THREE ESSAYS IN APPLIED TIME SERIES ECONOMETRICS

A Dissertation

Presented to

the Faculty of the Department of Economics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Taylor Collins

August 2017

THREE ESSAYS IN APPLIED TIME SERIES ECONOMETRICS

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Economics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Taylor Collins

August 2017

Abstract

This dissertation is composed of four chapters. Chapter 1 provides an introduction to the paper by highlighting some of the key economic questions, econometric methods, and conclusions that this paper chronicles.

In Chapter 2, I conducte a range of unit root tests on the unemployment rates of 30 OECD countries. The objective of these tests are to use modern data and methods to update an old line of research that endeavors to uncover the most appropriate way to model unemployment. I find less evidence supporting Structural theories of unemployment than have prior studies in this field.

In Chapter 3, I turn my attention to US monetary policy. Specifically, I utilize a new estimation technique called the Beverage-Nelson Filter to construct output gaps for use in an introductory Taylor Rule study. I revisit a marquee paper from John Taylor, conduct a structural change test of Bai and Perron, and utilize a wide modeling of monetary policy rules. I find that the Federal Funds Rate displayed as strong an adherence to baseline Taylor Rules through the 1960s as in any other era.

Chapter 4 then turns the focus to New Zealand monetary policy and their role as the world's first inflation targeting country. In this chapter, I study the effects of the inflation rate and it's forecasted value for the following two years on New Zealand's Official Cash Rate and the country's Industrial Production Index. Using a set of Threshold Regressions and VAR Regressions, I find that New Zealand's interest rate responds much more strongly to the medium-run projected inflation than it does to inflation that is realized or projected to occur in the short run. I also find evidence that production in New Zealand is more responsive to changes in projected inflation than to changers in the interest rate.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This paper is compromised of three essays that I have worked on in my time at the University of Houston. Each of these three essays are concerned primarily with applying well-established econometric methods to answering macroeconomic and/or monetary policy questions. The first essay concerns the proper modeling of the unemployment rate of 30 OECD countries, a question I explore primarily through the use of linear and nonlinear unit root tests. The second paper revisits a marquee Taylor Rule paper by using the newly developed Beverage-Nelson Filter to construct output gaps rather than the Hodrick-Prescott Filter. The final paper utilizes Threshold Regressions and VARs to study the effects of New Zealand's inflation targeting regime, the first of its kind.

In the first essay of this dissertation, I apply a set 5 unit root tests on 30 OECD countries in order to uncover evidence on the best way to model the unemployment rate of these countries. There are three primary classes of theories into how the unemployment rate evolves (natural rate theories, hysteresis theories, and structural theories) and they each have very distinct statistical properties. Unit root tests are of immense value when determining the class of theory to use because they allow us to look for the specific statistical properties that we would expect to be present according to each of theses classes of theories. However, various advances in the field of unit root testing have been made since a comparable study was conducted on 16 countries in 2000. So this essay is meant to update an old line of research with extended datasets and modern unit root testing techniques.

In this exploration, I find less evidence for structural theories of unemployment than has traditionally been concluded. To do this, I conduct two separate unit root testing procedures to allow for a "high-end" and "low-end" estimate of the number of permanent breaks that can be considered consistent with structuralist theories of unemployment. Using a Markov-Switching Augmented Dickey Fuller test to serve as the high-end estimate, I conclude that 15 of the 30 countries should reject the unit root null, lending credibility to the claim that these countries could be better modeled by structuralist theories of unemployment than by hysteresis theories. In order to cover the "low-end" estimate of the number of breaks that structuralist theories allow, I utilize a testing procedure laid out in Silvestre, Kim, and Perron that allows a maximum of two permanent structural breaks. This tests concludes that the unemployment rates of only one third of all countries considered are consistent with structural theories of unemployment. Both of these tests provide less evidence supporting structural theories of unemployment that has older literature.

In the second essay, I begin by arguing that constructing output gaps for use in Taylor Rule research using widely established filtering techniques is an inferior approach to using the newly developed Beveridge-Nelson Filter (Kamber, Morley and Wong 2017). Specifically, I argue that the commonly used filter of Hodrick and Prescott (1980) poses a multitude of theoretical issues which threaten the accuracy and reliability of any output gap found using this technique. These issues are alleviated through the use of the BN Filter, a tool which has yet to be utilized in any Taylor Rule research to my knowledge.

I then turn my attention to utilizing these output gaps in three different ways. I first focus on reworking a key John Taylor 1999 paper. I find that use of the BN Filter leads to the conclusion that the tail end of the Bretton Woods era was the stretch during which baseline Taylor Rules were followed most closely, in direct contrast to John Taylor's results. I then attempt to estimate break dates endogenously using a structural change test from Bai and Perron. This test concludes that monetary policy in the 90s and the 60s followed the baseline Taylor Rule to approximately the same degree, though policy in the 60s appears to be slightly too aggressive while policy in the 90s appears slightly too conservative. I lastly apply a wide modeling approach that incorporates several developments made to the Taylor Rule over time. This wide modeling approach matches the conclusions of the revisited Taylor study in finding that the tail end of the Bretton Woods era was the time in which the Taylor Rule most strongly tracked the Federal Funds Rate. In all, there is ample evidence that the 60s were a period of strong baseline Taylor Rule policy adherence.

In my final essay, I take an introductory look at monetary policy in New Zealand. New Zealand was chosen because they were the first nation to introduce an official inflation targeting policy. They also hold the idea of transparency in particularly high regard, publishing several years worth of forward looking inflation forecasts every quarter. I seek to make use of these forecasts to answer two questions. Is the Reserve Bank of New Zealand acting in accordance with their stated monetary policy goal of achieving medium run price stability? And is the New Zealand market responding to projected deviations from that monetary policy goal in accordance with expectations?

I find the answer to the first question to be a resounding yes. Using Threshold Regressions and VARs, I find that the New Zealand interest rate responds much more strongly to inflation projected to miss its target in the medium run than it responds to projected deviations in the short run. The Reserve Bank of New Zealand appears to be systematically working to achieve their primary policy agenda.

The response of the overall New Zealand economy is significantly less straight forward however. Using VAR Regressions, I show that the Industrial Production Index actually increases when inflation is projected to exceed its target at medium run horizons, despite this being the time frame at which the interest rate responds most strongly. It appears that producers in New Zealand are ratcheting up production before expected inflation kicks in with the corresponding increased interest rates only applying marginal downward pressure.

The three essays all share the common thread of dealing in applied macroeconomic econometrics. The goal of these essays is not to develop new tools to look at interesting economic questions, but rather to find new applications of various tools which have already been established. The first essay attempts to answer the pure macroeconomic question of how best to model an unemployment series, while I turn my attention to monetary policy in the United States and New Zealand in the second and third essays.

Chapter 2

The Structure of OECD Unemployment

2.1 Introduction

The purpose of this paper is to test for the presence of unit roots in the unemployment rates of 30 OECD countries. Unit root testing of unemployment series has been a common endeavor ever since Blanchard and Summers' (1986) Hysteresis paper cast doubts on the longstanding dominance of the natural rate theories of Phelps (1967, 1968) and Friedman (1968), particularly amongst the unemployment rates of European countries. This paper will follow in the line of Papell, Murray, and Ghiblawi (2000) in that I will utilize unit root testing in an attempt to find evidence supporting one of three general sets of unemployment theories: traditional natural rate theories, hysteresis theories, and structural theories. But where PMG find strong evidence for structural theories of unemployment, I find significantly less.

The traditional theories of Phelps and Friedman describe the unemployment rate as fluctuating in the short term, but always gravitating back to a "natural rate" of unemployment, generally defined as some market equilibrium. Since the unemployment rate is always gravitating back to some natural level, all shocks to the economy only have a temporary effect on unemployment. Thus, these theories imply that the unemployment rate should be stationary.

Blanchard and Summers concluded that these theories failed to well-describe what was happening to European unemployment rates. As a result, they proposed an alternative theory for unemployment, one they termed hysteresis. As defined by Blanchard and Summers, hysteresis is "a very high dependence of current unemployment on past unemployment." Rather than shocks only having temporary effects, hysteresis posits that temporary shocks will permanently alter the level of unemployment. Empirically, this should mean that the unemployment series contains is non-stationary.

Finally, structuralist theories of unemployment are characterized by "endogenizing the natural rate of unemployment - defined now as the current equilibrium steady state rate, given the current capital stock and any other state variables," as defined by Phelps (1994). As with the traditional theories of unemployment, structuralist theories maintain that the equilibrium path of unemployment is approaching a natural rate. However, in the structuralist theories of unemployment, the natural rate can move over time. According to the structuralist theories, most shocks to unemployment will only cause temporary movements around the natural rate, but there are some which will cause a permanent change in the unemployment rate itself. Empirically, structuralist theories expect to see the unemployment rate be stationary around a process that is subject to a few structural breaks.

This sets up a natural test to perform. If statistical evidence of a unit root is found in an unemployment series, this can be taken as evidence that a hysteresis model might most appropriately describe the behavior of that unemployment series. If evidence of stationarity is found, this would lead researchers to chose models more consistent with natural rate theories. And if stationarity around a process with structural breaks is found, it will point towards structuralist models being most appropriate. This paper is meant to contribute to the field of unit root testing unemployment by applying some of the most modern unit root testing techniques to a set of 30 OECD countries. I conduct a total of 5 separate unit root tests for each of the 30 countries in my sample. The unemployment series of each of the 30 countries can be found in Figure 2.1

The analysis of the results of these tests can occur across two different spectrums. First, if any given test yields the same conclusion across an overwhelming majority of countries, it can be taken as evidence that the theory associated with that particular outcome is likely to be a best fit when choosing the appropriate unemployment model for some arbitrary country. Second, if any given country yields the same conclusion across the spectrum of unit root tests conducted, this can be taken as particularly strong evidence that the given country should be modeled according to the corresponding model.

To conduct this analysis, I use a total of 5 unit root tests in all. I begin with the commonly used Augmented Dickey Fuller test and Dickey Fuller Generalized Least Squares test as a starting point. Using the ADF test, I can reject the unit root null for only 9 of the 30 countries. The DF-GLS test does not yield a much different conclusion, rejecting the unit root null for only 7 countries. Using these widely established tests alone would lead one to conclude that there is very little evidence supporting natural rate of unemployment theories over hysteresis theories.

However, these tests are known to potentially fail to reject the null when there is misspecification of the trend. Further, if the series is well described with a trend, hypothesis testing on the coefficient of that trend is dependent upon whether the noise component is $I(0)$ or $I(1)$. But this information is not likely known by the researcher beforehand. As such, I utilize a unit root test created by Perron and Yabu (2009) that is valid under either scenario, and requires no a priori knowledge. The results of this test are largely consistent with the older unit root tests as only 8 of 30 countries reject the unit root null, once again indicating that unemployment data tends to be inconsistent with natural rate of unemployment theories.

Then, in order to allow for the inclusion of structural theories of unemployment, I next conduct a Markov Switching unit root test. None of the lags in this regression are regime dependent, only the intercept. This allows me to conduct a unit root test on the series while also allowing two distinct "natural rates" to which the unemployment rate might gravitate towards over time. Incorporating this nonlinearity increases the number of countries for which I reject the unit root null to 15, providing some evidence for structuralist theories of unemployment dominating those countries (though still less than Papell, Murray and Glibwali).

Finally I conduct a unit root test from Silvistre, Kim and Perron (2009) which allows the possibility that a trend could be present and changing over time at a limited number of breaks. This test utilizes a structural break version of Perron and Yabu's procedure as a pre-test to determine the appropriate number of breaks to include in the trend. The pretest concludes that 10 of the 30 countries should have no structural break, 4 of the 30 countries should have a single structural break, and the remaining 16 countries should have two structural breaks. Amongst the 20 countries experiencing at least one structural break, the unit root null is rejected for only 10. Thus, whereas the Markov Switching Unit Root test provided a "high-end" estimate for the prevalence of structural theories, the Silvestre et. al procedure provides a "low-end" estimate.

Taken together, there is evidence to indicate that the unemployment rates of between

50% and 33% of the countries studied are best described by structural theories. The high-end estimate found from the Markov Switching test comes close to the results from Papell, Murray, and Glibwali, who rejected the unit root null for 63% of the the countries they considered (10 of 16). The more conservative conclusion found from applying the Silvestre et. al procedure is a significant departure from that finding, indicating that structural theories of unemployment may be less appropriate than previously thought.

The rest of the section proceeds as follows. Subsection 2.2 establishes a baseline of understanding for the paper by conducting the well established ADF and DF-GLS tests. In subsection 2.3, I conduct a unit root test from Perron and Yabu with better size and power than these older procedures. Subsection 2.4 incorporates non-linearity using a Markov Switching Augmented Dicky Fuller test. And in subsection 2.5 I conduct another version of this Perron and Yabu test (a version which allows for structural change) as a pretest for a Silvestre, Kim and Perron unit root testing procedure. Subsection 2.6 concludes.

2.2 Linear Unit Root Tests

I want to begin this study by testing for unit roots using some of the oldest and most widely known unit root tests available. In this section, two tests will be conducted in order to establish a baseline around which we can compare the results of more modern unit root testing techniques. The two tests I utilize in this section will be the Augmented Dickey Fuller test (ADF) and the Dickey Fuller Generalized Least Squares test (DF-GLS). I run these tests on 30 individual OECD countries with quarterly data extending through either 2016Q3 or 2016Q4 for each. The start date of each series varies depending on data availability, with the lengthiest series beginning in 1955Q1.

Both of these tests do not allow there presence of structural change, so they are ill-equipped to evaluate the appropriateness of structural theories of unemployment. Rather, these tests will be useful in evaluating Hysteresis theories against Natural Rate of Unemployment theories directly, as Blanchard and Summers first explored. The purpose of this section and the one that follows will be to see whether or not modern testing techniques would have lead researches to different conclusions from those in the early 90's that found that natural rate of unemployment theories were ill-equipped to describe the dynamics of European unemployment rates. This section establishes a baseline of understanding by using long established tests that would have been used by researchers in the 90's. The following section then applies a modern technique with better properties to which we can compare these results.

2.2.1 The Augmented Dickey Fuller Test

The ADF test I utilize is conducted by first estimating the following regression:

$$\Delta u_t = c + \alpha u_{t-1} + \sum_{k=1}^{n} \rho_k \Delta u_{t-i} + \epsilon_t \tag{2.1}$$

u_t is the unemployment rate. The variable of interest in this regression is α. If $\alpha = 0$, this regression will be indicative of a unit root containing process. If $\alpha < 0$, the regression reduces to a stationary process. This allows the construction of a simple 1-sided hypothesis test:

H_0 - Unemployment is nonstationary: $\alpha = 0$

H_1 - Unemployment is stationary: $\alpha < 0$

For my choice of the number of first difference lags to include, I use the General to Specific (GS) selection method utilized by Campbell and Perron (1991), Hall (1994), and Ng and Perron (1995). Starting with a maximum possible value for n as $n_{max} = 13$, I check the test statistic of the coefficient of the last lag for significance. If the last lag is significant, I set $n = 13$. If not, I lower n by 1 and repeat the significance test on the new last lag. This process is repeated until the last lag is significant (or there are no lags left to test, in which case n is set equal to 0).

The key results of the ADF test are reported in Table 2.1. Using critical values of MacKinnon (1996), I find that the unit root null can be rejected at the 10% significance level for only 10 of the 30 countries observed. In addition, table 1 reports the half lives for each of the 30 estimated regressions, defined as $HL_i = \frac{ln(0.5)}{4ln(1+\alpha_i)}$ where i denotes a country. This measure is reported in years and is of value in helping us determine the degree of persistence of a shock to each individual country. The longer the half life, the greater the support for Hysteresis theories; the shorter the half life, the greater the support for natural rate of unemployment theories. The half lives vary from about three quarters (Estonia) to about twenty seven and a half years (Germany).

2.2.2 Dickey Fuller-Generalized Least Squares Test

The second linear unit root test I conduct is the DF-GLS test first proposed by Elliot, Rothernberg, and Stock (1996). This test is similar to the ADF test with the key difference being that the DF-GLS test requires first demeaning the time series via Generalized Least Squares before running the ADF regression. This test is appealing because it is known to have substantially improved power over the ADF test when an unknown mean or trend is present.

The test proceeds as follows. First, let $z_t = 1$ and define T as the number of observations in a given country's unemployment rate time series. Then, demean the time series by regressing $[u_1, (1-aL)u_2, \ldots, (1-aL)u_T]$ on $[z_1, (1-aL)z_2, \ldots, (1-aL)z_T]$ to yield a $\hat{\beta}_{GLS}$. In this regression, $a = 1 + \bar{c}/T$ and $\bar{c} = -7$. The demeaned unemployment rate is then calculated as $u_t^r = u_t - z_t\hat{\beta}_{GLS}$ This demeaned unemployment rate is then used to estimate an ADF regression.

$$\Delta u_t^r = \alpha u_{t-1}^r + \sum_{k=1}^{n} \rho_k \Delta u_{t-i}^r + \epsilon_t \tag{2.2}$$

This regression has the same variable of interest, α, as the ADF test. The hypothesis test of value is the same as the ADF test:

H_0 - Unemployment is nonstationary: $\alpha = 0$

H_1 - Unemployment is stationary: $\alpha < 0$

For the choice of the number of lags included, I use the Modified Akaike Information Criteria (MAIC), with the maximum number of lags allowed again set to 13. Critical values come from MacKinnon (1996). Key results of the DF-GLS test are reported in Table 2.2. Using this test, the unit root null can only be rejected at the 10% significance level for 7 of the 30 countries considered. Half lives and their 90% confidence intervals are again calculated and are largely significantly higher than those found by the ADF test. As a whole, this unit root test appears to find much more persistence in the shocks to the unemployment series of these 30 countries than does the first.

As with the ADF test, the DF-GLS test finds little evidence supporting Natural Rate of Unemployment theories. This is not surprising; these tests have been used to test for a unit root in unemployment series often and many papers have found that they tend to fail to reject the unit root null. But these two tests do establish a baseline. The key question of the next section is whether or not a modern test with better properties leads to the same conclusion.

2.3 Incorporating a Deterministic Time Trend Component

The two previous linear unit root tests provide a solid starting point in testing for the existence of unit roots in unemployment being that they are some of the oldest and most widely used unit root tests available. However, both tests have shortcomings that newer tests improve upon. This section is about utilizing a newer unit root test which has better size and power properties than the two baseline tests conducted in the previous section.

Consider now a new functional form[1]:

$$
\begin{aligned}
y_t &= \mu + \beta t + u_t && (2.3)\\
u_t &= \alpha u_{t-1} + A^*(L)\Delta u_{t-1} + e_t && (2.4)\\
\text{with} \quad A^*(L) &= \sum_{i=0}^{\infty} a_i^* L^i \quad \text{where} \quad a_i^* = -\sum_{j=i+1}^{\infty} a_j
\end{aligned}
$$

I use a technique from Perron and Yabu (2009a) in this section to conduct a unit root test while allowing for the existence a deterministic time trend.The key draw of using this Perron and Yabu approach is that their process allows me to "test for the slope of a trend function when it is a priori unknown whether the series is trend-stationary or contains an autoregressive unit root." This test was shown by Perron and Yabu to have better size and power properties than the linear DF-GLS test. Further, it was shown to have better properties than similar tests of Bunzel and Vogelsang (2005) and Harvey et al (2007), both of which have the same critical values under both the

[1]The primary difference between this functional form and the one used in the ADF test is that I've now included a deterministic time trend component. In testing for unit roots in unemployment, the conventional wisdom is to omit this trend component as we don't typically think of the unemployment rate as containing any long term trend (if it did, the unemployment rate must eventually reach either the 0% lower bound or 100% upper bound). However, this Perron and Yabu estimation is not built to function without a time trend. In future editions of this paper, I will amend the PY test so that it can be estimated without a trend. It is worth noting however that the deterministic time trend was estimated to be insignificant for 29 of the 30 countries. Further, inclusion of the trend will skew results towards finding stationarity when none is present. Given that I already find overwhelming support for nonstationarity, amending the test should only strengthen this results.

$I(0)$ and $I(1)$ cases. After using this process to estimate μ and β, I then remove the estimated intercept and trend from the time series and conduct a unit root test on the residuals.

Perron and Yabu propose the following 5 step process for estimating the coefficients μ and β. First, detrend the time series by OLS to obtain a series of residuals $\hat{u}_t$. Second, construct a truncated version of a Weighted Symmetric Least-Squares (WSLS) estimate of α in accordance with Roy and Fuller (2001). This requires first obtaining the WSLS estimate of α (call it $\hat{\alpha}_W$). This estimate is defined as $\hat{\alpha}_W = [\sum_{t=2}^{T-1} \hat{u}_t^2 + T^{-1} \sum_{t=1}^{T} \hat{u}_t^2]^{-1} \sum_{t-2}^{T} \hat{u}_t \hat{u}_{t-1}$. An estimate of the variance is defined as $\hat{\sigma}_W^2 = [\sum_{t=2}^{T-1} \hat{u}_t^2 + T^{-1} \sum_{t=1}^{T} \hat{u}_t^2]^{-1} (T-3)^{-1} \sum_{t-2}^{T} (\hat{u}_t - \hat{\alpha}_W \hat{u}_{t-1})^2$. The t-ratio associated with testing $\alpha = 1$ is then $\hat{\tau}_W = (\hat{\alpha}_W - 1)/\hat{\sigma}_W$.

