Markov Regime-switching in Forecasting Models

by

BAO ANH NGUYEN

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs

in partial fulfilment of the requirements for the degree of

Master of Science

in

Mathematics and Statistics

Carleton University

Ottawa, Ontario

Abstract

Markov Regime-switching in Forecasting Models

by Nguyen Bao Anh

This thesis studies the Markov-switching model and its applications in commodity markets with a long historical data series in the international context. The background theory of Markov chain is summarized, and the Markov switching model is introduced and discussed. Then, literature surrounding the Markov-switching model is reviewed from the earliest iterations of Hamilton to recent developments. Using this model, we show applicability in forecasting commodity prices; in particular, the gold spot price. The research imposes the features of Markov regime-switching models, considering gold as a financial asset to offer a comprehensive methodology for forecasting commodity price. We show that applying Markov regime-switching could significantly improve the forecast abilities in commodity prices. The analysis indicates that the abnormal increases of gold price in history always resulted from special economic conditions. This study makes a novel contribution to the field by demonstrating that the impact of CPI change to gold price is subject to the regimes, which is more sophisticated than what has been commonly accepted in economics literature to date.

Key words : Markov-switching Model, Gold Price, MS Auto-regression, Regime-switching

Acknowledgements

I would like to express my thanks and appreciation to the following individuals and organizations for their ongoing contributions and support during my research:

School of Mathematics and Statistics, Professors Yiqiang Zhao and Steven Wang, who supported and admitted me to the Master program in Mathematics and Statistics.

Professor Yiqiang Zhao, my supervisor, who provided continuous support to my study progress and was a source of great motivation and inspiration. His guidance helped me in researching and writing this thesis.

Professors Shirley E. Mills, Song Cai, who introduced me to new and valuable topics in the courses they taught. Dr. Peter Wesolowski, for his help with proofreading and valuable advice on thesis presentation. Professor Patrick Boily who set a high standard and provided insightful guidance in writing a mathematical paper, toward a perfection.

Madam Snezana Vucic, and my colleagues in Business Intelligence Research and Analysis Directorate of the Canada Revenue Agency, for their major support in career development during my studies.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC   Akaike Information Criterion

AR   Autoregression (Model)

ARCH   Autoregressive Conditional Heteroskedasticity

ARIMAX   Auto Regressive Integrated Moving Average with eXogeneous Input

BIC   Bayesian Information Criterion

CPI   Consumer Price Index

GARCH   Generalized Autoregressive Conditional Heteroskedasticity

HMM   Hidden Markov Model

LBMA   London Bullion Market Association

LL   Log Likelihood function

MAPE   Mean Absolute Percentage Error

MLE   Maximum Likelihood Estimate

MS   Markov-Switching

MSM   Markov-Switching Model

RMSE   Root Mean Square Error (RMSE)

List of Symbols

α, β, γ Coefficients of the models

$\vec{\eta}_{ij}$ Coefficients on the set of dependent and exogenous variables

ε_t Residuals, random, or irregular terms

$f(A|B)$ Probability density function of A, given B, 2 states

$g(A|B)$ Probability density function of A, given B, N-states

$\mathcal{F}_t$ Information available at time t, gold return model

$L(\vec{\theta})$ Likelihood function of parameters

Ω_t Information available at time t, gold price model

p_{ij} The transition probability from state i to j

p_{ij}^n The n-step transition probability from state i to j

$p_{ij}(t)$ The transition probability from state i to j, in time-varying model

P   Transition Probability Matrix

s_t State at time t (value)

S_t State at time t (variable)

σ_t^2 Variance of the data at time t

$\vec{\theta}$ The parameters' vector of the model

$\vec{\mathbf{x}}$ The predictors' vector of the model, including the lags if any

y_t Gold price at time t, also known as information available at time t

$\hat{y}_t$ Forecasting price at time t

Chapter 1

Introduction

The following chapter presents the importance and justification of studying the research topic 'Markov regime-switching in Forecasting Models', followed by the background of the research area, and the novel contributions to literature offered by the thesis. The structure of this thesis is introduced before advancing to the next chapter.

1.1 Motivation

In the first half of 2020, we observed extraordinary socio-economic situations and indicators of economic development. The economic growth rate of Canada was projected to be between 1.6 to 1.8% in 2020 [1] but turned out to have an actual drop to annualized -2.1% in the first quarter[2]. The Canadian oil patch faced the worst crisis in its history due to a confluence of factors, including a collapse in demand for all types of fuel caused by the Covid-19 shut down, and an unprecedented oil surplus due to the failure of the Organization of Petroleum Exporting Countries and their allies to reach an agreement on new production quotas until

---

[1]Source: Conference Board of Canada https://www.conferenceboard.ca/topics/economics
[2]Source: Statistics Canada https://www150.statcan.gc.ca

late April. The negative oil price in April, which actually shocked the statisticians, economists, and even computer application programmers, those who set positive value constraints for transaction price in software. The appreciation of precious metals and stock indexes was also observed. The gold spot price reached 1,715.90 USD/ounce, which is the highest price since the end of 2015 where it stood at 1,060.00 USD/ounce. Supposing that one bought and held gold from that year, it would be possible to see a gain of 70% from the initial investment. The Bank of America projected that the Covid-19 pandemic will drive gold to a lofty record by October 2021. The average gold price in 2020 will reach 1,695 USD/ounce before soaring demand pushes it to 2,063 USD/ounce the following year [3]. Above all, gold and other natural resource commodities in normal economic circumstances remain volatile as an individual asset. In a larger perspective, the downturn has affected all aspects of Canada's economy; though some sectors have been particularly hard hit in a manner equivalent to the 2008-09 financial crisis.

The interesting question emerging from this concerns how statisticians, economists, and researchers, among others, can detect the different behaviours of socio-economic data in times of crisis. Naturally, we consider the potential methodologies that could learn and analyze the changes of states in the past, apply the different parameters in different periods and estimate the possibility to switch between states in the forecast models. Markov Regime-switching would give us a chance to detect different regimes throughout history.

In times of crisis, while stock markets crash in general, some physical assets

---

[3]Source: https://www.msn.com/en-au/money/markets/

actually turn out to be on the top of the list of items that investors would like to hold for maintaining their value in their portfolios. Precious metals such as gold, silver, and platinum are always in an investor's mind. There is the obvious implication that the stocks of precious metal miners and producers would appreciate because their products bring increased profits. Another advantage of gold stocks is that they are subject to depreciation and appreciation like other stocks. So if there is a positive change in stock values, driven by either an increase in gold prices or market shocks, gold mining stocks go up. Because these appreciate like stocks, the extent of appreciation can be even more significant. That increases the justification for researchers and scholars to develop the application of Markov regime-switching in forecasting models on the price and returns of gold in the international or in a particular market.

1.2 Background

First introduced in 1989 in the seminal work of Hamilton [6], Markov switching models are a well-studied family of models in probability theory imposing time-varying parameters in different periods of time, known as state or regime-specific values. Statistically, this time-variation is governed by a discrete-valued latent stochastic process in the past. Specifically, as a feature of the Markov chain, the current value of the state is determined only by its value from the immediate previous period, and the transition probability matrix known as a system of dynamic transformation. The transition probability matrix regulates the properties of the Markov process by determining the probability that each of the states can be visited in the next period given the current state . The estimates

of the transition probability entries and the matrix can be evaluated using filtering and smoothing methods as well as the estimation of state-specific parameters. These two features make it possible to have improved interpretations of the parameters associated with specific regimes combined with the corresponding regime probabilities, as well as for improved forecasting of performance based on persistent regimes and parameters characterizing them.

The characteristics of the physical gold commodity are high liquidity, consumption, and convenient value storage. It is also a homogenous good which can be easily converted in transactions or by contract. Therefore, not surprisingly, gold has a special position among other precious metals as an investment item and a type of currency to prevent inflation. Previous studies in the literature provide limited evidence that the fluctuation of the gold market is affected by the same macroeconomic factors as is the case for other metals (in the paper Trends of Base Metals Prices [13] we show that macroeconomic factors such as technological progress, resource scarcity, natural resource taxes, and interest rates are the determinants of base metal prices). The risk-mitigating features of gold have been discussed adequately in prior literature, which has noted the increasingly important role played by gold as a hedge against inflation.

Practically, gold price forecasting is performed by an average prediction of a diverse panel of experts and gold market analysts. Their assessments of gold price trends are based on a variety of methods including: expert technical analysis, market fundamentals, current market sentiment, and an analysis of global economic and political events [4].

---

[4]https://www.gold-eagle.com/forecasts_predictions
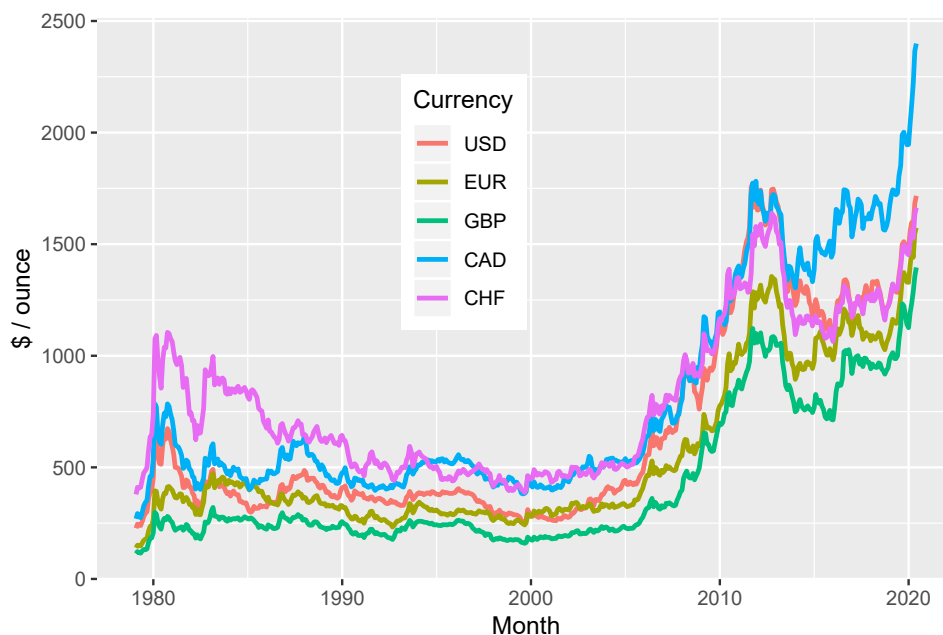
FIGURE 1.1: Gold Price in Major Currencies

In Figure 1.1 above we can see how the Gold spot price in major currencies changes throughout history [5]. In some periods, the price of gold in CAD and USD are quite close (from 2005 to 2013), while in others (from 2015 to 2020 or from 1980 to 1990) there are a gaps between their co-movements.

