while there has been a substantial improvement in the overall readability of SARB communications over time, this trend is primarily driven by a decline in the complexity of the *Forecasting* topic, which has also come to occupy an increasing share of the statements. In contrast, readability scores for the remaining topics have remained relatively stable over time.

To determine the impact of topic-specific linguistic complexity, I construct an ultra-high frequency series of the liquidity mean using tick-by-tick Bid and Ask data for the 10-year South Africa government bond, the country's most liquid bond indicator. I then compare the distribution of its liquidity on MPC announcement days with that observed 10-trading days prior. This is done for periods in which the level of complexity is low and when it is high on both the 10-day control and the announcement day. Consistent with the literature, I observe that greater overall complexity in central bank statements tends to correlate with reduced market liquidity. However, I also observe that liquidity is generally low for the government bond on the announcement day relative to the control days. Density distributional comparisons reveal relatively lower liquidity in the 10-year bond particularly, when linguistic complexity is high.

However, the aggregate document complexity effect on liquidity masks a clear divergence between topics. That is because a granular approach that identifies key topics and their respective complexities suggest that when forward-looking discussions, covering projections of growth, inflation, and potential upside or downside risks, are presented in a complex register, investors tend to pull liquidity sharply - based on the mean comparisons. In contrast, complexity in sections addressing the current economic and financial environment has only a muted impact on liquidity. These findings suggest that to bolster market stability, policymakers should prioritise simplifying the language used in forward-looking segments of their communications rather than uniformly simplifying the overall readability. The analytical results are based on high-frequency intra-day data within a tight window on the announcement day confirm this relationship.

I also establish that the volatility of liquidity is also sensitive to linguistic complexity of central bank communication as the market experiences greater spread variability. The results show that there is relatively higher and longer lasting volatility of up to two hours (up to close of business) in the bond market in the time in which the statements are complex on announcement day compared to one hour when the statements are easy. High

linguistic complexity of forecasting, monetary policy and financial markets are associated with right skewed spreads in liquidity variability bond market suggesting high uncertainty under these conditions. This investigation helps to highlight the value of topic specific clarity in understanding the influence of monetary policy communication.

My contribution to the body of knowledge is to demonstrate that improving the overall clarity of central bank statements does not automatically mitigate market responses. I show that complex language on specific topics can still trigger market surprises and therefore, need to be accounted for. As a result, my findings suggest that the SARB should devote more attention to refining the discussion of particular topics within MPC meetings.

To the best of my knowledge, this is the first study to attempt an analysis of the per-topic complexity of central bank communications and explicitly link the complexity of individual topics discussed to fluctuations of the bond liquidity in South African financial markets.

The rest of this study is organised as follows: Section 2 provides an overview of monetary policy in South Africa; Section 3 presents the literature review; Section 4 details the data, methodology and descriptive analysis; Section 5 discusses empirical results; and lastly, Section 6 offers a conclusion.

2 Monetary Policy in South Africa

2.1 About the SARB

The SARB is the central bank of South Africa. Like many central banks around the world, its primary mandate is to defend the value of the local currency (Rand or ZAR) with the aim of achieving balanced and sustainable economic growth. Some additional responsibilities include issuing the local currency, regulating and supervising financial institutions, managing the payments system, and managing the exchange rate control systems among others².

2. I present brief details on key features in Appendix A, but comprehensive description can be found on the bank website <https://www.resbank.co.za/en/home/>

The bank migrated from Monetary Aggregate Targeting (MAT) to Inflation Targeting (IT) monetary framework in 2000³. Under this monetary policy framework, the SARB uses interest rate (repo rate) adjustments as its primary policy instrument. Since changing the operational monetary framework, the bank has had 3 governors including the current. The governors are Tito Mboweni (1999–2009) whose leadership is widely considered to have been synonymous with the adoption of inflation targeting in 2000, adopting a framework that is currently in place. Gill Marcus (2009–2014) is credited with strengthening the SARB’s communication strategy, and improving the clarity of the policy decisions so that the public would easily comprehend the policy stance. Lesetja Kganyago (2014–present) is considered to place emphasis on the financial stability and management of systemic risks.⁴

With regards the MPC, the governors lead the committee which is responsible for setting the policy rate (the repurchase rate). The MPC at the SARB is comprised of seven members, namely, the Governor, the three deputy governors and senior officials appointed to the committee by the Governor.

Since one of the key defining features of the current framework is its emphasis on transparency and communication, the SARB has accordingly increased its transparency, publishing policy decisions six times a year. In the current schedule, the MPC statement is issued six times in each year in January, March, May, July, September and November⁵. The SARB holds its press conferences at 3:00 PM local time and typically lasts no more than an hour on announcement day. Each MPC meeting concludes with a policy statement and press release which typically ends with a Q & A session, reflecting the SARB’s commitment to clear communication. The SARB broadcasts the live press conference and publishes the videos on its website⁶.

3. The IMF classifies Ghana, Kenya, South Africa and Uganda as the only inflation-targeting countries in SSA. However, for more details on the history and other past monetary policy strategies of the SARB, see <https://www.resbank.co.za/en/home/about-us/history>

4. Visit official website for other SARB heads. See <https://www.resbank.co.za/en/home/about-us>

5. An exception is observed in 2009 during the Global Financial Crisis when it published nine MPC statements and in 2020 during Covid-19 crisis when it published 7 MPC statements.

6. The SARB publishes its monetary policy statements on its website immediately following every monetary policy meeting and the governor reads the statement during the press conference

The implementation strategy of the current framework requires the SARB to target consumer inflation of between 3% and 6% in order to stabilise the economy and I observe considerable communication on this subject in the policy statements. It is evident by the amount of space dedicated to the “inflation subject” in the official statements which covers inflation outcomes, inflation forecasts and risks to inflation forecasts. The SARB’s commitment to the secondary goals is also evident during the 2008–2009 Global Financial Crisis and the COVID-19 crisis, where it responded to both crises by cutting the policy rate and introducing liquidity measures to support economic growth and financial stability during the COVID-19 crisis⁷.

2.2 The SARB MPC Policy Communication

The SARB’s main communication tool for monetary policy are the MPC statements issued following every MPC meeting. The monetary policy statements typically provide valuable information on how decisions are arrived at, showing how the bank is adapting to changing economic challenges. In particular, the statements cover important information about the domestic and global growth concerns, inflation, international commodity prices and the exchange rate, employment, trade conditions, key risks to forecasts, key considerations, policy decisions and the framework. As such, the statement covers various specific topics. The monetary policy statements are published on pre-scheduled dates which are usually announced in the last quarter of the preceding year.

Over time, the language in the SARB’s monetary policy statements has evolved in part as a reflection of the changing economic challenges. For example, regarding forecasting which is an essential aspect of communication under the inflation-targeting regime—the phrase “inflation forecast” did not appear in SARB policy statements until the mid-2000s. Meanwhile, the term “inflation outlook” was consistently used between 2006 and 2010, after which the language shifted back to the use of “inflation forecast” and “inflation forecasts”.

7. The SARB cut the repurchase by 650 basis points between November 2008 and November 2010. In the first four months of 2020, it cut the repurchase rate by 225 basis points. See <https://www.resbank.co.za/en/home/about-us/history>



Figure 4.1: Governor-Specific Word Clouds

Furthermore, as indicated above, it is evident from the statements that changes in the vocabulary or terminology used coincide with the state of the economy or key economic issues faced by the SARB. For instance, during the early 2000s, rising unit labour costs posed a significant risk to projected inflation and was a key driver of inflation outcomes (by raising the production cost which manifested in relatively higher consumer prices), resulting in frequent references to “labour cost” in communications of inflation during the time which has since disappeared from the statements in the last decade. And under Governor Kganyago, for example, the word clouds in Figure 4.1 prominently feature terms like “forecast” and “risk” reflecting the current depth of the SARB’s inflation-targeting framework. These terms highlight the forward-looking nature of the policy framework, while emphasising the role of predictive analysis and risk management as key considerations for monetary policy decision-making. Coco, Viegi et al. (2020) make similar observations noting that, for example, Governor Kganyago’s communication put greater emphasis on expectations and inflation forecast, signaling a communication that is more forward-looking in nature. The evolution of word clouds also demonstrate the SARB’s responses to addressing specific macroeconomic challenges and its commitment to transparent and effective communication.

A typical Monetary Policy Statement of the SARB starts with global economic developments and ends with key risks to the inflation outlook followed by the policy decision summarising the interest hikes or cuts in response to inflation or recessions. Here are excerpts from one MPC statement released on 28 March, 2019, as an example:

“Since the previous meeting of the Monetary Policy Committee, global growth concerns have increased, with particular weakness visible in some of our major trading partners.” (...)

⋮

“The overall risks to the inflation outlook are assessed to be more or less evenly balanced. Key upside risks are rising administered prices including electricity and water tariffs, rising domestic food prices in the outer years and higher international oil prices. Downside risks include lower global inflation and an extended period of monetary accommodation in advanced economies.”

“Against this backdrop, the MPC unanimously decided to keep the repurchase rate unchanged at 6.75% per year. The Committee continues to assess the stance of monetary policy to be accommodative.”

3 Literature Review

Central bank communication involves actively disseminating information regarding their evaluation of the current state of the economy, the policy decisions they have made, and the future reaction function they will implement in order to influence the expectations of households, businesses, and financial markets. This is done in order to influence the expectations of the economy. As opposed to relying solely on adjustments to its policy tool, which is typically the short-term interest rate, a central bank has the ability to influence expectations regarding the future trajectory of rates, inflation, and growth through the utilisation of verbal guidance, forward guidance, published forecasts, speeches, press conferences, and published forecasts (Blinder et al. (2008)). The bank is able to increase the efficacy of its interventions, lower the amount of ambiguity that exists, and reinforce its accountability to the public and elected authorities so that it may achieve these goals through the process of making its policy standards and objectives more transparent.

It was not until a few decades ago that the central bank started putting in place a communication infrastructure that was organised in a methodical manner. In the years leading up to the 1990s, the majority of central banks would only publish annual reports or make brief announcements after the conclusion of their meetings. For example, despite the fact that it had started publishing detailed minutes five years after the meetings of the Federal Open Market Committee (FOMC) of the Federal Reserve, the Federal Reserve of the United States did not issue a formal statement until 1994, as stated by Mishkin (2007). This was the case despite the fact that it had begun providing this information. Immediately after its very first meeting in June 1998, the European Central Bank (ECB) started holding its now-familiar monthly press conferences and issuing “Introductory Statements” along with comprehensive economic projections (Ehrmann and Fratzscher (2007)). These practices have since become part of the institution’s standard operating procedure. The European Central Bank was compelled to uphold a stringent legislative requirement for transparency, and this was done in order to represent that mandate.

Another significant event that took place in 1997 was the Bank of England achieving operational independence. This was a significant event that occurred throughout the year. A full explanation of the decision-making process was included in the Monetary Policy Report, which was released on a quarterly basis. In addition, the report included estimates of inflation and an analysis of the economy. In their research, Mishkin (2007) and Goodfriend and King (2005) discovered that other central banks that target inflation quickly adopted this method of often presenting predictions and minutes. To re-establish credibility in the wake of the financial volatility that happened in the early 2000s, other authorities from emerging markets, like the central bank of Japan, followed suit. This was done in an effort to regain trust.

At the same time as these institutional transformations were taking place, there was a surge in the amount of academic interest in communication between central banks. The early theoretical study that addressed the ways in which transparency could promote well-being was conducted, as stated by Svensson (2006). This action was taken with the idea of better aligning the expectations of private individuals with the goals of policymakers. Communication shocks, which are sudden shifts in the tone or content of statements, have

the potential to influence stock prices, exchange rates, and long-term yields by amounts that are either equal to or greater than those caused by outright rate changes (Bernanke and Kuttner (2005) and Ehrmann and Fratzscher (2007)). This is the conclusion that can be drawn from empirical research that has showed that communication shocks have the potential to influence these factors. When it comes to communication shocks, unexpected variations in the tone or content of utterances are what distinguish them from other types of communication shocks. As a consequence of these discoveries, communication has been recognised as an instrument that is vital to the process of policymaking.

Furthermore, the literature shows that the techniques that central banks have evolved throughout the course of time have been established with the intention of finding a balance between flexibility and transparency. Markets are left wondering when there is an excessive amount of ambiguity in forward guidance, and an excessive degree of clarity can unnecessarily limit future policy responses (Mishkin (2007)). In the event that economic conditions undergo a major shift, markets are left guessing. Modern best practices maximise transparency while maintaining the flexibility to change course when it is essential. This is accomplished by the utilisation of behavioural insights, such as the use of language that is straightforward and consistent and the periodic summaries of important uncertainties. This can be achieved by having the ability to make course corrections when it is absolutely necessary to do so. After everything has been taken into account, communication between the central bank and the general public has progressed from being almost non-existent to being a diverse toolset that is based on evidence and is now an essential component of modern monetary policy.

Following from this, the most frequently cited rationales for central bank communication to financial markets and other stakeholders are well documented and published by the International Monetary Fund in 1999. These rationales continue to serve as a driving force behind contemporary central bank communication all over the world ⁸.

8. Della Valle et al. (2022) defines major target audiences of monetary policy communication. These audiences include, but are not limited to, the general public, the media, the government, parliament, players in the financial market, and central bank “watchers”. The primary justification for communication between the central bank and the public is that it is a sound governance practice that works to guarantee that the monetary authority continues to be accountable, particularly in situations where it is allowed independence. In the second place, it is considered that communication can improve the efficiency of

In fact, monetary policy communication has become so integral to modern monetary policy - both MAT and IT central banks have on boarded and embraced communication, committing to publish their policy decisions regularly. Geraats (2009) confirms this by showing the rise in central banks' communication about their decisions and actions. They show that accordingly, transparency has been increasing for both inflation targeters and monetary targeters regardless of the income levels of the host country. In line with Geraats (*ibid.*), Casiraghi and Perez (2022) also recently showed that regardless of the monetary framework, the key message of central bank communication is its policy regime and this is enough to reveal the logic behind policy decisions. This argument indeed reinforces the essence of communication as a central part of central banks' monetary policy operation toolkit of all monetary policy frameworks. I also note that the choice of the adopted monetary policy frameworks differ across different countries. Houben and Houben (2000) and Casiraghi and Perez (2022) show two broad considerations that influence central banks choices. First, they argue that the choice of monetary policy strategy is determined by the relationships underlying the economy and second, that the central banks consider their desired reaction function to economic developments⁹.

As alluded to earlier, indeed, just as communication between the central bank and the general public has been progressing, research also on central bank communication and its market implications on the economy present a diverse and evolving literature. Romer and Romer (2004) are among the first to identify monetary policy shocks from central bank statements. Lucca and Trebbi (2009) propose an index based on "hawkish" and "dovish" language to track policy rate changes around FOMC announcements, and Hansen and McMahon (2016) extract topics from FOMC minutes to evaluate the impact of forward guidance on economic and financial markets. Other related works include Hansen et al. (2018), Rosa and Verga (2007), Rosa (2011b), Jegadeesh and Wu (2017) and Picault and Renault (2017), and Smales and Apergis (2017a)¹⁰.

monetary policy by influencing the expectations and actions of market participants. In a recent reiteration of this rationale, Unsal et al. (2022) and Casiraghi and Perez (2022) have stated that the purpose of communication is to guarantee that the actions of central banks do not generate uncertainty in the markets, which ultimately contributes to the enhancement of the effectiveness of monetary policy.

9. Sullivan (2024) discusses in detail very specific factors within the broad considerations.

10. See Woodford (2005) and Blinder et al. (2008) and Blinder et al. (2024) for a comprehensive review.

The literature that examines market reactions to central bank announcements is similarly extensive. Seminal work by Bernanke and Kuttner (2005) demonstrates the significant impact of Fed policy decisions on stock prices, while later studies such as Lucca and Moench (2015), Neuhierl and Weber (2019), Rigobon and Sack (2004), Campbell et al. (2020) and Hanson and Stein (2015), and Bernanke and Kuttner (2005) expand on these findings. This research relates closely to studies quantifying policy surprises via asset price changes in short windows around announcements Kohn and Sack (2003), Gürkaynak et al. (2005), Chung et al. (2013), Smales and Apergis (2017a), Nakamura and Steinsson (2018), Cieslak and Schrimpf (2019), Leombroni et al. (2021) and Swanson (2021).

In what follows, I continue to highlight various contributions by examining several other dimensions of the literature on specific aspects of central bank communication and the textual data mining work that is associated with it. That is because the current study uses both text data and hard data to understand specific aspects of central bank communication and the accompanying implications on financial instruments. So, I try to present empirical literature that specifically focus on reviewing studies related to the transparency, and the significance of communication clarity. In addition, I investigate the literature concerning the influence that communication strategies regarding monetary policy have on the ability of the central bank to achieve its own goals, as well as on the financial markets and households¹¹.

Almost all the dimensions of monetary policy communication have ample empirical evidence and well established findings, especially on developed economies and, to a relatively moderate extent, in emerging economies. This means that research from AEs naturally provides the majority reference understanding and evidence on the effectiveness of communication as a modern monetary policy tool. Various researchers have studied and empirically demonstrated the impact of monetary policy communication. With regards how communication impacts the central bank's own objectives, for example, Woodford (2005) for the U.S. shows that communication of future policy by the FED increases the predictability of policy in ways that improved its own ability to achieve its stabilisation objectives. Ehrmann and Fratzscher (2007) reiterated this observation by showing that even with dif-

11. See Appendix E for a supplementary brief literature on research linking market emotions to monetary policy communication.

ferent communication strategies, central bank communication improved the predictability of policy decisions and the responsiveness of financial markets to communication for both the FED and the ECB. Other studies, such as Rosa and Verga (2007), take an outward look and pursue evidence of how market expectations respond to information released by the ECB in its monthly press conferences. Similarly, Hayo and Neuenkirch (2010) for the U.S. finds that the FED communication contributes to understanding federal funds target rate decisions.

Although monetary policy communication has been found to have a multi-sectoral impact in the economy, that is, financial market, households, etc, Blinder et al. (2008) noted that virtually all the research up to 2008 had focused on central bank communication with the financial markets. Recently, however, researchers have also been paying attention to communication with the public or households (Binder (2017)). Therefore, there are broadly two specific strands of research on the impact of central bank communication, a branch of literature that concerns itself with understanding the impact of central bank communication on financial markets and a fairly recent literature that focuses on understanding the impact of central bank communication on the households.

The literature documenting the impact of monetary policy communication on households is still in its infancy. However, central banks' desire for an all-encompassing communication strategy is on a rise. Recently Binder (ibid.) noted that as central banks increasingly recognise that low public informedness limits their ability to communicate with households, many are frantically trying to tailor their communication strategies for a broader audience. There is good reason for such endeavours because, for example, in a study with a focus on central bank communication with households and using the U.S. consumer survey data, they find that consumers' expectations are imperfectly anchored. Additional evidence indicate that anchoring of more informed consumers' expectations increased more than the anchoring of less informed consumers' expectations following the FED's announcement. This evidence is important in as far as putting weight on the importance

of clear, consistent and concise communication. Ehrmann and Wabitsch (2022) for Germany also finds interesting evidence that - as tweets by non-experts become more factual with ECB communication, their subjectivity becomes less pronounced and less dispersed, thus also underscoring the importance of communication.

Therefore, whether it is financial markets or households, I note that understanding monetary policy contents plays a key role in market players' understanding of policymakers' communication. It indeed helps anchor market expectations and helps central banks achieve their set objectives. Recent literature, has taken a huge research interest focused on clarity of monetary policy communication and with it a fast-growing literature on clarity of monetary policy communication to both financial markets and the public or households. The ECB (2002) noted and put forward that "transparency means more than simply releasing information, as this does not by itself translate into a better understanding of monetary policy". The Bank remarked that the potential problem of information overload contrasts with the need for clarity, which becomes even more important when information is to be communicated to different audiences across different environments.

The arguments above suggest that clarity helps to make the monetary policy decision process less noisy and helps to enhance the capability of agents and markets to attach the correct weight to the contents of the monetary policy communication. To support this, Jansen (2011), for example, applies readability statistics to the testimonies given by the FED Chairman and tests whether the clarity of central bank communication affects volatility in financial markets and shows among other things, that clarity matters and it has a diminishing effect on volatility. Again in the lines of Jansen (*ibid.*), Smales and Apergis (2017a), also provides evidence of the importance of monetary policy communication clarity in the market for 10-year Treasury note futures. Their results show that the differences of opinion as a result of interpreting complex language in policy communication result in more trading volume, even when the markets have relatively low levels of liquidity (a negative liquidity-volume relationship). Their result suggests that the central bankers may aid market liquidity by issuing statements that are easier to understand

with shorter sentences and more familiar words. In related studies Smales and Apergis (2017b), Anderson (2010) and Kohn and Sack (2003), find that more complex language in communication significantly increases the trading volume, and volatility of returns, in stock, bond, and currency markets.

Bulř et al. (2013) in another related study examines whether the clarity of central bank communication about inflation varies with the economic environment by analysing the clarity of various communication channels, in particular, inflation reports, press releases, and statements by seven central banks¹². Their main results suggest that first, readability was at different levels in different countries. Second, they find little evidence suggesting that the clarity of central bank communication is affected by higher uncertainty or with a less favourable inflation outlook hence suggesting that clarity of communication in the sample countries was independent of the external adverse economic developments. This result is important in as far as emphasising that while the economic situation maybe difficult, the monetary authority's communication must remain clear regardless.

In summary, the literature shows an evolving and active research landscape on central bank communication and its multifaceted impact on financial markets and households. I begin by highlighting the evolution and rationale behind increased central bank communication and its transparency, emphasising the importance of accountability, effectiveness of monetary policy, and avoidance of market uncertainty. The IMF and others, for example, Unsal et al. (2022) have underscored the pivotal role of clear communication in enhancing monetary policy efficacy and market stability. Additionally, Geraats (2009) and Casiraghi and Perez (2022) point out that communication strategies extend beyond inflation targeting frameworks to encompass various monetary policy frameworks, with a consistent emphasis on conveying policy regime and logic to stakeholders. Despite variations in monetary policy strategies across countries, central banks share common objectives which underscores the universal importance of communication and clarity in achieving central bank goals.

12. Sample included the following countries, Chile, Czech Republic, Euro Area, Poland, Sweden, Thailand and UK

Furthermore, the literature review delves into empirical evidence regarding the impact of central bank communication on financial markets and households. Studies such as Woodford (2005), Ehrmann and Fratzscher (2005), and Rosa and Verga (2007) demonstrate the influence of central bank communication on market expectations, interest rates, and asset volatility, particularly in advanced economies. Moreover, emerging research, exemplified by Binder (2017) and Ehrmann and Wabitsch (2022), highlights the nascent but growing importance of central bank communication with households. These studies underscore the significance of clear and concise communication in anchoring expectations and shaping economic outcomes. Overall, the literature suggests that clarity in central bank communication not only enhances market predictability and stability but also fosters public understanding and confidence in monetary policy decisions, ultimately contributing to more effective economic governance. It is clear from past studies that the assessments on the complexity and clarity of monetary policy communication and its impact have been based on the entirety of the policy documents rather than on specific topics covered by the decision-makers. This is the main blind spot of the research on this subject.

Finally, in addition to the overview of central bank communication provided in the literature review, it is essential that I reiterate the scarcity of research specific to sub-Saharan Africa. Despite the wealth of empirical evidence and theoretical frameworks available for advanced economies, studies focusing on the region's central banks are notably scarce. While the IMF provides insights into global monetary policy frameworks, including those in sub-Saharan Africa, there remains a gap in understanding the unique challenges and dynamics facing central banks concerning communication in this region. The lack of research on the communication of African sub-Saharan central banks interferes with efforts to fully grasp its impact on local financial markets, policymakers, and households. Therefore, there is a need for investigations tailored to sub-Saharan Africa. By addressing this research gap, my aim is to contribute valuable insights to enhance the effectiveness of central bank communication and promote economic stability and growth throughout the region.