Using these estimates, the third step is to next construct a bias-corrected estimate of α. This is done by first making the transformation $\hat{\alpha}_{TW} = \hat{\alpha}_W + C(\hat{\tau}_W)\hat{\sigma}_W$. Where

$$
\begin{aligned}
C(\hat{\tau}_W) &= -\hat{\tau}_W & if \quad \hat{\tau}_W > \tau_{pct} \\
C(\hat{\tau}_W) &= T^{-1}\hat{\tau}_W - 3[\hat{\tau}_W + K(\hat{\tau}_W + 5)]^{-1} & if \quad -5 < \hat{\tau}_W \leq \tau_{pct} \\
C(\hat{\tau}_W) &= T^{-1}\hat{\tau}_W - 3[\hat{\tau}_W]^{-1} & if \quad -(3T)^{1/2} < \hat{\tau}_W \leq -5 \\
C(\hat{\tau}_W) &= 0 & if \quad \hat{\tau}_W \leq -(3T)^{1/2}
\end{aligned}
$$

"Where $K = [3T - \tau_{pct}^2(I_p + T)][\tau_{pct}(5 + \tau_{pct})(I_p + T)]^{-1}$, where I_p is the integral part of $(p+1)/2$ and τ_{pct} is the percentile of the limiting distribution of $\hat{\tau}_W$ when $\alpha = 1$. If $\tau_{pct} = -1.96$, the median of the distribution of $\hat{\tau}_W$ when $\alpha = 1$, then $\hat{\alpha}_{TW}$ is approximately median unbiased, in the sense that it is nearly unbiased when $\alpha < 1$

and has a median of 1 when $\alpha = 1$" (Perron and Yabu). When using this value, I shall denote the corrected estimate by $\hat{\alpha}_{MU}$.

The fourth step is to apply the following truncation: $\hat{\alpha}_{MS} = \hat{\alpha}_{MU}$ if $|\hat{\alpha}_{MU} - 1| > dT^{-1/2}$ and 1 otherwise. Fifth and finally, I apply a GLS procedure using $\hat{\alpha}_{MS}$ to obtain an estimate of the trend parameter and construct the standard t-statistic, which we shall denote by $t_{\beta}^{FS}(MU)$

This process is preformed for each of the 30 countries considered. The results of the estimation are displayed in Table 2.3. Of critical importance, note that the estimated β for every country but Austria is found to be statistically insignificant. As conventional wisdom would suggest, there seems to be no long run time trend of statistical significance in the unemployment rate of the vast majority of the 30 countries.

Now that we have a good estimate of the trend and intercept, I test for the presence of a unit root around these estimates. For any given country, denote the estimated intercept and trend coefficients as $\hat{\mu}$ and $\hat{\beta}$. I then make the transformation $\tilde{u}_t = u_t - \hat{\mu} - \hat{\beta}t$ and estimate the following equation using least squares

$$\tilde{u}_t = \alpha\tilde{u}_{t-1} + \sum_{k=1}^{n} \rho_k \Delta\tilde{u}_{t-i} + \epsilon_t$$

The number of lags is chosen using the Modified Akaike Information Criteria. The key variable of this unit root test is again the α, with the key hypothesis test being

H_0 - Unemployment is nonstationary: $\alpha = 0$

H_1 - Unemployment is stationary around the estimated trend: $\alpha < 0$

Results of this estimation are contained in Table 2.4. If we look to this table and we compare it to the DF-GLS test results, we see that most of the countries which rejected the unit root test in the DF-GLS test also rejected the unit root null in Perron and Yabu's new test. In all, 9 of the 30 countries reject the unit root null in favor of stationarity. However, the fact remains that the majority of countries fail to reject the unit root null. When not allowing for the possibility of structural change, it appears that very little evidence exists supporting natural rate of unemployment theories.

This conclusion omits the possibility of structural models best describing the evolution of the unemployment rates however. Structural theories of unemployment hold that most shocks to unemployment are temporary, but only "a few" are permanent. The remainder of this essay will be conducting unit root tests which allow for such a dynamic to be found.

2.4 Markov Switching Augmented Dickey Fuller Test

The first problem in looking for evidence of structural theories is in how to define the term "few". I divide my remaining unit root testing into two sections so that I can allow for a "few permanent shifts" to occur in a couple of different ways.

In this section, I employ a version of a Markov Switching Augmented Dickey Fuller (MS-ADF) test for each country. This test allows a "high-end" estimate of what it means for there to be a few shocks that change the long run market dynamics. The test will be run specifically so that two distinct states-of-the-world exist for each country; one with a high-unemployment long run equilibrium level of unemployment and the other with a low-unemployment long run equilibrium level of unemployment. The transition between the two states of the world are dictated by a first-order Markov process so as to probabilistically determine which of the two states of the world are most likely. This allows the number of permanent shocks to the unemployment rate to be relatively high as we could see significant switching between states of the world. But the types of switching is still constrained to be small, so that the world is always described by one of only two distinct states. This dynamic could well describe markets which, for example, fluctuate between periods of recessions and expansions, each of which being characterized by distinct long run equilibrium levels of unemployment if we were to remain in that state in perpetuity.

The regime switching unit root test was first proposed by Hall, Psaradakis, and Sola (1999) to test for periodically collapsing bubbles and has since been used in a wide range of papers (Holmes 2008; Chen 2008; Camacho 2011 to name a few). The MS-ADF test takes the same basic functional form as the linear ADF test, but it incorporates the existence of multiple regimes, or states of the world, in order to capture dynamics that potentially change across time.

The MS-ADF test I use takes the form

$$\Delta u_t = c_{s_t} + \alpha u_{t-1} + \sum_{k=1}^{n} \rho_k \Delta u_{t-i} + \epsilon_t \quad \epsilon_t \sim NID(0, \Sigma) \tag{2.5}$$

where u_t is the unemployment rate, $s_t \in \{1, 2\}$ is the unobservable regime, c_{s_t}, α, and ρ_k are parameters to be estimated, and Σ is the error variance.

I assume that the state variable follows a two regime Markov process. In accordance with Hamilton (1994), a first order Markov switching process dictates the evolution of the unobserved state variable.

$$\begin{aligned} P[s_t = 1 | s_{t-1} = 1] &= p \\ P[s_t = 2 | s_{t-1} = 1] &= 1 - p \\ P[s_t = 2 | s_{t-1} = 2] &= q \\ P[s_t = 1 | s_{t-1} = 2] &= 1 - q \\ 0 < p < 1 \quad 0 < q &< 1 \end{aligned} \tag{2.6}$$

As with the linear ADF test, the key variable of interest in determining stationarity of the series is α. Stationarity can be tested with the following one-sided hypothesis test:

H_0 - Unemployment is nonstationary: $\alpha = 0$

H_1 - Unemployment is stationary: $\alpha < 0$

I estimate equation (6) using maximum likelihood estimation based on the Expectations Maximization (EM) algorithm of Hamilton (1994) and Krolzig (1997). Each iteration of the EM algorithm consists of two steps. First, the expectations step requires estimation of the unobserved states using their smoothed probabilities. Using the Baum-Hamilton-Lee-Kim filter and smoother, conditional probabilities are calculated from the estimated parameter vector of the last maximization step. The maximization step then requires updated estimation of the parameter vector as a solution of the first order conditions, where conditional regime probabilities are replaced with the smoothed probabilities of the last expectations step.

The number of first difference lags (n) was chosen to be equal to the linear ADF test so that I may run a Likelihood Ratio test on the fit of the linear regression compared to the regime switching one. The goal of this test is to determine whether or not the regime switching test provides a statistically significantly improved fit over the linear unit root test. In this test, the null hypothesis is that the linear model provides as good a fit as the non-linear one, so rejection of the null indicates a statistically significantly superior fit for the regime-switching regression. Of the 30 countries considered, 7 countries failed to reject the null of the Likelihood Ratio test (Germany, Hungary, Italy, Netherlands, Norway, Poland, and Slovenia). All of seven of these countries failed to reject the unit root null in the linear test, providing no sufficient evidence for stationarity.

For the remaining 23 countries, Table 2.5 reports the coefficient of the first lag of the level of unemployment (α), the estimated probabilities of remaining in each regime in period t when already in that regime in period $t-1$ (p for regime 1, q for regime 2), the estimated intercept and mean reverting level of unemployment for each regime (c_i and μ_i, respectively), and the calculated half-life of a shock to each country's unemployment rate. As with previous tests, finding α significantly less than 0 indicates rejection of the null of persistence of shocks to unemployment in favor of stationarity. Further, if the given time series is stationary, we can then calculate the level of unemployment that the process is reverting to in each regime as $\mu_i = \frac{c_i}{\alpha}$ (where $i \in \{1, 2\}$ indicates the potential regimes).

Running this test, I now reject the null of a unit root in favor of stationarity with breaks for 15 of the 30 countries at the 10% significance level. This portion is slightly lower than the findings of Papell, Murray and Ghiblawi, who reject their unit root null for 10 of 16 countries considered, though it's not far off.

In the 15 countries in which I reject the unit root null, shocks to unemployment are estimated to dissipate over time and the process will revert to some "mean" level of unemployment. The only thing this test does differently than the original ADF test that I ran is that it allows the economy's mean unemployment rate to fluctuate between 2 different levels. Thus, for these 15 countries rejecting the unit root null, a structuralist theory of unemployment is likely to be the most fitting type.

Figure 2.2 provides a grid of the estimated probability of being in the low-unemployment regime for each of the 15 countries that rejected the unit root null. Figure 2.3 provides a similar grid of the time dependent probability of being in the low unemployment regime for the 8 remaining countries that rejected the Likelihood Ratio test null.

Note the regime probabilities for the United States in Figure 2.2. The MS-ADF estimation concludes a high probability of the low unemployment regime occurring during US expansionary periods. In this regime, the unemployment rate that the series is reverting to in the long run is estimated to be 4.89%. We also see a spike in the probability of being in the high unemployment regime during during most recessionary periods, and in this regime the mean-reverting level of the unemployment rate is 19.79%. If we were to maintain recessionary dynamics into the long run, this is a good estimate for the level to which the unemployment rate may approach. According to this test, the United States' mean-reverting unemployment rate switches between low-unemployment and high-up employment regimes which closely coincide with expansions and recessions. And as can be seen in Figure 2.1, many countries appear to follow a similar dynamic.

2.5 Unit Root Test Allowing Structural Breaks

The objective of this final section is to test for stationarity when there are a limited number of structural breaks. Rather than modeling the evolution of unemployment rate as probabilistically switching between distinct regimes, I now will assume that there are (at most) 2 specific points in time in which shocks to the economy might possibly alter the long run trend path of the unemployment rate. This will cover the low-end estimate of what it means to allow a "few" shocks to have permanent effects. In this framework, I want to again consider the possibility that a deterministic time trend affects the unemployment rate, though now that trend will be allowed to change its behavior across time.

Conducting such a unit root test will be done in two parts. In Section 2.5.1, I will utilize a pre-test to determine the number of structural breaks that should be included. The pretest is a form of Perron and Yabu's earlier test, only now also allowing for multiple structural breaks over time. In Section 2.5.2, I will apply the number of breaks chosen from Perron and Yabu's pretest in a unit root testing procedure from either Kim and Perron (2009) or Silvestre, Kim, and Perron (2009), both of which are valid with structural breaks under both the null and alternative hypothesis. The only difference between the two tests is that the latter allows for a larger number of breaks.

2.5.1 The Pretest

The pretest comes in two parts. Section 2.5.1.1 will outline the Perron and Yabu (2009) procedure to test for a single structural break in the data. Section 2.5.1.2 will apply a sequential procedure from Kejriwal and Perron (2010) to determine the appropriate number of breaks.

2.5.1.1 Perron and Yabu Procedure

Consider a new functional form:

$$\begin{aligned} y_t &= x_t'\Psi + u_t \\ u_t &= \alpha u_{t-1} + A^*(L)\Delta u_{t-1} + e_t \\ with \quad A^*(L) &= \sum_{i=0}^{\infty} a_i^* L^i \quad where \quad a_i^* = -\sum_{j=i+1}^{\infty} a_j \end{aligned}$$

Perron and Yabu discuss 3 separate models, each of which hold different values for x_t' and Ψ.

Model 1 - Structural Change in Intercept: $x_t = (1, DU_t, t)'$ and $\Psi = (\mu_1, \mu_2, \beta_1)'$, where $DU_t = 1$ for $t > T_1$ (T_1 is the timing of the break). This allows for a one-time change in the intercept

Model 2 - Structural Change in Slope: $x_t = (1, t, DT_t)'$ and $\Psi = (\mu_1, \beta_1 \beta_2)'$, where $DT_t = 1(t - T_1)$ for $t > T_1$. This allows for a one-time change in the trend

Model 3 - Structural Change in Both Intercept and Slope: $x_t = (1, DU_t, t, DT_t)'$ and $\Psi = (\mu_1, \mu_2, \beta_1, \beta_2)'$. This allows for a one-time change in both the intercept and the trend

I estimate models 2 and 3 for all 21 countries in my sample. These two models are chosen because they both allow a break in the time trend of the series. Those estimates are obtained through the following 5 step procedure. First, detrend the data by OLS to obtain residuals $\hat{u}_t$. Then run an Augmented Dickey Fuller test (with no intercept or trend) on the residuals. The number of lags is selected using the BIC with the number of lags chosen constrained to the range $[0, 12(T/100)^{1/4}]$. From this estimation, the key coefficient to be estimated is $\bar{\alpha}$ (where $\bar{\alpha}$ is the coefficient associated with the lagged level of the unemployment rate in the ADF regression).

This estimated $\bar{\alpha}$ is known to be biased downward, so the third step requires construction of a bias-corrected version of $\bar{\alpha}$, denoted $\bar{\alpha}_M$, in accordance with Roy and Fuller (2001), before applying a truncation in accordance with Perron and Yabu. The bias-corrected estimate is given by $\bar{\alpha}_M = \hat{\alpha} + C(\hat{\tau})\bar{\sigma_\alpha}$ where

$$
\begin{aligned}
C(\hat{\tau}) &= -\hat{\tau} && if \quad \hat{\tau} > \tau_{pct} \\
C(\hat{\tau}) &= I_p T^{-1}\hat{\tau} - (1+r)[\hat{\tau} + c_2(\hat{\tau}+10)]^{-1} && if \quad -10 < \hat{\tau} \leq \tau_{pct} \\
C(\hat{\tau}) &= I_p T^{-1}\hat{\tau} - (1+r)\hat{\tau}^{-1} && if \quad -c_1^{1/2} < \hat{\tau} \leq -10 \\
C(\hat{\tau}) &= 0 && if \quad \hat{\tau} \leq -c_1^{1/2}
\end{aligned}
$$

where $c_1 = (1+r)T$ (with r being the number of parameters estimated in the trend function), $c_2 = [(1+r)T - \tau_{pct}^2(I_p + T)][\tau_{pct}(a + \tau_{pct})(I_p + T)]^{-1}$ (with p being the order of the AR process considered for the noise component), and τ_{pct} is a percentile of the limit distribution of $\hat{\tau}$ when $\alpha = 1$. The critical values for this term are given by Perron (1989). Per Perron and Yabu's recommendation, I use $\tau_{0.99}$ due to the break date being unknown. Completing the third step then requires applying the truncation $\bar{\alpha}_{MS} = \bar{\alpha}_M$ if $|\bar{\alpha}_M - 1| > T^{-1/2}$ and $\bar{\alpha}_{MS} = 1$ otherwise.

The fourth step is to apply Perron and Yabu's quasi-GLS procedure with $\bar{\alpha}_{MS}$ to obtain an estimate of Ψ and construct Wald statistics (denote it W_{RQF} or W^*_{RQF}, depending on the model and the value of $\bar{\alpha}_{MS}$). The quasi-GLS procedure first requires running the regression

$$(1-\tilde{\alpha}_{MS}L)y_t = (1-\tilde{\alpha}_{MS}L)x_t^{'}\Psi + (1-\tilde{\alpha}_{MS}L)u_t \tag{2.7}$$

for $t = 2, \ldots,$ T and $y_1 = x_t^{'}\Psi + u_1$. Denote the resulting estimates as $\tilde{\Psi}$. The specific form of the Wald test then depends on whether the errors are $I(0)$ of $I(1)$.

If the errors are $I(0)$, the robust quasi-FGLS Wald statistic is $W_{RQF} = [R(\tilde{\Psi} - \Psi)]^{'}[\hat{h}_v R(X'X)^{-1}R']^{-1}[R(\tilde{\Psi} - \Psi)]$ with $X = \{x_t^{\tilde{\alpha}_{MS}}\}, x_t^{\tilde{\alpha}_{MS}} = (1-\tilde{\alpha}_{MS}L)x_t$ for $t = 2, \ldots, T$ and $x_1^{\tilde{\alpha}_{MS}} = x_1$. $\hat{h}_v$ is an AR spectral density estimate at frequency zero. This spectral density estimate is found by utilizing the residuals of equation (6), $(1-\tilde{\alpha}_S L)u_t$. To obtain a consistent estimate, approximate the following regression:

$$y_t - \tilde{\alpha}_S y_{t-1} = x_t^{'}\Psi^* + \sum_{i=1}^{k} \rho_i \Delta y_{t-i} + e_{tk}$$

where $\hat{e}_{tk}$ are the corresponding OLS residuals. The estimate of the spectral density is then $\hat{\sigma}^2 = \hat{h}_v = (T-k)^{-1}\sum_{t=k+1}^{T} \hat{e}_{tk}$

If the errors are $I(1)$, we need to consider each of the two models separately. For model 2, the Wald statistic is the same as the $I(0)$ models, but the spectral density estimate will be different. To get that estimate, we first must estimate the regression

$$\hat{v}_t = \sum_{i=1}^{k} \psi_i \hat{v}_{t-i} + e_{tk}$$

Denote the estimate by $\hat{\psi}(L) = (1-\hat{\psi}_1 L - \ldots - \hat{\psi}_k L^k)$ and $\hat{\sigma}^2_{ek} = (T-k)^{-1}\sum_{t=k+1}^{T} \hat{e}^2_{tk}$. Then, $\hat{h}_v = \hat{\sigma}^2_{ek}/\hat{\phi}(1)^2$.

For model 3, a change in intercept is also involved making things slightly more complex. For this model, the Wald statistic to use is
$W^*_{RQF} = [R(\tilde{\Psi}^* - \Psi)]'[\hat{h}_v R(X'X)^{-1}R']^{-1}[R(\tilde{\Psi}^* - \Psi)]$ where $\tilde{\Psi}^* = (\tilde{\mu}_0, \mu_1^*, \tilde{\beta}_0, \tilde{\beta}_1)'$.

With the Wald statistics calculated, the fifth and final step when we have a break in the data at an unknown time is to evaluate the test statistic for each break date candidate. The Exp function provided by Perron and Yabu is a function of the Wald statistic found in the previous step, which is evaluated to determine the optimal location of the structural break. The Exp function is defined as Exp-$W_{RQF} = log[T^{-1}\sum_{\Lambda} exp(\frac{1}{2}W_{RQF}(\lambda_1'))]$ where $\Lambda = \{\lambda_1'; \epsilon \leq \lambda_1' \leq 1 - \epsilon\}$ for some $\epsilon > 0$. I choose $\epsilon = 0.1$. The estimated break date corresponding to the highest Exp-W_{RQF} is the one I use.

Table 2.6 provides key outputs by country for each of these 2 models. For model 2, $y_t = a_0 + b_0 t + b_1 DT + u_t$ where $DT = 1(t > TB)(t - TB)$ and TB is the estimate of the break date. For model 3, $y_t = a_0 + a_1 DU + b_0 t + b_1 DT + u_t$ where $DU = 1(t > TB)$.

Looking at the Model 2 results on Table 2.6, we see that only Hungary, Italy, Japan, Portugal and Russia have a Wald statistic sufficiently high to reject the null hypothesis that the trend function is stable in favor of the alternative hypothesis of there being a shift in the trend only. However, Table 2.6 also displays vastly different results for Model 3. For this model, two thirds of the 30 countries can reject the stable trend and intercept null in favor of a shifting one. These results indicate that it is primarily changes in the intercept term that will provide the regression with an increased fit; the trend function seems to be less important, as conventional wisdom suggests.

2.5.1.2 The Sequential Procedure

The sequential procedure portion of the pretest is conducted in a specific to general sequential manner in accordance with Kejriwal and Perron. The procedure first performs the Perron and Yabu estimation with a single structural break described in section 2.5.1.1. As per the results of the Perron and Yabu tests reported in Table 6, I decide to only focus on model 3 from this point forward. As Table 2.6 shows, 10 countries fail to reject the 0-break null at the 90% confidence level. Each of these countries will be tested for a unit root without allowing for structural change.

For the remainder of the countries in which the Perron and Yabu test does find a significantly large Wald value to reject the null, the data series is then split at the estimated break date. A 15% trimming is applied on each end of each sub-sample, and the Perron and Yabu test is rerun on each sub-sample of trimmed data. The maximum Wald value from each of the two sub-sampled tests is then compared to the critical values in Kejriwal and Perron. These results are provided in Table 2.7.

As can be seen in Table 2.7, of the countries finding a first significant break, a second significant break was found for 16 of the 20 countries. Only Finland, Denmark, South Korea and Mexico failed to find a second significant break after the first had been established. For these four countries, I will conduct a unit root test allowing a single break in the data at an unknown time. For the remaining 16 countries, I conduct a unit root test designed for two structural breaks at unknown times.

2.5.2 The Unit Root Tests

The testing procedure described in section 2.5.1 leaves each country in one of three possible scenarios. For the 10 countries that failed to find a single significant break in the data, an Augmented Dickey Fuller test is run in section 2.5.2.1. For the 4 countries found to have a single significant break, a unit root test from Kim and Perron is conducted in section 2.5.2.2. For the remaining 16 countries found to have two breaks, a unit root test from Silvestre, Kim, and Perron is conducted in section 2.5.2.3.

