uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Firouz Fallahi

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Economics)

GRADE / DEGREE

Department of Economics

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Three Essays on Applied Time Series Econometrics

TITRE DE LA THÈSE / TITLE OF THESIS

Gabriel Rodriguez

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Serge Coulombe

Marcel Voia

Lynda Khalaf

Greg Tkacz

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Three Essays on Applied Time Series Econometrics

A thesis

submitted to the School of Graduate Studies and Research of the University of Ottawa
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in

Economics

by

Firouz Fallahi

September, 2007

Your file Votre référence
ISBN: 978-0-494-49346-5
Our file Notre référence
ISBN: 978-0-494-49346-5

To my beloved wife, Najibeh,
my daughter, Tara,
and my son, Yashar.

Acknowledgements

I am deeply indebted to Professor Gabriel Rodríguez, my thesis supervisor, for his patience, inspiring advice, support, and guidance during the preparation of my thesis. Without his help, this work would not be possible. His advice and rich knowledge of econometrics opened my eyes to a new world of econometrics.

I would like also express my thanks to Professors Lynda Khalaf, Marcel Voia, Serge Coulombe, Gamal Atallah, and Dr. Greg Tkacz for their constructive and valuable comments and suggestions.

Also many thanks go to the staff of the Department of Economics at the University of Ottawa for their invaluable support.

Last but not least, I am forever thankful to my wife who provided consistent support in the good times and the bad.

Table of Contents

Abstract

This dissertation consists of three independent essays in applied econometrics. **Essay One: Unemployment and Criminality: Using Markov-Switching Models to Identify a Link.** This paper studies the existence of a relationship between the unemployment rate and crime in the U.S. Using Markov Switching Autoregressive models the behaviour of four crime variables and unemployment rate during the period of study is investigated and different regimes for each variable are determined. Using some non-parametric measures such as the Concordance Index (Harding and Pagan 2002) and Independence of Chronologies (Bodman and Crosby 2005), among others, the independency of cycles of unemployment rate and crime variables is tested. The results of this phase show that there is no relationship between the unemployment rate and burglary and motor. However, for larceny and robbery the results are mixed. At the second phase, Markov Switching Vector Autoregressive models are also used to determine the states for both unemployment rate and each of the crime variable simultaneously. The results of this stage show that the effect of unemployment rate on larceny and motor depends on the state of the variables. For larceny this effect is either positive or null, and for motor it fluctuates between negative, null, and positive. Also the result shows that regardless of the state of the variables, the effect of unemployment on burglary and robbery is negative and null, respectively.
Essay Two: Convergence in the Canadian Provinces: Evidence using Unemployment Rate. Quarterly time series data from Canada and the Canadian provinces for the period 1976:1-2005:3 are examined to determine if the unemployment rates in the Canadian provinces are converging to the national rate of unemployment. The paper checks for existence of stochastic convergence using recent unit root statistics, see Perron and Rodríguez (2003) and Rodríguez (2007); furthermore, it verifies the existence of $\beta-$convergence using recently proposed methods by Vogelsang (1998), Perron and Yabu (2005, 2006), and Bai and Perron (1998, 2003).

Results from different unit root tests, without and with structural breaks, confirm that stochastic convergence holds in all provinces except British Columbia. Next, the existence of β-convergence is tested using different approaches to estimate the trend function. The results show that deterministic convergence holds and the unemployment rates of the Canadian provinces are converging to the unemployment rate

of Canada. This conclusion is stronger when multiple breaks are allowed in the trend function, that is, using the approach of Bai and Perron (1998, 2003).

Essay Three: Persistence in Unemployment: Evidence from Canada and the Canadian Provinces. This paper addresses the persistence of unemployment in Canada. To test this issue, we use quarterly unemployment rate data from Canada and the Canadian provinces from 1976-2005. First, we utilize different unit root tests, without and with structural breaks, and confidence intervals for the autoregressive coefficients, to test the null hypothesis of unit root. The results of these tests show that the unemployment rates in Canada and the Canadian provinces, except Nova Scotia and Quebec, do not have a unit root and are therefore, not persistent.

Next, we check for stability of the persistence property of the data, using the Bai and Perron approach (1998, 2003), which allows for multiple breaks in data. The results from this section show that the persistence is not constant and is regime dependent. In addition, except for the unemployment rate in British Columbia, which is persistent in all regimes, the other series are not persistent in most regimes.

Introduction

This dissertation is a collection of three independent essays in applied econometrics. The first essay examines the relationship between the unemployment rate and crime in the U.S, using Markov Switching models. The second and third essays deal with the unemployment rate in Canada and the ten Canadian provinces. The second essay examines the issue of convergence in the Canadian provinces and the third investigates the issue of persistence in the unemployment rates of Canada and the ten Canadian provinces.

Economic theory anticipates the existence of a positive relationship between the unemployment rate and crime. However, results from past studies are mixed. Several studies conclude that there is a positive effect of unemployment rate on crime; simultaneously, other investigations show that this effect is negative and even null. In the first essay, we investigate the relationship between the unemployment rate and four types of crime in the U.S. These crimes are burglary, larceny, motor-vehicle theft, and robbery. Past studies tried to test the existence of a relationship between the unemployment rate and crime, directly using parametric estimations through multiple regression, vector autoregression, and cointegration models. In the first essay, we attempt to identify the link between the unemployment rate and crime not only directly from the estimated coefficients in the models, but also indirectly, from cycles of these variables. First, we use Markov Switching Autoregressive (MSAR) models to investigate the behavior of each time series, the crime variables and the unemployment rate. The results from these models show the cycles of each variable during the time span under study, which can be used to determine different regimes for each variable. These regimes represent the periods of high and low crime and the unemployment rates. Next, different indices can be used to identify the synchronization of the cycles of the unemployment rate and crimes, (that is, to see whether or not the cycles of the unemployment rate and the cycles of each crime are independent from each other). We use three different methods to test the independence of these cycles, and find that the unemployment rate has a significant relationship with robbery and larceny, but has no effect on motor-vehicle theft or burglary. Furthermore, the relationship between robbery and the unemployment rate and its lagged values, is positive. At the same time, larceny has a negative relationship with the current rate of unemployment; however, the lagged values of the unemployment rate has a positive effect on larceny.

Indeed, these results still ignore one possibility about the relationship between unemployment and crime. It is possible that the unemployment rate has different effects on crime, in different regimes of the unemployment rate. For example, a significant relationship between the unemployment rate and crime might exist only when the unemployment rate is high and there might be no relationship when the unemployment rate is low. To address this possibility in our study, we try to reinvestigate the link between the unemployment rate and crime, using Markov Switching Vector Autoregressive (MSVAR) models. These models allow us to determine different cycles for the unemployment rate and each one of the crime variables, simultaneously. The results from the MSVAR models show that the impact of the unemployment rate on larceny and motor-vehicle theft depends on the regime of the variables. The effect for larceny is either positive or null, and for motor-vehicle theft, it fluctuates between negative, null, and positive. In addition, regardless of the regime, the influence on burglary and robbery is negative and null, respectively.

The second essay examines regional disparities and convergence in the Canadian provinces. Regional disparities have been in the center of policy makers concern for many years since these disparities are capable of creating imbalances in the development process of any country and of creating political problems. The disparity problem is important in any country; however, it is more significant in diversified countries, like Canada. As Gunderson (1996, p.1) stated, "[b]eing a geographically large country with substantial heterogeneity in resource endowments and accesses to markets, regional issues have always been important in Canada".

Convergence is a simple notion, which comes from neoclassical growth models for a closed economy.[1] In these models, if the regions only differ from each other in the initial level of per-capita income and capital, they will reach the same steady-state level. This is because the productivity of capital in the poor regions is higher than in the rich regions. Therefore, the underprivileged regions will experience a higher growth rate than the prosperous areas. If the poor regions grow faster than the rich do, they will eventually catch-up, and the disparity gap will shrink and eventually close. On the other hand, if the growth rates of the rich regions are higher than that of poor regions, the disparity gap will increase over time. In this case, we say that these regions are diverging from each other. This catch-up concept of convergence is known in the literature, as deterministic convergence. Stochastic convergence is another form of convergence. Simply, stochastic convergence exists if the disparity variable is stationary. That is, shocks will have temporary effects on the disparity and it will return to its original level.

Disparities can be defined using different variables, such as the relative per-capita income, per-capita earning, wage, unemployment rate, and so forth. In essay two, we use the provincial unemployment rates of Canada and the ten Canadian provinces to study the disparity issue. We use several recently developed methodologies to investigate the convergence, or divergence, in the unemployment rates of these provinces. We use different unit root tests, without and with structural breaks, to test for stochastic convergence. Furthermore, we use

---

[1] see Solow (1956) for example.

various methods to estimate the trend function, to address the deterministic convergence in data. These methods are proposed by Vogelsang (1998), and Perron and Yabu (2005, 2006), which allow for structural breaks in data. The results of this essay show that stochastic convergence holds for all the Canadian provinces, except British Columbia. In addition, we find deterministic convergence in these provinces, which indicates that the unemployment rates of the Canadian provinces are converging to the national rate of unemployment. This conclusion is stronger when multiple break points are allowed in the trend function.

The last essay of this dissertation examines the persistence of the unemployment rate. The degree of persistence is the key factor in distinguishing the traditional theory of the unemployment rate, i.e. the natural unemployment rate theory of Friedman (1968) and Phelps (1967), from the hysteresis theory of unemployment rate (Blanchard and Summers, 1987). According to the former theory, the unemployment rate is stationary about its mean. However, the latter theory suggests that the unemployment rate is persistent and is not stationary. The hysteresis theory of unemployment is developed to explain persistently high unemployment rates in 1970s and 1980s, especially in Europe. In addition, it is very important for policy makers to know whether the unemployment rate is persistent. When the evidence of persistence is found, any interference in the labor market that changes the unemployment rate will affect it for a lengthy period. Alternatively, when the unemployment rate is not persistent, the effect of this interference will quickly disappear and it will return to its original starting point, therefore, the interference will not be justified.

Existence of persistence in the unemployment rate can be studied using the root of the time series data or using the autoregressive coefficients. When the series has a root equal to unity, then it will be nonstationary and will support the hysteresis theory. This is also true when the autoregressive coefficient or the sum of the autoregressive coefficients, is larger than or equal to one. Unit root tests and confidence intervals for the largest autoregressive coefficient, or the sum of the autoregressive coefficients, can be utilized to test the integration degree of the data. If any time series is integrated, of order one, $I(1)$, or the confidence interval of its autoregressive root(s) contains unity, then it is not stationary and any shock to this data will have a permanent effect. This type of data will not return to its mean, in other words, it is a persistent time series. We use these tools to study the existence and degree of persistence in the unemployment rates of Canada and the ten Canadian provinces. The results show that the unemployment rates in Canada and provinces are not persistent, the exception being the unemployment rates in Nova Scotia and Quebec.

Next, an interesting question regarding these unemployment rates is whether their persistence property is stable or has changed over time. We use the approach of Bai and Perron (1998, 2003) to answer this question. This approach allows for multiple break points in data.

Results show that all series, except the unemployment rate in Quebec, have at least one structural break. In addition, persistence is regime dependent. Furthermore, except for the unemployment rate in British Columbia, which is persistent in all regimes, the other series are not persistent in most regimes.

References

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

[2] Bai, J., and Perron, P., (2003), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics* **18**, 1-22.

[3] Blanchard, O. J. and Summers, L. H., (1987), "Hysteresis in Unemployment," *European Economic Review*. **31**, 288-95.

[4] Friedman, Milton, (1968), "The Role of Monetary Policy," *American Economic Review* **58**, 1–17.

[5] Gunderson, morley (1996),"Regional Productivity and Income Convergence in Canada Under Increasing Economic Integration," *Canadian Journal of Regional Science* **19 (1)**, 1-23.

[6] Perron, P. and Yabu, T. (2005), "Testing for Shifts in Trend with an Integrated or Stationary Noise Components," Working Paper,Department of Economics, Boston University.

[7] Perron, P. and Yabu, T. (2006), "Estimating Deterministic Trends with an Integrated or Stationary Noise Components," Working Paper, Department of Economics, Boston University.

[8] Phelps, E.S., (1967), "Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time," *Economica* **34**, 254–81.

[9] Solow, Robert M., (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*. **70**, 65-94.

[10] Vogelsang, T.J. (1998), "Trend Function Hypothesis Testing in the Presence of Serial Correlation," *Econometrica*.**66**, 123-48.

Chapter 1
Unemployment and Criminality: Using Markov-Switching Models to Identify a Link

1.1 Introduction

Although economic theory anticipates the existence of a positive relationship between unemployment and crime, empirical works in this regard found mixed results. Chiricos (1987) reviewed 68 studies about the relationship between crime and the unemployment rate and found that only less than half of these studies have found positive significant effects of the unemployment on crime rates. That is, most of these studies show a negative or no relationship between crime and unemployment.

Some authors argue that the source of disagreement among the results comes from the type of data that has been used; see Kapuscinski, Braithwaite, and Chapman (1998).

Cook and Zarkin (1985) presented an analysis of the business cycle and its impact on homicides, robbery, burglary, and auto theft using U.S. data for the period 1933-1981. Using a nonparametric test based on the changes in criminal activity during the entire business cycle, they showed that an increase in robbery and burglary is higher during economic contractions than expansions. Furthermore, more auto theft occurs during expansions relative to contractions. Using other set of tools, they found a positive relationship between the unemployment rate and robbery and burglary but this effect was negative for auto theft.

Hale and Sabbagh (1991) investigated the effect of unemployment on eight types of crime in England and Wales. They found that there is a significant relationship between unemployment and five kinds of crime. Their results also show that there is no relationship between unemployment and auto theft. Property crimes including theft, burglary, and robbery had positive relationships with changes in the unemployment rate. Robbery had negative relationship with changes in the unemployment rate at the previous period.

Using age-specific arrest rates and an age-specific unemployment rate, during 1958-1995, Britt (1997) found that unemployment has a negative effect on homicide and aggravated assault for the younger age groups and a positive effect for older age groups. Witt, Clarke and Fielding (1999) found a positive relationship between crime and the male unemployment rate.

Entrot and Spengler (2000) used panel data methodology to study the effect of socio-economic factors on crime. They used two data sets, one only for West Germany and the other for unified Germany. The result of the first data set shows that the effect of unemployment is small, often insignificant, with ambiguous signs. On the other hand, the results from a unified Germany (1993-1996) indicate that the impact of unemployment becomes higher and unambiguously positive.

Greenberg (2001) used time series data, cointegration, and error correction models to investigate the relationship between divorce rates and unemployment on robbery and homicide rates in U.S. He concluded that lagged values of unemployment and unemployment duration have a negative effect on robbery. His finding is consistent with the finding of a previous study by Cantor and Land (1985). Raphael and Winter-Ebmer (2001) analyzed U.S. data and their findings show that the unemployment rate has a significantly positive effect on property crimes, but not on the violent crimes.

The same conclusion is obtained by Levitt (2001) using a state-level panel of annual data for the period 1950-1990 in U.S. Imrohoroglu et. al. (2004) analyzed the trends in the aggregate property crime rate in the U.S. for the period 1975-1996, using a dynamic equilibrium model. They tried to investigate the factors that determine the pattern of this kind of crime. They found a negligible effect of the unemployment rate on crime. Edmark (2005) using a panel of Swedish counties over 1988-1999 showed that there is a strong positive relationship between the unemployment rate, burglary, car theft and bike theft.

Recently Lee and Holoviat (2006) used the cointegration approach to identify a long run relationship between unemployment and a set of crime variables in three Asian-Pacific countries. They found a long-run relationship, in particular between unemployment of young males and crime.

Most of these studies were carried out using multiple regression models, vector autoregression or error correction models. All of these methods assume a stable behavior of the variable under examination. However, the unemployment rate directly related to business cycles and has a cyclical pattern so linear models may provide a weak fit. Consequently, one needs to use non-linear models, such as the Markov Switching (MS) models (Hamilton, 1989).These models have the capability to capture changes in the behavior of time series by allowing the switching between regimes or states. MS models have been used widely in the literature and have proven their ability to explain the changes in pattern of time series. Hamilton (1990) introduced the expectation maximization (EM) algorithm for estimating the parameters of these kind of models. Kim (1994) extended the concept of the MS model to state-space models and suggested an approximate smoothing algorithm. Albert and Chib (1993) proposed Gibbs sampling methods. Some authors tried to extend the original model proposed by Hamilton (1989) by including time-varying or duration dependent transition probabilities. For example, Durland and McCurdy (1994) studied the duration dependent case and Filardo and Gorden (1998) estimated transition probabilities as a function of exogenous information.

This paper has two goals. The first goal is to study the behavior of different types of crime and the second goal deals with the potential relationship between the unemployment rate and the different crime rates in the U.S. Our sampling period covers 1975:1 until 2004:4 and four types of crimes are analyzed: burglary, larceny, motor-vehicle theft (motor), and robbery.

The Markov-Switching methodology is used to investigate the behavior of the unemployment rate and crime variables and link them, using the data above mentioned. The remaining portion of this paper is organized in two phases. The first phase consists in using MS-AR models to investigate the behavior of the crime variables and the unemployment rate during the period of study and determine different regimes for each variable. Next, using some non-parametric measures such as the Concordance Index (Harding and Pagan 2002) and the Independence of Chronologies (Bodman and Crosby 2005), the independency of cycles of the unemployment rate and crime variables is tested. In the second phase, the Markov-Switching Vector Autoregressive models will be used to determine the states for both the unemployment and crime variable simultaneously. These models will also be used to investigate the effect of unemployment on crime variables at different regimes of these variables. The reason for using this method is the following: the unemployment rate may influence the crime rates differently in separate states and disregarding this fact may cause the mixed results reported in the previous studies.

1.2 Methodology

The switching model was motivated by Quandt (1972) and Goldfeld and Quandt (1973). It has been popularized by Hamilton (1989) with an application to business cycles. In the switching model introduced by Quandt (1972) the switching mechanisms are independent from each other, whereas in the switching model of Quandt and Goldfeld (1973) and Hamilton (1989), switching is governed by a first order Markov chain.

Consider the case of a single[2] stationary variable y_t, and assume that the realized values of this variable for $t = 1, 2, ..., T_1$ can be described with a first-order autoregression:

$$y_t = c_1 + \rho_1 y_{t-1} + \epsilon_t \tag{1.1}$$

where $\epsilon_t \sim N(0, \sigma^2)$. Now assume that there is a jump or structural change in this variable at time T_1 which leads to the model

$$y_t = c_2 + \rho_2 y_{t-1} + \epsilon_t \tag{1.2}$$

for $t = T_1{+}1, T_1{+}2, ..., T$. These two models could be compressed into one equation as follows

$$y_t = c_1 + \rho_1 y_{t-1} + \delta D_t + \gamma D_t y_{t-1} + \epsilon_t, \tag{1.3}$$

where the variable D_t takes values of 0 and 1 for $t \leq T_1$ and $t > T_1$, respectively, or alternatively

$$y_t = c_{s_t} + \rho_{s_t} y_{t-1} + \epsilon_t, \tag{1.4}$$

where s_t takes values of 1 and 2 before and after a change in y_t, respectively. In other words, the period $t \leq T_1$ is represented by $s_t = 1$ and the period after jump, $t > T_1$ by $s_t = 2$.

---

[2] Although the case of a single variable is discussed here, the framework may be easily extended for more than one variable.

Nevertheless, these models have three shortcomings. First, the exact date of the jump has to be known a priori to be able to use the dummy variable. Second, it is not satisfactory to use this model to forecast the behavior of y_t. Third, s_t is treated as a deterministic variable and is perfectly predictable, which is not a realistic assumption.

Therefore, to solve these problems and to complete the data generating process one needs to include a description of the probability law for s_t. In the MS model, the switching mechanism is controlled by an unobservable state variable called s_t. This state variable follows a first order Markov chain. In other words, the value of this state variable at period t depends only on its value at period $t-1$.[3]

The switching model for y_t is given by

$$y_t = \left\{ \begin{array}{l} c_1 + \rho_1 y_{t-1} + \epsilon_t, \ s_t = 1 \\ c_2 + \rho_2 y_{t-1} + \epsilon_t, \ s_t = 2 \end{array} \right\}.$$

This model shows two different dynamic structures, depending on the value of the state variable s_t. Different assumptions about s_t generate different models. When $s_t = 1$ for $t = 1, 2, ..., T_1$ and $s_t = 2$ for $t = T_1 + 1, T_1 + 2, ..., T$, this model is nothing but a model with a single structural change at time T_1. When s_t are independent Bernoulli random variables, this model corresponds to the random switching model of Quandt (1972). If s_t is considered as an indicator variable such that $s_t = 1$ or $s_t = 2$ respectively for $\phi \leq c$ and $\phi > c$, where c is the threshold value, this model becomes a threshold model. When s_t follows a Markov process this model is the MS model.

In the MS model, the properties of y_t are determined jointly by the characteristics of ϵ_t and the state variable s_t. The state variable generates random and frequent changes in the pattern of the model. To describe the full dynamics of the variable, a probabilistic description of how the variable s_t moves from one state to another is needed. First order Markov chain implies that these probabilities are:

$$\Pr[s_t = j | s_{t-1} = i, s_{t-2} = k, ...; y_{t-1}, y_{t-2}, ...] = \Pr[s_t = j | s_{t-1} = i] = p_{ij}.$$

---

[3] Since the realization of the state variable is not directly observable, sometimes these models are called Hidden Markov models.

The transition between states or regimes can be represented using a transition probability matrix. For a simple model, which has only two regimes, this matrix is

$$\begin{aligned} P &= \begin{bmatrix} Pr(s_t = 1|s_{t-1} = 1) & Pr(s_t = 2|s_{t-1} = 1) \\ Pr(s_t = 1|s_{t-1} = 2) & Pr(s_t = 2|s_{t-1} = 2) \end{bmatrix} \\ &= \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \end{aligned}$$

where p_{ij} $(i, j = 1, 2)$ indicates the transition probabilities of $s_t = j$ given that $s_{t-1} = i$ and $p_{i1} + p_{i2} = 1$.

There are different ways to estimate the MS models. Examples include, the maximum likelihood estimation (MLE) of Hamilton (1989), the expectation maximization (EM) algorithm proposed by Hamilton (1990), and the Gibbs sampling approach of Albert and Chib (1993). See also Krolzig (1997) and Kim and Nelson (1999). The EM algorithm has been designed to estimate the parameters of a model where the observed time series depends on an unobserved or a hidden stochastic variable. As mentioned earlier, y_t is directly observable but the state variable is unobservable and one can only make inference about its value based on the realized values of y_t. Formally, let $\xi_{it} = \Pr[s_t = i|\Omega_t; \theta]$, for $i = 1, 2$, where Ω_t refers to the information set (i.e., the set of observations available at period t) and θ is the vector of parameters to estimate. In general, inference is iterative for $t = 1, 2, ..., T$, in the sense that the previous value $\xi_{it-1} = \Pr[s_{t-1} = i|\Omega_{t-1}; \theta]$ is taken as input.

For inference purpose the densities under the different states are needed. Assuming normal fundamental they are given by

$$\begin{aligned} \eta_{it} &= f(y_t|s_t = i, \Omega_{t-1}; \theta) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{(y_t - c_i - \rho y_{t-1})^2}{2\sigma^2}]. \end{aligned}$$

The conditional density of the t-th observation can thus be calculated as

$$f(y_t|\Omega_{t-1}; \theta) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij} \xi_{jt-1} \eta_{it},$$

and therefore:

$$\xi_{it} = \frac{\sum_{j=1}^{2} p_{ij} \xi_{jt-1} \eta_{it}}{f(y_t|\Omega_{t-1}; \theta)}.$$

Using these results and iteration one can get the conditional log likelihood of observed data as following:

$$Log\ f(y_1, y_2, ..., y_T|y_0; \theta) = \sum_{t=1}^{T} Log\ f(y_t|\Omega_{t-1}; \theta)$$

for the given values of θ. To estimate θ, numerical optimization is used to maximize the conditional log likelihood using some starting values for ξ_{j0}. Assuming that the Markov chain is ergodic[4] the unconditional probabilities can be used as initial values. They are defined by

$$\begin{aligned} \xi_j &= \Pr[s = j] \\ &= \frac{1 - p_{ii}}{2 - p_{ii} - p_{jj}}. \end{aligned}$$

Once the coefficients of the model are estimated and a transition matrix is calculated, one can calculate the probability of being in state j at each period of time, based on the information of the whole sample. This series of probabilities is known as "smoothed probabilities". In addition, one can calculate the probability of being at state j at each time only based on the information up to that date (not the whole sample), the corresponding series of probabilities is known as "filtered probabilities".

In an autoregressive framework, there are several options concerning the regime-dependency of parameters. Regime-dependent parameters could be the intercept (I), the autoregressive coefficients (A), the variance (H), or their combinations. For example, a MSI model denotes a MS model with a regime dependent intercept. A MSIA model denotes a MS model where the intercept and the autoregressive coefficients are regime dependent.

The MS models have been used particularly in modeling business cycles. One interesting issue in this context is the analysis of asymmetries. There exists different measures of asymmetries depending whether the model is parametric or non parametric. In the last case, the most popular measures of asymmetries are steepness, deepness, and sharpness. Sichel (1993) defined an asymmetric cycle as one in which some phases of the cycle are different from the mirror image of the opposite phase.[5]

---

[4] Its transition matrix has at least one eigenvalue equal to unity. Recall that the two regime Markov chains are ergodic, provided that $p_{11} < 1, p_{22} < 1, p_{11} + p_{22} > 0$,

[5] Other types of asymmetry such as asymmetric persistence to shocks (see Hess and Iwata, 1997) and duration dependence (see Sichel, 1991, Filardo and Gordon, 1998) are widely used in the parametric

Consider a stationary univariate process y_t with mean μ_y and standard deviation σ_y. The process y_t is called symmetric if the following condition holds for all $\varepsilon \in \Re$:

$$\Pr[y_t < \mu_y - \varepsilon] = \Pr(y_t > \mu_y + \varepsilon].$$

In the literature on asymmetries, different ways have been suggested for testing for steepness and deepness. Sichel (1991) suggests the use of the skewness as an indicator of asymmetry, defined by

$$\tau_y = \frac{E[y_t - \mu_y]^3}{\sigma_y^3}.$$

One can infer about the type of asymmetry based on the sign of τ_y. The type of this asymmetry is named deepness if $\tau_y < 0$. There is tallness if $\tau_y > 0$.

Steepness of the distribution deals with the changes in the process over time, namely Δy_t. Since y_t is assumed to be stationary therefore its mean would be equal to zero, $\mu_{\Delta y} = 0$ and to satisfy the symmetric condition the following must hold for all $\varepsilon \in \Re$:

$$\Pr[\Delta y_t < -\varepsilon] = \Pr[\Delta y_t > \varepsilon].$$

Similarly, the degree of asymmetry can be measured by using the skewness coefficient of Δy_t. It is denoted by $\tau_{\Delta y}$. This coefficient shows that when $\tau_{\Delta y} < 0$ the asymmetry exists and is called negative steepness. On the other hand, $\tau_{\Delta y} > 0$ shows positive steepness.

Neftci (1984) used the duration of increases and decreases as an indicator for steepness of the time series. When the duration of increases is higher than the duration of decreases in the time series, the contractions are for sure steeper than the expansions.

The last measure of asymmetries for the non-parametric models is sharpness which has been introduced by McQueen and Thorley (1993) and implies that the shape of troughs and peaks are not the same (e.g. one is round and the other is sharp). Formally, a process is "non-sharp" if and only if the condition $p_{n1} = p_{nN}$ and $p_{1n} = p_{Nn}$ is satisfied for all $n \neq 1, N$; and $p_{1N} = p_{N1}$. So, in a two state model, non-sharpness requires to have $p_{12} = p_{21}$. In a MS model with three regimes, non-sharpness is obtained if $p_{12} = p_{32}$, $p_{13} = p_{31}$, and $p_{21} = p_{23}$.

---

models.

1.3 Data

For this study, quarterly data of four crimes series (Burglary, Larceny, Motor-Vehicle Theft, and Robbery) for the U.S. is used. The data is obtained from the Uniform Crime Reports of the Federal Bureau of Investigation (FBI). This data consists of 120 observations for the period of 1975:1 to 2004:4, and shows the number of each crime per 100,000 of population. The data for the quarterly unemployment rates is taken from the Bureau of Labor Statistics.

The data is seasonally non-adjusted. We use the TRAMO/SEATS procedure of Gomez-Maravall (1992) to remove the seasonality and detect the presence of outliers in the series.

The presence of unit roots is formally tested by using the ADF, GLS-based Augmented Dickey-Fuller statistic, and the feasible point optimal statistic suggested by Elliott, Rothenberg, and Stock (1996). The null hypothesis of a unit root is not rejected for any of the crime series. It means that all crime time series are I(1) processes. On the other hand, the unemployment rate appears to be stationary, that is, I(0).

1.4 Empirical Results of the Univariate MS-AR Models

Given previous results of the unit root tests, the crime series are modeled in first differences and the unemployment rate enters in level. A generalized version of the model suggested by Hamilton (1989) is estimated. All estimations are carried out using the MSVAR class for Ox by Krolzig (2005).

Before estimating the MS models, a linear model is specified. The lag length is determined by using information criteria such as AIC, BIC, and HQ. Using the selected lag structure, MS models with two and three states are estimated, allowing for changes in the intercept, variance and autoregressive parameters. Based on the information criteria and the LR statistic, we select the best fitting MS models. The null hypothesis of linearity is rejected for all selected models.

As Krolzig (1997) argues, a general test to compare two models with different number of regimes is unavailable. The issue is that regular asymptotic theory fails because of the presence of unidentified nuisance parameters and violation of the non singularity conditions.

A close examination of these models reveals that the selected models for burglary, larceny, and robbery can satisfactorily capture the movement of these series. However, the selected models of motor-vehicle theft and the unemployment rate do not fit well, so other models were selected to be able to explain the movements of these series.

1.4.1 Burglary

For this time series two MS models have been selected. The two selected MS models for burglary rates are the MSIH(2)-AR(1) and the MSIH(3)-AR(8). Table 1 shows results for the model MSIH(2)-AR(1). The null hypothesis of linearity for this time series is rejected. Overall, the results show that at the first regime, a decrease in the rate of burglary is present, and in the second regime, this rate increases or remains constant. Therefore, the first regime stands for lower and the second regime is for higher burglary rates in the U.S.

Based on the transition probabilities, it is clear that both regimes are very persistent with probabilities of 90% and 96%, respectively for the first regime (lower burglary rates) and the second regime (higher burglary rates). In addition, the unconditional probabilities of being at the lower and higher burglary rate regimes are 29% and 71% respectively. The transition probabilities reveal that the duration of the first regime is 10.2 quarters, whereas the second regime has a duration of 24.7 quarters. This indicates the presence of asymmetries in duration of the regimes of lower and higher burglary rates. It means that the periods of reduction in the rate of burglary are much shorter than the periods with positive change at this rate.

The results also show that only 34% of the observations for burglary are categorized in the first regime. It means that between the second quarter of 1977 and the end of 2004, in only 37.6 quarters the rate of burglary was declining. Concerning the standard errors, the model allows for different variances for each regime. In fact, the results show that the second regime is much less volatile than the first regime (standard error of 9.46 compared to 22.6).