4 Data, Methodology and Descriptive Analysis

This study focuses on South Africa and the primary selection criteria takes into account three main factors. First, the South African financial market is widely regarded as a forerunner in the SSA, and it is regarded by many as a benchmark for the economic and financial developments that have occurred in the region (for example, see Odhiambo (2011), Ojah and Kodongo (2015), Berg et al. (2015), and Ahmed (2016)). Because it is distinguished by its modern infrastructure, diverse economy, and powerful regulatory framework, it is easily distinguished as one of the most developed and liquid markets in the region. This is the reason why it is described by these characteristics. Due to the fact that the majority of the theory and accompanying data is primarily developed and documented on AEs and certain EMEs, South Africa provides a unique perspective in this regard.

I take into consideration the availability of data for both policy statements and high-frequency intra-day data, which is an essential component for successfully executing my analytical approach. This is the second and most crucial factor that I take into account.

In order to initially estimate the complexity time series of the overall policy communication and that of specific topics therein, my input data consists of 151 monetary policy statements with a total of 300,323 words. The monetary policy statements are published in English by the SARB covering the period between October 1999 and July 2023 (see Appendix Table 4.7). The statements are available on the official SARB website. The shortest MPC statement was issued on 15 January 2002 with just 135 words and it was a special and emergency meeting to change the monetary policy stance due to a huge depreciation of the Rand, which changed the inflationary expectations in South Africa. The Monetary Policy Committee decided to increase the repo rate by 1 percentage point at this meeting. In general, Governor Marcus issued relatively long statements by word count averaging 2,606 words compared to averages of Governor Mboweni's 1,760 and Governor Kganyago's 1,930 words (see Table 4.1). Therefore, the trend in the size of the monetary statements peaks in July 2014 during Governor Marcus before trending downwards during Governor Kganyago.

These documents are downloaded in PDF format from the SARB website¹³ and converted to plain text while preserving original headings, sections, and paragraph breaks. Then I apply a cleaning pipeline to strip boilerplate material, standard introductions, footers, and annexes in order to isolate only the substantive policy discussion. These cleaned statements form the basis for the topic-level analysis of communication complexity.

Table 4.1: Summary of SARB MPC Statements

	Count	Period
Number of statements	151	13 Oct 1999 to 20 Jul 2023
Average length - whole sample	1,989	13 Oct 1999 to 20 Jul 2023
Minimum length	135	15 Jan 2002
Maximum length	3,401	17 Jul 2014
Total words full sample	300,323	13 Oct 1999 to 20 Jul 2023
Average length Gov. Mboweni	1,760	13 Oct 1999 to 22 Sep 2009
Average length Gov. Marcus	2,606	22 Oct 2009 to 18 Sep-2014
Average length Gov. Kganyago	1,930	20 Nov 2014 to 20 Jul 2023

Notes: The table reports on number of monetary policy statements, length of statements, governor regime averages and total number of words.

I also downloaded hard data of the 10-year South African government bond available at highest-frequency, minute-by-minute Ask, Bid, and Price data for the period between July 2015 and July 2023. The minute-by-minute intra-day data is retrieved from the Bloomberg Terminal for each trading day coinciding with the SARB's monetary policy announcement. In addition, I also downloaded 10 working days prior to the announcement day which serves as a control in subsequent analyses. For example, on 20 July 2023, the SARB held a press conference starting at 3:00 PM local time and the press briefing typically lasts no more than an hour. This means that I download minute-by-minute data for the 10-year government bond for the same day - announcement day¹⁴. This approach is particularly common in the literature, for example, Kuttner (2001), Chung et al. (2013), Gürkaynak et al. (2005), Smales and Apergis (2017a) and Nakamura and Steinsson (2018) and allows me to make precise assessments of financial market reactions

13. <https://www.resbank.co.za/en/home/publications>

14. The monetary policy statements have been accessed at <https://www.resbank.co.za/en/home/publications/>

in real-time, capturing immediate price movements and volatility around the monetary policy announcements with regard to specific key policy topics. In the literature, the 10-year government bond is also used to appreciate long-term expectations, and its liquidity can signal investor confidence or uncertainty in response to policy news.

4.1 Measuring Text Complexity

A large body of the literature that analyses the quality of monetary policy communication and its impact on financial markets highlights the importance of text complexity, readability, and clarity (see for example, Li (2008), Biddle et al. (2009), Miller (2010), Lehavy et al. (2011), Rennekamp (2012), Lawrence (2013), Bulíř et al. (2013), Lundholm et al. (2014), De Franco et al. (2015), Guay et al. (2016), Hayo et al. (2022) and Vyshnevskiy et al. (2024)). Some studies measure these aspects using basic numerical indicators, such as sentence count, sentence length, and syllable count. However, various studies typically assess readability or complexity by approximating the educational level required to comprehend a written document. A widely used metric in analysing the communication complexity of central bank statements is the Flesch–Kincaid Grade Level (FK), which integrates sentence length, word length, and syllable counts within a regression framework (Smales and Apergis (2017a), and Hayo et al. (2022))¹⁵.

4.1.1 Overall Document Complexity

The Flesch–Kincaid grade level (Kincaid (1975)) has become particularly useful for evaluating readability in monetary policy documents. This popularity is attributed to its straightforward interpretability as simply the number of years of education required to understand written communication. The metric is also attractive because it is easy to

15. In the literature, there are a number of measures of text complexity, for more on the different metrics, see for example, Flesch (1948), Gunning (1952) and Mc Laughlin (1969) among others. Appendix F provides a brief on the common alternative measures.

apply to financial communication. Higher numbers mean higher text complexity and the converse is true. In terms of quantitative formulation, the FK assesses the readability of English text based on two key factors: the average sentence length (ASL) and the average number of syllables per word (ASW). It is represented as follows:¹⁶

$$FK = 0.39 * \left(\frac{ASL}{ASW} \right) + 11.8 * \left(\frac{ASW}{ASL} \right) - 15.59 \quad (4.1)$$

$$ASL = \frac{\text{Number of words}}{\text{Number of sentences}}$$

$$ASW = \frac{\text{Number of syllables}}{\text{Number of words}}$$

where the parameters were estimated by regressing the number of years of education of 569 survey participants from the U.S. Navy on the independent variables computed from the passages administered in the survey. Although the FK index is easy to interpret, it is also criticised for having several limitations. First, the FK index is based on a U.S. experiment and is calibrated to the U.S. education system, taking into account only the reading levels and school grades defined in the United States. Second, the original study involved participants from the U.S. Navy, and as such, it does not consider a reader's background knowledge or familiarity with the subject matter—factors that can significantly influence comprehension. In particular, longer sentences are often necessary to articulate complex macroeconomic concepts. However, such structures do not inherently imply that the content is difficult for economic agents to understand.

16. Typical grade level categories based on FK scores are: 0-4: Very easy 4th-grade student or lower, 5-6: Easy 5th or 6th-grade student, 7-8: Fairly easy 7th or 8th-grade student, 9-10: average difficulty 9th or 10th-grade student, 11-12: Fairly difficult 11th or 12th-grade student, 13-15: Difficult college-level and 16-18: Very difficult postgraduate-level.

4.1.2 Determining Topic Complexity

Traditional readability measures are typically applied to entire documents, such as policy statements, reports, or inflation updates (Smales and Apergis (2017a) and Smales and Apergis (2017b)), resulting in a single overall complexity score. However, central bank statements often cover multiple topics, each varying in detail, vocabulary, and relevance to market conditions. As a result, some topics, and the clarity with which they are discussed, may have a greater impact on market responses than others. To capture these nuances, I evaluate the clarity and complexity of each individual topic instead of treating the document as a homogeneous whole.

Topic identification. To extract the underlying topical structure, the study employs the Latent Dirichlet Allocation (LDA) model proposed by Blei et al. (2003). While alternative methods such as the Louvain algorithm Blondel et al. (2008), which clusters words into non-overlapping groups by maximising network modularity, could be considered, LDA is preferred because it allows a word to belong to multiple topics. This flexibility is crucial when a word such as *rate* appears in different contexts (e.g., inflation rate, unemployment rate, or growth rate).

LDA functions as a dimensionality reduction tool that transforms high-dimensional text data into a set of interpretable topics. In this study, text data consists of the SARB monetary policy statements. Notably, upon manual inspection, I observed that each paragraph in these statements typically focuses on a single topic. Although the same topic might recur in different paragraphs, there is generally no overlap of topics within the same paragraph. For this reason, each paragraph is treated as an independent document and topic estimation is performed at a paragraph level.

Prior to applying LDA, common terms (e.g., “SARB”, “quarter(s)”) and extraneous symbols (often introduced during the PDF-to-text conversion) are eliminated, and the remaining words are reduced to their common roots. Formally, for each paragraph i (denoted as w_i), I assume that the word distribution follows a multinomial model:

$$w_i \sim \text{Mult}(\Phi' \theta_i, N_i), \quad (4.2)$$

where, given a fixed number K of topics, $\theta_i = [\theta_{i,1}, \dots, \theta_{i,K}]$ is drawn from a Dirichlet distribution $\text{Dir}(\alpha)$ and represents the topic proportions for paragraph i . This satisfies $\theta_{i,k} \geq 0$ and $\sum_{k=1}^K \theta_{i,k} = 1$. $\Phi = [\phi_1, \dots, \phi_K]'$ is the word distribution for each topic, with each ϕ_k drawn from $\text{Dir}(\beta)$, ensuring $\phi_k > 0$ and $\sum_{k=1}^K \phi_k = 1$. N_i is the total number of words in paragraph i .

The outcome is a probabilistic mapping of K topics to each paragraph. Although K is an exogenous parameter, I determine its plausible range (between five and ten topics) through manual reading. To refine my choice, I recursively estimate LDA models for $K = 1, \dots, 10$ and select $K = 5$ based on the highest topic coherence, which measures interpretability through word co-occurrence patterns.

Topic Dictionaries. After identifying topics and assigning them probabilistically to paragraphs, I develop an algorithm to assess language complexity at the topic level. This process involves two steps. First, I construct a dictionary capturing the key terms and phrases associated with that topic. During this stage, I account for the evolution of language over time in the SARB monetary policy statements by manually reviewing each statement. This ensures that vocabulary shifts, where some terms may drop out and new ones emerge, are accurately captured. The dictionaries include expressions ranging from unigrams to five-word phrases. Figure 4.2 provides a visual summary of the constructed dictionaries for each topic.

5 Empirical Results

5.1 Overall and Per-Topic Complexity

An evaluation of the level and evolution of the linguistic complexity in the SARB statements indicate that the bank has made progress in simplifying communication across respective governor regimes. Since adopting the IT monetary framework in 2000, Table 4.2 and Figure 4.3 indicate that the SARB has made notable positive gains in communication.

Most significant communication improvements are observed under the leadership of Governors Mboweni (1999 to 2009) and Kganyago's (2014 to date). The mann-kendall trend test of the FK trend for individual governors are statistically significant with a p-value $p < 0.01$ (see Table 4.2), thus confirming the sub-sample trends in the linguistic complexity in communication. Governor Kganyago, in particular, has demonstrated the steepest, most consistent and substantial reductions in linguistic complexity of the MPC statements with an average FK grade score of 12.28 compared to 12.83 and 13.57 for Governor Mboweni and Governor Marcus, respectively. The regime of Governor Marcus (2009 to 2014) was, however, characterised by increasing linguistic complexity in the MPC statements.

Table 4.2: The Mann-Kendall Trend Test

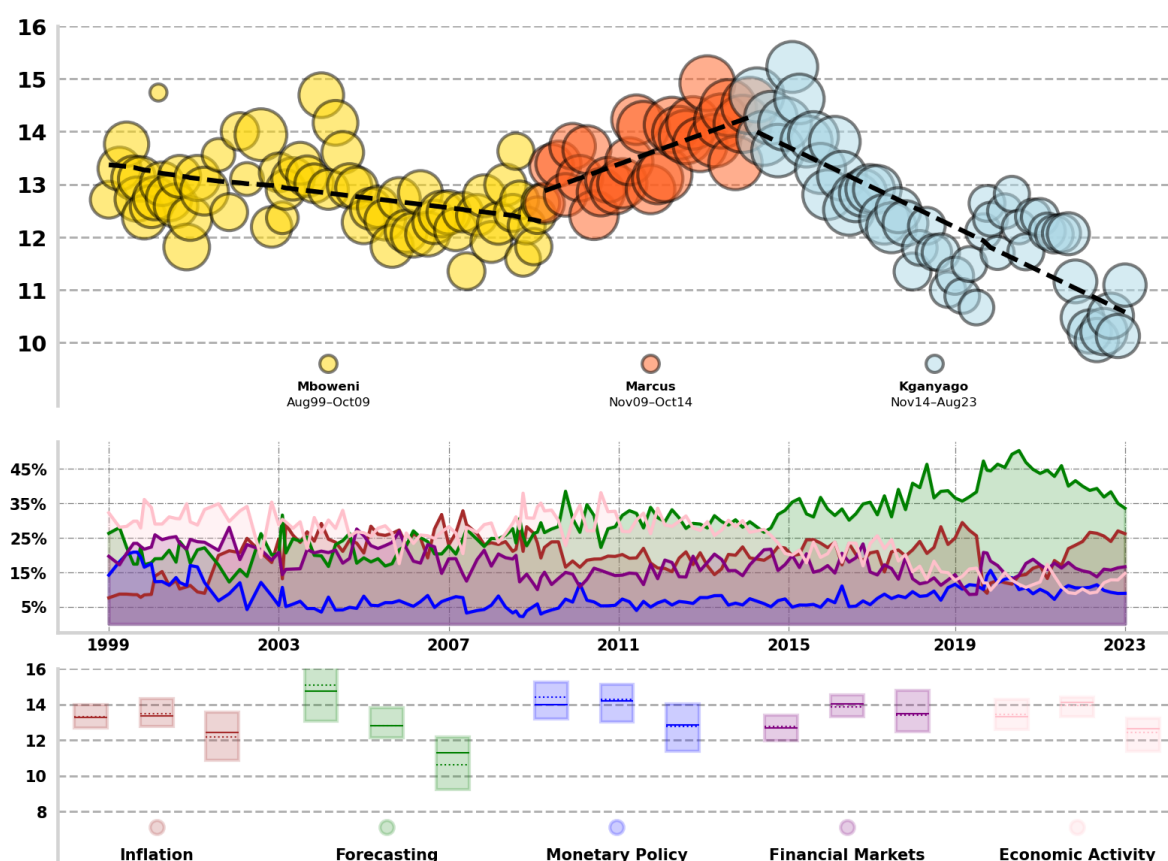
	Mboweni	Marcus	Kganyago	Full Sample
Number of MPCs	67	30	54	151
Mean FK	12.83	13.57	12.28	12.78
Minimum FK	11.35	12.41	10.04	10.04
Maximum FK	16.97	14.92	15.23	16.97
Slope Coefficient	-10.22	7.23	-17.20	-15.66
P-Value	0.00021***	0.00012***	0.00000***	0.00014***
Z-Statistic	-3.701	3.854	-6.923	-3.803
Comment	Decreasing	Increasing	Decreasing	Decreasing

Notes: The table summarises full sample and governor specific statistics including the Mann-Kendall test which is used to establish if communication complexity shows a significant and consistent upward or downward trend. The test confirms declining linguistic complexity during the regimes of Governors Mboweni and Kganyago and a rise during Governor Marcus. Over time, however, the SARB's communication shows declining linguistic complexity which suggests improvement in clarity. Statistical significance levels are $p < 0.01$ (***), $p < 0.05$ (**), and $p < 0.1$ (*).

Despite improvements in communication during Governor Kganyago's tenure, there have been periods of rising linguistic complexity, for example, the complexity of the communication statements increased during the COVID-19 pandemic in 2020. The deterioration in communication was driven by the increasing complexity of communication in forecasting, financial markets and economic activity. The temporal rise in complexity perhaps reflect the difficult of communicating pessimistic forecasts and inefficient banking operations in an environment heavily impacted by external and domestic challenges which adversely impacted international trade, supply chains and travel services.

Following this period, the downward trajectory in complexity resumed, reflecting less complex communication as the crisis ended. Another episode of rising complexity was as reflected by the FK grade level in 2023 and was due to rising linguistic complexity in financial markets. When I compare the per-topic complexity across governors, I observe that Governor Kganyago has improved communication of all the topics in the statements. In general, although the individual topics complexity have trended together, it is the linguistic complexity improvements in inflation, financial markets, and economic activity that drove the complexity index during Governor Mboweni. In contrast, all topic complexities deteriorated under Governor Marcus, while under Governor Kganyago's leadership, each topic's linguistic complexity improved, resulting in a downward trend in the overall communication complexity. I also see that while forecasting is the best communicated under Governor Kganyago, it was once the most linguistically complex topic to communicate during the early IT adoption years of Governor Mboweni. This observation immediately raises the question of how topic-specific linguistic complexity impacts market liquidity; a focus that is explored in the subsequent sections.

Furthermore, although some studies in the literature such as Smales and Apergis (2017a) support the claim that there is positive correlation between size of document and its complexity, I fail to establish the same in the case of the SARB in a discernible way as I observe short statements associated with high FK grade levels and vice-versa. This observation underscores the importance of analysing language complexity at the topic



Notes: The figure shows the trend in overall linguistic complexity, topic proportions and comparisons of governor performance in linguistic complexity (measured through the Flesch–Kincaid reading grade level) from the Reserve Bank of South Africa. In the Top Panel, the area of each circle is proportional to the length (total number of words) of the respective statement, with color representing the specific central bank governor regimes. The Middle Panel illustrates the time series of five topics, namely, Inflation, Forecasting, Monetary Policy, Financial Markets and Economic Activity. The Boxplot, in the Bottom Panel depicts the difference in performance of governors across five topics with results indicating that Governor Kganyago had consistently lower median in majority of topics meaning superior performance.

Figure 4.3: Overall and Per-Topic Communication Features

level rather than relying solely on document length. Overall, my findings indicate that the observed improvement in the overall readability of SARB communications is primarily driven by a decline in the complexity of almost all the topics, but with forecasting having come to occupy an increasing portion of the statements since 2010.

My topic identification strategy indicates that the SARB discusses five key topics, namely, forecasting, inflation, financial markets, economic activity and monetary policy. The observed increase in discussion on forecasting is consistent with the importance that the IT monetary framework places on forward-looking guidance. Other notable dynamics are that forecasting proportion declined immediately following the global pandemic in 2020

while inflation and economic activity proportions in the statements increased during and after the COVID-19 pandemic after a remarkable dip prior to the crisis. This is perhaps not surprising as the two topics, inflation and economic activity, relate to the primary and secondary objectives of the SARB and as such, this observation merely reflect the key issues that the central bank was grappling with. Monetary policy and financial markets topics remained relatively stable over time.

5.2 Identifying the SARB Communication Effect

To assess the presence of the SARB announcement effect in financial markets, I analyse the behavior of liquidity using the bid-ask spread of the 10-year government bond on the announcement days. In particular, I determine the sensitivity of the 10-year government bond liquidity to the SARB communication and its complexity by comparing the means and the distributional differences of liquidity on the announcement days with a 10-day control. Relatively higher observed mean quoted spread values imply lower levels of liquidity in the market and the converse being true. I purposely choose to calculate and use averages of 30 minutes intervals in this exercise so that I create a balance between capturing meaningful smoothed out liquidity and price movements and avoiding excessive noise due to individual trade deviations from the expected.

In the literature, the communication effect is initially identified by observing the evolution of a financial instrument liquidity within a tight window before and after the announcement (see, for example, Chung et al. (2013) and Smales and Apergis (2017a)). However, my approach to identifying the communication effect slightly departs from the literature and presents a unique extension as follows. First, I divide the full sample of policy statements into two categories; statements categorised as easy to understand and the hard to understand. In addition, I use density distributions, which are superior to traditional line charts in a variety of ways when it comes to highlighting the essential characteristics and features of data. They are particularly useful for detecting characteristics like modal patterns, skewness, kurtosis, the presence of outliers, or multi-modality that line charts

may overlook or misunderstand. Because overlapping density curves might reveal subtle changes in central tendency, spread, or shape, revealing the true structure and variability of the data, I compare the liquidity on announcement day and 10-day control distributions at the same time.

Second, the categorisation is applied to both the overall statements complexity and complexity of individual topics within the statements. The aim is to help appreciate market liquidity outcomes under different communication complexity conditions. More specifically, I investigate how liquidity of the 10-year government bond reacts when specific topics, such as inflation, forecasting or monetary policy, are presented with varying degrees of complexity, low or high. This approach enables me to identify and gain initial insights into financial market responses and appreciate the specific topics whose linguistic complexity the financial markets are more sensitive to. This approach, therefore, provides a more appreciable understanding of the central bank communication effect while offering valuable perspectives on the interplay between central bank communication content-complexity and financial markets movements.

5.2.1 Communication Effect on the Bond Liquidity

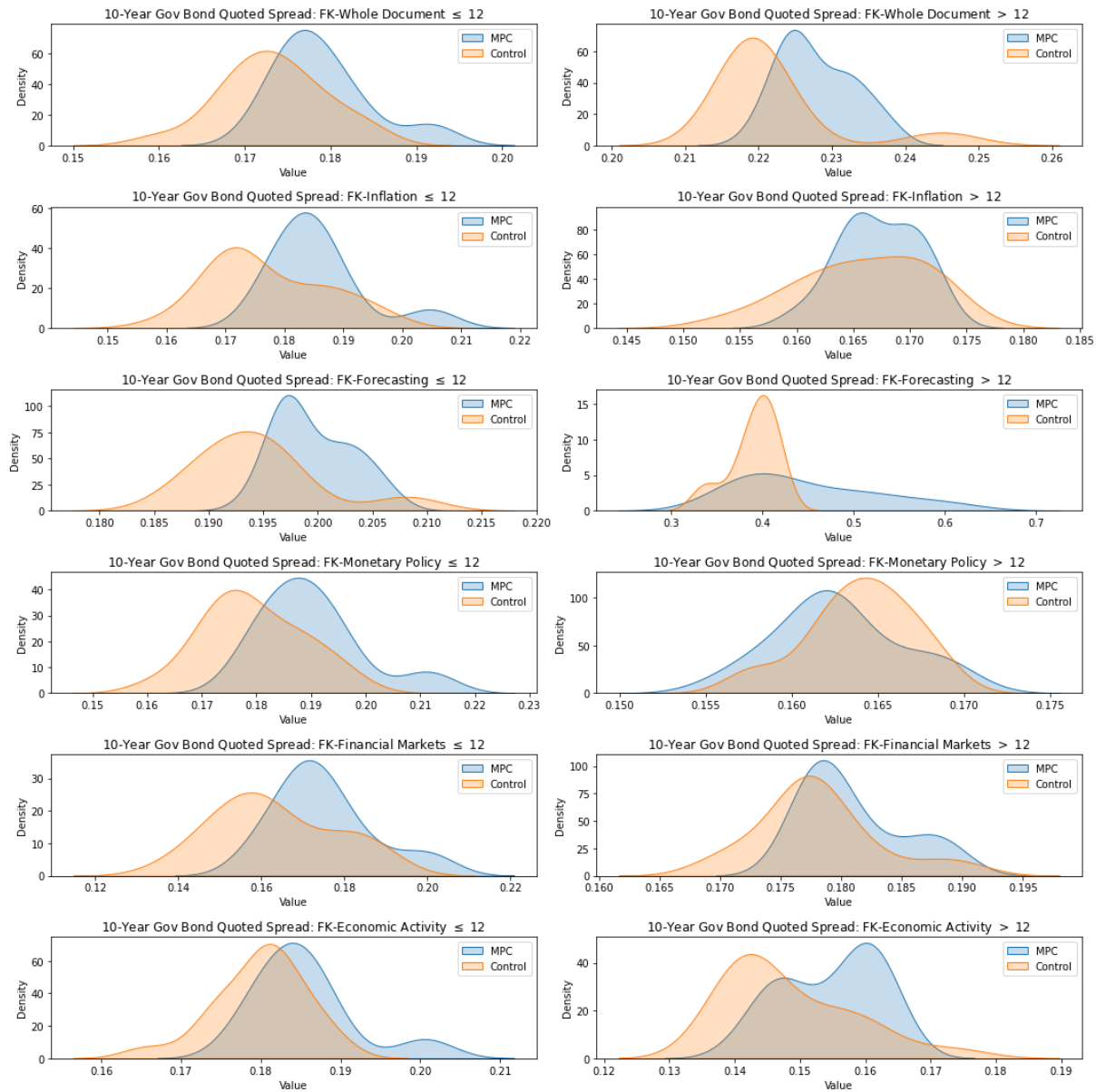
The expected evolution of liquidity of the South African 10-year government bond market on the day of MPC announcement versus 10-day control, and when document and topic linguistic complexity is low and high is depicted in Appendix Figure 4.11. First, I establish a key feature of the financial markets that liquidity declines sharply at around 1:30 PM (-3) in both the 10-day control and the day of announcement day.