2.5.2.1 The Case with No Breaks

The 10 countries who failed find even a single significant break are Austria, Belgium, Chile, Estonia, Netherlands, New Zealand, Norway, Slovenia, Spain and the UK. The appropriate unit root test to run on these four countries is thus the Augmented Dickey Fuller test, the results of which have already been provided in Table 2.1. Referring to this table, note that only Belgium, Chile and Estonia can reject the unit root null in favor of stationarity, providing evidence that these three countries might best be modeled by natural rate of unemployment theories.

2.5.2.2 The Case with 1 Break

The next task is to conduct a unit root test for all countries finding exactly 1 break date to be appropriate. To accomplish this, I use a procedure proposed by Kim and Perron. This procedure is of particular value because it allows for a break to occur under both the null and alternative hypotheses, a recent development in unit root testing which offers a substantial improvement over the long-used Zivot and Andrews unit root tests. When there is a break present, Kim and Perron's test is designed to have the same limiting distribution regardless of whether the break date is known or unknown, allowing increased power while maintaining correct size.

Consider the following equation:

$$y_t = \mu + \beta t + \mu_b DU_t + \beta_b B_t + u_t$$

where $DU_t = B_t = 0$ if $t < T$ (where T corresponds to the break date) and $DU_t = 1$, $B_t = t - T$ if $t > T$. The noise series $\{u_t\}$ is such that $A(L)u_t = \epsilon_t$ where $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$ and $A(L)$ is a lagged polynomial of order p.

Kim and Perron's unit root test proceeds as follows. First, detrend the unemployment series using the deterministic components to yield a residual series $\tilde{u}_t$. Then estimate the following regression for every break date candidate.

$$y_t = \hat{\mu}_b DU_t + \hat{\beta}_b B_t + \hat{u}_t$$

The estimated break date is obtained by minimizing the Sum of Squared Residuals over all possible break dates. Once the break date is chosen, the full model can be estimated as

$$\tilde{u}_t = \hat{\alpha}\tilde{u}_{t-1} + \mu + \beta t + \mu_b DU_t + \beta_b B_t + \sum_{j=1}^{k} \hat{d}_j \Delta\tilde{u}_{t-j}$$

where the number of lags (k) is chosen using the Bayes Information Criteria. The key results of running this test on the unemployment rates of Denmark, Finland, South Korea and Mexico are displayed in Table 2.8. As can be seen, Finland, Korea and Mexico reject the unit root null at the 90% confidence level while only Denmark fails to reject it. This provides support for the claim that structuralist theories of unemployment are appropriate for Finland, South Korea and Mexico.

2.5.2.3 The Case with 2 Breaks

The final task of this paper is to generalize the procedure of section 2.5.2.2 to accommodate multiple structural breaks. This process is laid out in a paper by Silvestre, Kim and Perron. Consider the following equations:

$$\begin{aligned} y_t &= d_t + u_t \\ u_t &= A(L)u_t + v_t \end{aligned}$$

where the deterministic component is $d_t = z_t^{'}(T_0)\psi_0 + z_t^{'}(T_1)\psi_1 + z_t^{'}(T_2)\psi_2 = z_t^{'}(\lambda^0)\psi$. T_1 and T_2 correspond to the two break dates. This general notation allows 3 models to be considered by Silvestre, Kim and Perron. I am only interested in estimating their Model 2 however (this model corresponds with the pretest preformed in section

2.5.1). I thus define $z_t(T_0) = (1, t)$ and $z_t(T_j) = (DU_t(T_j), DT_t(T_j))$ for $j = \{1, 2\}$. $\psi_j = (\mu_j, \beta_j)'$. More precisely, the full deterministic component of this model is defined as $d_t = \mu_0 + \beta_0 t + \mu_1 DU_t(T_1) + \beta_1 DT_t(T_1) + \mu_2 DU_t(T_2) + \beta_2 DT_t(T_2)$, where $DU_t(T_j) = 1$ for $t > T_j$ and 0 otherwise. $DT_t(T_j) = (t - T_j)$ for $t > T_j$ and 0 otherwise.

The first step in the process is to apply a GLS detrending technique. Unit root statistics are going to be based on the use of transformed data, $u_t^{\bar{\alpha}} = (u_1, (1 - \bar{\alpha}L)y_t)$ and $z_t^{\bar{\alpha}} = z_1(\lambda^0)$, $(1 - \bar{\alpha}L)z_t(\lambda^0))$ with $\bar{\alpha} = 1 + \bar{c}/T$. $\bar{c}$ is the non centrality parameter.

Using this transformed data, the deterministic parameters (ψ) can be estimated by minimizing the objective function:

$$S^*(\psi, \bar{\alpha}, \lambda^0) = \sum_{t=1}^{T} (y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}}(\lambda^0))^2.$$

The minimum of this function can be denoted $S(\bar{\alpha}, \lambda^0)$

A key point in this estimation is the non-centrality parameter, the definition of which is based on the point optimal statistic used in Elliott, Rotherberg, and Stock (1996). It's based on a test in which the null is that the true value of α is equal to 1 verse the alternative that the true α is equal to $\bar{\alpha}$. The point optimal statistic is defined as

$$P_T^{GLS}(c, \bar{c}, \lambda^0) = \{S(\bar{\alpha}, \lambda^0) - \bar{\alpha}S(1, \lambda^0)\}/s^2(\lambda^0)$$

where $S(\bar{\alpha}, \lambda^0)$ and $S(1, \lambda^0)$ are the sum of squared residuals from a GLS regression with $\alpha = \bar{\alpha}$ and $\alpha = 1$ respectively. $s^2(\lambda^0)$ is an estimate of the spectral density at frequency 0 of v_t. In accordance with Ng and Perron (2001) and Perron and Ng (1998), the autoregressive estimate is defined by

$$s^2(\lambda^0) = s_{ek}^2/(1 - \sum_{j=1}^{k} \hat{b}_j)^2$$

where $s_{ek}^2 = (T-k)^{-1} \sum_{t=k+1}^{T} \hat{e}_{t,k}^2$ and $\{\hat{b}_j = \hat{e}_t, k\}$ are obtained from an Ordinary Least Squares regression

$$\Delta\tilde{y}_t = b_0\tilde{y}_{t-1} + \sum_{j=1}^{k} b_j\Delta\tilde{y}_{t-j} + e_{t,k}$$

with $\tilde{y}_t = y_t - \hat{\Psi}' z_t(\lambda^0)$ where $\hat{\Psi}$ minimizes (14). The number of lags, k, are chosen by the Bayes' Information Criteria.

Because the test statistic P_T^{GLS} is non-standard, Silvestre, Kim and Perron simulate critical values. Table 2.9 provides results of their unit root test when imposing 2 structural breaks on each of the 16 remaining countries. Of the 16 countries tested, 7 of them now reject the unit root null.

In all, this section concludes that natural rate of unemployment theories are appropriate for 3 of 30 countries, while structural theories of unemployment are better suited 10 of 30 countries. This is a sharp divergence from the literature. Whereas Papell et. al rejected the unit root null with structural change for over 60% of the countries in their sample, this modern unit root testing technique puts that percentage at only one-third. This is significant in that it indicates that hysteresis models of unemployment may be a more accurate description of reality than has been previously concluded.

2.6 Conclusions

This paper provides an updated answer to an old question: how should we model the unemployment rate? While this question has been studied through the use of unit root testing by Papell, Murray, and Ghiblawi in the past, I update that paper using modern techniques, newer data, and more countries.

I conduct two separate unit root testing procedures to allow for a "high-end" and "low-end" estimate of the number of permanent breaks that can be considered consistent with structuralist theories of unemployment. Using a Markov-Switching Augmented Dickey Fuller test to serve as the high-end estimate, I conclude that 15 of the 30 countries should reject the unit root null, lending credibility to the claim that these countries could be better modeled by structuralist theories of unemployment than by hysteresis theories. What's more, the half-life of the shocks to the unemployment rate fell significantly when allowing for this switching behavior, as opposed to the linear unit root tests I conducted as a baseline. In not allowing the dynamics of the world to change, linear unit root tests find a much larger persistence in unemployment shocks than what I found in the MS-ADF test. The results of this "high-end" estimate are nearly consistent with previous literature.

Then, in order to cover the "low-end" estimate of the number of breaks that structuralist theories allow, I utilize a testing procedure laid out in Silvestre, Kim, and Perron that yields decidedly different results. This procedure utilizes a pre-test from Perron and Yabu to determine the number of structural breaks to include in the data. If no break is found, a linear ADF test is conducted. If one break is found, a unit root test from Kim and Perron is conducted. And if two breaks are found, the unit root testing procedure detailed in Silvestre, Kim, and Perron is conducted. This multi-step process resulted in the rejection of the unit root null with structural breaks for only 10 of the 30 countries, while 17 of the 30 countries failed to reject the unit root null (the final 3 countries rejected the null but with no structural breaks). These results are a sharp divergence from the literature. Whereas previous studies have largely concluded hysteresis theories to be less compelling than those of the structuralist school, this result casts doubt on that conclusion. The results from this test indicate that structural models of unemployment may be far less appropriate than has previously been concluded.

2.6.1 Results Across Tests

Lastly, we can alternatively view the results on a country by country basis across the 5 unit root tests. If the same conclusion is reached for a particular country across a class of tests, this can be taken as particularly strong evidence that the unemployment rate of that country has evolved in a manner consistent with that conclusion.

For example, when considering the class of linear unit root tests, South Korea, Mexico, and the US each reject the unit root null for all three tests. This is particularly strong evidence that these countries would have their unemployment rates be better modeled by natural rate of unemployment theories than by hysteresis theories.

When considering the class of unit root tests allowing structural change, 5 countries reject the unit root null for both tests (Australia, Canada, South Korea, Mexico, US) whereas 9 countries fail to reject the unit root null for both tests (Austria, Czech Republic, Hungary, Japan, Netherlands, Norway, Poland, Slovenia, Spain). This finding is roughly consistent with the conclusion of Blanchard and Summers (1987); they found that it was European economies that seemed most compatible with the idea of unemployment hysteresis. The two groups above are largely clustered on these grounds. It is only European economies and Japan that consistently fail to reject the unit root null, while it is the US and some of their major trading partners that consistently find evidence of stationarity. A natural point of extension for future research will be to look for common factors for each of the countries in each of these two categories so as to uncover evidence on the potential causes of hysteresis.

2.7 References

Blanchard, O., and L. Summers, "Hysteresis and the European Unemployment Problem," in S. Fischer (Ed.), NBER Macroeconomics Annual (Cambridge, MA: MIT Press, 1986), 15-78.

Bunzel, H., Vogelsang, T.J., 2005. "Powerful trend function tests that are robust to strong serial correlation with an application to the Prebish-Singer hypothesis." Journal of Business and Economic Statistics 23, 381-394.

Camacho, M., 2011. "Markov-switching models and the unit root hypothesis in real US GDP." Economics Letters 112, 161-164.

Campbell, J., and P. Perron, "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots," in O. Blanchard and S. Fischer (Eds.), NBER Macroeconomic Annual (Cambridge, MA: MIT Press, 1991), 141-201.

Chen, S.-W., 2008. "Non-stationarity and non-linearity in the stock prices: evidence from the OECD countries." Economics Bulletin 3, 1-11.

Dickey, D.A., Fuller, W.A., 1979. "Distribution of the estimators for autoregressive time series with a unit root." Journal of the American Statistical Association 74, 427-431.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. "Efficient tests for an autoregressive unit root." Econometrica 64, 813-836.

Friedman, M., "The Role of Monetary Policy," American Economic Review 58 (1968), 1-17.

Fuller, W.A., 1996. Introduction to Statistical Time Series, 2nd ed.. Wiley, New York, NY.

Hall, A., "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection," Journal of Business and Economic Statistics 12 (1994), 461-70.

Hall, S.G., Psaradakis, Z., Sola, M., 1999. "Detecting periodically collapsing bubbles: a Markov-switching unit root test." Journal of Applied Econometrics 14, 143-154.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Harvey, D.I., Leybourne, S.J., Taylor, A.M.R., 2007. "A simple, robust and powerful test of the trend hypothesis." Journal of Econometrics 141, 1302-1330.

Holmes, M.J., 2008. "Real exchange stationarity in Latin America and relative purchasing power parity: regime switching approach." Open Economics Review 19, 261-275.

Kejriwal, M., and Perron, P. (2010), "A Sequential Procedure to Determine the Number of Breaks in Trend with an Integrated or Stationary Noise Component." Journal of Time Series Analysis, 31, 305-328

Kim, D., and Perron, P. (2009), "Unit Root Tests Allowing for a Break in the Trend Function at an Unknown Time Under Both the Null and Alternative Hypothesis." Journal of Econometrics, 148, 1-13

Krolzig, H., 1997. "Markov-Switching Vector Autoregressions Modeling, Statistical Inference and Application to Business Cycle Analysis." Springer, Berlin.

MacKinnon, James G, 1996. "Numerical Distribution Functions for Unit Root and Cointegration Tests," Journal of Applied Econometrics, John Wiley and Sons, Ltd., vol. 11(6), pages 601-18, Nov.-Dec.

Ng, S. and Perron, P. (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," Econometrics 69, 1519-1554.

Papell, D., Murray C., and Ghiblawi, H. (2000), "The Structure of Unemployment," The Review os economics and Statistics 82(2), 309-315.

Perron, P. (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," Econometrica, 57, 1361-1401; Corr., 61 (1993), 248249.

Perron, P. and Ng, S. (1998), "An Autoregressive Spectral Density Estimator at Frequency Zero for Nonstationary Tests," Econometric Theory 14, 560-603.

Perron, P., and Yabu, T. (2009), "Estimating Deterministic Trends With an Integrated or Stationary Noise Component," Journal of Econometrics, 151, 56-69.

Perron, P., and Yabu, T. (2009), "Testing for Shifts in Trend with an Integrated or Stationary Noise Component," Journal of Business and Economic Statistics 27 (2009), 369-396.

Perron, P., and Qu, Z. (2007), "A simple modification to improve the finite sample properties of Ng and Perron's unit root tests," Economic Letters 94, 12-19.

Phelps, E., "Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time," Economica 34 (1967), 254-281.

Phelps, E., "Money Wage Dynamics and Labor Market Equilibrium," Journal of Political Economy 76 (1968), 678-711.

Phelps, E., "Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest, and Assets" Cambridge, MA: Harvard University Press, (1994).

Ng, S., and P. Perron, "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag," Journal of the American Statistical Association 90 (1995), 268-281.

Roy, A., and Fuller, W.A., 2001. "Estimation for autoregressive processes with a root near one." Journal of Business and Economic Statistics 19, 482-493.

Roy, A., Falk, B., and Fuller, W. A. (2004), "Testing for Trend in the Presence of Autoregressive Error," Journal of the American Statistical Association, 99, 1082-1091.

Silvestre, J., Kim, D., and Perron, P. (2009), "GLS-Based Unit Root Tests with Multiple Breaks Under Both the Null and Alternative Hypotheses." Econometric Theory, 25, 1754-1792.

Zivot, E., and Andrews, D.W.K. (1992), "Further Evidence on the great crash, the oil price shock and the unit root hypothesis." Journal of Business and Economic Statistics, 10, 251-270.