Figure 1.2 shows the relative contradiction of gold price and Consumer Price Index (CPI) in Canada from 1979 to 2000 [6]. Gold price moves in the opposite direction of inflation. This reflects the basic function of gold as an asset insofar as it stores value against inflation. Through a macroeconomics foundation, the relationship between a pair of currencies, or in other words, the exchange rate, depends on six factors: Differentials in Inflation, Differentials in Interest Rates, Current Account Deficits, Public Debt, Terms of Trade, and Strength of

---

[5]Data: World Gold Council. https://www.gold.org/goldhub
[6]Data: Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/

Economic Performance.
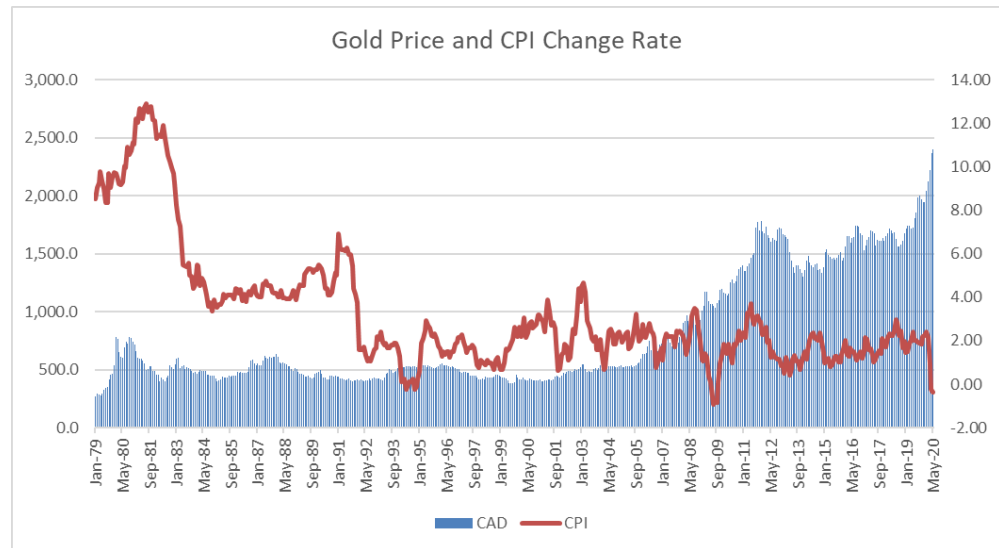
FIGURE 1.2: Gold Price and CPI

All of the many factors and relationships among them impacting hundreds of economies around the world suggest that it is not realistic to build a model capable of accounting for such an extensive number of determinants. In the present document we use the gold prices in another currency (USD) and Consumer Price index to predict the price in CAD. This model reflects the reality that Canada is not a major consumer of gold and that the gold price in CAD depends on the tendencies of markets in other economies.

The study of gold price and returns benefits investors of gold stocks, with the exception of physical gold. When an investor selects a gold stock, he or she is not buying an amount of physical gold, gold bullion or gold coins. Consider,

for example, mining stocks. The benefit is quick gains if the mining stocks perform well as for-profit economic entities. In a relative interaction, physical gold is better suited for anti-inflation investments. If we are investing in miners, we are investing in gold mining stocks, and therefore in particular companies. We assume that they are well managed as the standard in the industry. Indeed, one stock presents for the financial health or profit of one particular company. This research also refers to the index of gold stock in NYSE to study and illustrate the implication.

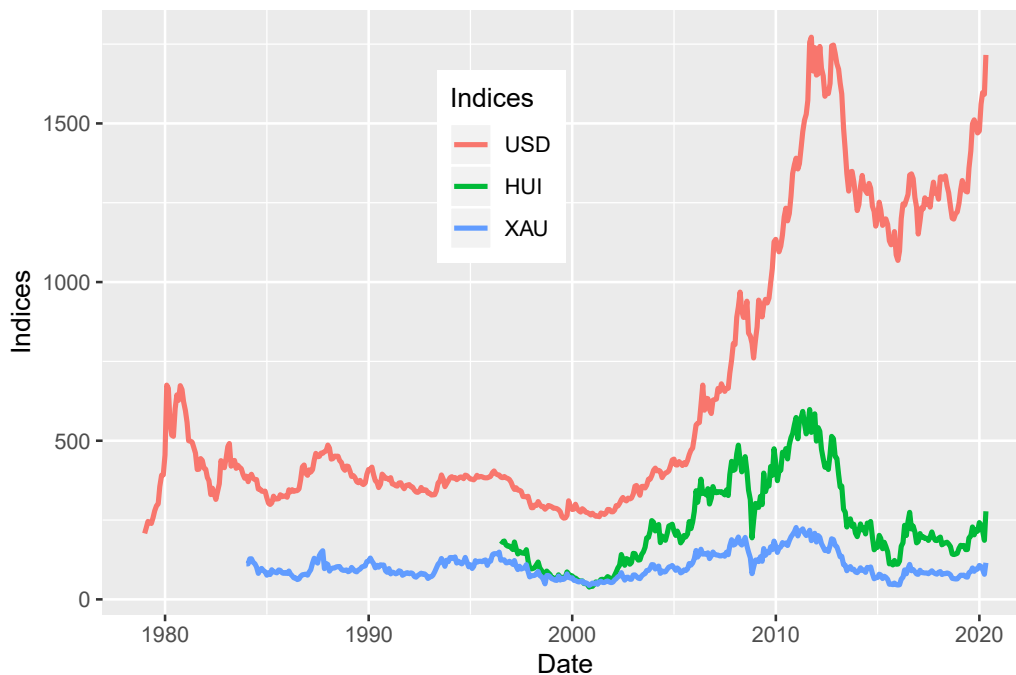FIGURE 1.3: Indices of Gold Stocks vs. Gold Price in USD

It may be of interest to commodity investors and stock exchanges that the price of gold would impact the indices of gold stock. If gold price is high, then the gold mines would gain better profits and hence, the stock price would increase. Figure 1.3 shows the co-movement of gold price in USD and the HUI, XAU

indices - the two most well-known gold indices in the financial markets [7].

1.3   Main Contributions

This study involves literature on Statistics of Financial Markets: Markov Chain, Auto-regression, and Regime-switching models. We illustrate the applications of Markov regime-switching model in an analysis of historical gold price, using the availability of data from the World Gold Council. The following are the major contributions of the work to literature:

- It contributes to the understanding of the development in Markov-switching Models studies, and their applications in forecasting works.

- The thesis provides a novel contribution, according to our best knowledge, introducing and presenting a forecasting model for gold spot price using Markov Regime-switching auto-regression and the selected exogenous variables: USD value and CPI change. We figure out that, change in CPI impacts gold price in a sophisticated manner, subject to the regimes.

- In doing so we are able to show that applying Markov regime-switching could significantly improve the forecast abilities in commodity prices.

- The study also analyzes the periods of time that gold price switches; abnormal increases of gold price always result from special economic conditions. Additionally, we apply the MS GARCH model on the gold returns series. The main results are interpreted and presented with proposals for

---

[7]Data: Yahoo Finance. https://ca.finance.yahoo.com/

potential future research. The study offers novel results for forecasting a particular type of asset in financial theory or fundamental statistics.

1.4 Organization

The remaining portion of the thesis is organized as follows:

- Chapter 2 reviews theoretical background, concepts, and methods for estimating the models' parameters. We systematize the literature of Markov regime-switching models, and summarize the development of studies in the field.

- Chapter 3 compares popular models and our MS auto-regression model on gold price forecasting works .

- Chapter 4 concludes and discusses potential research in order to fully capture these concepts in future research.

- We also introduce our first results in forecasting volatility of gold as an asset return, using the MS GARCH model in Appendix B.

Chapter 2

Theory & Literature Review

This chapter introduces the theoretical background in Markov chain, the concept and illustration example of transition probability matrix that are the foundation for the regime-switching models. The strategy to estimate parameters of the regime-switching model using maximum likelihood method is presented. The literature on Markov regime-switching models is summarized and the development of the studies in the field over the last 30 years is discussed.

2.1 Markov Chain & Regime-switching

Regime-switching models are the models that can characterize time series properties in different regimes. Models in which the switching among regimes occurs stochastically according to a Markov process are called Markov regime-switching models or Markov switching models (MS, MSM) for short.

We start with the definitions of basic terms and concepts that will be used to develop a unified approach for building our Markov switching model. Only the concepts of relevant terms are defined in this section because there are sufficient

definitions in textbook and literature.

2.1.1 Definitions and Concepts

In the below definitions, the state of a stochastic process $\{S_t, t = 0, 1, ...\}$ at time t is the value of S_t, denoted by s_t, indexed in its value space by i, j.

Definition 2.1: Markov Chain

A discrete time Markov chain is defined as a stochastic process $\{S_t, t = 0, 1, ...\}$ such that, at all time t, the probability of the future state S_{t+1}, is only dependent on the present state and independent of the past states, that is:

$$P(S_{t+1} = j | S_t = i_t, \ldots, S_1 = i_1, S_0 = i_0) = P(S_{t+1} = j | S_t = i_t)$$

this probability, denoted by p_{ij}, is the transition probability from i to j, the probability that the state will next be j when the current state is i.

Definition 2.2: Transition Probability Matrix

$$\mathbf{P}_{k,k} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,k} \end{bmatrix}$$

The Transition Probability Matrix of k states is a matrix whose (i,j)-entry is the transition probability to state j given the previous state is i.

The multi-step transition probabilities can be calculated recursively. In general, the $n-$step transition probability is

$$P\{S_{t+n} = j | S_t = i\} = P_{ij}^n \tag{2.1}$$

This leads to the formula for calculating multi-step transition probability, from state i to state j via state k, known as Chapman-Kolmogorov equation.

Theorem 2.1: Chapman-Kolmogorov equation

$$P_{ij}^{m+n} = \sum_{k=1}^{N} P_{ik}^n P_{kj}^m \quad i, j = 1, 2, ...N \tag{2.2}$$

The following properties of stochastic matrices hold:

Lemma: Properties of Transition Probability Matrix

- *All entries take values in the range [0,1], and each row sums to 1.*

- *In a Markov chain, the process moves from one state to another with the same probabilities at each step. The transition probability matrix is a stochastic matrix whose (i, j)-entry gives the probability that it moves from the ith state to the jth state during the next step of the process.*

- *The probability vector after n steps of a Markov chain is $\mathbf{P}^n\vec{v}$, where $\vec{v}$ is the initial probability vector and $\mathbf{P}$ is the transition probability matrix. A limit vector for a Markov chain is always a fixed point (a vector $\vec{v}$ such that $\mathbf{P}\vec{v} = \vec{v}$.).*

Example: Assuming that we have two regimes (or states), normal condition of the economy and the economy in crisis time, then the transition matrix would

be:

$$\mathbf{P}_{k,k} = \begin{bmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{bmatrix}$$

where $p_{1,1} + p_{1,2} = 1$ and $p_{2,1} + p_{2,2} = 1$, and

$$\mathbf{P}_{k,k} = \begin{bmatrix} P(\text{Normal at t+1} \mid \text{Normal at t}) & P(\text{Crisis at t+1} \mid \text{Normal at t}) \\ P(\text{Normal at t+1} \mid \text{Crisis at t}) & P(\text{Crisis at t+1} \mid \text{Crisis at t}) \end{bmatrix}$$

As an illustration, we would follow these steps to compute the transition matrix to model the transition between the normal condition and crisis time: Suppose that there are 200 observations, 125 observations show that they keep remaining in normal condition, 05 observations out of them exhibit the shift from normal condition to crisis time, 60 observations move from crisis to normal, 10 observations show that they remain in crisis given that the previous period of time is in crisis.

The frequencies, or the numbers of observations for the different transitions:

$$\mathbf{O}_{k,k} = \begin{bmatrix} 125 & 05 \\ 60 & 10 \end{bmatrix}$$

We then normalize the rows such that their sum is equal to 1:

$$\mathbf{P}_{k,k} = \begin{bmatrix} 0.962 & 0.038 \\ 0.857 & 0.143 \end{bmatrix}$$

In this example, the probability of switching from regime 1 normal condition

to regime 2 crisis time is 3.8%, the probability of staying in crisis is 14.3% if the previous state is crisis.

Definition 2.3: Conditional, Filtered, and Smoothing Probabilities

Conditional probability is the probability that S_t take the value s_t based on the information prior to time t (i.e $t - 1$), Filtering probability is the probability that S_t take the value s_t based on the information in the past and current information, and Smoothing probability is the probability that S_t take the value s_t based on the full sample information.

That is

$$P_{\text{conditional}} = P(S_t = s_t | \Omega_{t-1})$$

$$P_{\text{filtered}} = P(S_t = s_t | \Omega_t) \tag{2.3}$$

$$P_{\text{smoothing}} = P(S_t = s_t | \Omega_T)$$

In this thesis we study the estimation method using the maximum likelihood function. To assess the likelihood of the state variable S_t, it is critical to evaluate its optimal forecasts (conditional expectations) of $S_t = s_t$, based on different information sets. These forecasts include the predictions of $P_{conditional}$, $P_{filtered}$, and $P_{smoothing}$.

To estimate the parameters of the Markov-switching model we consider the joint conditional probability of each future states, as a function of the joint conditional probabilities of current states and the transition probabilities.

Definition 2.4: Filtering Process

A filtering process, with a system of dynamic transformation, is a process where the input is the conditional probabilities of the current states. The system of dynamic transformation is the transition probability matrix. The output is the conditional probability of future states.