This is because on weekdays, businesses in South Africa open from 8:00 PM to 1:00 PM and 2:00 PM to 5:00 PM. So 1:30 PM falls right at the center of the lunch hour and during lunch hours there is reduced to no activity of buying and selling which lowers market liquidity as traders take a break hence my observation.

Second, and more substantive, is that I compare the average liquidity outcomes in two categories, the period in which the statement is easy ($FK \leq 12$) and the period in which the statement is hard ($FK > 12$) with respective 10-day controls¹⁷. The categorisation is informed by the educational structure of South Africa and the practical needs of the audience that is targeted by the central bank communication which include journalists, analysts, and informed members of the public. Most of these groups require having completed secondary education. In addition, the average grade score for the whole document and entire sample is $FK = 12.78$ and is slightly higher than 12. The overall average grade score is different from the individual topic average grade scores as well, so the choice cut-off is applied in all scenarios. However, I do not expect the results to be overly sensitive to the exact cut-off.

From Figure 4.4, the mean of the realised liquidity of the 10-year government bond is higher than the control for the whole document and across all topics when the $FK \leq 12$ suggesting a general lower liquidity outcomes on the day of announcement. This is confirmed by the t-test p-values reported in Table 4.3 for the differences in announcement means and control conditions. The mean liquidity, however, is lower than the control when the $FK > 12$ for the whole document and forecasting. This is an example of the limitation of identifying the communication effect using the mean approach. It can be misleading because the actual evolution of liquidity when the $FK > 12$ for both whole document and forecasting exhibit big differences between the control and the observed on day of announcement after the announcement, which is masked in the mean calculation.

17. I subjectively decided that the number of years above 12 years (secondary school) - vocational and tertiary education is “high complexity” in accordance with the South African education system. As such, statement with the FK value of 12 and less are considered “low complexity”.



Notes: The plots show the differences in averages and variability of the liquidity of the 10-year government bond market when the linguistic complexity of the MPC communications from the Reserve Bank of South Africa is easy or hard. This is for the day of announcement and 10-day control for the period between 08:00 AM and 17:00 PM. The reported values are the means of 30-minute intervals on the day of announcement and the 10-day control. Therefore, the figures represent the density distributions in the periods when the central bank communication of the overall statement and that of specific topics, namely, Inflation, Forecasting, Monetary Policy, Financial Markets and Economic Activity, is easy and when hard. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.4: Liquidity Density Distributions

Therefore, I test the distributional differences of the liquidity time series. The distributional shift test indicate that liquidity when forecasting is $FK > 12$ shows the largest difference in the mean distribution shapes as indicated by Kolmogorov-Smirnov (KS)-tests $p\text{-value} = 0.0028^{18}$. Similarly, the mean does not shift in the remaining cases when the $FK > 12$ indicating that the expected liquidity remains more or less unchanged, but again the KS confirms that the distributional of liquidity is significantly different in all cases. The standard deviation, however, is easily seen to be larger on the announcement day relative to the control when $FK > 12$ for the whole document and forecasting suggesting high variability in the liquidity when the communication is unclear and complex. The MPC announcements and its linguistic complexity generally appear to drive the liquidity. In general, the KS test p -values of distributional shifts indicate that the distributions of the mean liquidity differ significantly between MPC communication and control conditions. With an exception of inflation, and monetary policy when $FK > 12$, statistically significant distributional differences suggest that the linguistic complexity of specific topics have a relationship with not only the average liquidity but also its overall shape or variability on the day announcement.

In summary, this exercise establishes that the densities generally overlap and shifted to the right reflecting relatively lower average liquidity levels on MPC announcement days even when the complexity is low, $FK \leq 12$. The distributions are also relatively narrow, suggesting lower variability in liquidity around its mean under low communication complexity conditions. The liquidity distributions shift even further to the right compared to the control, reflecting increasingly lower liquidity levels on MPC days when communication complexity is high. The distribution under the MPC announcement conditions are more spread out and right skewed, suggesting extreme low liquidity levels and greater variability in liquidity under high linguistic complexity conditions. The results of the formal tests are reported in Table 4.3 below.

18. Kolmogorov-Smirnov (KS) test is a nonparametric test of the equality of continuous one-dimensional probability distributions.

Table 4.3: Comparison of Mean Liquidity

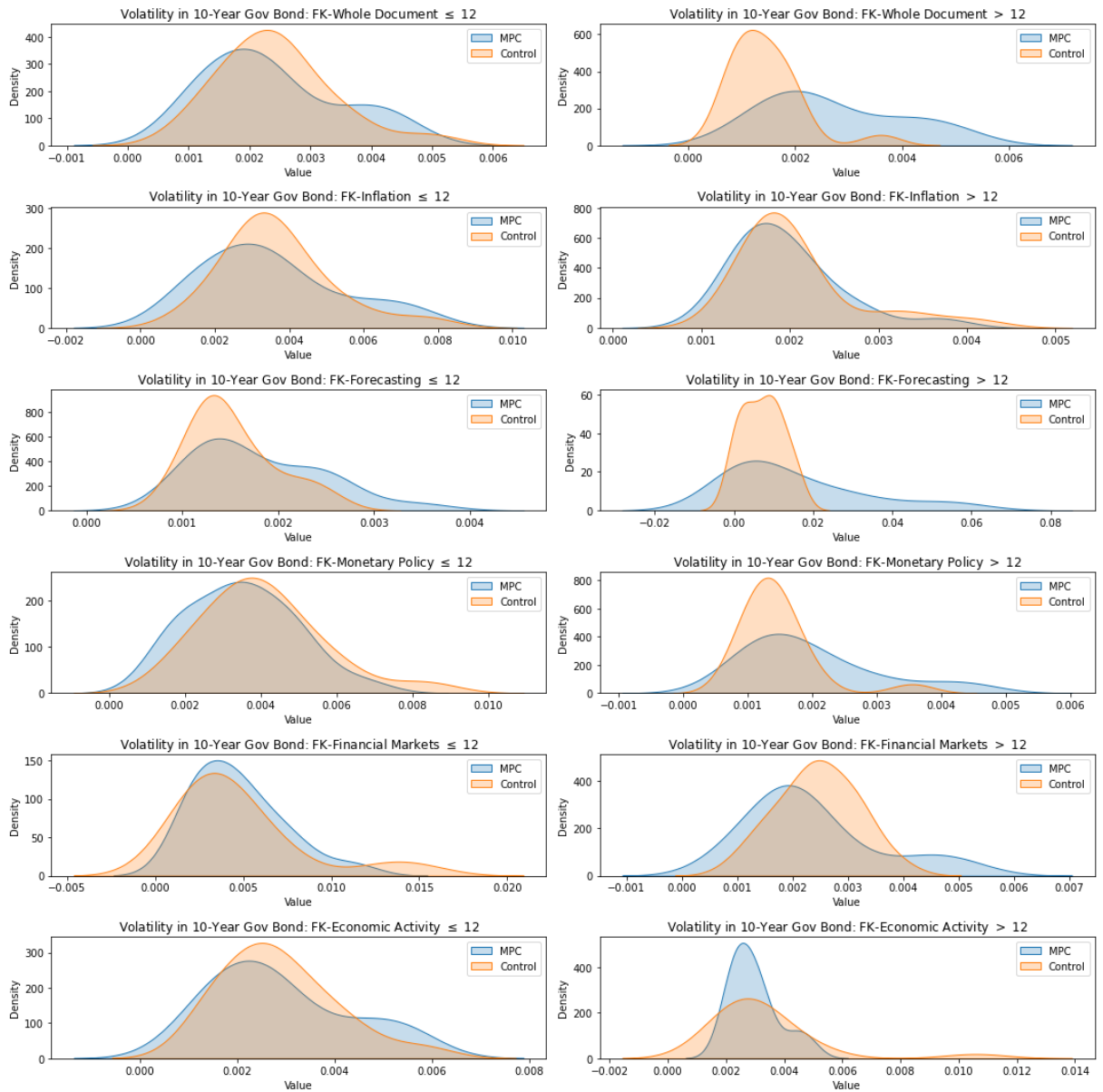
Topic & Complexity	Condition	Mean (MPC)	Mean (10-Day Control)	Std Dev (MPC)	Std Dev (Control)	T-Test P-Value	KS-Test P-Value
FK-Whole Document ≤ 12	MPC	0.18	0.17	0.0056	0.0060	0.0040***	0.0092***
FK-Whole Document > 12	MPC	0.23	0.22	0.0050	0.0084	0.0147**	0.0000***
FK-Inflation ≤ 12	MPC	0.19	0.18	0.0079	0.0098	0.0063***	0.0007***
FK-Inflation > 12	MPC	0.17	0.17	0.0034	0.0056	0.3781	0.5379
FK-Forecasting ≤ 12	MPC	0.20	0.19	0.0034	0.0057	0.0027***	0.0002***
FK-Forecasting > 12	MPC	0.45	0.39	0.0754	0.0257	0.0034***	0.0028***
FK-Monetary Policy ≤ 12	MPC	0.19	0.18	0.0092	0.0091	0.0010***	0.0007***
FK-Monetary Policy > 12	MPC	0.16	0.16	0.0036	0.0030	0.2629	0.1532
FK-Financial Markets ≤ 12	MPC	0.18	0.16	0.0118	0.0144	0.0092***	0.0092***
FK-Financial Markets > 12	MPC	0.18	0.18	0.0042	0.0049	0.0616*	0.0681*
Economic Activity ≤ 12	MPC	0.19	0.18	0.0063	0.0056	0.0068***	0.0267**
Economic Activity > 12	MPC	0.15	0.15	0.0074	0.0098	0.0308**	0.0092***

Notes: For the liquidity of the 10-year government bond, I report the means, standard deviations, T-Test P-Value and KS-Test P-Value for categories easy, FK ≤ 12 and hard FK > 12 for the whole document and specific topics on MPC announcement day and 10-day control. Mean (MPC), Mean (10-Day Control) and respective standard deviations are averages and standard deviations, respectively, of each category on announcement day and 10-day control. The T-Test P-Value measures if differences in means between MPC and Control for each category is statistically significant ($p\text{-value} < 0.05$), similarly, a small KS-Test P-Value indicates if distributions of liquidity differ significantly between MPC and Control conditions. Statistical significance levels are $p < 0.01$ (***), $p < 0.05$ (**), and $p < 0.1$ (*).

5.2.2 Communication Effect on the Bond Volatility

Similarly, in this section, the study compares the average liquidity volatility on the announcement day and the 10-day control. The exercise establishes that when the Fk ≤ 12 for the whole document, there is no material difference in the liquidity volatility with the $p\text{-value} = 0.8524$. A statistical test of the distributional shifts under similar conditions also yields similar results with $p\text{-value} = 0.5379$ (see Table 4.4). However, when the FK > 12 , the mean volatility on announcement day is higher than the control (mean = 0.0027 vs 0.0014) and is statistically significant with $p\text{-value} = 0.0006$. The distributional difference is also highly statistically significant with $p\text{-value} = 0.0007$. This result suggests that complex overall communication of the SARB is associated with declining trading activities suggestive of raising uncertainty in the bond market.

With regards to specific topics and varying complexity levels, forecasting and monetary policy communication only show higher volatility during MPC communication days when $FK > 12$, with statistical significance at 10% suggesting bond market sensitivity to linguistic complexity of forecasting and monetary policy stance elements. This can also be seen from the spread and right skewness of the respective charts in Figure 4.5. Similarly, inflation, financial markets and economic activity show no statistical significant differences under both complexity conditions.



Notes: The plots show the differences in averages and variability of the liquidity volatility of the 10-year government bond market when the linguistic complexity of the MPC communications from the Reserve Bank of South Africa is easy or hard. This is for the day of announcement and 10-day control for the period between 08:00 AM and 17:00 PM. The reported values are the means of 30-minute intervals on the day of announcement and the 10-day control. Therefore, the figures represent the density distributions in the periods when the central bank communication of the overall statement and that of specific topics, namely, Inflation, Forecasting, Monetary Policy, Financial Markets and Economic Activity, is easy and when hard. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.5: Volatility Density Distributions

Table 4.4: Comparison of Liquidity Volatility

Topic & Complexity	Condition	Mean (MPC)	Mean (10-Day Control)	Std Dev (MPC)	Std Dev (Control)	T-Test P-Value	KS-Test P-Value
FK-Whole Document ≤ 12	MPC	0.0024	0.0024	0.0011	0.0009	0.8524	0.5379
FK-Whole Document > 12	MPC	0.0027	0.0014	0.0013	0.0007	0.0006***	0.0007***
FK-Inflation ≤ 12	MPC	0.0036	0.0037	0.0019	0.0015	0.8392	0.8081
FK-Inflation > 12	MPC	0.0020	0.0021	0.0006	0.0007	0.5397	0.5379
FK-Forecasting ≤ 12	MPC	0.0018	0.0015	0.0007	0.0004	0.1346	0.5379
FK-Forecasting > 12	MPC	0.0154	0.0071	0.0173	0.0051	0.0581*	0.1532
FK-Monetary Policy ≤ 12	MPC	0.0035	0.0040	0.0014	0.0016	0.2569	0.8081
FK-Monetary Policy > 12	MPC	0.0020	0.0015	0.0010	0.0006	0.0641*	0.1532
FK-Financial Markets ≤ 12	MPC	0.0047	0.0048	0.0025	0.0036	0.9585	0.8081
FK-Financial Markets > 12	MPC	0.0024	0.0025	0.0012	0.0007	0.7679	0.1532
Economic Activity ≤ 12	MPC	0.0029	0.0029	0.0014	0.0012	0.9200	1.0000
Economic Activity > 12	MPC	0.0029	0.0033	0.0008	0.0020	0.4596	0.8081

Notes: For the liquidity volatility of the 10-year government bond, I report the means, standard deviations, T-Test P-Value and KS-Test P-Value for categories easy, FK ≤ 12 and hard FK > 12 for the whole document and specific topics on MPC announcement day and 10-day control. Mean (MPC), Mean (10-Day Control) and respective standard deviations are averages and standard deviations, respectively, of each category on announcement day and 10-day control. The T-Test P-Value measures if differences in means between MPC and Control for each category is statistically significant ($p\text{-value} < 0.05$), similarly, a small KS-Test P-Value indicates if distributions of liquidity differ significantly between MPC and Control conditions. Statistical significance levels are $p < 0.01$ (***), $p < 0.05$ (**), and $p < 0.1$ (*).

5.3 Factors Explaining the SARB Communication Effect

In this section, I want to make an attribution to different factors that explain the government bond liquidity. This is with particular interest to the role of linguistic complexity of monetary policy communication on the liquidity of the 10-year bond during the day of announcement. I follow the methodology used in two key studies in the literature, namely, Chung et al. (2013) and most recently Smales and Apergis (2017a) in which the former analyses the effects of monetary policy announcements on stock market liquidity using intra-day data and the latter studies factors driving the liquidity of the 10-year treasury note futures on the day of monetary policy announcements. Given the novel topic-specific linguistic complexity time series generated from my approach, I initialise with a hypothesis, that the linguistic complexities of the SARB specific topics in the policy statements are important factors for explaining the liquidity of the South African 10-year bond on the day of announcement.

In the previous section, I demonstrated that government bond liquidity experiences significant distributional shifts to the right on announcement days. Moreover, the distributional shifts are markedly larger compared to control days, particularly when the overall linguistic complexity of the policy statement is high and, notably, the complexity of specific topics. So, to examine the quantitative relationship between liquidity and linguistic complexity while controlling for the microstructure of the market and macroeconomic influences, the study takes care to align with event study conventions in the finance literature that use the liquidity measure of the difference between the liquidity on announcement day and the 10-day control.

To identify the time-varying effect of communication complexity, both at the overall and topic-specific level, I do so through the temporal dummy variable $I_{j,t}$ multiplied by the complexity per topic factor. This captures the interaction between the post-announcement time interval and the complexity of respective topic i , allowing an assessment of whether SARB announcements have a greater effect on liquidity due to the linguistic complexity of specific topics. So, I estimate regression models at 10- and 15-minute time intervals specified as follows:

$$\Delta L_{j,t} = c + \lambda \Delta V_{j,t} + \sum_{j=1}^{18} \theta_j I_j + \gamma P_{j,t} + \sum_{j=7}^{18} \rho_j (I_j \times S_t) + \sum_{j=7}^{18} \phi_j (I_j \times Topic_{i,t}) + \varepsilon_{j,t}, \quad (4.3)$$

where Δ denotes the announcement-day value minus its control-day counterpart. $L_{j,t}$ is the average quoted bid-ask spread in the j th ten-minute interval on announcement day t . V captures intra-interval liquidity variability, computed as the standard deviation of the minute-by-minute quoted spreads within the ten-minute interval. $I_1, \dots, I_{18}$ are binary indicators for each ten-minute segment covering the window from two hours before the press briefing to market close (14:00-17:00). P denotes the yield on the ten-year government bond. S measures the policy surprise, defined as the unexpected change in the SARB's target interest rate. $Topic_{i,t}$ reflects the prominence of thematic element i in the briefing, as extracted from topic-model analysis, $\phi_j (I_j \times Topic_{i,t})$, reflects the interaction between a given time interval and a specific topic, ε is the usual error term.

Although I control for market factors and the monetary policy rate, it is often the case that high frequency intra-day liquidity measures will show high autocorrelation. In addition, following the policy announcement, the financial markets may perhaps react in a way that increases volatility thus changing the variance of residuals over time. In light of this, the literature shows that the standard OLS standard errors will be biased due to serial correlation (OLS assumes homoskedasticity). In this regard, I specify my regression models with Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors consistent with choice by others in the literature, for example, Chung et al. (2013) and Smales and Apergis (2017a).

5.3.1 Panel Analysis

I report the 10-minute intervals regression results in Table 4.5 and the associated interval dummies in Appendix Table 4.9. With an exception of the sign on the coefficient on $Surprise_{j,t}$, all the other results are consistent with results of Chung et al. (2013) and Smales and Apergis (2017a). That is to say that volatility and bond price have an effect of reducing liquidity in the market. The results suggest that high volatility increases uncertainty around the policy events thus, raising the quoted spreads making it harder to efficiently. In addition, the increases in bond prices reduces liquidity perhaps to reflect falling yields.

The surprise, however, which is determined as the difference between the new policy rate and previous policy rate has a negative sign and is strongly statistically significant indicating that an increase in the repo rate has an effect of increasing liquidity. I am unable to determine whether the improvement in the liquidity is due to improving sentiment about the economy leading to improved investment in the financial instrument or simply banks lending more aggressively with the expectation of higher future returns. This result is, however, sensible if I assume that I have not omitted an important variable that is influencing both liquidity and the surprise simultaneously since policy stance is often decided based on inflation projections thus, independent of market developments on a particular announcement day.

Furthermore, the results indicate that linguistic complexity of the whole monetary policy communication has an effect of disrupting liquidity. Now, related to my second hypothesis, I run a regression in Model 4, with topic-specific complexity indices and the results show that with an exception of the coefficient on complexity of monetary policy, complexities of all other topics have the expected sign that suggest that they have an effect of lowering liquidity in the market albeit only complexity of forecasting is statistically significant.

In order to study the evolution of liquidity due to the monetary policy communication, I monitor the 10-minute intervals time dummies related liquidity averages from 2:00 PM to 5:00 PM. In the 10-minute interval regression, I omit interval, *I7* dummy (*I5* in the 15-minute interval regression) being the first interval between 3:00 PM and 3:09 PM at the start of the press briefing. This is my reference interval for understanding the evolution of liquidity.

Results show that the coefficients of the interval dummies, θ_{js} , in all the models are positive but change signs to negative relative to the 7th interval (at the beginning of the press release) suggesting that monetary policy communication has an immediate effect of improving market liquidity. The dummies are however, statistically insignificant which indicates that the overall liquidity developments are driven by fundamental explanatory variables, in general, the market micro-structure factors, linguistic complexity of communication and economic conditions rather than the specific moments after the announcement.

The trend in the liquidity reverses after interval *I14* (4:20 PM) which indicate that liquidity starts to deteriorate 20 minutes after the monetary policy announcement (see Appendix Table 4.9). I draw qualitatively similar conclusions from 15-minute interval regressions (see Table 4.6 and Appendix Table 4.11 for time interval dummies).

Table 4.5: Results of 10-Minute Regression

	Model 1	Model 2	Model 3	Model 4
Constant	0.198*** (0.013)	-0.747*** (0.080)	-1.270*** (0.117)	-1.279*** (0.132)
Liquidity Volatility	3.232*** (0.728)	1.852*** (0.486)	2.077*** (0.476)	1.893*** (0.405)
Price		0.413*** (0.035)	0.536*** (0.038)	0.565*** (0.045)
Surprise			-0.156*** (0.011)	-0.139*** (0.011)
FK-Document			0.020*** (0.004)	
FK-Inflation				0.002 (0.003)
FK-Forecasting				0.007*** (0.001)
FK-Monetary Policy				-0.001 (0.002)
FK-Financial Markets				0.003 (0.002)
FK-Economic Activity				0.005 (0.004)
Adjusted R ²	0.082	0.586	0.630	0.634
F-statistic	1.314	18.303	26.645	21.498
AIC	-1368.99	-1876.73	-1948.48	-1952.12

Notes: The table reports results of regression models estimating the effect of SARB communication on bond liquidity. The dependent variable is liquidity, measured as the average bid-ask spread during the interval. Time dummies ($I1 \rightarrow I18$) are included in the 10-minute intervals models but omitted for brevity. They represent time intervals around the SARB announcement which is usually at 3:00 PM. Volatility is the standard deviation of the quoted spread during a given interval. FK measures the level of linguistic complexity of the monetary policy statement meaning the number of years of education required to comprehend the statement text and the Surprise is simply the difference between new policy rate and the previous policy rate. Sample Period: MPC meetings between July 2015 - July 2023. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. A detailed table of time dummy coefficients is provided in the appendix.