2.8 Tables and Figures

Figure 2.1: *Unemployment Series*

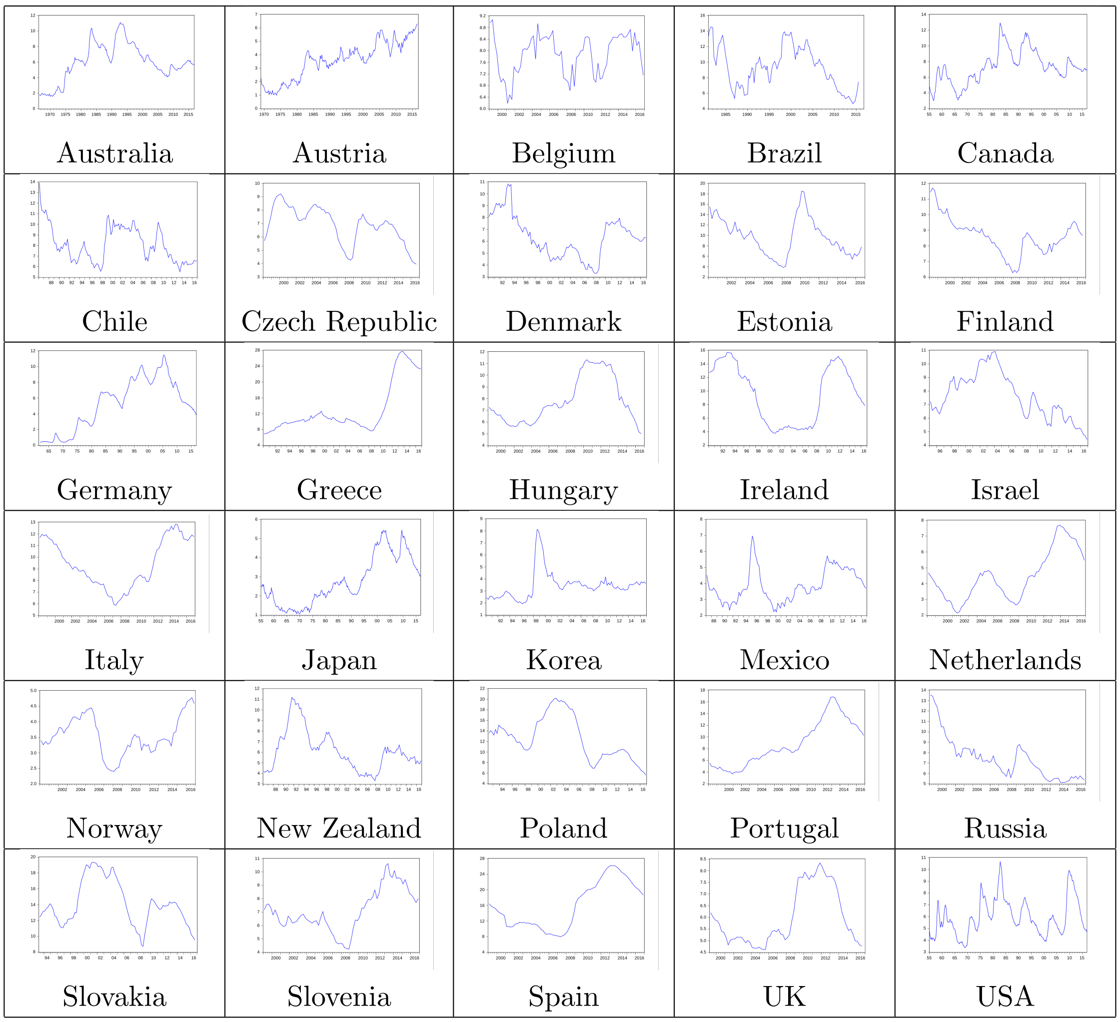

Figure 2.2: *Estimated Probability of the Low Unemployment Regime*

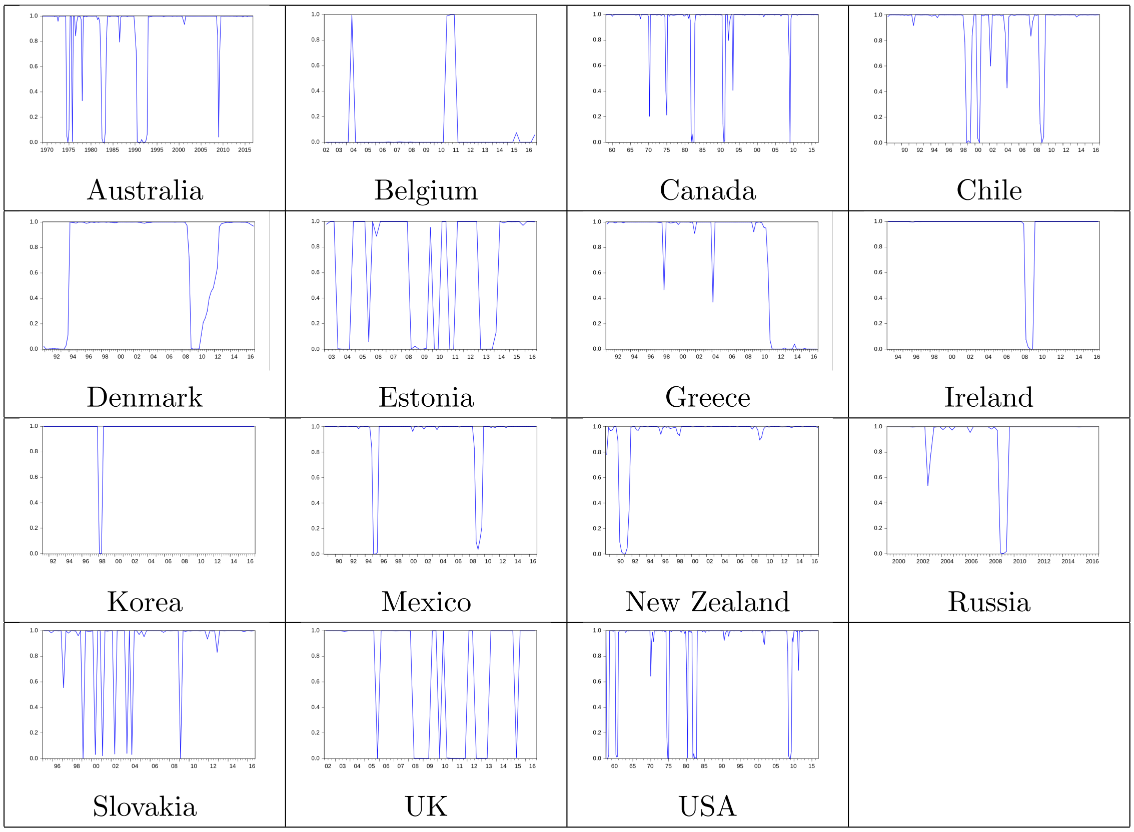

Figure 2.3: *Estimated Probability of the Low Unemployment Regime*

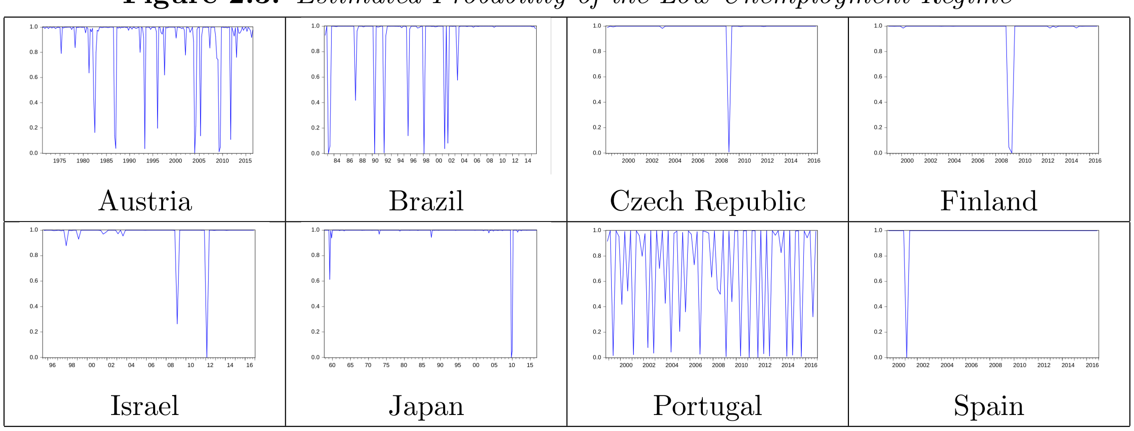

Table 2.1: *Key ADF Test Results*

Country	α	Lags	HL	Country	α	Lags	HL
Australia	-0.0197 (0.0087)	9	8.71	Italy	-0.0336 (0.0202)	7	5.07
Austria	-0.0129 (0.0151)	8	13.35	Japan	-0.0081 (0.0056)	10	21.31
Belgium	-0.6141** (0.1801)	13	0.18	South Korea	-0.1055** (0.0315)	4	1.55
Brazil	-0.0330 (0.0232)	5	5.16	Mexico	-0.0721* (0.0257)	5	2.32
Canada	-0.0174 (0.0101)	12	9.87	Netherlands	-0.0293 (0.0127)	2	5.83
Chile	-0.0903* (0.0343)	7	1.83	Norway	-0.0827 (0.0349)	2	2.01
Czech Republic	-0.0507 (0.0230)	1	3.33	New Zealand	-0.0440 (0.0183)	9	3.85
Denmark	-0.0476 (0.0262)	2	3.55	Poland	-0.0131 (0.0102)	11	13.14
Estonia	-0.2156** (0.0635)	11	0.71	Portugal	-0.0171 (0.0102)	2	10.05
Finland	-0.0948** (0.0310)	2	1.74	Russia	-0.1018*** (0.0279)	1	1.61
Germany	-0.0063 (0.0037)	13	27.42	Slovakia	-0.0355 (0.0195)	6	4.79
Greece	-0.0126 (0.0059)	3	13.67	Slovenia	-0.0723 (0.0417)	13	2.31
Hungary	-0.0509 (0.0235)	9	3.32	Spain	-0.0133 (0.0104)	1	12.94
Ireland	-0.0428** (0.0141)	12	3.96	UK	-0.0656 (0.0266)	11	2.55
Israel	-0.0208 (0.0244)	1	8.24	USA	-0.0428** (0.0129)	9	3.96

Table 2.2: *Key DF-GLS Test Results*

Country	α	Lags	HL	Country	α	Lags	HL
Australia	-0.0038 (0.0047)	4	45.52	Italy	-0.0153 (0.0138)	2	11.24
Austria	-0.0008 (-0.0933)	4	216.52	Japan	-0.0074 (0.0056)	12	23.33
Belgium	-0.0979 (0.0619)	1	1.68	South Korea	-0.0577** (0.0263)	3	2.92
Brazil	-0.0200 (0.0158)	1	8.58	Mexico	-0.0507** (0.0215)	1	3.33
Canada	-0.0072 (0.0065)	12	23.98	Netherlands	-0.0209* (0.0128)	1	8.20
Chile	-0.0042 (0.0107)	1	41.17	Norway	-0.0829** (0.0347)	2	2.00
Czech Republic	-0.0394** (0.0190)	1	4.31	New Zealand	-0.0207 (0.0129)	4	8.28
Denmark	-0.0239 (0.0216)	0	7.16	Poland	-0.0071 (0.0116)	1	24.32
Estonia	-0.0343 (0.0303)	1	4.96	Portugal	-0.0126 (0.0092)	2	13.67
Finland	-0.0174 (0.0180)	2	9.87	Russia	-0.0014 (0.0149)	0	123.69
Germany	-0.0016 (0.0023)	13	108.22	Slovakia	-0.0257* (0.0149)	1	6.66
Greece	-0.0067 (0.0056)	4	25.78	Slovenia	-0.0397 (0.0297)	4	4.28
Hungary	-0.0266 (0.0193)	4	6.43	Spain	-0.0134 (0.0103)	1	12.85
Ireland	-0.0131 (0.0097)	2	13.14	UK	-0.0269 (0.0195)	1	6.35
Israel	-0.0160 (0.0245)	3	10.74	USA	-0.0217** (0.0099)	12	7.90

Table 2.3: *Perron and Yabu Part 1*

Country	μ	β	$\hat{\alpha}_{TW}$	$\bar{\alpha}_{MS}$	**Lags**
Australia	1.781 (0.539)	0.019 (0.038)	0.985	1	5
Austria	1.973 (0.222)	0.019 (0.002)	0.875	1	5
Belgium	8.989 (0.358)	0.025 (0.042)	0.826	1	2
Brazil	13.397 (0.931)	-0.043 (0.079)	0.943	1	2
Canada	4.792 (0.421)	0.009 (0.027)	0.977	1	13
Chile	13.965 (0.291)	-0.060 (0.062)	0.950	1	2
Czech Republic	5.744 (0.824)	-0.024 (0.095)	0.914	1	2
Denmark	8.092 (0.464)	-0.017 (0.045)	0.953	1	1
Estonia	15.643 (1.519)	-0.117 (0.187)	0.922	1	2
Finland	11.483 (0.457)	0.037 (0.053)	0.931	1	3
Germany	0.380 (0.526)	0.016 (0.035)	0.993	1	14
Greece	6.714 (2.140)	0.154 (0.206)	0.978	1	3
Hungary	7.295 (0.817)	-0.032 (0.097)	0.971	1	5
Ireland	12.777 (1.349)	-0.046 (0.130)	0.980	1	3
Israel	7.261 (0.442)	-0.032 (0.047)	0.958	1	4
Italy	11.692 (0.639)	0.001 (0.073)	0.977	1	3
Japan	2.462 (0.220)	0.002 (0.014)	0.977	1	13
Korea	2.422 (0.536)	0.011 (0.052)	0.911	1	4
Mexico	4.517 (0.475)	-0.007 (0.043)	0.929	1	2
Netherlands	4.670 (0.586)	0.011 (0.067)	0.965	1	2
New Zealand	4.192 (0.675)	0.008 (0.061)	0.961	1	5
Norway	3.377 (0.350)	0.018 (0.042)	0.913	1	3
Poland	13.594 (1.230)	0.005 (0.124)	0.949	1	2
Portugal	5.497 (1.204)	0.063 (0.138)	0.952	1	3
Russia	13.650 (0.457)	-0.115 (0.054)	0.966	1	1
Slovak	12.463 (1.238)	-0.031 (0.128)	0.998	1	2
Slovenia	7.172 (0.522)	0.012 (0.062)	0.931	1	5
Spain	16.373 (1.863)	0.032 (0.220)	0.980	1	2
UK	6.214 (0.456)	-0.021 (0.054)	0.970	1	2
US	4.734 (0.454)	-0.000 (0.029)	0.958	1	13

Table 2.4: *Perron and Yabu Part 1 Unit Root Test Results*

Country	α	Lags	HL	Country	α	Lags	HL
Australia	-0.0069 (0.0059)	4	25.03	Italy	-0.0098 (0.0111)	2	17.60
Austria	-0.0780** (0.0307)	1	2.13	Japan	-0.0099 (0.0063)	10	17.42
Belgium	-0.0845 (0.0546)	1	7.85	South Korea	-0.0703** (0.0286)	3	2.38
Brazil	-0.0408* (0.0211)	1	4.16	Mexico	-0.0465** (0.0204)	1	3.64
Canada	-0.0120 (0.0081)	12	14.35	Netherlands	-0.0213* (0.0132)	1	8.05
Chile	-0.0187 (0.0144)	1	9.18	Norway	-0.0452* (0.0251)	2	3.75
Czech Republic	-0.0196 (0.0128)	1	8.75	New Zealand	-0.0177 (0.0120)	4	9.70
Denmark	-0.0318 (0.0243)	0	5.36	Poland	-0.0018 (0.0110)	1	96.18
Estonia	-0.0514 (0.0332)	1	4.96	Portugal	-0.0243* (0.0134)	2	13.67
Finland	-0.0203 (0.0191)	2	8.45	Russia	-0.0154 (0.0209)	0	11.17
Germany	-0.0025 (0.0028)	13	69.22	Slovakia	-0.0156 (0.0110)	1	11.02
Greece	-0.0186** (0.0077)	4	9.23	Slovenia	-0.0436 (0.0305)	4	3.89
Hungary	-0.0148 (0.0131)	4	11.62	Spain	-0.0144 (0.0107)	1	11.95
Ireland	-0.0166 (0.0105)	2	10.35	UK	-0.0022 (0.0039)	1	78.68
Israel	-0.0158 (0.0179)	3	10.88	USA	-0.0184** (0.0091)	12	9.33

Table 2.5: *MS-ADF Test Results*

Country	c_L	c_H	α	μ_L	μ_H	**p**	**q**	**HL**
Australia	0.144	0.779	-0.0315*** (0.0075)	4.57	24.73	0.964	0.705	5.41
Austria	0.072	0.625	-0.0234 (0.0139)	3.08	26.71	0.937	0.284	7.32
Belgium	5.263	6.314	-0.7953*** (0.1514)	6.62	7.94	0.481	0.961	0.11
Brazil	0.302	2.046	-0.0458 (0.0186)	12.64	44.67	0.936	0.119	3.70
Canada	0.139	0.914	-0.0239* (0.0090)	5.82	38.24	0.975	0.509	7.16
Chile	0.505	1.675	-0.0811* (0.0311)	6.22	20.65	0.959	0.544	2.05
Czech Republic	0.183	1.109	-0.0301 (0.0223)	6.08	10.55	0.986	0.000	5.67
Denmark	0.612	1.332	-0.1263*** (0.0333)	4.85	10.55	0.977	0.939	1.28
Estonia	2.169	4.037	-0.2986*** (0.0384)	7.26	13.52	0.834	0.668	0.49
Finland	0.446	1.395	-0.0590 (0.0287)	7.56	23.64	0.985	0.469	2.85
Greece	0.668	1.533	-0.0653*** (0.0144)	10.23	23.48	0.991	0.984	2.57
Ireland	0.276	1.708	-0.0425*** (0.0121)	6.49	40.19	0.989	0.737	3.99
Israel	-0.007	1.180	-0.0050 (0.0238)	-1.4	236	0.975	0.000	34.57
Japan	-0.335	0.018	-0.0044 (0.0055)	-76.14	4.09	0.370	0.993	32.30
South Korea	0.347	2.286	-0.1080*** (0.0223)	3.21	21.17	0.990	0.495	1.52
Mexico	0.315	0.982	-0.0921*** (0.0231)	3.42	10.66	0.980	0.691	1.79
New Zealand	0.299	0.914	-0.0541** (0.0184)	5.53	16.89	0.985	0.765	3.12
Portugal	-0.021	0.525	-0.0147 (NA)	-1.43	35.71	0.573	0.000	11.70
Russia	0.621	1.886	-0.1147*** (0.0239)	5.41	16.44	0.575	0.977	1.42
Slovakia	0.636	1.741	-0.0531** (0.0152)	11.98	32.79	0.903	0.000	3.18
Spain	-2.225	0.296	-0.0156 (0.0088)	-142.63	18.97	0.000	0.985	11.02
UK	0.572	1.005	-0.1168*** (0.0143)	4.90	8.60	0.853	0.658	1.40
USA	0.223	0.874	-0.0460*** (0.0108)	4.85	19.00	0.967	0.647	3.70

Table 2.6: *Perron and Yabu Pretest*

Country	**M2 W_{RQF}**	**M2 Break Date**	**M3 W_{RQF}**	**M3 Break Date**
Australia	0.400	1985Q3	11.939***	1994Q3
Austria	0.282	1984Q1	1.2489	1981Q3
Belgium	0.134	2011Q3	1.111	2002Q1
Brazil	0.317	2003Q4	3.462*	1997Q4
Canada	0.123	1991Q4	3.922**	1982Q1
Chile	0.804	1989Q2	2.072	1998Q3
Czech Republic	0.220	2014Q2	3.686**	2009Q2
Denmark	0.438	2001Q4	30.951***	2009Q1
Estonia	0.127	2005Q2	1.760	2008Q4
Finland	0.855	2007Q4	9.978***	2008Q4
Germany	1.023	2005Q2	12.2575***	2006Q4
Greece	0.040	2008Q3	3.077*	2011Q2
Hungary	7.517***	2012Q4	12.698***	2009Q2
Ireland	0.253	2002Q2	8.507***	2008Q4
Israel	0.835	2002Q3	6.677***	2004Q4
Italy	3.032**	2007Q2	7.701***	2006Q4
Japan	1.4677*	2964Q4	8.300***	1998Q1
Korea	0.035	1991Q1	16.333***	1997Q4
Mexico	0.079	2001Q4	3.379**	1997Q2
Netherlands	0.115	2008Q3	1.300	2012Q2
New Zealand	1.366	1990Q4	2.588	1994Q1
Norway	0.184	2012Q3	1.724	2006Q2
Poland	0.495	2002Q3	3.837**	2005Q4
Portugal	2.244**	2014Q1	18.810***	2011Q4
Russia	2.315**	2002Q1	4.525**	2002Q1
Slovak	0.370	2001Q3	6.915***	2005Q2
Slovenia	0.445	2007Q4	2.083	2009Q4
Spain	0.437	2005Q3	1.139	2008Q4
UK	1.106	2013Q1	2.150	2008Q4
US	0.048	1982Q1	3.300*	2008Q4

Table 2.7: *Perron and Yabu Structural Break Test on Trimmed Data*

Country	Max W_{RQF}	Trimmed Data Span
Australia	9.211***	1998Q1-2013Q2
Brazil	6.935***	2000Q3-2013Q1
Canada	3.981**	1988Q1-2011Q1
Czech Republic	4.121**	2000Q1-2007Q3
Denmark	0.711	1993Q1-2006Q2
Finland	1.614	1999Q3-2007Q1
Germany	6.727***	1968Q4-2000Q1
Greece	6.131***	1993Q2-2008Q1
Hungary	15.611***	2000Q3-2007Q3
Ireland	3.488**	1992Q4-2006Q1
Israel	7.802***	2006Q2-2015Q1
Italy	10.809***	1992Q1-2005Q3
Japan	5.578***	1961Q3-1991Q3
Korea	1.701	2000Q4-2013Q4
Mexico	1.562	2000Q1-2013Q4
Poland	21.901***	2007Q2-2015Q1
Portugal	3.682**	2000Q1-2009Q4
Russia	7.078***	2004Q3-2014Q3
Slovak	9.740***	2007Q1-2015Q1
US	6.587***	1961Q1-2000Q3

Table 2.8: *Single Break Unit Root Test*

Country	Break Date	Lags	$\hat{\alpha}$
Denmark	2009Q1	0	-2.9484
Finland	2008Q4	2	-3.6767*
Korea	1997Q4	1	-3.9538**
Mexico	1995Q1	1	-3.6239*

Table 2.9: *Two Break Unit Root Test*

Country	P_T^{GLS}
Australia	7.463**
Brazil	7.912*
Canada	6.104**
Czech Republic	59.657
Germany	3.833***
Greece	72.399
Hungary	26.959
Ireland	27.583
Israel	7.378**
Italy	1.742***
Japan	17.147
Poland	32.860
Portugal	18.495
Russia	17.852
Slovak	70.959
US	6.538**

Chapter 3

An Alternative Historical Analysis of Monetary Policy Rules

3.1 Introduction

This paper begins by arguing that constructing output gaps for use in Taylor Rule research using traditional filtering techniques is an inferior approach to using the newly developed Beveridge-Nelson Filter (Kamber, Morley and Wong 2017). Specifically, I argue that the commonly used filter of Hodrick and Prescott (1980) poses a multitude of theoretical issues which threaten the accuracy and reliability of any output gap found using this technique. Not only does this filter potentially lead to spurious estimates of the output gap, even non-spurious estimates of the output gap are at risk of being poorly approximated in real time due to endpoint approximation issues with the filter. I propose to instead construct the output gap using an amended version of the Beveridge-Nelson decomposition that, contrary to the commonly used HP filter, threatens no spurious results and has no issues with accurate endpoint approximations. This amended version of the BN decomposition was proposed by Kamber, Morley and Wong (hereafter referred to as KMW) and will be called the Beveridge-Nelson Filter. To my knowledge, this paper presents the first use of the Beverage Nelson Filter in Taylor Rule research.

This paper's primary focus is to re-evaluate key aspects of John Taylor's 1999 paper *A Historical Analysis of Monetary Policy Rules* and to extend his analysis to include modern data. This is done by first re-estimating all output gaps using the Beveridge-Nelson Filter rather than the Hodrick-Prescott Filter employed by Taylor. Taylor Rules are then estimated for various eras of monetary policy, including a Bretton Woods, Post-Bretton Woods, and Zero Lower Bound eras. Baseline policy rules are then compared to actual federal funds rates. In contrast to Taylor's results, use of the BN Filter in constructing output gaps leads to the conclusion that monetary policy followed a baseline model more closely in the 1960s relative to the widely cited late 80's/early 90's.

I then move on the estimating the break dates of various eras of monetary policy endogenously. I create a deviation variable measuring the extent to which the baseline Taylor Rule exceeds the actual federal funds rate and I utilize this variable in a Bai and Perron (1998, 2003) structural change test in the spirit of Rzhevskky, Papell, and Prodan (2017). Doing this, I find an additional fifth era of monetary policy.

I conclude the paper with a simple wide modeling approach that incorporates several common baseline Taylor Rule models chosen to represent the evolution of the Taylor Rule over the years. This approach again concludes the Federal Funds Rate more closely followed a baseline Taylor Rule in the 60s than it did in any other stretch of time.

The paper is organized as follows. Section 3.2 provides a motivation for the paper. I set up the Taylor Rule framework, discuss issues with using the Hodrick-Prescott Filter to create output gaps, and discuss the Beveridge-Nelson Filters fix to some of these issues. This section is essentially a literature review extended to interpret filtering concerns specifically in the context of Taylor Rule research. Section 3.3 contains all results of reworking Taylor's 1999 paper by using the BN filter. Section 3.4 contains the Bai and Perron structural change test. In section 3.5, I do a simple wide modeling of Taylor Rule baseline models using the two different output gaps. Section 3.6 concludes.

3.2 Motivation and Literature Review

3.2.1 The Basic Taylor Rule

The Taylor Rule is a monetary policy framework that describes the short-term nominal interest rate to be a function of the inflation rate, the Feds target inflation rate, an output gap, and the equilibrium level of the real interest rate. The appeal in this type of equation is the simplicity through which the Taylor Rule reflects the idea that the Federal Reserve follows a dual mandate: as output rises above (below) potential, or as inflation rises above (below) target, the Fed will raise (lower) the interest rate to stabilize the economy. The relationship was first described in Taylor (1993) and has been the focus of a significant number of papers since. Though many of these papers offered modifications over the years, Taylors original rule is

$$i_t = \pi_t + \phi(\pi_t - \bar{\pi}) + \gamma y_t + R$$

where i_t is the short-term nominal interest rate, π_t is the year-over inflation rate, $\bar{\pi}_t$ is the target inflation rate, R is the equilibrium real interest rate, and y_t is the output gap. This equation can then be trivially rearranged to be denoted

$$i_t = c + \lambda\pi_t + \gamma y_t$$

where $\lambda = 1 + \phi$ and $c = R - \phi\bar{\pi}$. This simplest Taylor Rule form has been the foundation for years of Taylor Rule research. Using this notation, the logic of the Taylor Rule dictates that γ should be greater than 0 and λ should be greater than 1. If γ were less than zero, it would dictate that the federal funds rate falls as the output gap increases. If λ were less than 1, it would likewise dictate that the federal funds rate should fall as the actual inflation rate outpaces the target inflation rate. Both of these realities would indicate that the Federal Reserve is violating its dual mandate to maximize output while keeping inflation low and stable.