That is

$$P(S_t = s_t | y_{t-1}) = P(S_t = s_t | S_{t-1} = s_{t-1}) \times P(S_{t-1} = s_{t-1} | y_{t-1}) \qquad (2.4)$$

where the output $P(S_t = s_t | y_{t-1})$ is the conditional probability of being in state s_t given information at $t - 1$. The input $P(S_t = s_t | S_{t-1} = s_{t-1})$ stands for the conditional probabilities of the current states. The system of dynamic transformation $P(S_{t-1} = s_{t-1} | y_{t-1})$ is the transition probability matrix.

2.1.2 Estimating Parameters

Two-state Markov Chain Process

The parameters of the system can be estimated by solving the maximizing likelihood function problem. The likelihood function is the probability density of the data, viewed as a function of the parameters, takes the dataset as a given. The conditional likelihood function is the product of conditional probability density functions. Basically, if we have a Markov chain process: $y_t = \sum_{i=1}^{k} I_i \vec{\theta_i} + \varepsilon_t$,

where $i \in [1,2]$, $I_i = 1$ in state i and 0 otherwise, and ε_t is a white noise residual. We want to evaluate the conditional probability $P(S_t = s_t | y_{t-1})$, which is the probability of being in state s_t given the information available at time $t - 1$. Using the transition probability and property, this task is quite possible by the filtering process and equation 2.4. Eventually we are able to estimate the filtering probability, the probability of being in state s_t, $P(S_t | y_t)$ by imposing the information we know at time t.

The joint density distribution of y_t and S_t

$$
\begin{aligned}
f(y_t, S_t = s_t | y_{t-1}) &= f(y_t | S_t = s_t, y_{t-1}) \times P(S_t = s_t | y_{t-1}) \\
&= f(y_t | S_t = s_t, y_{t-1}) \times P(S_t = s_t | S_{t-1} = s_{t-1}) \\
&\quad \times P(S_{t-1} = s_{t-1} | y_{t-1})
\end{aligned}
\tag{2.5}
$$

The density distribution of y_t

$$
f(y_t | y_{t-1}) = \sum_{s_t=1}^{2} f(y_t, S_t = s_t | y_{t-1})
\tag{2.6}
$$

The updated joint probability of y_t and S_t is

$$
P(S_t = s_t | y_t) = \frac{f(y_t, S_t = s_t | y_{t-1})}{f(y_t | y_{t-1})}
\tag{2.7}
$$

N-state Markov Chain Process

In reality we are facing many forecasting problems that have more than one states. They can be solved generally by an auto-regressive intergrated moving average with exogenous variables model where the data series is auto-regressive

of order p with exogenous variables x_{it} (ARIMAX). Let Ω_{t-1} denotes the available information set $\{x_k, y_k, k = t-1, t-2, \ldots, t-p\}$. The strategy to estimate the model's parameters is well studied in literature (for example [6], and [9]). Considering the data series as a Markov chain process, then the joint conditional probability of the current state and p previous states is

$$P(S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p} = s_{t-p}|\Omega_{t-1}) \tag{2.8}$$

The joint density distribution of y_t and S_t is

$$
\begin{aligned}
g(y_t, & S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}|\Omega_{t-1}) \\
&= g(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}, \Omega_{t-1}) \\
&\quad \times P(S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}|\Omega_{t-1}) \\
&= g(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}, \Omega_{t-1}) \\
&\quad \times P(S_t = s_t|S_{t-1} = s_{t-1}) \times P(S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}|\Omega_{t-1})
\end{aligned} \tag{2.9}
$$

The density distribution of y_t

$$g(y_t|\Omega_{t-1}) = \sum_{s_t=1}^{k} \sum_{s_{t-1}=1}^{k} \cdots \sum_{s_{t-p}=1}^{k} f(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}, \Omega_{t-1})$$

The updated joint probability of y_t and S_t at time t

$$P(S_t = s_t|y_t) = \frac{g(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1}|\Omega_{t-1})}{g(y_t|\Omega_{t-1})}$$

The summation over the states at lag p is just the output of the filter

$$P(S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p} = s_{t-p} | \Omega_t)$$

$$= \sum_{s_{t-p-1}=1}^{N} P(S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p-1} = s_{t-p-1} | \Omega_t)$$

The probability of the states at time t is obtained

$$P(S_t = s_t | \Omega_t) = \sum_{s_t=1}^{N} \sum_{s_{t-1}=1}^{N} \cdots \sum_{s_{t-p}=1}^{N} P(S_t = s_t, \ldots, S_{t-p-1} = s_{t-p-1} | \Omega_t) \quad (2.10)$$

The likelihood function, where $\vec{\theta}$ denotes the vector of the set of parameters

$$L(\vec{\theta}) = \prod_{t=1}^{T} g(y_t | \Omega_{t-1}; \vec{\theta}) \quad (2.11)$$

Logarithmizing the likelihood function yields

$$\mathscr{L}(\vec{\theta}) = \sum_{t=1}^{T} log \; g(y_t | \Omega_{t-1}; \vec{\theta}) \quad (2.12)$$

Solving the problems of maximizing the likelihood function, yields the MLE for parameters $\vec{\theta}$. For simplicity we illustrate the two-state Markov chain process of equation $y_t = \sum_{i=1}^{k} I_i * \vec{\theta}_i + \varepsilon_t : i = 1, 2$. Let $P_{tL}(1) = P(S_{t-1} = 1 | y_{t-1})$ and $P_{tL}(2) = P(S_{t-1} = 2 | y_{t-1})$. Assuming that the residual is normally distributed, then the maximum log-likelihood function is

$$L(\vec{\theta}) = \sum_{t=1}^{T} f(y_t | \Omega_{t-1}; \vec{\theta})$$

$$= \sum_{t=1}^{T} \sum_{s_t=1}^{2} f(y_t | S_t = s_t, y_{t-1}; \vec{\theta}) \times P(S_t = s_t | y_{t-1})$$

$$= \sum_{t=1}^{T} \sum_{s_t=1}^{2} \sum_{s_{t-1}=1}^{2} \{ f(y_t | S_t = s_t, y_{t-1}; \vec{\theta})$$

$$\times P(S_t = s_t | S_{t-1} = s_{t-1}) \times P(S_{t-1} = s_{t-1} | y_{t-1}) \} \qquad (2.13)$$

$$= \sum_{t=1}^{T} \Big\{ \frac{1}{\sqrt{2\pi}\sigma} \exp[\frac{-(y_t - \mu_1)^2}{2\sigma^2}] \times [p_{11} \times P_{1L}(1) + p_{21} \times P_{tL}(2)]$$

$$+ \frac{1}{\sqrt{2\pi}\sigma} \exp[\frac{-(y_t - \mu_2)^2}{2\sigma^2}] \times [p_{21} \times P_{tL}(1) + p_{22} \times P_{tL}(2)] \Big\}$$

where μ_1, μ_2 are the mean values of y_t in the two states.

Smoothing

At time t we have a set of information about what happened until the last period $t-1$, which we base our estimate for state S_t. Whenever the new information arrives, we take the chance to update the estimated S_t. We can also use the information at time $t+1$ or further in the future to smooth the data series. The smoothing process revises $P(S_t = s_t | \Omega_t)$ to reflect the most accurate estimate possible. Denote the revised smoothing probability by $P(S_t = s_t | \Omega_T)$

$$P(S_{t-p}, \ldots, S_t, S_{t+1} | \Omega_T) = P(S_{t-p+1}, \ldots, S_t, S_{t+1} | \Omega_T) \times P(S_{t-p} | S_{t-p+1}, \ldots, S_{t+1}, \Omega_T)$$

$$= P(S_{t-p}, \ldots, S_{t+1} | \Omega_T) = P(S_{t-p+}, S_{t-p}, \ldots, S_{t+1} | \Omega_T) \times P(S_{t-p} | S_{t-p+1}, \ldots, S_{t+1}, \Omega_t)$$

$$= \frac{S_{t-p+1}, \ldots, S_t, S_{t+1} | \Omega_T) \times P(S_{t-p}, S_{t-p+1}, \ldots, S_t, S_{t+1}, \Omega_t)}{P(S_{t-p+1}, S_{t-p}, \ldots, S_t, S_{t+1} | \Omega_t)} \qquad (2.14)$$

$$= \frac{S_{t-p+1}, \ldots, S_t, S_{t+1} | \Omega_T) \times P(S_{t-p}, S_{t-p+1}, \ldots, S_t, S_{t+1}, \Omega_t) \times P(S_{t+1} | S_t}{P(S_{t-p+1}, S_{t-p}, \ldots, S_t, S_{t+1} | \Omega_t)}$$

Assuming that the probability density function remains unchanged, or approximating over the two consequent periods $t - p$ to $t - p + 1$, that is

$$f(y_{t+1}, \Omega_{Tt} | S_{t-p}, S_{t-p+1}, \ldots, S_t, S_{t+1}, \Omega_t) \approx f(y_{t+1}, \Omega_{Tt} | S_{t-p+1}, \ldots, S_t, S_{t+1}, \Omega_t)$$

(2.15)

where $\Omega_{Tt} = \Omega_T - \Omega_t$ then the second equality holds, meaning that we will have

$$P(S_{t-p} | S_{t-p+1}, \ldots, S_t, S_{t+1}, \Omega_T) = P(S_{t-p} | S_{t-p+1}, \ldots, S_t, S_{t+1}, \Omega_t) \qquad (2.16)$$

Now summing over $S_{t+1} = 1, 2, \ldots N$ to obtain the probabilities of visiting $s_{t-p}, \ldots, s_t$, given the whole sample:

$$P(S_{t-p}, \ldots, S_t | \Omega_T) = \sum_{S_{t+1}=1}^{N} P(S_{t-p}, \ldots, S_t, S_{t+1} | \Omega_T) \qquad (2.17)$$

These are the smoothed states given no serial correlation. In the cases where there are no lags y_t involved, we obtain the smoothing probability

$$P(S_t | \Omega_T) = \sum_{S_t=1}^{N} \sum_{S_{t-1}=1}^{N} \cdots \sum_{S_{t-p}=1}^{N} P(S_t, \ldots, S_{t-p} | \Omega_T) \qquad (2.18)$$

That implies $P(S_t | \Omega_T)$ is equal the summation over all probabilities in states S_t to S_{t-p} in the whole space of state values.

Time-varying Transition Probabilities

We want to add a flexibility to the model with time-varying transition probabilities, which is the probability that from $S_t = i_t$ at time t the state moves to S_{t+1},

given the set of information at time $t + 1$ is Ω_{t+1}:

$$P(S_{t+1} = j | S_t = i_t | \Omega_{t+1}) = p_{ij}(t + 1)$$

. The transition probability matrix of k states is

$$P_t = \begin{bmatrix} p_{1,1}(t) & p_{1,2}(t) & \cdots & p_{1,k}(t) \\ p_{2,1(t)} & p_{2,2}(t) & \cdots & p_{2,k}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,1}(t) & p_{k,2}(t) & \cdots & p_{k,k}(t) \end{bmatrix}$$

As the output is binary variables they can be represented by probabilities in the form of logit or probit function. Depending on features of these functional forms we are able to set the range of the probability. To simulate time-varying switching we introduce $\vec{\eta}_{ij}$, a vector of coefficients on the set of dependent and exogeneous variables. The logit functional form of transition probabilities is:

$$p_{ij}(t) = \frac{1}{1 + \exp[-\Omega_t \vec{\eta}_{ij}]} \tag{2.19}$$

Now suppose that the term $-\Omega_t \vec{\eta}_{ij} = \alpha_{ij0} - \gamma_{ij} y_{t-1}$, a linear equation of y_{t1}, then

$$p_{ij}(t) = \frac{1}{1 + \exp[\alpha_{ij0} - \gamma_{ij} y_{t-1}]} \tag{2.20}$$

Because $\exp[-\Omega_t \vec{\eta}_{ij}] = \exp[\alpha_{ij0} - \gamma_{ij} y_{t-1}] \in [0, +\infty])$; we must have $p_{ij}(t) \in [0, 1]$. The exponential type transition probability is

$$p_{ij}(t) = 1 - \exp[-(\Omega_t \vec{\eta}_{ij})^2] \tag{2.21}$$

and supposing the same format of $-\Omega_t \vec{\eta}_{ij}$, the transition probability is

$$p_{ij}(t) = 1 - exp[-(\alpha_{ij0} - \gamma_{ij} y_{t-1})^2] \tag{2.22}$$

Similarly, because $\exp[-(\alpha_{ij0} - \gamma_{ij} y_{t-1})^2] \in [0, 1]$ thus we must have $p_{ij}(t) \in [0, 1]$.