Table 4.6: Results of 15-Minute Regression

	Model 1	Model 2	Model 3	Model 4
Constant	0.193*** (0.013)	-0.742*** (0.094)	-1.268*** (0.135)	-1.294*** (0.152)
Liquidity Volatility	3.493*** (0.915)	2.103*** (0.624)	2.333*** (0.610)	2.147*** (0.525)
Price		0.409*** (0.041)	0.531*** (0.044)	0.567*** (0.052)
Surprise			-0.155*** (0.012)	-0.137*** (0.012)
FK-Document			0.020*** (0.004)	
FK-Inflation				0.002 (0.003)
FK-Forecasting				0.007*** (0.001)
FK-Monetary Policy				-0.001 (0.002)
FK-Financial Markets				0.003 (0.002)
FK-Economic Activity				0.006 (0.005)
Adjusted R ²	0.102	0.605	0.649	0.653
F-statistic	1.319	19.366	27.042	20.472
AIC	-976.21	-1344.90	-1397.07	-1399.05

Notes: The table reports results of regression models estimating the effect of SARB communication on bond liquidity. The dependent variable is liquidity, measured as the average bid-ask spread during the interval. Time dummies ($I1 \rightarrow I18$) are included in the 15-minute intervals models but omitted for brevity. They represent time intervals around the SARB announcement which is usually at 3:00 PM. Volatility is the standard deviation of the quoted spread during a given interval. FK measures the level of linguistic complexity of the monetary policy statement meaning the number of years of education required to comprehend the statement text and the Surprise is simply the difference between new policy rate and the previous policy rate. Sample Period: MPC meetings between July 2015 - July 2023. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. A detailed table of time dummy coefficients is provided in the appendix.

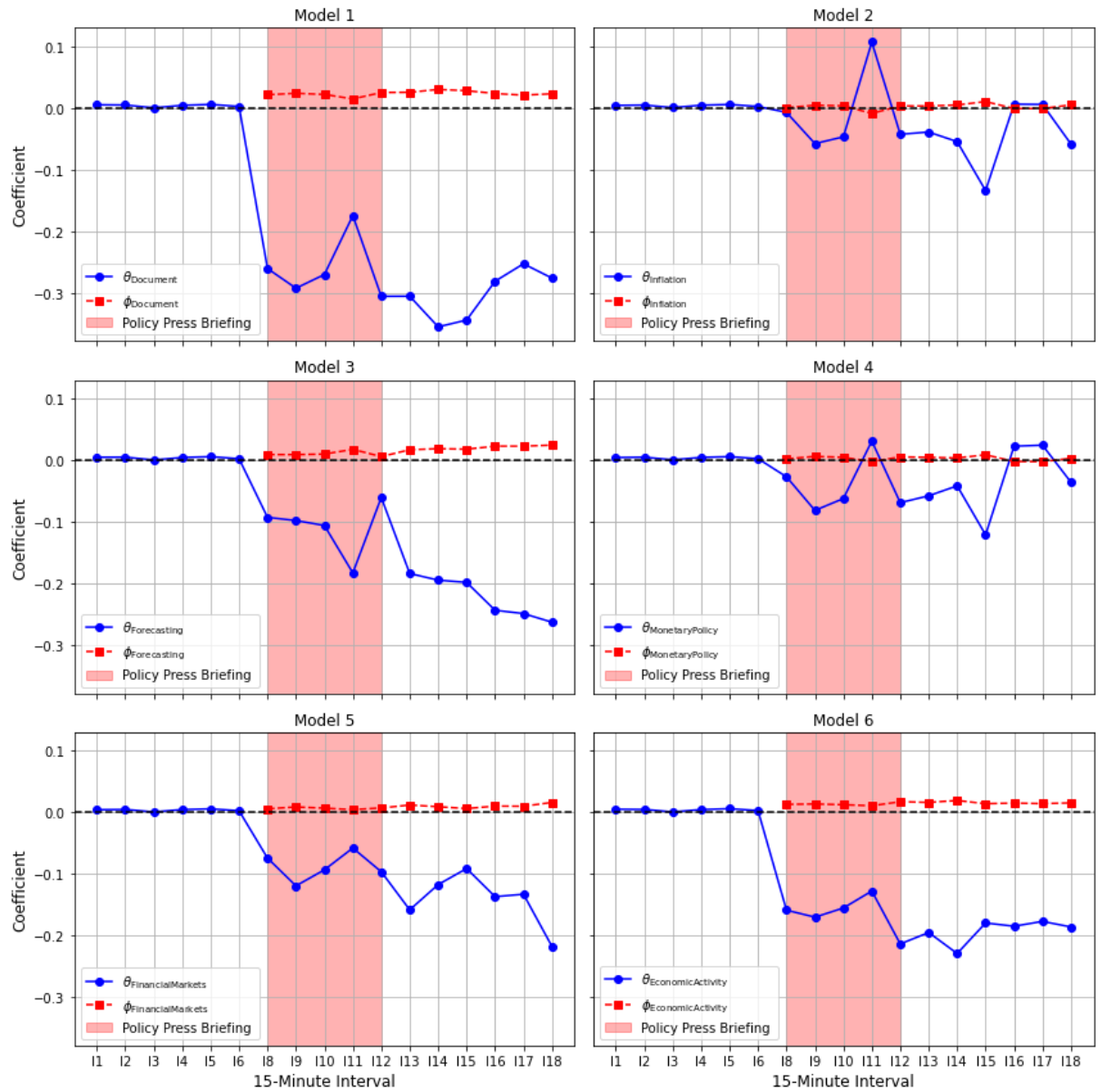
5.3.2 Identifying Time Varying Effect of Topic Complexity

In order to identify the time varying effect of communication complexity, at both overall and topic-specific level, I estimate full models which include the financial market micro-structure factors and the monetary policy surprise factor. Using 10-minute intervals data, I introduce in separate models the complexities of the overall communication (document), and key topics, namely, inflation, forecasting, monetary policy, financial markets and economic activity. The models include time interval dummies ($I1 \rightarrow I18$) and interactions, $(I_j \times Topic_{i,t})$ for $I8 \rightarrow I18$ with specific *FK* indices. The time interval dummies are for the periods before, during and after the press briefing till the financial markets close (2:00 PM to 5:00 PM).

The results show that, first, liquidity is relatively lower in the periods before the MPC press conference and begins to improve as the conference progresses. The results indicate that the time interval coefficient dummies change sign from positive to negative, which suggests an improvement in liquidity. However, the dummy coefficients are not statistically significant. Second, to answer the question: How does liquidity evolve across different time intervals when the individual topics' complexity is introduced? The estimates of interaction coefficients ϕ_j s have well-defined positive signs which suggest that high linguistic complexity has the effect of lowering liquidity. These results are broadly consistent with the distributional shifts in the evidence already presented earlier.

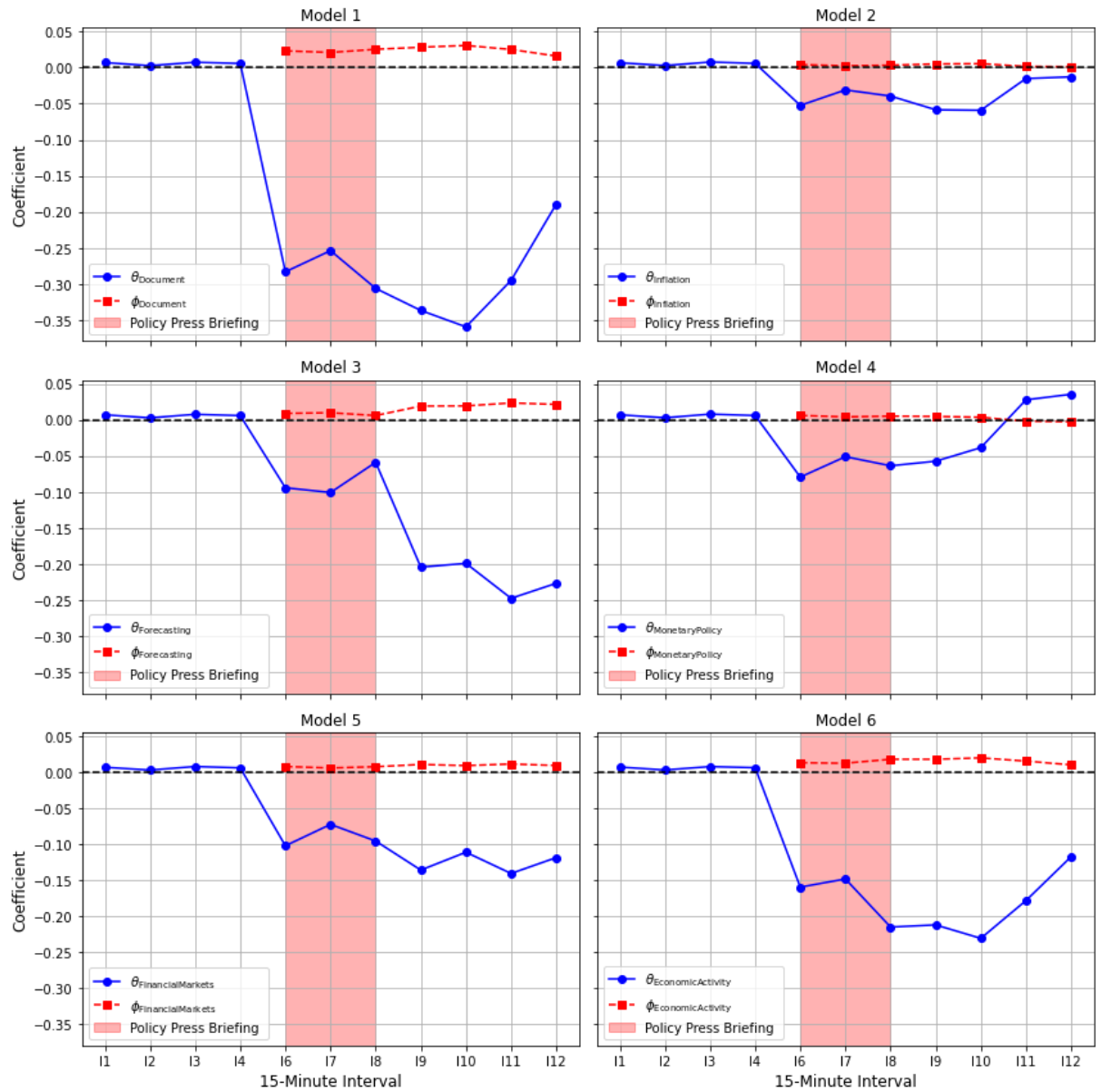
More specifically, the results indicate that when the monetary policy statement or specific topics such as forecasting are complex, market participants trade less, suggesting that they struggle to interpret policy signals, leading to increased uncertainty and reduced liquidity. Other topics namely, financial markets and economic activity also have significant interaction interval dummies which suggests similar effects. The estimated interaction coefficients, ϕ_j s after *I7* are plotted in Figure 4.6 and Figure 4.7 for the 10- and 15-minute intervals, respectively. The blue line plots θ_j s, which displays time dummy coefficients without interaction terms, illustrating the briefing's overall effect on liquidity. The red line plots ϕ_j s, which isolate the effect of communication complexity on liquidity

by interacting communication complexity with time dummies. Furthermore, estimation results show that the model with individual topics complexities has the highest explanatory power of all models – as demonstrated by the higher adjusted R^2 and lower AIC. For full regression results, see Appendix Table 4.5 and Table 4.6.



Notes: The plots show the expected evolution of liquidity over time before and after the SARB MPC press conference. Coefficients of interval dummies and interaction dummies in models help to explain liquidity in the 10-year Government Bond market on announcement day. Interactions are with the linguistic complexities of the whole document in Model 1, Model 2 with Inflation, Model 3 with Forecasting, Model 4 with Monetary Policy, Model 5 with Financial Markets and Model 6 with Economic Activity. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.6: 10-Minute Dummy Coefficients



Notes: The plots show the expected evolution of liquidity over time before and after the SARB MPC press conference. Coefficients of interval dummies and interaction dummies in models help to explain liquidity in the 10-year government bond market on announcement day. Interactions are with the linguistic complexities of the whole document in Model 1, Model 2 with Inflation, Model 3 with Forecasting, Model 4 with Monetary Policy, Model 5 with Financial Markets and Model 6 with Economic Activity. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.7: 15-Minute Dummy Coefficients

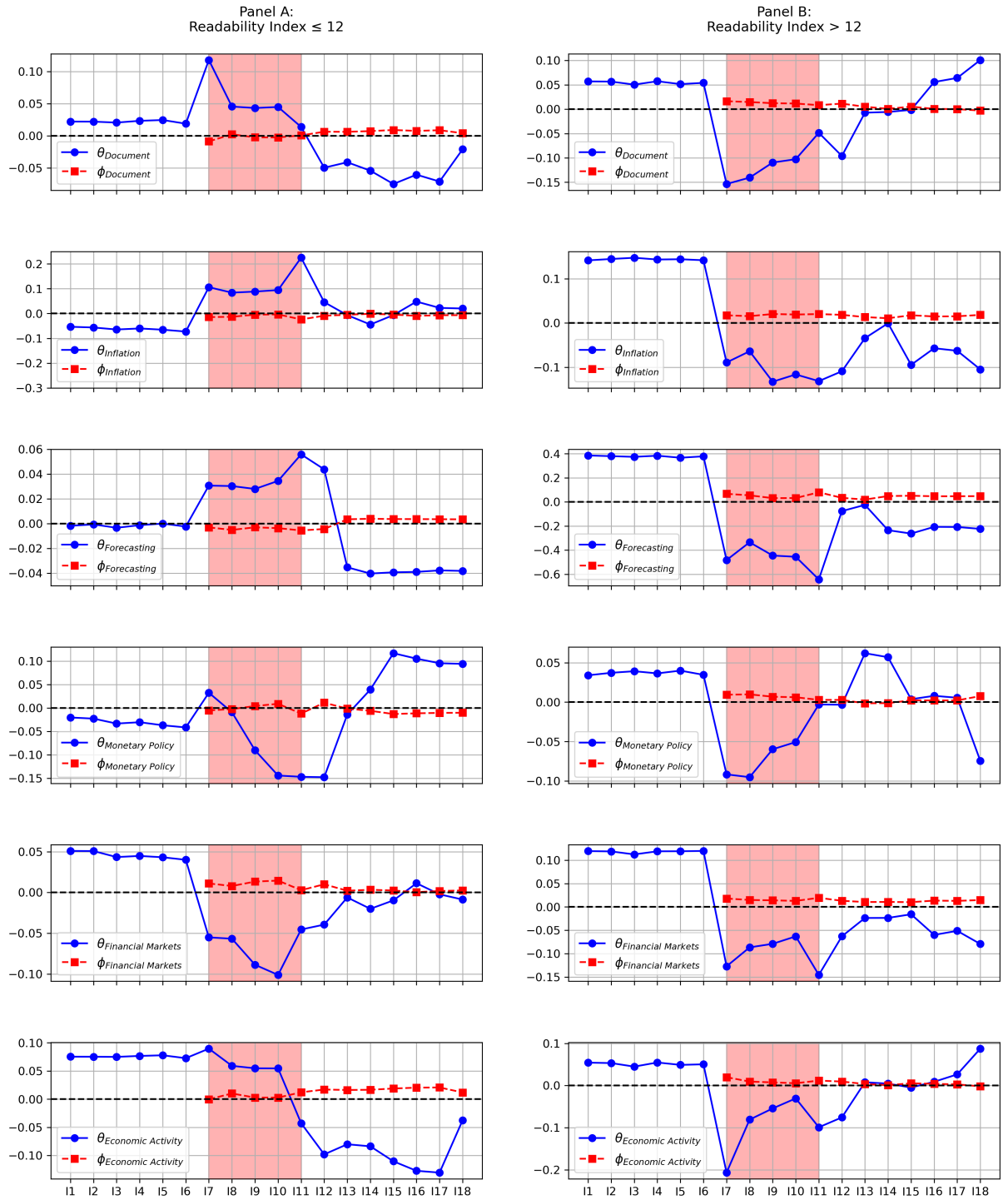
5.4 Additional Results

In order to evaluate the robustness of the regression results presented above, I create panel datasets that use 10-minute time intervals to divide both overall document complexity and topic-specific complexity into two categories: easy and hard. For example, a specific data set is created for times when the overall complexity of communication is low, and the same is true for times when the complexity is high. For every distinct topic, this procedure is repeated. The categorisation approach follows a similar strategy to the distributional shift analysis approach above. Crucially, the primary findings are consistent across these subsets, corroborating qualitatively comparable findings (see Figure 4.8). Panel A displays the findings for relatively easy communication (readability index ≤ 12), while Panel B displays the findings for hard communication (readability index > 12).

My results generally show that the blue lines for all issues falls around the time of the press briefing, indicating that market liquidity improves, probably as a result of reduced uncertainty. However, as the red lines illustrate, these advantages are either eliminated or inverted when communication complexity is introduced. This suggests that the market may find it more challenging to understand and react appropriately to central bank announcements if they become more complex. The results provide credence to the idea that while central bank communication usually enhances market functioning, the benefits vary depending on the communication's delivery quality. Complex language diminishes liquidity and makes it harder to grasp the market, especially when talking about forward-looking subjects like forecasting. Moreover, the interaction coefficients ϕ_{js} are larger in the high complexity conditions, indicating that market liquidity deteriorates and is relatively lower when language is complex.

In summary, my results indicate that consistent with the literature, it is important to investigate the impact of central bank communication on financial markets. However, by disentangling communication and estimating the complexity of topics covered, I show that even if the SARB has a decreasing trend of communication complexity, there is still work to be done when they discuss the forward-looking topic as there is evidence that when communication on this topic is complex, there is a significant negative effect on liquidity in the financial markets.

Finally, the advantage of this approach is to provide for the per-topic complexity scores for the case of the SARB. However, the approach is generalisable. So, the study provides monetary authorities with practical insights that can enable them to focus on improving specific areas of communication by using a more appropriate complexity model. My findings support the significance of looking into how central bank communication affects financial markets, which is in line with the body of existing literature.



Coefficient estimates for average liquidity over 18 consecutive 10-minute intervals before, during and after the SARB MPC press conference. The solid blue line shows the interval-dummy coefficients (θ_j), and the dashed red line shows the topic-specific complexity coefficients (ϕ_j). The shaded region marks the duration of the press briefing.

Figure 4.8: Dummy Coefficients by Complexity Category

In summary, my results differ from previous studies in two key ways. First, while research in other countries typically finds a positive relationship between liquidity and monetary policy surprise, those studies measure the surprise as the difference between the new policy rate and the corresponding futures rate capturing market expectations. In contrast, the current study uses an alternative measure of surprise that does not incorporate expectations, primarily because I lacked access to South African futures rate data. This difference in measurement may explain the divergence in liquidity responses.

Second, whereas existing literature generally assesses the overall linguistic complexity of central bank communication, my approach distinguishes between the complexity of the entire policy statement and the complexity of its key topics. This topic-specific analysis allows for a more nuanced examination of how different aspects of communication affect liquidity. And a key weakness of the study is the use of the FK metric whose parameters are calibrated based on U.S data as this may not accurately represent the comprehension capabilities of the South African public. Finally, the advantage of this targeted approach is that it provides monetary authorities with an approach that focuses on improving specific areas of communication that may contribute to liquidity disruptions, ultimately leading to more stable market outcomes.

6 Conclusions

Despite having a strong foundation, research on the complex relationship between monetary policy communication and financial markets is still evolving. By offering fresh empirical evidence that highlights South Africa's financial market, which has been widely overlooked in this area of research, the current study contributes to the current and ongoing scholarly debate on the effects of central bank communication on financial instruments. By showing how South Africa's 10-year government bond liquidity declines on monetary policy announcement days in comparison to control days, my analysis supports earlier studies from advanced economies. In line with the literature, my results also suggest that complex communication reduces liquidity because investors face higher information asymmetry or uncertainty. This observation is also consistent with the theory that unclear public signals make it harder for agents to coordinate or react confidently. The results

suggest that the degree of liquidity reduction varies according to the complexity of the central bank's communications. The study also shows that negative market reactions are caused by the complexity of various topics covered in policy documents, especially related to economic forecasting.

The novel disaggregated approach for evaluating communication complexity contributes insightful information to ongoing research discussions, as well as policy-related debates. This study evaluates the complexity of monetary policy statements and identifies individual topical components, thus going beyond aggregate measures commonly used in the literature. The degree of complexity in the main thematic areas of policy documents is evaluated using a well-known readability metric, the Flesch-Kincaid grade level, and advanced text analysis techniques. The study identifies five main topics that the SARB discusses, namely, forecasting, financial markets, monetary policy, inflation, and economic activity.

This study proceeds to provide insights of how the complexity of different topics covered in monetary policy statements affects the financial markets in South Africa. This is done by analysing the complexity of the internal dimensions of policy communication and the overall communication measure of complexity. An assessment of the 10-year government bond market prior to, during, and following announcements is made using time dummies and interaction time interval dummies to examine liquidity responses over a range of time periods. This also helped to capture how market sensitivity to policy communication evolves over time.

The key findings suggest that although overall SARB communication has become more transparent in recent years, this improvement is driven almost exclusively by the handful of topics, primarily forward-looking economic outlooks, that receive greatest emphasis at each meeting. Themes relating to upside and downside financial risks have remained comparatively opaque, and when discussed in a complex manner they substantially impair bond-market liquidity. The study demonstrates this through density-plot comparisons of

quoted spreads on announcement versus control days, the Kolmogorov-Smirnov (KS)-tests and via time-varying regression analyses at 10- and 15-minute intervals, which reveal persistent, right-skewed spread spikes in both the mean and volatility for high-complexity of forecasting, monetary policy and financial markets.

These results carry clear policy implications. Central banks should look beyond document-level readability and target clarity improvements in those topics most prone to opacity and market sensitivity. In practice, SARB and other central banks can use this topic-specific language complexity approach to prioritise simpler language in key areas of discussions, thereby smoothing intra-day liquidity dynamics and reducing uncertainty. The study provides a data-based framework that enables in-depth examination of the distinct effects on market liquidity of monetary policy communication by subject.

Chapter 4: Appendix

A The SARB Key Features

The SARB was established by an Act of Parliament on June 30, 1921, and is the oldest central bank in Africa. Further, it is also noteworthy because it is the fourth oldest central bank outside Europe and the United Kingdom (preceded only by the U.S. Federal Reserve, Bank of Japan, and Bank of Java). It is owned by private individuals who are also members of the public and this is one of the features that sets it apart from other central banks located all over the world. The SARB has issued a total of two million shares, and no single shareholder has more than 10,000 shares for themselves. Furthermore, the South African government is eligible to receive an annual dividend on any earnings that remain after the dividend has been disbursed. This payment is based on any earnings that result from the distribution of the dividend.

The primary aim of the SARB is; “To protect the value of the currency in the interest of balanced and sustainable economic growth in the Republic”. This is stated explicitly in the Constitution of the Republic of South Africa. The SARB follows a flexible inflation-targeting system that was achieved in agreement with the National Treasury. In its implementation framework, it is the job of the Monetary Policy Committee of the SARB to establish the policy rate in order to steer inflation towards the middle of a target band, which is between three and six percent, throughout the course of the medium term. This is done with considerable caution in order to decrease the probability of excessive output and employment instability.

The South African Reserve Bank is subject to the supervision of the Board of Directors, which is comprised of the Governor, three Deputy Governors, and eleven non-executive directors. The Board of Directors is led by the Governor. The Minister of Finance has suggested that the President should be given the task of selecting the Governor and Deputies for a term of five years. As a result, the President is the one who is accountable for making this selection. On the other hand, the non-executive directors of the bank are appointed for a period of three years by either the shareholders of the bank or the President of the

bank. As part of its efforts to achieve a balance between democratic responsibility and operational autonomy, the Bank keeps an open line of communication with the Minister of Finance. In spite of the fact that the Bank is constitutionally independent and possesses the right to act “without fear, favour, or prejudice”.