The smoothed probabilities and the filtered probabilities of being at each regime at each point of time are visualized at Figure 1a. It shows that the first regime (lower burglary rates) is formed by the periods 1977:2-1981:3, 1983:4-1985:1, and 1988:3-1990:3. The same figure also shows that observations since 1990:4 until the end of the sample form the second regime (higher burglary rates).

Testing for asymmetries, we obtain that the hypothesis of non-sharpness, non-deepness, and non-steepness cannot be rejected for this model. Therefore, the cycles are symmetric.

The other selected model is the MSIH(3)-AR(8). In this case we have three regimes and a larger lag structure. The null hypothesis of linearity is again strongly rejected. As in the previous model, this model allows for heteroscedasticity among the regimes. Table 1 presents the results. The first two regimes have negative intercepts, however, the first one is not significant so the first regime shows the periods with no significant changes at the rate of burglary and the second regime stands for the periods with negative changes at the rate of this kind of crime. The third regime has positive intercept and stands for higher burglary rates. In this model, all the autoregressive coefficients are statistically significant.

The transition probabilities show that the third regime (higher burglary rates) is the most persistent regime with a probability of 94%, and the second regime (lower rate) is less persistent with 35%. Furthermore, there is a 25% chance to move from the first regime (constant rate of burglary) to the second one and only 2% to move to the third regime. Moreover, if this variable is at the second regime at time t, there is a higher chance to move to the first regime than to the third regime. As for the third regime, the probability of moving from this regime to the second regime is only 0.02%; this indicates that the burglary rates almost never changes from the third regime to the second one directly. In other words, if there is a positive change in the burglary rates at any period, then it would be very unlikely to see any decreases in this rate at next period.

As in the previous model, the third regime (higher burglary rates) has the longest duration and it continues more than 17 quarters, while the second regime lasts less than two quarters. These duration periods can be translated to the share of each regime in the sample; the higher the duration the higher is the share of that regime. In this case, 33% of observations fall into the first regime, 54% in third regime, and only 13% into the second regime. Therefore, in most of the period under study burglary was high in the U.S.

According to the standard errors of the regimes, it is observed that the first regime has the highest volatility (21.37) and the second regime the lowest (0.33). At the same time, the standard error of the third regime is 8.5.

Figure 1b shows the smoothed and the filtered probabilities of being at each regime. It is observed that the burglary rate fluctuates among regimes between 1977:2 and 1993:3. However, it has stayed at the third regime since 1993:4, which means that there was not any significant decline at the rate of burglary since this period.

Tests of asymmetry (with a χ^2 values of 0.712, 0.268, and 0.900, respectively) do not show any evidence of asymmetries in cycles. Therefore, different duration of regimes is the only source of asymmetry for this variable.

1.4.2 Larceny

Application of the TRAMO/SEATS procedure of Gomez and Maravall (1992) indicates the presence of an outlier at the observation 2003:2. A dummy variable (impulse type) is used to take this observation into account.

All information criteria indicate that the preferred model is the MSI(2)-AR(1). Notice that this is the simplest MS model. This model has only two regimes, no heteroscedasticity and a simple lag structure.

Table 1 presents the results for this model. Results show that the first regime has a negative intercept, the intercept of the second regime is positive and both of them are significant. The coefficient of the dummy variable for the outlier is negative and significant. According to these results, the first regime stands for the periods with decreasing rates of larceny and the second regime shows the periods with increasing rates of larceny.

The durations of the regimes are 2.93 and 12.17 quarters, respectively. From 111 observations included in the estimation, the first regime accounts for only 20% of the observations (22.6 quarters) while the second regime accounts for 88.4 quarters. The transition probabilities indicate that the second regime is more persistent than the first regime with a probability of 92% compared to 66%, respectively. It means that, if there is an increasing rate of larceny at any period, it will be very difficult to return to a regime with a decreasing rate. The unconditional probabilities of being at the regimes with lower and higher rates of larceny are 19% and 81% respectively.

Figure 2 shows the smoothed and the filtered probabilities of larceny in the two regimes. According to these results, the first regime is consists of 1977:2-1978:1, 1980:1-1980:4, 1991:1-1991:2, 1996:4-1997:1, 2001:1-2002:1, 2004:2-2004:3. Notice that most of these periods last only for two quarters.

The values of the χ^2 statistic for the hypothesis of non-sharpness and non-deepness are 6.3 and 5.8, respectively. It means that the null hypothesis can be rejected in favor of the alternative. Consequently, the cycles are not symmetric on these features. However, there is no evidence of steepness.

1.4.3 Motor-Vehicle Theft

For this variable, the AIC and the LR test suggest that the best model is a MSIAH(2)-AR(1). However, as mentioned earlier, a close investigation of this model reveals that it does not fit the data and it cannot capture the movements of this time series. At the same time, the SIC suggests a MSI(2)-AR(1) as the best model which matches the data and its movement well. Therefore, this model is selected as the best model for the motor- vehicle theft variable.

Table 1 shows the results for this model. The intercepts are negative and positive respectively for the first and second regimes. However, only the second intercept is significant. The coefficient of the lagged variable is not significant. The first regime shows the periods with negative change at the rate of motor-vehicle theft, and the second regime stands for the periods with positive changes of this rate.

Both of the regimes are very persistent and there is a 4% probability of moving from the first (lower rates of motor-vehicle theft) to the second regime (higher rates of motor-vehicle theft). The opposite direction has an 8% probability.

The results also indicate that 35% of the observations are classified in the higher rate regime. The expected duration of this regime is 13 quarters compared to 24 for the first regime.

Figure 3 shows the smoothed and the filtered probabilities of being at each regime. It shows that periods 1978:2-1983:1, 1990:4-2000:1, and since the second quarter in 2002, the variable is at the first regime. That is, the change at the rate of motor-vehicle theft in the U.S is negative. For the rest of the sample, the changes in motor-vehicle theft have been positive and the rate of motor-vehicle theft was higher.

According to the tests for asymmetry, all the cycles are symmetric at the 5% level.

1.4.4 Robbery

Two additive outliers are identified by using the procedure TRAMO/SEATS. They are located at observations 1993:1 and 1994:1. All information criteria select the model MSIH(2)-AR(1).

Table 1 presents the estimates. The first regime stands for all the periods with negative changes at the rate of robbery (lower robbery rates) and the second regime shows all the periods with positive changes in this rate (higher rates of robbery).

The second regime is more volatile than the first regime. It is reflected by the corresponding standard errors. The expected durations are 9.0 and 6.8 quarters, respectively for the first and second regimes. The lower robbery rates regime includes 67% of the observations (62.6 quarters). Transition probabilities indicate that both of the regimes are persistent, with probabilities equal to 89% and 85%, respectively. It means that if there is a negative change (positive change) in this rate at any period, then there is a high probability to have a negative change (positive change) at the next period.

Figure 4 visualizes the smoothed and the filtered probabilities for both regimes. The longest period with a decreasing rate of robbery is 1993:4-2002:1, which contains 34 quarters. The rest of the study period is oscillating between the regimes.

The only form of asymmetry is the different durations. All the other asymmetry tests indicate symmetry of the cycles, indicating that the cycles are not sharp, deep, or steep.

1.4.5 Unemployment Rate

Although the information criteria have selected the MSI(2)-AR(3) and the MSI(3)-AR(3) as the best models, these models cannot capture the movements of the series. The model that can fit the data well and describe the changes of the unemployment rate in the U.S. is the MSIAH(2)-AR(3). The LR linearity test strongly supports the non-linearity in this variable.

Table 2 shows the results of estimation of this model. This model allows for regime dependent autoregressive parameters and variance, so there are different estimations for coefficients of lagged variables in regimes. The estimated standard errors are 0.119 and 0.285, respectively, for regime 1 and 2. Therefore, the regime with higher unemployment rates is more volatile than the first regime.

Regime 1 shows all the periods with negative changes at the unemployment rate in the U.S and the second regime stands for all the periods with positive changes at this variable.

The expected duration of low unemployment is almost three times larger than the duration of higher rates of the unemployment. During the period of study, 79 quarters form the regime of low unemployment rate and the rest (32 quarters) is classified as the second regime (higher unemployment). Regime 1 is more persistent than the second regime based on the transition probabilities, with probabilities of 94% and 84%, respectively, to stay at the same state at the next period.

Figure 5 shows the smoothed and the filtered probabilities for both regimes. The only periods with high unemployment rates are 1979:1-1984:3, 1990:3-1992:1 and 2001:1-2001:4. These periods almost perfectly corresponds to the recession periods in U.S. At the rest of study period the unemployment rate in the U.S. was at a lower regime.

1.5 Synchronization of Cycles

So far, MS models for each single time series has been estimated and their properties, such as expected durations, transition probabilities, and asymmetries have been investigated individually. One of the questions of interest in this paper is the existence of any relationship between the unemployment rate (as an indicator of the condition of the economy) and the crime variables. In order to answer this question one can test whether the timing of the cycles of the unemployment rate and the crimes under study in this paper are similar or not. In other words, one can check for synchronization of cycles; that is, to see whether higher unemployment rate periods are independent from the higher crime rates regimes or not.[6]

In the following section the unemployment rate will be considered as the reference variable and its cycles as the reference cycle. These cycles will be compared with cycles of the crime time series to test for common cycles. For this purpose, the Concordance Index (CI), proposed originally by Harding and Pagan (2002), a method proposed by Bodman and Crosby (2005), and Pearson's contingency coefficient will be used.

The CI is a non-parametric statistic that shows the proportion of time that two series are in the same regime. For two series x_t and y_t for $t = 1, 2, ..., T$, let $S_{xt}(S_{yt})$ be a dummy variable that takes the value of unity when $x_t(y_t)$ is in regime 1 and value of zero when it is not in regime 1. Then the CI for x_t and y_t is given by the following expression

$$CI = T^{-1}\{\sum_{t=1}^{T} S_{x_t} S_{y_t} + \sum_{t=1}^{T}(1 - S_{x_t})(1 - S_{y_t})\}.$$

For example, a value of 0.8 for this index means that two series of x_t and y_t are in the same regime 80 percent of the time. Since CI is defined as the proportion of time that two series are in the same state, this index is bounded between zero and unity. The CI index has a value of unity when $S_{x_t} = S_{y_t}$ and value of zero when $S_{y_t} = (1 - S_{x_t})$. These two series are called pro-cyclical if $CI = 1$. They are counter cyclical if $CI = 0$.

It is natural to say that having a high concordance index means high common cycle. However, the question is how high should it be to interpret that as pro-cyclical? Even for two unrelated series, the expected value of the concordance index may be 0.5 or higher.[7]

---

[6] It should be mentioned that there is no agreement about the definition of synchronized cycles.

[7] For example, consider tossing two fair coins, the probability that both coins are in the same state (either heads or tails) is 0.5.

The above formula for concordance index can be written in a different way as follows

$$
\begin{aligned}
CI &= 1 + 2T^{-1}\sum_{t=1}^{T} S_{x_t}S_{yt} - \mu_{S_x} - \mu_{S_y} \\
&= 1 + 2\rho_S\sigma_{S_x}\sigma_{S_y} + 2\mu_{S_x}\mu_{S_y} - \mu_{S_x} - \mu_{S_y},
\end{aligned}
$$

where ρ_S is the estimated correlation coefficient between S_{x_t} and S_{y_t}. If $S_{x_t} = S_{y_t}$ or $S_{y_t} = (1 - S_{x_t})$, then $\sigma_{S_x}\sigma_{S_y} = \sigma^2_{S_x}$ so the value of unity for this index corresponds to $\rho_S = 1$ and value of zero to $\rho_S = 0$. Therefore, $\rho_S = 1$ ($\rho_S = -1$) shows that two cycles are perfectly positively (negatively) synchronized, and they are unsynchronized when $\rho_S = 0$.

Assuming that the two series are statistically uncorrelated ($\rho_S = 0$), the expected value of this index will be:

$$E(CI) = 1 + 2\mu_{S_x}\mu_{S_y} - \mu_{S_x} - \mu_{S_y}.$$

The expected value of being at each regime can be measured by dividing the number of periods at that regime by T. Now, this expected value can be compared with the calculated value from the series. If the former is smaller than the latter, one can say that there is a link between the cycles. This says that the number of periods where the series are in the same state is higher than if they were uncorrelated. If the former is larger than the latter, one can conclude that these series are counter-cyclical. The significance of this result has to be checked, though, to see whether the ratio of these two is statistically different from 1 or not.

Another problem that exists with using this index is that it depends on the expected values of S_{x_t} and S_{y_t}, that is their mean. Suppose that the mean of S_{x_t} and S_{y_t} is 0.5 and these two series are unsynchronized, then the expected value of the concordance index will be 0.5 which confirms the assumption that they were unrelated. But if the regime that takes value of one has higher duration than the other, the mean values of the series will be higher than 0.5. Now, assume that the means are 0.8, therefore, the expected concordance index will be 0.68 which is higher than 0.5, and one may think that these two series have common cycle even though they are not related. Therefore, the mean value of the series has to be taken into account. For this purpose the mean corrected concordance index (Artis et al, 2004) is considered.

Let $\overline{S}_x = T^{-1}\sum_{t=1}^{T} S_{x_t}$ indicate the estimated probability of being at regime 1. Then the mean corrected concordance index will be

$$CI^{corr} = 2T^{-1}\sum_{t=1}^{T}(S_{x_t} - \overline{S}_x)(S_{y_t} - \overline{S}_y).$$

As mentioned before, one of the shortcomings of the concordance index is that it does not allow for a statistical testing of the result. Harding and Pagan (2002) suggested that one can use a regression model to deal with this problem. To do so, the following regression can be used

$$\sigma_{s_y}^{-1} S_{y_t} = \alpha_1 + \rho_S \sigma_{S_x}^{-1} S_{x_t} + u_t.$$

Now, the hypothesis that $\rho_S = 0$ can be tested using the t-ratio of the coefficient of the $\sigma_{S_x}^{-1} S_{x_t}$. In this regression, when the null hypothesis is true, the error term inherits the serial correlation properties of S_{y_t}. In addition, S_{y_t} is strongly serially correlated, so robust estimated standard errors have to be used [such as HAC Newey-West method].

Using the regime classifications based on the MS models the calculated CIs are less than the expected CIs (under the assumption that the series are uncorrelated) except for robbery. Also the estimated correlation between the unemployment rate and burglary, larceny, motor-vehicle theft is negative; however, it is positive between the unemployment rate and robbery. This shows that the relationship between the unemployment rate and burglary, larceny and motor-vehicle theft is counter-cyclical. However, this relationship is pro-cyclical between the unemployment rate and robbery. According to the robust t-ratios reported at Table 3, only the results for larceny and robbery are statistically significant at 5%; so only two of the crime variables are significantly contemporaneously concordant with the unemployment rate cycle.

In summary we find no contemporaneous relationship between the unemployment rate and the following crimes: burglary and motor-vehicle theft. There is a negative and a positive contemporaneous relationship between the unemployment rate and larceny, and robbery, respectively.

The previous results were based on the assumption that the relationship between crimes and the unemployment rate, if there is any, is contemporaneous. We next investigate the concordance index of the lagged values of the unemployment rate and crime variables, assuming that changes in the regime of the unemployment rate at time $t - i$ may related with change at the regime of the crime series at time t, where i shows the number of lagged periods.

The results are the same for the relationship between the unemployment rate, burglary, motor-vehicle theft, and robbery. However, results change for larceny: correlation between the unemployment rate with one quarter lag and larceny is the same as for the contemporaneous case (negative and statistically significant at 10%) while it is not significant with two and three lags. With introducing higher lags the results change, the unemployment rate with four and five quarters lag show a positive and significant correlation with larceny. This is the case even with higher lags of the unemployment rate. Therefore, there is a positive relationship between larceny at time t with past values of the unemployment rate.

Another method used to test for any similarity or dissimilarity between the cycles of unemployment and the crime variables is the method proposed by Bodman and Crosby (2005). Bodman and Crosby used a simple but intuitive statistical method. Assume that the probability of being at regime 1 (lower regime) for the unemployment rate and robbery is 10%. Therefore, these two variables will be in regime 1 jointly only with 1% probability if they are independent from each other.

Hence, define a dummy variable S_x to be 1 when the series is in regime 1 and 0 otherwise, and take these as the realization of independent Bernoulli trials with length T (sample size). Now assume that the probability of being at regime 1 is p_i for each time series. Under the null hypothesis of independence, the expected number of periods that two series jointly spend in regime 1 will be $p_i \times p_j \times T$, with a standard deviation equal to $[(p_i p_j \times (1 - p_i p_j)]^{1/2}$.

Using the Markov Switching regime classification from the previous part of the paper, dummy variables are defined for being at the lower regime for each time series. None of the results are statistically significant even at the 10% of significance level. Therefore, based on this method the unemployment rate and all the crime variables under study are independent, and there is no significant relationship between the unemployment rate and crimes.

The conventional contingency table statistics, so called Pearson's contingency coefficient can be used to test for any form of relationship between the unemployment rate and the crime. The classification of the regimes for each single variable from the MS models can be used to create the contingency table as following

		Variable y		
		Regime 1	Regime 2	Subtotal
	Regime 1	n_{11}	n_{12}	$n_{1.}$
Variable x	Regime 2	n_{21}	n_{22}	$n_{2.}$
	Subtotal	$n_{.1}$	$n_{.2}$	N

This table can be used to check for any association between the unemployment rate and the crimes. Pearson's contingency coefficient is defined by

$$CC = \sqrt{\frac{\widehat{\chi}^2}{N + \widehat{\chi}^2}},$$

where

$$\widehat{\chi}^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{[n_{ij} - (n_{i.}n_{.j}/N)]^2}{n_{i.}n_{.j}/N},$$

and n_{ij} is the number of the periods where variables x_t and y_t are respectively at regime i and j, at the same time. This coefficient is associated to the conventional correlation coefficient for continuous data. For finite dimensions, a corrected version of this statistic has to be used, in these cases, the coefficient is bounded from above and is biased. The limit of this coefficient is proportional to the dimension of the table. For a 2×2 table, the correction factor is $\sqrt{0.5}$, so the corrected contingency coefficient will have the following form

$$CC^{corr} = \frac{CC}{\sqrt{0.5}}.$$

The CC^{corr} ranges between 0 and 1. The value of 0 denotes that two series are uncorrelated, and the value one shows a complete correlation. Larceny and robbery have the highest correlation of 0.396 and 0.349, respectively. Assuming 0.5 as the threshold level, none of the crime series shows any kind of commonality with the unemployment rate.

1.6 Empirical Results of MSVAR Models

In this section, bivariate MS models are estimated to investigate the common regime changes of the unemployment rate and each one of the crime series under investigation.

The optimal number of lags have been selected using AIC and SIC criteria for the system. The selected lag for larceny and robbery is two. For burglary and motor-vehicle theft, the information criteria have selected two optimal lags, two and three. Because we are interested to see whether the unemployment rate has a different effect on crime in separate regimes, only models that allow for regime-dependent autoregressive parameter are estimated.

The MSIA and MSIAH models are estimated for each system using the selected optimal lags assuming two and three regimes. To choose the best MS-VAR model for each system AIC, SIC, HQ criteria, and LR (Likelihood Ratio) test are used.

The selected model for the system containing burglary and the unemployment rate is a MSIAH(2)-VAR(3) model, with a very strong rejection of the null hypothesis of linearity.

The estimation results of this model are summarized in Table 7. In this model the intercepts of both regimes are positive, however, the intercept is statistically significant only at the second regime. Hence, the second regime stands for the periods with positive changes in the rate of burglary and the first regime shows the periods with no significant changes in this rate.

The estimated coefficients of lagged values of the unemployment rate in the equation for burglary are significant at regime 2. However, in regime 1 only the unemployment rate with three lags is significant. The overall effect of the unemployment on burglary at regime 1 is -1.7 and at regime 2 is -3.5. Therefore, these results show that the unemployment rate has a negative effect on burglary regardless of the regime.

The expected duration of regime 1 is about two years. In addition, the estimated expected durations of regimes in this model show that the period with positive changes at the rate of burglary (regime 2) lasts much less than the periods with no changes at this rate. According to the results, only 17 quarters (15%) of the 111 quarters included in this study are categorized in the second regime (higher rates of burglary). For the rest of the sample period, this variable was statistically constant. The transition probabilities for these regimes show that only the first regime is persistent. The probabilities of moving from/to the first regime are different. Moving from regime 2 to regime 1 is more likely than the opposite.

Based on the smoothed and the filtered probabilities for each regime at this model for burglary, which is illustrated at Figure 6, all the periods classified as regime 2 are peaks except from the last quarter in 1981 to last quarter in 1982 and the third and fourth quarters of 1984. From the first quarter in 1985 to the end of sample, except for three quarters, all are in regime 1.

The selected model for larceny and unemployment rates is a MSIA(3)-VAR(2). Table 8 shows the estimated results. The estimates for larceny show that the intercept of regime 1 is significant and negative and the intercept of the regime 2 is positive and significant at border of 10%, however, the other intercept is not significant. Therefore, the regimes 1, 2, and 3 stand for negative, positive, and no significant changes in larceny.

Regarding the equation for larceny, except for larceny with one lag and the unemployment rate with two lags at regime 3 the other lagged variables are not significant at conventional levels of significance. The coefficient of the unemployment rate with one lag at regime 3 is significant at the border of 10%. Therefore, the overall effect of the unemployment rate is significant only at regime 3 and this effect is equal to 0.76.

The transition probability matrix shows that all the regimes are persistent but the persistency of the second and third regimes are higher than for regime 1 (lower rate of larceny). Regime 3 never changes to regime 1 and larceny moves to regime 2 only with a 14.1% probability. The chance of having a regime with lower (constant) rate of larceny after a regime with higher rate is only 5.2% (1.4%). The expected duration of the regimes indicate that regime 2 has the highest expected duration and regime 1 the lowest. The durations of regime 2 and 3 are about 15 and 7 quarters, respectively.

Figure 7 shows the smoothed and the filtered probabilities of being at each regime. For 1980:4-1982:4, 1990:3-1992:1, 2004:3-2004:4, the rate of larceny remains constant and this variable is in regime 3. The second regime, which is for high rates of larceny, consists of 1977:4, 1983:1-1983:2, 1984:2-1990:2, 1992:2-2001:2, and 2002:1-2004:1. The rest of the sample is in regime 1 and the rate of larceny in these periods is low.

The selected MS model for motor-vehicle theft and unemployment rate is a MSIA(3)-VAR(3). The estimation result for this model is reported in Table 9. The intercept of the regime 3 is -0.816 and this is the only significant intercept in all regimes. At the equation for motor-vehicle theft, in regime 1 (3) only the coefficient of the unemployment rate with three (two and three) periods lag is significant. In regime 2 the lagged unemployment rates are not significant at conventional levels. The overall effect of the unemployment rate in regime 1 and 3 is -0.044 and 0.166, respectively. Therefore, the unemployment rate has different effects in separate regimes.

At the same time, regime 1 has the highest duration and lasts more than 6 years; in addition, regime 3 lasts less than half a year. Regime 1, 2, and 3 includes 58.3, 34.1, and 18.5 quarters of the sample. Furthermore, regime 1 is very persistent with a probability equal to 96.2%, regime 2 is persistent as well but the last regime is not persistent. Regime 2 never changes to regime 1 and it changes to regime 3 with a probability of 37.7%. At the same time regime 1 never alters to regime 3.

Classification of the regimes is illustrated in Figure 9. Based on the estimated smoothed probabilities for each regime, regime 1 consists of 1978:1-1979:4, 1991:1-2000:4, and 2002:1-2004:4. The rest of the sample are peaks that oscillate between regime 2 or 3.

The information criteria and the LR test select a MSIA(2)-VAR(2) model for the variables robbery and unemployment rate. This model has 2 regimes and intercept and autoregressive parameters are regime dependent. The estimated results are presented in Table 10.

The results show that both intercepts are positive, however the intercept is significant only at regime 2. Therefore, it could be said that regime 1 shows a regime with constant rate of robbery and regime 2 stands for higher rates of robbery. The estimated coefficients of lagged values of robbery and the unemployment rate are not significant at none of the regimes. Consequently, the unemployment rate does not have any significant effect on the rate of robbery.

The transition probabilities show that both regime 1 and 2 are very persistent. There is only a 1.2% chance of moving from regime 1 to regime 2 and the probability of moving in opposite direction is 10.3%. Out of 111 observation included in this study 98.9 quarters are in regime 1 and only 12.1 quarters in regime 2. The duration of the first regime is also much higher than the second regime. The estimated smoothed probabilities at Figure 9 show that regime 2 begins from the first quarter in 1980 and stops at the end of 1982. The rest of the sample is in regime1.

1.7 Conclusions

The results of univariate MS-AR models jointly with different measures of association have been used to test the existence of any relationships between the unemployment rate and various crime variables. Results confirm that there is no relationship between the unemployment rate, burglary and motor-vehicle theft. Results for larceny and robbery are mixed. The CI is the only measure that shows a relationship between the unemployment rate and robbery and larceny. The contemporaneous relationship is positive for robbery and negative for larceny; however, it turns to be positive between the lagged values of the unemployment rate and larceny.

The objective of second phase, MS-VAR models, was to test for the existence of different influences of the unemployment rate in separate states of the crime variable. Therefore, the overall effects of the unemployment rate in different regimes on the crime series are estimated. Results are mixed. Results show that the unemployment rate has a negative effect on burglary regardless of the state of the burglary rate, and it does not have any significant effect on robbery during the periods under study. As for the larceny and motor-vehicle theft, the effect of the unemployment rate depends on the regime of these variables. For larceny and motor-vehicle theft this effect is positive and zero in regimes 3 and 2, respectively. At regime 1 for these crimes, the overall effect of the unemployment rate on larceny is zero; however, it is negative for motor-vehicle theft.

References

[1] Albert, J., S. Chib (1993), "Bayesian Analysis via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business and Economic Statistics*, **11(1)**, 1-15.

[2] Artis, M. ,H. M. Krolzig and J. Toror(2004),“ The European Business Cycle,” *Oxford Economic Papers*, **56(1)**, 1-44.

[3] Bodman, P. and M. Crosby (2005), “Are Business Cycles Independent in the G7?,” *International Economic Journal* **19 (4)**, 483-499.

[4] Britt, C. L. (1997), “Reconsidering the Unemployment and Crime Relationship: Variation by Age Group and Historical Period,” *Journal of Quantitative Criminology* **13 (4)**, 405-428.

[5] Cantor, D., and K. C. Land (1985), “Unemployment and crime rates in the post-World War II United States: A theoretical and empirical analysis,” *American Sociological Review*, **50(3)**, 317–332.

[6] Chamlin, M. B. and J. K. Cochran (2000), “Unemployment, Economic Theory, and Property Crime: A Note on Measurement,” *Journal of Quantitative Criminology* **16 (4)**, 443-455.

[7] Chiricos, T. (1987), “Rates of Crime and Unemployment: An Analysis of Aggregate Research Evidence,” *Social Problem* **34,** 187-212.

[8] Clements, M. P. and H. M. Krolzig (2003), “Business Cycle Asymmetries: Characterization and Testing Based on Markov-Switching Autoregressions,” *Journal of Business & Economic Statistics* **21 (1)**, 196-211.

[9] Cook P. J. and G. A. Zarkin (1985), “Crime and the Business Cycle.” *Journal of Legal Studies* **14 (1)**, 115-128.

[10] Doornik, J. A. (2001), *Object-Oriented Matrix Programming using Ox,* 4th Edition. London: Timberlake Consultants Press.

[11] Durland, J. M. and T. H. McCurdy (1994), “Duration-Dependent transitions in a Markov Model of US GNP Growth,” *Journal of Business and Economic Statistics*, **12(3)**, 279 – 88.

[12] Edmark, K. (2005), “Unemployment and Crime: Is there a Connection?,” *Scandinavian Journal of Economics* **107 (2)**, 353-373.

[13] Elliott, G., T. J. Rothenberg, and J. H. Stock (1996),“Efficient Tests for an Autoregressive Unit Root,” *Econometrica*, **64(4)**, 813–836.

[14] Entorf, H. and H. Spengler (2000), "Socioeconomic and Demographic Factors of Crime in Germany - Evidence from Panel Data of the German States," *International Review of Law and Economics* **20 (1)**, 75-106.

[15] Filardo, A. J. and S. F. Gordon (1998), "Business Cycle Durations," *Journal of Econometrics* **85 (1)**, 99-123.

[16] Goldfeld, S. M. and R. E. Quandt (1973), "A Markov Model for Switching Regressions," *Journal of Econometrics* **1(1)**, 3-15.

[17] Gomez, V., and A. Maravall (1992), "Time Series Regression with Arima Noise and Missing Observations - Program Tram," Working Paper 92/81, European University Institute, Florence.

[18] Greenberg, D. (2001), "Time series Analysis of Crime Rate," *Journal of Quantitative Criminology* **17(4)**, 291-327

[19] Hale, C. and D. Sabbagh (1991), "Testing the Relationship between Unemployment and Crime: A methodological comment and empirical analysis using time series data from England and Wales," *Journal of Research and Delinquency* **28 (4)**, 400-417.

[20] Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, **57(2),** 357-384.

[21] Hamilton, J. D. (1990), "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics* **45(1,2),** 30-70.

[22] Hamilton, J. D. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

[23] Hamilton, J. D. and B. Raj (2002), "New Directions in Business Cycle Research and Financial Analysis," *Empirical Economics* **27(2)**, 149-162.

[24] Harding, D. and A. Pagan (2002), "Dissecting the Cycle: A Methodological Investigation," *Journal of Monetary Economics* **49(2)**, 365-381.

[25] Harding, D. and A. Pagan (2006), "Synchronization of Cycles," *Journal of Econometrics* **132(1)**, 59-79.

[26] Hess, G. D. and S. Iwata (1997), "Asymmetric Persistence in GDP? A Deeper Look at Depth," *Journal of Monetary Economics* **40(3)**, 535-554.

[27] Imrohoroglu, A., A. Merlo and P. Rupert (2004), "What Accounts for the Decline in Crime?," *International Economic Review* **45(3)**, 707-729.

[28] Kapuscinski C. A., J. Braithwaite and C. Bruce (1998), "Unemployment and Crime: Toward Resolving the Paradox," *Journal of Quantitative Criminology* **14(3)**, 215-243.

[29] Kim, C. J. (1994), "Dynamic Linear Models with Markov-Switching," *Journal of Econometrics* **60(1-2)**, 1-22.

[30] Kim, C. J. and C. R. Nelson (1999), *State Space Models with Regime Switching, Classical and Gibbs Sampling Approaches with Applications*, Cambridge, MA: MIT Press.

[31] Krolzig, H. M. (1997), *Markov-Switching Vector Autoregressions: Modeling, Statistical Inference, and Application to Business Cycle Analysis*, Springer-Verlag.

[32] Krolzig, H. M. (1998), Econometric Modeling of Markov-Switching Vector Autoregressions using MSVAR for OX. Institute of Economics and Statistics and Nuffield College, Oxford.

[33] Krolzig, H. M. (2005). MSVAR 1.31k for Ox 3.4, http://www.kent.ac.uk/economics /staff/hmk/index.html.

[34] Krolzig, H. M. and J. Toro (2005), "Classical and Modern Business Cycle Measurement: The European Case," *Spanish Economic Review* **7(1)**, 1-21.

[35] Levitt, S. D. (2001), "Alternative Strategies for Identifying the Link between Unemployment and Crime," *Journal of Quantitative Criminology* **17(4)**, 377-390.

[36] McQueen, G. and S. Thorley (1993), "Asymmetric Business Cycle Turning Points," *Journal of Monetary Economics* **31(3),** 341-363.