2.2 Literature Review

We can divide the literature of Markov-switching models into two periods: from its introduction until the year 2000 it played a major role in solving problems in business cycle and financial research; after 2000, interest in switching models became broader with applications in technical issues and complex models of regime-switching. Within the scope of this thesis we are able to review some notable studies and applications.

2.2.1 Literature of the Field From 1989 to 2000

Initially originating from the publications of Hamilton (1989, 1990), Markov switching models have drawn much attention in technical, financial and economic data series analysis. In the long run there are changes in business environment, behaviors of customers or even the natural conditions. Collectively, these movements can be defined as regime shifts, where the parameters of the process changed. The Markov switching model exhibits more than one state or regime; it can present the dynamic process of the studied variables and shows us how these variables are evolving over times, and hence, gives us the perspective of the process in the future. Probably the majority of applied research

is in the area of business cycles where recent studies are still burgeoning. In the time of digital electronic devices booming with the availability of large datasets, the Markov switching model has been widely applied because there are more chances to discover different regimes in historical data [6].

Sola, Martin and Driffill, and Edward (1994), tested the term structure of interest rates using a stationary vector auto-regression with regime switching. In another paper Driffill and Sola (1998) apply the Markov switching model to justify whether there is an intrinsic bubble in stock prices so that stock prices deviate too far from the values predicted by the common predicted models: present value model or the fundamental asset evaluation using income and asset value. The paper concludes that a Markov switching model is a more appropriate representation of dividends because the dividends switch between regimes while stock prices can be better explained than through the bubble hypothesis.

In attempt to characterize the business environment and the switching between different regimes, Asea and Blomberg (1998) studied the lending behaviour of banks over lending cycles, using the Markov switching model with a panel data set consisting of approximately two million commercial and industrial loans granted by 580 banks between 1977 and 1993. They demonstrate that banks change their lending standards from tightness to laxity systematically over the cycle [5]. Town (1992), based on the well observed phenomenon that mergers take place in waves, fits the merger data into a Markov switching model with shifts between two states of high and low levels of activity. The results show an improvement over ARIMA models. [1].

The changing pattern of interest rates is indicative of business cycle conditions and could be subject to regime shifts itself. To investigate how real interest rates shift, Bekdache (1999) adopts a time varying parameter model with Markov switching conditional heteroscedasticity to capture two sources of shifts in real interest rates: shifts in coefficients and shifts in variance. The former relates the *ex ante* real rate to the nominal rate, the inflation rate and a supply shock variable, while the latter has unconditional shifts in the variance of the stochastic process. The results support a time varying parameter model over Markov switching with limited states [2]. Dewachter (1996) studies interest rate volatility by examining both regime shifts in the variance and links between volatility and levels of the interest rate. While regime shifts are found in the variance, the contribution of volatility-level links cannot be ignored. The above findings suggest that univariate or single element regime shifts in interest rate modelling fail to fully characterise interest rate dynamics [3].

Research conducted by Kim and Nelson (1999) classifies the economy into two states of booms and recessions, then investigates whether there has been a structural break in post-war US real GDP growth towards stabilisation. They use a Bayesian approach to identify a structural break at an unknown change-point in a Markov-switching model. Their empirical results suggest a break in GDP growth toward stabilization at the first quarter of 1984, and a narrowing gap between growth rates during recessions and booms [7]. In 2000, Maheu and McCurdy use a Markov switching model to classify returns into a high-return stable state and a low-return volatile state. The two states' bull and bear markets are respectively defined in the research question 'Identifying bull and bear markets in stock returns'. The paper's finding is that although bull markets

have a declining hazard function, the best market gains come at the start of a bull market. The paper also finds that volatility increases with duration in bear markets, which is intuitively foreseeable.

2.2.2 Literature of the field from 2000 to present

After the year 2000, the Hidden Markov models featured issues concerning mixture modeling. In 'A Markov switching model for annual hydrologic time series' by Akintug and Rasmussen (2005), the study assumes that the climate is switching between M states and that the state sequence can be described by a Markov chain. Observations are assumed to be drawn from a normal distribution whose parameters depend on the state variable. The paper presents the stochastic properties of this class of models along with procedures for model identification and parameter estimation. There are strong similarities between MS and ARMA models, as those are applied to the time series of the mean annual discharge of the Niagara River. Although it is difficult to draw any general conclusion from a single case study, it appears that MS models (and ARMA models derived from MS models) generally have stronger auto-correlation at higher lags than ARMA models estimated by conventional maximum likelihood [10].

In his PhD thesis 'Volatility Estimation and Price Prediction Using a HMM with Empirical Study', Pei Yin (2007) applies the HMM estimates to a price model and develops a prediction formula. On an empirical level, the study uses Monte Carlo simulation analysis and applies the model to 73 strategically selected data sets of historical security prices. The results suggest great applicability of the

HMM. Later, the author compares HMM and the well-established GARCH(1,1) with the same data sets to show the better perfomance of HMM [11].

Marco Bazzi et al., (2014) propose a new Markov switching model with time varying probabilities for the transitions. The novelty of the model is that the transition probabilities evolve over time by means of an observation driven model. The innovation of time varying probability is generated by the score of the predictive likelihood function. This paper also studies the dynamic mean and variance behaviour of U.S. Industrial Production growth. The findings include empirical evidence of changes in the regime switching probabilities, with more persistence for high volatility regimes in the earlier part of the sample.

N. Sopipan et al (2014) prove that the MRS-GARCH model is the best performing model for gold price volatility according to some loss functions. The paper forecasts closing prices of gold price to trade future contracts using data from 2007 to 2011 in the Thai market [12].

Shih-Tang Hwuy et al. (2016) propose a novel $N-$state Markov-switching regression model in which the state indicator variable is correlated with the regression disturbance term. The model admits a wide variety of patterns for this correlation, while maintaining computational feasibility. The paper considers two applications of the $N-$regime endogenous switching model, one to an empirical model of U.S. business cycles, and the other to a switching volatility model of U.S. equity returns [14].

Peter Nystrup et al. (2017) present an adaptive estimation approach that allows

for the parameters of the estimated models to be time varying. It is shown that a two state Gaussian hidden Markov model with time varying parameters is able to reproduce the long memory of squared daily returns that was previously believed to be the most difficult aspect to reproduce with a hidden Markov model. Capturing the time varying behavior of the parameters also leads to improved one step density forecasts. It is shown that the forecasting performance of the estimated models can be further improved using local smoothing to forecast the parameter variations [17].

Yoosoon Chang et al (2017) introduce a new approach to model regime switching using an auto-regressive latent factor, which determines regimes depending upon whether it takes a value above or below some threshold level. The paper develops a modified Markov switching filter to estimate the mean and volatility models with Markov switching that are frequently analyzed, and find that the presence of endogeneity in regime switching is indeed strong and ubiquitous [16].

Maciej Augustyniak, Luc Bauwens & Arnaud Dufays (2019) propose the factorial hidden Markov volatility (FHMV) process to model financial returns or realized variances. Its dynamics are driven by a latent volatility process specified as a product of three components: a Markov chain controlling volatility persistence, an independent discrete process capable of generating jumps in the volatility, and a predictable (data-driven) process capturing the leverage effect. An empirical study on six financial time series shows that the FHMV process compares favorably to state-of-the-art volatility models in terms of in-sample fit and out-of-sample forecasting performance over time horizons ranging from 1

to 100 days [21].

In summary, the Markov switching methodology was introduced by the seminal work of Hamilton (1989). Before the year 2000, Markov regime-switching is directly applicable to time series analysis in economics and finance for its dynamic nature. After the year 2000 there are significant studies to extend the scope of its applications, methodologies, and enhance its forecasting abilities.

Chapter 3

Forecasting Models & Results

This chapter presents the application and results of the MS auto-regression model with exogenous variables on the gold spot price data series. To illustrate the application, the two-regime model is compiled. We show that the MS auto-regression model can improve forecast ability, indicated by the measures of MAPE, and RMSE. The other criteria such as Akaike Information Criterion, R-squared are used to select the parameters and evaluated the goodness of predictors. The transition probability matrix, which characterizes the switch between regimes in historical gold price data, is evaluated.

3.1 Time Series of Gold Price

We begin by taking a look at the time series of gold price in Canadian dollars. Details about the datasets are described in the appendix A. Figure 3.1 demonstrates the monthly average gold price from Jan. 1979 to May 2020. Intuitively it can be seen that, an auto-regressive model would not yield a rational forecasting result because the impacts of economic factors are ignored. But the price in the previous month could provide a clue as to performance in the following month

with a certain range of errors. Monthly average price is calculated by the total monthly revenue, divided by the quantity traded in the whole month. Using the monthly average price helps excluding the temporary shocks or short-term tendencies of the market.
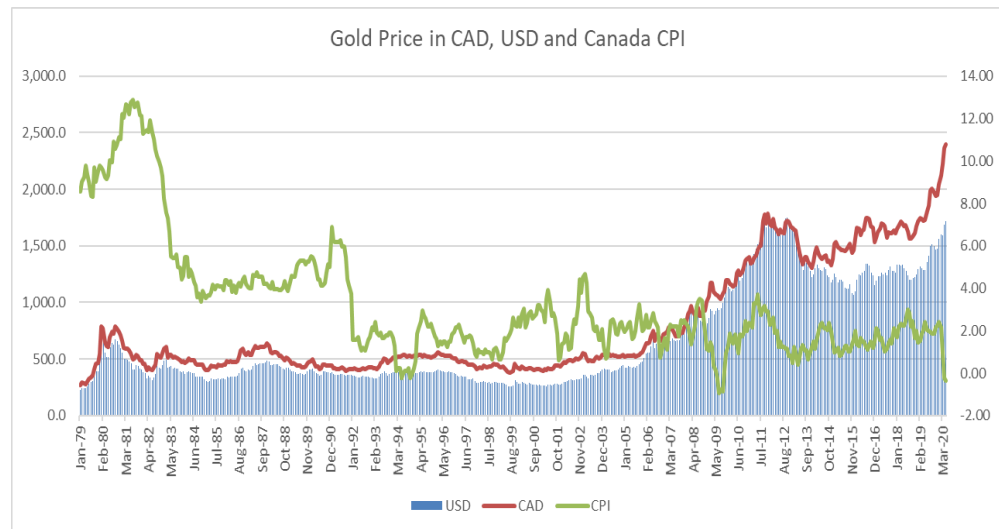

FIGURE 3.1: Gold Price in CAD, USD, and the CPI change rate

There are significant ever-changing variables and factors that make forecasting gold price accurately a challenge. Economically, the forecasters and analysts depend on primary factors that they think will impact the price of gold: the US dollar, investment demand, Central banks buying, exchange trading volumes, technical indicators, new mine supply, and other economic and monetary factors. Financially, gold and other asset prices are forecasted based on asset pricing models, which describes the expected rates of return of financial assets.

In the scope of this study we follow the advantageous notion that gold is a hedge against inflation of currencies, and that Canada is not a major gold buyer globally, to forecast and analyze gold price. As a result, the study covers the relationships with the transactions and values of USD – the relevant currency and domestic inflation. With the same statistical methodology, it is possible to use some other predictors accordingly for other research questions.