This paper is largely interested in following the line of literature focused on the values of the two coefficients λ and γ. In accordance with Orphanides (2001), virtually all work over the last 15 years with this objective has used real-time data so that the equation will reflect only the information available to policy makers at the time of their decisions. The data for the inflation rate and the output gap both come from The Real-Time Dataset for Macroeconomists maintained by the Philadelphia Fed. The short-term interest rate data series comes from compiling the Feds overnight lending rate into quarterly data for 1958-2008. For the years 2009-2015, I instead us the Wu and Xia (2016) shadow interest rate to deal with the Federal Funds Rate reaching the zero lower bound. These rate, first introduced by Black (1995), are estimated using end-of-month Nelson-Seigel-Svennson yield curve parameters (Svennson 1994) estimated from the Gurkaynak, Sack, and Wright (2006) dataset. I return to using the official Federal Funds Rate beginning in 2016 when the shadow rates return to positive values.

For the purpose of this paper, the primary variable of interest is the output gap, y_t, defined by Taylor as the being the percentage deviation of actual output from potential. For a researcher working with this model, the task is then left to the individual to choose which method to use in calculating what exactly the potential of output is. This has been a long standing problem with using monetary policy rules, identified as early as 1997 by Alan Greenspan. For a researcher wanting to detrend output using a filter such as Hodrick-Prescott, the common approach is to identify the estimated trend of the filtered series to be potential output. In the next section, I will show why this is problematic.

3.2.2 A Quick Background in Detrending Data with the HP Filter

Arguably the most common method currently used for detrending output is to run the log of output through a filter such as the Hodrick-Prescott filter. The goal of this filter is to separate a series (in this case, output) into two distinct components: a stochastic trend component and cyclical component. For the remainder of this paper, I will denote this relationship with the following equation.

$$x_t = \mu_t + \psi_t$$

In this notation, x_t is the log of output, μ_t is the stochastic trend and ψ_t is the cyclical component.

The Hodrick-Prescott filter approaches the task of computing the trend by minimizing the loss function

$$\Sigma_{t=1}^{T}\left[(x_t - \mu_t)^2 + \lambda(\mu_{t+1} + 2\mu_t + \mu_{t-1})^2\right]$$

to obtain an optimal trend which we can call μ_t^*. This essentially minimizes the cycle component while imposing a penalty for variation in the second difference of the growth component. As such, the first squared term of this loss function captures the accuracy of the estimated trend while the second squared term captures the smoothness of the trend. The term λ determines the weights of these two components and should be an estimate of the variance of the underlying stochastic components (though standard procedure is to set $\lambda = 1600$ for quarterly data, in line with Hodrick and Prescotts original work on the subject). Once the optimal trend is computed, the cycle is trivially backed out as $\psi_t = x_t - \mu_t^*$. To use the results of this filter, researchers then set the trend element of the filtered data to be equal to potential output. This interpretation then indicates that the cyclical component of an HP filtered log-output series will be the output gap if we use choose to use the Hodrick Prescott filter to detrend our data.

3.2.3 Potential Issues Associated with HP Filtering a Time Series

With the output gap then being defined as the cycle of the filtered series, we need to pay particular attention to the cycles that arise from these filtering processes. One such concern arises when filtering time series with the typical spectrum shape identified by Granger (1966) as being common across most economic series measured in levels (this typical spectrum has a smooth, strictly decreasing, convex shape). When applied to these spectrums of low frequency, the HP filter provides distorted cycles, performing poorly (Guay and St-Amant (2005)). Cogley and Nason (1995) find that for a difference stationary time series (as our series of output is) the Hodrick Prescott filter is likely to generate spurious cyclical structure at business cycle frequencies. This means that we must question whether or not any output gaps created by the HP filter are spurious. Harvey and Jager (1993) find that applying the HP filter to a unit root process also produces spurious cycles. Running an Augmented Dickey Fuller test on my series of output, I find a p-value of 0.4028, meaning I cannot reject a unit root being present in the series of output. This again draws question to the accuracy of the output gaps implied from running the HP filter.

The Hodrick Prescott filter also has additional problems associated with estimating endpoints of a time series. Baxter and King (1999) note that the Hodrick Prescott filter results in unusual behavior of the cyclical components near the end of the sample. This results from the fact that total output in any given period is set equal to a function of trend values that span periods both prior to and following the given date. This means that as this filter approaches the end-points of the time series, it must alter the defined relationship between actual output and trend to create an approximation. This approximation is attractive in business cycle studies as it allows the HP filter to create an estimate of the cycle for every time period in which we have data. Business cycle studies usually span many years of data, so a little extra variation at the edges might be minimally invasive enough that it does not drastically change the overall results.

However, in creating output gaps for the use of studying Taylor Rules, this approximation is likely to be far more damaging. Reason being, virtually all Taylor Rule studies use real time or quasi real time data. This means that to determine the output gap at any given time period (lets say time T for the sake of discussion), we will filter the period T vintage of the data and record only the output gap computed for period T. We would then move on to filter the T+1 vintage of the data to record the output gap for only that period, and so on for the entirety of the sample. What this results in is the reality that our series of output gaps computed by the HP filter is a series of only endpoint estimates of that gap at every time period; all of which are susceptible to this trap of poor fit. This could be theoretically fixed if we dropped a certain number of endpoints from the sample. But doing so would defeat a primary purpose of using real-time data, as we want to be able to derive an output gap for any given quarter using only information available up to that given quarter.

Figure 3.1 highlights this issue by applying a procedure from Giles (2011) to HP filter the 1984Q2 vintage of output with confidence intervals. Removing 2 years from each end of the series, the average span of the confidence interval for the middle bulk of the estimated output gaps is 2.73. However, at 1984Q2 (the only point we care about when using real-time data), the span of the confidence interval is 77% larger at 4.85. Constructing an output gap series with the HP filter in real time will thus lead to low relatively confidence estimates for the entire series.

3.2.4 Motivating Use of the Beveridge-Nelson Filter

The basic purpose of the Beveridge-Nelson Decomposition (1980) is the same as that of the HP filter; it still attempts to separate a single time series into two components: a cycle and a trend. We start with the same basic breakdown as before.

$$x_t = \mu_t + \psi_t$$

Where the variable μ_t is a deterministic component and ψ_t is a stochastic component. Neither component is required to adhere to a specific functional form across all decompositions, but a common definition of the deterministic trend is to simply make it

$$\mu_t = \kappa + \delta t$$

where κ and δ are constants. I will stick with this definition for the sake of easier communication; despite it not being the only form μ theoretically can take. The decompositions stochastic element is found only inside the second term. The BN decomposition defines this term to be the sum of a stochastic trend component and an I(0) cyclical component

$$\psi_t = TS_t + C_t$$

This is designed so that the stochastic trend, TS, captures all shocks that permanently change the level of x_t and the stationary component, C_t, captures only shocks with a temporary impact on the level of x_t. We can plug this relationship back into the original equation for x_t to get

$$x_t = \mu_t + TS_t + C_t$$

where $(\mu_t + TS_t)$ is the total trend and C_t represents deviation around it. What is convenient about this form is that Beveridge and Nelson showed that if our output series is stationary in first differences (as we have already shown it is), the BN decomposition would break that series down so that the estimated total trend is a random walk with a drift which captures growth (the stochastic portion of total trend, TS, is a pure random walk and the drift comes from the deterministic component μ), and the cycle is stationary. To do this, they propose to set the long horizon conditional expectation of the time series $\{x_t\}$ to be the long-horizon conditional expectation of the trend component (assuming that the long-horizon conditional expectation of the cycle is zero).

$$\mu_t + TS_t = lim_{j \to \infty} E[x_{t+j} - jE[\Delta x]]$$

Then the implied cycle can then be calculated in every period as

$$C_t = x_t - \mu_t - TS_t$$

With this basic framework of the BN Decomposition, we now want to consider how this structure of decomposing output alleviates some of the concerns of the HP filter. First, we should note that the BN Decomposition does not suffer from the endpoint approximation shortcomings of the HP filter. BN computes its trend by creating a limiting infinite future forecast of the output level around which the trend component is partially estimated. This means no trimming will be necessary to maintain consistent endpoint approximations when we are using real-time data. Second, the HP Filter is susceptible to create spurious cycles when the original data is difference stationary, as output is. The BN Decomposition on the other hand actually utilizes difference stationarity in its formulation of the random walk trend component. The Beveridge-Nelson Decomposition thus has the advantage that it avoids spurious cycles all together, including for series with the typical Granger spectrum.

The BN Decomposition is not without its faults however, particularly when used to create output gaps. In particular, this process has been shown to produce output gaps that are small in amplitude and lacking in persistence. In addition, the calculated output gaps often miss the reference cycle of US expansions and recessions (KMW 2017). Kamber, Morely and Wong attribute these failures to the BN decomposition's tendency to imply "a very high signal-to-noise ratio in terms of the variance of trend shocks as a fraction of the overall quarterly forecast error variance." They provide a method for imposing a low variance-to-noise ratio and coin the term 'Beveridge-Nelson Filter' to describe the amended process. When applying this BN Filter to U.S. log real GDP, they find output gaps that are largely persistent, large in amplitude, and matching well with the NBER business cycles. In addition, they find that real-time estimates are subject to smaller revisions and that they out perform out-of-sample forecasts of output growth and inflation compared to other trend cycle decomposition methods, including the HP filter.

To create a series of output gaps using the BN Filter, I follow KMW's approach of estimating the output gap using the Beveridge-Nelson Decomposition based on a Bayesian estimation of an AR(12) while imposing a signal to noise ratio that optimizes the tradeoff between amplitude and fit. I estimate this signal to noise ratio to be 0.2400. For comparison, the standard Beveridge-Nelson Decomposition using an AR(1) estimates the signal to noise ratio to be 2.22. Such a result would indicate that trend shocks are much more volatile than quarter to quarter forecast errors in log real GDP. A signal to noise ratio of 0.2400 implies that trend shocks explain 24% of the quarterly forecast error variance.

The results of estimating the BN Filter on real time data of the log of real US

GDP is presented in Figure 3.2, along with the real time estimates of the output gap using the Hodrick Prescott Filter. Note that both series capture NBER recessions and expansions well. Two important distinctions should be noted between these two series. First, the HP Filter indicates more extreme movement of the output gaps. In every recession, and all but 1 expansionary periods, the HP Filtered estimated output gap reaches a greater absolute amplitude than does the BN Filter.

Second, the greatest period of divergence between the two output gaps occurs during and in the aftermath of the Great Recession. The HP Filter paints a decidedly prettier picture of the effect on output from this time than does the BN Filter. Relative to previous recessions, the fall in output estimated to occur during the Great Recession by the HP Filter is not as severe or as persistent as estimated by the BN Filter. The HP Filter estimates that the output gap reaches a trough of 4.37% below potential output during this period. The average for all previous recessions was a trough of 3.04% below potential, making the maximum estimated loss of output during a single quarter of the Great Recession 44% more severe than the average of the maximum single quarter loss of all previous recessions. In comparison, the BN Filter estimates that the trough of the Great Recession was 86% greater than the average trough of all previous recessions (3.91% during the Great Recession vs. a 2.10% average during all other recessions). This discrepancy is potentially caused by the endpoint approximation issues of the HP Filter.

The persistence of the negative effect on output is also estimated to be greater when using the BN Filter as opposed to the HP Filter during the Great Recession. From 2007Q4 (the NBER estimated start date of the Great Recession), the HP Filter estimated that it took only 11 quarters for the output gap to return to a positive figure. The average length of time across all previous recessions required to return to a positive output gap from the official NBER recession start date was 8.57 quarters. The BN Filter on the other hand estimates that the output gap has not returned to a positive figure since the Great Recession's onset. The dataset ends in 2016Q4, so this is a total of 36 quarters. The average comparable length across all previous recessions was estimated to be 9.29 quarters. So whereas the HP Filter identifies the Great Recession to look a lot like a slightly worse than average recession, the BN Filter identifies it as a significant deviation of the recessions of the past. Given what we know about this period, the BN Filtered estimates of the output gap feel more appropriate.

3.3 *Historical Analysis* Revisited

John Taylor's 1999 paper *A Historical Analysis of Monetary Policy* takes a historical approach to monetary policy evaluation. However, in his creation of his output gap variable, Taylor uses the Hodrick-Prescott Filter to detrend his output series. As a result, his estimation of the output gap is susceptible to the issues outlined in the previous section. Here, I repeat some of Taylor's key tests while instead creating the output gaps using the Beveridge-Nelson Filter and extending the data to include information through 2016. I also use quasi-real time data in accordance with Orphanides and Van Norden (2002), whereas Taylor used only revised data.

3.3.1 Estimation Exercise

Taylor starts by estimating the Taylor Rule discussed in section 3.2.1 across intervals of time.

$$i_t = c + \lambda\pi_t + \gamma y_t$$

Taylor estimates this equation over various time periods and concluded that "a monetary policy rule in which the interest rate responds to inflation and real output more aggressively than during the 1960s and 1970s... and more like the late 1980s and 1990s-is a good policy rule." I divide the late 50's, 60's, 70's and early 80's into a Bretton Woods era (1958-1973) and a Post-Bretton Woods era (1974-1987). I also extend Taylor's late 80's and 90's period to run through 2008, right up to the beginning of the "Zero Lower Bound era". The results of these baseline estimations across eras are presented in Tables 3.1 (with HP output gaps) and 3.2 (with BN output gaps).

The results here are largely consistent with John Taylor's. Using either the HP or BN generated output gaps results in the FFR's response to inflation and output gap shocks being significantly greater during the late 80's through late 2000's than do shocks during the late 50's through early 80's. This reality is even greater when using the BN Filter output gaps rather than the HP Filter output gaps. However, this reality breaks down completely at the onset of the Zero Lower Bound Era. During this time period, the estimated coefficients are all negative, implying that the Fed is operating in direct opposition to the intuition of the Taylor Rule. This estimation indicates that as output or inflation falls, the Fed responded by lowering the Federal Funds Rate, surpressing output and inflation further. The inflation rate and Federal Funds Rate are plotted in Figures 3.3 for reference.

3.3.2 Plotting Baselines Exercise

Taylor then goes on the observe two baseline policy rules and plots these baselines against the actual Federal Funds Rate for the Bretton Woods/Post-Bretton Woods eras. The two policy rules he uses are

$$Rule1 : i_t = 1 + 1.5\pi_t + 0.5y_t$$

$$Rule2 : i_t = 1 + 1.5\pi_t + y_t$$

The intuition behind this exercise is to interpret the baseline Taylor Rules as preferred policy rules. Using these rules, we can look at spans of history during which the Federal Funds Rate deviated from the preferred policy rules to the most significant degree. These deviations can be used as policy mistakes, and we can check the degree to which these mistakes effected the economy.

I replicate Taylor's original plots using the HP Filtered output gaps as well the BN Filtered output gaps for my all 4 of my time periods. The period spanning 1958 to 1973 are plotted in Figures 3.4 and 3.5. Figures 3.6 and 3.7 plot the period spanning 1974 to 1986. Figures 3.8 and 3.9 plot the periods spanning 1987 to 2008. And Figures 3.10 and 3.11 plot the Zero Lower Bound era comprising 2009-2016.

These Figures display a decidedly different reality from the one found by Taylor. He had concluded that there were three periods in which the Federal Funds Rate significantly deviated from the preferred policy rules: 1960-1964 (where the Fed's policy was too tight), 1965-1980 (where the Fed's policy was too loose), and 1982-1984 (where the Fed's policy again too tight). In all, Taylor concluded that the preferred policy rules were poor approximations of reality for nearly the entirety of the 60s, 70s and early 80s. The major contribution Taylor made with this analysis however was finding an exceptionally strong fit for the time period spanning 1987 to 1997. This was a time period in which the economy preformed exceptionally well, so finding the baseline Taylor Rules to so closely approximate the actual Federal Funds Rate was associated with the improved economic reality.

However, all it takes is real time data to show that this relationship breaks down to some degree, regardless of which Filter was used to create the output gaps. To show this, I define the Mean Minimum Absolute Deviation (MMAD) as

$$MMAD_i = \frac{1}{n}\sum_{t=0}^{n} min\{|TR1_{t,i} - FFR_t|, |TR2_{t,i} - FFR_t|\}$$

Where n is the number of quarters in the time span I am considering and the index i indicates whether the output gap was created by the HP Filter or the BN Filter.

Starting with the period in which Taylor found the strongest preferred policy adherence (1987-1997), I find that $MMAD_{HP} = 1.396$ and $MMAD_{BN} = 1.235$. Thus, the output gaps found using the BN Filter leads to an even stronger baseline policy adherence than the output gaps found using the HP Filter, strengthening Taylor's conclusion about this period. Further, this 'strong fitting period' can be extended all the way into the late 2000s. When considering the time period 1987 through 2008, I find $MMAD_{HP} = 1.441$ and $MMAD_{BN} = 1.318$. Again I conclude that the baseline Taylor Rules more closely approximate reality when using the BN Filter to create output gaps rather than the HP Filter.

However, this analysis deviates greatly from John Taylor's in that I no longer find this period to be the one of greatest preferred policy adherence. Rather, baseline Taylor Rules are an even better marker of reality for the Bretton Woods era spanning 1958 through 1973. During this time period, I find $MMAD_{HP} = 1.047$ and $MMAD_{BN} = 1.178$. Part of what made Taylor's paper so groundbreaking at the time was that the period of strong economic performance (1987-1997) displayed very close baseline policy adherence whereas the periods of poor economic performance (such as the pre-Bretton Woods era) seemed to correspond to greater baseline policy divergence. I see here that the baseline policies were in fact more closely followed in this pre-Bretton Woods era than they were during economy's strong stretch starting in the late 80s.

In addition, I find significant periods of deviation running from 1974 through 1986 and for the Zero Lower Bound era. From 1974-1986, I find $MMAD_{HP} = 2.849$ and $MMAD_{BN} = 3.050$. The Zero Lower Bound era finds and even greater divergence when using the HP Filter to construct output gaps. During this period, I find

$MMAD_{HP} = 3.961$ and $MMAD_{BN} = 3.003$.

Considering these results in sum, the first thing that sticks out is the extreme period of preferred policy divergence that has occurred in the aftermath of the great recession. The BN Filtered output gaps lead to baseline Taylor Rules that are closer approximations of reality than those utilizing HP Filtered output gaps during this period, but even those missed the mark significantly. With the crises so extreme, it appears that the Fed nearly entirely abandoned whatever approach they help during the prior 20 years in which they largely mimicked baseline policy rules.

Finally, it is worth noting that the BN Filtered output gaps resulted in baseline Taylor Rules that more closely mimicked the results originally found by Taylor than did using the HP Filtered output gaps. The baseline rules using the HP Filtered output gaps found closer adherence to the baseline policy rules from 1958 through 1986 than did the baseline rules utilizing the BN Filter, whereas the BN Filtered output gaps has lead to better fitting Taylor Rules since 1987.

3.4 Estimating Rules vs. Discretionary Eras of Monetary Policy

One shortcoming with the approach taken to this point is that the cutoff dates were determined exogenously. The periods chosen correspond to well known changes in the US macroeconomy, but exogenous cutoff dates still run the risk of being poorly chosen. The next task of this paper will be to identify regimes in which monetary policy closely follows a baseline Taylor by estimating break dates endogenously.

In this endeavor, I will emulate a key test from Nikolsko-Rzhevskyy et. al (2015) in which the authors define rules-based eras and discretionary-based eras of monetary policy solely from the data. Where I will differ from Rzhevskyy et. al is that they opted to use quadratic detrending to estimate their output gaps, concluding that the Hodrick Prescott Filter was ill-suited to the study. However, given that the BN Filter addresses some of the key issues associated with the HP Filter, I opt to repeat Rzhevskyy's key structural change test using this BN Filter.

The objective is to identify periods of time in which monetary policy closely followed a prescribed baseline rule (rules-based eras) and periods of time in which monetary policy deviates from a baseline policy rule greatly (discretionary-based eras). They find break dates at 1979Q4, 1987Q2, and 2000Q4, creating a total of 4 distinct eras. By instead using the Beverage Nelson Filter to construct output gaps, and with the benefit of a few extra years worth of data, I find that the 1987 break estimated by Rzhevskky occurred 4 years later, as well as finding an additional significant break date. Here's how...

To conduct this test, I first define a deviation variable, d_t as being the extent to which a prescribed baseline Taylor Rule exceeds the actual Federal Funds Rate. The baseline Taylor Rule I consider is the first one provided in section 3.3.2, John Taylor's original rule:

$$TR1_t = 1 + 1.5\pi_t + 0.5y_t$$

The deviation variable is then defined as $d_t = TR1_t - FFR_t$. This deviation variable is then used as the dependent variable in a Bai and Perron (1998, 2003) test for multiple structural breaks. Consider then the following linear regression model consisting of m structural breaks:

$$d_t = \gamma_0 + \gamma_1 DU_{1t} + \gamma_2 DU_{2t} + \ldots + \gamma_m DU_{tm} + u_t$$

d_t is the previously defined inflation variable and $DU_{mt} = 1$ if $t > TB_t$ and 0 otherwise, for all breakpoint TB_t.

This test is conducted by sequentially considering l vs. $l+1$ breaks, where the estimated breakpoints are obtained by a global minimization of the SSR. The Bai and Perron test proceeds as follows: First run a test of zero structural breaks against an alternative of one structural break. In the event that the zero break null is rejected in favor of the 1 break alternative, move forward one iteration and conduct a test with a null of one structural break against an alternative of two structural breaks. This process is repeated until no further significant breaks are found (where the maximum number of potential breaks is set equal to 5). A 15% trimming is also applied to each sub sample in accordance with Bai and Perron's (2003) recommendation the achieve the correct size.