[37] Neftci, S. N. (1984), "Are Economic Time Series Asymmetric Over the Business Cycle?," *The Journal of Political Economy* **92(2),** 307-328.

[38] Quandt, R. E. (1972), "A new approach to estimating switching regressions," *Journal of the American Statistical Association* **67**, 306-310.

[39] Raphael, S. and R. Winter-Ebmer (2001), "Identifying the Effect of Unemployment on Crime," *Journal of Law and Economics* **44(1)**, 259-283.

[40] Sichel, D. E. (1991), "Business Cycle Duration Dependence: A Parametric Approach," *The Review of Economics and Statistics* **73(2)**, 254-260.

[41] Sichel, D. E. (1993), "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry* **31 (2)**, 224-236.

[42] Sichel, D. E. (1994), "Inventories and the Three Phases of the Business Cycle," *Journal of Business & Economic Statistics* **12(3)**, 269-277.

[43] Witt R., A. Clarke and F. Nigel (1999), "Crime and Economic Activity: A Panel Data Approach," *British Journal of Criminology* **39(3)**, 391-400.

Tables

Table 1. Results of Univariate MS-AR models for Crimes

	Burglary		Robbery	Larceny	Motor
	MSIH(2)-AR(1)	MSIH(3)-AR(8)	MSIH(2)-AR(1)	MSI(2)-AR(1)	MSI(2)-AR(1)
Intercepts					
μ_1	-6.049 (-1.571)	-5.797 (-1.517)	-3.289 (-1.270)	-2.773 (-4.098)	-0106 (-1.032)
μ_2	0.677 (0.481)	-8.397 (-49.224)	-1.478 (-0.319)	2.106 (5.729)	0.197 (2.001)
μ_3		2.983 (2.561)			
Autoregressive Parameters					
ϕ_1	0.061 (0.511)	-0.047 (-6.623)	-0.015 (-0.194)	0.094 (1.038)	-0.104 (-0.923)
ϕ_2		-0.027 (-2.911)			
ϕ_3		0.069 (10.958)			
ϕ_4		-0.052 (-6.797)			
ϕ_5		0.107 (18.192)			
ϕ_6		0.109 (17.025)			
ϕ_7		0.196 (27.136)			
ϕ_8		-0.126 (-17.474)			
Dummy 1			98.852 (5.370)	-10.989(-4.971)	
Dummy 2			-97.897 (-6.482)		
Standard Errors					
σ_1	22.605	21.369	13.359	2.172	0.301
σ_2	9.464	0.334	27.824		
σ_3		8.501			
Transition Probabilities					
P11	0.902	0.745	0.889	0.658	0.959
P12	0.098	0.253	0.111	0.342	0.041
P13		0.002			
P21	0.040	0.369	0.147	0.082	0.076
P22	0.960	0.347	0.853	0.918	0.924
P23		0.285			
P31		0.058			
P32		0			
P33		0.942			
Durations					
Regime 1	10.18	3.93	9.01	2.93	24.33
Regime 2	24.73	1.53	6.80	12.17	13.20
Regime 3		17.12			
Log Likelihood					
Nonlinear	-450.050	-421.027	-490.750	-268.123	-31.197
Linear	-461.335	-452.363	-495.241	-273.131	-34.740

t-statistics are in paranthesis.

Table 2. Results of MSIAH(2)–AR(3) for Unemployment Rate

	Regime 1	Regime 2
μ	0.139 (1.085)	0.607 (2.271)
ϕ_1	1.189 (9.757)	1.613 (8.630)
ϕ_2	-0.056 (-0.302)	-0.730 (-2.221)
ϕ_3	-0.166 (-1.500)	0.036 (0.192)
Standard errors	0.119	0.285
Duration	16.48	6.14
Log Likelihood		
Nonlinear	37.311	
Linear	19.840	
Transition Probabilities		
	Regime 1	Regime 2
Regime 1	0.939	0.061
Regime 2	0.163	0.837

t-statistics are in parenthesis.

Table 3. Concordance Index

	Unemployment rate & Burglary 2 regimes	Unemployment rate & Burglary 3 regimes	Unemployment rate & Larceny	Unemployment rate & Motor	Uemployment rate & Robbery
CI	0.333	0.288	0.261	0.550	0.667
E(CI)	0.414	0.349	0.361	0.551	0.555
$\hat{\rho}$	-0.195	-0.192	-0.292	-0.004	0.255
t-ratio	-1.196	-1.473	-2.212	-0.027	1.643
CI^{corr}	-0.081	-0.061	-0.100	-0.002	0.111

Table 4. The Estimated Correlations between Crime Variables and Unemployment rate;

from regressing crime variables on lagged values of unemployment rate

	Unemployment rate & Burglary 2 regimes	Unemployment rate & Burglary 3 regimes	Unemployment rate & Larceny	Unemployment rate & Motor	Unemployment rate & Robbery
1 lag	-0.161	-0.076	-0.305*	0.003	0.305*
t-ratio	-0.932	-0.639	-1.851	0.017	1.875
2 lag	-0.171	-0.131	-0.213	0.011	0.355*
t-ratio	-0.983	-1.213	-1.552	0.062	2.337
3 lag	-0.137	-0.089	-0.067	0.018	0.406*
t-ratio	-0.834	-0.759	-0.666	0.107	2.699
4 lag	-0.103	-0.087	0.133*	-0.015	0.416*
t-ratio	-0.658	-0.753	1.853	-0.088	2.676
5 lag	-0.069	-0.085	0.244*	-0.02	0.384*
t-ratio	-0.455	-0.706	2.809	-0.108	2.373
6 lag	-0.034	-0.141	0.247	0.018	0.352*
t-ratio	-0.231	-1.308		0.096	2.206
7 lag	0	-0.139	0.251	0.055	0.277*
t-ratio	0	-1.091		0.303	1.74
8 lag	0.034	-0.078	0.254*	0.093	0.245
t-ratio	0.22	-0.689	2.998	0.511	1.603

An * indicates the significant correlations at 10% or lower.

Table 5. Calculate CIs between Crime Variables and Lagged Values of The unemployment Rate

	Unemployment rate & Burglary 2 regimes	Unemployment rate & Burglary 3 regimes	Unemployment rate & Larceny	Unemployment rate & Motor	Unemployment rate & Robbery
1 lag	0.342	0.324	0.252	0.55	0.685
2 lag	0.333	0.306	0.279	0.55	0.703
3 lag	0.342	0.315	0.324	0.55	0.721
4 lag	0.351	0.315	0.387	0.532	0.721
5 lag	0.36	0.315	0.423	0.523	0.703
6 lag	0.369	0.297	0.423	0.532	0.685
7 lag	0.378	0.297	0.423	0.541	0.649
8 lag	0.387	0.315	0.423	0.55	0.631

Table 6. Results from Bodman and Crosby Method

	Unemployment rate & Burglary 2 regimes	Unemployment rate & Burglary 3 regimes	Unemployment rate & Larceny	Unemployment rate & Motor	Unemployment rate & Robbery
Actual number of quarters jointly at regime 1	19	8	8	49	56
Number of quarters at regime 1 assuming independency	23.486	11.387	13.523	49.108	49.820
Standard deviation	4.303	3.197	3.446	5.233	5.240

Table 7. Estimation Results of MSIAH(2)-VAR(3) Model of Burglary

	Regime 1		Regime 2	
	Dburglary	unemployment	Dburglary	unemployment
μ	7.415(1.209)	0.351(5.159)	30.554(2.448)	0.229(0.402)
$\phi_{1,1}$	0.236(2.500)	-0.003(-3.222)	-1.007(-10.788)	0.008(1.722)
$\phi_{1,2}$	-0.120(-1.405)	-0.002(-2.449)	0.531(3.881)	-0.004(-0.596)
$\phi_{1,3}$	0.190(2.248)	-0.001(-1.129)	-0.394(-3.889)	-0.005(-1.029)
$\phi_{2,1}$	-6.303(-0.847)	1.230(14.891)	-40.687(-6.496)	1.688(5.471)
$\phi_{2,2}$	19.352(1.525)	-0.137(-0.958)	71.390(7.299)	-0.801(-1.664)
$\phi_{2,3}$	-14.749(-2.038)	-0.160(-1.935)	-34.240(-6.721)	0.100(0.399)
σ	12.018	0.131	5.582	0.275
		Log Likelihood		
Nonlinear		-391.945		
Linear		-432.636		
		Transition Probabilities		
	Regime 1	Regime 2		Duration
Regime 1	0.875	0.125		8.03
Regime 2	0.695	0.306		1.44

t-statistics in parentheses.

Table 8. Estimation Results of MSIA(3)-VAR(2) Model of Larceny

	Regime 1		Regime 2		Regime 3	
	Dlarceny	unemployment	Dlarceny	unemployment	Dlarceny	unemployment
μ	-12.420(-3.598)	0.828(4.271)	2.199(1.641)	0.189(2.539)	-2.846(-0.723)	0.099(0.486)
$\phi_{1,1}$	-0.384(-1.548)	-0.049(-4.129)	0.147(1.311)	-0.003(-0.648)	0.645(3.528)	0.051(5.524)
$\phi_{1,2}$	-0.039(-0.097)	0.113(4.188)	0.054(0.570)	0.007(1.508)	-0.104(-0.533)	0.020(1.991)
$\phi_{2,1}$	-1.262(-0.408)	1.861(9.803)	-1.566(-0.904)	1.020(12.491)	-4.405(-1.624)	1.034(7.039)
$\phi_{2,2}$	2.630(0.904)	-0.972(-5.795)	1.402(0.795)	-0.247(-2.474)	5.175(1.760)	-0.022(-0.140)
D2003	-2.621(-0.151)	0.236(0.267)	-10.342(-4.455)	0.335(2.829)	-11.891(-0.282)	-0.036(-0.017)
σ	2.266	0.116	2.266	0.116	2.266	0.116

Log Likelihood	
Nonlinear	-199.930
Linear	-249.360

Transition Probabilities

	Regime 1	Regime 2	Regime 3	Duration
Regime 1	0.555	0.297	0.149	2.25
Regime 2	0.052	0.934	0.014	15.23
Regime 3	0	0.141	0.859	7.110

t-statistics are in parenthesis.

Table 9. Estimation Results of MSIA(3)-VAR(3) Model of Motor

	Regime 1		Regime 2		Regime 3	
	Dmotor	unemployment	Dmotor	unemployment	Dmotor	unemployment
μ	0.148(0.733)	0.079(0.609)	-0.050(-0.235)	0.727(5.637)	-0.816(-2.100)	1.493(5.444)
$\phi_{1,1}$	-0.165(-1.576)	-0.058(-0.900)	-0.246(-1.982)	-0.092(-1.382)	1.028(3.481)	-0.953(-4.871)
$\phi_{1,2}$	-0.207(-2.028)	-0.117(-1.886)	0.416(3.336)	-0.138(-1.999)	-0.882(-4.285)	-0.150(-1.160)
$\phi_{1,3}$	-0.107(-1.051)	-0.005(-0.075)	0.401(3.790)	-0.094(-1.472)	-0.155(-0.761)	-0.194(-1.466)
$\phi_{2,1}$	0.297(1.538)	1.221(10.303)	0.248(1.458)	1.203(10.783)	0.231(0.725)	1.148(5.001)
$\phi_{2,2}$	0.078(0.238)	0.069(0.351)	-0.414(-1.514)	-0.352(-2.008)	-1.489(-2.447)	0.499(1.144)
$\phi_{2,3}$	-0.419(-2.203)	-0.312(-2.724)	0.184(1.131)	0.033(0.355)	1.424(4.080)	-0.832(-3.360)
σ	0.215	0.131	0.215	0.131	0.215	0.131

	Log Likelihood
Nonlinear	39.614
Linear	-8.384

Transition Probabilities

	Regime 1	Regime 2	Regime 3	Duration
Regime 1	0.962	0.038	0	26.56
Regime 2	0	0.623	0.377	2.66
Regime 3	0.163	0.552	0.285	1.4

t-statistics are in parenthesis.

Table 10. Estimation Results of MSIA(2)-VAR(2) Model of Robbery

	Regime 1		Regime 2	
	Drobbery	unemployment	Drobbery	unemployment
μ	10.044(1.078)	0.326(4.508)	117.856(2.697)	0.438(1.399)
$\phi_{1,1}$	-0.078(-0.988)	0.001(2.308)	-0.032(-0.131)	0.002(0.882)
$\phi_{1,2}$	0.096(1.176)	0.000(0.145)	-0.244(-1.109)	-0.011(-6.434)
$\phi_{2,1}$	5.770(0.581)	1.386(17.868)	5.625(0.330)	1.436(10.632)
$\phi_{1,2}$	-8.104(-0.847)	-0.448(-5.990)	-20.919(-1.128)	-0.456(-3.119)
D1993	102.968(5.249)	-0.073(-0.481)	77.160(338.796)	-0.752(-0.204)
D1994	-100.258(-5.124)	0.092(0.606)	-117.797(-2411.194)	0.041(0.010)
σ	19.200	0.149	19.200	0.149
		Log Likelihood		
Nonlinear		-439.875		
Linear		-464.712		
		Transition Probabilities		
	Regime 1	Regime 2		Duration
Regime 1	0.988	0.012		85.84
Regime 2	0.103	0.897		9.74

t-statistics are in parenthesis.

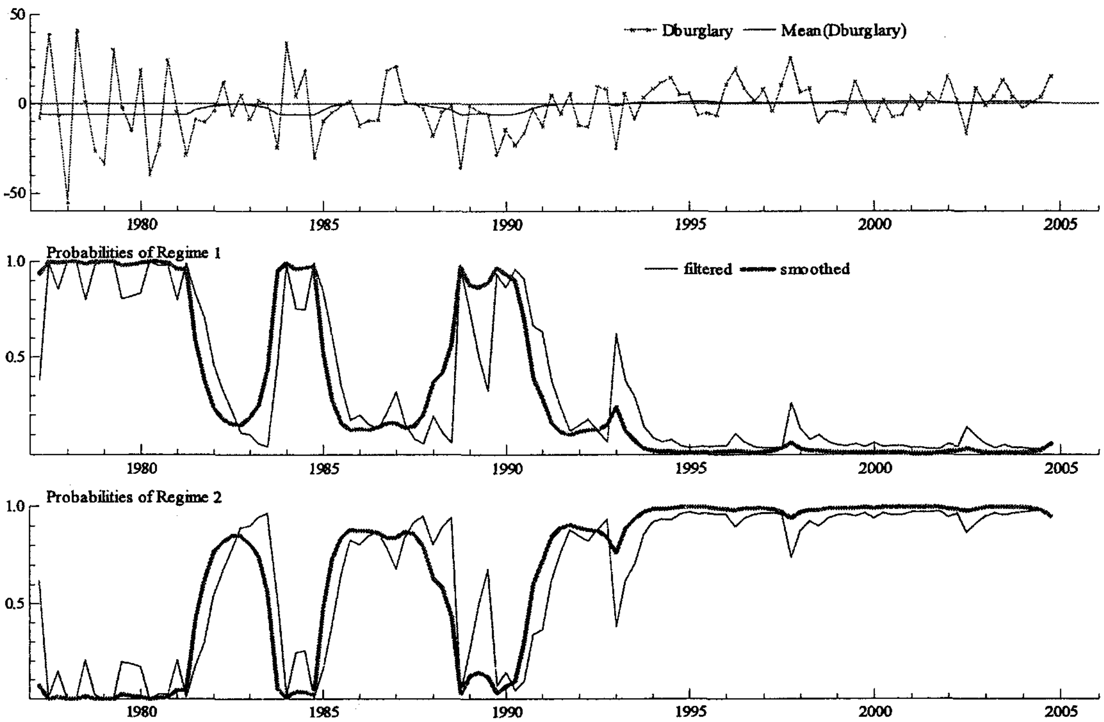

Figure 1a. Mean, smoothed, and filtered probabilities of Burglary; MSIH(2)-AR(1) model

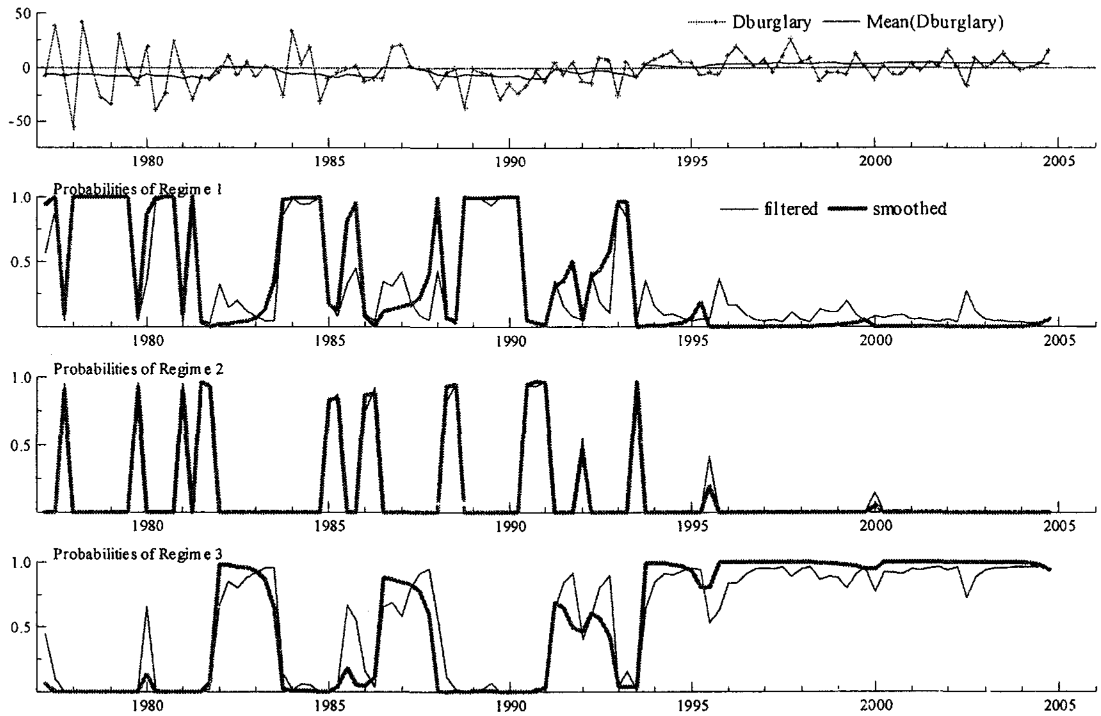

Figure 1b. Mean, smoothed, and filtered probabilities of Burglary; MSIH(3)-AR(8) model

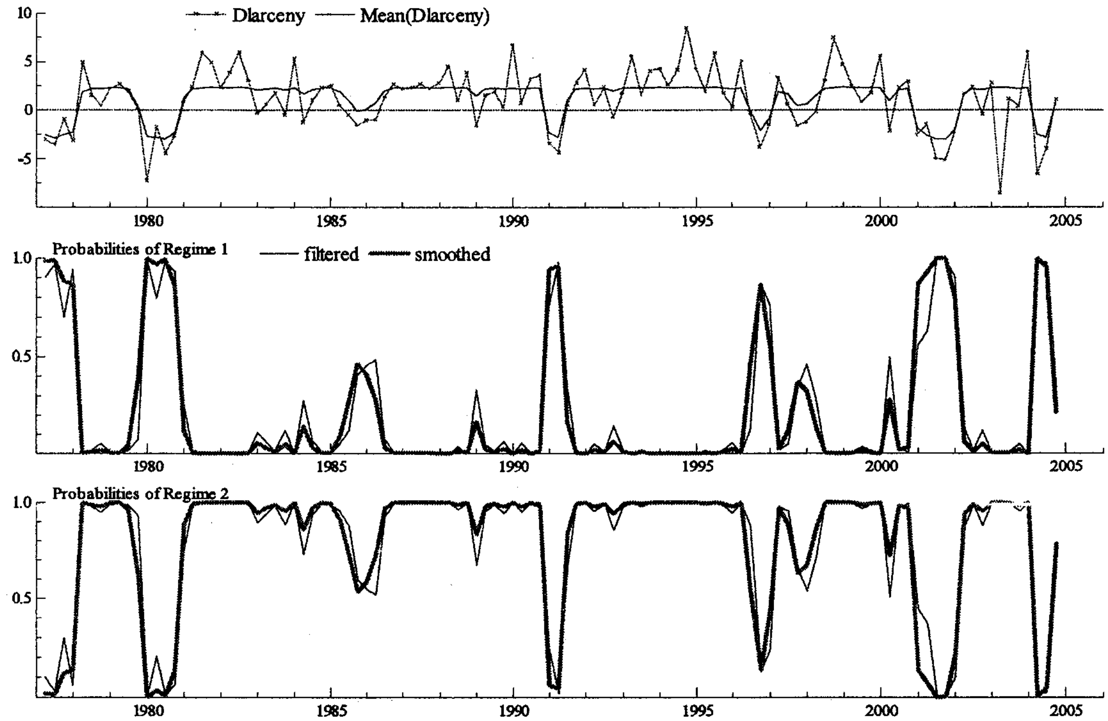

Figure 2. Mean, smoothed, and filtered probabilities of Larceny; MSI(2)-AR(1) model

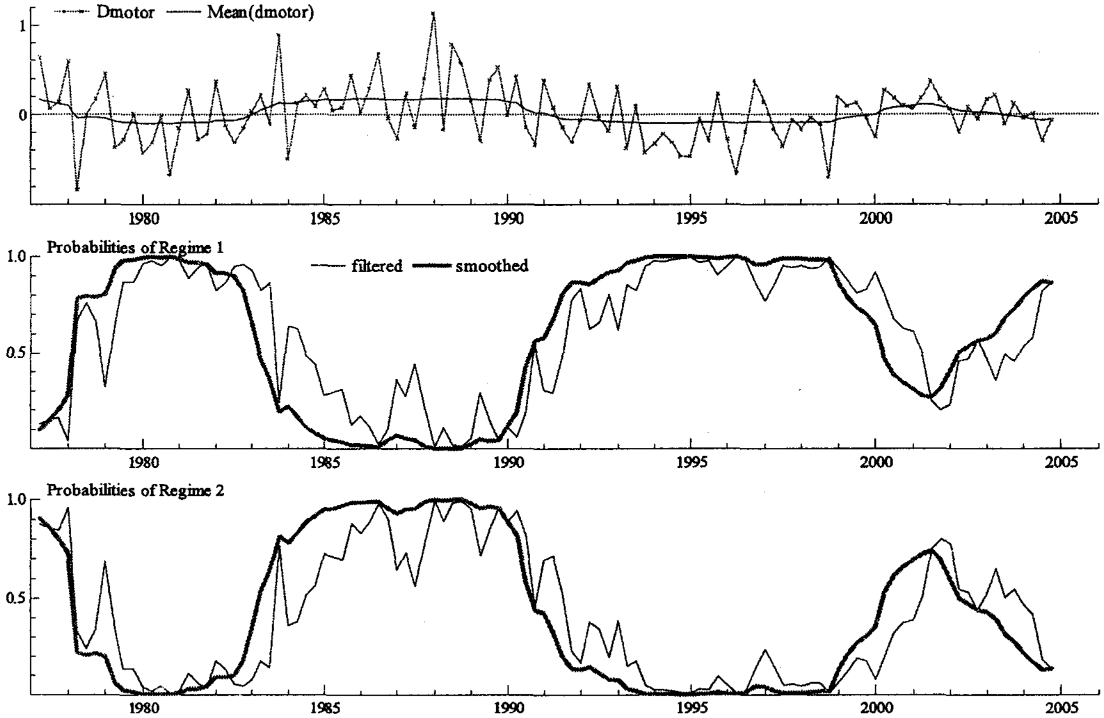

Figure 3. Mean, smoothed, and filtered probabilities of Motor; MSI(2)-AR(1) model

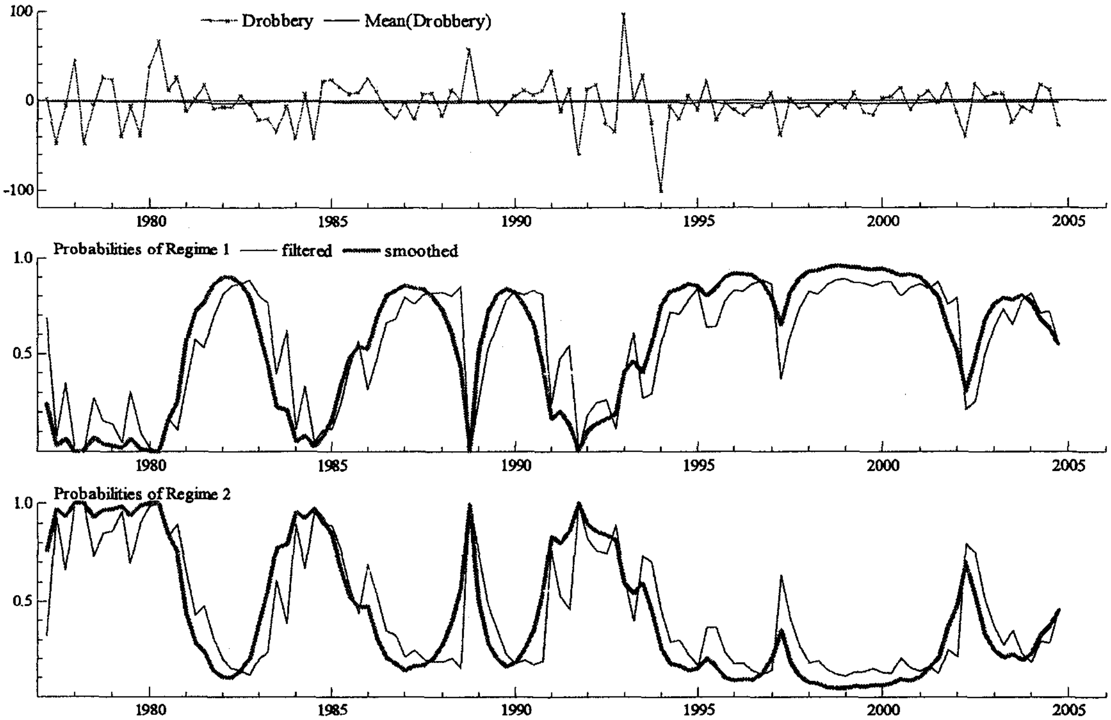

Figure 4. Mean, smoothed, and filtered probabilities of Robbery; MSIH(2)-AR(1) model

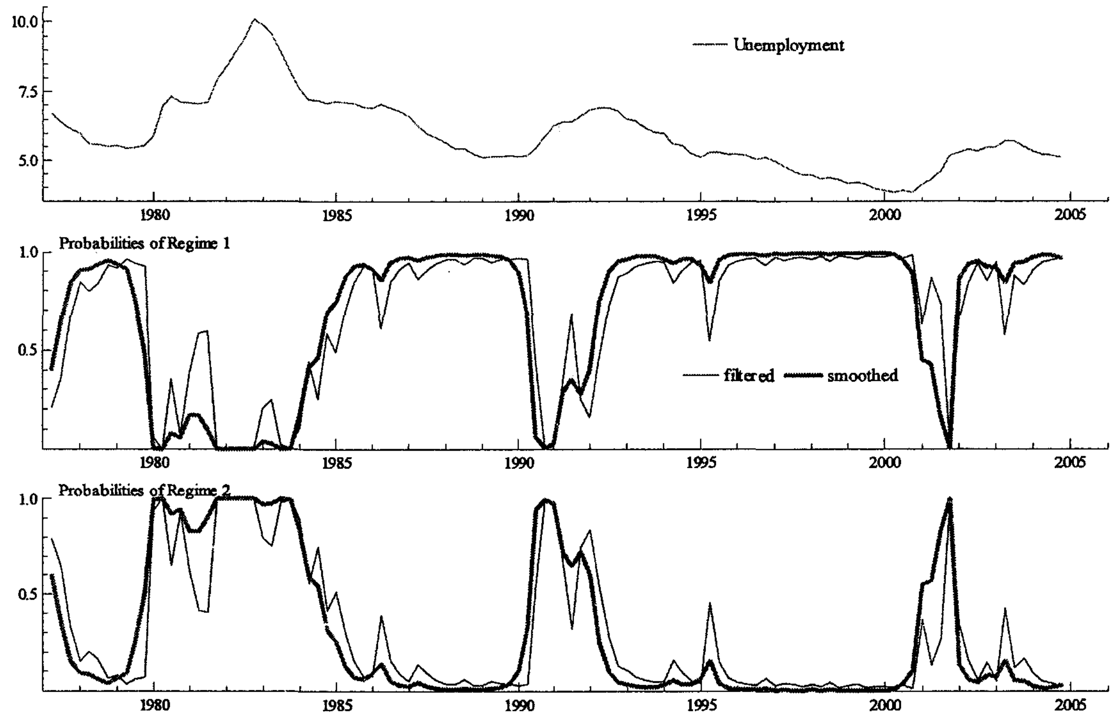

Figure 5. Series, smoothed, and filtered probabilities of the Unemployment Rate; MSIAH(2)-AR(3)

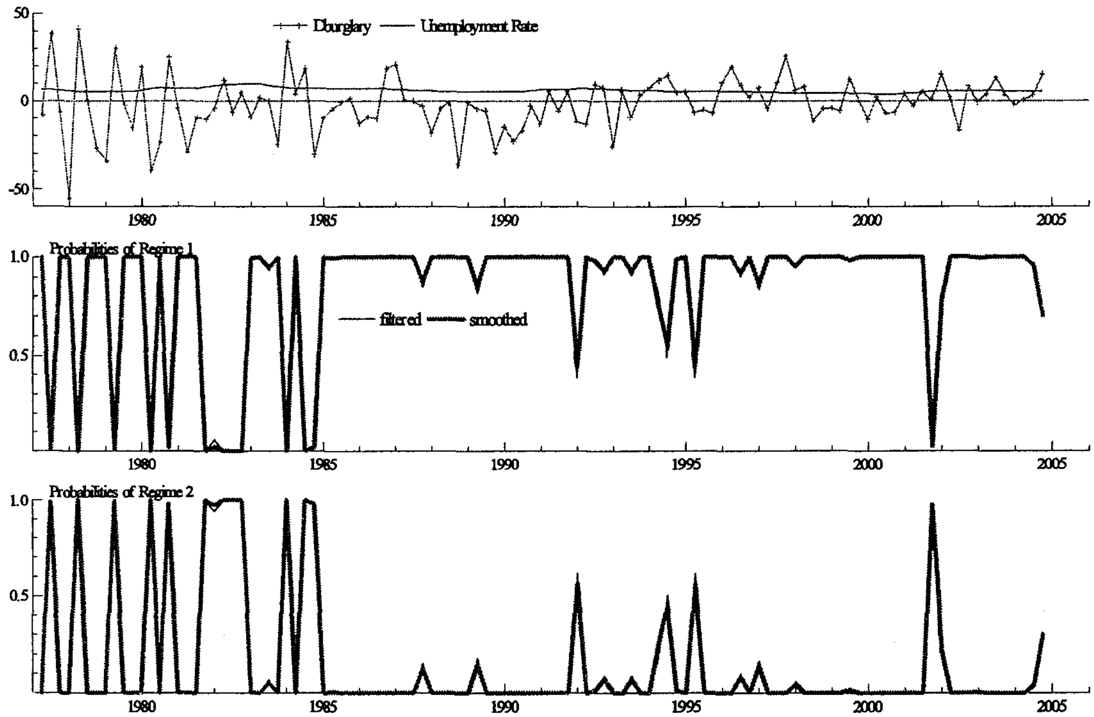

Figure 6. Series, smoothed and filtered probabilities of bivariate model for Burglary and the Unemployment Rate, MSIAH(2)-VAR(3)