Extensive multi-disciplinary efforts have been spent on building a reliable model to forecast gold price. They include: the Economic Explanatory Model, Autoregressive Integrated Moving Average (ARIMA), Adaptive Neuro-Fuzzy Inference System (ANFIS), Multilayer Perceptron (MLP) Neural Network, Radial Basis Function (RBF) Neural Network, and Generalized Regression Neural Networks (GRNN). In the scope of our study we develop the application of Markov regime switching in auto-regression model, and the performance of the model is evaluated. Our proposed MS auto-regression model will be based on components of the explanatory and auto-regression models, therefore, we apply these two models first to obtain the necessary foundation for comparison.

3.2   The Explanatory Model

The gold price in Canadian dollar can be explained by its price in USD, a major currency in the world economy, and the change in CPI. Therefore, it is modeled as a linear function of price in USD and the rate of change of CPI in the Canadian economy. We can always take into account one or many other exogenous variables using a similar approach. Collinearity is an issue when using USD price as predictor, but that is a matter of fact; we could not reject its impacts to

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	112.56754	12.07728	9.321	< 2e-16 ***
USD	1.07277	0.01167	91.911	< 2e-16 ***
CPI	-6.72388	1.80322	-3.729	0.000215 ***

Residuals	Min	1Q	Median	3Q	Max
Value	-277.93	-42.23	2.13	28.93	445.81

Residual standard error: 108.5 on 494 degrees of freedom
Multiple R-squared: 0.951
Adjusted R-squared: 0.9508
F-statistic: 4792 on 2 and 494 DF
p-value: < 2.2e-16

TABLE 3.1: Fitted Parameters of the Explanatory Regression Model

the gold price in Canada and must analyze the price in USD as it is the most powerful currency globally.

$$\hat{y}_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t \tag{3.1}$$

where $\hat{y}_t$ is the forecasting spot price of gold at time t in CAD, x_{1t}, x_{2t} is the price in USD and CPI change rate at time t, respectively, ε_t is a random term, or forecasting residual which is normally distributed. In this model, x_{1t}, x_{2t} are exogenous variables; α, β_1, β_2 are the coefficients that need to be estimated.

The regression of $\hat{y}_t$ with explanatory variables x_{1t}, x_{2t} yields the outcome as in Table 3.1. Note the Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

That implies the explanatory model of gold price can be written as

$$\hat{y}_t = 112.56754 + 1.07277x_{1t} - 6.72388x_{2t} + \varepsilon_t \tag{3.2}$$

The negative coefficient of x_{2t} implies that CPI increases negatively impact to the gold price as we can roughly see in the Figure 3.2 above.

R-squared, which is a statistical measure, represents the proportion of the variance for a dependent variable that is explained by independent variables, takes a very high values, 0.951. Therefore we can say that the price can be well explained through the USD value and the CPI change. However, we can see that the residuals may not be normally distributed (the quartile values are -277.93; -42.23; 2.13; 28.93; 445.81). In spite of this we are confident that the two exogenous variables are good candidates for our MS auto-regression model, which will be presented later in this chapter.

3.3 The Auto-regression Model

Gold price is an auto-regressive process. The price depends on 'itself' based on the previous time periods, with the residuals normally distributed.

$$\hat{y}_t = \alpha + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t \tag{3.3}$$

where $\hat{y}_t$ is the forecasting spot price of gold at time t, y_{t-i} is the gold price at time $t - i$, α, γ_i are parameters that need to be estimated, p is the lag, $i \in [1, p]$, and ε_t is a random term, or forecasting residual which is normally distributed.

	AR1	Intercept
Coefficients:	0.9994	826.7293
S.E.	0.0011	1223.7952

σ^2 estimated as 1381: LL = -2505.31, AIC = 5016.61

TABLE 3.2: Fitted Parameters of the Auto-regression Model, AR(1)

The outcome of the auto-regression in R gives

Using lags $p = 2, 3, 4$ for the auto-regression model yields similar AIC values (4994.58, 4993.25, 4995 respectively). Therefore we choose AR(1) with the lag $p = 1$ for simplification, because the higher p does not improve significantly forecast ability. This reflects our intuition and feature of the Markov chain, that the current state of the process only depends on the state of immediately preceding period. Table 3.3 below shows the outcome of AR(4) model with similar AIC.

	AR1	AR2	AR3	AR4	Intercept
Coefficients:	1.2355	-0.3234	0.1030	-0.0158	834.4876
S.E.	0.0449	0.0715	0.0716	0.0457	1118.6860

σ^2 estimated as 1306:

	LL = -2491.57	AIC = 4995.13

TABLE 3.3: Fitted Parameters of the Auto-regression Model, AR(4)

The outcome of the model implies $y_t = (1 - 0.9994) * 826.7293 + 0.9994 y_{t-1}$ [15]. Hence, the auto-regression model is

$$\hat{y}_t = 0.4960 + 0.9994 y_{t-1} + \varepsilon_t \tag{3.4}$$

3.4 The MS Auto-regression Model

Suppose that gold price is a random process following two regimes: in one regime, gold price depends only on an exogenous variables; in another regime, gold price is an auto-regressive process, where the price at time t depends on the price at time $t - 1$. That means, we impose both of the two models above and evaluate the switch between them described by a Markov process.

The model

$$\hat{y}_t = \sum_{i=1}^{2} I_i \vec{\theta}_i \vec{\mathbf{x}}_i + \varepsilon_t \tag{3.5}$$

where $\vec{\theta}_i$ is the parameters' vetor of the regime $i, i \in [1, 2]$

$$\vec{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \end{bmatrix}$$

and $\vec{\mathbf{x}}_i$ is the predictors' vector of the regime i

$$\vec{\mathbf{x}} = \begin{bmatrix} 1 & x_1 & x_2 & y_{t-1} \end{bmatrix}^T$$

where I_i takes value 1 if the process is in regime i, 0 otherwise, and ε_t is the residual term at time t as usual. The Markov regime-switching model yields the following coefficients.

REGIME 1

Residual standard error:	52.59953		
Multiple R-squared:	0.9883		

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)(S)	3.4230	15.4871	0.2210	0.825092
USD(S)	0.0918	0.0282	3.2553	0.001133 **
CPI(S)	1.5468	1.7671	0.8753	0.381411
$CAD_1(S)$	0.9265	0.0239	38.7657	< 2.2e-16 ***

Standardized Residuals	Min	1Q	Median	3Q	Max
Value	-139.4999	-5.4652	-0.3805	0.6528	203.4511

REGIME 2

Residual standard error:	12.88244		
Multiple R-squared:	0.9962		

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)(S)	44.4512	4.2992	10.3394	< 2.2e-16 ***
USD(S)	0.1998	0.0235	8.5021	< 2.2e-16 ***
CPI(S)	-2.1165	0.3151	-6.7169	1.856e-11 ***
$CAD_1(S)$	0.7676	0.0249	30.8273	< 2.2e-16 ***

Standardized Residuals	Min	1Q	Median	3Q	Max
Value	-32.0437	-3.8700	0.0138	5.2002	35.7343

INFORMATION CRITERIA FOR BOTH REGIMES

AIC	BIC	LL
4611.161	4694.466	-2297.58

TABLE 3.4: Fitted Parameters of the MS Auto-regression Model, MS AR(1)

The outcome of the MS auto-regression in R gives equations for the two regimes, 3.6a represents Regime 1 and 3.6b represents Regime 2 :

$$\hat{y}_{t1} = 3.4230 + 0.0918x_1 + 1.5468x_2 + 0.9265y_{t-1} \tag{3.6a}$$

$$\hat{y}_{t2} = 44.4512 + 0.1998x_1 - 2.1165x_2 + 0.7676y_{t-1} \tag{3.6b}$$

The Markov-switching model has both regimes significant, where the R-squared values are 0.9883 and 0.9962.

Equation 3.2 describes the impact of CPI change to gold price in the explanatory model. But in 3.6a and 3.6b observe the more sophisticated effects of the hidden mechanism that controls this impact. In some periods of time the CPI change has positive marginal effects on gold price.

Table 3.5 at the end of this section shows the transition probability matrix: the transition probability matrix has high determinant values, which indicates that it is difficult to change from one regime to the other. The model perfectly detects the periods of each state. The residuals are white noise and show a better fit to the normal distribution in each regime than in the explanatory model (see Table 3.1).
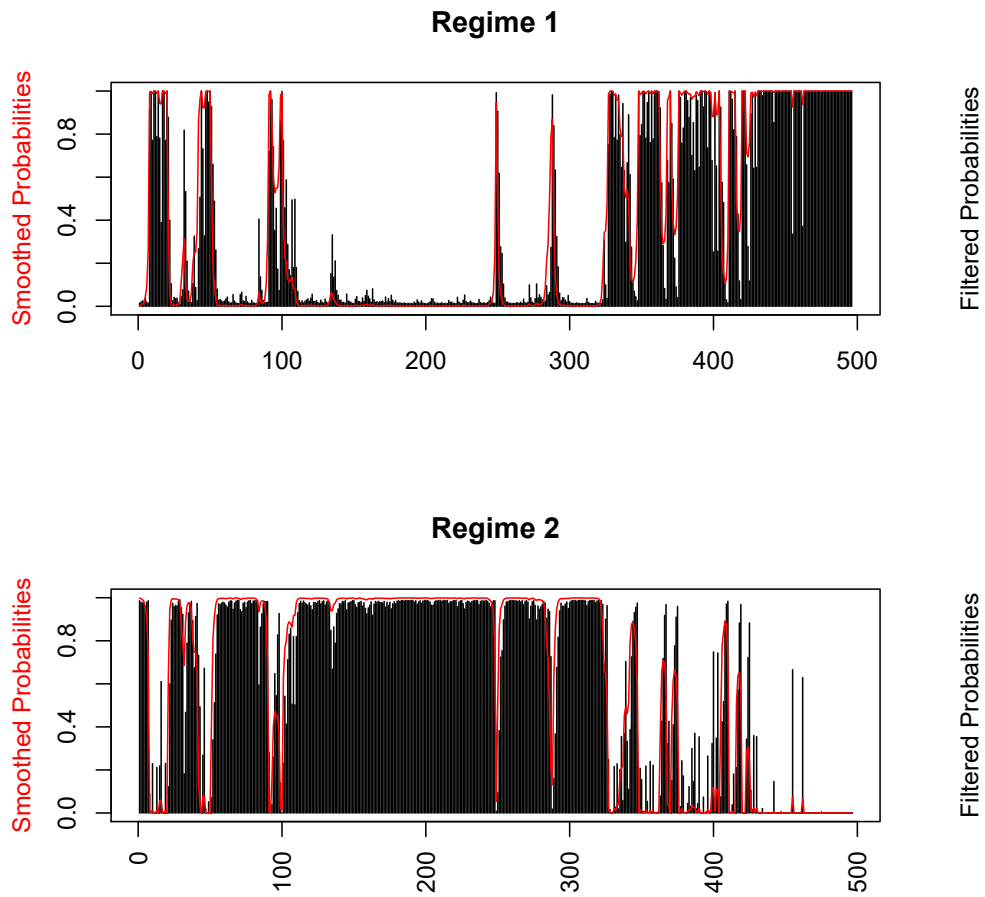
FIGURE 3.2: Graph of Filtered and Smoothed Probabilities