In the month of June in the year 2022, the SARB commenced the implementation of a tiered-floor method for the aim of carrying out monetary policy. According to the SARB, the deposit rate acts as a floor for the rates that are used in the money market. Furthermore, banking institutions get interest on surplus reserves that are kept overnight at the SARB. As a result of this, the costs associated with implementation have been decreased, and at the same time, the dependability of the transmission of decisions made by the MPC into market rates has been ensured. The former shortage system, which required the South African Reserve Bank to purposely reduce liquidity in order to produce financing scarcity, has been replaced by this structure.

Not only is the SARB responsible to safeguard the stability of prices, but it is also obligated to protect the stability of the financial system, as stipulated by the legislation. The South African Financial Stability Department is responsible for continuously monitoring macroprudential policy tools, identifying systemic risks, and working in collaboration with both domestic and international partners in order to mitigate potential vulnerabilities to the South African financial system. This is done in order to ensure that the South African financial system remains economically stable. This is done in order to guarantee that the South African financial system will continue to be secure and stable. In addition, the Bank is responsible for the supervision of the nation’s payment system, which includes the obligation of ensuring the safety, efficiency, and robustness of retail payments, interbank settlement, and the Real-Time Gross Settlement (SAMOS) system. This responsibility falls under the Bank’s scope.

In its capacity as the issuer of the rand (ZAR), the SARB is responsible for the production, printing, and distribution of coins and banknotes. This responsibility includes the distribution of the rand. The SARB is also accountable for regularly improving security features in order to prevent counterfeiting. In addition to its role as the government’s banker and advisor, the Reserve Bank of South Africa is also responsible for the upkeep

of the country's official gold and foreign exchange reserves. In difficult situations, the Reserve Bank of South Africa may also act as a lender of last resort. A further responsibility that falls under the scope of the Reserve Bank of South Africa is the administration of the nation's gold reserves. The primary objective of these reserves is to mitigate the effects of fluctuations in the currency rate and fulfill the obligations that are linked to the balance of payments.

In summary, the SARB is responsible for the creation of comprehensive economic information and research. In the course of the Financial Stability Review that is carried out by the Bank, a comprehensive analysis of vulnerabilities across a wide variety of sectors and counterparties is presented. On the other hand, the Monetary Policy Review, which is carried out by the Bank, is responsible for assessing inflation projections, financial risks, and the economic conditions that are taking place both locally and internationally. The SARB wants to improve the accountability, transparency, and efficiency of its policy instruments by disseminating these releases and maintaining regular communication with the general public, markets, and academic institutions. This will be accomplished through the dissemination of these releases.

B The SARB MPS (1999 - 2023)

The Monetary Policy Committee statements that were issued by the South African Reserve Bank between October 1999 and July 2023 had a total of 300,323 words generated. Whenever it comes to the dissemination of economic analysis, policy decisions, and future recommendations, each statement typically consists of an average of 1,988 words, which displays a considerable and consistent degree of depth across the board. The cumulative volume is evidence that the SARB has been committed to transparency and providing a comprehensive justification for its monetary policy over the course of more than twenty years. This commitment has been demonstrated during the course of the data collection process.

Given the passage of time, it is becoming increasingly apparent that there is a distinct trend towards increasing the length of remarks. In the early years (1999–2002), when statements frequently ranged in length from 1,400 to 2,200 words, a format that was more compact was represented. Beginning around the year 2003, the word counts gradually began growing into the region of 2,000–3,000. This growth continued until the year 2009. By the middle of the 2010s, the majority of statements comfortably topped 2,500 words, and by July 2014, they had reached a peak of over 3,400 words (Statement 96). There are a number of factors that have likely contributed to this growth. Some of these factors include the increasing complexity of the domestic and international economic landscapes, the increased range of data and analytical tools that are available to the SARB, and the aspiration of the central bank to provide its economic judgements with a more nuanced context.

Despite the fact that there is a general trend towards growing, the series includes a number of noteworthy outliers that are dispersed over the entirety of the series. Statement 10, which was made in January of 2002, is the statement that is the shortest, with only 135 words being included in it. This is a unique occurrence that raises the possibility that a release will be issued in a particular manner or that a shortened summary will be utilised in place of the comprehensive analytical narrative that is often utilised for other meetings. Statement 96, on the other hand, which was generated on July 17, 2014, has a total of 3,401 words, and a number of other statements from that time period had a cumulative total of more than 3,000 words. In times of increasing global unpredictability, when markets and decision-makers required exceptionally precise communication, such as the aftermath of the Eurozone debt crisis and shifting commodity prices, these peaks correspond with periods of heightened global unpredictability. In other words, these peaks coincidentally occur at the same time.

By comparing the averages from one year to the next, the SARB provides an additional indication of the transformation of its communication approach. At the beginning of the decade, in the early 2000s, annual averages were less than 2,200 words. The number of words, on the other hand, had increased to somewhere between 2,500 and 2,800 by the time the decade of 2010–2020 turned around. This statement has been lengthened in step

with broader industry trends among intelligent central banks, which are moving toward announcements that are more narratively driven and packed with data. This statement has been lengthened in accordance with these trends. Because they offer a helpful longitudinal record of South Africa's monetary policy framework, economic outlooks, and the shifting balance of risks over time, researchers and practitioners can benefit from the increasing breadth and volume of SARB communications. This is because these communications give a useful record of historical data (see Appendix Table 4.7).

Table 4.7: SARB MPC Statements Word Count, 1999 - 2023

Statement No.	MPC Dates	Word Count	Statement No.	MPC Dates	Word Count	Statement No.	MPC Dates	Word Count
1	13-Oct-99	1454	52	11-Oct-07	1872	102	23-Jul-15	2828
2	24-Nov-99	1913	53	06-Dec-07	1582	103	23-Sep-15	2599
3	13-Jan-00	2090	54	31-Jan-08	1406	104	19-Nov-15	2433
4	02-Mar-00	2082	55	10-Apr-08	1722	105	28-Jan-16	2617
5	06-Apr-00	2076	56	12-Jun-08	1815	106	17-Mar-16	2513
6	19-May-00	2080	57	14-Aug-08	1714	107	19-May-16	2542
7	15-Jun-00	2047	58	09-Oct-08	1815	108	21-Jul-16	2496
8	11-Aug-00	1901	59	11-Dec-08	1759	109	22-Sep-16	2561
9	21-Sep-00	1995	60	05-Feb-09	1408	110	24-Nov-16	2403
10	16-Oct-00	294	61	24-Mar-09	1353	111	24-Jan-17	2600
11	16-Nov-00	1994	62	30-Apr-09	1530	112	30-Mar-17	2412
12	19-Jan-01	2249	63	28-May-09	1265	113	25-May-17	2212
13	16-Mar-01	2054	64	25-Jun-09	1443	114	20-Jul-17	2336
14	25-Apr-01	1894	65	13-Aug-09	1416	115	21-Sep-17	2254
15	14-Jun-01	2259	66	22-Sep-09	1411	116	23-Nov-17	2466
16	26-Jul-01	2099	67	22-Oct-09	1495	117	18-Jan-18	2528
17	20-Sep-01	2179	68	17-Nov-09	1611	118	28-Mar-18	2058
18	15-Nov-01	2418	69	26-Jan-10	1951	119	24-May-18	2060
19	15-Jan-02	135	70	25-Mar-10	2061	120	19-Jul-18	1410
20	14-Mar-02	1111	71	13-May-10	1861	121	20-Sep-18	1246
21	13-Jun-02	1442	72	22-Jul-10	2105	122	22-Nov-18	1357
22	12-Sep-02	1439	73	09-Sep-10	2125	123	17-Jan-19	1280
23	28-Nov-02	1141	74	19-Nov-10	2347	124	28-Mar-19	1430
24	20-Mar-03	2822	75	20-Jan-11	2408	125	23-May-19	1296
25	12-Jun-03	1800	76	24-Mar-11	2303	126	18-Jul-19	1391
26	14-Aug-03	1607	77	12-May-11	2080	127	19-Sep-19	1232
27	10-Sep-03	1057	78	21-Jul-11	2529	128	21-Nov-19	1245
28	16-Oct-03	1673	79	22-Sep-11	2554	129	16-Jan-20	1228
29	11-Dec-03	1858	80	10-Nov-11	2281	130	19-Mar-20	1280
30	26-Feb-04	1887	81	19-Jan-12	2492	131	14-Apr-20	1311
31	22-Apr-04	1762	82	29-Mar-12	2789	132	21-May-20	1148
32	10-Jun-04	1580	83	24-May-12	2438	133	23-Jul-20	1109
33	12-Aug-04	2111	84	19-Jul-12	2652	134	17-Sep-20	1174
34	14-Oct-04	1667	85	20-Sep-12	3030	135	19-Nov-20	1239
35	09-Dec-04	2060	86	22-Nov-12	3090	136	21-Jan-21	1136
36	10-Feb-05	1887	87	24-Jan-13	2850	137	25-Mar-21	1336
37	14-Apr-05	2031	88	20-Mar-13	2930	138	20-May-21	1230
38	09-Jun-05	1849	89	23-May-13	2807	139	22-Jul-21	1410
39	11-Aug-05	1937	90	18-Jul-13	3051	140	23-Sep-21	1321
40	13-Oct-05	1756	91	19-Sep-13	3112	141	18-Nov-21	1487
41	08-Dec-05	1759	92	21-Nov-13	2874	142	27-Jan-22	1530
42	02-Feb-06	1695	93	29-Jan-14	3048	143	24-Mar-22	1913
43	13-Apr-06	1869	94	27-Mar-14	3227	144	19-May-22	2010
44	08-Jun-06	1910	95	22-May-14	2967	145	21-Jul-22	2033
45	03-Aug-06	1844	96	17-Jul-14	3401	146	22-Sep-22	2072
46	12-Oct-06	2042	97	18-Sep-14	3222	147	24-Nov-22	2101
47	07-Dec-06	1996	98	20-Nov-14	3002	148	26-Jan-23	2374
48	15-Feb-07	2018	99	29-Jan-15	3043	149	30-Mar-23	2253
49	12-Apr-07	1976	100	26-Mar-15	2857	150	25-May-23	2061
50	07-Jun-07	2120	101	21-May-15	2792	151	20-Jul-23	1952
51	16-Aug-07	1995					Total	300,323

C Monetary Policy Cycles and Key Drivers

Between June 2003 and April 2006, the South African Reserve Bank (SARB) had a generally accommodating approach, which resulted in a reduction of the repurchase rate. This was done in response to a good medium-term outlook and a benign inflation backdrop. Price pressures around the world remained modest, monetary growth was maintained at a moderate pace, and inflation itself was showing a declining trend. Because both core and headline inflation remained comfortably below goal and showed few indications of overheating, the Bank was able to loosen policy and increase domestic demand without running the danger of a return of price pressures. This was made possible by the fact that the Bank was able to raise domestic demand.

On the other hand, by the middle of 2006, there were indications of increased price risks, which led to a turn to tightening. As domestic inflation began to increase and the outlook became more pessimistic, the SARB increased interest rates between June 2006 and October 2008. The rise in the price of oil, the increase in the cost of electricity, and the concurrent pressures on currency rates all contributed to a reduction in real incomes and threatened to alter overall consumer price measures. As a result of the continuing, albeit patchwork, downturn in global commodity prices and the uncertainties surrounding the policy of the United States Federal Reserve regarding interest rates, the picture became even more difficult. This caused the Monetary Policy Committee to adopt a more restrictive stance in order to anchor inflation expectations.

As a result of the global financial crisis, the South African Reserve Bank shifted its course of action and implemented a considerable policy easing from December 2008 to November 2013. There was a significant decrease in imported inflationary pressures as a result of a sudden slowdown in global demand and steep declines in oil prices. In addition, a stronger rand significantly lowered cost-push impulses. In response to weak conditions on a worldwide scale and a stretched economy in the United States, the South African Reserve Bank decreased interest rates in an effort to be supportive of economic expansion. The easing cycle that occurred during these years is a reflection of the South African Reserve Bank's twin mandate, which is to stimulate output while also stabilising prices in the event that external shocks need countercyclical action.

Most recently, in the period following COVID-19, the Bank was able to successfully handle two distinct incidents. During the period beginning in January 2020 and ending in September 2021, it once again adopted an accommodation bias, which reflects dovish attitudes taken by central banks of advanced economies, a muted outlook, strength in the rand, and moderated global price pressures. The Monetary Policy Committee decided to tighten monetary policy from November 2021 to September 2023. This decision was made because the effects of social unrest, pandemic-related budgetary stresses, and disruptions in global supply chains were becoming more pronounced, resulting in persistent pricing pressures. A return to restrictive policy was required in order to re-anchor inflation expectations after a number of factors contributed to inflation that was high. These factors included rising oil prices, decreased commodity exports, a depreciating currency, and the shock of the conflict between Russia and Ukraine. The Appendix Table 4.8 summarises the SARB policy stance in bullet presentation.

Table 4.8: Monetary Policy Cycles and Key Drivers

Period	Policy	Key Drivers
Jun 2003 – Apr 2006	Loosening	• Falling inflation and inflation outlook • Low monetary growth • Low global inflation
Jun 2006 – Oct 2008	Tightening	• Rising inflation and unfavourable outlook • Exchange rate pressures • Rising oil prices • Low commodity prices • Rising electricity prices • Uncertain US interest rates
Dec 2008 – Nov 2013	Loosening	• Falling inflation and favourable outlook • Declining oil prices • Support to domestic economy (GFC) • Ailing world economy • Appreciating domestic currency
Jan 2014 – May 2017	Tightening	• Unfavourable inflation outlook • Exchange rate pressures • Ailing world economy • Droughts in 2016 impacting food prices
Jan 2020 – Sep 2021	Loosening	• Stable inflation and inflation outlook • Strong exchange rate • Support to domestic economy • Low global inflation • Accommodative policies in advanced central banks
Nov 2021 – Sep 2023	Tightening	• Unfavourable inflation outlook • Weaker commodity export prices vs. rising oil prices • Weakening local currency • Lasting impact of the pandemic • Social unrest • Ukraine–Russia war disruption of production

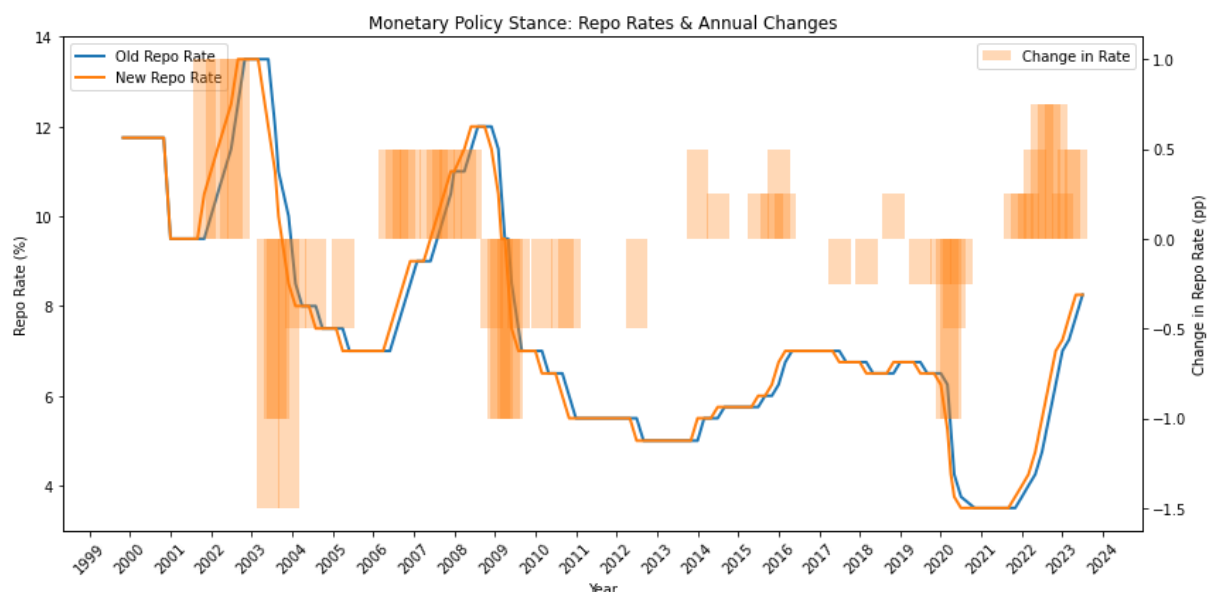
D SARB Policy Rate Overtime

The policy rate changes were as Follows:

The South African Reserve Bank made some of the most significant decisions about the policy rate throughout the time period beginning in November 1999 and ending in July 2023, and the dataset was created to commemorate those decisions. Every single observation is made up of three different fields, and each one is connected to a policy-review date, which occurs either on a monthly or quarterly basis.

In my hard data, when referring to the policy (repo) rate that was in place immediately before to the decision, the term “Old repo rate” is used. “New repo rate” is the name given to the policy rate that was decided upon in that particular decision. Change is defined as the difference between the new rates and the old rates, expressed as a percentage point difference; for instance, a rate increase of one point would be represented by a figure of +1.00.

The rates are always stated as annualised percentages, regardless of the circumstances. For each and every policy date, there is a pre- and post-decision rate that corresponds to it, in addition to the computed change. Throughout the entirety of the series, there are no missing values anywhere. Not only does this framework give us the ability to see the progression of the level of monetary policy, but it also gives us the ability to see the magnitude and progression of policy changes through the utilisation of bars on a secondary axis for Change. This is accomplished through the utilisation of line charts for both Old repo and New repo (see Appendix Figure 4.9).



Notes: The plot show the decisions about the policy rate throughout overtime. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.9: SARB Policy Rate Changes

E Supplementary Related Literature

Given that affective responses to policy pronouncements have the potential to either magnify or minimise economic volatility, the relationship between market emotions and the clarity of communication is of the utmost importance. To evaluate the influence of central bank communication, some studies in the literature have used text mining techniques on various monetary policy documents. For example, according to Blommestein et al. (2012) and Khatri et al. (2014), the language of central banks has the ability to arouse animal spirits, also known as investor optimism and pessimism, which in turn have an effect on trading volumes and equities returns. Through the creation of a worldwide investor sentiment index based on survey data and the demonstration that excessive optimism anticipates lower future returns, Schmeling (2009) draws attention to the feedback loop that exists between the tone of the central bank and the attitude found in the market. Furthermore, Da et al. (2015) investigates the discourse in the media on monetary policy and finds that a change toward more negative rhetoric in the financial press enhances bond market volatility. This is further supported by the findings of the aforementioned researchers.

Quantifications of sentiment and emotion based on text have become increasingly sophisticated in recent years. For the purpose of calibrating sentiment, early studies (Tetlock (2007) and Garcia (2013)) utilised the percentage of positive terms in financial texts as opposed to the percentage of negative words. Loughran and McDonald (2011) conducted subsequent research that resulted in the creation of domain-specific lexicons with the purpose of improving sentiment classification in corporate filings. Apergis (2015) and Tumala and Omotosho (2019) utilised this method in order to forecast fluctuations in output growth and exchange rates. Researchers in the field of psychology, such as Ekman (1973), identified six fundamental emotions that extend beyond simple polarity. These emotions are happiness, sorrow, anger, fear, surprise, and disgust. Researchers Consoli et al. (2021) and Khatri et al. (2014) applied multivariate emotion detection algorithms to macroeconomic news and discovered that different negative emotions (for instance, fear versus rage) have differential forecasting power for credit risk and sovereign yield spreads. This was discovered by comparing the predictive power of these two emotions. As a whole, this body of research demonstrates that the characteristics of decision makers and the intricate emotional content of their messages, which markets decipher using a range of text analytic techniques, are essential components of policy communication that is both clear and trustworthy.

F Other Readability Measures

It is possible to evaluate the readability of monetary policy papers by utilising the Flesch Reading Ease (FRE) score, which ranges from 0 (very difficult) to 100 (extremely easy). The computation of it is

$$\text{FRE} = 206.835 - 1.015 \left(\frac{\text{words}}{\text{sentences}} \right) - 84.6 \left(\frac{\text{syllables}}{\text{words}} \right) \quad (4.4)$$

According to Flesch (1948), simpler text has a higher FRE. Despite the fact that it is well-known for its straightforward 0–100 scale, automated implementations may be inaccurate due to the fact that they are dependent on precise syllable counts; they do not directly transform into a grade-level equivalent.

With the help of the Flesch-Kincaid Grade Level (FKGL), these same inputs are transformed into an approximate level of education for students in the United States:

$$\text{FKGL} = 0.39 \left(\frac{\text{words}}{\text{sentences}} \right) + 11.8 \left(\frac{\text{syllables}}{\text{words}} \right) - 15.59 \quad (4.5)$$

If a student's score is 12.3, for example, they should be able to understand the subject (Kincaid (1975)). These findings were published in the year 1975. Despite the fact that it is dependent on robust syllable parsing, FKGL is straightforward, easy to compute, and incorporated into a number of other text-analysis programs.

The Gunning Fog Index (GFI) takes into consideration polysyllabic words, which are also referred to as “complex words”, by reducing the frequency with which they are used.

We define “complex words” as words that have three or more syllables, and the formula for calculating G F I is as follows:

$$\text{GFI} = 0.4 \left[\frac{\text{words}}{\text{sentences}} + 100 \frac{\text{complex words}}{\text{words}} \right] \quad (4.6)$$

Despite the fact that GFI is sensitive to domain-specific vocabulary due to its concentration on jargon, it may inflate the difficulty of technical documents that contain a large number of legitimate multi-syllable words and neglect intricacies in sentence structure that go beyond the average length.

According to Mc Laughlin (1969), the SMOG Index, also known as the Simple Measure of Gobbledygook, is intended to have a significant association with comprehensive comprehension tests that consist of at least thirty sentences:

$$\text{SMOG} = 1.0430 \sqrt{\frac{\# \text{polysyllables} \times 30}{\# \text{sentences}}} + 3.1291 \quad (4.7)$$

Despite the fact that its use is restricted to shorter policy statements due to its limited scope and minimum-length constraint, it frequently outperforms alternatives in academic environments because it focuses solely on polysyllabic counts (Mc Laughlin (ibid.)).

By replacing syllables with character counts, the Coleman–Liau Index (CLI) completely sidesteps the process of linguistic parsing:

$$L = 100 \times \frac{\text{letters}}{\text{words}}, \quad S = 100 \times \frac{\text{sentences}}{\text{words}}, \quad (4.8)$$

$$\text{CLI} = 0.0588L - 0.296S - 15.8 \quad (4.9)$$

According to Coleman and Liau (1975), the number of characters per word is quite automatable, despite the fact that it may not adequately reflect the conceptual or phonetic complexity of the term. This is because it only employs the counts of letters and punctuation.

In conclusion, the Automated Readability Index (ARI) provides yet another character-based grade approximation, which is defined as follows:

$$\text{ARI} = 4.71 \left(\frac{\text{characters}}{\text{words}} \right) + 0.5 \left(\frac{\text{words}}{\text{sentences}} \right) - 21.43 \quad (4.10)$$

In contrast to grade-level measures (FKGL, GFI, SMOG, CLI, and ARI), which do allow for direct educational readings, comparative studies have shown that FRE does not allow for such readings. In spite of the fact that it requires lengthy passages, SMOG frequently demonstrates the strongest empirical link with comprehension tests. CLI and ARI perform exceptionally well in entirely automated and comprehensive investigations, but they may overlook subtleties in language. In addition to having a high predictive validity across all genres, FKGL does not have a minimum-length limit and provides quick grade-level.

For my research on central bank communications, where passages vary in length and evident policy implications need to be communicated to both professionals and non-specialist audiences, the Flesch–Kincaid Grade Level to be the most appropriate tool to use.