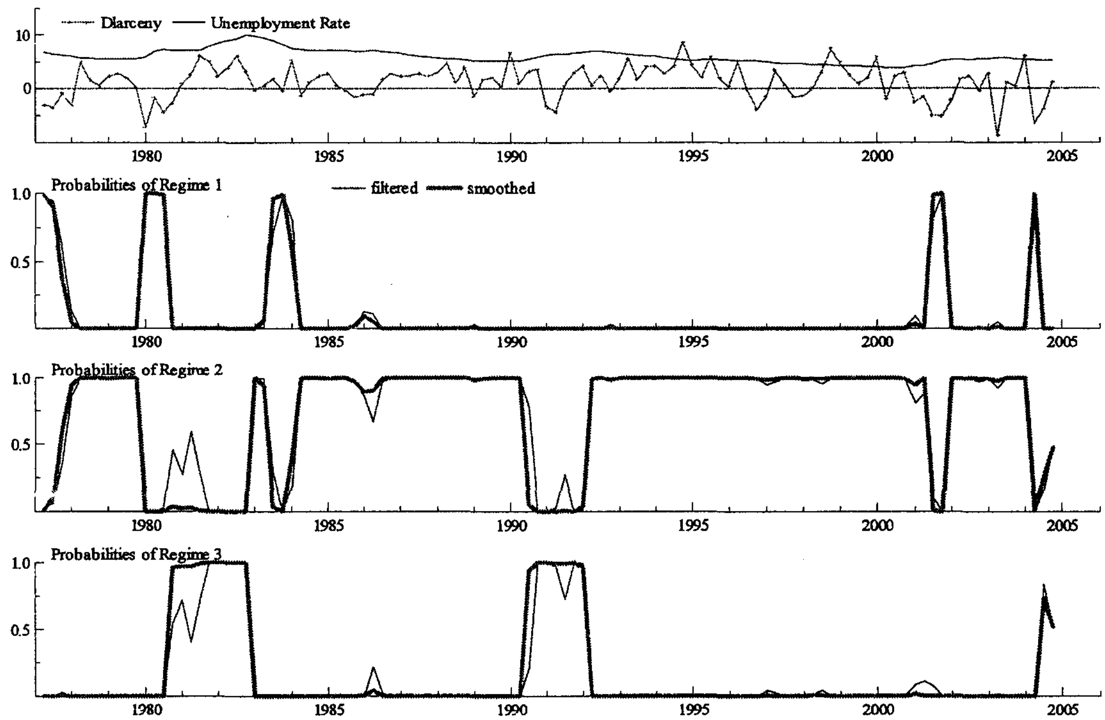

Figure 7. Series, smoothed and filtered probabilities of bivariate model for Larceny and the Unemployment Rate, MSIA(3)-VAR(2)

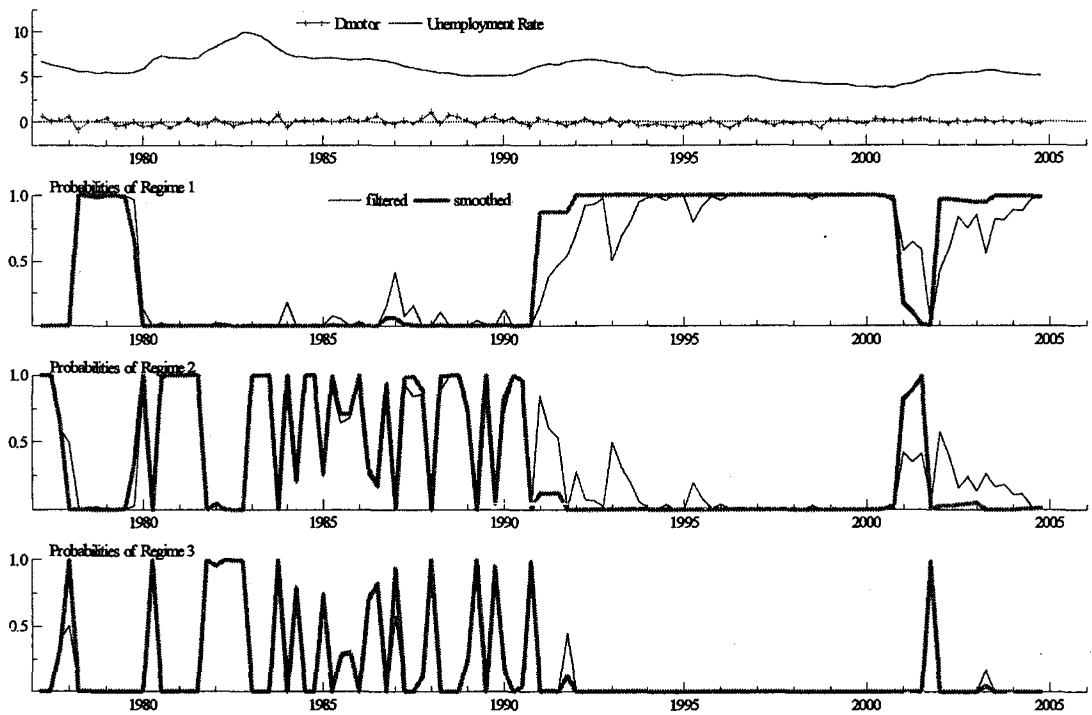

Figure 8. Series, smoothed and filtered probabilities of bivariate model for Motor and the Unemployment Rate, MSIA(3)-VAR(3)

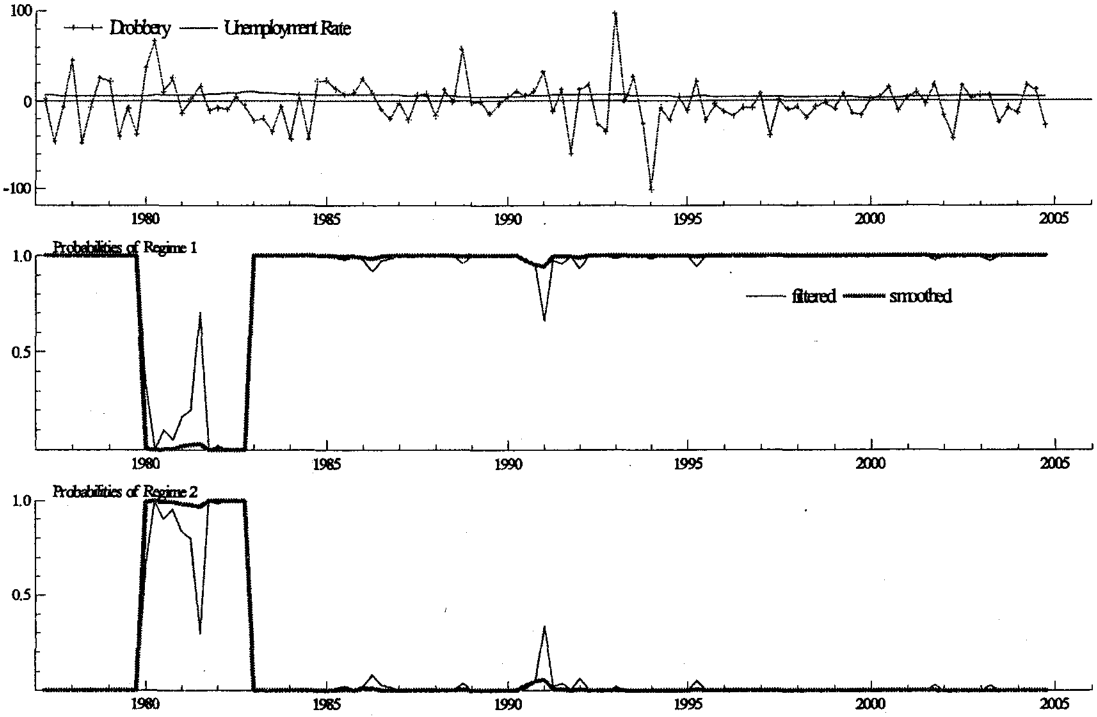

Figure 9. Series, smoothed and filtered probabilities of bivariate model for Robbery and the Unemployment Rate, MSIA(2)-VAR(2)

Chapter 2
Convergence in the Canadian Provinces: Evidence using Unemployment Rate

2.1 Introduction

Regional disparities have been at the centre of the policy makers concern for many years and have been studied by many researchers all over the world. Regional disparities are capable of creating imbalances in the development process of any country and in some countries, such issues obviously create political problems. In Indonesia, for example, the rich regions are trying to separate from the central government because people believe that their wealth is used to subsidize the poor regions. Any study of regional disparities has to start by considering two issues: 1) the definition of a region and 2) the definition of disparity. The study must define the term "region" and also how disparity will be measured. Researchers use different measures of regional disparities, such as the relative per-capita income, per-capita earning, wage, unemployment rate, and so on. The main objective of this paper is to study the disparities in the unemployment rates in the Canadian provinces using recent methodologies and data. Therefore, we use the difference in the unemployment rates of these provinces from the unemployment rate of Canada as our disparity indicator. We also want to investigate the trend of these disparities to see whether or not it has changed over last three decades.

The disparity problem is important in any country; however, it is more significant in diversified countries, like Canada. As Gunderson (1996, p.1) stated, " [b]eing a geographically large country with substantial heterogeneity in resource endowments and accesses to markets, regional issues have always been important in Canada". Canada spent a massive amount of money during the 1970s and 1980s on so called "economic development policy", to improve the economic conditions of the less developed regions.[8] In 1968, the federal government created the Department of

[8] Over 6 billion dollars during 1969-1983 (McGee, 1992, p.41).

Regional Economic Expansion (DREE) which was active until 1981. The aim of this department was to promote economic expansion to adjust for the disparities in poor provinces. By doing so the federal government tried to minimize the disparity gaps in growth rate and economic conditions among the Canadian provinces. In other words, the aim of the federal government of Canada was to create convergence among the Canadian provinces. One of the factors in this policy was the unemployment rate.

Convergence has a simple definition which comes from neoclassical growth models; see Solow (1956), for example, for a closed economy. In these models, if the regions only differ from each other in the initial level of per-capita income and capital, they will reach the same steady-state level, since the productivity of capital in the poor regions is higher than the productivity of capital in the rich regions. Therefore, the poor regions will experience a higher growth rate than the rich regions. If the poor regions grow faster than the rich regions, the former will eventually catch-up, the disparity gap will shrink and eventually close over time. On the other hand, if the growth rates of the rich regions are higher than the rates for poor regions, the disparity gap will increase over time. In this case, we say that these regions are diverging from each other.

Barro and Sala-i-Martin (1991) divided the concept of convergence into three categories: σ-convergence, absolute β-convergence and conditional β-convergence. The σ and β–convergence are the main concepts and measure two different phenomena. σ-convergence occurs when the dispersion of, say, per-capita income or unemployment rate shrinks over the span of the study. In contrast β-convergence examines whether or not the poor regions are catching-up with the rich regions. Absolute β-convergence, implies that the regions converge to the same level of, say, income or unemployment rate, whereas conditional β-convergence takes into consideration the differences in the regions, such as different endowment, weather, infrastructure, and so on. Therefore, according to conditional β-convergence, the regions could have different steady sate levels and do not need to converge to the same steady state level.

Another type of convergence considered in the literature is called "stochastic convergence", developed by Bernard and Durlauf (1995). This concept deals with the effect of shocks on the variable under study, mainly using unit root tests and cointegration techniques. To have stochastic convergence in any variable, the variable must be stationary, i.e., the shocks to the variable must die-out exponentially.

Researchers, in general, adopted the convergence definition of Sala-i-Martin (1991) and used cross sectional data and methods to test for the existence of convergence in different regions. However, such models of convergence received critiques from scholars like Friedman (1992), Quah (1993, 1996), and Bernard and Durlauf (1995). These authors criticized the definition and the use of cross section models in testing for convergence. Since the middle of the 1990s, new methods in time series, panel data models and Markov chain analysis were developed and used to study convergence. In most such studies, researchers tried to divide the sample into different sub-samples based on some major shocks, such as the Second World War, energy crises, etc. Clearly, this raised problems since the allocation of subsamples was done following prior beliefs. Consequently, other methods were developed to allow for structural breaks in the variable under study. Vogelsang (1998), for example, provides some tools to determine the break dates endogenously instead of imposing them.

Barro and Sala-i-Martin (1992), used personal income since 1840 and gross state product for the U.S. states since 1963. Using cross section models, they found that the states had converged during the span of study. Barro (1991) applied cross section models to the GDP data of 98 countries from 1960 to 1985 and concluded that the trend of growth for the countries that had lower GDP in 1960 were steeper than that of the countries with high GDP in 1960. In other words, these countries were converging to each other. The results of these two papers show that the speed of convergence is about 2% both in the U.S. states and in 98 countries included in the study. However, they found that this speed is about 1% for 20 original members of OECD, which has been considered a homogenous group of countries in the study. This finding is one of the main critiques of Quah (1996) which showed that the speed of convergence is very sensitive to the cross sectional and time dimensions of the sample. Using Monte-Carlo simulations he showed that he could get 2% convergence rate for two independent random walk variables by changing the number of cross sectional units and the time span of the sample. He concluded that the results from the convergence studies for heterogenous group of regions might be misleading. He developed a theoretical model in which the poor regions create one polar and rich regions another, and over time rich regions become more prosperous, the poor become more underprivileged and the middle class disappears.

Carlino and Mills (1993) developed necessary conditions for convergence in per-capita income in the U.S. regions; these authors identified two conditions which underly a convergence concept consistent with: 1) the series must be stationary to satisfy the stochastic convergence condition and 2) the series must also satisfy β-convergence. In other words, both stochastic and β-convergence are necessary for having convergence in any series. Myatt (1992), using annual data from 1966 to 1990 and Zellner's system of seemingly unrelated regression (SUR), studied the Canadian provincial unemployment rate disparities. He studied the trend, causes, if and remedies for the provincial unemployment rate in Canada. Results showed that the unemployment rate disparity had an upward trend, which was worsened by federal to provincial transfer payments. He concluded that there was divergence in the unemployment rates in the Canadian provinces.

Coulombe and Lee (1995) used different measurements of per-capita income and output to investigate convergence in the Canadian provinces. Their study used a pooled cross section time series data of the Canadian provinces from 1961 to 1991. They concluded that the results are in favor of the existence of convergence among the Canadian provinces. Coulombe and Day (1999) studied the evolution of the relative regional income disparities in Canada and compared it with the evolution of twelve US states bordered with Canada. They confirmed that convergence is occurring in the Canadian regions.

DeJuan and Tomljanovich (2005), using a method proposed by Vogelsang (199 8), investigated the existence of convergence in the Canadian provinces using personal income and earning from 1926 to 1996. Their results support the existence of both stochastic and β-convergences for most of the provinces, after allowing for one break in the data. In addition, using the data of personal earnings they found that the governmental redistribution program had an accelerating effect on the convergence of the provinces.

Recently, Rodríguez (2006) used two sets of data for per-capita income of the Canadian provinces from 1926 to 1999, one set including the interprovincial transfers and the other excluding these transfers, to examine convergence and assess the effects of interprovincial transfers on the convergence process. Using a class of tests proposed by Vogelsang (1998) and allowing for a break in the data, he studied convergence in two cases of known and unknown break points. The results showed that the relative per-capita income of the Canadian provinces were converging and, in addition, the interprovincial transfers did not have any role in accelerating the convergence in these provinces.

Rodríguez-Oreggia (2005) studied the convergence process in the Mexican regions using per-capita GDP. He found that after liberalization in Mexico, the northern regions, bordering the U.S., have moved away from a "falling-behind" situation to a situation the author called a "winner". On the other hand, the southern regions fell behind. In addition, he concluded that human capital played an important role in these disparities. Costanini and Lupi (2006) using panel unit root tests, investigated properties of the Italian regional unemployment rates. They found some evidence of stochastic convergence in these regions; however, there was not enough evidence to conclude that stochastic convergence existed among these regions. The authors believed that a better regional disaggregation would perhaps make it easier to form a clear and strong conclusion.

In this paper, we use some very recent econometric methods to study convergence in the unemployment rates of the ten Canadian provinces using quarterly data from the period 1976-2005. After defining the data, we check for stochastic convergence using unit root tests without and with breaks. We use three unit root tests that allow for one and two break points in the data; see Perron and Rodríguez (2003), Rodríguez (2007), and Lumsdaine and Papell (1997). Next, we use three different methods to investigate β-convergence in the data. These methods are introduced by Vogelsang (1998), Perron and Yabu (2005, 2006), and Bai and Perron (1998, 2003). The first two methods are robust to the presence of unit roots, however, the last method is not. In addition, except for the latter method, which allows for multiple breaks, the other methods allow for only one break point in the data. We find strong evidence of convergence in most of the provinces by introducing structural breaks in the data.

The remainder of this paper is organized as follows. The next section provides the methodological issues and the class of tests that will be used in this paper, including different unit root tests to study stochastic convergence. We also describe a class of tests proposed by Vogelsang (1998), Perron and Yabu (2005, 2006), and Bai and Perron (1998, 2003). Section 3 provides a brief look at the data and the empirical results from various methods. Finally, section 4 concludes.

2.2 Methodology

2.2.1 Stochastic convergence

Let y_t^i denote the difference between the unemployment rate of the province i at period t (u_t^i) and the unemployment rate of Canada (u_t^{can}). Stochastic convergence implies that shocks to y_t^i will have temporary effects. Formally, stochastic convergence exists if y_t^i is stationary. Consequently, a common test for stochastic convergence will be testing for a unit root in y_t^i . Rejection of the unit root hypothesis supports the existence of stochastic convergence. Conversely, failure to reject the unit root hypothesis will be an indicator for stochastic divergence. There are many unit root tests in the literature. However, in this paper we use the following unit root tests: Augmented Dickey-Fuller (ADF) test (Dickey and Fuller (1979) and Said and Dickey (1984)), ADF^{GLS} test of Elliott, Rothenberg, and Stock (ERS) (1996), a unit root test proposed by Perron and Rodríguez (2003) which considers one break under the alternative hypothesis, and a minimum LM unit root test proposed by Lumsdaine and Papell (1997) which considers two breaks under the alternative hypothesis.

The Augmented Dickey-Fuller (ADF) test

This test has been proposed by Dickey and Fuller (1979), Said and Dickey (1984). The test consists of estimating the following regression:

$$\Delta\tilde{y}_t = \alpha\tilde{y}_{t-1} + \sum_{i=1}^{k}\gamma_i\Delta\tilde{y}_{t-i} + \epsilon_t$$

where $\tilde{y}_t = y_t - \hat{\mu}$ or $\tilde{y}_t = y_t - \hat{\mu} - \hat{\beta}t$, y_t is the time series being tested, t is the time trend, Δ is the first difference operator and k is the number of lags, which must be determined in such a way that the ϵ_t become serially uncorrelated. Using this estimation, we test for a null hypothesis that $\alpha = 0$. If we cannot reject this hypothesis, then we have a unit root and the series is not stationary.

ADF^{GLS} test

To improve the power of the unit root test, ERS (1996) have proposed a new method for unit root testing called ADF^{GLS}. This is an ADF type test, which uses GLS detrended data. The ADF^{GLS} statistic can be constructed using the following equation

$$\Delta\tilde{y}_t = \alpha\tilde{y}_{t-1} + \sum_{i=1}^{k}\gamma_i\Delta\tilde{y}_{t-i} + e_{tk}$$

with $\tilde{y}_t = y_t - \hat{\psi}^{GLS} z_t$ and the set of coefficients $\hat{\psi}^{GLS}$ minimizes

$$S(\psi,\bar{\alpha}) = \sum_{t=1}^{T}[y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}}]^2$$

where $y_t^{\bar{\alpha}} = [y_1, (1-\bar{\alpha}L)\, y_t]$, $z_t^{\bar{\alpha}} = [z_1, (1-\bar{\alpha}L)z_t]$, for $t = 2, ..., T$; and $\bar{\alpha} = 1 + \bar{c}/T$ representing a value under the alternative hypothesis. In addition, we have $z_t = \{1\}$ or $z_t = \{1, t\}$, which represent the deterministic components included in the detrending regression. ERS define a new time series as

$$y_t^{\bar{\alpha}} = \begin{cases} y_t, & \text{if } t = 1 \\ y_t - \bar{\alpha}y_{t-1}, & \text{if } t > 1, \end{cases}$$

where $\bar{\alpha} = 1 + \bar{c}/T$. ERS recommend the use of $\bar{\alpha} = 1 - 7/T$ for a model with only an intercept and $\bar{\alpha} = 1 - 13.5/T$ for a model with an intercept and a time trend.

At the next step, an ADF test can be carried out using the transformed data as follows

$$\tilde{y}_t = \alpha\tilde{y}_{t-1} + \sum_{i=1}^{k}\gamma_i\tilde{y}_{t-i} + \eta_t.$$

As with the ADF test, we consider the $t-$ratio of $\hat{\alpha}$ from this regression. This $t-$ratio follows a Dickey-Fuller distribution when only an intercept has entered in this equation. However, when we include both an intercept and a time trend, the distribution differs. ERS simulated the critical values for the latter case.

Perron (1989) showed that ADF type tests fail to reject the unit root null hypothesis when the series is stationary with a broken trend. In other words, the failure to reject the null hypothesis of unit root might arise from exogenous changes in the data generating process. Hence, he proposed allowing for one break in the ADF test, assuming that this break point is known and exogenous. However, the problem with this method is that we must know the break date prior to testing for a unit root. To overcome this problem other unit root tests, including Zivot and Andrews (1992), Perron (1997), Lumsdaine and Papell (1997), and Perron and Rodríguez (2003), *inter alia*, have been developed, which allow determination of the break date as an endogenous parameter in the model.

An ADF^{GLS} unit root test with 1 break

Perron and Rodríguez (2003) extended the GLS detrending approach of ERS (1996) and Ng and Perron (2001) to the context of structural change. They considered a break, under the alternative hypothesis, in the slope or a break in both the intercept and slope of the trend function of y_t. The break point is determined endogenously. As in the case of the ADF^{GLS} without structural change, we have $z_t = \{1, t, 1(t \geq T_B) \times (t - T_B)\}$ or $z_t = \{1, 1(t \geq T_B), t, 1(t \geq T_B)(t - T_B)\}$ which represents the deterministic components included in the deterending regression. Using simulations, Perron and Rodríguez (2003) selected $\bar{c} = -22.5$ as the optimal non-centrality parameter. Therefore, this value is used in this paper to carry out this test. To determine the optimal lag we can use information criteria. Perron and Rodríguez (2003) also introduced the structural break in a group of unit root tests called M-tests, which were originally suggested by Stock (1999).[1]

Recently, Rodríguez (2007) constructed a new version of these tests, which only allow a break in the intercept. We also use these tests to check for the existence of unit root in the data under study.

Unit root tests with 2 breaks

The previous test allows for only one break in each time series; however, it is possible to have multiple structural breaks in the series. Therefore, considering only one endogenous break point may not be sufficient and this could cause a loss in information when there is more than one break (Lumsdaine and Papell, 1997). Lumsdaine and Papell (1997) proposed a new procedure that allows one to examine the significance of two structural breaks. They use a modified version of the ADF

---

[1] Further details about these unit root tests can be found in Perron and Rodríguez (2003).

test with two endogenous breaks as follows:

$$\Delta y_t = \mu + \beta t + \theta DU1_t + \delta DT1_t + \omega DU2_t + \varphi DT2_t + \alpha y_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta y_{t-i} + \epsilon_t$$

where $DU1_t = 1$ if $t > TB1$ and zero otherwise; $DU2_t = 1$ if $t > TB2$ and zero otherwise; $DT1_t = t - TB1$ if $t > TB1$ and zero otherwise; and finally $DT2_t = t - TB2$ if $t > TB2$ and zero otherwise. $DU1_t$ and $DU2_t$ capture the structural breaks in the intercept at time $TB1$ and $TB2$, respectively. The other two dummy variables, $DT1_t$ and $DT2_t$, capture the structural changes in the slope at time $TB1$ and $TB2$, respectively. The optimal lag length can be determined based on the general to specific approach using t tests as suggested by Ng and Perron (1995). The break points are selected based on the minimum value of the t statistic for α.

2.2.2 β-convergence

In this section, we discuss the methodologies that can be used to test for β-convergen ce. These methods use a linear deterministic trend function. Let y_t define the measure of the disparities. We define the disparities as the difference between the unemployment rate of each province and the unemployment rate of Canada. β-convergence implies that the provinces that initially had an unemployment rate lower than the national rate of unemployment will have a positive trend in the unemployment rate and the provinces with a higher initial unemployment rate than the national rate will have a negative trend. These requirements can be directly linked to the parameters of the deterministic trend function of y_t.

Suppose that y_t is modeled as

$$y_t = \mu + \beta t + \epsilon_t$$

where μ represents the initial level of y_t, β represents the average change at y_t, and ϵ_t is a serially correlated random process with mean zero. β-convergence requires to have a negative relationship between μ and β, i.e. if $\mu > 0$ then $\beta < 0$ and if $\mu < 0$ then $\beta > 0$. Therefore, the evidence on β-convergence can be obtained from the estimates of the trend function.

If ϵ_t is serially uncorrelated, then we can use the OLS estimates of μ and β. These estimates will be unbiased and efficient. However, in practice ϵ_t might be autocorrelated and may even have a unit root. In addition, the interpretation of the trend parameters in the autoregressive representation of y_t depends on the nature of ϵ_t. When ϵ_t is an $I(0)$ process, then the inference about β can be obtained from the estimate of the slope. But if ϵ_t is an $I(1)$ process this coefficient is zero and inference must be made from an estimate of the intercept in the autoregressive representation of y_t. Therefore, the possibility of a unit root in ϵ_t can make the interpretation of the coefficients of the trend function very difficult. To overcome this problem, we need to use methods that are robust to the statistical properties of ϵ_t.

In this paper, we use two such methods. These methods are proposed by Vogelsang (1998) and Perron and Yabu (2005, 2006). They give the possibility of estimating and making inference about β-convergence. Furthermore, we apply a method proposed by Bai and Perron (1998, 2003) to estimate the trend function. This method is not robust to the existence of unit root, and therefore can only be applied to time series that are stationary.

The method of Vogelsang

Vogelsang (1998), proposed a class of statistics that consists in estimating two OLS regressions. The first, y_t regression, is given as follows

$$y_t = \mu_1 DU_{1t} + \beta_1 DT_{1t} + \mu_2 DU_{2t} + \beta_2 DT_{2t} + \epsilon_t \tag{2.1}$$

where $DU_{1t} = 1$ if $t \leq T_B$ or zero otherwise; $DU_{2t} = 1$ if $t > T_B$ or zero otherwise; $DT_{1t} = t$ if $t \leq T_B$ or 0 otherwise, and $DT_{2t} = t - T_B$ if $t > T_B$ and 0 otherwise. T_B is the date of a shift in the parameters of the trend function of y_t and is considered either known or unknown. However, if this date is treated as unknown it can be estimated from the data. The parameters μ_1 and μ_2 show whether the unemployment rate of the province is higher or lower than the unemployment rate of Canada in periods 1 and T_B, respectively.

The second regression, z_t regression, is given by:

$$z_t = \mu_1 DT_{1t} + \beta_1 SDT_{1t} + \mu_2 DT_{2t} + \beta_2 SDT_{2t} + S_t \tag{2.2}$$

where $z_t = \sum_{j=1}^{t} y_j$, $SDT_{it} = \sum_{j=1}^{t} DT_{ij}$, $i = 1, 2$, $S_t = \sum_{j=1}^{t} \epsilon_j$ and DT_{it} is as defined above. This regression is obtained using the partial sums of y_t.

To test for convergence, we need to assess whether μ_1, μ_2, β_1, and β_2 are statistically significant and have the appropriate signs. For this purpose, a modified class of statistics proposed by Vogelsang (1998) can be used. Let t_y and t_z denote the $t-$statistics for testing the null hypothesis that the individual parameters in the y_t and z_t regressions are zero. The modified $t-$statistics for the y_t regression is $T^{-1/2}t_y$, where T is the sample size. On the other hand, the modified $t-$statistics for the z_t regression are defined as $t - PS_T = T^{-1/2}t_z \exp(-bJ_T)$, where b is a constant (to be calculated) and J_T is T^{-1} multiplied by the Wald statistic for testing $c_2 = c_3 = ... = c_9 = 0$ in the following regression[9]

$$y_t = \mu_1 DU_{1t} + \beta_1 DT_{1t} + \mu_2 DU_{2t} + \beta_2 DT_{2t} + \sum_{i=2}^{9} c_i t^i + \epsilon_t. \qquad (2.3)$$

The J_T term can be calculated as $(RSS_Y - RSS_J)/RSS_J$, where RSS_Y is the sum of the squared residuals from regression (2.1), and RSS_J is the sum of the squared residuals from regression (2.3). For a given significance level, the constant value of b can be computed in such a way that the critical values of $t-$statistics are the same whether ϵ_t is $I(0)$ or $I(1)$. Consequently, the modified $t-$tests by J_T from the z_t regression are robust to $I(1)$ errors. Note that if $b = 0$, the J_T modification will not have any effect on t-tests. Therefore, the distribution of $t - PS_T$ is different when ϵ_t is $I(1)$ compared to the case when ϵ_t is $I(0)$. Hence, it is recommended to only use $b = 0$ if we are sure that the errors are $I(0)$. There is no need for any modification in the y_t regression because the $T^{-1/2}t_y$ statistics has a well defined asymptotic distribution when ϵ_t is $I(1)$ and when ϵ_t is $I(0)$ this statistic converges to zero. Therefore, test based on the $T^{-1/2}t_y$ statistic is conservative when the errors are stationary.

The asymptotic distributions of $T^{-1/2}t_y$ and $t - PS_T$ are not standard and they depend on the break date used in the regressions. Therefore, the break date will affect the critical values for these statistics. The break date can be assumed known or unknown. In the latter case, it must be estimated from the data. To prevent the estimation of break points too close to the beginning and the end of the sample, we need to use trimming. Depending on the sample size, it is possible to use 5%, 10%, up to 25% trimming. For example, if we use 10% trimming from the beginning and

[9] J_T-statistic is a unit root statistic, which was proposed by Park and Choi (1988) and Park (1990).

the end of sample, the new sample will be $(0.1T, 0.9T)$. Next, from each regression we compute T^{-1} multiplied by the Wald statistics in order to test the joint hypothesis that $\mu_1 = \mu_2$ and $\beta_1 = \beta_2$. In other words, to test the null hypothesis that there is no break in the trend function of the time series y_t. The estimated break date is the break date that has the largest normalized Wald statistic. The critical values of the $T^{-1/2}t_y$ and $t - PS_T$ statistics are taken from Vogelsang (1998).

The approach of Perron and Yabu

Perron and Yabu (2005, 2006) have proposed another method to estimate the trend function without any knowledge about the statistical characteristics of the error terms. Their test is based on Feasible quasi GLS. Let us assume that the univariate time series y_t, $t = 1, ..., T$, has been generated by

$$y_t = x_t'\Psi + \epsilon_t,$$

$$\epsilon_t = \alpha\epsilon_{t-1} + e_t,$$

for $t = 1, ..., T$ where $e_t \sim i.i.d.(0, \sigma^2)$, x_t is a $(r \times 1)$ vector of deterministic components and Ψ is a $(r \times 1)$ vector of unknown parameters. In addition, $-1 \leq \alpha \leq 1$ so that both stationary and integrated errors are allowed in the data generating process. The null hypothesis is $R\Psi = \omega$ where R is a $(q \times r)$ full rank matrix, ω is a $(q \times 1)$ vector, and q is the number of the restrictions.