Figure 3.2 describes filtered and smoothed probabilities. Since the smooth probabilities are evaluated using the whole dataset while the filtered probability at time t are evaluated using data up to and including time t (but excluding time $t + 1, ..., T$), we realize a slight difference between them.
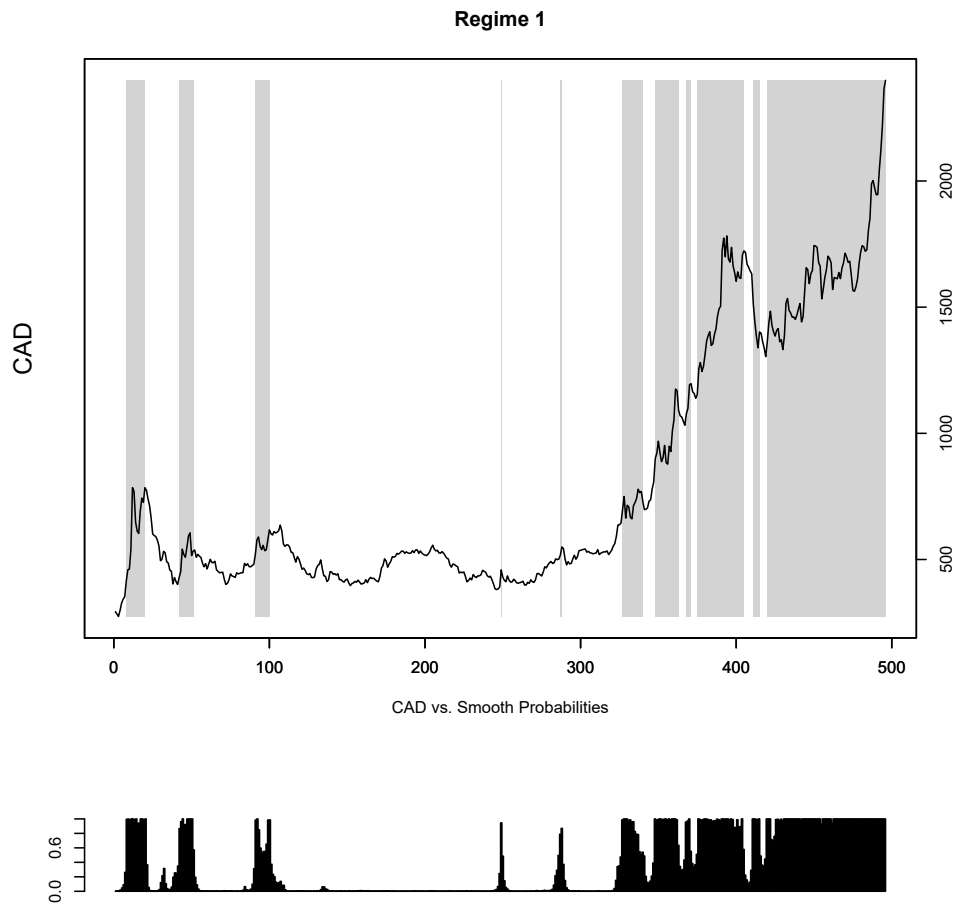
Regime 1

FIGURE 3.3: Graph of Price vs. Smoothed Probabilities

In Figure 3.3 the regimes and price are represented in the same graph. It shows that almost all the substantial increases of gold price are associated with regime 1. This implies that the abnormal increases of, or the shocks in gold price, resulted from special economic conditions.
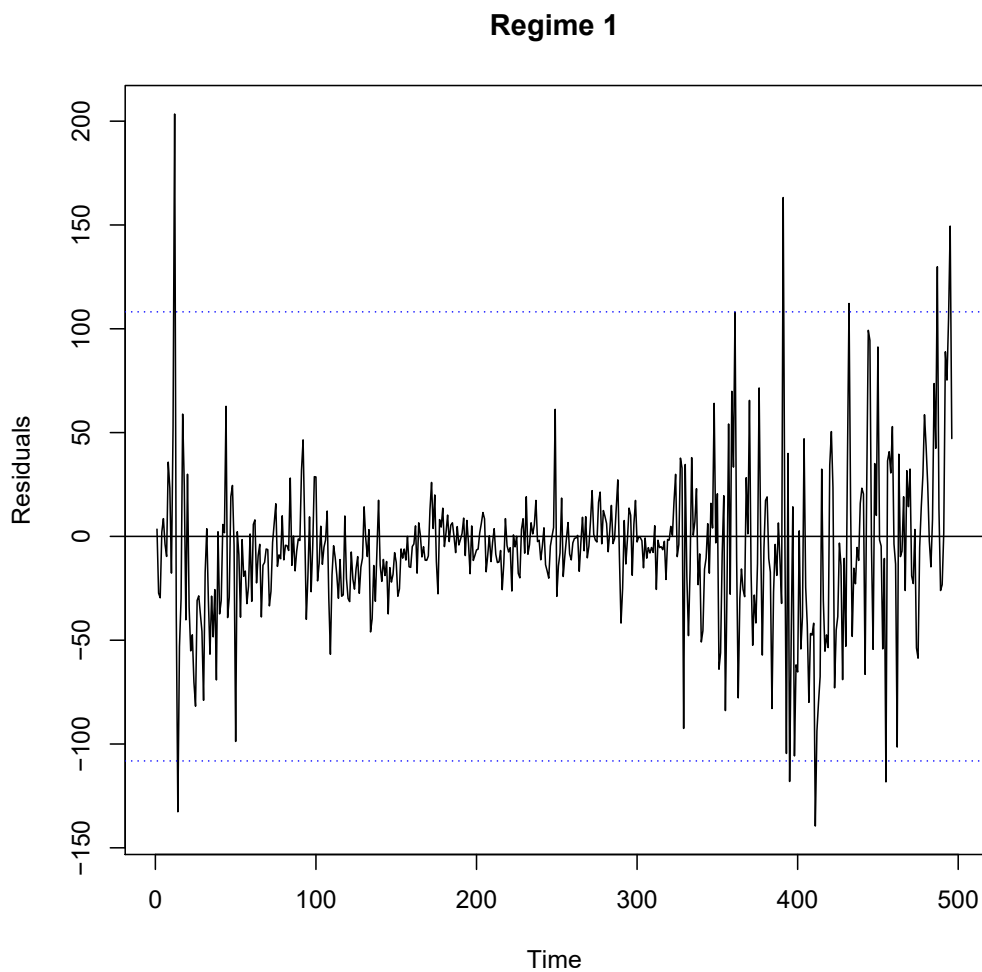
Regime 1

FIGURE 3.4: Graph of Residuals, regime 1

Figure 3.4 introduces the residual terms of the model with regime 1. This reveals, for example, in the periods from Jul. 2005 to May 2020 (observations from 330 to 497), gold price follows regime 1, and the residual is a white noise.
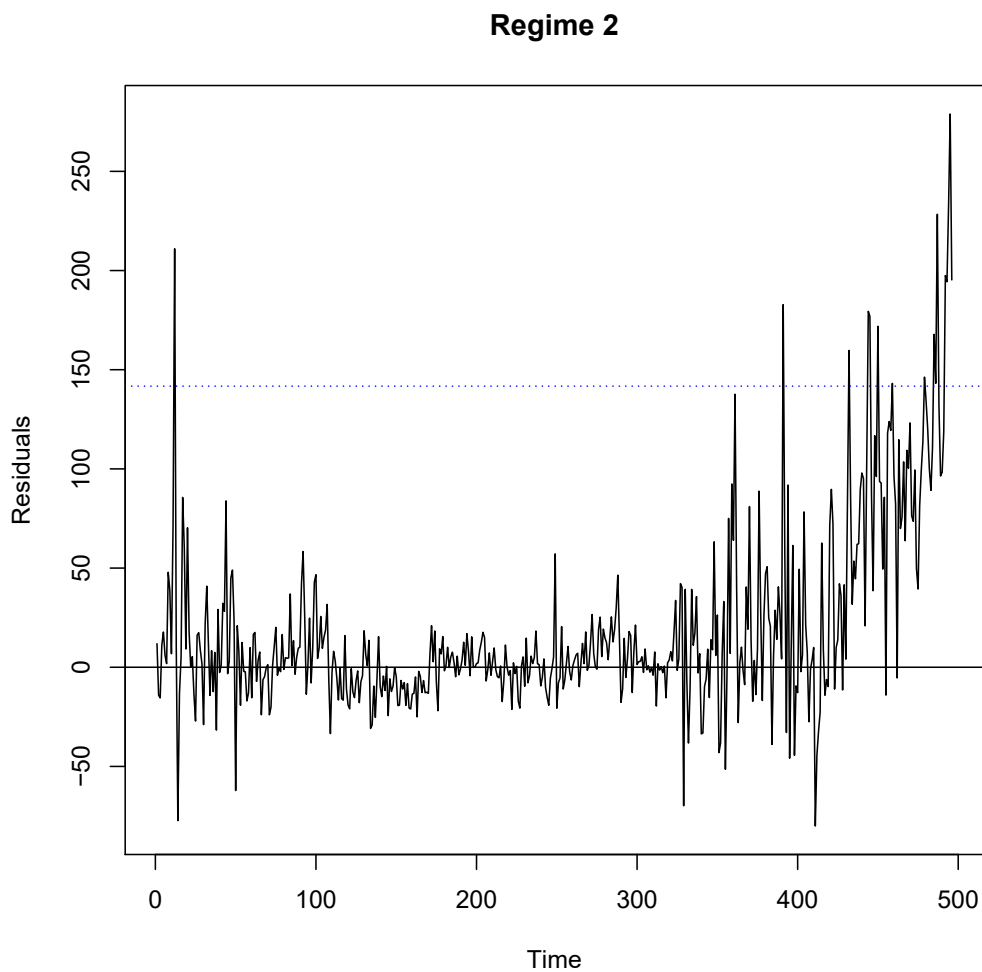
Regime 2

FIGURE 3.5: Graph of Residuals, regime 2

Figure 3.5 describes the residual terms of the model with regime 2. It can be clearly seen that, in the period from Jul. 2005 to May 2020 (observations from 330 to 497), the residual has a tendency with a mean greater than 0; it is a non-stationary process. This confirms the price follows regime 1.
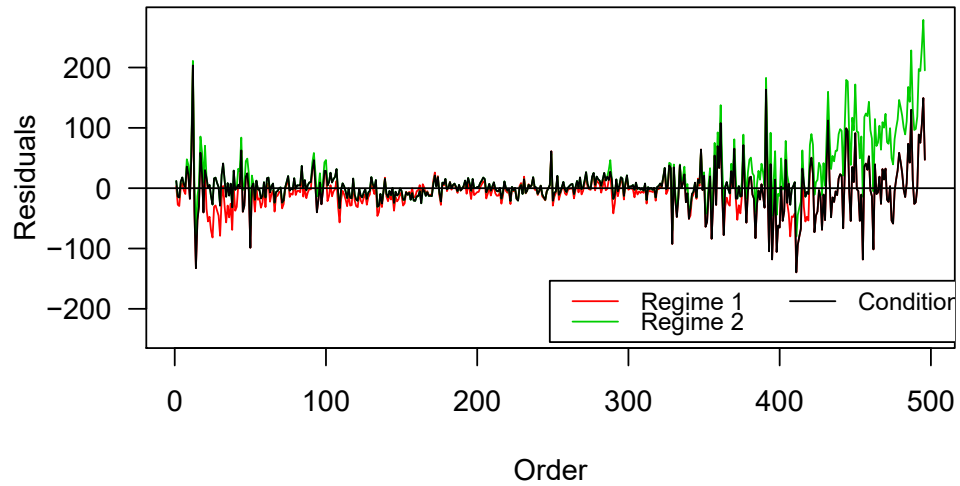
FIGURE 3.6: Graph of Residuals, the two regimes

By combining both of the two regimes in the model, the graph of residual terms fluctuates around zero. Visibly it looks like white noise. The calculation shows that the mean value of the residuals is $1.9744 \times 10^{-8} \approx 0$.

	Regime 1	Regime 2
Regime 1	0.93419448	0.06580552
Regime 2	0.03776094	0.96223906

TABLE 3.5: Transition Probability Matrix P(2,2)

3.5 Forecasting Work

Suppose that this month, gold price is in regime 1, probability that next month it will stay in regime 1 is 93.42% and visit regime 2 is 6.58%. Probability that gold price process stays in regime 2 for the two consecutive months is 3.78% while probability that it moves from regime 2 to regime 1 is 96.22%.

The dataset used as sample is from Jan. 1979 to May 2020. Now given that the spot price in USD for Jun. 2020 is 1732.2, the CPI change rate is 0.5472, the price in CAD in May 2020 is 2398.9. Using equations 3.6a and 3.6b obtains

$$\hat{y}_{t1} = 3.4230 + 0.09181732.2 + 1.54680.5472 + 0.92652398.9 = 2385.887$$

$$\hat{y}_{t2} = 44.4512 + 0.19981732.2 - 2.11650.5472 + 0.76762398.9 = 2230.794$$

The graph in Figure 3.3 shows that gold price is in regime 1. Using the transition probability matrix P(2,2) above yields the expected price in June $\hat{y}_t = 0.93422385.887 + 0.06582230.794 = 2375.672$

The actual price of gold in June, according to World Gold Council [1] is CAD 2347.5 per ounce. The forecasting error is $|1 - \frac{2375.672}{2347.5}| \approx 1.198\%$. Similarly, we use the formula 2.1 to obtain the $k-$step transition probability, then forecasting for k months ahead $\hat{y}_{t_0+k}$.