Results of the estimation are provided in Table 3.3. As can be seen, I find 4 breaks at the 5% significance level, resulting in 5 eras. The first era spanning 1958 through

1970 is characterized by the Federal Funds Rate closely adhering to the baseline Taylor Rule, with the rate set only 0.81 percentage points below the prescribed Taylor Rule on average. This is consistent with a rules-based era of monetary policy. The break date occurring at 1971Q1 is a break that was not identified by Rzhevskyy et. al.

The first discretionary era seen is the next era found, running from 1971 through 1979Q3. The 1979Q4 break date matches exactly the break date estimated by Rzhevskyy et. al. With an estimated $\gamma_1 = 3.35$, the average deviation for the era is estimated to be 4.162 percentage points. This period displays the most extreme deviations from the baseline Taylor Rule of any era with the Federal Funds Rate consistently set below the baseline rule. This era in US history is also characterized by the most extreme inflation we have seen in the last 60 years.

The next estimated era spans 1979Q4 through the end of 1990. This 1991Q1 break date is a significant deviation from the prior literature, with Rzhevskyy et. al estimating this break to occur in 1987Q2. This era appears to be another discretionary era where the Federal Funds Rate exceeds the prescribed Taylor Rule by 1.859 percentage points on average. The deviation variable remains positive for the entirety of this stretch.

The fourth era spans 1991Q1 through 2001Q3. The previous paper found the break date at 2000Q4, so this result is a very minimal change. And like their paper, I conclude that this is a rules based era, one in which the Federal Funds Rate exceeds the prescribed Taylor Rule by 0.765 percentage points on average, the smallest average deviation of any era, even if negligibly different in average absolute amplitude from the 1958-1970 era. The final era runs from 2001Q4 to the end of the sample, 2016Q4. During this discretionary period of monetary policy, the prescribed Taylor Rule exceeded the Federal Funds Rate by 2.540 percentage points on average.

3.5 Wide Modeling Approach

Moving past mimicking the evaluations of prior studies, I lastly turn my attention to modernizing the Taylor Rules that I consider. Since Taylor's original work on the subject, a number of extensions to his simple baseline rules have been shown to improve upon the fit of those rules. In this section, I track the Taylor Rule's ability to capture the true Federal Funds Rate across time when using the BN Filtered estimates of the output gap by considering a number of these extended models. To do this, I employ a "thick" modeling approach wherein I use four alternate baseline Taylor Rule models to develop a span of the predicted Federal Funds Rates. The four baseline Taylor Rules I consider are

$$i_t = 1 + 1.5\pi_t + 0.5s_t \tag{3.1}$$

$$i_t = 0.15(1 + 1.5\pi_t + 0.5s_t) + 1.35i_{t-1} - 0.5i_{t-2} \tag{3.2}$$

$$i_t = 0.15(1 + 1.5\pi^e_{t+4} + 0.5s_t) + 1.35i_{t-1} - 0.5i_{t-2} \tag{3.3}$$

$$i_t = 0.15(1 + 1.5\pi^e_{t+4} + 0.5s_t + 0.75(s^e_{t+4} - s_t)) + 1.35i_{t-1} - 0.5i_{t-2} \tag{3.4}$$

These four baseline models are all versions of the benchmark model described in the section 3.2.1 with various amendments. These four equations are appealing because they display some of the key innovations of the Taylor Rule since its first appearance. Equation (3.1) is the original Taylor Rule. Equation (3.2) adds interest rate smoothing. Equation (3.3) adds a forward looking inflation rate. Equation (3.4) adds a forward looking measure of output gap.

For each of these four models, I input data for inflation and the output gap to develop four different predictions of the Federal Funds Rate. Figure 3.12 displays the span of the maximum and minimum predicted rates along with the actual Federal Funds Rate. In all, there are 44 quarters in which the actual Federal Funds rate fell inside the span of the maximum and minimum FFR predicted by the four baseline rules. Of those 44 quarters, 22 of them occurred before 1970. This figure supports the findings of the previous sections which concluded that the 60s are the period of time in which the Federal Funds Rate most closely approximated the baseline Taylor Rule.

3.6 Conclusions

The ultimate goal of this paper was to provide a first look into the usefulness of the Beveridge-Nelson Filter in Taylor Rule analysis. For a process that possesses so many theoretical advantages over the oft-used Hodrick-Prescott filter, the BN

Decomposition had not yet found its way into Taylor Rule research, largely due to inconsistencies with the estimated output gaps created by such a procedure. However, the recently created Beverage Nelson Filter of Kamber, Morely, and Wong addressed these issues by imposing a lower signal to noise ratio. When using this process, the calculated output gaps display favorable properties over those created by the Hodrick Prescott Filter. To my knowledge, this is the first paper which uses this Beverage Nelson Filter to create output gaps for use in Taylor Rule research.

By reworking John Taylor's 1999 paper *A Historical Analysis of Monetary Policy Rules* I find that use of the BN Filter in constructing baseline Taylor Rules suggests that the Fed's policy decisions did not mimic these baselines to nearly the degree that the traditional literature has implied. It was precisely the close adherence of the Federal Funds Rate of these baseline models in the late 80s/early 90's that lead to the Taylor Rule gaining notoriety. However, when using the BN Filter, this adherence weakens to some degree. It now appears that the Federal Funds Rate followed a baseline Taylor Rule just as closely in the 60s as they did during the economic expansion beginning in the late 80s. This is supported by a Mean Minimum Absolute Deviation variable, a wide modeling approach of 4 baseline Taylor Rules, and a structural change test of Bai and Perron.

3.7 References

Bai, J. and Perron, P. (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," Econometrics, 66, 47-78.

Bali, J. and Perron, P. (2003a), "Computation and Analysis of Multiple Structural Change Models," Journal of Applied Econometrics, 18, 1-22.

Bai, J. and Perron, P. (2003b), "Critical Values for Multiple Structural Change Tests," Econometrics Journal, 6, 72-78.

Baxter, M. and King, R.G. (1999), Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series, The Review of Economics and Statistics, 81, 575-593.

Belke, A. and lose, J. (2017), "Does the ECD rely on a Taylor Rule? Comparing Real-Time with Ex-Post Data", Bank and Bank Systems, 6(2)

Beveridge, S. and Nelson, C.R. (1981), A New Approach to the Decomposition of Economic Time Series Into Permanent and Transitory Components with particular Attention to the Business Cycle, Journal of Monetary Economics, 7, 151-174.

Black, F. (1995), "Interest Rates as Options", The Journal of Finance, Vol. 50, No. 5, 1371-1376

Cogley, T. and Nason, J. (1995), Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research, Journal of Economic Dynamics and Control, 19, 253-278.

Granger, Clive W.J. (1996), The Typical Spectral Shape of an Economic Variable, Econometrica, Volume 34, Number 1 (January) 150-161.

Greenspan, A. (1997), Remarks at the 15th anniversary conference of the Center for Economic Policy Research. Stanford University, 5 September.

Guay, A. and St-Amant, P. (2005), Do The Hodrick-Prescott and Baxter-King Filters Provide a Good Approximation of Business Cycles?, Annales dEconomie et de Statistique, Issue 77, 133-155.

Gurkaynak, R., Sack, B., and Wright, J. (2006), "The U.S. Treasury Yield Curve: 1961 to the Present", Working Papers Series, The Federal Reserve Board.

Harvey, A.C. and Jager, A. (1993), Detrending, Stylized Facts, and the Business Cycle, Journal of Applied Econometrics, Vol. 8, No. 3, 231-247.

Hodrick, R.J. and Prescott, E.C. (1980), Post-war U.S. Business Cycles; an Empirical Investigation, Working paper, Carnegie Mellon.

Hodrick, R.J. and Prescott, E.C. (1997), Post-war U.S. Business Cycles; an Empirical Investigation, Journal of Money, Credit, and Banking, 29, 1-16.

Ivanov, L. (2005), Is The Ideal Filter Really Ideal: The Usage of Frequency Filtering and Spurious Cycles, South Eastern Journal of Economics, 1, 79-96.

Kaiser, R. and Maravall, A. (1999), Estimation of the Business Cycle: A Modified Hodrick-Prescott Filter, Spanish Economic Review, Rev. 1, 175-206.

Kamber, G., Morley J., and Wong, B (2017), "Intuitive and Reliable Estimates of the Output Gap from a Beveridge-Nelson Filter", CAMA Working Papers, No 3/2017.

Mach, M. (2016), "Can the Taylor Rule be a Good Guidance for Policy? The Case of 2001-2008 Real Estate Bubble", Prague Economic Papers, 25, 4.

Meltzer, A.H. (2011), Federal Reserve Policy in the Great Recession, Remarks presented at Cato Institute Monetary Conference, November.

Levy, D. and Dezhbakhsh, H. (2003), On the Typical Spectral Shape of and Economic Variable, Applied Economic Letters, 10(7), 417-423.

Molodtsova, T., Rzhevskyy, A., Papell, D. (2008), "Taylor Rules with Real-Time Data: A Tale of Two Countries and One Exchange Rate", Journal of Monetary Economics, 55, October 2008, 63-79.

Morley, J. (2011), The Two Interpretations of the Beveridge-Nelson Decomposition, Macroeconomic Dynamics, 15, June, 419-439.

Murray, C. (2003), Cyclical Properties of Baxter-King Filtered Time Series, The Review of Economics and Statistics, 85, 472-476.

Nelson, C.R. (2008), The Beveridge-Nelson Decomposition in Restrospect and Prospect, Journal of Economics, 146, 202-206.

Orphanides, A. (2001), Monetary Policy Rules Based on Real-Time Data, American Economic Review, 91(4), September, 964-985.

Orphanides, A. and Van Norden S. (2002), "The Unreliability of Output Gap Estimates in Real Time", Review of Economics and Statistics, 84, 569-583

Rzhevskyy, A., Papell, D., and Prodan, R. (2013), (Taylor) Rules versus Discretion in U.S. Monetary Policy, Working Papers, University of Houston

Rzhevskyy, A., Papell, D., and Prodan, R. (2017), The Taylor Principles, Working Papers, University of Houston

Taylor, J.B. (1993), Discrection versus Policy Rules in Practice. Carnegie Rochester Conference Series on Public Policy, 39, 195-214.

Taylor, J.B. (1999), "A Historical Analysis of Monetary Policy Rules." Monetary Policy Rules, 319-348

Taylor, J.B. (2012), Monetary Policy Rules Work and Discretion Doesnt: A Tale of Two Eras, Journal of Money, Credit, and Banking, Vol. 44, No. 6, 1017-1032.

Wu, J. and Xia, F. (2016), "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound", Journal of Money, Credit and Banking, Vol 48

3.8 Tables and Figures

Figure 3.1: *Cycle Resulting from Applying HP Filter to 1984Q2 Vintage of Real GDP*