To test for a structural change in the trend function three models can be considered:

1. Structural change in intercept : In this case, $x_t = (1, DU_t, t)'$ and $\Psi = (\mu_1, \mu_2, \beta_1)$ where $DU_t = 1(t > T_B)$, T_B is the break date and $T_B = [\lambda_1 T]$. The hypothesis of interest in this case will be $\mu_2 = 0$.

2. Structural change in slope: In this case, $x_t = (1, t, DT_t)'$ and $\Psi = (\mu_1, \beta_1, \beta_2)$ where $DT_t = \mathbf{1}(t > T_B)(t - T_B)$. The null hypothesis in this case will be $\beta_2 = 0$.

3. Structural change both in the intercept and the slope. In this case, $x_t = (1, DU_t, t, DT_t)'$ and $\Psi = (\mu_1, \mu_2, \beta_1, \beta_2)$. The hypothesis of interest in this case is $\mu_2 = \beta_2 = 0$.

In addition, $\mathbf{1}(.)$ is the indicator function, $\lambda_1 \in (0,1)$, and $[.]$ denotes the largest integer. In all these models, we allow for only one break in the intercept or slope or in both. Now consider the case that the break date and the autoregressive coefficient, α, is known. Then, the GLS estimate of the parameters can be obtained applying OLS to the following regression

$$(1-\alpha L)y_t = (1-\alpha L)x_t'\Psi + (1-\alpha L)\epsilon_t$$

for $t = 2, ..., T$ together with $y_1 = x_1'\Psi + \epsilon_1$.

In this case, the Wald statistic for testing the $R\Psi = \omega$, is asymptotically distributed as a χ^2 random variable. However, in practice the break date and α are unknown. Therefore, the methodology must be extended accordingly, in order to determine the break date. Following Andrews (1993) and Andrews and Ploberger (1994), Perron and Yabu (2005, 2006) consider the following three functionals of the Wald test for different break dates:

$$\begin{aligned} Mean - W_{FS} &= T^{-1}\sum\nolimits_{\Lambda} W_{FS}(\lambda_1') \\ Exp - W_{FS} &= \log[T^{-1}\sum\nolimits_{\Lambda} \exp(\frac{1}{2}W_{FS}(\lambda_1'))] \\ \sup - W_{FS} &= \sup_{\Lambda} W_{FS}(\lambda_1') \end{aligned}$$

where $\Lambda = \{\lambda_1'; \epsilon \leq \lambda_1' \leq 1-\epsilon\}$ for some $\epsilon > 0$. Notice that T_B' denotes a generic break date, which is used to construct the Wald statistic and T_B denotes the true break date. These statistics are optimal only when the errors are stationary. The asymptotic critical values are calculated by Perron and Yabu (2005). Results suggest that although the limiting distribution of these statistics are different, when considering the $Exp - W_{FS}$ functional, the relevant quantiles are very similar for the $I(0)$ and $I(1)$ cases. On the other hand, simulation results show that the Mean and Sup versions of the test are less useful than the Exp version.

We know that the OLS estimate of the autoregressive coefficient is biased downward especially when α is near unity. Therefore, to improve finite sample properties, Perron and Yabu (2005) propose two methods: 1) use median unbiased estimates as described by Andrews (1993), 2) use a modified version of the weighted symmetric least square estimates described by Roy and Fuller (2001). When using the first method, we replace the OLS estimate of $\hat{\alpha}$ by a value that would create $\hat{\alpha}$ as the median of distribution, assuming normal errors. This specific correction depends on the nature of the deterministic components and the values for the linear trend case are tabulated in Andrews (1993). The procedure recommended by Perron and Yabu (2005, 2006) is to use the modified version of Roy and Fuller (2001) method as following:

1. For any given break date, use OLS to detrend the data and obtain the residuals $\hat{\epsilon}_t$;

2. Estimate an $AR(1)$ for $\hat{\epsilon}_t$ in order to find the estimate of $\hat{\alpha}$ and the $t-$ratio, $\hat{\tau}$;

3. Use $\hat{\alpha}$ and $\hat{\tau}$ to get the Roy and Fuller (2001) bias corrected estimate $\hat{\alpha}_M$;

4. Apply the following truncation

$$\hat{\alpha}_{MS} = \begin{cases} \hat{\alpha}_M, & \text{if } |\hat{\alpha}_M - 1| > T^{-1/2}, \\ 1, & \text{if } |\hat{\alpha}_M - 1| \leq T^{-1/2}. \end{cases}$$

5. Use $\hat{\alpha}_{MS}$ and GLS method to estimate of the coefficients of the trend function and obtain the estimate of the variance of the residuals. Then create the standard Wald statistic, denoted W_{FMS}.

6. For a case with an unknown break date, repeat the previous steps for all possible break dates and construct the Exp-Wald statistic, denoted $Exp - W_{FMS}$.

Throughout the paper the subscript RQF stands for Robust Quasi Feasible GLS. For a further treatment of these tests, see Perron and Yabu (2005, 2006).

The approach of Bai and Perron

Both of the above methodologies allow for only one break, however, it is possible for some time series to have multiple breaks. To this end, Bai and Perron (1998, 2003) proposed a method to estimate the multiple structural changes in a linear model using OLS. Their method gives the possibility of testing for break against stability of the model. In addition, it is possible to test for a given number of breaks using different tests, including information criteria, such as BIC and LWZ (Liu et al, 1997).

Following the same notation as Bai and Perron (1998, 2003), consider the following multiple linear regression with m breaks:

$$y_t = \begin{cases} x_t'\beta + z_t'\delta_1 + \epsilon_t & t = 1, 2, \ldots, T_1 \\ x_t'\beta + z_t'\delta_2 + \epsilon_t & t = T_1 + 1, \ldots, T_2 \\ \vdots & \\ x_t'\beta + z_t'\delta_{m+1} + \epsilon_t & t = T_m + 1, \ldots, T \end{cases}$$

where x_t $(p\times 1)$ and z_t $(q\times 1)$ are vectors of covariates and β and δ_j $(j = 1, 2, ..., m+1)$ are the corresponding vectors of coefficients and ϵ_t is the disturbance term at time t. This method attempts to estimate the unknown coefficients along with the unknown break dates. For a pure structural change model, where all the parameters can have shifts, $p = 0$, whereas, for a partial structural change model, where some of the parameters do not have any break or shift, we have $p > 0$ and $q > 0$.

In the estimation process we use OLS. To obtain the estimates of the coefficients and the break dates, we can use an iterative procedure as follows. Assuming $\theta = (\delta,\, T_1, ..., T_m)$, define the sum of squared residuals as $SSR(\beta,\, \theta)$. Using the method introduced by Sargan (1964), minimize $SSR(\beta,\, \theta)$ by iteration. First, assume β is fixed and minimize $SSR(\beta,\, \theta)$ with respect to θ and then keep θ constant and minimize $SSR(\beta,\, \theta)$ with respect to β. Iterate this procedure until the decrease in the objective function becomes less than a given number. Bai and Perron (1998, 2003) recommend, for efficiency purpose, to keep $\{T_1, ..., T_m\}$ constant at the second step and to minimize the objective function with respect to β and δ simultaneously.

To estimate the number of breaks Perron and Yabu propose different statistics. The first statistic is the sup F type test for no break, $m = 0$, against an alternative hypothesis of $m = k$ breaks. Let R be a matrix such that $(R\delta)' = (\delta_1' - \delta_2', ..., \delta_k' - \delta_{k+1}')$; define

$$F_T(\lambda_1, \ldots, \lambda_k; q) = \frac{1}{T}(\frac{T-(k+1)q-p}{kq})\hat{\delta}' R'(R\hat{V}(\hat{\delta})R')^{-1}R\hat{\delta}$$

where $\hat{V}(\hat{\delta})$ is the estimate of the variance covariance matrix of $\hat{\delta}$ which is robust to serial correlation and heteroscedasticity. Now, define the $\sup F_T(k;q) = F_T(\hat{\lambda}_1, ..., \hat{\lambda}_k; q)$ where $\hat{\lambda}_1, ..., \hat{\lambda}_k$, $\hat{\lambda}_i = \hat{T}_i/T$, minimize the global $SSR(.)$.[10] The asymptotic distribution of this statistic depends on the minimal length h for every segment or in another words, it depends on the trimming parameter. The critical values can be found in Bai and Perron (2003). In addition, the maximum number of the breaks depends on the trimming parameter, for example for a 15% trimming, the maximum number of breaks will be 5.

In using the above-mentioned statistic, the researcher must test no break versus a pre-specified number of breaks. However, Bai and Perron propose two other tests, called double maximum tests, which can be used to test for no break against an unknown number of breaks given some upper limit for the number of breaks M. The first test is defined as $UD_{MAX}F_T(M,q) = \max_{1\leq m\leq M} F_T(\hat{\lambda}_1, ..., \hat{\lambda}_m; q)$. The second test denoted by $WD_{MAX}F_T(M,q)$ is defined as $WD_{MAX}F_T(M,q) = \max_{1\leq m\leq M} \frac{c(q,\alpha,1)}{c(q,\alpha,m)} F_T(\hat{\lambda}_1, ..., \hat{\lambda}_m; q)$, where $\hat{\lambda}_j$ are obtained using global minimization of the $SSR(.)$ and $c(q, \alpha, m)$ is the critical value of the $\sup F_T(k;q) = F_T(\hat{\lambda}_1, ..., \hat{\lambda}_k; q)$ for a given significance level α. The critical values can be found in Bai and Perron (2003).

Another statistic is proposed to test for a given number of breaks l, against $l+1$ breaks, which denoted $\sup F_T(l+1|l)$. A model with $l+1$ break will be selected if the overall minimal value of the $SSR(.)$ is sufficiently smaller than the $SSR(.)$ from a model with l breaks. We can repeat this test for $l = 1$ to $l = M$. Additionally, the number of breaks can be determined using the information criteria, Bayesian Information Criteria (BIC) and Liu, Wu, and Zidek (LWZ) (Liu et al, 1997) defined

---

[10] This is equivalent to maximizing the $F-$test assuming spherical errors.

as follows:

$$
\begin{aligned}
BIC(m) &= \ln \hat{\sigma}(m) + g \ln(T)/T \\
LWZ(m) &= \ln(\frac{S_T(\hat{T}_1, ..., \hat{T}_m)}{T-g}) + (\frac{g}{T} \times c_0(\ln(T))^{2+\delta_0})
\end{aligned}
$$

where $g = (m+1)q + m + p$, and $\hat{\sigma}^2(m) = T^{-1}S_T(\hat{T}_1, ..., \hat{T}_m), c_0 = 0.299$ and $\delta_0 = 0.1$. Any number of breaks that minimizes these information criteria is the selected number of breaks for that time series. Finally, repartition and sequential procedures can be used to determine the exact number of breaks.

2.3 Empirical results

Quarterly data on unemployment rates from 1976-2005 for Canada and the ten Canadian provinces are used in this study. The ten provinces are Alberta, British Columbia (BC), Manitoba, New Brunswick (NB), Newfoundland (NF), Nova Scotia (NS), Ontario, Prince Edward Island (PEI), Quebec, and Saskatchewan.

We denote the unemployment rate of each province with ur_t^i and the unemployment rate of Canada with ur_t^{can}. We are interested in examining the trend of the regional disparities of the unemployment rate in the Canadian provinces. We define disparities as the difference between the unemployment rates of each province and the unemployment rate of Canada as follows

$$y_t^i = ur_t^i - ur_t^{can}.$$

Here, the unemployment rate of Canada, ur_t^{can}, is the reference variable. Therefore, convergence means that the unemployment rates of the provinces move toward the unemployment rate of Canada. Figure 1 and 2 show y_t^i for each province during 1976-2005. Figure 1, shows y_t^i for the other group of provinces, which includes Atlantic Canada and Quebec. Figure 2, represents the y_t^i series for a group of provinces where unemployment rates were close and usually less than the unemployment rate in Canada. The unemployment rate in this group was higher than the national rate of unemployment during the sample under study.

Following Carlino and Mills (1993), convergence exists if stochastic convergence and β-convergence are verified, as explained in the previous section. Stochastic convergence implies that shocks to the deviation of the unemployment rate of a given province from Canada's unemployment rate will be temporary and will only have a temporary effect. Consequently, a common test for stochastic convergence will be testing for a unit root in the deviation of the unemployment rate of the province from the national rate of unemployment in Canada. Rejection of the unit root supports stochastic convergence; on the other hand, failure to reject the unit root is an indicator for stochastic divergence. In other words, the stochastic convergence holds if y_t^i does not have unit root.

2.3.1 Stochastic Convergence

Table 1 presents the results obtained from the application of ADF and ADF^{GLS} unit root tests on these variables. Lag length is selected using BIC information criterion. Results show that the null hypothesis of unit root can be rejected for Manitoba, NB, NF, NS, and Quebec by both tests. However, for PEI and Saskatchewan the null hypothesis can be rejected only using ADF^{GLS} statistics, which is more powerful than ADF. At the same time, neither ADF nor ADF^{GLS} can reject the existence of unit root for Alberta, BC, and Ontario.

On the other hand, the graphs of these time series suggest some breaks in the variable, so ADF and ADF^{GLS} tests may fail to reject the existence of unit root due to misspecification of the deterministic trend. To surmount this problem we use unit root tests which allow for break at the trend function. We use the BIC and modified AIC information criteria and MSB and P$_T$ statistics from Perron and Rodríguez (2003) and Rodríguez (2007); in addition, we denote the Lumsdaine and Papell (1997) test statistic by LP. $T_{B_1}$and T_{B_2} show the estimated break dates. Also, we allow for break in both the intercept and the trend in Perron and Rodríguez (2003) test and break in the intercept in Lumsdaine and Papell (1997).

The results from applying these tests[11] are shown at Table 2. Now, we can reject the unit root for Alberta and Ontario by allowing for two break points in the trend function. The only province that we cannot reject the null hypothesis of unit root is BC. Therefore, stochastic convergence holds for all provinces except for BC.

[11] The results from applying test of Rodríguez (2007), not reported here, do not differ from the results from the test of Perron and Rodríguez (2003).

2.3.2 β-Convergence

Tables 3–5 present the results obtained using the $t - PS_T$ without the J_T correction, the $t - PS_T$ with the J_T correction, and the $T^{-1/2}t_y$ statistic, respectively. These statistics are calculated considering an unknown break date in the regressions. The last column in these tables shows the estimated break dates.

Table 3 reports the estimated μ, β and the break dates using regression z_t (2.2). The $t - PS_T$ statistics without the J_T correction are given in parenthesis below each coefficient and the asymptotic critical values are given in the bottom two rows. The estimates of μ_1 for provinces are statistically different from zero for most of the provinces. This indicates that disparities existed between the unemployment rate of the provinces and the national level of the unemployment rate in 1976. In addition, the results reveal that there is more evidence for deterministic convergence after the break than before, in the Canadian provinces. However, we should be cautious in using these results since these statistics are obtained assuming that the residuals are $I(0)$ process (Tomljanovich and Vogelsang, 2002). Even clearly stationary residuals could spuriously increase the $t - PS_T$ statistics in a small sample.

The $t - PS_T$ statistics with J_T correction are shown in table 4. The estimated coefficients are the same as table 3, but the statistics are smaller now. In this table $t - PS_T$ statistics are reported for 5% and 10% in parenthesis below each coefficient. These results show that before the break, deterministic convergence is found in Manitoba, NF, NS, and Saskatchewan and there is a divergence in NB. On the other hand, after the estimated break, only PEI and NS demonstrate deterministic convergence and NF shows deterministic divergence.

Table 5 reports the results using the $T^{-1/2}t_y$ statistics. In the periods before the break, there was convergence in NF, NS, and Saskatchewan; at the same time Alberta, Manitoba, NB, and Quebec have experienced divergence. After the break, there was convergence in PEI and divergence in NF.

Using the same notation as Rodríguez (2006) Table 6, summarizes the results of Tables 3, 4, and 5. In this table, a (capital) C denotes point estimates consistent with β-convergence, that is, $\mu > 0$ and $\beta < 0$, or $\mu < 0$ and $\beta > 0$. Also, in this case both estimates are statistically significant at least at the 10 percent level. A (small) c denotes point estimates consistent with β-convergence but only with one coefficient statistically significant at least at the 10 percent level. Divergence is indicated using the D and d symbols, where D indicates that both coefficients are statistically significant and d signifies that only one coefficient is statistically significant. The symbol E denotes point estimates that are not statistically different from zero and are small in magnitude, which implies that β-convergence has already occurred. Lastly, the symbol u means that no conclusion is possible using all information from Tables 3-5. This situation is characterized when the coefficients are not significant but they are not small enough in magnitude to be considered as equilibrium (E).

As may be checked from Table 6, there is no evidence of β-convergence at all for NB and, in fact, results are in favor of divergence in this province. This divergence is observed for this province in both periods before and after the break date. On the other hand, β-convergence has or is occurring for NS and Saskatchewan all the time. At the same time in NF we observe that β-convergence was occurring before the break and after the break there is strong evidence of divergence, that is the disparities between the unemployment rate of this province and national unemployment rate of Canada is increasing. Divergence is observed for PEI before the estimated break date; however, evidence indicates β-convergence after the break. We see β-convergence in Alberta only after the estimated break. This is also the case for Quebec. The results for Ontario are inconclusive and, finally, for BC the results support β-convergence and/or equilibrium.

These results are consistent with the results of another paper by Rodríguez (2006). In this paper, the author used the same methodology to study convergence in the per-capita personal income of the Canadian provinces and to investigate the effect of interprovincial transfers on divergence or convergence. He found that the most of provinces are converging or have converged into the unweighted provincial average. In addition, results of this study showed that the interprovincial transfers did not have any significant role on the convergence of the provinces.

Results from applying the method of Perron and Yabu (2005, 2006) are presented in Table 7. We reject the null hypothesis that the trend function is stable in favor of a trend function with a break in all provinces with the exception of Manitoba, NB, PEI, and Saskatchewan. This table also shows the estimated break date obtained by minimizing the sum of squared residuals from a regressing y_i on a constant, a time trend, and dummies for shift in level and slope; these break dates are the same as the break dates estimated using the method of Vogelsang (1998). This table also shows the estimated μ and β for the periods before and after the estimated breaks. According to these results, Ontario and Quebec experience divergence from the national rate of unemployment in Canada during the sample period (the results from this test are summarized in Table 6). On the other hand, results support deterministic convergence for Alberta, BC, and NS. Finally, the unemployment rate in NF was approaching the national rate before the break in 1983:3 so there is evidence of β-convergence in this province before the break date. However, in the period after the break, we see divergence in NF.

For Manitoba, NB, PEI, and Saskatchewan, this test cannot reject the null hypothesis of stable trend function; therefore, we cannot use these results to discuss the divergence or convergence in these provinces.

Up to this point, we used the methodologies that allow for only one break in the trend function of the unemployment rate disparities of the Canadian provinces. However, it is possible to have more than one break in these series; therefore, we use a methodology proposed by Bai and Perron (1998, 2003) to deal with deterministic convergence in presence of multiple breaks in the series.

The results of following this procedure for the disparities in the unemployment rate of the Canadian provinces are reported in Table 8. This method allows for breaks in μ and β in trend function. In the estimation process we use 15% trimming from both ends of the sample (17 observations) and we also impose the maximum number of breaks equal to 5. According to this table UD_{MAX}, WD_{MAX} , and $SupF(0|l)$ suggest the existence of at least one break in each series. The exact number of breaks can be determined using information criteria, such as BIC (Schwarz) and LWZ (Liu, *et. al.* (1997)), using $SupF(l+1|l)$, and the outputs from the sequential procedure. The number of breaks for each series is reported in the bottom three rows of Table 8. Based on the sequential procedure at 5% level of significance, there are 2 or 3 breaks in each series except for Manitoba with no break and Quebec with 4 breaks. However, at 10% level of significance, the results of this procedure show that there are 3 break points in the series for Manitoba, that is, there are 4 different regimes in the time series of disparities in Manitoba. For all series, we use the results from the sequential procedure at 5% level of significance, the exception being Manitoba, for which we use the result at 10% level.

The estimated break dates and associated 90% confidence interval along with the estimated coefficients and $t-$statistics are reported in Table 9. The estimated break dates are very precise (at 90% confidence interval, standard errors of confidence bands are less than 1.5 years). The point estimates of the first break dates for most provinces are around 1982-1984 and the second and the third break dates are around 1989-1992 and 1996-1998, respectively. The last break date for Quebec is in 2001.

The results from Table 9 are summarized in Table 10. In this table we use the same notation as Table 6 to investigate the deterministic convergence or divergence in the series. We also call the period between two breaks, a regime. According to this table, there is more evidence of β-convergence in the Canadian provinces.

In all regimes, we see that BC[12], Manitoba, NB, and NS have experienced β-convergence. However, the other provinces had experienced convergence in some period and divergence in others. Except for the second regime, Alberta shows divergence and NF shows β-convergence. In Quebec and PEI, we observe that the rate of unemployment of these provinces are converging to the Canadian rate of unemployment after the first break in 1982:3 and 1982:2, respectively. Before the second break in 1990:2, Ontario was diverging; however, there is clear convergence after this break date. Finally, convergence in Saskatchewan changes to divergence after the second break in 1990:3.

2.4 Conclusion

This paper has examined the stochastic and deterministic convergence of the unemployment rates in the Canadian provinces for the period of 1976-2005, by using recently proposed econometric methodologies. Using the ADF and ADF^{GLS} unit root tests, we found evidence of the stochastic convergence for most, but not all of the provinces. Therefore, we used some new tests of unit root, which allow for the break in the data. By applying these tests and allowing for one and two structural breaks in the data, we could reject the null of unit root and conclude that the stochas-

[12] The results for BC should be approached cautiously, since basd on the results from section 3.1 it has unit root and the approach of Bai and Perron is not robust to the presence of unit root.

tic convergence exists for all the provinces except British Columbia. Next, we test for deterministic convergence in the data, using tests proposed by Vogelsang (1998), Perron and Yabu (2005, 2006), and Bai and Perron (1998, 2003). The results of these tests show that the by introducing multiple structural breaks in the data, we get more evidence of deterministic convergence in the Canadian provinces.

Overall, combining the results from the study of stochastic convergence in section 3.1 with the results that we obtained using methodologies of Vogelsang, Perron and Yabu, and Bai-Perron, there is strong evidence that the unemployment rates of these provinces are converging to the national level of the unemployment rate. This conclusion is stronger when we allow for more than one break in the trend function.

References

[1] Andrews, D. W. K., (1993), "Exactly Median-unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica*, **61**, 139-65.

[2] Andrews, D. K. W. and Ploberger, W. (1994), "Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative," *Econometrica*, **62**, 1383–1414.

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

[4] Bai, J., and Perron, P., (2003), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, **18**, 1-22.

[5] Barro, Robert J., (1991), "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics*, **106**, 407-43.

[6] Barro, R. and Sala-i-Martin, X., (1992), "Convergence," *Journal of Political Economy*, **100**, 223-51.

[7] Bernard, A. B. and Durlauf, S. D. (1995), "Convergence in International Output," *Journal of Applied Econometrics*, **10**, 97–108.

[8] Carlino, C. A. and Mills, L. O., (1993), "Are US regional incomes converging?" *Journal of Monetary Economics*, **32**, 335-46.

[9] Costantini, Mauro and Lupi,C., (2006), "Divergence and long-run equilibria in Italian regional unemployment," *Applied Economics Letters*, **13**, 899-904.

[10] Coulombe, Serge and Lee F. C.,(1995), "Convergence across Canadian Provinces, 1961 to 1991," *The Canadian Journal of Economics*, **28(4a)**, 886-98.

[11] Coulombe, Serge and Day, K. M., (1999), "Economic Growth and Regional Income Disparities in Canada and the Northern United States." *Canadian Public Policy*, **25(2)**, 155-78.

[12] DeJuan, J. and Tomljanovich, M.,(2005), "Income convergence across Canadian provinces in the 20th century: Almost but not quite there," *The Annals of Regional Science*, **39**, 567-92.

[13] Dickey, D. and Fuller, W. (1979), "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association,* **74**, 427-431.

[14] Elliot, G., Rothenberg, T. and Stock, J. H. (1996), "Efficient tests for an autoregressive unit root," *Econometrica*, **64**, 813-39.

[15] Friedman, M. (1992), "Do Old Fallacies Ever Die?" *Journal of Economic Literature*, **30**, 2129–2132.

[16] Gunderson, morley (1996),"Regional Productivity and Income Convergence in Canada Under Increasing Economic Integration," *Canadian Journal of Regional Science*, **19(1)**, 1-23.

[17] Liu, J., Wu, S., and Zidek, JV. (1997), "On segmented multivariate regressions," *Statistica Sinica*, **7**, 497-525.

[18] Lumsdaine, R. and Papell, D., (1997), "Multiple trend breaks and the unit root hypothesis," *Review of Economics and Statistics,* **79**, 212–8.

[19] McGee, R. H., (1992), *Getting It Right: Regional Development in Canada*, Montreal & Kingston: McGill-Queen's University Press.

[20] Myatt, Anthony,(1992), "Provincial unemployment Rate Disparities: A Case of No Concern?" *Canadian Journal of Regional Science*, **25(1)**, 101-119.

[21] Ng, S. and Perron, P., (2001), "Lag length selection and the construction of unit root tests with good size and power," *Econometrica*, **69**, 1519-54.

[22] Park, J.Y., and Choi, B. (1988), "A new approach to testing for a unit root." CAE Working Paper No. 88-23, Cornell University, Ithaca, NY.

[23] Park, J.Y. (1990), "Testing for unit root and cointegration by variable addition," in Fomby, bt. T. and Rhodes, F., *Advances in Econometrics: Cointegration, Spurious Regression and Unit Roots*, JAI Press, London, 107-34.

[24] Perron, P. (1989), "The great crash, the oil price shock and the unit root hypothesis," *Econometrica* **57**, 1361–1401.

[25] Perron, P., (1997), "Further evidence on breaking trend functions in macroeconomic variables," *Journal of Econometrics,* **80**, 355–85

[26] Perron, P. and Rodríguez, G. (2003), "GLS detrending, efficient unit root tests and structural change," *Journal of Econometrics*, **115**, 1-27.

[27] Perron, P. and Yabu, T. (2005), "Testing for Shifts in Trend with an Integrated or Stationary Noise Components." Working Paper,Department of Economics, Boston University.

[28] Perron, P. and Yabu, T. (2006), "Estimating Deterministic Trends with an Integrated or Stationary Noise Components." Working Paper, Department of Economics, Boston University.

[29] Quah, D. (1993), "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics*, **95**, 427–443.

[30] Quah, Danny, T. (1996), "Empirics for economic growth and convergence," *European Economic Review*, **40**, 1353-75.

[31] Rodríguez-Oreggia, Eduardo,(2005), "Regional disparities and determinants of growth in Mexico," *The Annals of Regional Science*, **39**, 207-220.

[32] Rodríguez. G. (2006), "The role of the interprovincial transfers in the β-convergence process: further empirical evidence for Canada." *Journal of Economic Studies,* **33(1)**,12-29.

[33] Rodríguez. G. (2007), "Finite Sample Bahavior of the Level Shift Model Using Quasi-Difference Data", forthcoming in the Journal of Statistical Computation and Simulation.

[34] Roy, A., and W.A. Fuller (2001), "Estimation for Autoregressive Processes With a Root Near One," *Journal of Business and Economic Statistic*, **19**, 482-93.

[35] Said E. and David A. Dickey (1984), "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order," *Biometrika*, **71**, 599–607.

[36] Sargan JD., (1964), "Wages and prices in the United Kingdom: a study in econometric methodology," in *Econometric Analysis for National Economic Planning,* Hart PE, Mills G., Whitaker JK. (Eds), Butterworths, London, 25-54.

[37] Solow, Robert M., (1965), "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, **70**, 65-94.

[38] Stock, J. H. (1999), "A class of tests for integration and cointegration," in Engle, R.F. and White, H. (Eds), *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, Oxford University Press, Oxford, 137-67.

[39] Vogelsang, T.J. (1998), "Trend function hypothesis testing in the presence of serial correlation." *Econometrica*, **66**, 123-48.

[40] Zivot, E. and Andrews, D. W. K., (1992), "Further evidence of the great crash, the oil-price shock and the unit-root hypothesis," *Journal of Business and Economic Statistics,* **10**, 251–70.

Tables

Table1: Results from unit root tests without break

	ADF			ADF^{GLS}	
Province	no regressors	with c	with c & t	with c	with c & t
Alberta	-0.937	-1.707	-1.661	-1.547	-1.601
British Columbia	-1.535	-1.449	-1.558	-1.223	-1.585
Manitoba	-1.048	-3.057^{c}	-3.050	-3.050^{a}	-3.061^{c}
New Brunswick	-1.051	-3.057^{c}	-3.370^{d}	-2.828^{a}	-3.389^{c}
Newfoundland	0.008	-2.287	-3.326^{d}	-0.972	-3.038^{c}
Nova Scotia	-1.081	-4.337^{a}	-4.335^{a}	-2.077^{c}	-2.937^{d}
Ontario	-0.876	-1.217	-1.741	-1.232	-1.352
PEI	-0.842	-2.472	-2.840	-2.117^{c}	-2.873^{d}
Quebec	-0.725	-3.295^{c}	-4.016^{c}	-2.334^{c}	-2.763^{d}
Saskatchewan	-1.102	-1.952	-2.263	-1.806^{d}	-2.254

a, b, c and d show significance at 1%, 2.5%, 5%, and 10% level of significance.

c and t denote the intercept and the time trend.

Table 2: Results of unit root tests with break

Province	Statistic	$t_{\hat{\alpha}}$	k	T_{B1}	T_{B2}
Alberta	MSB	0.174	1	1983:2	
	P_T	17.081	1	1983:2	
	LP	-6.357^{c}	0	1982:4	1989:3
British Columbia	MSB	0.273	1	1989:4	
	P_T	41.069	1	1989:4	
	LP	-4.486	0	1981:3	1997:1
Manitoba	MSB	0.169	1	1987:1	
	P_T	15.115	1	1987:1	
	LP	-5.999^{d}	0	1987:2	1990:2
New Brunswick	MSB	0.125^{b}	1	1991:4	
	P_T	7.988^{a}	1	1991:4	
	LP	-5.578	7	1991:4	1996:2
Newfoundland	MSB	0.110^{a}	3	1983:2	
	P_T	6.543^{a}	3	1983:2	
	LP	-5.959^{d}	3	1979:2	1983:4
Nova Scotia	MSB	0.125^{b}	1	1997:4	
	P_T	8.852^{b}	1	1985:2	
	LP	-7.135^{c}	0	1984:2	1992:4
Ontario	MSB	0.225	1	1989:3	
	P_T	27.312	1	1984:1	
	LP	-6.786^{c}	7	1983:2	1990:2
PEI	MSB	0.138^{d}	1	1989:2	
	P_T	10.211^{c}	1	1987:2	
	LP	-5.219	0	1984:3	1988:4
Quebec	MSB	0.150	1	1977:1	
	P_T	11.419^{d}	1	1977:1	
	LP	-4.689	7	1983:4	1999:3
Saskatechwan	MSB	0.196	1	1993:1	
	P_T	19.724	1	1993:1	
	LP	-4.508	8	1985:1	1990:3

$a, b, c,$ and d show significance at 1%, 2.5%, 5%, and 10% level of significance.