---

[1]https://www.gold.org/goldhub/data/gold-prices

3.6   Evaluation of Models' Performance

The most popular performance measure for forecasting models is the mean squared-error. There are two variants of this measure: Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

The Mean Absolute Percentage Error defined as

$$\text{MAPE} = \frac{\sum^N \frac{|\hat{y}_t - y_t|}{y_t}}{N}$$

where $\hat{y}_t, y_t$ are the forecasted price and actual price at time t, respectively. By our calculation, the auto-regression model AR(1) has an MAPE of 2.9862%. The explanatory model yields an MAPE of 7.3063%. Using the MS auto-regression model we obtain a forecasting error average of 2.6085%.

Using the Root Mean Square Error (RMSE) measure

$$\text{RMSE} = \sqrt{\frac{\sum^N (\hat{y}_t - y_t)^2}{N}},$$

our proposed MS auto-regression model has RMSE of 33.6233, one of the best performance among the models evaluated in literature [20].

Chapter 4

Conclusion & Potential Research

4.1 Conclusion

In this research, we studied the theoretical background of the Markov chain Regime-switching model, and its application in forecasting data series. The thesis proposes a Markov Regime-switching auto-regression model for forecasting gold price. The results show that the MS auto-regression model is very efficient in the analysis of historical data series, specifically, gold price in the Canadian market. The specific application of Markov regime-switching in forecasting gold price with two regimes is examined and compared with other popular analyzing methods. We show that the MS auto-regression performs dominantly. The forecasting model for gold price in two regimes is proposed

$$\hat{y}_{t1} = 3.4230 + 0.0918x_1 + 1.5468x_2 + 0.9265y_{t-1}$$

$$\hat{y}_{t2} = 44.4512 + 0.1998x_1 - 2.1165x_2 + 0.7676y_{t-1}$$

The R-squared values in the two regimes of the model are 0.9883 and 0.9962. Using Mean Absolute Percentage Error to evaluate the forecasting error, using the MS auto-regression model we can obtain the forecasting error average of

2.6085%. This is a significant improvement in comparison with other popular models in the literature. The model with exogenous variables gives MAPE = 7.3063% and the model with auto-regression yields MAPE = 2.9862%. The Transition Probability Matrix, which is considered as an important property of the specific Markov chain, is evaluated.

	Regime 1	Regime 2
Regime 1	0.93419448	0.06580552
Regime 2	0.03776094	0.96223906

We argue and present the statistical evidence that the dramatic changes in gold price likely happen in regime 1 which is explained by macroeconomic factors. Socio-economic conditions substantially impact gold price. The other commodity price series can be analyzed using the same method. We can also partition a long period of time into many regimes by giving $N \geq 2$ or applying $N = 2$ multiple times.

The forecasted price series fits very well with the actual gold price series as per the Figure 4.1 in the next page.

4.2   Potential Research

We know that, when the transition matrix **P** for a Markov chain is regular, then the Markov chain has a unique limit vector (known as a steady-state vector), regardless of the values of the initial probability vector by the property of the

Markov chain. If the transition matrix **P** for a Markov chain is regular, the positive powers of **P** approach a limit (which is a matrix) all of whose columns equal the chain's steady-state vector. This might be particularly beneficial to some type of commodity prices or asset returns (including stock prices) because then we will have a steady state, meaning stable prices over time.
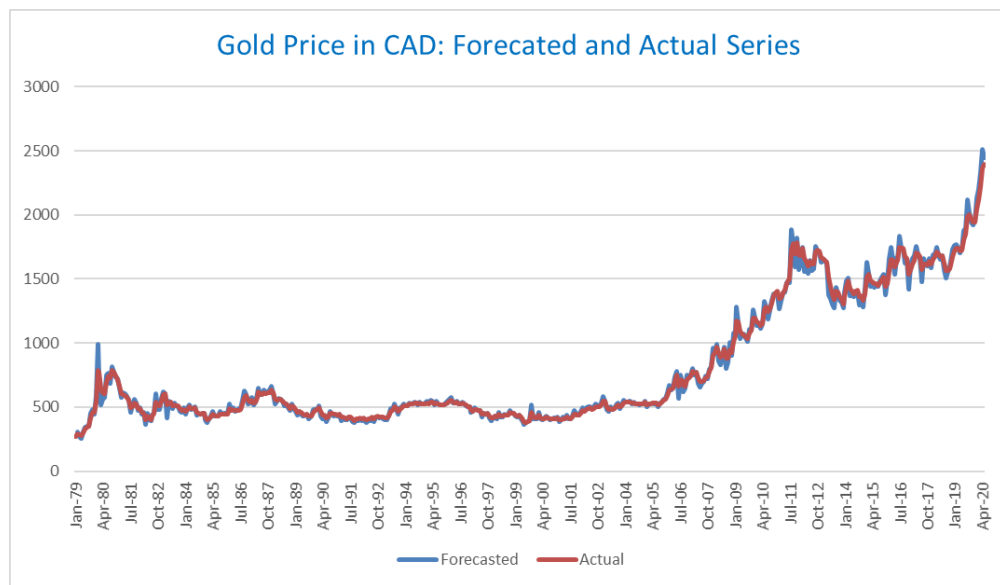

FIGURE 4.1: Gold price in CAD: Forecasted and Actual Series

The steady state is quite possible in reality when the demand and supply of a commodity are saturated in the market. In a business and legal environment, a study in historical data series can address the time and the conditions where entries to an industry would be stable. Suppose a transition probability matrix **P**, initial state vector $\vec{v}_0$, steady state vector $\vec{v}_\infty$; **I** denotes the identity matrix.

After k periods, the state vector is:

$$\vec{v}_0 \mathbf{P}^k = \vec{v}_\infty$$

Solving the following equation yields the steady state vector $\vec{v}_\infty$.

$$\vec{v}_\infty (\mathbf{P} - \mathbf{I}) = 0$$

In federal or provincial administration, the tax debts have significant impacts to government budgeting and public policies. Studies in steady-states of these processes would benefit the economy and social welfare by figuring out the circumstances that there would be no large deviation in tax collection and hence, there would be no more struggling with budget deficit.

In a world that is ever-changing, the stability or steady state of particular processes, including technical and natural phenomena, or our love to this life, would be something of extreme enjoyment.

Ottawa, August 2020

Appendix A

Data Cleaning and Manipulation

Data Sources

1. Gold Price: Thank to the availability of real time data on the spot price of gold in the international market by the World Gold Council [1] where the data is updated daily in major currencies. While this thesis studies the price on the current market at which the trading batch is bought or sold for immediate payment and delivery, gold not only has a spot price, but it also has the London Bullion Market Association (LBMA) gold price, as well as several regional prices. LBMA trading is conducted amongst members of the LBMA, loosely overseen by the Bank of England. Most of the members are major international banks or bullion dealers and refiners. The LBMA gold price is used as an important benchmark throughout the gold market, while the other regional gold prices are important to local markets. This data set provides the gold price over a range of time frames (daily, weekly, monthly, annually) going back to 1978, and in the major trading, producer, and consumer currencies.

---

[1]https://www.gold.org/goldhub/data/gold-prices

In this thesis we are interested in the monthly price series because it is appropriate for the long-term and relevant to our research question. Monthly data is calculated as an average of transaction price, total revenue divided by the total transaction volume of the month. Therefore it reflects the real expected value more closely, ruling out the emotion or some bubble, and hence it fluctuates less. Monthly data also exposes the seasonality, if any, more clearly while the seasonality of weekly data distorts from year to year, and daily data is not continuous, interrupted by weekends and the holidays.

2. Consumer Price Index: Thanks to the availability of the data series on CPI change from the Federal Reserve Bank of St. Louis[2] (Its Research Division is in the top 1% of all economics research departments worldwide). This is a US-based think tank with the mission to promote stable prices, encourage maximum sustainable economic growth and support financial stability throughout the United States [22].

The data is growth rate, or the rate of change, in comparison with the same period of the previous year. They are not seasonally adjusted. Unit is percentage, monthly.

Data Manipulation

Some figures use the index in the horizontal axis to present graphs tidily. The Table A.1 below refer index to the time labels (months)

---

[2]https://fred.stlouisfed.org/series/CPALTT01CAM659N

Table A1: Time & Index Reference

Index	1	50	100	200	300	350	400	497
Time	01/79	02/83	06/91	08/95	12/03	02/08	04/12	05/20

Data obtained is of high-quality. We check the set of quality criteria:

- Data-Type and Range Constraints: There are no missing data which make the algorithms and calculation undefined.

- Accuracy: Considering the degree of conformity of a measure to a standard or a true value, we take 1 decimal number for price series and 2 decimal numbers for CPI change rate.

- Consistency and Uniformity: The data series presents the nominal price. Original data remains unchanged, no conversion or changes are made. The data set contains prices that are quoted in currency unit (CAD and USD) per troy ounce.

We don't have any issue in mandatory constraints, regular expression patterns in the data.

Appendix B

MS GARCH Model on Gold Return

This Appendix presents the results of the MS GARCH model without logarithmizing the returns. In literature many studies use the log returns of the assets, if doing so, we're essentially making the assumption that there is no conditional variation in the mean.

Figure B.1 - Monthly gold return from Jan 1979 to May 2020
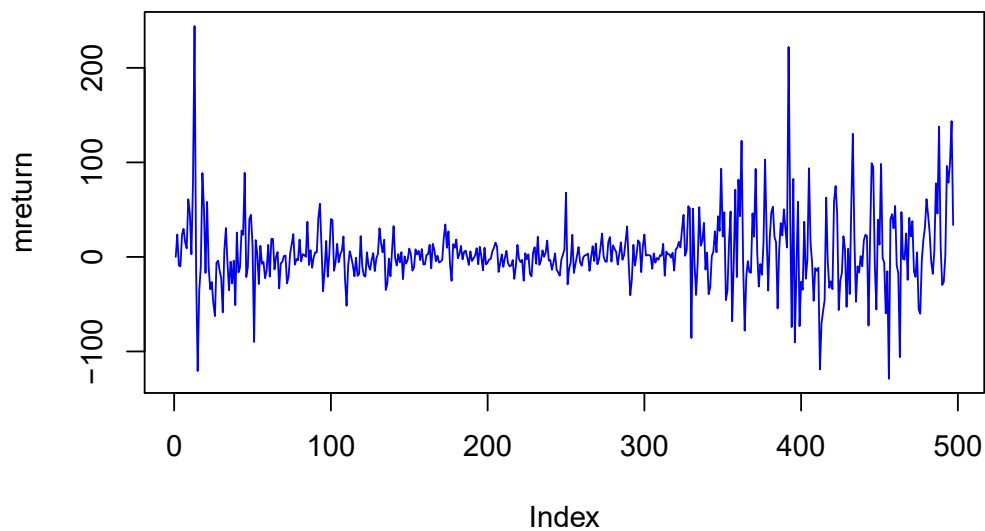
The graph above shows monthly gold return from Jan 1979 to May 2020. It addresses a regime of higher volatility from Jan. 1979 to Apr. 1987 (indices 1 to 100) and Aug. 2005 to May 2020 (indices 320 to 497) and another regime of low volatility.

Model GARCH(p, q) for gold return: GARCH(p, q) is the generalized form of the ARCH(q) model where p is the number of lag of the auto-regressive terms and q is the number of lag of irregular terms.