G Liquidity, Volatility, and Complexity Trend

G.1 Liquidity of Financial Variables

In the lines of Smales and Apergis (2017a), I define liquidity as the *Quoted Spread*. The *Quoted Spread* is an average time-weighted bid-ask spread during a given interval j . A narrower quoted spread represents higher liquidity, while a wider spread means lower liquidity. The bid-ask spread for interval j at time t is calculated as:

$$\text{Bid-Ask Spread}_{j,t} = \text{Ask Price}_{j,t} - \text{Bid Price}_{j,t}.$$

Where:

- Ask Price $_{j,t}$: The lowest price at which a seller is willing to sell during interval j at time t .
- Bid Price $_{j,t}$: The highest price a buyer is willing to pay during the same interval and time.

G.2 Volatility of Financial Variables

Definition: Taylor (2011) defines volatility of prices as the rate at which prices change. When referring to financial market variables, volatility means the extent to which the price of a financial asset such as stocks, bonds, commodities, or currencies varies (see Clements et al. (2015)). Many have shown that measuring and understanding volatility of financial asset returns is important for managing portfolios, managing risk and option pricing (see Martens and Van Dijk (2007) and Daly (2008)). Therefore, volatility is itself used to represent the extent of price fluctuations—whether they are large or small, frequent or infrequent.

In the literature, volatility is often quantified using a number measures but its commonly calculated as the standard deviation of a financial variable over a specific time period¹⁹. In this study, I measure volatility by its standard deviation. Therefore, let liq_i represent the liquidity on day i , and $\bar{liq}$ represent the expected liquidity. The volatility of historical liquidity or standard deviation σ , is given by:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (liq_i - \bar{liq})^2} \quad (4.11)$$

where:

- N is the total number of liquidity in the period.
- liq_i is the liquidity for each day i .
- $\bar{liq}$ is the expected liquidity.

There are, however, other popular measures of volatility in the literature, for example, the so called realised range. It was introduced as an improvement over traditional realised volatility measures such as the variance and standard deviation that rely solely on closing prices (see Martens and Van Dijk (2007)). The realised range captures intra-day price extremes (the highest and lowest prices). In the literature, it is argued that this measure provides a more refined estimate of volatility compared to that of the daily closing prices, as it captures intra-day price movement.

As indicated above, I adopt the standard deviation as the preferred measure of volatility. This choice is driven by the nature of my analysis, which does not involve intra-day extremes (highest and lowest values) but rather focuses on single-point estimates representing the liquidity of the 10-year government bond market.

19. see Daly (2008) for some additional measures of volatility.

G.3 The Mann-Kendall Test

The Mann-Kendall test is a popular non-parametric test used in the literature to identify trends in time series data without requiring the data to be normally distributed. It tests the null hypothesis H_0 that there is no trend versus the alternative H_1 that there is a trend (either upward or downward) (Mann (1945) and Kendall (1948)). I use this test in the current study to establish the statistical significance of the linguistic complexity trend of the SARB's monetary policy communication. The test is formulated as follows:

Mann-Kendall Statistic (S)

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i)$$

where:

$$\text{sgn}(x_j - x_i) = \begin{cases} +1, & \text{if } x_j - x_i > 0 \\ 0, & \text{if } x_j - x_i = 0 \\ -1, & \text{if } x_j - x_i < 0 \end{cases}$$

The variance of S under the null hypothesis (no trend) is given by:

$$\text{Var}(S) = \frac{n(n-1)(2n+5)}{18}$$

For datasets with ties, the variance is adjusted as:

$$\text{Var}(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^m t_i(t_i-1)(2t_i+5)}{18}$$

where t_i is the number of tied ranks in the i -th group, and m is the number of tied groups.

The standardised test statistic Z is calculated as:

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}}, & \text{if } S > 0 \\ 0, & \text{if } S = 0 \\ \frac{S+1}{\sqrt{\text{Var}(S)}}, & \text{if } S < 0 \end{cases}$$

The Mann-Kendall test is performed by comparing Z to the standard normal distribution. The null hypothesis H_0 (no trend) is rejected if:

$$|Z| > Z_{\alpha/2}$$

where $Z_{\alpha/2}$ is the critical value of the standard normal distribution at a significance level α . The study uses this test to confirm that linguistic complexity has declined significantly during the regimes of Governors Mboweni (1999–2009) and Kganyago (2014–present), while it increased during Governor Marcus’s tenure (2009–2014).

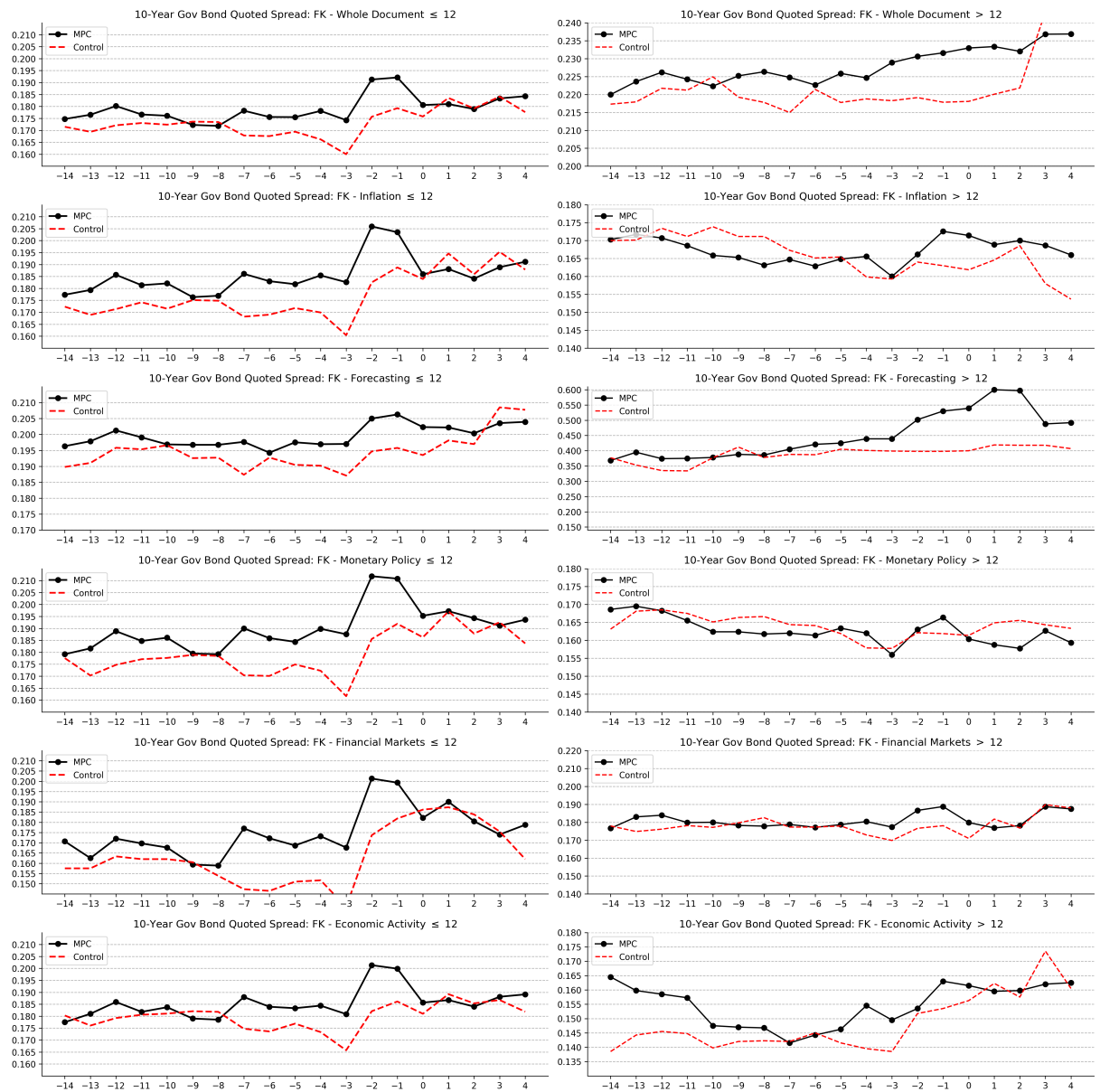
H Topic-Term Representation

An illustration of the representative terms for each topic is provided in Appendix Figure 4.10. Before applying LDA, I removed common and extraneous terms typical of the transformation from PDF to text documents, and reduced words to their roots, treating each paragraph as an independent document. The topic estimation assumes a multinomial model with Dirichlet priors, and the number of topics K is exogenously set. To determine a plausible range, I recursively estimated models for $K = 1, \dots, 10$ and selected $K = 5$ based on the highest topic coherence score, which reflects superior word co-occurrence patterns and interpretability. The five topics derived are *Inflation*, *Forecasting*, *Monetary Policy*, *Financial Markets*, and *Economic Activity*.



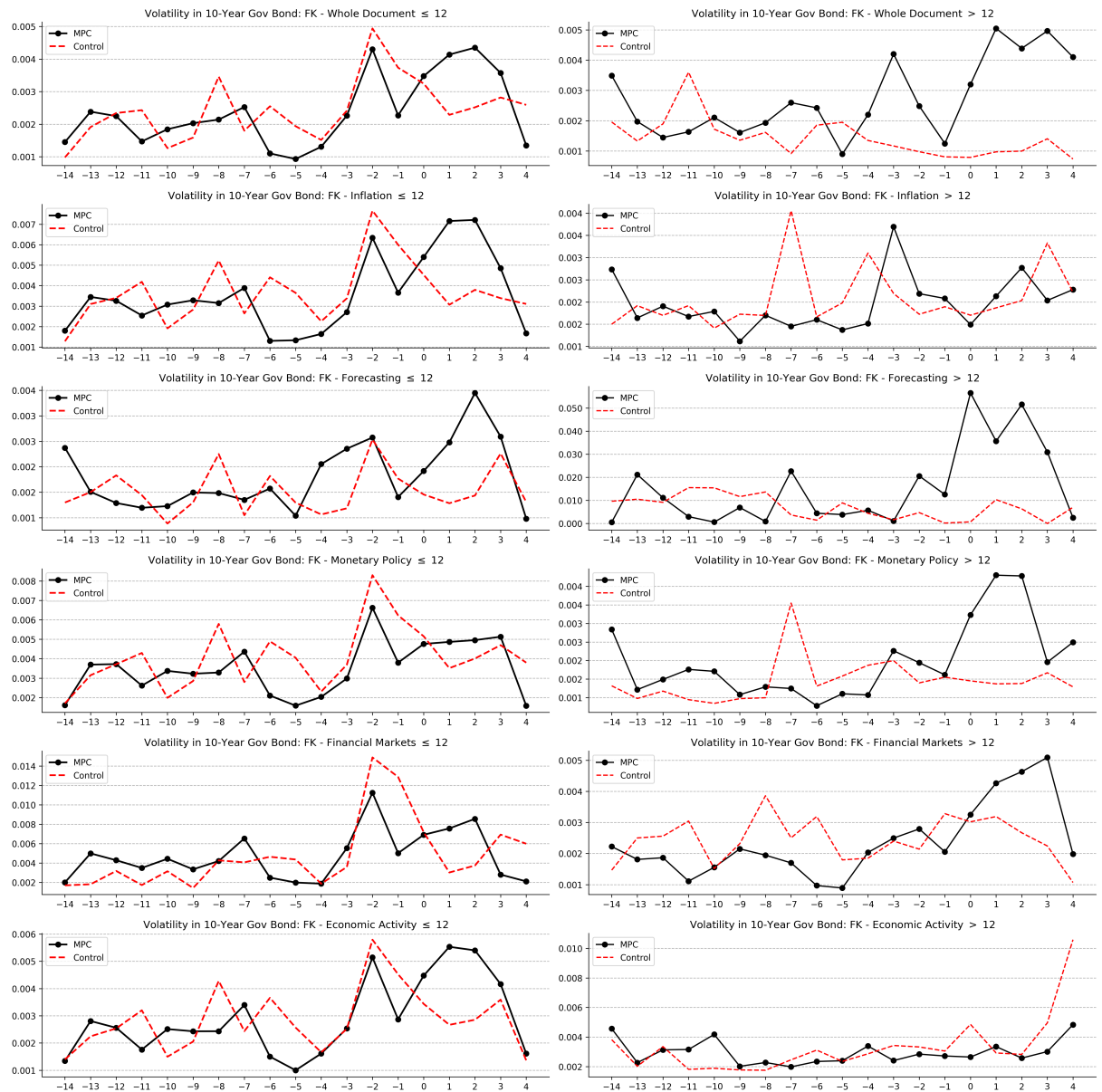
Word clouds of the topics extracted from the SARB MPC statements using LDA. Words shown in larger font appear more prominently within each topic.

Figure 4.10: Word Clouds of the Topics



Comparison of liquidity in the 10-year government bond market on announcement day and 10-day control. Note: The plots show the expected evolution of liquidity in the 10-year government bond market when the linguistic complexity of the MPC communications from the Reserve Bank of South Africa is easy or hard, in the period around the MPC announcement ($t = 0$). The reported values are the means of 30-minute intervals on the day of announcement and the 10-day control, specifically, during periods when the central bank communication is easy, ($FK \leq 12$) and when hard ($FK > 12$). Therefore, the figures represent the evolution of liquidity in the periods when the central bank communication of the overall statement and that of specific topics, namely, Inflation, Forecasting, Monetary Policy, Financial Markets and Economic Activity, is easy and when hard. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.11: Liquidity Comparisons



Comparison of volatility in the 10-year government bond market on announcement day and 10-day control. Note: The plots show the expected evolution of the volatility in the 10-year government bond market when the linguistic complexity of the MPC communications from the Reserve Bank of South Africa is easy or hard, in the period around the MPC announcement ($t = 0$). The reported values are the means of volatility of 30-minute intervals on the day of announcement and the 10-day control, specifically, during periods when the central bank communication is easy, ($FK \leq 12$) and when hard ($FK > 12$). Therefore, the figures represent the evolution of volatility in the periods when the central bank communication of the overall statement and that of specific topics, namely, Inflation, Forecasting, Monetary Policy, Financial Markets and Economic Activity, is easy and when hard. Sample Period: MPC meetings between July 2015 - July 2023.

Figure 4.12: Volatility Comparisons

Table 4.9: 10-Minute Regression Addendum

	Model 1	Model 2	Model 3	Model 4
I1	0.001 (0.019)	0.004 (0.013)	0.003 (0.012)	0.004 (0.012)
I2	0.004 (0.018)	0.005 (0.012)	0.003 (0.012)	0.004 (0.012)
I3	-0.005 (0.018)	0.001 (0.012)	-0.001 (0.011)	0.000 (0.011)
I4	0.002 (0.018)	0.005 (0.012)	0.003 (0.011)	0.004 (0.011)
I5	0.001 (0.017)	0.006 (0.012)	0.003 (0.011)	0.005 (0.011)
I6	-0.002 (0.013)	0.003 (0.010)	0.000 (0.009)	0.002 (0.009)
I8	-0.003 (0.013)	0.000 (0.009)	-0.001 (0.008)	-0.000 (0.008)
I9	-0.010 (0.017)	-0.005 (0.011)	-0.006 (0.010)	-0.006 (0.010)
I10	-0.008 (0.017)	-0.004 (0.012)	-0.005 (0.011)	-0.005 (0.011)
I11	0.002 (0.020)	0.002 (0.014)	0.003 (0.013)	0.000 (0.012)
I12	-0.001 (0.018)	-0.002 (0.011)	-0.004 (0.010)	-0.002 (0.010)
I13	0.002 (0.021)	-0.001 (0.013)	0.001 (0.011)	-0.003 (0.011)
I14	0.007 (0.023)	0.003 (0.015)	0.005 (0.013)	0.001 (0.013)
I15	-0.004 (0.022)	-0.011 (0.012)	-0.008 (0.011)	-0.012 (0.011)
I16	0.009 (0.026)	-0.004 (0.016)	0.002 (0.015)	-0.004 (0.014)
I17	-0.002 (0.022)	-0.006 (0.015)	-0.001 (0.014)	-0.007 (0.013)
I18	0.015 (0.025)	0.001 (0.016)	0.004 (0.015)	-0.002 (0.014)
Adjusted R ²	0.082	0.586	0.630	0.634
F-statistic	1.314	18.303	26.645	21.498
AIC	-1368.99	-1876.73	-1948.48	-1952.12

Notes: The table reports the coefficients and standard errors for the time dummies (I1–I18) from the regression models. The dependent variable is bond liquidity. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4.10: Results of 10-Minute Interaction Regression

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-0.923*** (0.078)	-0.779*** (0.084)	-0.844*** (0.079)	-0.788*** (0.077)	-0.782*** (0.078)	-0.925*** (0.088)
Liquidity Volatility	1.967*** (0.469)	1.946*** (0.481)	2.043*** (0.402)	2.002*** (0.486)	1.847*** (0.452)	1.770*** (0.477)
Price	0.489*** (0.034)	0.427*** (0.037)	0.454*** (0.035)	0.430*** (0.034)	0.428*** (0.034)	0.489*** (0.038)
Surprise	-0.1447*** (0.010)	-0.1550*** (0.011)	-0.1479*** (0.009)	-0.155*** (0.011)	-0.148*** (0.010)	-0.148*** (0.010)
I1	0.005 (0.013)	0.004 (0.013)	0.005 (0.013)	0.004 (0.013)	0.004 (0.013)	0.005 (0.013)
I2	0.005 (0.012)	0.005 (0.012)	0.005 (0.012)	0.005 (0.012)	0.005 (0.012)	0.005 (0.012)
I3	0.000 (0.012)	0.001 (0.012)	0.000 (0.012)	0.001 (0.012)	0.001 (0.012)	0.001 (0.012)
I4	0.004 (0.012)	0.005 (0.012)	0.004 (0.012)	0.005 (0.012)	0.004 (0.012)	0.004 (0.012)
I5	0.006 (0.012)	0.006 (0.012)	0.006 (0.012)	0.006 (0.012)	0.006 (0.012)	0.006 (0.012)
I6	0.003 (0.010)	0.003 (0.010)	0.002 (0.010)	0.002 (0.010)	0.003 (0.010)	0.003 (0.010)
I8	-0.260** (0.087)	-0.007 (0.073)	-0.093** (0.037)	-0.027 (0.070)	-0.074 (0.069)	-0.156** (0.071)
I9	-0.292** (0.102)	-0.058 (0.071)	-0.098** (0.040)	-0.081 (0.064)	-0.119 (0.066)	-0.169** (0.078)
I10	-0.270** (0.103)	-0.047 (0.075)	-0.106*** (0.036)	-0.062 (0.072)	-0.093 (0.078)	-0.155** (0.079)
I11	-0.174* (0.088)	0.107 (0.088)	-0.183** (0.079)	0.031 (0.077)	-0.058 (0.078)	-0.128 (0.088)

Continued on next page

Table continued from previous page

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	(0.096)	(0.100)	(0.077)	(0.102)	(0.094)	(0.083)
I12	-0.305***	-0.043	-0.061	-0.070	-0.096**	-0.213***
	(0.080)	(0.059)	(0.035)	(0.051)	(0.045)	(0.060)
I13	-0.305**	-0.039	-0.184***	-0.058	-0.158**	-0.195**
	(0.100)	(0.064)	(0.053)	(0.062)	(0.069)	(0.079)
I14	-0.355***	-0.055	-0.194**	-0.042	-0.117	-0.229**
	(0.108)	(0.082)	(0.084)	(0.091)	(0.089)	(0.096)
I15	-0.344***	-0.134**	-0.198***	-0.121**	-0.091	-0.179**
	(0.096)	(0.051)	(0.052)	(0.049)	(0.069)	(0.080)
I16	-0.281*	0.006	-0.243**	0.023	-0.137	-0.185*
	(0.119)	(0.109)	(0.089)	(0.097)	(0.101)	(0.110)
I17	-0.252*	0.006	-0.249**	0.024	-0.133	-0.177
	(0.114)	(0.114)	(0.097)	(0.094)	(0.125)	(0.110)
I18	-0.276**	-0.058	-0.263***	-0.036	-0.217**	-0.186**
	(0.111)	(0.096)	(0.076)	(0.120)	(0.092)	(0.091)
I8_inter	0.022**	0.001	0.009**	0.002	0.006	0.013**
	(0.007)	(0.006)	(0.004)	(0.005)	(0.005)	(0.006)
I9_inter	0.024**	0.004	0.009**	0.006	0.009*	0.014**
	(0.008)	(0.006)	(0.004)	(0.005)	(0.005)	(0.006)
I10_inter	0.022**	0.004	0.010**	0.005	0.007	0.012*
	(0.008)	(0.006)	(0.003)	(0.005)	(0.006)	(0.006)
I11_inter	0.015*	-0.009	0.018**	-0.002	0.004	0.011
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
I12_inter	0.025***	0.003	0.006*	0.005	0.007**	0.017***
	(0.007)	(0.005)	(0.003)	(0.004)	(0.004)	(0.005)
I13_inter	0.025**	0.003	0.017***	0.005	0.012*	0.016**
	(0.008)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
I14_inter	0.030**	0.005	0.019**	0.004	0.009	0.019**
	(0.009)	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)

Continued on next page

Table continued from previous page

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
I15_inter	0.028*** (0.008)	0.011** (0.004)	0.018*** (0.005)	0.009** (0.004)	0.006 (0.006)	0.014** (0.007)
I16_inter	0.024* (0.011)	-0.001 (0.009)	0.023** (0.009)	-0.002 (0.007)	0.010 (0.008)	0.015 (0.010)
I17_inter	0.021* (0.010)	-0.001 (0.009)	0.023** (0.009)	-0.002 (0.007)	0.010 (0.010)	0.014 (0.010)
I18_inter	0.023** (0.010)	0.005 (0.008)	0.024*** (0.007)	0.003 (0.009)	0.016** (0.007)	0.015* (0.008)
Adjusted R ²	0.616	0.583	0.637	0.584	0.595	0.601
F-statistic	14.78	13.14	17.31	13.83	13.40	12.51
AIC	-1914.46	-1862.17	-1951.23	-1863.62	-1880.52	-1890.35

Notes: This table presents full regression results across six models analysing bond liquidity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.11: 15-Minute Regression Addendum

	Model 1	Model 2	Model 3	Model 4
I1	0.005 (0.018)	0.006 (0.012)	0.006 (0.011)	0.006 (0.011)
I2	-0.002 (0.018)	0.003 (0.012)	0.002 (0.011)	0.002 (0.011)
I3	0.004 (0.018)	0.008 (0.012)	0.007 (0.011)	0.007 (0.011)
I4	-0.000 (0.013)	0.006 (0.010)	0.004 (0.009)	0.005 (0.009)
I6	-0.007 (0.012)	-0.003 (0.008)	-0.004 (0.008)	-0.004 (0.007)
I7	-0.004 (0.016)	-0.000 (0.011)	-0.001 (0.011)	-0.001 (0.010)
I8	0.003 (0.018)	0.001 (0.011)	-0.000 (0.010)	0.000 (0.010)
I9	0.005 (0.021)	0.001 (0.012)	0.004 (0.011)	-0.001 (0.011)
I10	0.011 (0.024)	0.005 (0.015)	0.008 (0.014)	0.003 (0.013)
I11	0.015 (0.024)	0.004 (0.014)	0.008 (0.013)	0.001 (0.012)
I12	0.015 (0.024)	0.001 (0.014)	0.006 (0.013)	-0.001 (0.012)
Adjusted R ²	0.102	0.605	0.649	0.653
F-statistic	1.319	19.366	27.042	20.472
AIC	-976.21	-1344.90	-1397.07	-1399.05