Table 3: Emprical results using the z_t regression and $t - PS_T$ statistics without J_T correction

Province	$\hat{\mu}_1$ $(t\text{-}stat)$	$\hat{\beta}_1$ $(t\text{-}stat)$	$\hat{\mu}_2$ $(t\text{-}stat)$	$\hat{\beta}_2$ $(t\text{-}stat)$	$\hat{T}_B$
Alberta	-3.319^d	-1.557	0.258	-4.054^c	1982:4
	(-1.956)	(-0.120)	(0.437)	(-3.176)	
British Columbia	-0.806	8.869^c	-1.495^d	3.891^d	1989:3
	(-0.967)	(2.523)	(-1.295)	(1.062)	
Manitoba	-2.208^c	0.267	-0.840^d	-2.837^c	1987:1
	(-3.427)	(0.083)	(-1.484)	(-1.852)	
New Brunswick	3.777^c	0.508	1.976^c	2.491	1990:3
	(6.902)	(0.234)	(2.186)	(0.810)	
Newfoundland	7.486^c	-6.405	8.526^c	0.929	1983:3
	(6.295)	(-0.771)	(17.064)	(0.831)	
Nova Scotia	2.439^c	-2.151	2.898^c	-0.824	1984:3
	(3.004)	(-0.422)	(6.751)	(-0.814)	
Ontario	-0.470^d	-4.450^c	-0.606^d	-0.486	1990:2
	(-1.575)	(-3.704)	(-1.284)	(-0.308)	
PEI	1.930^c	2.132	7.136^c	-4.595^c	1989:1
	(2.572)	(0.652)	(7.501)	(-1.575)	
Quebec	1.847^d	4.029	1.927^c	-0.074	1982:3
	(1.578)	(0.436)	(5.045)	(-0.091)	
Saskatchewan	-4.434^c	5.722^c	-3.193^c	2.763	1990:3
	(-5.705)	(1.856)	(-2.488)	(0.633)	
10% critical value	$\pm$1.570	$\pm$1.330	$\pm$1.140	$\pm$0.936	
5% critical value	$\pm$2.190	$\pm$1.760	$\pm$1.500	$\pm$1.270	

$a, b, c,$ and d show significance at 1%, 2.5%, 5%, and 10% level of significance.

Table 4: Emprical results using the z_t regression and $t - PS_T$ statistics with J_T correction

Province	$\hat{\mu}_1$ (5% t-stat) (10% t-stat)	$\hat{\beta}_1$ (5% t-stat) (10% t-stat)	$\hat{\mu}_2$ (5% t-stat) (10% t-stat)	$\hat{\beta}_2$ (5% t-stat) (10% t-stat)	$\hat{T}_B$
Alberta	-3.319	-1.557	0.258	-4.054^{d}	1982:4
	(-1.212)	(-0.028)	(0.064)	(-0.782)	
	(-1.302)	(-0.043)	(0.106)	(-1.158)	
British Columbia	-0.806	8.869	-1.495	3.891	1989:3
	(-0.073)	(0.001)	(0.000)	(0.001)	
	(-0.108)	(0.01)	(-0.001)	(0.005)	
Manitoba	-2.208^{d}	0.267	-0.84	-2.837	1987:1
	(-2.016)	(0.016)	(-0.177)	(-0.392)	
	(-2.182)	(0.026)	(-0.309)	(-0.606)	
New Brunswick	3.777^{c}	0.508	1.976	2.491	1990:3
	(4.848)	(0.080)	(0.531)	(0.288)	
	(5.111)	(0.109)	(0.769)	(0.385)	
Newfoundland	7.486^{c}	-6.405	8.526^{c}	0.929	1983:3
	(4.881)	(-0.354)	(6.162)	(0.395)	
	(5.070)	(-0.446)	(8.039)	(0.486)	
Nova Scotia	2.439^{c}	-2.151	2.898^{c}	-0.824	1984:3
	(2.392)	(-0.210)	(2.711)	(-0.418)	
	(2.475)	(-0.258)	(3.440)	(-0.504)	
Ontario	-0.470	-4.450	-0.606	-0.486	1990:2
	(-0.294)	(-0.022)	(-0.002)	(-0.002)	
	(-0.378)	(-0.100)	(-0.009)	(-0.009)	
PEI	1.930^{d}	2.132	7.136^{d}	-4.595	1989:1
	(1.499)	(0.125)	(0.864)	(-0.324)	
	(1.625)	(0.204)	(1.519)	(-0.505)	
Quebec	1.847	4.029	1.927	-0.074	1982:3
	(0.802)	(0.055)	(0.336)	(-0.013)	
	(0.888)	(0.101)	(0.681)	(-0.022)	
Saskatchewan	-4.434^{d}	5.722	-3.193	2.763	1990:3
	(-1.831)	(0.058)	(-0.026)	(0.023)	
	(-2.170)	(0.160)	(-0.086)	(0.058)	
10% critical value	$\pm$1.570	$\pm$1.330	$\pm$1.140	$\pm$0.936	
5% critical value	$\pm$ 2.190	$\pm$1.760	$\pm$1.500	$\pm$1.270	

$a, b, c,$ and d show significance at 1%, 2.5%, 5%, and 10% level of significance.

Table 5: Emprical results using the y_t regression and $T^{-1/2}t_y$ statistics

Province	$\hat{\mu}_1$ $(t-stat)$	$\hat{\beta}_1$ $(t-stat)$	$\hat{\mu}_2$ $(t-stat)$	$\hat{\beta}_2$ $(t-stat)$	$\hat{T}_B$
Alberta	-3.392^c	-0.722	0.044	-3.457	1982:4
	(-1.223)	(-0.043)	(0.029)	(-1.213)	
British Columbia	-0.389	6.882	-0.797	1.826	1989:3
	(-0.130)	(0.738)	(-0.287)	(0.246)	
Manitoba	-2.149^c	-0.278	-0.838	-2.495	1987:1
	(-1.189)	(-0.041)	(-0.598)	(-0.769)	
New Brunswick	3.755^c	0.700	2.040	1.883	1990:3
	(1.629)	(0.105)	(0.893)	(0.289)	
Newfoundland	7.471^c	-6.713	8.775^c	0.254	1983:3
	(2.093)	(-0.345)	(4.208)	(0.063)	
Nova Scotia	2.452^c	-2.342	3.066	-1.351	1984:3
	(1.075)	(-0.212)	(2.110)	(-0.455)	
Ontario	-0.570	-4.022	-0.828	0.350	1990:2
	(-0.514)	(-1.229)	(-0.766)	(0.115)	
PEI	1.666	3.516	6.917^d	-4.359	1989:1
	(0.515)	(0.337)	(2.392)	(-0.581)	
Quebec	1.725^c	5.249	1.995	-0.383	1982:3
	(0.985)	(0.480)	(2.145)	(-0.220)	
Saskatchewan	-4.510^c	6.191	-3.342	3.009	1990:3
	(-2.063)	(0.977)	(-1.542)	(0.487)	
10% critical value	$\pm$0.671	$\pm$1.470	$\pm$2.370	$\pm$1.480	
5% critical value	$\pm$0.875	$\pm$2.000	$\pm$3.000	$\pm$2.010	

$a, b, c,$ and d show significance at 1%, 2.5%, 5%, and 10% level of significance.

Table 6: Summary of emprical results from Volgelsang and Perron and Yabu methodologies

	Vogelsang						Perron & Yabu	
	$t - PS_T$ $I(0)$ Errors Assumed		$t - PS_T$ Robust to $I(1)$ Errors		$T^{-1/2}t_y$ Robust to $I(1)$ Errors		W_{RQF} Robust to $I(1)$ Errors	
Province	Pre-break	Post-break	Pre-break	Post-break	Pre-break	Post-break	Pre-break	Post-break
Alberta	d	c	u	c	d	u	c	C
British Columbia	c	C	u	E	u	u	c	C
Manitoba	c	D	c	u	d	u	d	C
NB	d	d	d	u	d	u	d	c
NF	c	d	c	d	c	d	C	D
Nova Scotia	c	c	c	c	c	u	C	c
Ontario	D	d	u	E	u	u	D	D
PEI	d	C	d	c	u	c	D	C
Quebec	d	c	u	u	d	u	D	D
Saskatchewan	C	c	c	u	c	u	C	C

Table 7: Results from applying Perron and Yabu

			Intercept		Slope	
Province	W_{RQF}	$\hat{T}_B$	Pre-break	Post-break	Pre-break	Post-break
Alberta	11.778^{a}	1982:4	-3.392^{c}	3.638^{c}	-0.007	-0.035^{c}
British Columbia	3.752^{b}	1989:3	-0.389	-4.193^{c}	0.069^{c}	0.018^{c}
Manitoba	0.769	1987:1	-2.149^{c}	1.437^{c}	-0.003	-0.025^{c}
New Brunswick	0.923	1990:3	3.755^{c}	-2.127^{c}	0.007	0.019
Newfoundland	2.799^{c}	1983:3	7.471^{c}	3.386^{c}	-0.067^{c}	0.003^{c}
Nova Scotia	2.537^{c}	1984:3	2.452^{c}	1.435^{c}	-0.023^{c}	-0.014
Ontario	3.614^{b}	1990:2	-0.571^{c}	2.075^{c}	-0.040^{c}	0.004^{c}
PEI	0.972	1989:1	1.667^{c}	3.387^{c}	0.035^{c}	-0.044^{c}
Quebec	3.965^{b}	1982:3	1.725^{c}	-1.148^{c}	0.053^{c}	-0.004^{c}
Saskatchewan	1.196	1990:3	-4.511^{c}	-2.485^{c}	0.062^{c}	0.030^{c}

a, b, c, and d show significance at 1%, 2.5%, 5%, and 10% level of significance.

Table 8: Results from applying Bai and Perron approach

Statistics	Alberta	BC	Manitoba	NB	NF	NS	Ontario	PEI	Quebec	Saskatchewan
UD_{MAX}	248.248^{a}	474.371^{a}	70.803^{a}	75.874^{a}	60.010^{a}	76.882^{a}	291.746^{a}	175.221^{a}	47.552^{a}	110.638^{a}
WD_{MAX} 5%	396.022*	930.092*	102.582*	121.039*	88.647*	105.483*	572.021*	206.132*	90.680*	132.228*
sup $F_T(1)$	69.456^{a}	442.603^{a}	10.454^{d}	19.115^{a}	55.658^{a}	16.829^{a}	84.442^{a}	134.016^{a}	36.540^{a}	16.947^{a}
sup $F_T(2)$	241.268^{a}	64.336^{a}	25.538^{a}	35.846^{a}	50.541^{a}	38.732^{a}	103.881^{a}	175.221^{a}	27.765^{a}	110.638^{a}
sup $F_T(3)$	112.647^{a}	255.582^{a}	70.803^{a}	39.906^{a}	63.010^{a}	76.882^{a}	99.120^{a}	83.132^{a}	31.758^{a}	96.375^{a}
sup $F_T(4)$	248.248^{a}	240.970^{a}	64.304^{a}	75.874^{a}	55.468^{a}	63.020^{a}	207.213^{a}	65.726^{a}	47.552^{a}	67.788^{a}
sup $F_T(5)$	200.624^{a}	474.371^{a}	48.805^{a}	59.799^{a}	45.212^{a}	44.803^{a}	291.746^{a}	67.467^{a}	46.249^{a}	56.284^{a}
sup $F_T(2\|1)$	24.862^{a}	41.604^{a}	35.450^{a}	34.389^{a}	23.225^{a}	50.859^{a}	39.805^{a}	13.547^{c}	18.549^{a}	93.789^{a}
sup $F_T(3\|2)$	1.822^{d}	80.208^{a}	18.474^{a}	18.566^{a}	8.505	3.623	69.349^{a}	4.57	26.431^{a}	43.756^{a}
sup $F_T(4\|3)$	16.808^{b}	82.495^{a}	24.572^{a}	6.816	16.776^{b}	2.658			26.431^{a}	6.506
sup $F_T(5\|4)$		32.167^{a}		1.453						
BIC	5	5	4	3	3	3	4	3	4	4
LWZ	4	4	3	3	1	3	4	2	3	3
Sequential	$2^{a,b,c,d}$	$3^{a,b,c,d}$	$0^{a,b,c}\ 3^{d}$	$3^{a,b,c,d}$	$2^{a,b,c,d}$	$2^{a,b,c,d}$	$3^{a,b,c,d}$	$2^{a,b,c}\ 3^{d}$	$4^{a,b,c,d}$	$3^{a,b,c,d}$

^{a,b,c,}and ^{d} show significance at 1%, 2.5%, 5%, and 10% level of significance.

* significant at 5% level of significance.

Table 9: Estimation results from Bai-Perron method

	Alberta (*t−stat*)	BC (*t−stat*)	Manitoba (*t−stat*)	NB (*t−stat*)	NF (*t−stat*)	NS (*t−stat*)	Ontario (*t−stat*)	PEI (*t−stat*)	Quebec (*t−stat*)	Saskatchewan (*t−stat*)
μ_1	-3.392	0.437	-2.433	4.616	7.471	2.452	-0.992	1.519	1.725	-3.513
	(-16.793)	(1.691)	(-13.714)	(21.068)	(23.467)	(13.427)	(-10.464)	(4.394)	(15.131)	(-18.306)
β_1	-0.007	-0.012	0.029	-0.055	-0.067	-0.023	-0.002	0.080	0.052	0.012
	(-0.594)	(-0.803)	(2.445)	(-5.046)	(-3.865)	(-2.648)	(-0.4)	(3.577)	(7.376)	(0.89)
μ_2	-1.896	5.634	-4.043	4.891	8.478	4.911	-2.254	-4.898	3.478	-8.551
	(-3.453)	(6.744)	(-6.738)	(5.937)	(20.247)	(9.614)	(-8.417)	(-5.681)	(9.209)	(-25.971)
β_2	0.044	-0.059	0.046	-0.011	0.006	-0.045	-0.006	0.188	-0.046	0.152
	(3.39)	(-3.105)	(2.754)	(-0.616)	(0.845)	(-4.624)	(-1.093)	(8.897)	(-4.789)	(19.945)
μ_3	-0.475	2.597	2.349	3.518	15.822	5.841	-1.471	9.228	2.170	-2.731
	(-1.511)	(2.616)	(7.046)	(2.658)	(7.825)	(12.209)	(-3.725)	(18.832)	(6.179)	(-3.954)
β_3	-0.019	-0.048	-0.057	-0.022	-0.065	-0.036	0.011	-0.044	-0.003	-0.004
	(-5.507)	(-3.435)	(-11.931)	(-1.196)	(-3.37)	(-7.216)	(2.04)	(-7.88)	(-0.504)	(-0.45)
μ_4		2.455	-1.509	6.197			-4.386		8.103	-0.723
		(1.88)	(-1.416)	(6.061)			(-8.27)		(7.828)	(-0.679)
β_4		-0.020	-0.006	-0.031			0.035		-0.065	-0.010
		(-1.623)	(-0.585)	(-3.081)			(6.849)		(-5.839)	(-0.984)
μ_5									-0.142	
									(-0.09)	
β_5									0.013	
									(0.932)	
					Estimate Break Dates and their 90% confidence interval					
T_{B1}	1982:4	1983:2	1982:1	1984:2	1983:3	1984:3	1983:2	1982:2	1982:3	1981:4
$C.I.$	1982:2-1983:1	1983:1-1983:2	1981:3-1982:2	1983:4-1985:2	1983:1-1984:1	1983:2-1984:4	1982:4-1983:4	1981:4-1982:3	1982:1-1983:4	1981:1-1982:1
T_{B2}	1989:3	1989:3	1987:1	1990:3	1998:2	1992:4	1990:2	1989:1	1988:1	1990:3
$C.I.$	1989:1-1991:1	1988:2-1990:1	1986:2-1987:2	1989:4-1990:4	1997:1-1998:4	1992:2-1993:3	1989:4-1990:3	1988:3-1989:2	1987:3-1989:2	1990:1-1990:4
T_{B3}		1997:2	1998:3	1996:3			1997:4		1996:2	1998:2
$C.I.$		1996:4-1998:3	1997:4-2003:3	1995:4-1997:1			1997:1-1998:3		1995:3-1996:3	1997:4-1998:4
T_{B4}									2001:2	
$C.I.$									2000:3-2001:3	
$\bar{R}^2$	0.846	0.801	0.648	0.666	0.644	0.449	0.900	0.800	0.709	0.839
F	109.038^a	60.349^a	28.015^a	30.270^a	36.414^a	16.855^a	134.361^a	79.625^a	29.693^a	77.823^a

^{a,b,c,}and ^{d} show significance at 1%, 2.5%, 5%, and 10% level of significance.

Table 10: Summary of Bai-Perron

Province	Regime 1	Regime 2	Regime 3	Regime 4	Regime 5
Alberta	d	C	d		
BC	c	C	C	C	
Manitoba	C	C	C	u	
NB	C	c	c	C	
NF	C	d	C		
NS	C	C	C		
Ontario	d	d	C	C	
PEI	D	C	C		
Quebec	D	C	C	C	u
Saskatechwan	c	C	d	u	

Figures

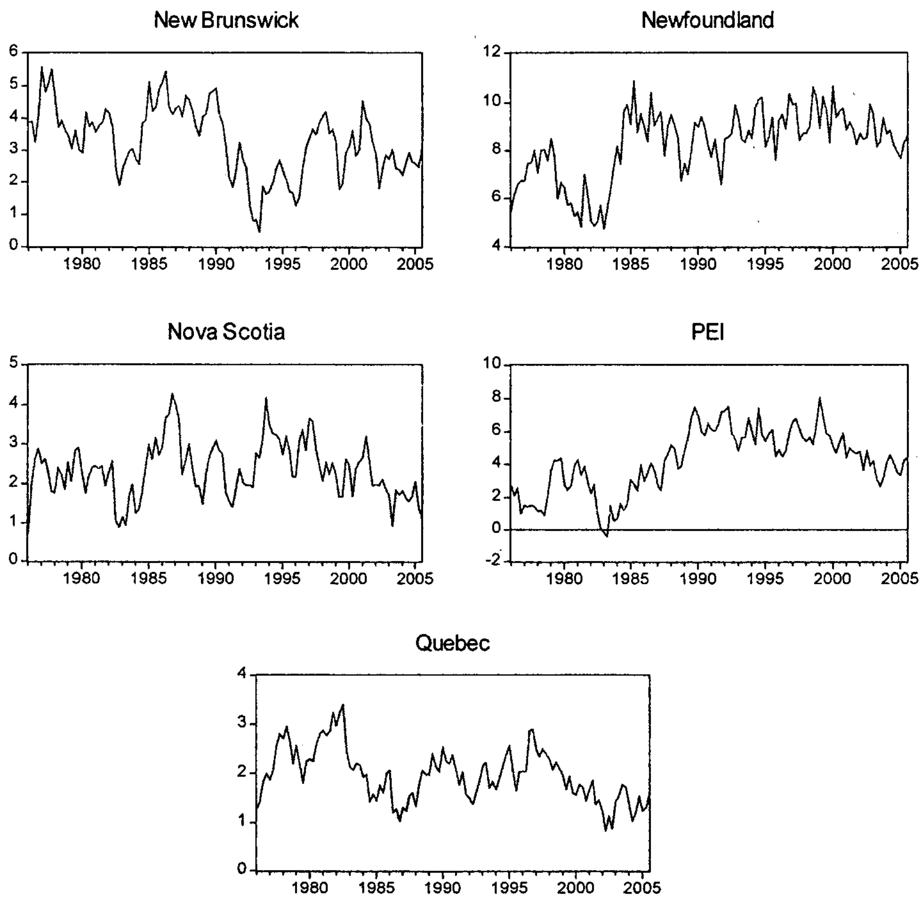

Figure 1: Deviation of the unemployment rates of the Canadian provinces from the unemployment rate of Canada.

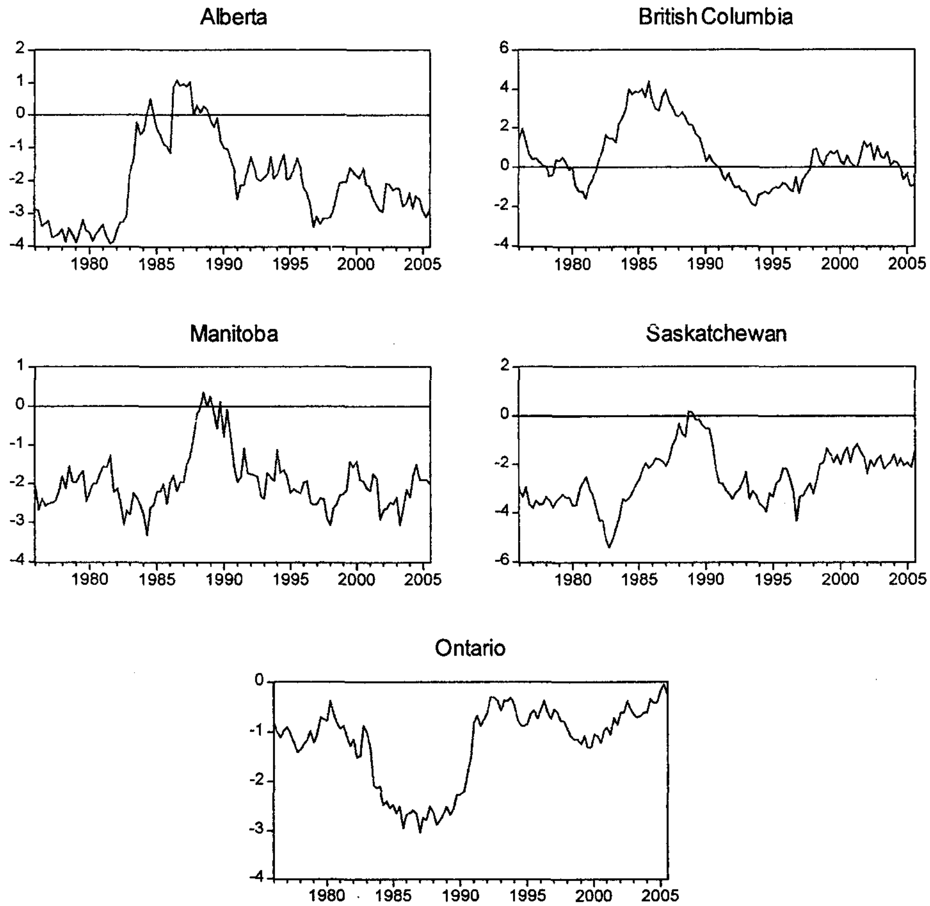

Figure 2: Deviation of the unemployment rates of the Canadian provinces from the unemployment rate of Canada.

Chapter 3
Persistence in Unemployment: Evidence from Canada and the Canadian Provinces

3.1 Introduction

One of the main theories of macroeconomics is the natural rate of unemployment theory, by Friedman (1968) and Phelps (1967). According to this theory, the unemployment rate is stationary around the natural rate of unemployment and any shock to the unemployment rate will have a temporary effect and the unemployment rate will return to its natural rate. However, the behavior of the unemployment rates in 1970s and 1980s, i.e., persistent and high for a long period of time, has lead to questioning the validity of this theory. New theories were needed to describe this phenomenon. Blanchard and Summers (1987) proposed an alternative theory called the hysteresis theory of unemployment rate, which states that shocks to the unemployment rate will have a permanent effect on the unemployment rate. They define hysteresis as, "a very high dependence of current unemployment rate on past unemployment rates". The hysteresis effect can be explained by the fact that individuals who have been unemployed for long periods, are likely to remain unemployed, since their skills eventually decline. This means that the past unemployment affects the current and future unemployment rates.

It is very important for policy makers to know whether the unemployment rate is persistent. When evidence of persistence is found, any interference in the labor market that changes the unemployment rate will affect it for an extended period. Conversely, when the unemployment rate is not persistent, the effect of this interference will disappear in the short run and unemployment will return to its original starting point. Therefore, the interference is not justified.

In this paper, we investigate the degree of persistence in the unemployment rates of Canada and the ten Canadian provinces. In fact, we try to answer the following questions: Are the unemployment rates in Canada and the Canadian provinces persistent? Is the persistence constant or has it changed over time? We are also interested to see which one of the above mentioned theories describes the movements in these unemployment rates, the traditional theory of unemployment rate or the hysteresis theory.

Unit root tests provide one way to assess persistence. Using these tests, the researcher checks for the integration order of the variable. When the variable is integrated of order one, $I(1)$, the variable has unit root and is persistent. On the other hand, when the variable is $I(0)$, it is stationary and not persistent. However, most unit root tests have low power to reject the null hypothesis of unit root, when the root is close to unity. In other words, it is hard to distinguish a unit root from a close to unit root case.

Another tool that we can use to examine the persistence of variable is building the confidence interval for the autoregressive coefficient. This interval can be constructed for the largest autoregressive root, or for the sum of the autoregressive roots. In both cases, when the interval contains unity, we can conclude that the variable is difference stationary, $I(1)$, and when it does not include unity, it is trend stationary, $I(0)$. Being a difference stationary, implies that the variable has unit root and is persistent. In the literature, several methods are available to construct the confidence intervals: the asymptotic method, local to unity of Stock (1991), percentile-t of Hall (1992), percentile bootstrap of Efron and Tibshirani (1993), grid bootstrap of Hansen (1999), amongst many.

Finally, we want to check the stability of the persistence in the unemployment rate series to see whether it has remained constant or has changed over time. To this end, we use a method proposed by Bai and Perron (1998, 2003), which allows multiple breaks in the data. The number of break points and their locations are determined endogenously in the model. Using this approach, we are able to test the stability of the persistence in the data.

The behavior of unemployment rates has been studied by many researchers in numerous countries. Blanchard and Summer (1986), estimate an AR(1) model for the unemployment rates in the U.K. and the United States. Based on the autoregressive coefficient, they concluded that the unemployment rates in these two countries are persistent and that they have a weak tendency to return to their mean. Fortin (1989, 1991) checked for the existence of hysteresis in the unemployment rate in

Canada from 1957-1999. By estimating a Philips curve, which included expected inflation, he concluded that a negative hysteresis was present for data from 1957-1972. During 1973-1990, he found evidence of positive hysteresis in the data. Milbourne, Purvis, and Scoones (1991) built a labor market model to investigate the effect of the Canadian unemployment insurance system on unemployment. Using this model and the monthly unemployment data for Canada from 1970-1988, they showed that features of the Canadian unemployment insurance system makes the natural rate of unemployment dependent on its past values, which means persistence in the data.

Jaeger and Parkinson (1994), used the unemployment rate data from Canada, Germany, the U.K., and the U.S. to test for hysteresis. They used augmented Dickey-Fuller (ADF) unit root tests and unobserved component model and concluded that hysteresis effects existed in the data. Jones (1995) investigated hysteresis in the unemployment rate of Canada and concluded that there is no firm evidence of hysteresis, though he did not rule out the possibility of hysteresis. Roed (1996), examined the unemployment rates in 16 OECD countries, including Canada. The results from ADF unit root tests and the Kwiatkowsky, Phillips, Schmidt, and Shin (1992) test for stationarity (KPSS), confirmed the hysteresis effect in the unemployment rates in Canada, Australia, Japan, and several European countries over the period of 1970-1994. Koustas and Veloce (1996) used disaggregated unemployment data by gender and age in Canada and the USA. They applied long memory models to the data and concluded that persistence in Canada is higher than in the USA. In addition, persistence is higher for adult males than adult females. Nott (1996), using a nonlinear Philips curve, did not find the hysteresis effect in Canadian unemployment. In addition, she estimated the linear Philips curve used by Fortin (1991) during 1954-1995. Her results contradicted the findings of Fortin (1991). Hence, she concluded that there was no hysteresis in the Canadian unemployment rate.

Song and Wu (1997), studied hysteresis using annual data for 48 U.S. states from 1962-1993. They reported that although standard unit root tests cannot reject the hypothesis of unit root, they could reject this hypothesis using more powerful panel-based tests. Binachi and Zoega (1998) pointed out that it is very important to take into consideration the occasional shifts in the mean of unemployment. They used Markov Switching models to study persistence in unemployment in OECD countries. They found that when considering the possibility of mean shifting in the modelling process, the results show that there is no evidence of persistence in unemployment rates. Their data for Canada was from 1970-1996, and they determined three different regimes for this data. Keil and Pantuosco (1998) have studied the Canadian and US regional unemployment rates to find an explanation for persistent gap between the unemployment rates in Canada and US, since 1982. In this paper, they also briefly address persistence in the unemployment rates. Using the coefficient of the lagged dependent variable in a panel data model as an indicator of persistence, they find some indications of persistence in Canada. Their results show that this coefficient varies between 0.73 and 0.92, depending on the specification of the model.

Arestis and Mariscal (2000) applied the structural break univariate unit root test of Perron (1997) to the unemployment rates from twenty-two OECD countries. The test allows for one endogenously determined structural break in each unit root test. Although their results are mixed, they mostly reject the unit root null hypothesis and hysteresis. The results imply that the unemployment rate in Canada, during the time span of 1960:1-1997:2, was trend stationary with no break points, which means that the unemployment rate in Canada fluctuated around its mean and there was no sign of hysteresis. Papell, Murray, and Ghiblawi (2000) examined the postwar unemployment rates for sixteen OECD countries. They applied various unit root tests and concluded that when a one-time structural break is allowed, the unit root hypothesis can be rejected for most countries and the degree of persistence falls significantly. Next, they tested for multiple break points and found one or two breaks for most of the data. These break points clustered around 1973-75, 1980-82, and 1988-91. They were able to reject hysteresis for most of countries, while allowing for breaks in data. Gil-Alana (2002) analyzed the persistence in the unemployment in Canada from 1966 to 2000, using ARFIMA models. He determined an ARFIMA(0, 1.37, 0) model as the best specification and concluded that unemployment in Canada was highly persistent during this period. Wu (2003) studied persistence in the regional unemployment rates in China. He used panel unit root tests and found that provincial unemployment is more persistent than the national unemployment. In addition, persistence in the total unemployment rate is higher than persistence in the unemployment rate of the youth.

Camarero and Tamarit (2004) applied multivariate SURE unit root tests to the unemployment rate of nineteen OECD countries, from 1956-2001. Their results indicate that there is no evidence of hysteresis in most of these countries, including Canada. Mikhail, Eberwein and Handa (2005), investigated the persistence in aggregate and sectoral Canadian unemployment from 1976 to 1999. They used "corrected for short-range dependency modified rescaled-range test statistic" to address this issue. The results provided strong evidence of persistence and they concluded that the fluctuations in aggregate and sectoral Canadian unemployment, are persistent. Recently, they reinvestigated persistence in the aggregate Canadian unemployment, using ARFIMA models. Using data from 1976-1998, they found that the persistence in the unemployment rate of Canada holds in the short and medium run. However, in the long run, persistence is uncertain. Gustavsson and Osterholm (2006) used the unit root test of Kapetanios, Shin, and Snell (2003) to address the hysteresis problem in monthly unemployment data from Australia, Canada, Finland, Sweden, and the USA. They concluded that by allowing non-linearity in the tests, much less evidence of hysteresis can be found in the data, with respect to a standard ADF test. They could reject hysteresis in all countries, except Australia. Their time span for Canada was from 1976:1 -2005:1.