Let ε_t denotes a process of gold return, $\mathcal{F}_t$ denotes the information set at time t, which encompasses ε_t and all the past realizations of the process ε_t. The process ε_t is GARCH(p, q), if $E[\varepsilon_t|\mathcal{F}_{t-1}] = 0$,

$$\sigma^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

where,

- $\sigma^2 = Var(\varepsilon_t|\mathcal{F}_{t-1})$ and $Z_t = \varepsilon_t/\sigma$ is iid (strong GARCH)
- $\sigma^2 = Var(\varepsilon_t|\mathcal{F}_{t-1})$ (semi-strong GARCH), or
- $\sigma^2 = P(\varepsilon_t^2|1, \varepsilon_{t-1}, \varepsilon_{t-2} \ldots, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2)$ (weak GARCH)

Table B.1 - Fitted parameters of MS GARCH model

	Estimate	Std. Error	t value	Pr(>|t|)
alpha0_1	6.6514	3.7481	1.7746	3.798e-02
alpha1_1	0.0483	0.0374	1.2907	9.840e-02

```
beta_1        0.9055       0.0229 39.5062       <1e-16

alpha0_2 1649.9528 64394.0768   0.0256 4.898e-01

alpha1_2      0.0003       0.0112   0.0240 4.904e-01

beta_2        0.6145      15.0423   0.0409 4.837e-01

P_1_1         0.9595       0.0527 18.1947       <1e-16

P_2_1         0.1421       0.0290   4.9080 4.600e-07
```
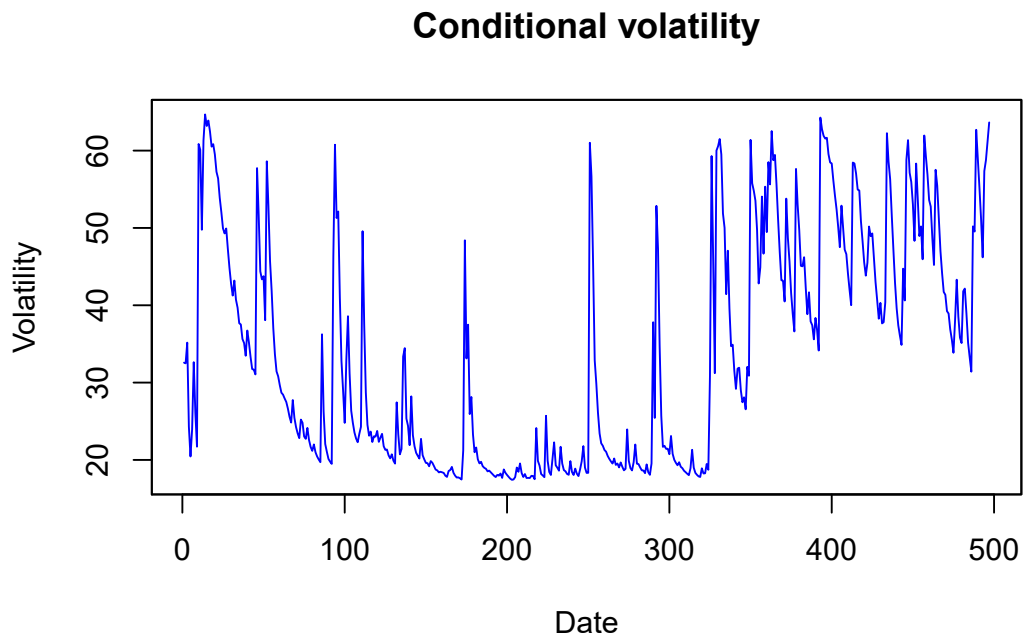
Table B.2 -Transition Probability Matrix

	Regime 1	Regime 2
Regime 1	0.9595	0.0405
Regime 2	0.1421	0.8579

Figure B.2 - Smoothed Probabilities
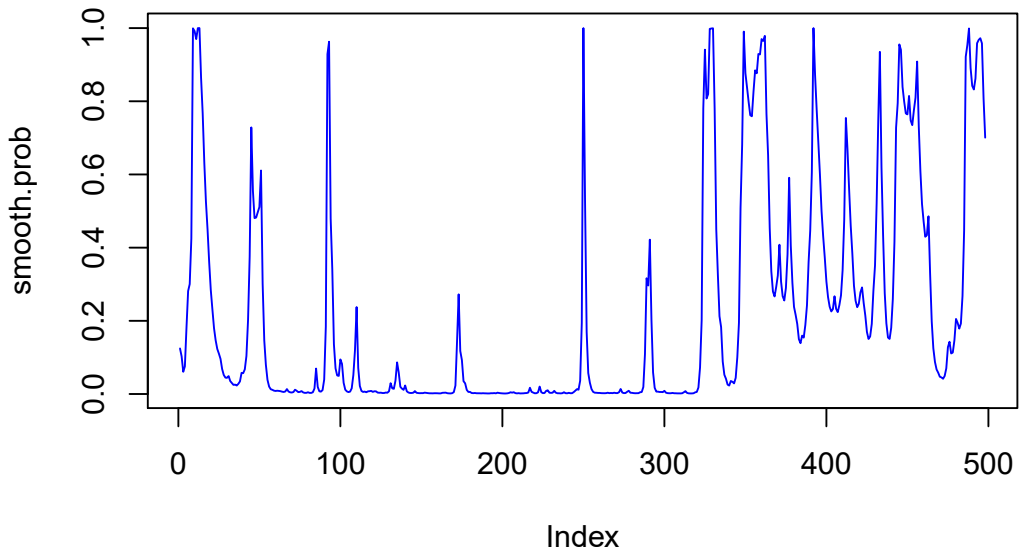
Figure B.3 - Conditional Volatility

Conditional volatility

Figure B.4 - Filtered Probabilities

Filtered for State 2
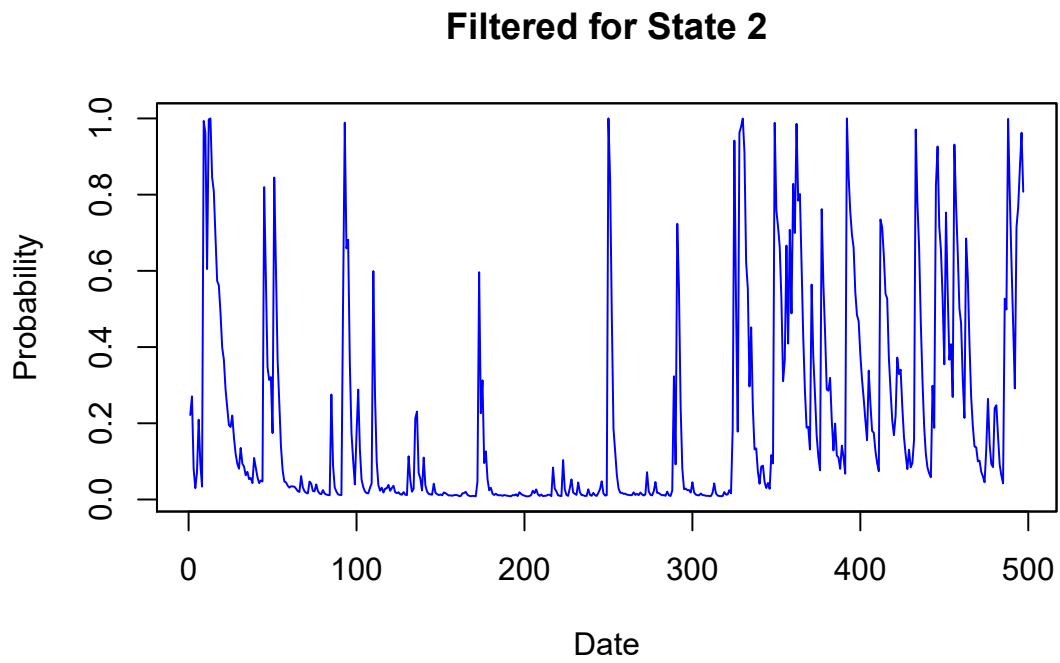
Bibliography

[1] Town, Robert, (1992), Merger Waves and the Structure of Merger and Acquisition Time-Series, Journal of Applied Econometrics, 7, issue S, p. S83-100.

[2] Bekdache, Basma, (1998), Alternative Approaches to Modeling Time Variation in the Case of the U.S. Real Interest Rate, Computational Economics, 11, issue 1-2, p. 41-51.

[3] Dewachter, H. (1996), Modelling interest rate volatility: Regime switches and level links. Weltwirtschaftliches Archiv 132, 236–258 . https://doi.org/10.1007/BF02707806

[4] Sola, Martin and Driffill, Edward, (1994), Testing the term structure of interest rates using a stationary vector autoregression with regime switching, Journal of Economic Dynamics and Control, 18, issue 3-4, p. 601-628.

[5] Asea, Patrick K. and Blomberg, Stephen, (1998), Lending cycles, Journal of Econometrics, 83, issue 1-2, p. 89-128, https://doi.org/10.1016/S0304-4076(97)00066-3

[6] J. D. Hamilton, (1989) A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica, Vol. 57, No. 2, , pp.357-384. doi:10.2307/1912559

[7] Kim, Chang-Jin and Nelson, Charles. (1999) Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle, The Review of Economics and Statistics volume 81 - 4, pages 608-616.

[8] John M. maheu Thomas H. McCurdy (2000) Identifying Bull and Bear Markets in Stock Returns, Journal of Business Economic Statistics, 18:1, 100-112, DOI: 10.1080/07350015.2000.10524851

[9] Wang, P. (2002). Financial Econometrics. London: Routledge, https://doi.org/10.4324/9780203892879

[10] Akintug, B., and Rasmussen, P. F. (2005), A Markov switching model for annual hydrologic time series, Water Resour. Res., 41, W09424, doi:10.1029/2004WR003605.

[11] Pei Yin (2007) Volatility Estimation and Price Prediction Using a HMM with Empirical Study, PhD thesis, University of Missouri-Columbia.

[12] N. Sopipan, P. Sattayatham and B. Premanode, (2014) Forecasting Volatility of Gold Price Using Markov Regime Switching and Trading Strategy, Journal of Mathematical Finance, Vol. 2 No. 1, 2012, pp. 121-131. doi: 10.4236/jmf.2012.21014.

[13] Nguyen Bao Anh, Aggey Semenov (2015) Trends of Base Metals Prices. Theoretical Economics Letters, 05, 531-540. doi: 10.4236/tel.2015.54062

[14] Shih-Tang Hwuy et al., (2016) N-State Endogenous Markov-Switching Models, Working paper, Department of Economics, University of Washington.

[15] Robert H. Shumway, David S. Stoffer, (2017) Time Series Analysis and Its Applications With R Examples, Springer International Publishing. eBook ISBN 978-3-319-52452-8.

[16] Yoosoon Chang et al., (2017) A new approach to model regime switching, Journal of Econometrics, Volume 196, Issue 1, January 2017, Pages 127-143. doi: doi.org/10.1016/j.jeconom.2016.09.005.

[17] Peter Nystrup, Henrik Madsen, Erik Lindström, (2017) Long Memory of Financial Time Series and Hidden Markov Models with TimeVarying Parameters, Jurnal of Forecasting, Volume36, Issue 8 December 2017 Pages 989-1002. https://doi.org/10.1002/for.2447

[18] David Ardia et al., (2017) Markov-Switching GARCH Models in R: The MSGARCH Package, Journal of Statistical Software, October 2019, Volume 91, Issue 4. doi:10.18637/jss.v091.i04.

[19] Marco Bazzi et al., (2017) Time Varying Transition Probabilities for Markov Regime Switching Models. Journal of Time Series Analysis, Volume38, Issue3, May 2017 Pages 458-478 https://doi.org/10.1111/jtsa.12211

[20] R. Hafezi, A. N. Akhavan, (2018) Forecasting Gold Price Changes: Application of an Equipped Artificial Neural Network, AUT J. Model. Simul. Eng., 50(1) 71-82. DOI: 10.22060/miscj.2018.13508.5074

[21] Maciej Augustyniak,Luc Bauwens Arnaud Dufays, (2019) A New Approach to Volatility Modeling: The Factorial Hidden Markov Volatility Model. Journal of Business Economic Statistics, Volume 37, 2019 - Issue 4. https://doi.org/10.1080/07350015.2017.1415910

[22] Organization for Economic Co-operation and Development, Consumer Price Index: Total All Items for Canada [CPALTT01CAM659N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPALTT01CAM659N, July 21, 2020