Notes: The table reports the coefficients and standard errors for the time dummies (I1–I12) from the regression models. The dependent variable is bond liquidity. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.12: Results of 15-Minute Interaction Regression

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-0.915*** (0.091)	-0.787*** (0.099)	-0.837*** (0.093)	-0.778*** (0.092)	-0.774*** (0.091)	-0.914*** (0.102)
Liquidity Volatility	2.221*** (0.604)	2.195*** (0.621)	2.235*** (0.508)	2.280*** (0.624)	2.102*** (0.590)	2.043*** (0.607)
Price	0.483*** (0.040)	0.428*** (0.043)	0.449*** (0.041)	0.424*** (0.040)	0.423*** (0.040)	0.483*** (0.045)
Surprise	-0.143*** (0.011)	-0.153*** (0.012)	-0.146*** (0.010)	-0.153*** (0.012)	-0.147*** (0.011)	-0.146*** (0.012)
I1	0.007 (0.012)	0.007 (0.012)	0.007 (0.012)	0.007 (0.012)	0.006 (0.012)	0.007 (0.012)
I2	0.003 (0.012)	0.003 (0.012)	0.003 (0.012)	0.003 (0.012)	0.003 (0.012)	0.003 (0.012)
I3	0.007 (0.012)	0.008 (0.012)	0.008 (0.012)	0.008 (0.012)	0.008 (0.012)	0.007 (0.012)
I4	0.006 (0.010)	0.006 (0.010)	0.006 (0.010)	0.006 (0.010)	0.006 (0.010)	0.006 (0.010)
I6	-0.283*** (0.096)	-0.052 (0.064)	-0.094** (0.033)	-0.079 (0.059)	-0.102 (0.063)	-0.160** (0.073)
I7	-0.254** (0.102)	-0.031 (0.074)	-0.101** (0.037)	-0.051 (0.070)	-0.073 (0.072)	-0.148* (0.078)
I8	-0.306*** (0.079)	-0.039 (0.060)	-0.059 (0.036)	-0.064 (0.052)	-0.096** (0.046)	-0.215*** (0.060)
I9	-0.337*** (0.089)	-0.058 (0.061)	-0.204*** (0.046)	-0.057 (0.063)	-0.136* (0.071)	-0.212** (0.076)
I10	-0.359** (0.112)	-0.059 (0.087)	-0.199** (0.088)	-0.039 (0.095)	-0.111 (0.094)	-0.231** (0.102)
I11	-0.295**	-0.015	-0.248**	0.028	-0.141	-0.178

Continued on next page

Table continued from previous page

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	(0.107)	(0.097)	(0.080)	(0.093)	(0.097)	(0.104)
I12	-0.295**	-0.015	-0.248**	0.028	-0.141	-0.178
	(0.107)	(0.097)	(0.080)	(0.093)	(0.097)	(0.104)
I12	-0.189*	-0.013	-0.227**	0.035	-0.119	-0.118
	(0.089)	(0.082)	(0.085)	(0.092)	(0.082)	(0.075)
I6_inter	0.023**	0.004	0.009**	0.006	0.008*	0.013**
	(0.008)	(0.005)	(0.003)	(0.005)	(0.005)	(0.006)
I7_inter	0.021*	0.003	0.010**	0.004	0.006	0.012*
	(0.008)	(0.006)	(0.004)	(0.005)	(0.005)	(0.006)
I8_inter	0.025***	0.003	0.006*	0.005	0.007**	0.018***
	(0.007)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
I9_inter	0.028***	0.005	0.019***	0.005	0.010*	0.018**
	(0.008)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
I10_inter	0.031**	0.005	0.019**	0.004	0.009	0.019**
	(0.010)	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)
I11_inter	0.025**	0.002	0.023**	-0.002	0.010	0.015*
	(0.009)	(0.007)	(0.009)	(0.007)	(0.008)	(0.009)
I12_inter	0.016*	0.001	0.021**	-0.003	0.009	0.010
	(0.008)	(0.006)	(0.008)	(0.007)	(0.006)	(0.007)
Adjusted R ²	0.636	0.600	0.651	0.603	0.613	0.621
F-statistic	16.18	14.24	18.89	14.55	14.02	13.59
AIC	-1375.22	-1333.21	-1394.92	-1336.66	-1348.16	-1357.53

Notes: This table presents full regression results across six models analysing bond liquidity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Chapter 5

Concluding Remarks

Summary of Contributions

This thesis expands upon existing economic literature through multiple distinct contributions that span three main chapters. All chapters are integrated perspectives on macroeconomic dynamics, dealing essentially with data measurements and improvements, cross-sectional differences, and policy information flow. Chapter 3 builds on Chapter 2's insights through the application of a temporal disaggregation method introduced earlier to new data. Chapter 4 deviates in its exact interest but maintains a connection to the thesis's central theme which focuses on generating data and performing economic analyses at a granular level instead of the traditional aggregate level. Utilising a granular approach becomes essential because it enables us to gain comprehensive insights into economic relationships which, in turn, supports the development of precise and effective policy measures. Precise economic data analysis at this nuanced level remains essential to develop policy responses that ensure both accuracy and suitability to specific economic contexts. In what follows, I will outline the main discoveries, contributions and respective policy relevance presented in each chapter along with the innovative approaches and study outcomes.

The first study focuses primarily on developing and demonstrating the performance gains from a new data-driven method for temporal disaggregation of macroeconomic time series data. Temporal disaggregation transforms low-frequency time series, such as annual statistics, into high-frequency time series such as quarterly figures. The study develops the new approach using Partial Least Squares (PLS) regression which although uncommon in this area proves highly effective for accurate and reliable temporal disaggregation because it does well in uncovering latent factors driving the target variable. Technically, the application of PLS regression strengthens the connection between predictors and the

target variable, resulting in significant improvements to the predictive accuracy of the disaggregation process. The study's statistical analysis demonstrates that this method offers a reliable way to transform low-frequency time series data into high-frequency data while maintaining strong accuracy and consistency.

Economic analysts appreciate this process because high-frequency data on broad indicators necessary for short-term analysis and forecasting can come to existence after converting low-frequency economic indicators through such methods. This process helps economists to analyse economic trends more efficiently and promptly. The European Statistical System has advocated for the PLS approach as a potential advancement in temporal disaggregation while this thesis endorses its utilisation in real-world economic analyses. In this study's empirical research, I analyse macroeconomic data from the U.S. where true high-frequency series exist, obtained from the Federal Reserve Economic Data (FRED) database for the years between 1960 and 2020. Country-specific macroeconomic data supplements this dataset to enable targeted temporal disaggregation and nowcasting experiments. The empirical studies demonstrate how the PLS approach successfully creates high-frequency data that maintains both accuracy and informative quality.

The second study of this thesis first, resolves an essential data gap in U.S. economic analysis by estimating a historical quarterly GDP growth time series dataset for each state. Prior to 2005 there were no official quarterly GDP statistics available for U.S. states which severely limited regional economic analysis capabilities as seen from the literature. The lack of accessible data has hindered researchers from studying how national economic events affect different states individually. Therefore, to address the data challenge, the study applies the PLS temporal disaggregation method, which allows me to create state-specific quarterly output growth time series from available annual GDP growth data and state-specific quarterly economic indicators. The newly generated data set is used to perform multiple essential economic analyses. First, the study examines how an aggregate financial shock influences state-level economic activities and how state economic activities become interconnected and form economic clusters. This research provides critical insights because the previous literature commonly used national-level aggregates or high-frequency proxies such state specific employment data to investigate state-level economic activity. The study's Vector Autoregression (VAR) impulse response analysis demonstrates that

different states output growth respond differently to the same aggregate financial shock. The economic output of some states drops significantly due to a financial shock while others demonstrate greater resilience, which showcases the varying regional capacities to handle economic disruptions. The study's analysis of economic interconnectedness demonstrates that U.S. states have become more linked together with three major clusters and two additional smaller clusters now forming. While previous studies relied on employment data to estimate state economic activity, this study utilises high-frequency output data for a more precise depiction of state-level economic activity. The results have a policy implication that underscores the limitations of a one-size-fits-all monetary or fiscal policy, since state heterogeneity leads to uneven outcomes, which policymakers should account for.

The third study in this thesis sheds light on a significant research void with respect to the importance of topic-specific linguistic complexities in central bank communications. Existing studies in this field generally measure central bank communication's overall linguistic complexity without distinguishing the complexity levels of different topics within these communication documents. The common approach in the literature presents limitations because it operates under the assumption that each topic discussed in central bank communications has equal complexity and influence on the financial markets. The main contribution of this research offers a novel approach to measuring and considering the linguistic complexity of monetary policy communication based on specific topics. Rather than replacing the existing approach, this approach introduces a dimension that deepens understanding through its detailed analysis of central bank communications. Creating topic-specific complexity time series enables me to investigate the relationship between linguistic complexity in policy communications and market liquidity specifically concerning 10-year government bond movements on announcement days. The event study regression framework shows that high linguistic complexity in specific topics leads to reduced liquidity in the bond market. The increase in market uncertainty as measured by liquidity volatility correlates with the reduced trading activity manifested in larger quoted spreads. The findings show that participants in financial markets respond to how clear both the general central bank messages and the specific topics they address are. This research finding carries important consequences since it demonstrates that financial market stability

can be enhanced by central banks when they simplify the language of topics that affect market participants. My approach provides central banks with the ability to customise their communications which helps them achieve better clarity and less market uncertainty, thereby resulting in more stable economic results.

Future Research

Related to Chapter 2: Analysts and policymakers in economics face one of their biggest challenges when dealing with missing data problems. Missing data creates obstacles for precise economic analysis while reducing the impact of policy measures. Temporal disaggregation offers significant opportunities for improving prediction accuracy in missing time series data analysis. Traditional statistical models have laid the groundwork for tackling temporal disaggregation challenges yet they require further development for enhanced performance. Upcoming studies could investigate the implementation of advanced machine learning techniques to improve the management of missing data. Although machine learning approaches such as LASSO (Least Absolute Shrinkage and Selection Operator) and PLS regression demonstrate potential for temporal disaggregation they could achieve better accuracy and efficiency through broader investigation of other machine learning models to process high-dimensional data. Machine learning approaches could broaden temporal disaggregation applications to incorporate detailed and unstructured datasets when used alongside traditional statistical methods. The creation of these new approaches will enable economists to perform more thorough analyses which will support policymakers to make more informed decisions and adapt to economic changes better.

Related to Chapter 3: Based on the evidence of connectedness, future research should examine if states which are highly susceptible to an aggregate financial shock display any common characteristics. The value of studying the economic resilience of states lies in understanding if those with strong dependence on volatile industries and minimal economic diversity or those burdened by high debt levels experience greater negative impacts from financial disruptions through the financial accelerator. Fiscal imbalances together with demographic challenges like an aging population dependence on federal funds and natural disaster vulnerability may represent possible factors. Future research

that recognises these characteristics will enable policymakers to implement more precise interventions by focusing on states that face the highest risk of negative economic impacts after financial disruptions. Researching state-level policies during financial disruptions can help identify which regional measures successfully mitigate economic shocks.

Related to Chapter 4: Future research will benefit from exploring country-specific linguistic complexity measurement strategies in respective countries. The consequences of educational and linguistic variations between nations on readability measurement remain an unexplored research topic. The standard Flesch-Kincaid grade score used to determine text readability derives from the U.S. educational system which limits its effectiveness in other countries that have distinct educational systems. Upcoming studies could focus on creating Flesch-Kincaid-like models that match the unique educational and linguistic situations of each country. Developing these localised models will enable researchers to assess linguistic complexity more precisely in non U.S. countries. Readability measures would achieve broader global applicability which would serve policymakers and central banks operating in various economic and educational settings. In addition, future research could focus on building models that link how the structure and content of central bank language evolves (via the network model) with how markets actually move at high frequency. This would help to map the dynamic transmission of narrative-driven shocks across financial asset classes, for instance, how an abrupt focus on “inflation” in a speech transmits to financial instruments. This has the potential to increase both the analytical tool kit and the policy relevance of such studies on central bank communication.

Bibliography

- Acemoglu, Daron et al. (2012). ‘The network origins of aggregate fluctuations’. In: *Econometrica* 80.5, pp. 1977–2016.
- Ahmed, Abdullahi D (2016). ‘Integration of financial markets, financial development and growth: Is Africa different?’ In: *Journal of International Financial Markets, Institutions and Money* 42, pp. 43–59.
- Ahn, Seung C and Juhee Bae (2022). ‘Forecasting with partial least squares when a large number of predictors are available’. In: *Available at SSRN 4248450*.
- Ajao, Isaac O, Femi J Ayoola and Joseph O Iyaniwura (2016). ‘Temporal disaggregation methods in flow variables of economic data: Comparison study’. In: *Editorial Board* 5.1, p. 36.
- Al-Osh, M (1989). ‘A dynamic linear model approach for disaggregating time series data’. In: *Journal of Forecasting* 8.2, pp. 85–96.
- Anderson, Magnus (2010). ‘Using Intraday Data to Gauge Financial Market Response to Federal Reserve and ECB Monetary Policy Decision’. In: *International Journal of Central Banking* 6.2, pp. 107–116.
- Andersson, Malin, Hans Dillén and Peter Sellin (2006). ‘Monetary policy signaling and movements in the term structure of interest rates’. In: *Journal of Monetary Economics* 53.8, pp. 1815–1855.
- Ando, Mitsuyo and Fukunari Kimura (2012). ‘How did the Japanese Exports Respond to Two Crises in the International Production Networks? The Global Financial Crisis and the Great East Japan Earthquake’. In: *Asian Economic Journal* 26.3, pp. 261–287.
- Apergis, Nicholas (2015). ‘Forecasting Credit Default Swaps (CDSs) spreads with news-wire messages: Evidence from European countries under financial distress’. In: *Economics Letters* 136, pp. 92–94.
- Armesto, Michelle T, Kristie M Engemann, Michael T Owyang et al. (2010). ‘Forecasting with mixed frequencies’. In: *Federal Reserve Bank of St. Louis Review* 92.6, pp. 521–36.

- Bai, Jushan (2003). ‘Inferential theory for factor models of large dimensions’. In: *Econometrica* 71.1, pp. 135–171.
- Bai, Jushan and Serena Ng (2008). ‘Forecasting economic time series using targeted predictors’. In: *Journal of Econometrics* 146.2, pp. 304–317.
- Barbaglia, Luca, Sergio Consoli and Sebastiano Manzan (2023). ‘Forecasting with economic news’. In: *Journal of Business & Economic Statistics* 41.3, pp. 708–719.
- (2024). ‘Forecasting GDP in Europe with textual data’. In: *Journal of Applied Econometrics* 39.2, pp. 338–355.
- Bean, Charles R (2016). *Independent review of UK economic statistics*. HM Treasury.
- Belloni, Alexandre, Victor Chernozhukov and Lie Wang (2011). ‘Square-root lasso: pivotal recovery of sparse signals via conic programming’. In: *Biometrika* 98.4, pp. 791–806.
- Berg, Andrew et al. (2015). ‘Monetary policy issues in sub-Saharan Africa’. In: *The Oxford Handbook of Africa and Economics* 2, pp. 62–87.
- Bernanke, Ben S and Kenneth N Kuttner (2005). ‘What explains the stock market’s reaction to Federal Reserve policy?’ In: *The Journal of Finance* 60.3, pp. 1221–1257.
- Biddle, Gary C, Gilles Hilary and Rodrigo S Verdi (2009). ‘How does financial reporting quality relate to investment efficiency?’ In: *Journal of Accounting and Economics* 48.2-3, pp. 112–131.
- Binder, Carola (2017). ‘Fed speak on main street: Central bank communication and household expectations’. In: *Journal of Macroeconomics* 52, pp. 238–251.
- Blei, David M, Andrew Y Ng and Michael I Jordan (2003). ‘Latent dirichlet allocation’. In: *Journal of Machine Learning Research* 3.Jan, pp. 993–1022.
- Blinder, Alan S et al. (2008). ‘Central bank communication and monetary policy: A survey of theory and evidence’. In: *Journal of economic literature* 46.4, pp. 910–945.
- Blinder, Alan S et al. (2024). ‘Central bank communication with the general public: Promise or false hope?’ In: *Journal of Economic Literature* 62.2, pp. 425–457.
- Bloem, Mr Adriaan M, Mr Robert Dippelsman and Mr Nils Øyvind Mæhle (2001). *Quarterly national accounts manual: concepts, data sources, and compilation*. International Monetary Fund.
- Blommestein, Hans J, Sylvester CW Eijffinger and Zongxin Qian (2012). ‘Animal spirits in the Euro area sovereign CDS market’. In.

- Blondel, Vincent D et al. (2008). ‘Fast unfolding of communities in large networks’. In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.10, P10008.
- Boivin, Jean and Serena Ng (2006). ‘Are more data always better for factor analysis?’ In: *Journal of Econometrics* 132.1, pp. 169–194.
- Boot, Johannes Cornelius Gerardus, Walter Feibes and Johannes Hubertus Cornelius Lisman (1967). ‘Further methods of derivation of quarterly figures from annual data’. In: *Applied Statistics*, pp. 65–75.
- Brownlees, Christian, Guðmundur Stefán Guðmundsson and Gábor Lugosi (2022). ‘Community detection in partial correlation network models’. In: *Journal of Business & Economic Statistics* 40.1, pp. 216–226.
- Bulíř, Aleš, Martin Čihák and David-Jan Jansen (2013). ‘What drives clarity of central bank communication about inflation?’ In: *Open Economies Review* 24, pp. 125–145.
- Buono, Dario et al. (2018). *ESS guidelines on temporal disaggregation, benchmarking and reconciliation*. Publications Office of the European Union.
- Cabred, Bernardí and Jose Manuel Pavía (1999). ‘Estimating $J (> 1)$ quarterly time series in fulfilling annual and quarterly constraints’. In: *International Advances in Economic Research* 5.3, pp. 339–349.
- Campbell, John Y, Carolin Pflueger and Luis M Viceira (2020). ‘Macroeconomic drivers of bond and equity risks’. In: *Journal of Political Economy* 128.8, pp. 3148–3185.
- Carlino, Gerald and Robert DeFina (1998). ‘The differential regional effects of monetary policy’. In: *Review of economics and statistics* 80.4, pp. 572–587.
- Casiraghi, Marco and Leonardo Pio Perez (2022). ‘Central Bank Communications’. In.
- Chamberlin, Graeme (2010). ‘Methods Explained: Temporal Disaggregation’. In: *Economic & Labour Market Review* 4.11, pp. 106–121.
- Chen, Baoline et al. (2007). *An empirical comparison of methods for temporal distribution and interpolation at the national accounts*. BEA.
- Cholette, Pierre A and Estela Bee Dagum (1994). ‘Benchmarking time series with autocorrelated survey errors’. In: *International Statistical Review/Revue Internationale de Statistique*, pp. 365–377.

- Chow, Gregory C and An-loh Lin (1971). ‘Best linear unbiased interpolation, distribution, and extrapolation of time series by related series’. In: *The review of Economics and Statistics*, pp. 372–375.
- Chung, Kee H, John Elder and Jang-chul Kim (2013). ‘Liquidity and information flow around monetary policy announcement’. In: *Journal of Money, Credit and Banking* 45.5, pp. 781–820.
- Cieslak, Anna, Adair Morse and Annette Vissing-Jorgensen (2019). ‘Stock returns over the FOMC cycle’. In: *The Journal of Finance* 74.5, pp. 2201–2248.
- Cieslak, Anna and Andreas Schrimpf (2019). ‘Non-monetary news in central bank communication’. In: *Journal of International Economics* 118, pp. 293–315.
- Clemen, Robert T and Robert L Winkler (1986). ‘Combining economic forecasts’. In: *Journal of Business & Economic Statistics* 4.1, pp. 39–46.
- Clements, AE, AS Hurn and VV Volkov (2015). ‘Volatility transmission in global financial markets’. In: *Journal of Empirical Finance* 32, pp. 3–18.
- Clements, Michael P and Ana Beatriz Galvão (2008). ‘Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States’. In: *Journal of Business & Economic Statistics* 26.4, pp. 546–554.
- (2009). ‘Forecasting US output growth using leading indicators: An appraisal using MIDAS models’. In: *Journal of Applied Econometrics* 24.7, pp. 1187–1206.
- Coco, Alberto, Nicola Viegi et al. (2020). *The monetary policy of the South African Reserve Bank: stance, communication and credibility*. Economic Research and Statistics Department, South African Reserve Bank.
- Cohen, Kalman J, Wolfgang Müller and Manfred W Padberg (1971). ‘Autoregressive approaches to disaggregation of time series data’. In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 20.2, pp. 119–129.
- Coleman, Meri and Ta Lin Liao (1975). ‘A computer readability formula designed for machine scoring’. In: *Journal of Applied Psychology* 60.2, p. 283.
- Consoli, Sergio, Luca Tiozzo Pezzoli and Elisa Tosetti (2021). ‘Emotions in macroeconomic news and their impact on the european bond market’. In: *Journal of International Money and Finance* 118, p. 102472.

- Cuartas, Blanca Moreno, Esteban Fernández Vázquez and Geoffrey JD Hewings (2020). ‘Regional temporal disaggregation on economic series with macroeconomic balance: An entropy econometrics-based model’. In: *Innovations in Urban and Regional Systems: Contributions from GIS&T, Spatial Analysis and Location Modeling*, pp. 243–256.
- Da, Zhi, Joseph Engelberg and Pengjie Gao (2015). ‘The sum of all FEARS investor sentiment and asset prices’. In: *The Review of Financial Studies* 28.1, pp. 1–32.
- Daly, Kevin (2008). ‘Financial volatility: Issues and measuring techniques’. In: *Physica A: statistical mechanics and its applications* 387.11, pp. 2377–2393.
- De Franco, Gus et al. (2015). ‘Analyst report readability’. In: *Contemporary Accounting Research* 32.1, pp. 76–104.
- De Lombaerde, Philippe et al. (2018). ‘Using network analysis to study globalization, regionalization, and multi-polarity—Introduction to special section’. In: *Network Science* 6.4, pp. 494–516.
- De Mol, Christine, Domenico Giannone and Lucrezia Reichlin (2008). ‘Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?’ In: *Journal of Econometrics* 146.2, pp. 318–328.
- Della Valle, Guido, Darryl King and Romain Veyrune (2022). ‘Monetary and Capital Markets Department: Technical Assistance Handbook-Reserve Requirements’. In.
- Demiralp, Selva and Oscar Jorda (2004). ‘The response of term rates to Fed announcements’. In: *Journal of Money, Credit and Banking*, pp. 387–405.
- Denton, Frank T (1971). ‘Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization’. In: *Journal of the american statistical association* 66.333, pp. 99–102.
- Di Fonzo, Tommaso (2003). ‘Temporal disaggregation of economic time series: towards a dynamic extension’. In.
- Diebold, Francis X and Jose A Lopez (1996). ‘8 Forecast evaluation and combination’. In: *Handbook of statistics* 14, pp. 241–268.
- Diebold, Francis X and Peter Pauly (1990). ‘The use of prior information in forecast combination’. In: *International Journal of Forecasting* 6.4, pp. 503–508.