Persistence in the national unemployment rate of Canada has been studied by many researchers; however, to the best of our knowledge, this paper is the first to attempt to study the hysteresis effect or persistence in the unemployment rates of the Canadian provinces. We conduct extensive tests in order to check for the existence of persistence and to determine the degree of persistence.

The remainder of this paper is organized as follows. The next section provides the methodological issues used in this paper, including different unit root tests and various methods to construct a confidence interval. We also describe the approach of Bai and Perron (1998, 2003), which allows multiple breaks in the data. Section 3 provides a brief look at the data and the empirical results from various methods. Finally, section 4 concludes.

3.2 Methodology

One of the most important problems in studying the time series variables is the persistence and, as mentioned earlier, the autoregressive root of time series data plays a very important role in the literature about persistence. If the data has a root equal to one, then that variable is persistent and shocks will have a permanent effect. Therefore, we can use the root of the unemployment rates to investigate the hysteresis effect in the unemployment rates. It is possible to use some unit root tests to examine the null hypothesis, that the root is equal to one. If this is the case, we can conclude that there is a hysteresis effect in the unemployment rates. On the other hand, if we can reject the null hypothesis of unit root, we can say that there is no hysteresis effect in the data.

Another tool that we can use to check for hysteresis in the unemployment rates is the confidence interval of the root(s). In this method, we can construct a confidence interval for the largest autoregressive coefficient or a confidence interval for the sum of the autoregressive coefficients. When the confidence interval contains unity, we can conclude that the variable is very persistent and any shock will have a permanent effect.

3.2.1 Unit root Tests

The Augmented Dickey-Fuller (ADF) test

This test has been proposed by Dickey and Fuller (1979), Said and Dickey (1984). The test consists of estimating the following regression:

$$\Delta\tilde{y}_t = \alpha\tilde{y}_{t-1} + \sum_{i=1}^{k}\gamma_i\Delta\tilde{y}_{t-i} + \epsilon_t$$

where $\tilde{y}_t = y_t - \hat{\mu}$ or $\tilde{y}_t = y_t - \hat{\mu} - \hat{\beta}t$, y_t is the time series being tested, t is the time trend, Δ is the first difference operator and k is the number of lags, which must be determined in such a way that the ϵ_t become serially uncorrelated. Using this estimation, we test for a null hypothesis that $\alpha = 0$. If we cannot reject this hypothesis, then we have a unit root and the series is not stationary.

ADF^{GLS} test

To improve the power of the unit root test, ERS (1996) have proposed a new method for unit root testing called ADF^{GLS}. This is an ADF type test, which uses GLS detrended data. The ADF^{GLS} statistic can be constructed using the following equation

$$\Delta\tilde{y}_t = \alpha\tilde{y}_{t-1} + \sum_{i=1}^{k}\gamma_i\Delta\tilde{y}_{t-i} + e_{tk}$$

with $\tilde{y}_t = y_t - \hat{\psi}^{GLS} z_t$ and the set of coefficients $\hat{\psi}^{GLS}$ minimizes

$$S(\psi,\bar{\alpha}) = \sum_{t=1}^{T}[y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}}]^2$$

where $y_t^{\bar{\alpha}} = [y_1, (1-\bar{\alpha}L)\, y_t]$, $z_t^{\bar{\alpha}} = [z_1, (1-\bar{\alpha}L)z_t]$, for $t = 2, ..., T$; and $\bar{\alpha} = 1 + \bar{c}/T$ representing a value under the alternative hypothesis. In addition, we have $z_t = \{1\}$ or $z_t = \{1, t\}$, which represent the deterministic components included in the detrending regression. ERS define a new time series as

$$y_t^{\bar{\alpha}} = \begin{cases} y_t, & \text{if}\, t = 1 \\ y_t - \bar{\alpha}y_{t-1}, & \text{if}\, t > 1, \end{cases}$$

where $\bar{\alpha} = 1 + \bar{c}/T$. ERS recommend the use of $\bar{\alpha} = 1 - 7/T$ for a model with only an intercept and $\bar{\alpha} = 1 - 13.5/T$ for a model with an intercept and a time trend.

At the next step, an ADF test can be carried out using the transformed data as follows

$$\tilde{y}_t = \alpha\tilde{y}_{t-1} + \sum_{i=1}^{k}\gamma_i\tilde{y}_{t-i} + \eta_t.$$

As with the ADF test, we consider the $t-$ratio of $\hat{\alpha}$ from this regression. This $t-$ratio follows a Dickey-Fuller distribution when only an intercept has entered in this equation. However, when we include both an intercept and a time trend, the distribution differs. ERS simulated the critical values for the latter case.

Perron (1989) showed that ADF type tests fail to reject the unit root null hypothesis when the series is stationary with a broken trend. In other words, the failure to reject the null hypothesis of unit root might arise from exogenous changes in the data generating process. Hence, he proposed allowing for one break in the ADF test, assuming that this break point is known and exogenous. However, the problem with this method is that we must know the break date prior to testing for a unit root. To overcome this problem other unit root tests, including Zivot and Andrews (1992), Perron (1997), Lumsdaine and Papell (1997), and Perron and Rodríguez (2003), *inter alia*, have been developed, which allow determination of the break date as an endogenous parameter in the model.

An ADF^{GLS} unit root test with 1 break

Perron and Rodríguez (2003) extended the GLS detrending approach of ERS (1996) and Ng and Perron (2001) to the context of structural change. They considered a break, under the alternative hypothesis, in the slope or a break in both the intercept and slope of the trend function of y_t. The break point is determined endogenously. As in the case of the ADF^{GLS} without structural change, we have $z_t = \{1, t, 1(t \geq T_B) \times (t - T_B)\}$ or $z_t = \{1, 1(t \geq T_B), t, 1(t \geq T_B)(t - T_B)\}$ which represents the deterministic components included in the deterending regression. Using simulations, Perron and Rodríguez (2003) selected $\bar{c} = -22.5$ as the optimal non-centrality parameter. Therefore, this value is used in this paper to carry out this test. To determine the optimal lag we can use information criteria. Perron and Rodríguez (2003) also introduced the structural break in a group of unit root tests called M-tests, which were originally suggested by Stock (1999).[1]

Recently, Rodríguez (2007) constructed a new version of these tests, which only allow a break in the intercept. We also use these tests to check for the existence of unit root in the data under study.

Unit root tests with 2 breaks

The previous test allows for only one break in each time series; however, it is possible to have multiple structural breaks in the series. Therefore, considering only one endogenous break point may not be sufficient and this could cause a loss in information when there is more than one break (Lumsdaine and Papell, 1997). Lumsdaine and Papell (1997) proposed a new procedure that allows one to examine the significance of two structural breaks. They use a modified version of the ADF

[1] Further details about these unit root tests can be found in Perron and Rodríguez (2003).

test with two endogenous breaks as follows:

$$\Delta y_t = \mu + \beta t + \theta DU1_t + \delta DT1_t + \omega DU2_t + \varphi DT2_t + \alpha y_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta y_{t-i} + \epsilon_t$$

where $DU1_t = 1$ if $t > TB1$ and zero otherwise; $DU2_t = 1$ if $t > TB2$ and zero otherwise; $DT1_t = t - TB1$ if $t > TB1$ and zero otherwise; and finally $DT2_t = t - TB2$ if $t > TB2$ and zero otherwise. $DU1_t$ and $DU2_t$ capture the structural breaks in the intercept at time $TB1$ and $TB2$, respectively. The other two dummy variables, $DT1_t$ and $DT2_t$, capture the structural changes in the slope at time $TB1$ and $TB2$, respectively. The optimal lag length can be determined based on the general to specific approach using t tests as suggested by Ng and Perron (1995). The break points are selected based on the minimum value of the t statistic for α.

3.2.2 Building Confidence Intervals

Many researchers have studied the measurement of the persistence of time series variables, such as Stock (1991), Hall (1992), Efron and Tibshirani (1993), Hansen (1999), to name few. To examine the persistence of any variable, it is possible to use unit root tests or construct the confidence interval of the autoregressive coefficient of the variable. However, it is well known that unit root tests may fail to reject the null hypothesis of unit root when the root is close to the unity. Hence, an alternative approach such as confidence interval, might be more informative than the unit root tests. If the confidence interval contains the unity, then the variable is integrated and persistent. When this interval does not include unity, the variable is $I(0)$.

In literature about unemployment rate, researchers usually assume that the unemployment rate (ur) follows an autoregressive process of order k

$$ur_t = \mu + \sum_{i=1}^{k} \rho_i ur_{t-i} + \varepsilon_t, \qquad \varepsilon_t \sim i.i.d.(0, \sigma^2).$$

The largest autoregressive coefficient (ρ) or the sum of the autoregressive coefficients (P) can be used to measure the persistence of the unemployment rate. When $\rho = 1$ or $P = 1$ shocks will have a permanent effect on the unemployment rate, which is known as hysteresis.

There are several ways of building the confidence interval in the literature. Asymptotic confidence interval is the conventional way, in this approach the lower and upper bounds of the 90% confidence interval can be calculated as the point estimate of the root ± 1.645 standard error. However, as noted by researchers, this way of calculating the confidence interval is not appropriate when the root is near to unity and its value is large. In addition, this method is not useful in constructing the confidence interval when the variable is integrated and has a unit root, since the traditional asymptotic theory is discontinuous in this case (Torous *et al.*, 2004, p. 943). At the same time, according to the empirical work on the persistency of the time series variables, it is well known that most macroeconomic variables have a root close or equal to unity. Therefore, the standard asymptotic method of constructing a confidence interval is problematic.

In order to overcome this deficiency, new methods have been developed that are robust to the presence of a root close or on the unit circle. Stock (1991), using the local to unity framework of Bobkoski (1983), Cavanagh (1985), and Philips (1987), developed a method to construct the confidence interval of the largest autoregressive root by inverting ADF $t-$tests. According to this theory, the largest autoregressive root ρ modeled as $\rho = 1 + c/T$, where c is a real valued constant, which shows the deviation from unit root case and T is the sample size. Here, as the sample size increases the autoregressive parameter converges to unity. In this framework, when c is zero, the series has unit root and is integrated. When c is negative, the series is near to integrated and has large but stationary root. On the other hand, when c is positive the series has explosive root.

Since ρ is close to unity, in practice, it will be very difficult to distinguish stationarity of the series from non-stationarity. However, given the distribution of ADF statistic under the local to unity theory, we can get the confidence interval of the root. Suppose that the value of the ADF statistics is $\hat{\tau}(c)$, then a 100(1-α)% confidence interval for this value will be:

$$\{f_{0.5\alpha}(c) \leq \hat{\tau}(c) \leq f_{1-0.5\alpha}(c)\}$$

where $f_{0.5\alpha}(c)$ and $f_{1-0.5\alpha}(c)$ are the 0.5α and $1 - 0.5\alpha$ percentiles of τ as a function of c. Since f is strictly monotone in c, therefore, we can build the confidence interval of c by inverting f. Hence the $100(1-\alpha)\%$ confidence interval of c will be

$\{f_{0.5\alpha}^{-1}(c) \leq c \leq f_{1-0.5\alpha}^{-1}(c)\}$, see Torous *et al.*, 2004, p. 943.Confidence interval of ρ can be build using the confidence interval for c, since $\rho = 1 + c/T$. Originally, Cavanagh (1985) proposed this method to build a confidence interval solely for a first order autoregressive model with no intercept. However, Stock (1991) extended the original work to a general case, a higher order autoregressive model with intercept, or an intercept and a time trend. He showed that the confidence intervals built using this method are asymptotically valid.

Hall (1992) proposed another method to build the confidence interval. This method is called percentile-t. In this method, the estimated value of the autoregressive coefficient, using OLS, is considered as the true value and is used to evaluate the sampling distribution of the t-statistic and build the confidence interval. However, the percentile-t is useful for location statistics[13] such as median, trimmed mean, and sample percentile and the results from applying that to other statistics will not be accurate (Efron and Tibshirani, 1998, p. 162). In addition, recall that these confidence intervals are just approximations. To get a better approximation it is possible to apply bootstrapping techniques. These techniques improve the performance of the conventional asymptotic approximations.[14] Hence, Efron and Tibshirani (1993) propose to use bootstrapping methods to improve the credibility of percentile-t method and its coverage. They propose an improved version of percentile method called bias-corrected and accelerated bootstrap percentile (BC_a). The percentile interval of ξ can be constructed as $\left(\hat{\xi}^{*(\alpha)}, \hat{\xi}^{*(1-\alpha)}\right)$, where, * shows the bootstrap replicates and $\hat{\xi}^{*(\alpha)}$ shows the $100.\alpha th$ percentile of the bootstrap distribution. Therefore, when $\alpha = 0.05$ and the number of replications (B) is 2000, the $\left(\hat{\xi}^{*(0.05)}, \hat{\xi}^{*(0.95)}\right)$ is the interval that contains the $100th$ to the $1900th$ ordered values of the 2000 numbers $\hat{\xi}^{*}(b)$. Now, consider the problem of constructing a confidence interval for ξ using BC_a method. In this case, for any α, the confidence interval will be

$$\left(\hat{\xi}^{*(\alpha_1)}, \hat{\xi}^{*(\alpha_2)}\right)$$

[13] A location statistic is a statistic such that a decrease in the variable by a constant value of c decreases the statistic itself by c.

[14] Instead of using the asymptotic sampling distribution, the bootstrapping technique uses an exact distribution that acts as real distribution of the population.

where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}\left(\hat{z}_0 + z^{(\alpha)}\right)}\right)$$

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}\left(\hat{z}_0 + z^{(1-\alpha)}\right)}\right).$$

Here, $\Phi(.)$ is the standard normal cumulative distribution function and $z^{(\alpha)}$ is the $100\alpha th$ percentile point of a standard normal distribution (Efron, Tibshirani, 1993, p. 185). For example, $z^{(0.90)} = 1.285$ and $\Phi(1.285) = 0.90$.

The results from the percentile-t and BC_a methods will be equal when $\hat{a}$ and $\hat{z}_0$ are zero. However, this is not generally the case. The value of bias-correction, $\hat{z}_0$, can be directly obtained by counting the number of the bootstrap cases that give a value less than the original estimate for the $\hat{\xi}$, i.e.,

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\xi}^*(b) < \hat{\xi}}{B}\right),$$

here, $\hat{z}_0$ measures the median bias of $\hat{\xi}^*$.

There are several ways to compute the acceleration parameter $\hat{a}$. The easiest way is to calculate it using *jackknife* values of $\hat{\xi}$. In the *jackknife* method, $\hat{\xi}_{(i)}$ shows the value of the statistic calculated from a sample where the ith observation is deleted. The expression of $\hat{a}$ is

$$\hat{a} = \frac{\sum_{i=1}^{n}\left(\hat{\xi}_{(.)} - \hat{\xi}_{(i)}\right)^3}{6\{\sum_{i=1}^{n}\left(\hat{\xi}_{(.)} - \hat{\xi}_{(i)}\right)^2\}^{3/2}},$$

where $\hat{\xi}_{(.)} = \sum_{i=1}^{n}\hat{\xi}_{(i)}/n$.

Kilian (1998), introduced a biased-corrected percentile bootstrapping method to calculate the confidence interval for the impulse responses. Assume that in a VAR context, $\varphi_{kl,i}$ shows the response of variable k to a one-time impulse in variable l, in i periods ago. The asymptotic distribution of $\hat{\varphi}_{kl,i}$ is normal; however, the small sample distribution of this estimate could be biased and skewed (Kilian, 1998, p. 219). Bootstrapping methods allow for skewness by not imposing symmetry. However, these methods assume that $\hat{\varphi}_{kl,i}$ is unbiased, which might not be the case in small samples. Using the biased estimated coefficients will cause a bias in the estimated confidence interval. Therefore, it is necessary to use an unbiased estimation of the coefficients in constructing the intervals. Kilian (1998), attempts to tackle this problem in constructing the confidence intervals, based on bootstrapping methods, by proposing an indirect way of correcting for the bias. Kilian's biased-corrected bootstrap method consists of the following steps. First, estimate the $VAR(k)$ model and generate 1000 bootstrapping replications $\hat{\beta}^*$ from

$$y_t^* = \hat{\nu} + \hat{B}_1 y_{t-1}^* + \cdots + \hat{B}_k y_{t-k}^* + \varepsilon_t^* \tag{3.1}$$

using standard bootstrapping techniques. Then approximate the bias term Ψ by $\hat{\Psi} = \hat{\beta}^* - \hat{\beta}$, where * shows the bootstrap replications. Second, calculate the modulus of the largest root and the companion matrix associated with $\hat{\beta}$ and denote that by $m(\hat{\beta})$. If $m(\hat{\beta}) \geq 1$, use $\hat{\beta}$ without any adjustments. However, if $m(\hat{\beta}) < 1$, use the bias corrected coefficient $\tilde{\beta} = \hat{\beta} - \hat{\Psi}$. If $m(\tilde{\beta}) \geq 1$, let $\hat{\Psi}_1 = \hat{\Psi}$ and also $\varpi_1 = 1$ and define the bias as $\hat{\Psi}_{i+1} = \varpi_i \hat{\Psi}_i$ and $\varpi_{i+1} = \varpi_i - 0.01$. Use the new biased corrected $\tilde{\beta}$ for $i = 1, 2, \ldots$ until $m(\tilde{\beta}_i) < 1$. Now, substitute $\tilde{\beta}$ for $\hat{\beta}$ in equation (**??**) and using bootstrapping techniques, generate 2000 replicates $\hat{\beta}^*$. $\hat{\Psi}$ from the previous step can be used as a proxy for $\hat{\Psi}^*$. Third, calculate $\tilde{\beta}^*$ and $\hat{\Psi}^*$ by following the algorithm in the second step. Finally, calculate the α and $1 - \alpha$ percentile interval endpoints of the distribution of $\hat{\varphi}_{kl,i}^*$. This approach is designed for stationary models and it is not asymptotically valid for non-stationary models (Kilian, 1998, p.255). Nevertheless, it has good coverage for non-stationary cases without time trends.

The Percentile-t method of Hall (1992) uses OLS to estimate the value of the autoregressive coefficient and uses this estimated value as the true value to evaluate the sampling distribution of the t-statistic and build the confidence interval. However, as noted by Hansen (1999, p. 595), the confidence interval built based on the percentile-t method has incorrect first order asymptotic coverage. That is, the interval constructed using this approach cannot properly control type I error. This is even true for large samples. In addition, the conventional bootstrap cannot satisfy the first-order asymptotic consistency since the asymptotic distributions from this approach depend on c which is not estimable (Basawa et al. 1991). Therefore, Hansen (1999) proposed using a grid bootstrapping procedure. This procedure gives a confidence interval that is asymptotically valid and can control for type I error, even in the cases where the root is close to one.

Hansen (1999), using Monte Carlo simulations, found that the grid point bootstrapping way of constructing a confidence interval is asymptotically valid and provides good coverage in finite samples. Two types of grid intervals proposed by Hansen (1999) are grid-α and grid-t, which are defined based on the type of statistics chosen for testing the hypothesis. In the former case, the statistic is $b(\xi) = \hat{\xi} - \xi$ and in the latter case, this statistic is a t-statistic $t(\xi) = (\hat{\xi} - \xi)/s(\hat{\xi})$. Contrary to conventional bootstrapping methods the grid search method does not use the point estimate of $\hat{\xi}$. Instead, the bootstrap quantile function is a function free of ξ (Hansen , 1999 p. 596).

To construct the confidence interval, we must have the bootstrap quantile function $q_n^*(\alpha|\xi)$.However, in practice this function is unknown, therefore, it must be estimated. To estimate this function we can use simulation and bootstrapping as follows: Suppose the bootstrap distribution of the sample for a given ξ is $G_n^*(x|\xi) = G_n(x|\xi, \hat{\eta}(\xi))$,where ξ is the parameter of interest and η is the nuisance parameters. Then generate B random samples X_n^* from this distribution. For each generated sample, calculate the statistic $S_n^*(\xi)$, where $S_n^*(\xi)$ is a nondegenerate test statistic to test the hypothesis $H_0 : \xi = \xi_0$, and sort the B calculated $S_n^*(\xi)$.The simulation estimate of $q_n^*(\alpha|\xi)$ will be $100\alpha\%$ order statistic $\hat{q}_n^*(\alpha|\xi)$. Now, construct $q_n^*(\alpha|\xi)$ for different possible ξs in a grid, say $A_G = [\xi_1, \xi_2, ..., \xi_G]$.Then use kernel regression to smooth the estimated $\hat{q}_n^*(\alpha|\xi)$. The kernel estimate for every given ξ is

$$\tilde{q}_n^*(\alpha|\xi) = \frac{\sum\limits_{j=1}^{G} K(\frac{\xi-\xi_j}{h})\hat{q}_n^*(\alpha|\xi)}{\sum\limits_{j=1}^{G} K(\frac{\xi-\xi_j}{h})}.$$

In this function $K(.)$ is a kernel function and h is a bandwidth. These smoothed $\tilde{q}_n^*(\alpha|\xi)$ can be used to construct the confidence set C_g :

$$\hat{C}_g = \{\xi \in R : \tilde{q}_n^*(\alpha_1|\xi) \leq S_n(\xi) \leq \tilde{q}_n^*(\alpha_2|\xi)\}.$$

The intersection of $\tilde{q}_n^*(\alpha_1|\xi)$ and $S_n(\xi)$ will give the lower bound of the confidence interval for ξ and the intersection of $\tilde{q}_n^*(\alpha_2|\xi)$ and $S_n(\xi)$ will give the upper bound. Now, assume that the observed data y_t can be explained by an $AR(k)$ model

$$y_t = \mu_0 + \beta t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_k y_{t-k} + e_t$$

or in ADF form as

$$y_t = \mu_0 + \beta t + P y_{t-1} + \rho_2 \triangle y_{t-1} + ... + \rho_k \triangle y_{t-k+1} + e_t.$$

In this representation, the parameter of interest (ξ) is P , the sum of the autoregressive coefficients, and η denotes the remaining nuisance parameters. In addition, k is the optimal number of lags, to be determined in such a way that no autocorrelation remains in the residuals. The confidence interval for P can be constructed using grid bootstrapping by following the above steps.

Hansen (1999) compared the performance of several ways of constructing the confidence interval. Using Monte Carlo simulations, Hansen (1999, p. 600) showed that for an AR(1) case, the percentile-t method works well, only when the sample size is large and the roots are inside the unit circle, i.e., stationary. Kilian's biased-corrected percentile method works in the same way as percentile-t methods. However, its performance is poor when the sample size is small or the roots are near unity. Stock's method to construct the confidence interval performs well in the cases where the root is unity, close to unity, and even when the root is explosive. However, its performance decreases when the root is small and stationary. Hansen (1999) concludes that the grid bootstrap methods have the best performance in constructing the confidence intervals.

The approach of Bai and Perron

Bai and Perron (1998, 2003) proposed a method to estimate the multiple structural changes in a linear model using OLS. Their method gives the possibility of testing for break against stability of the model. In addition, it is possible to test for a given number of breaks using different tests, including information criteria, such as BIC and LWZ (Liu et al, 1997). Applying this method to an autoregressive model will provide estimations for the autoregressive coefficients before and after the break points, which can be used to compare persistence in different regimes.[15]

[15] One shortcoming with this method is the fact that it provides a point estimation for the autoregressive coefficients not a confidence interval for them.

Following the same notation as Bai and Perron (1998, 2003), consider the following multiple linear regression with m breaks:

$$y_t = \begin{cases} x_t'\beta + z_t'\delta_1 + \epsilon_t & t = 1, 2, \ldots, T_1 \\ x_t'\beta + z_t'\delta_2 + \epsilon_t & t = T_1 + 1, \ldots, T_2 \\ \vdots & \\ x_t'\beta + z_t'\delta_{m+1} + \epsilon_t & t = T_m + 1, \ldots, T \end{cases}$$

where x_t $(p \times 1)$ and z_t $(q \times 1)$ are vectors of covariates and β and δ_j $(j = 1, 2, ..., m+1)$ are the corresponding vectors of coefficients and ϵ_t is the disturbance term at time t. This method attempts to estimate the unknown coefficients along with the unknown break dates. For a pure structural change model, where all the parameters can have shifts, $p = 0$, whereas, for a partial structural change model, where some of the parameters do not have any break or shift, we have $p > 0$ and $q > 0$.

In the estimation process we use OLS. To obtain the estimates of the coefficients and the break dates, we can use an iterative procedure as follows. Assuming $\theta = (\delta, T_1, ..., T_m)$, define the sum of squared residuals as $SSR(\beta, \theta)$. Using the method introduced by Sargan (1964), minimize $SSR(\beta, \theta)$ by iteration. First, assume β is fixed and minimize $SSR(\beta, \theta)$ with respect to θ and then keep θ constant and minimize $SSR(\beta, \theta)$ with respect to β. Iterate this procedure until the decrease in the objective function becomes less than a given number. Bai and Perron (1998, 2003) recommend, for efficiency purpose, to keep $\{T_1, ..., T_m\}$ constant at the second step and to minimize the objective function with respect to β and δ simultaneously.

To estimate the number of breaks Perron and Yabu propose different statistics. The first statistic is the $\sup F$ type test for no break, $m = 0$, against an alternative hypothesis of $m = k$ breaks. Let R be a matrix such that $(R\delta)' = (\delta_1' - \delta_2', ..., \delta_k' - \delta_{k+1}')$; define

$$F_T(\lambda_1, \ldots, \lambda_k; q) = \frac{1}{T}(\frac{T - (k+1)q - p}{kq})\hat{\delta}' R'(R\hat{V}(\hat{\delta})R')^{-1}R\hat{\delta}$$

where $\hat{V}(\hat{\delta})$ is the estimate of the variance covariance matrix of $\hat{\delta}$ which is robust to serial correlation and heteroscedasticity. Now, define the $\sup F_T(k; q) = F_T(\hat{\lambda}_1, ..., \hat{\lambda}_k; q)$ where $\hat{\lambda}_1, ..., \hat{\lambda}_k$, $\hat{\lambda}_i = \hat{T}_i/T$, minimize the global $SSR(.)$.[16] The asymptotic distribution of this statistic depends on the minimal length h for every

[16] This is equivalent to maximizing the $F-$test assuming spherical errors.

segment or in another words, it depends on the trimming parameter. The critical values can be found in Bai and Perron (2003). In addition, the maximum number of the breaks depends on the trimming parameter, for example for a 15% trimming, the maximum number of breaks will be 5.

In using the above-mentioned statistic, the researcher must test no break versus a pre-specified number of breaks. However, Bai and Perron propose two other tests, called double maximum tests, which can be used to test for no break against an unknown number of breaks given some upper limit for the number of breaks M. The first test is defined as $UD_{MAX}F_T(M,q) = \max_{1\leq m\leq M} F_T(\hat{\lambda}_1,...,\hat{\lambda}_m;q)$. The second test denoted by $WD_{MAX}F_T(M,q)$ is defined as $WD_{MAX}F_T(M,q) = \max_{1\leq m\leq M} \frac{c(q,\alpha,1)}{c(q,\alpha,m)} F_T(\hat{\lambda}_1,...,\hat{\lambda}_m;q)$, where $\hat{\lambda}_j$ are obtained using global minimization of the $SSR(.)$ and $c(q,\alpha,m)$ is the critical value of the sup $F_T(k;q) = F_T(\hat{\lambda}_1,...,\hat{\lambda}_k;q)$ for a given significance level α. The critical values can be found in Bai and Perron (2003).

Another statistic is proposed to test for a given number of breaks l, against $l+1$ breaks, which denoted sup $F_T(l+1|l)$. A model with $l+1$ break will be selected if the overall minimal value of the $SSR(.)$ is sufficiently smaller than the $SSR(.)$ from a model with l breaks. We can repeat this test for $l = 1$ to $l = M$. Additionally, the number of breaks can be determined using the information criteria, Bayesian Information Criteria (BIC) and Liu, Wu, and Zidek (LWZ) (Liu et al, 1997) defined as follows:

$$\begin{aligned} BIC(m) &= \ln\hat{\sigma}(m) + g\ln(T)/T \\ LWZ(m) &= \ln(\frac{S_T(\hat{T}_1,...,\hat{T}_m}{T-g}) + (\frac{g}{T} \times c_0(\ln(T))^{2+\delta_0}) \end{aligned}$$

where $g = (m+1)q + m + p$, and $\hat{\sigma}^2(m) = T^{-1}S_T(\hat{T}_1,...,\hat{T}_m), c_0 = 0.299$ and $\delta_0 = 0.1$. Any number of breaks that minimizes these information criteria is the selected number of breaks for that time series. Finally, repartition and sequential procedures can be used to determine the exact number of breaks.

3.3 Empirical results

Quarterly data on unemployment rates from 1976:1-2005:3 for Canada and the ten Canadian provinces are used in this study to examine the hysteresis hypothesis. The ten provinces are Alberta, British Columbia (BC), Manitoba, New Brunswick (NB), Newfoundland (NF), Nova Scotia (NS), Ontario, Prince Edward Island (PEI), Quebec, and Saskatchewan.

Here, we want to study the properties of these unemployment rates. We would like to see whether the traditional theory of unemployment or the hysteresis theory of unemployment can describe the movements of these time series. In the former case, the shocks to the unemployment rates are temporary and they will return to their long-run level, which is the natural rate of unemployment. On the other hand, if they can be described by hysteresis theories, then any shock to these variables will have permanent effect and they will never return to the original steady state level. To distinguish between these theories, we can use the integration order of the series. Any stationary series will satisfy the condition for the former theory and a non-stationary series conforms with hysteresis theories.

Consequently, unit root tests have been popular tool to empirically test the hysteresis hypothesis. Hence, in this paper, we start with testing for existence of unit root in the unemployment rates of Canada and the Canadian provinces. To that end, we use different unit root tests including ADF, ADF^{GLS} test of Elliot, et. al. (1996), ADF^{GLS} test with one break (Perron and Rodríguez, 2003 and Rodríguez, 2007), and a minimum LM unit root test with two breaks, proposed by Lumsdaine and Papell (1997).

Table 1 presents the results obtained from the application of ADF and ADF^{GLS} unit root tests on the unemployment rates. Lag length is selected using Bayesian information criterion (BIC). The results show that the null hypothesis of unit root can be rejected for NB, NF, Ontario, and Canada. However, following Perron (1989), now it is well known that ADF type unit root tests may fail to reject the null hypothesis of unit root if there is an exogenous change in the deterministic part of the series. In other words, if the series is stationary around a broken trend, then these tests will fail to reject the hypothesis of unit root, even though the series is trend stationary. Therefore, the possibility of structural breaks has to be taken into consideration when we are testing for unit root. Here, we use an ADF^{GLS} test, which allows for a shift in level or broken trend, or both a shift in level and a broken trend, proposed by Perron and Rodríguez (2003) and Rodríguez (2007). The results of applying the unit root test of Perron and Rodríguez (2003) on the unemployment rates are shown in Table 2. We allow for a shift in the intercept and also a broken time trend in testing for unit root in these series. The results, using the P_T and MSB statistics, show that by introducing one structural break in the time series of the unemployment rates, we can reject the hypothesis of unit root in Alberta, Manitoba, NB, NF, Ontario, PEI, Saskatchewan, and Canada. The estimated break dates, T_{B_1}, are shown in the last column of the table. In addition, we use a unit root test recently proposed by Rodríguez (2007), which allows for testing the crash models, that is, a shift in level of the series. This test uses two methods to select the break date: 1) the SUPREMUM method 2) the INFIMUM method. Results from applying this test are reported in Table 3. Again, we use P_T and MSB statistics to test for the unit root hypothesis. According to table 3, using this test we can only reject the null hypothesis of unit root for NB and Ontario.