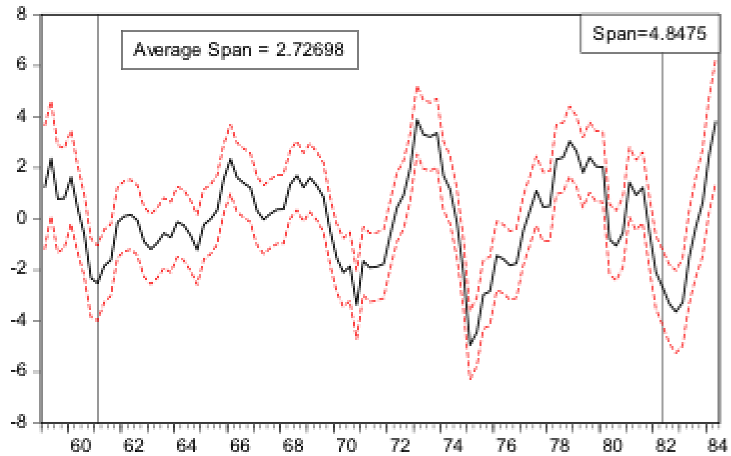

Figure 3.2: *HP and BN Output Gaps*

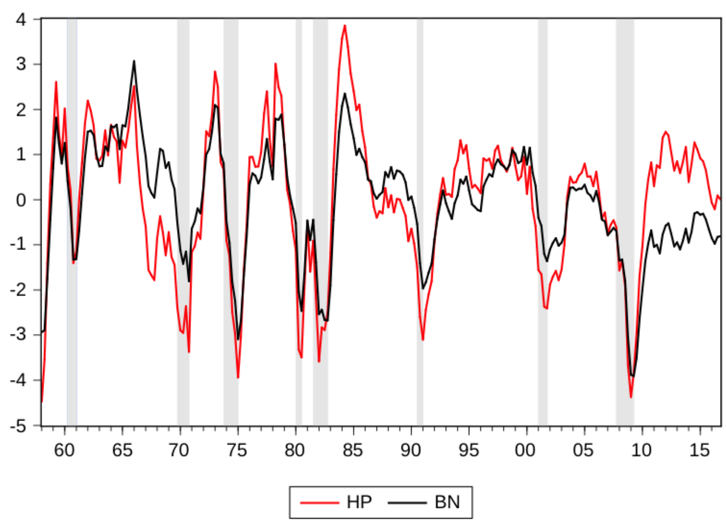

Figure 3.3: *Inflation and Federal Funds Rates*

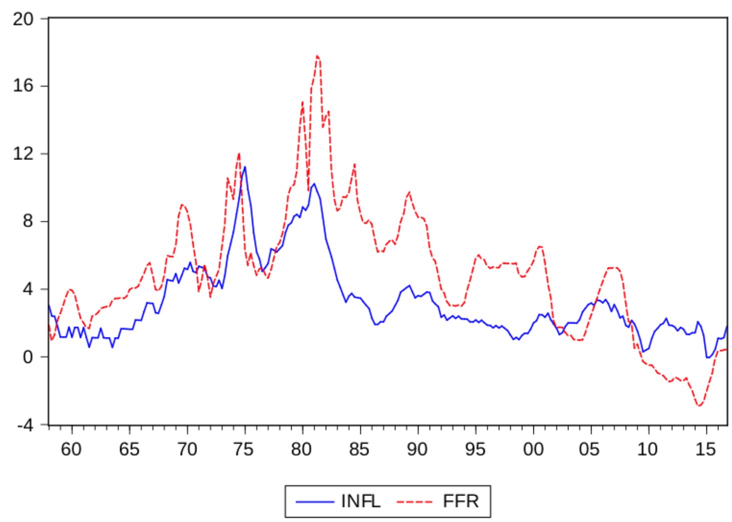

Figure 3.4: *Deviations from Baseline Policy Rule: 1959-1973 using HP OGs*

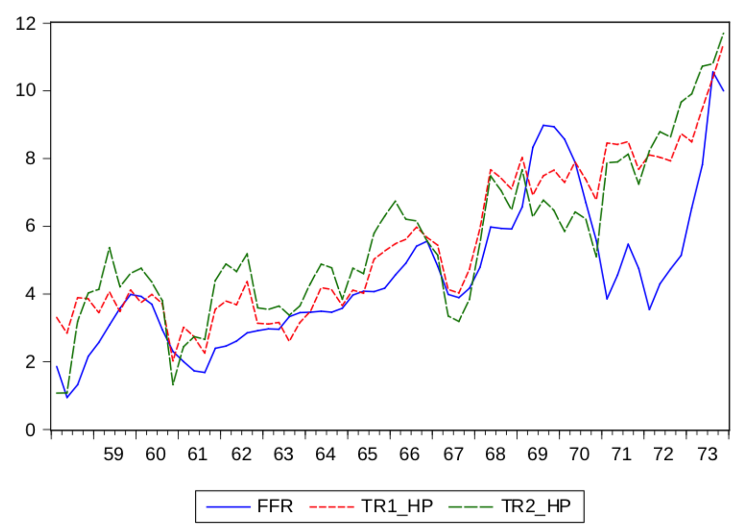

Figure 3.5: *Deviations from Baseline Policy Rule: 1959-1973 using BN OGs*

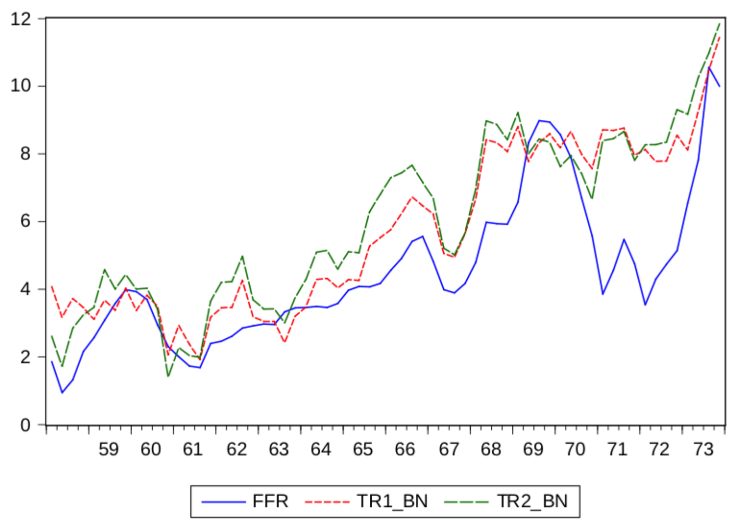

Figure 3.6: *Deviations from Baseline Policy Rule: 1975-1986 using HP OGs*

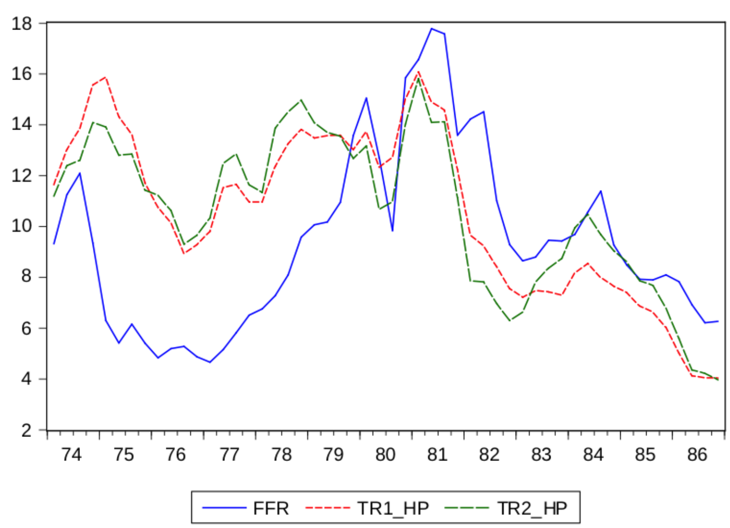

Figure 3.7: *Deviations from Baseline Policy Rule: 1975-1986 using BN OGs*

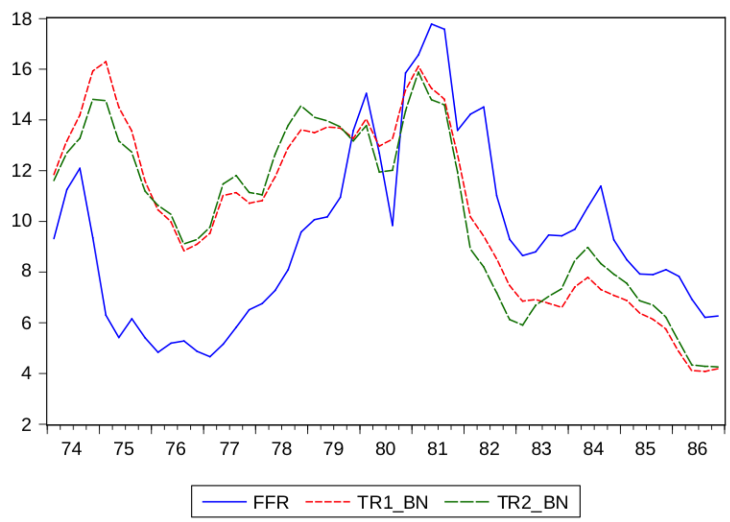

Figure 3.8: *Deviations from Baseline Policy Rule: 1987-2008 using HP OGs*

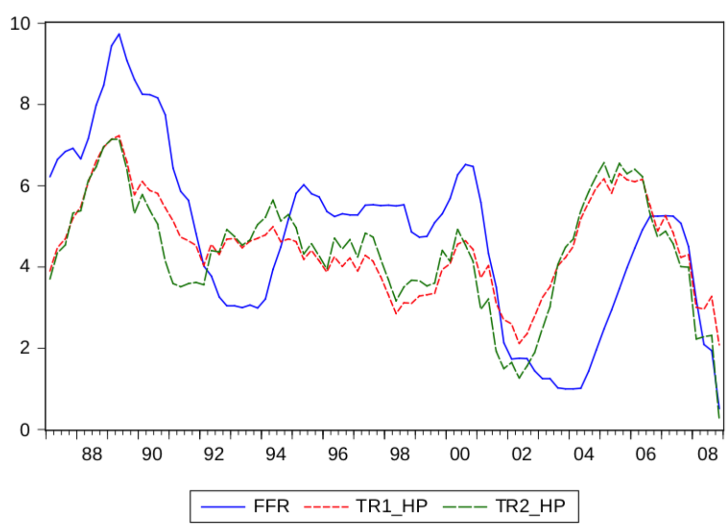

Figure 3.9: *Deviations from Baseline Policy Rule: 1987-2008 using BN OGs*

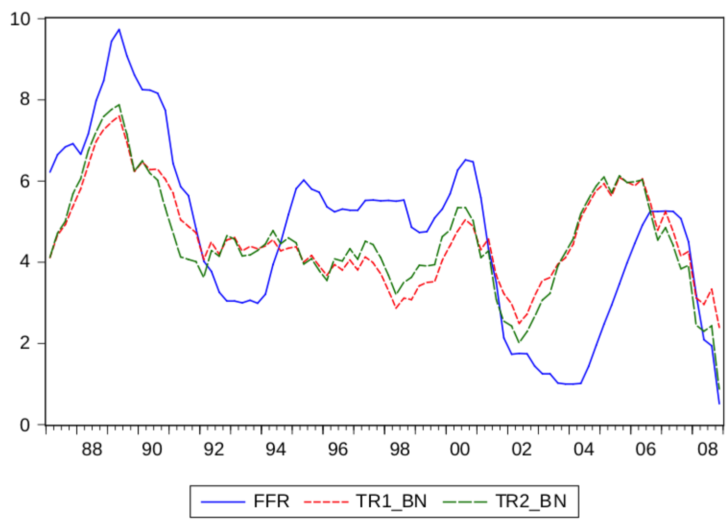

Figure 3.10: *Deviations from Baseline Policy Rule: 2009-2016 using HP OGs*

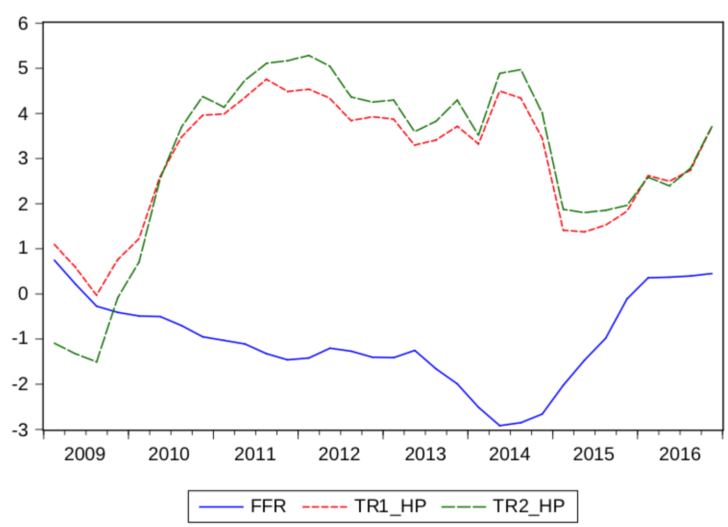

Figure 3.11: *Deviations from Baseline Policy Rule: 2009-2016 using BN OGs*

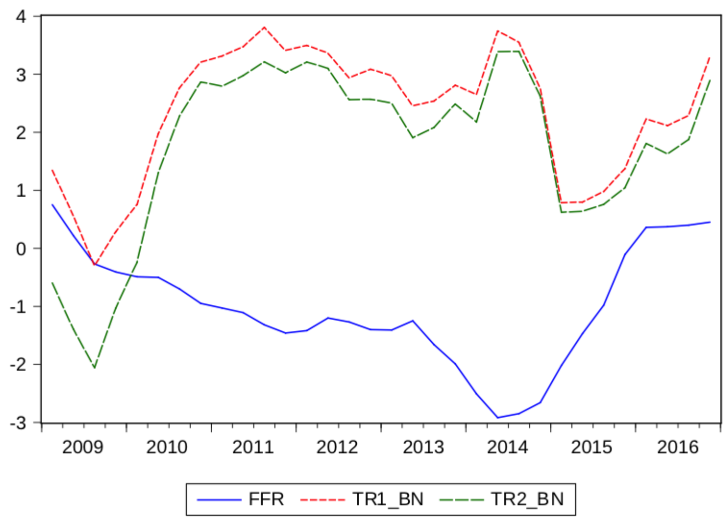

Figure 3.12: *Thick Modeling using BN Slack*

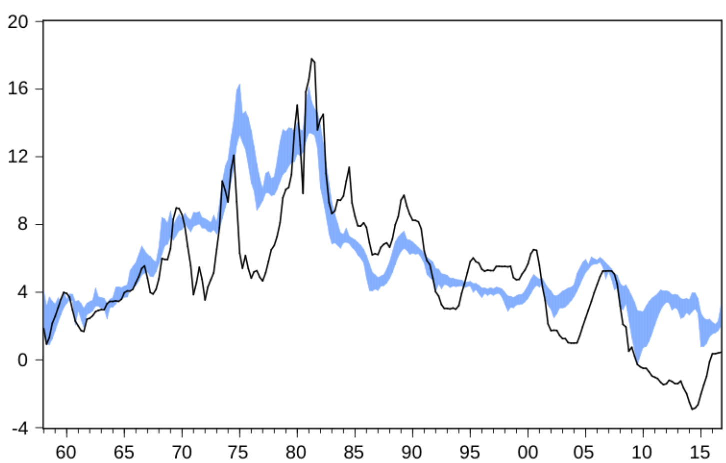

Table 3.1: *Output Gaps Created with HP Filter*

Variable	1958:1-1973:4	1974:1-1986:4	1987:1-2008:4	2009:1-2016:4
c	1.077***	6.605***	1.184*	-0.870
	(0.347)	(01.491)	(0.673)	(0.326)
λ	1.150	0.443	1.508*	-0.074
	(0.104)	(0.222)	(0.673)	(0.234)
γ	0.293***	-0.232	0.435**	-0.424
	(0.103)	(0.269)	(0.187)	(0.103)
R^2	0.677	0.188	0.278	0.402

Table 3.2: *Output Gaps Created with BN Filter*

Variable	1958:1-1973:4	1974:1-1986:4	1987:1-2008:4	2009:1-2016:4
c	0.901***	6.139***	1.178*	-1.334
	(0.312)	(1.381)	(0.606)	(0.417)
λ	1.115	0.511	1.515**	-0.226
	(0.087)	(0.210)	(0.239)	(0.244)
γ	0.555***	-0.162	1.118***	-0.517
	(0.117)	(0.359)	(0.226)	(0.160)
R^2	0.732	0.180	0.404	0.305

Table 3.3: *Bai and Perron Test for Multiple Structural Changes*

Era	Break Date	Coefficients	Average Deviation
1958Q1-1970Q4		γ_0 = 0.812	0.812
1971Q1-1979Q3	1971Q1	γ_1 = 3.350	4.162
1979Q4-1990Q4	1979Q4	γ_2 = -6.021	-1.859
1991Q1-2001Q3	1991Q1	γ_3 = 1.094	-0.765
2001Q4-2016Q4	2001Q4	γ_4 = 3.305	2.540

Chapter 4

Inflation Targeting in New Zealand: Does Practice Match Policy?

4.1 Introduction

On December 20th 1989, New Zealand made monetary policy history by becoming the first country to adopt a formal inflation targeting policy with the passage of the Reserve Bank of New Zealand Act. This Act established that the primary objective of the Reserve Bank of New Zealand is to maintain price stability. The purpose of this paper is to begin the process of uncovering the statistical effects that this policy has had on New Zealand's economy since its passage by exploring two introductory questions. One, is the Reserve Bank of New Zealand acting in accordance with their stated monetary policy goal of achieving medium run price stability? And two, is the New Zealand market responding to projected deviations from that monetary policy goal in accordance with expectations?

I start answering question one though the use of a set of Threshold Regressions. The purpose of these regressions is to compare the estimated response of the Reserve Bank of New Zealand's primary policy tool, the Official Cash Rate, to projected deviations from the target inflation rate over various periods of time. The RBNZ is known as one of the more transparent monetary authorities in the world, providing real time projections of the inflation rate for each quarter up to at least two years into the future. This provides the opportunity to study the time frame at which projected deviations from the inflation target are responded to most strongly. If the RBNZ is in fact systematically working to achieve price stability over the medium term as their mandate dictates, logic would suggest that the OCR should respond more strongly to deviations from the inflation target projected to occur in 2 years over those that are projected to occur next quarter.

Fortunately, the data support that pattern. My series of Threshold Regressions show that the OCR responds most strongly to large deviations projected to occur further in the future. When looking 8 quarters ahead for example, the OCR responds very strongly when the projected inflation rate is at least 0.4 percentage points greater than the official target rate, but it does not seem to respond at all to small deviations. But as the forecast horizon shortens, the strength of the OCR response to projected deviations from inflation not only diminish in amplitude, but it also changes its response pattern. In the short run, the OCR seems to respond strongly to smaller projected deviations from the inflation target while not responding at all to larger deviations. These conclusions support the position that the RBNZ has been acting systematically to achieve price stability in the medium term above concerns for the present.

The second question concerns the market reaction to these deviations. Are markets responses consistent with the behavior we see from the Reserve Bank? To answer this question, I make use of a set Trivariate VARs where I again utilize projected inflation deviations from target over various time periods to determine how the market is responding relative to various forecast horizons. The approach is also advantageous because it provides an opportunity to simultaneously double check the conclusions from the Threshold Regressions.

The set of Trivariate VARs run utilize an industrial production index, the New Zealand 90-day nominal interest rate, and one of the projected inflation deviation variables from the Threshold Regressions. Supporting the conclusion of the Threshold Regression, the 90 interest rate responds to projected deviations at longer forecast horizons much more strongly than it responds to deviations at shorter horizons, as expected. However, we do not necessarily see the Industrial Production Index behave as expected. In fact, the Industrial Production Index is actually estimated to temporarily increase as a result of a projected inflation deviation occurring two years in the future despite this being the forecast horizon to which the interest rate is responding most significantly. One explanation is that industry is responding more strongly to the projected inflation spike itself than they are reacting to the RBNZ's monetary tool. If firms do not believe that the RBNZ will raise rates to combat the future inflation projected to occur, those firms could be ramping up production before that anticipated inflation kicks in.

The rest of the essay proceeds as follows. Section 4.2 discuses New Zealand monetary policy history as it regards their inflation targeting agenda. Section 4.3 makes use of a set of Threshold Regressions to uncover which forecast horizon is most important in the RBNZ's selection of the Official Cash Rate. Section 4.4 Makes use of a set of VAR regressions to explore how firms are responding projected future inflation deviations from target. Section 4.5 Concludes.

4.2 Monetary Policy History in New Zealand And the Reserve Bank of New Zealand Act of 1989

I'll begin my brief preview of the history of the New Zealand economy in the 1970s. Since the effects of inflation targeting are the primary points of interest in this paper, I begin with 1970 because I want to communicate an understanding of what lead NZ to inflation targeting in the first place. The 1970s are significant because they marked the start of an extended period of time in which New Zealand was exposed to persistently high inflation fluctuating between 10 and 15 percent. As can be seen in Figure 4.1, displaying New Zealand's inflation rate between 1967 and 2016, New Zealand experienced a stretch between 1974 and 1987 in which the inflation rate was above 10% for all but 2 years. And while New Zealand might have never experienced hyper inflation, this period of persistently high inflation coincided with a period of low and volatile economic growth, leaving many New Zealanders with a particularly negative perception of inflation.

This lead to a great deal of support for broad economic reform in the back half of the 1980s. In addition to the Reserve Bank of New Zealand Act of 1989, New Zealand also implemented a floating currency, privatized many state owned enterprises, implemented public sector reform and welfare reform, and liberalization of the labor market and trade markets. All of these coinciding reforms resulted in a significantly more stable New Zealand economy beginning in 1990 (Bollard et al 1996). Beginning in the 90's, New Zealand has experienced significantly lower inflation and a GDP growth rate that is both higher and more stable than in the period preceding 1989.

But beyond all of these broad reforms to the structure of New Zealand's economy, it is the position of the Reserve Bank of New Zealand that "the best contribution monetary policy can make to the economy is to keep inflation low and stable" (Spencer et. al 2006). With the passage of the Reserve Bank of New Zealand Act of 1989, New Zealand began making strides to embrace this policy objective by becoming the first country to adopt an official target range for inflation. This target range is determined jointly by the Government and the Governor according to the Policy Targets Agreement and is currently set at an inflation target band of 1-3 percent on average over the medium term. But when the RBNZ Act of 1989 was first announced, the target band for inflation was set between 0-2 percent. After shifting the target several times in the earlier years, the target band has remained stable since September of 2002 with a 1-3 percent target range, as measured by headline

Consumer Price Inflation. The primary policy tool currently used to achieve this goal is the Official Cash Rate, first going into practice in March 1999.

Further, while the Reserve Bank of New Zealand Act of 1989 assures the RBNZ a certain degree of monetary policy operational independence, an additional defining characteristic of the act is the stringent demand for transparency from the RBNZ. As part of this effort in transparency, the RBNZ must publish 4 Monetary Policy Statements a year and must wait no longer than 6 months between publications. Further, a Select Committee of Parliament reviews each one of the RBNZ's Monetary Policy Statements and implicity reviews the RBNZ's handling of monetary policy. This means that while the RBNZ has operational independence, they also must provide explanations for any policy decisions which deviate from their central objective of inflation stability over the medium term. This is meant to instill strong confidence in the general public that the RBNZ will work systematically to achieve their stated monetary policy goal, an essential component of a successful targeting regime (Bernanke et al 1996)

4.3 Threshold Effects in Projected Inflation Deviations from Target

The first objective of this essay is to identify the time frame at which the Reserve Bank of New Zealand responds most strongly. The RBNZ's stated goal is to be within one percentage point of the 2% target inflation rate in the medium term.

One factor that makes New Zealand unique compared to most inflation targeting countries is the degree of transparency that the RBNZ employs; they provide their official inflation forecasts for at least 2 years additional post-dating the release of every Monetary Policy Statement. In this section, I use these inflation forecasts in a threshold regression to test the timeframe at which the RBNZ responds to projected deviations from the target inflation rate most strongly.

I first define Dev_i to be the projected i-quarter-ahead inflation rate deviation from the target inflation rate defined in levels. The variables are constructed so that Dev_i is positive when the projected inflation rate is above the target rate. In this notation, Dev_0 is the deviation of the realized inflation rate from the target rate. Dev_8 is the most forward looking projection considered. Using these variables, I first run a set of 9 simple threshold regressions in order to determine which time lag leads the RBNZ to respond most strongly. The regression equation is

$$OCR_t = c + \delta_j Dev_{t,i} + \epsilon_t \tag{4.1}$$

where OCR is the quarterly average of the monthly Official Cash Rate and δ_j is the regime j coefficient where I restrict the maximum number of possible regimes to 6. Denoting the number of breaks estimated to exist to be k, there then exists a set of increasing threshold values $\gamma_1 < \ldots < \gamma_{k+1}$ such that we are in regime j only if $\gamma_j \leq Dev_{t,i} \leq \gamma_{j+1}$. For any number of estimated breaks k, γ_1 will always be $-\infty$ and γ_{k+1} will be ∞.

Equation 4.1 is estimated using the methodology of Bai and Perron (1998), where the estimated number of breaks is determined using a test of L+1 vs. L sequentially determined thresholds, with 15% sample trimming. For a given estimated number of breaks k, estimation of the thresholds are achieved by comparing the Sum of Square residuals for all possible sets of k threshold values. Running this regression separately for each of the 9 Dev_i variables, I find that no regression estimates the existence of more than two regimes separated by a single threshold break. When estimating the equation using the 6-quarter-ahead and 7-quarter-ahead projected deviation variables, I find no significant break at all.

The results of the 9 estimations are provided in Table 4.1. For the equations finding a single significant threshold break, γ_2 is the estimated level of Dev_i at which the break exists. δ_1 is the Dev_i coefficient when $Dev_i < \gamma_2$ and δ_2 is the Dev_i coefficient when $Dev_i \geq \gamma_2$. δ is the Dev_i coefficient across the entire sample when no threshold break is found.

The most important result of note is that these estimated equations indicate that the Reserve Bank of New Zealand is responding most strongly to projected inflation rates in the medium term over those of the short term. Dev_8 is the only variable that results in the OCR responding more strongly to large shocks than to small shocks. At this 2 year forecast horizon, it is estimated that the RBNZ will raise the OCR by 3.6 percentage points for every additional percentage point that the projected 8-quarter-ahead inflation rate rises above the target inflation rate when that deviation is above 0.4. When Dev_8 is less than 0.4, the coefficient on Dev_8 is statistically insignificant with a p value of 0.5. At this forecast horizon, the RBNZ seems to respond very strongly to large projected positive deviations from the inflation rate target, while they seem to not respond to small or negative deviations at all.

Contrasting this to the 6-quarter ahead and 7-quarter-ahead estimations, no significant break is found. Further, the estimated coefficients on Dev_6 and Dev_7 are lower in amplitude than δ_2 from the Dev_8 equation. While this estimation indicates that the RBNZ responds to all 6-quarter-ahead and 7-quarter-ahead projected deviations significantly and equally, that response is weaker than the response to large deviations projected to occur at the 2 year horizon.

When considering all remaining projected and realized inflation target deviations, a significant threshold break is found for each. Further, these equations display a very clear pattern: the degree of response falls as the forecast horizon is shortened. Each of these regressions indicate that the OCR responds more significantly to smaller shocks than to large ones, with the δ_1 coefficient larger than the δ_2 coefficient for each regression. At these shorter forecast horizons, the RBNZ does seem to respond to small projected inflation deviations from target, but they respond much less so to large projected deviations from target. And the shorter the time horizon, the less significantly the RBNZ responds.

These results should not be entirely surprising. It is the stated purpose of the RBNZ to stabilize inflation over the medium term, so a stronger response to inflation deviations forecasted to occur in two years should take precedence over realized inflation deviations from target occurring in the current period. Comparing these threshold regressions to one another thus provides evidence that the RBNZ is in fact working towards this end.

4.4 A Trivariate VAR: Which Inflation Forecasts Matter?

I next conduct an unrestricted VAR analysis in order to identify potential effects of the realized and projected inflation rates. Specifically, I want to know the relationship between New Zealand's projected inflation rate at various time lags and their 90-day Interest Rate (IR_t) and Industrial Production Index (IP_t). To uncover this, I estimate a set of unrestricted VARs of the form

$$\begin{bmatrix} IP_t \\ x_t \\ IR_t \\ IP_{t-1} \\ x_{t-1} \\ IR_{t-1} \\ IP_{t-2} \\ x_{t-2} \\ IR_{t-2} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} a^1_{11} & a^1_{12} & a^1_{13} & a^2_{11} & a^2_{12} & a^2_{13} & a^3_{11} & a^3_{12} & a^3_{13} \\ a^1_{21} & a^1_{22} & a^1_{23} & a^2_{21} & a^2_{22} & a^2_{23} & a^3_{21} & a^3_{22} & a^3_{23} \\ a^1_{31} & a^1_{32} & a^1_{33} & a^2_{31} & a^2_{32} & a^2_{33} & a^3_{31} & a^3_{32} & a^3_{33} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} IP_{t-1} \\ x_{t-1} \\ IR_{t-1} \\ IP_{t-2} \\ x_{t-2} \\ IR_{t-2} \\ IP_{t-3} \\ x_{t-3} \\ IR_{t-3} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where x_t is a placeholder for the official New Zealand inflation rate ($Infl_t$) or any of the set of previously created deviation variables (Dev_i). This VAR(1) representation is the companion form of a VAR(3) composed of the three stated endogenous variables. This VAR(1) can likewise be depicted as

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}\mathbf{Y}_{t-1} + \mathbf{e}_t$$

where $var(\mathbf{e}_t) = \Sigma$ and the structural identification of shocks is based on the standard Cholesky Decomposition as in Sims (1980).

I first set $x_t = Infl_t$ as a baseline estimation exercise. Impulse Response Functions of one-standard-deviation shocks from this estimated model are presented in Figure 4.3. In this figure, Var1 is the Industrial Production Index, Var2 is the Inflation Rate, and Var3 is the 90 day Interest Rate. IRFs are presented with 95% confidence intervals. This paper is most concerned with the impact of the inflation rate on industrial production and the interest rate. As these figures show, there is virtually none. This comes at some surprise. In a vaccuum, we would expect that higher inflation would lead to raising interest rates, so the fact that all impact is wiped out within 3 months comes as some surprise.

However, we see a decidedly different reality when we replace the inflation rate with Dev_i in the VAR. This is done for each of the 9 Dev_i variables and the corresponding VAR is estimated. Figure 4.4 plots the Impulse Response Functions of a 1-standard-deviation shock to each of the Dev_i variables on the 90 day interest rate. Figure 4.5 presents a comparable grid showing the IRFs of a 1-standard-deviation shock to each of the Dev_i variables on the industrial production index.

Figure 4.4 tells a clear story. As the inflation forecast horizon rises, the impact of a projected deviation from target rises in both amplitude and persistence. The interest rate displays nearly no response to a deviation in the realized inflation rate from target. The interest rate will respond more strongly to projected inflation target deviations for 1-quarter-ahead and 2-quarter-ahead forecasts, but these response are still small and temporary.

Starting at the 3-quarter-ahead forecast, the impact of a 1-standard-deviation shock to the projected deviation of inflation from target appears to become permanent. However, the amplitude is still small for shocks to the 3-quarter-ahead, 4-quarter-ahead, and 5-quarter ahead projected deviations, none of which reach a maximum impact beyond 0.3 percentage points.

Over the medium term, the impact of a shock to the projected inflation deviation only amplifies. The IRF from the Dev_6, Dev_7, and Dev_8 all reach a peak of at least 0.4 and display an substantial degree of persistence. This matches the conclusions of the Threshold Regression analysis in finding that it is projected deviations of inflation from target occurring at a medium run time frame that the RBNZ is responding most strongly to. Note: the primary determinant of the 90 day interest rate is the OCR rate set by the RBNZ - both rates are provided in Figure 4.6 and display an extremely high correlation of 0.997. The 90 day interest rate was used because it predates the origin of the OCR which did not go into use until 1999.