- Diebold, Francis X and Kamil Yilmaz (2014). ‘On the network topology of variance decompositions: Measuring the connectedness of financial firms’. In: *Journal of econometrics* 182.1, pp. 119–134.
- Durbin, James and Benoit Quenneville (1997). ‘Benchmarking by state space models’. In: *International Statistical Review* 65.1, pp. 23–48.
- ECB (2002). ‘Transparency in the monetary policy of the ECB’. In: *ECB Monthly Bulletin* November.
- Efron, Bradley et al. (2004). ‘Least angle regression’. In: *The Annals of statistics* 32.2, pp. 407–499.
- Ehrmann, Michael and Marcel Fratzscher (2005). ‘How should central banks communicate?’ In.
- (2007). ‘Communication by central bank committee members: different strategies, same effectiveness?’ In: *Journal of Money, Credit and Banking* 39.2-3, pp. 509–541.
- Ehrmann, Michael and Jonathan Talmi (2020). ‘Starting from a blank page? Semantic similarity in central bank communication and market volatility’. In: *Journal of Monetary Economics* 111, pp. 48–62.
- Ehrmann, Michael and Alena Wabitsch (2022). ‘Central bank communication with non-experts—A road to nowhere?’ In: *Journal of Monetary Economics* 127, pp. 69–85.
- Ekman, Paul (1973). ‘Universal facial expressions in emotion’. In: *Studia Psychologica* 15.2, pp. 140–147.
- Eurostat (2014). *Memobust handbook - Handbook on Methodology of Modern Business Statistics*.
- Eurostat, Unit F and Mr Hartmut Buchow (2013). ‘Information society statistics’. In.
- Evans, Kevin and Alan Speight (2010). ‘International macroeconomic announcements and intraday euro exchange rate volatility’. In: *Journal of the Japanese and international economies* 24.4, pp. 552–568.
- Falk, R Frank and Nancy B Miller (1992). *A primer for soft modeling*. University of Akron Press.
- Fernandez, Roque B (1981). ‘A methodological note on the estimation of time series’. In: *The Review of Economics and Statistics* 63.3, pp. 471–476.

- Ferrara, L and GL Mazzi (2017). ‘Parametric models for cyclical turning point indicators’. In: *Handbook on Cyclical Composite indicators, Eurostat and United Nations, editors*, pp. 327–356.
- Flesch, Rudolph (1948). ‘A new readability yardstick.’ In: *Journal of applied psychology* 32.3, p. 221.
- Garcia, Diego (2013). ‘Sentiment during recessions’. In: *The journal of finance* 68.3, pp. 1267–1300.
- Geladi, Paul and Bruce R Kowalski (1986). ‘Partial least-squares regression: a tutorial’. In: *Analytica chimica acta* 185, pp. 1–17.
- Geraats, Petra M (2009). ‘Trends in monetary policy transparency’. In: *International Finance* 12.2, pp. 235–268.
- Gertler, Mark and Simon Gilchrist (1993). ‘The role of credit market imperfections in the monetary transmission mechanism: arguments and evidence’. In: *The Scandinavian Journal of Economics*, pp. 43–64.
- Gilchrist, Simon and Egon Zakrajšek (2012a). ‘Credit spreads and business cycle fluctuations’. In: *American economic review* 102.4, pp. 1692–1720.
- (2012b). ‘Credit supply shocks and economic activity in a financial accelerator model’. In: *Rethinking the Financial Crisis* 1, pp. 37–72.
- Goldin, Ian and Tiffany Vogel (2010). ‘Global governance and systemic risk in the 21st century: Lessons from the financial crisis’. In: *Global Policy* 1.1, pp. 4–15.
- Gómez, Víctor and Agustín Maravall (1994). ‘Estimation, prediction, and interpolation for nonstationary series with the Kalman filter’. In: *Journal of the American Statistical Association* 89.426, pp. 611–624.
- Goodfriend, Marvin and Robert G King (2005). ‘The incredible Volcker disinflation’. In: *Journal of Monetary Economics* 52.5, pp. 981–1015.
- Groen, Jan JJ and George Kapetanios (2016). ‘Revisiting useful approaches to data-rich macroeconomic forecasting’. In: *Computational Statistics & Data Analysis* 100, pp. 221–239.
- Guay, Wayne, Delphine Samuels and Daniel Taylor (2016). ‘Guiding through the fog: Financial statement complexity and voluntary disclosure’. In: *Journal of Accounting and Economics* 62.2-3, pp. 234–269.

- Guerrero, Victor M and J Martínez (1995). ‘A recursive ARIMA-based procedure for disaggregating a time series variable using concurrent data’. In: *Test* 4, pp. 359–376.
- Gunning, Robert (1952). *The technique of clear writing*.
- Gürkaynak, Refet S, Brian Sack and Eric T Swanson (2005). ‘Do actions speak louder than words? The response of asset prices to monetary policy actions and statements’. In: *International Journal of Central Banking* 1.1, pp. 55–93.
- Gürkaynak, Refet S, Brian P Sack and Eric T Swanson (2004). ‘Do actions speak louder than words? The response of asset prices to monetary policy actions and statements’. In: *The Response of Asset Prices to Monetary Policy Actions and Statements (November 2004)*.
- Hall, Viv B and C John McDermott (2011). ‘A quarterly post-Second World War real GDP series for New Zealand’. In: *New Zealand Economic Papers* 45.3, pp. 273–298.
- Hamilton, James D and Michael T Owyang (2012). ‘The propagation of regional recessions’. In: *Review of Economics and Statistics* 94.4, pp. 935–947.
- Hansen, Stephen and Michael McMahon (2016). ‘Shocking language: understanding the macroeconomic effects of central bank communication’. In: *Journal of International Economics* 99, S114–S133.
- Hansen, Stephen, Michael McMahon and Andrea Prat (2018). ‘Transparency and deliberation within the FOMC: A computational linguistics approach’. In: *The Quarterly Journal of Economics* 133.2, pp. 801–870.
- Hanson, Samuel G and Jeremy C Stein (2015). ‘Monetary policy and long-term real rates’. In: *Journal of Financial Economics* 115.3, pp. 429–448.
- Harvey, Andrew and Chia-Hui Chung (2000). ‘Estimating the underlying change in unemployment in the UK’. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 163.3, pp. 303–309.
- Harvey, Andrew C and Richard G Pierse (1984). ‘Estimating missing observations in economic time series’. In: *Journal of the American statistical Association* 79.385, pp. 125–131.
- Hayo, Bernd and Matthias Neuenkirch (2010). ‘Do Federal Reserve communications help predict federal funds target rate decisions?’ In: *Journal of Macroeconomics* 32.4, pp. 1014–1024.

- Hayo, Bernd et al. (2022). ‘Complexity of ECB communication and financial market trading’. In: *Journal of International Money and Finance* 128, p. 102709.
- Helland, Inge S (1990). ‘Partial least squares regression and statistical models’. In: *Scandinavian journal of statistics*, pp. 97–114.
- Holland, Paul W, Kathryn Blackmond Laskey and Samuel Leinhardt (1983). ‘Stochastic blockmodels: First steps’. In: *Social networks* 5.2, pp. 109–137.
- Houben, Aerd CFJ and Aerd CFJ Houben (2000). ‘Key Considerations Governing the Choice of Monetary Policy Strategy’. In: *The Evolution of Monetary Policy Strategies in Europe*, pp. 11–72.
- Huang, Huiyu and Tae-Hwy Lee (2010). ‘To combine forecasts or to combine information?’ In: *Econometric Reviews* 29.5-6, pp. 534–570.
- Islam, Mohammad Rafiqul (2009). ‘Evaluation of different temporal disaggregation techniques and an application to Italian GDP’. In.
- (2013). ‘Comparison of temporal disaggregation techniques in the case of export of Bangladesh’. In: *International Journal of Applied Research and Studies* 2.11, pp. 1–9.
- Jacobson, Paul T et al. (2012). *Assessment and mapping of the riverine hydrokinetic resource in the continental United States*. Tech. rep. Electric Power Research Institute.
- Jansen, David-Jan (2011). ‘Does the clarity of central bank communication affect volatility in financial markets? Evidence from humphrey-hawkins testimonies’. In: *Contemporary Economic Policy* 29.4, pp. 494–509.
- Jarmin, Ron S (2019). ‘Evolving measurement for an evolving economy: thoughts on 21st century US economic statistics’. In: *Journal of Economic Perspectives* 33.1, pp. 165–84.
- Jegadeesh, Narasimhan and Di Wu (2017). ‘Deciphering Fedspeak: The information content of FOMC meetings’. In: *Monetary Economics: Central Banks–Policies & Impacts eJournal*.
- Kashyap, Anil K and Jeremy C Stein (1995). ‘The impact of monetary policy on bank balance sheets’. In: *Carnegie-rochester conference series on public policy*. Vol. 42. Elsevier, pp. 151–195.
- Kendall, Maurice George (1948). ‘Rank correlation methods.’ In.

- Khatri, Sunil Kumar, Himanshu Singhal and Prashant Johri (2014). ‘Sentiment analysis to predict Bombay stock exchange using artificial neural network’. In: *Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization*. IEEE, pp. 1–5.
- Kim, Sangmoon and Eui-Hang Shin (2002). ‘A longitudinal analysis of globalization and regionalization in international trade: A social network approach’. In: *Social forces* 81.2, pp. 445–468.
- Kincaid, JP (1975). ‘Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel’. In: *Chief of Naval Technical Training*.
- Klemenjak, Christoph, Stephen Makonin and Wilfried Elmenreich (2021). ‘Investigating the performance gap between testing on real and denoised aggregates in non-intrusive load monitoring’. In: *Energy Informatics* 4, pp. 1–15.
- Kohn, Donald L and Brian P Sack (2003). ‘Central bank talk: does it matter and why?’ In.
- Kolter, J Zico and Tommi Jaakkola (2012). ‘Approximate inference in additive factorial hmms with application to energy disaggregation’. In: *Artificial intelligence and statistics*. PMLR, pp. 1472–1482.
- Koop, Gary et al. (2020). ‘Regional output growth in the United Kingdom: More timely and higher frequency estimates from 1970’. In: *Journal of Applied Econometrics* 35.2, pp. 176–197.
- Korniyenko, Ms Yevgeniya, Magali Pinat and Brian Dew (2017). *Assessing the fragility of global trade: The impact of localized supply shocks using network analysis*. International Monetary Fund.
- Korobilis, Dimitris (2021). ‘High-dimensional macroeconomic forecasting using message passing algorithms’. In: *Journal of Business & Economic Statistics* 39.2, pp. 493–504.
- Kuttner, Kenneth N (2001). ‘Monetary policy surprises and interest rates: evidence from the Fed funds futures market’. In: *Journal of Monetary Economics* 47.3, pp. 523–544.
- Kuzin, Vladimir, Massimiliano Marcellino and Christian Schumacher (2011). ‘MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area’. In: *International Journal of Forecasting* 27.2, pp. 529–542.

- Lawrence, Alastair (2013). ‘Individual investors and financial disclosure’. In: *Journal of Accounting and Economics* 56.1, pp. 130–147.
- Lehavy, Reuven, Feng Li and Kenneth Merkley (2011). ‘The effect of annual report readability on analyst following and the properties of their earnings forecasts’. In: *The Accounting Review* 86.3, pp. 1087–1115.
- Leiva-Leon, Danilo (2017). ‘Measuring business cycles intra-synchronization in US: A regime-switching interdependence framework’. In: *Oxford Bulletin of Economics and Statistics* 79.4, pp. 513–545.
- Leombroni, Matteo et al. (2021). ‘Central bank communication and the yield curve’. In: *Journal of Financial Economics* 141.3, pp. 860–880.
- Li, Bin et al. (2020). ‘Estimation of regional economic development indicator from transportation network analytics’. In: *Scientific reports* 10.1, p. 2647.
- Li, Feng (2008). ‘Annual report readability, current earnings, and earnings persistence’. In: *Journal of Accounting and Economics* 45.2-3, pp. 221–247.
- Li, Jiahao and Weiye Chen (2014). ‘Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models’. In: *International Journal of Forecasting* 30.4, pp. 996–1015.
- Lima, Luiz Renato, Fanning Meng and Lucas Godeiro (2020). ‘Quantile forecasting with mixed-frequency data’. In: *International Journal of Forecasting* 36.3, pp. 1149–1162.
- Lisman, Johannes Hubertus Cornelius and Jan Sandee (1964). ‘Derivation of quarterly figures from annual data’. In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 13.2, pp. 87–90.
- Litterman, Robert B (1983). ‘A random walk, Markov model for the distribution of time series’. In: *Journal of Business & Economic Statistics* 1.2, pp. 169–173.
- Loughran, Tim and Bill McDonald (2011). ‘When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks’. In: *The Journal of finance* 66.1, pp. 35–65.
- Lucca, David O and Emanuel Moench (2015). ‘The pre-FOMC announcement drift’. In: *The Journal of Finance* 70.1, pp. 329–371.
- Lucca, David O and Francesco Trebbi (2009). *Measuring Central Bank communication: an automated approach with application to FOMC statements*. Tech. rep. National Bureau of Economic Research.

- Lundholm, Russell J, Rafael Rogo and Jenny Li Zhang (2014). ‘Restoring the tower of Babel: How foreign firms communicate with US investors’. In: *The Accounting Review* 89.4, pp. 1453–1485.
- Lütkepohl, Helmut (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Makonin, Stephen et al. (2015). ‘Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring’. In: *IEEE Transactions on smart grid* 7.6, pp. 2575–2585.
- Mann, Henry B (1945). ‘Nonparametric tests against trend’. In: *Econometrica: Journal of the econometric society*, pp. 245–259.
- Martens, Martin and Dick Van Dijk (2007). ‘Measuring volatility with the realized range’. In: *Journal of Econometrics* 138.1, pp. 181–207.
- Martin, Gassebner, Keck Alexander and Teh Robert (2006). ‘The impact of disasters on international trade’. In: *ETH Zurich (Swiss Federal Institute of Technology) Working Paper. O.*
- Mazzi, G, F Moauro and G Savio (2005). *Theory and applications of univariate and multivariate models for temporal disaggregation*. Tech. rep. Technical report, Statistical Office of the European Communities, Eurostat.
- Mc Laughlin, G Harry (1969). ‘SMOG grading-a new readability formula’. In: *Journal of reading* 12.8, pp. 639–646.
- Miller, Brian P (2010). ‘The effects of reporting complexity on small and large investor trading’. In: *The Accounting Review* 85.6, pp. 2107–2143.
- Miralles, Jose Manuel Pavía, Luis-Eduardo Vila Lladosa and Roberto Escuder Vallés (2003). ‘On the performance of the Chow-Lin procedure for quarterly interpolation of annual data: Some Monte-Carlo analysis’. In: *Spanish Economic Review* 5.4, pp. 291–305.
- Mishkin, Frederic S (2007). ‘Can central bank transparency go too far?’ In.
- Moauro, Filippo and Giovanni Savio (2005). ‘Temporal disaggregation using multivariate structural time series models’. In: *The Econometrics Journal* 8.2, pp. 214–234.
- Mosley, Luke, Idris A Eckley and Alex Gibberd (2022). ‘Sparse temporal disaggregation’. In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 185.4, pp. 2203–2233.

- Nakamura, Emi and Jón Steinsson (2018). ‘High-frequency identification of monetary non-neutrality: the information effect’. In: *The Quarterly Journal of Economics* 133.3, pp. 1283–1330.
- Neuhierl, Andreas and Michael Weber (2019). ‘Monetary policy communication, policy slope, and the stock market’. In: *Journal of Monetary Economics* 108, pp. 140–155.
- Newbold, Paul and David I Harvey (2002). ‘Forecast combination and encompassing’. In: *A companion to economic forecasting* 1, p. 620.
- Nijman, Theodoor Evert and Franz C Palm (1986). ‘The construction and use of approximations for missing quarterly observations: A model-based approach’. In: *Journal of Business & Economic Statistics* 4.1, pp. 47–58.
- Odhambo, Nicholas M (2011). ‘Financial intermediaries versus financial markets: A South African experience’. In: *The International Business & Economics Research Journal* 10.2, p. 77.
- Ojah, Kalu and Odongo Kodongo (2015). ‘Financial markets development in Africa’. In: *Monga, C. and Lin, JY (edd.) The Oxford handbook of Africa and economics* 2, pp. 401–423.
- Oliner, Stephen D, Glenn D Rudebusch et al. (1995). ‘Is there a bank lending channel for monetary policy?’ In: *Federal Reserve Bank of San Francisco Economic Review* 2.3, p. 20.
- Pavía-Miralles, Jose Manuel et al. (2010). ‘A survey of methods to interpolate, distribute and extra-polate time series’. In: *Journal of Service Science and Management* 3.04, p. 449.
- Pfeffermann, Danny (2015). ‘Methodological issues and challenges in the production of official statistics: 24th annual Morris Hansen lecture’. In: *Journal of Survey Statistics and Methodology* 3.4, pp. 425–483.
- Picault, Matthieu and Thomas Renault (2017). ‘Words are not all created equal: A new measure of ECB communication’. In: *Journal of International Money and Finance* 79, pp. 136–156.
- Pirouz, Dante M (2006). ‘An overview of partial least squares’. In: *Available at SSRN 1631359*.
- Polson, Nicholas (2017). ‘Teaching Notes for Probability 41901’. In.

- Proietti, Tommaso (2006). ‘Temporal disaggregation by state space methods: Dynamic regression methods revisited’. In: *The Econometrics Journal* 9.3, pp. 357–372.
- Quilis, Enrique M (2018). ‘Temporal disaggregation of economic time series: The view from the trenches’. In: *Statistica Neerlandica* 72.4, pp. 447–470.
- Reber, Ricci L and Sarah J Pack (2014). *Methods of Temporal Disaggregation for Estimating Output of the Insurance Industry*. BEA.
- Rennekamp, Kristina (2012). ‘Processing fluency and investors’ reactions to disclosure readability’. In: *Journal of Accounting Research* 50.5, pp. 1319–1354.
- Rigobon, Roberto and Brian Sack (2004). ‘The impact of monetary policy on asset prices’. In: *Journal of Monetary Economics* 51.8, pp. 1553–1575.
- Romer, Christina D and David H Romer (2004). ‘A new measure of monetary shocks: Rerivation and implications’. In: *American Economic Review* 94.4, pp. 1055–1084.
- Rosa, Carlo (2011a). ‘The high-frequency response of exchange rates to monetary policy actions and statements’. In: *Journal of Banking & Finance* 35.2, pp. 478–489.
- (2011b). ‘Words that shake traders: the stock market’s reaction to central bank communication in real time’. In: *Journal of Empirical Finance* 18.5, pp. 915–934.
- Rosa, Carlo and Giovanni Verga (2007). ‘On the consistency and effectiveness of central bank communication: Evidence from the ECB’. In: *European Journal of Political Economy* 23.1, pp. 146–175.
- Rossi, Nicola et al. (1982). ‘A note on the estimation of disaggregate time series when the aggregate is known’. In: *The Review of Economics and Statistics* 64.4, pp. 695–696.
- Rünstler, Gerhard and Franck Sédillot (2003). *Short-term estimates of euro area real GDP by means of monthly data*. Tech. rep. ECB working paper.
- Sainani, Kristin L (2014). ‘Explanatory versus predictive modeling’. In: *PM&R* 6.9, pp. 841–844.
- Sax, Christoph and Peter Steiner (2013). ‘Temporal disaggregation of time series’. In.
- Schmeling, Maik (2009). ‘Investor sentiment and stock returns: Some international evidence’. In: *Journal of empirical finance* 16.3, pp. 394–408.
- Silva, JMC Santos and FN Cardoso (2001). ‘The Chow-Lin method using dynamic models’. In: *Economic modelling* 18.2, pp. 269–280.

- Smales, LA and Nicholas Apergis (2017a). ‘Understanding the impact of monetary policy announcements: The importance of language and surprises’. In: *Journal of Banking & Finance* 80, pp. 33–50.
- Smales, Lee A and Nicholas Apergis (2017b). ‘Does more complex language in FOMC decisions impact financial markets?’ In: *Journal of International Financial Markets, Institutions and Money* 51, pp. 171–189.
- Snyder, David and Edward L Kick (1979). ‘Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions’. In: *American journal of Sociology* 84.5, pp. 1096–1126.
- Stock, James H and Mark W Watson (2002a). ‘Forecasting using principal components from a large number of predictors’. In: *Journal of the American statistical association* 97.460, pp. 1167–1179.
- (2002b). ‘Macroeconomic forecasting using diffusion indexes’. In: *Journal of Business & Economic Statistics* 20.2, pp. 147–162.
- (2006). ‘Forecasting with many predictors’. In: *Handbook of economic forecasting* 1, pp. 515–554.
- Sullivan, Megan (2024). ‘Understanding and predicting monetary policy framework choice in developing countries’. In: *Economic Modelling* 139, p. 106783.
- Svensson, Lars E O (2006). ‘Social value of public information: Comment: Morris and shin (2002) is actually pro-transparency, not con’. In: *American Economic Review* 96.1, pp. 448–452.
- Swanson, Eric T (2021). ‘Measuring the effects of federal reserve forward guidance and asset purchases on financial markets’. In: *Journal of Monetary Economics* 118, pp. 32–53.
- Taylor, Stephen J (2011). *Asset price dynamics, volatility, and prediction*. Princeton university press.
- Ter Ellen, Saskia, Vegard H Larsen and Leif Anders Thorsrud (2022). ‘Narrative monetary policy surprises and the media’. In: *Journal of Money, Credit and Banking* 54.5, pp. 1525–1549.
- Tetlock, Paul C (2007). ‘Giving content to investor sentiment: The role of media in the stock market’. In: *The Journal of finance* 62.3, pp. 1139–1168.

- Tibshirani, Robert (1996). ‘Regression shrinkage and selection via the lasso’. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288.
- Timmermann, Allan (2006). ‘Forecast combinations’. In: *Handbook of economic forecasting* 1, pp. 135–196.
- Trabelsi, Abdelwahed and Leila Hedhili (n.d.). ‘Methods Explained: Temporal Disaggregation Model Based Approaches to Benchmarking and Forecasting Economic Time Series: Current Status and Future Trends’. In: *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics* ().
- Trabelsi, Abdelwahed and Steven C Hillmer (1990). ‘Bench-marking time series with reliable bench-marks’. In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 39.3, pp. 367–379.
- Tumala, Mohammed Musa and Babatunde S Omotosho (2019). ‘A text mining analysis of central bank monetary policy communication in Nigeria’. In: *CBN Journal of Applied Statistics* 10.2.
- Unsal, D Filiz, Chris Papageorgiou and Hendre Garbers (2022). ‘Monetary policy frameworks: an index and new evidence’. In.
- Vyshnevskiy, Iegor, Wytone Jombo and Wook Sohn (2024). ‘The clarity of monetary policy communication and financial market volatility in developing economies’. In: *Emerging Markets Review* 59, p. 101121.
- Wasserman, Stanley and Katherine Faust (1994). ‘Social network analysis: Methods and applications’. In.
- Wei, William WS and Daniel O Stram (1990). ‘Disaggregation of time series models’. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 52.3, pp. 453–467.
- Winkler, Robert L and Spyros Makridakis (1983). ‘The combination of forecasts’. In: *Journal of the Royal Statistical Society: Series A (General)* 146.2, pp. 150–157.
- Wold, Herman (1982). ‘Soft modelling: the basic design and some extensions’. In: *Systems under indirect observation, Part II*, pp. 36–37.
- Wold, Svante et al. (1987). ‘Multi-way principal components-and PLS-analysis’. In: *Journal of chemometrics* 1.1, pp. 41–56.
- Woodford, Michael (2005). *Central bank communication and policy effectiveness*.

- Zaier, L Hedhili and M Abed (2014). ‘Temporal disaggregation of economic time series using artificial neural networks’. In: *Communications in Statistics-Theory and Methods* 43.8, pp. 1824–1833.
- Zheng, Isabel Yi and James Rossiter (2006). ‘Using monthly indicators to predict quarterly GDP’. In.
- Zou, Hui (2006). ‘The adaptive lasso and its oracle properties’. In: *Journal of the American statistical association* 101.476, pp. 1418–1429.