When looking to the IRFs in Figure 4.5, things get a little more murky. When looking at the short run forecasts, the IRFs follow expectations closely. The IRFs found using the Dev_0, Dev_1, Dev_2, Dev_3, and Dev_4 variables all suggest that a one standard deviation shock to the short term projected deviations from inflation target will cause no immediate impact on the industrial production index as firms take time to adjust. However, they all display a delayed reaction in which the industrial production index falls over time, consistent with the theory that higher prices and corresponding interest rate increases will apply downward pressure on firms' ability to produce. Thus deviations from the inflation target that are expected to occur in the short run behave largely as expected.

But as we transition to the medium run, that relationship breaks down. As seen in Figure 4.5, the IRF obtained using the Dev_5 variable (a time horizon which we might well consider the dividing point of the short run and the medium run) displays virtually no impact on the industrial production index.

However, strong positive reactions are found for every 'medium run' time horizon, Dev_6, Dev_7, and Dev_8. This is surprising because the results from the Threshold Regressions indicated that these are the time horizons at which the interest rate responds most sensitively to projected deviations. Consistent with that finding, it would be normal to expect that industrial production rolls back in response to medium-run projected inflation deviations from target. We however see the direct opposite, industrial production seems to tick up. This perhaps suggests that inflation expectations are driving industrial production decisions to a greater extent than is the interest rate. These IRFs are consistent with a story that says firms see the projected medium-run inflation rate deviations from target and decide to tick up production before those price hikes kick in, with the corresponding interest rate increases applying only marginal downward pressure on production.

4.5 Conclusions

This paper set out to answer two key questions. Does the Reserve Bank of New Zealand consistently act in accordance with their stated primary goal of medium-run price stability? And how is the New Zealand economy responding to these RBNZ policies?

The first question can confidently be answered yes. Using sets of both Threshold Regressions and Trivariate VAR regressions, I show that the New Zealand 90-day nominal interest rate and the Official Cash Rate both respond more strongly to inflation deviations from target projected to occur over the medium term than they respond to projected deviations over shorter time horizons.

The second question is answered less completely. My set of Trivariate VARs do however provide a starting point. These regressions displayed a pattern wherein industrial production diminishes as a result of inflation being projected to exceed its target rate at short time horizons, but industrial production is estimated to temporarily increase when these target rate deviations are projected to occur over medium-run time horizons (1.5-2 years). This suggests that the interest rate plays a relatively smaller role in determining industrial production compared to the inflation rate itself. When inflation is near, it puts immediate downward pressure on the ability of firms to produce. But when it is projected to occur 18+ months out, firms find it profitable to temporarily increase production before that projected future inflation becomes realized.

In all, this paper is meant to serve as a jumping off point for studying the effects of inflation targeting policies in New Zealand, the world's first official inflation targeter. This paper shows that the Reserve Bank of New Zealand is acting systematically to achieve their stated monetary policy goal of medium run price stability. However, this paper also shows that New Zealand firms may not be as concerned with the RBNZ's primary policy tool as they are with the actual inflation rate and the RBNZ's projections for its future.

4.6 References

Bai, J. and Perron, P. (1998) "Estimating and Testing Linear Models with Multiple Structural Changes", Econometrica, 66, 47-78

Ball, L. and Sheridan, N. (2004) "Does inflation Targeting Matter?", The Inflation Targeting Debate, University of Chicago Press

Bernanke, B., Laubach, T., Mishkin, F., Posen, A. (1999) "Inflation Targeting. Lessons from Internation Experience", Princeton University Press

Bernanke, B. and Mishkin F. (1997) "Inflation Targeting: A New Framework For Monetary Policy?", Journal of Economic Perspectives, 11, 97-116

Bollard, A., Lattimore, R., Silverstone, B. (1996) "A Study of Economic Reform: The Case of New Zealand", Elsevier, 518 pages

McDermott, J. (2017) "The Value of Forecasting in an Uncertain World", Reserve Bank of New Zealand Speeches

Odior, E. (2016) "Inflation Targeting in an Emerging Market VAR and Impulse

Response Function Approach", European Scientific Journal, vol. 8, No. 6

Sims, C. (1980) "Macroeconomics and Reality", Econometrica, 48, 1-48

Spencer, G., Bollard, A., Hodgetts, B., Hunt, C., and Reddell, M. (2006) "Inflation Targeting: The New Zealand Experience and Some Lesson", Reserve Bank of New Zealand Speeches

Svennson, L. (1997) "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets", European Economic Review, 41, Issue 6, 1111-1146

4.7 Tables and Figures

Figure 4.1: *New Zealand Historical Inflation Rate*

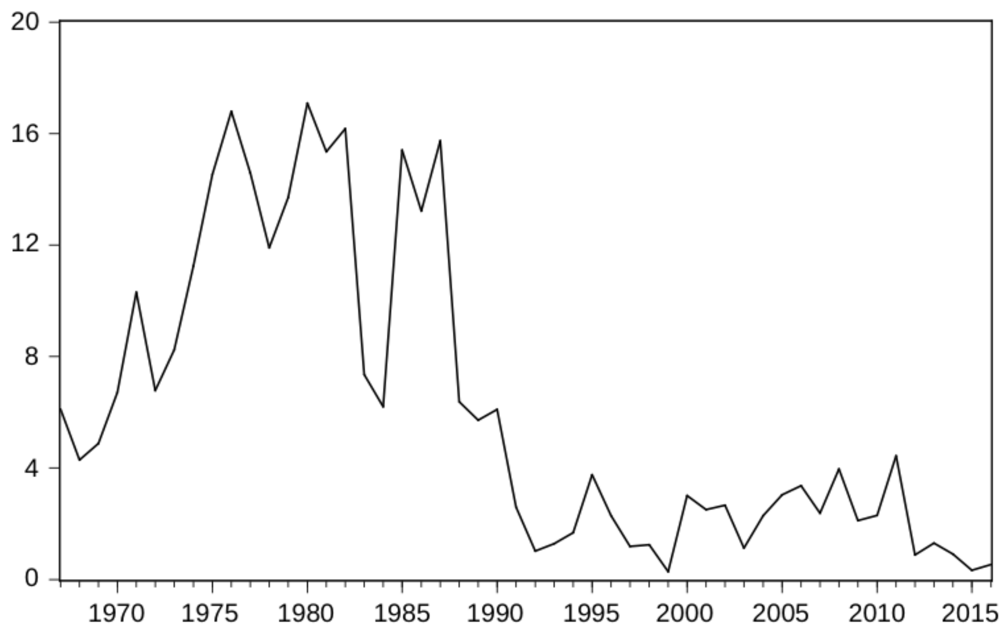

Figure 4.2: *New Zealand Inflation Rate, Target Inflation Rate, and Target Bands*

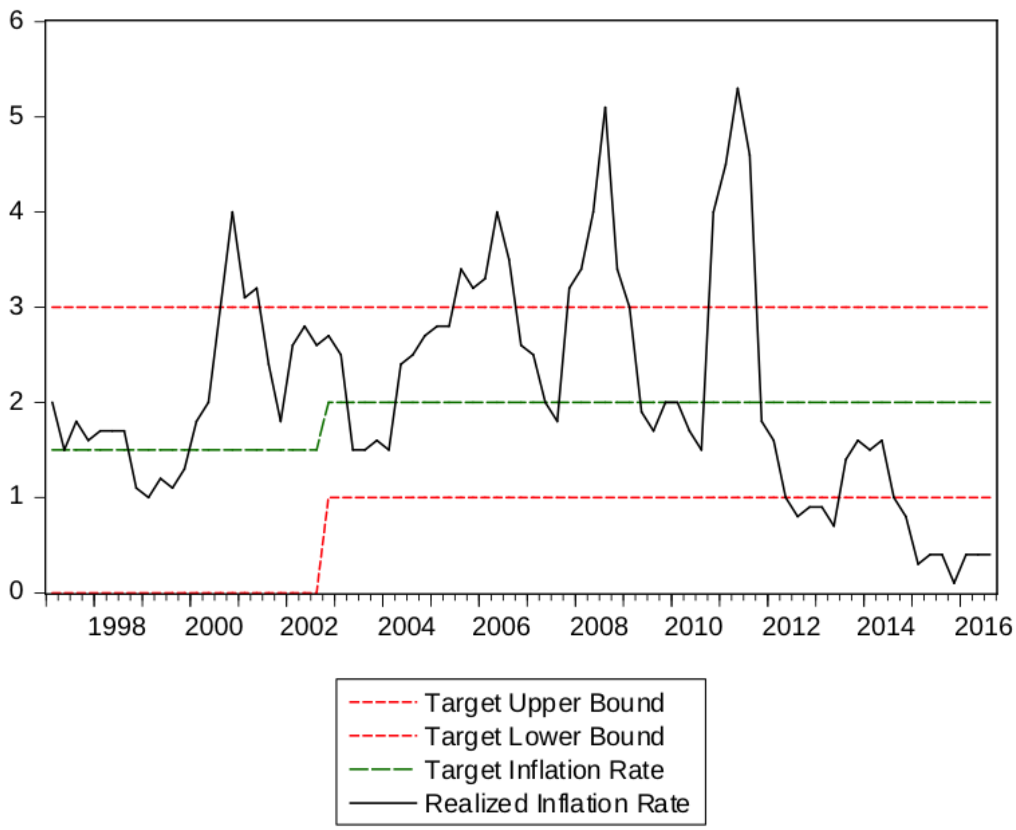

Figure 4.3: *Impulse Response Functions from Unrestricted VAR*

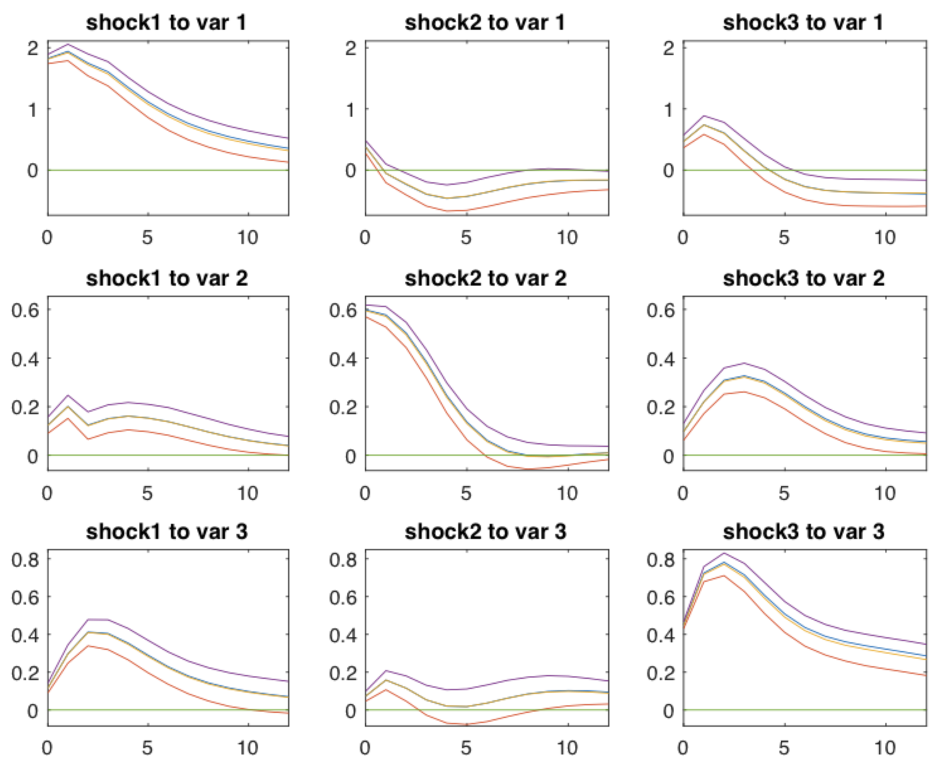

Figure 4.4: *IRFs: Impact of Projected Inflation Rate Deviations from Target on 90-Day Interest Rate with 90% Confidence Intervals*

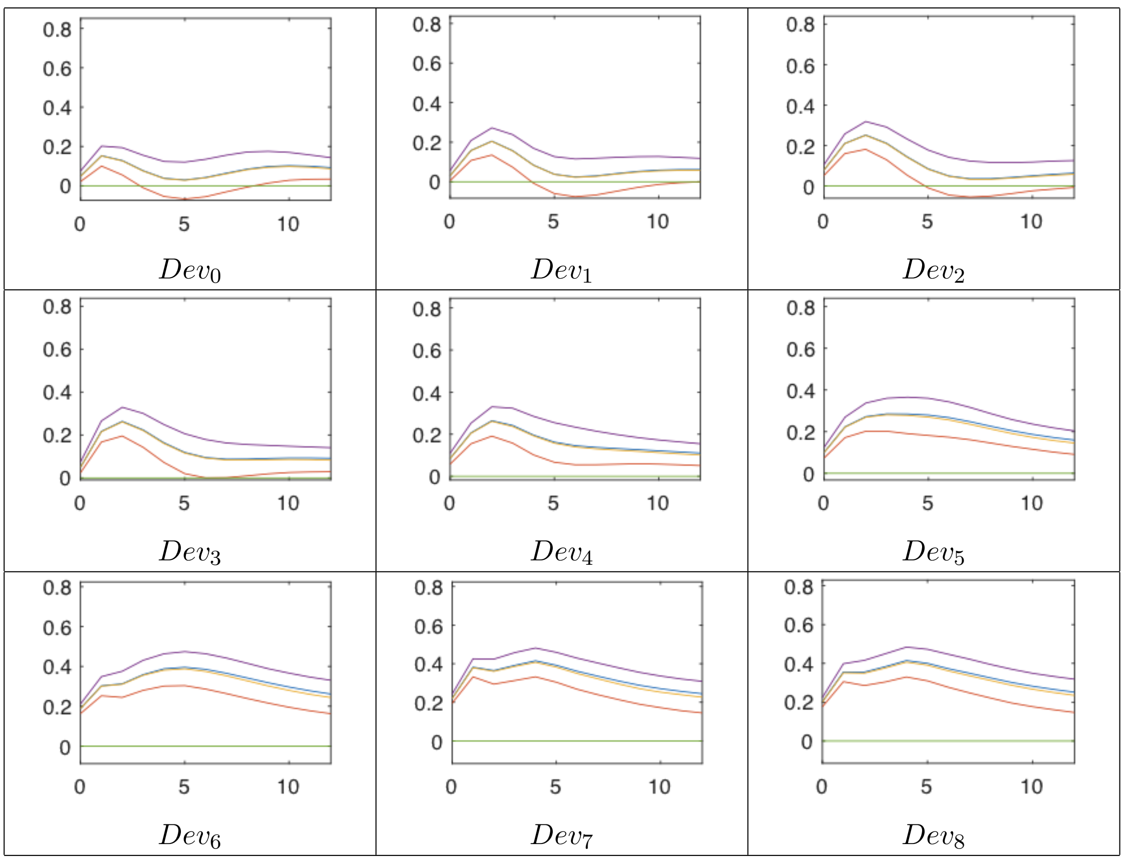

Figure 4.5: *IRFs: Impact of Projected Inflation Rate Deviations from Target on Industrial Production Index with 90% Confidence Intervals*

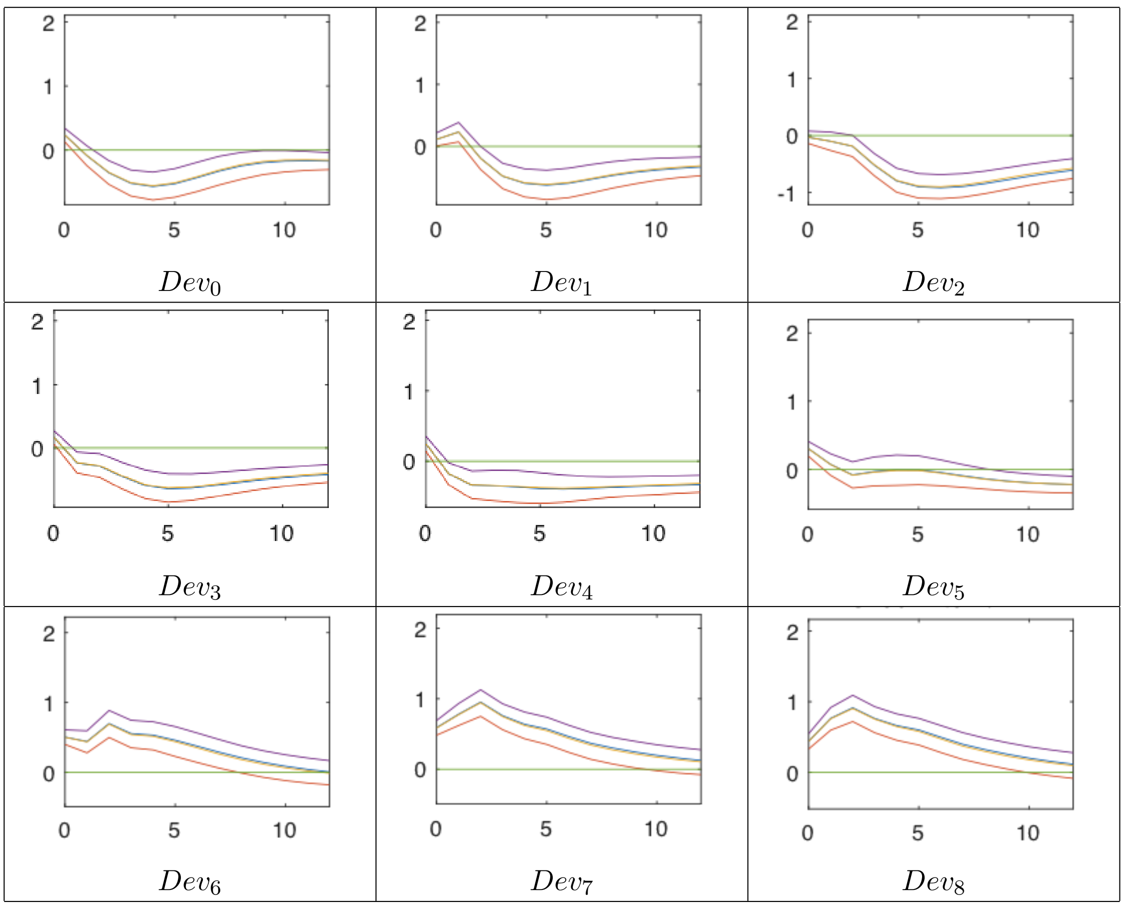

Figure 4.6: *The OCR and 90-day Nominal Interest Rate*

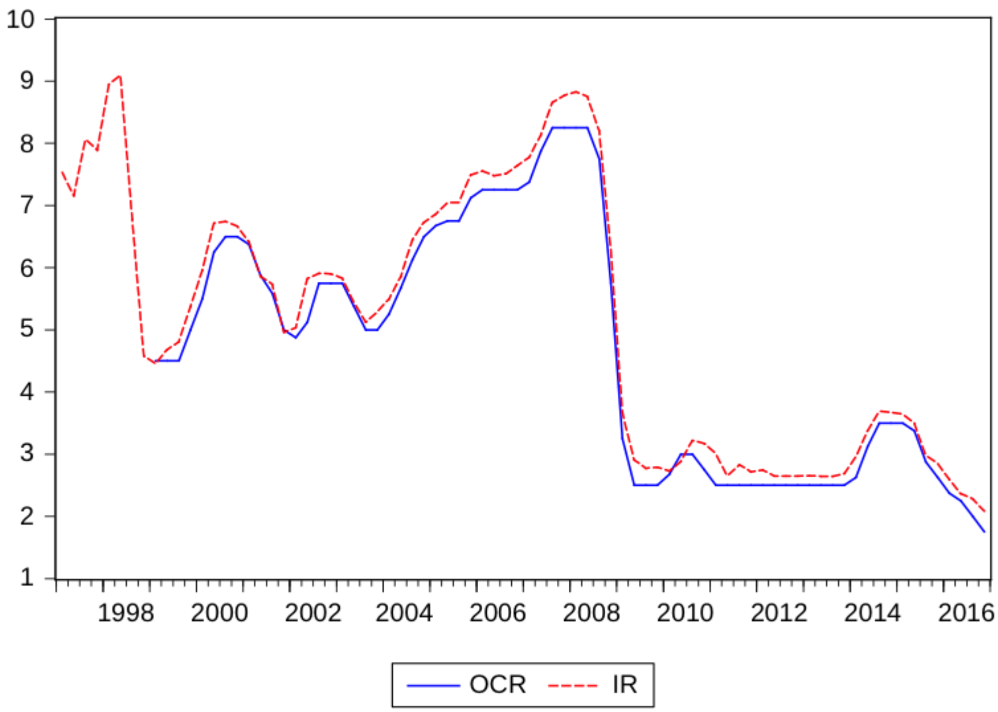

Table 4.1: *Projected Inflation Deviation From Target Threshold Regressions*

	c	γ_2	δ_1	δ_2	δ
Dev_0	4.72	1.5	1.491*** (0.219)	0.213 (0.217)	N/A
Dev_1	4.53	1.6	1.482*** (0.233)	0.330 (0.228)	N/A
Dev_2	4.49	1.9	1.631*** (0.254)	0.357 (0.232)	N/A
Dev_3	4.39	1.3	1.859*** (0.343)	0.643** (0.266)	N/A
Dev_4	4.35	1.0	2.922*** (0.476)	0.920*** (0.291)	N/A
Dev_5	4.30	1.0	3.223*** (0.468)	1.365*** (0.353)	N/A
Dev_6	3.95	None	N/A	N/A	3.005*** (0.386)
Dev_7	3.81	None	N/A	N/A	3.119*** (0.469)
Dev_8	3.72	0.4	-0.997 (1.478)	3.604*** (0.250)	N/A