The unit root tests are very sensitive to the specification of the deterministic part of the model. Therefore, if there is more than one break point in the data and we restrict the number of breaks to one, then we misspecify the model. Hence, we use the minimum LM unit root test proposed by Lumsdaine and Papell (1997), which allows for two breaks in the series. Two models are estimated for this unit root test: 1) a crash model and 2) a model with a shift in the intercept and a broken trend. The results are reported in Table 4. According to this table, we can reject the unit root hypothesis in Alberta, BC, Manitoba, NF, Saskatchewan, and Canada using the crash model. However, using the model with broken trend and intercept, we can reject this hypothesis only for Ontario, PEI, Saskatchewan, and Canada. The estimated break dates are shown by T_{B_1} and T_{B_2}.

In sum, the results of applying different unit root tests on the unemployment rate series indicate that except for NS and Quebec, the series for other provinces do not have a unit root. In other words, the unemployment rates in Canada and the Canadian provinces are not consistent with hysteresis hypothesis, the exception being the unemployment rates in NS and Quebec.

Alternatively, we can investigate the impulse response function of the variable. However, point estimates of an impulse response alone is not enough and we need to have the associated confidence interval of the autoregressive coefficient to test for existence of persistence in the unemployment rate series. There are several ways of obtaining the confidence interval; however, in this paper we use seven methods to construct the confidence intervals. These methods are conventional asymptotic confidence interval($\hat{\rho}_i \pm 1.645 \times s(\hat{\rho}_i)$), Stock's (1991) local to unity asymptotic confidence interval, percentile bootstrap of Efron and Tibshirani (1993), percentile-t bootstrap of Hall (1992), biased-corrected percentile bootstrap of Kilian (1998), and grid-α and grid-t of Hansen (1999)[17]. We consider two cases in constructing these confidence intervals: 1) a model with an intercept only and 2) a model with an intercept and a time trend. The estimated 90%-confidence intervals, based on the above mentioned methods, are shown in Table 5 and 6. If the calculated interval does not include the unity, we conclude that the series does not have a unit root. Therefore, it is not persistent and cannot satisfy the condition for hysteresis. In contrast, if this interval includes the unity, then the time series does have a unit root and is persistent.

The calculated confidence interval, using percentile bootstrap of Efron and Tibshirani (1993), is the only case that does not contain unity in any of the unemployment rate series. The average lower (upper) bound of the estimated Intervals are 0.787 (0.950) and 0.835 (0.968) for a model with a time trend and a model with no time trend, respectively. The estimated 90% confidence intervals based on the conventional asymptotic method do not include unity for most, but not all series. At the same time, the other methods show that the 90%-confidence intervals contain unity for most series, the exception being the unemployment rate in NF. The estimated confidence intervals for the unemployment rate in this province do not include unity, based on most methods.

As it is well known, the conventional asymptotic method of building a confidence interval is incorrect when the root is near to unity. One way to resolve this problem is by using the grid bootstrapping method proposed by Hansen (1999), which has superior performance, with respect to other methods.[18] Hence, we use the results from grid-t and grid-α of Hansen (1999) to make inference about the per-

---

[17] We use a GAUSS code by Hansen (1999) to carry out all confidence interval calculations.

[18] Hansen, 1999, p. 601.

sistence of these unemployment rates. We use 1999 replications and a grid with 200 gridpoints. According to the results of these methods (shown in the last two rows of the table 5 and 6), the 90% confidence intervals of all series contain unity. Therefore, all the unemployment rates in the Canadian provinces have unit root and are persistent. The only exception is the unemployment rate in NF, which a unit root is excluded when we include only an intercept in the model. The average widths of the intervals, based on Hansen's grid-α and grid-t in the models with an intercept and a time trend (only an intercept), are 0.106 and 0.098 (0.103 and 0.099).

These results are calculated assuming that the optimal autoregressive order of the series is 1; however, this may not always be the case. Therefore, we check for validity of this assumption. We use the Bayesian information criterion (BIC) to select the optimal autoregressive order of the series. The results show that it is 2 for Alberta, BC, Ontario, Quebec, and Canada, 4 for NF, and 1 for the rest. Using these autoregressive orders we calculate the 90% confidence interval, based on grid-t bootstrap method of Hansen (1999) and percentile-t bootstrap (Hall, 1992). The results are reported in Table 7 and 8, respectively, for a model with an intercept, a time trend, and a model with only an intercept. These tables show the estimated Ps and their standard deviations, along with the 90% confidence intervals. Notice that the estimated lower bound of $\hat{P}$ for all series are larger than 0.90, with exception of NB and NF in Table 7 and NF and PEI in Table 8.

The confidence intervals calculated based on grid-t bootstrap are wider than the confidence intervals of percentile-t bootstrap for all series. In addition, the right endpoints of the confidence intervals are larger using the grid-t bootstrap than the percentile-t. On the other hand, the left endpoints of the confidence intervals are larger using the percentile-t than the grid-t bootstrap. The 90% confidence intervals for P calculated using grid-t bootstrap, contain unity for all the unemployment rate series and indicate that all series are persistent, the exception being NF when we do not include a time trend. Nevertheless, the results from using percentile-t method show that the 90% confidence interval for unemployment rates in NB, Ontario, Quebec, and Canada do not contain unity, and therefore, are stationary. However, even if we accept the results from percentile-t method and conclude that the unemployment rate series in NB, Ontario, Quebec, and Canada are stationary, we may say that the series have a high degree of persistence since the lower bound of the 90% confidence interval for these series are larger than 0.90.

The other objective of this paper is to assess the stability of the persistency of unemployment rates, after allowing for breaks in these series. The benefits of allowing for break in the data are twofold. First, as Perron (1989) stated, if the data has a shift in the level and we ignore this fact, then we can mistakenly conclude that the data is persistent. Secondly, by allowing for break we can analyze the stability of the persistency in the data, i.e., whether or not it is the same for both regimes. Therefore, we use the approach of Bai and Perron (1998, 2003), which allows testing multiple breaks at unknown dates. The exact number of breaks can be determined using four different methods: the modified Schwarz criterion, LWZ (Liu, et. al., 1997); the Schwarz information criterion, BIC (Yao, 1988); repartitioning procedure and sequential procedure (Bai and Perron, 1998, 2003).

The results of applying the Bai and Perron (1998, 2003) approach to the unemployment rates, allowing up to 5 breaks, are shown in Table 9-11. Here, we use the optimal autoregressive orders, previously calculated. In table 9, UD_{MAX} , WD_{MAX}, and sup $F_T(.)$ tests are significant for all series, which indicate that at least one break is present in theses series. However, notice that the sup $F_T(1)$ is not significant for the unemployment rate in Quebec, therefore, there is no structural break in the unemployment rate in Quebec.

In order to determine the exact number of breaks, we use the sequential procedure. According to the results, there are 3 break points in the unemployment rates in NS and Ontario; only 1 break point in the unemployment rates in Alberta, BC, Manitoba, NF, PEI, and Saskatchewan; and no break in Quebec. The number of break points for Canada and NB is 2. The estimated break dates are very precise and for most provinces, the first break date is around 1981-1982 and the second break dates are around 1989-1990. Ontario and NS are the only provinces that have the third break date and it happens in 2000:4 and 1993:3, respectively. Seven breaks occur during 1981-1982, the recession period, five breaks are between 1988 and 1990, two breaks in 1993, and two breaks in 2000, which show some sort of clustering.

Table 10 reports the estimated intercepts, slopes, and the autoregressive coefficients, ρ_is, for each series. According to this table, ρ_is are not constant in all the regimes and the persistence of the unemployment rates is different in the separate regimes. Furthermore, among 16 break points that exist in these series, in 11 cases the new level of unemployment rate is higher than its level before break and only in 5 cases, the new unemployment rate is lower. In BC, the unemployment rate is very persistent ($P = 1.157$)[19] before the break in 1982:1, however, its persistency declines after this break point ($P = 0.895$). The persistency of the unemployment rates in Alberta, NB, NS, Ontario, and Saskatchewan increases in periods after the first break date. The estimated ρ_1 for Ontario is larger than unity for all regimes, except the regime 4, which starts from 2000:4. However, none of the regimes are persistent with Ps equal to 0.511, 0.639, 0.806, and 0.442, respectively, for regime one to four. For the unemployment rate in Canada, this approach has determined 2 break points and, therefore, 3 regimes. In all these regimes the estimated ρ_1 is larger than the unity but the series is not persistent. Furthermore, except for the ρ_1 in the third regime in NB, which is negative (and insignificant), all the estimated $\rho_1 s$ are positive. Finally, the estimated model with no break for the unemployment rate in Quebec shows that the α is 0.927, consequently, it is stationary but persistent. Overall, the results from applying the approach of Bai and Perron to the unemployment rate series show that the unemployment rate in BC is persistent in all regimes, and the unemployment rate in Manitoba and NS are persistent in regime 1 and 2, respectively. Finally, the unemployment rate in Quebec does not have any break and is persistent.[20] The unemployment rates in all other provinces and regimes are not persistent.[21]

Overall, the unemployment rates in Quebec and NS are the only series in which all methods confirm the existence of persistence. According to the results from Hansen's (1999) grid-t bootstrap method, all series are persistent. The persistence of the unemployment rate in BC, in both regimes, and Manitoba in regime 1 is also confirmed by the results from the approach of Bai and Perron (1998, 2003). On the other hand, the results from unit root tests and the approach of Bai and Perron (1998, 2003) show that the unemployment rates in the other provinces are not persistent, al-

---

[19] Notice that P is the sum of the autoregressive coefficients in the models.

[20] Since the approach of Bai and Perron is not robust against nonstationarity of the series and we could not reject the null hypothesis of unit root for unemployment rates in NS and Quebec, the results from the approach of Bai and Perron for these two provinces should be approached cautiously.

[21] Since the sum of the autoregressive coefficients are less than one in all regimes.

though they have at least one break. There is strong evidence against the traditional theory, which says there is a constant natural rate of unemployment, since all series have at least one break point or has a unit root. On the other hand, nine out of eleven unemployment series have, at most, two structural breaks, which indicates that there have been only a few permanent changes in the unemployment rates in Canada and the Canadian provinces.

3.4 Conclusion

Data on unemployment rates in Canada and the ten Canadian provinces are used to check for persistence of the unemployment rates. We test for null hypothesis of unit root against stationarity, first without allowing for any break. Results indicate that only four series are stationary and seven series have unit roots. Next, we apply unit root tests, which allows to have one and two break points in the data and find that all series are stationary, except for Nova Scotia and Quebec.

Furthermore, we attempt to test for persistence of these series by constructing the confidence interval using different methods and find that the calculated 90% confidence interval contains unity for all series, except the unemployment rate in Newfoundland, which does not include unity when we have only an intercept in the model. Hence, all series are difference stationary except the unemployment rate in Newfoundland, which is trend stationary.

Next, we proceed by testing for the existence of multiple break points and investigating the changes in the persistence in different regimes of these series. In all series, except Quebec, we can reject the null hypothesis of stability and find at least one break point in the data. The break dates are clustered around 1981-82, 1988-90, 1993, and 2000. Additionally, most of the regimes have a higher intercept, which interpret as an increasing unemployment rate. Also, the unemployment rate in British Columbia is persistent in all regimes and the unemployment rates in Manitoba and Nova Scotia are persistent in regime 1 and 2, respectively. Finally, the unemployment rates in all other provinces and regimes are not persistent.

These results indicate that hysteresis theory only seems relevant in Quebec and Nova Scotia, since we could not reject the possibility of unit root for the unemployment rate in these provinces.

References

[1] Arestis, P. and Marscal, I. B. F., (2000), "OECD unemployment: structural breaks and stationarity," *Applied Economics*, **32**, 399- 403.

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

[3] Bai, J., and Perron, P., (2003), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, **18**, 1-22.

[4] Basava, I. V., Mallik, A. K., Cormick, W. P., Reeves, J. H. and Taylor, R. L., (1991), "Bootstrapping Unstable First-Order Autoregressive Processes," *Annals of Statistics*, **19**, 1098-1101.

[5] Bianchi, M. and Zoega, G., (1998), "Unemployment Persistence: Does the Size of the Shock Matter?" *Journal of Applied Econometrics*, **13**, 283-304.

[6] Blanchard, O. J. and Summers, L. H., (1986), "Hysteresis and the European unemployment problem," NBER Macroeconomics Annual, Cambridge, MA, pp. 15–78.

[7] Blanchard, O. J. and Summers, L. H., (1987), "Hysteresis in Unemployment," *European Economic Review*, **31**, 288-95.

[8] Bobkoski M. J., (1983), "Hypothesis testing in nonstationary time series." Unpublished Ph.D. thesis, University of Wisconsin, Madison.

[9] Camarero, M. and Tamarit, C., (2004), "Hysteresis vs. natural rate of unemployment: new evidence for OECD countries," *Economics Letters*, **84**, 413–17.

[10] Cavanagh, C., (1985), "Root local to unity," Manuscript, Harvard University.

[11] Dickey, D. and Fuller, W. (1979), "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association,* **74**, 427-431.

[12] Efron, B., and Tibshirani, R. J., (1993), *An Introduction to the Bootstrap*, Chapman & Hall/CRC, CRC Press Inc.

[13] Elliot, G., Rothenberg, T. and Stock, J. H. (1996), "Efficient tests for an autoregressive unit root," *Econometrica*, **64**, 813-39.

[14] Fortin, Pierre, (1989), "How "Natural" is Canada's High Unemployment Rate?" *European Economic Review,* **33**, 89-110.

[15] Fortin, Pierre, (1991), "The Philips Curve, Macroeconomic Policy and the Welfare of Canadians," *Canadian Journal of Economics*, **24**, 774-803.

[16] Friedman, Milton, (1968), "The Role of Monetary Policy," *American Economic Review*, **58**, 1–17.

[17] Gil-Alana, L. A., (2002), "Modelling the Persistence of Unemployment in Canada," *International Review of Applied Economics*, **16**, 465-77.

[18] Gustavsson, M. and Osterholm, P., (2006), "Hysteresis and non-linearities in unemployment rates," *Applied Economics Letters*, **13**, 545–48.

[19] Hall, P., (1992), *The Bootstrap and Edgeworth Expansion*, Springer Series in Statistics, New York, Springer-Velrag New York Inc.

[20] Hansen, B. E., (1999), "The Grid Bootstrap and the Autoregressive Model," *The Review of Economics and Statistics*, **81**, 594-607.

[21] Jaeger, A. and Parkinson, M., (1994), "Some evidensce on hysteresis in unemployment rates," *European Economic Review*, **38**, 329-42.

[22] Jones, Stephen, R. G., (1995) , *The persistence of Unemployment: Hysteresis in Canadian Labour Market*, Montreal, McGill Queen's University Press.

[23] Kapetanios, G., Shin, Y. and Snell, A., (2003), "Testing for a unit root in the nonlinear STAR framework," *Journal of Econometrics*, **112**, 359–79.

[24] Keil, M. and Pantuosco, L., (1998), "Canadian and US Unemployment Rates: A Comparison Based on Regional Data," *Canadian Public Policy*, **24(S1)**, 38-55.

[25] Kilian, L., (1998), "Small sample confidence intervals for impulse response functions," *Review of Economics and Statistics*, **80**, 218-30.

[26] Koustas, Z. and Veloce, W., (1996), "Unemployment Hysteresis in Canada: an Approach Based on Long Memory Time Series Models," *Applied Economics*, **28**, 823-31.

[27] Kwiatkowsky, D., Phillips, P.C.B., Schmidt, P. and Shin, Y., (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root," *Journal of Econometrics,* **54**, 159–178.

[28] Lumsdaine, R. and Papell, D., (1997), "Multiple trend breaks and the unit root hypothesis," *Review of Economics and Statistics,* **79**, 212–8.

[29] Liu, J., Wu, S., and Zidek, JV. (1997), "On segmented multivariate regressions," *Statistica Sinica*, **7**, 497-525.

[30] Mikhail, O., Eberwein, C. J. and Handa, J., (2005), "Testing for persistence in aggregate and sectoral Canadian unemployment," *Applied Economics Letters*, **12**, 893–98.

[31] Mikhail, O., Eberwein, C. J. and Handa, J., (2006), "Estimating persistence in Canadian unemployment: evidence from a Bayesian ARFIMA," *Applied Economics*, **38**, 1809–19.

[32] Milbourne, R. D., Purvis, D. D. and Scoones, W. D., (1991), "Unemployment Insurance and Unemployment Dynamics," *The Canadian Journal of Economics*, **24**, 804-26.

[33] Ng, S. and Perron, P., (2001), "Lag length selection and the construction of unit root tests with good size and power," *Econometrica*, **69**, 1519-54.

[34] Nott, Loretta, (1996), "Hysteresis in the Canadian labour market: Evidence from the 1990s," Research Report no. 9605, Department of Economics, University of Western Ontario.

[35] Papell, D. H., Murray, C. J. and Ghiblawi, H., (2000), "The Structure of Unemployment," *The Review of Economics and Statistics*, **82**, 309-15.

[36] Perron, P., (1989), "The great crash, the oil price shock and the unit root hypothesis," *Econometrica,* **57**, 1361–1401.

[37] Perron, P., (1997), "Further evidence on breaking trend functions in macroeconomic variables," *Journal of Econometrics*, **80**, 355- 85.

[38] Perron, P. and Rodríguez, G. (2003), "GLS detrending, efficient unit root tests and structural change," *Journal of Econometrics*, **115**, 1-27.

[39] Phelps, E.S., (1967), "Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time." *Economica,* **34**, 254–81.

[40] Philips, P. C. B., (1987), "Toward a unified asymptotic theory for autoregression." *Biometrika*, **74**, 535-47.

[41] Rodríguez. G. (2007), "Finite Sample Bahavior of the Level Shift Model Using Quasi-Difference Data", forthcoming in the Journal of Statistical Computation and Simulation.

[42] Røed, K., (1996), "Unemployment hysteresis – macro evidence from 16 OECD countries," *Empirical Economics,* **21**, 589–600.

[43] Said E. and David A. Dickey (1984), "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order," *Biometrika*, **71**, 599–607.

[44] Sargan JD., (1964), "Wages and prices in the United Kingdom: a study in econometric methodology," in *Econometric Analysis for National Economic Planning,* Hart PE, Mills G., Whitaker JK. (Eds), Butterworths, London, 25-54.

[45] Song, F. M. and Wu, Y., (1997), "Hysteresis in unemployment: Evidence from 48 U.S. states" *Economic Inquiry*, **35**, 235-43.

[46] Stock, J., (1991), "Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series." *Journal of Monetary Economics,* **28**, 435–59.

[47] Stock, J., (1999), "A class of tests for integration and cointegration," in Engle, R.F. and White, H. (Eds), *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, Oxford University Press, Oxford, 137-67.

[48] Torous, W., Valkanov, R. and Yan, S., (2004), "On Predicting Stock Returns with Nearly Integrated Expanatory Variables," *Journal of Business*, **44**, 937-66.

[49] Wu, Z., (2003), "The persistence of regional unemployment: evidence from China," *Applied Economics*, **35**, 1417–21.

[50] Zivot, E. and Andrews, D. W. K., (1992), "Further evidence of the great crash, the oil-price shock and the unit-root hypothesis," *Journal of Business and Economic Statistics,* **10**, 251–70.

Tables

Table 1:Results from unit root tests without break

	ADF			ADF^{GLS}	
Series	no regressors	with c	with c & t	with c	with c & t
Alberta	-0.514	-1.597	-1.741	-1.213	-1.277
BC	-0.707	-1.515	-1.849	-1.551	-1.658
Manitoba	-0.365	-1.539	-1.912	-1.154	-1.217
NB	-0.391	-1.836	-2.687	-1.655^{d}	-1.869
NF	-0.015	-2.565^{d}	-2.225	-1.091	-1.529
NS	-0.221	-1.715	-2.143	-0.949	-0.929
Ontario	-0.530	-2.792^{d}	-2.792	-2.326^{c}	-2.552
PEI	-0.218	-2.195	-1.968	-1.423	-1.795
Quebec	-0.302	-2.067	-2.621	-1.388	-1.603
Saskatchewan	-0.105	-1.965	-1.808	-1.013	-1.236
Canada	-0.446	-2.492	-2.743	-1.885^{d}	-2.134

a,b,c, and d shows significant at 1%, 2.5%, 5%, and 10%.

Table 2: Results of Unit Root Tests with 1 Break in Intercept and Time Trend

Series	Criteria	MSB	P_T	k	T_B
Alberta	BIC	0.144	10.839^{d}	2	1984:1
	MAIC	0.181	16.835	1	1985:4
BC	BIC	0.184	17.893	1	1985:4
	MAIC	0.184	17.893	1	1985:4
Manitoba	BIC	0.099^{b}	5.093^{a}	3	1991:3
	MAIC	0.204	21.43	1	1992:4
NB	BIC	0.139	9.632^{c}	1	1986:2
	MAIC	0.139^{d}	9.632^{a}	1	1986:2
NF	BIC	0.136	9.993^{c}	3	1981:2
	MAIC	0.162	14.143	4	1997:2
NS	BIC	0.199	20.543	1	1984:1
	MAIC	0.199	20.543	1	1984:1
Ontario	BIC	0.156	12.446	1	1993:1
	MAIC	0.156	12.446^{d}	1	1993:1
PEI	BIC	0.120^{c}	7.563^{a}	1	1993:1
	MAIC	0.142^{d}	10.391^{b}	4	1995:3
Quebec	BIC	0.191	18.334	1	1985:1
	MAIC	0.191	18.334	1	1985:1
Saskatechwan	BIC	0.121^{c}	7.481^{a}	2	1988:1
	MAIC	0.163	13.503^{d}	1	1986:2
Canada	BIC	0.144	10.530^{d}	2	1982:2
	MAIC	0.155	12.624^{d}	1	1980:4

^{a,b,c,} and ^{d} shows significant at 1%, 2.5%, 5%, and 10%.

Table 3: Results of Unit Root Tests, Crash model

		Infimum method			Supremum method		
Series	Criteria	MSB	P_T	k	MSB	P_T	k
Alberta	BIC	0.233	12.093	2	0.325	40.596	1
	MAIC	0.233	12.093	2	0.325	40.596	1
BC	BIC	0.231	11.895	1	0.239	25.394	1
	MAIC	0.231	11.895	1	0.239	25.394	1
Manitoba	BIC	0.202	9.153	3	0.355	47.266	1
	MAIC	0.282	18.198	4	0.327	40.231	6
NB	BIC	0.185	7.873	1	0.185	11.98	1
	MAIC	0.185	7.873	1	0.185^d	11.98	1
NF	BIC	0.183	7.813	3	0.216	17.035	4
	MAIC	0.205	9.847	4	0.216	17.035	4
NS	BIC	0.277	20.144	1	0.288	33.231	1
	MAIC	0.277	20.144	1	0.288	33.231	1
Ontario	BIC	0.175	5.689	2	0.175^d	6.837^d	2
	MAIC	0.177	6.166	1	0.225	11.378	1
PEI	BIC	0.231	11.725	1	0.247	18.738	1
	MAIC	0.265	15.298	1	0.362	40.254	5
Quebec	BIC	0.250	14.733	1	0.285	26.389	1
	MAIC	0.204	9.985	8	0.285	26.389	1
Saskatechwan	BIC	0.278	18.603	1	0.278	37.782	1
	MAIC	0.278	18.409	1	0.278	37.782	1
Canada	BIC	0.177	6.871	2	0.177	10.819	2
	MAIC	0.192	8.528	1	0.198	13.488	3

^{a,b,c,} and ^{d} shows significant at 1%, 2.5%, 5%, and 10%.

Table 4: Results of Unit Root Tests with 2 breaks

	Crash model				Broken trend and intercept			
	Statistic	k	T_{B1}	T_{B2}	Statistic	k	T_{B1}	T_{B2}
Alberta	-6.536^{c}	7	1981:3	1990:3	-6.153	7	1982:1	1991:1
BC	-6.586^{c}	8	1981:3	1987:1	-6.156	6	1981:3	2001:2
Manitoba	-6.069^{d}	3	1981:4	1990:1	-6.273	3	1981:4	1990:3
NB	-5.939	3	1981:3	1991:2	-6.216	7	1984:2	1991:2
NF	-6.133^{d}	7	1982:2	1991:4	-6.066	8	1984:2	1991:4
NS	-5.261	1	1981:3	1990:4	-5.699	6	1987:1	1997:2
Ontario	-5.645	1	1981:2	1990:2	-6.638^{d}	5	1981:3	1990:2
PEI	-4.197	0	1982:1	1990:3	-7.167^{c}	0	1986:3	1990:3
Quebec	-5.765	7	1981:3	1990:3	-6.433	7	1981:3	1990:2
Saskatechwan	-7.213^{c}	2	1982:1	1996:1	-6.998^{d}	2	1982:1	1993:4
Canada	-6.407^{c}	6	1981:3	1990:2	-7.155^{c}	6	1982:1	1990:2

^{a,b,c,}and ^{d} shows significant at 1%, 2.5%, 5%, and 10%.

Table 5: 90% confidence intervals for measures of unemployment rate persistence

	Asymptotic	Stock	Percentile	Percentile-t	Kilian	Grid-α	Grid-t
Alberta	(0.941, 1.006)	(0.966, 1.037)	(0.820, 0.969)	(0.977, 1.033)	(0.844, 1.003)	(0.983, 1.039)	(0.979, 1.034)
BC	(0.922, 1.004)	(0.955, 1.037)	(0.806, 0.963)	(0.964, 1.036)	(0.856, 1.014)	(0.965, 1.038)	(0.965, 1.036)
Manitoba	(0.904, 0.993)	(0.916, 1.034)	(0.796, 0.955)	(0.942, 1.024)	(0.883, 1.028)	(0.936, 1.037)	(0.938, 1.032)
NB	(0.836, 0.961)	(0.836, 1.026)	(0.741, 0.922)	(0.872, 0.997)	(0.850, 1.001)	(0.866, 1.028)	(0.868, 1.026)
NF	(0.798, 0.944)	(0.809, 1.023)	(0.718, 0.904)	(0.833, 0.978)	(0.813, 0.983)	(0.840, 1.023)	(0.840, 1.025)
NS	(0.897, 0.987)	(0.894, 1.032)	(0.803, 0.955)	(0.928, 1.016)	(0.871, 1.021)	(0.924, 1.035)	(0.927, 1.029)
Ontario	(0.921, 1.003)	(0.949, 1.036)	(0.801, 0.960)	(0.963, 1.036)	(0.871, 1.015)	(0.960, 1.037)	(0.960, 1.035)
PEI	(0.872, 0.989)	(0.911, 1.034)	(0.771, 0.943)	(0.916, 1.027)	(0.875, 1.033)	(0.911, 1.033)	(0.911, 1.035)
Quebec	(0.893, 0.983)	(0.883, 1.031)	(0.794, 0.951)	(0.926, 1.011)	(0.874, 1.026)	(0.919, 1.034)	(0.922, 1.028)
Saskatchewan	(0.909, 0.996)	(0.926, 1.035)	(0.797, 0.960)	(0.947, 1.027)	(0.869, 1.027)	(0.943, 1.037)	(0.944, 1.033)
Canada	(0.928, 1.001)	(0.944, 1.036)	(0.815, 0.966)	(0.963, 1.029)	(0.859, 1.013)	(0.962, 1.038)	(0.963, 1.032)

An intercept and a time trend are included in the model.

Table 6: 90% confidence intervals for measures of unemployment rate persistence

	Asymptotic	Stock	Percentile	Percentile-t	Kilian	Grid-α	Grid-t
Alberta	(0.944, 1.010)	(0.980, 1.038)	(0.863, 0.986)	(0.969, 1.029)	(0.903, 1.021)	(0.962, 1.031)	(0.963, 1.031)
BC	(0.936, 1.015)	(0.988, 1.039)	(0.854, 0.985)	(0.963, 1.036)	(0.906, 1.025)	(0.961, 1.032)	(0.961, 1.036)
Manitoba	(0.915, 1.003)	(0.950, 1.036)	(0.847, 0.976)	(0.939, 1.022)	(0.908, 1.028)	(0.934, 1.028)	(0.934, 1.030)
NB	(0.883, 0.994)	(0.924, 1.035)	(0.824, 0.962)	(0.908, 1.012)	(0.883, 1.016)	(0.905, 1.022)	(0.905, 1.029)
NF	(0.797, 0.939)	(0.789, 1.002)	(0.742, 0.910)	(0.819, 0.961)	(0.808, 0.963)	(0.840, 0.985)	(0.840, 0.979)
NS	(0.908, 0.998)	(0.935, 1.036)	(0.853, 0.972)	(0.930, 1.017)	(0.899, 1.021)	(0.928, 1.026)	(0.926, 1.028)
Ontario	(0.920, 1.003)	(0.950, 1.036)	(0.840, 0.976)	(0.946, 1.023)	(0.922, 1.037)	(0.940, 1.029)	(0.942, 1.029)
PEI	(0.872, 0.982)	(0.890, 1.032)	(0.816, 0.954)	(0.895, 0.9997)	(0.879, 1.006)	(0.891, 1.019)	(0.891, 1.022)
Quebec	(0.912, 1.001)	(0.942, 1.036)	(0.847, 0.973)	(0.937, 1.019)	(0.910, 1.027)	(0.933, 1.027)	(0.933, 1.029)
Saskatchewan	(0.906, 0.992)	(0.912, 1.034)	(0.849, 0.969)	(0.926, 1.008)	(0.904, 1.016)	(0.921, 1.025)	(0.923, 1.023)
Canada	(0.937, 1.009)	(0.973, 1.038)	(0.855, 0.987)	(0.963, 1.029)	(0.909, 1.027)	(0.956, 1.030)	(0.957, 1.031)

Only an intercept is included in the model.

Table 7: Bootstrap confidence intervals of the unemployment rates

Series	k	$\hat{P}$	$s(\hat{P})$	Grid-t Bootstrap	Percentile Bootstrap
Alberta	2	0.966	0.020	(0.961, 1.025)	(0.963, 1.022)
BC	2	0.955	0.024	(0.945, 1.027)	(0.948, 1.022)
Manitoba	1	0.949	0.027	(0.938, 1.032)	(0.941, 1.024)
NB	1	0.898	0.038	(0.868, 1.027)	(0.873, 0.993)
NF	4	0.887	0.046	(0.850, 1.036)	(0.858, 1.006)
NS	1	0.942	0.027	(0.925, 1.030)	(0.930, 1.015)
Ontario	2	0.941	0.021	(0.923, 1.013)	(0.926, 0.993)
PEI	1	0.930	0.035	(0.911, 1.035)	(0.916, 1.025)
Quebec	2	0.928	0.027	(0.906, 1.021)	(0.910, 0.999)
Saskatchewan	1	0.952	0.026	(0.944, 1.033)	(0.946, 1.027)
Canada	2	0.950	0.018	(0.935, 1.012)	(0.937, 0.995)

An intercept and a time trend are included in the model.

Table 8: Bootstrap confidence intervals of the unemployment rates

Series	k	$\hat{P}$	$s(\hat{P})$	Grid-t Bootstrap	Percentile Bootstrap
Alberta	2	0.969	0.020	(0.951, 1.021)	(0.955, 1.015)
BC	2	0.964	0.023	(0.944, 1.024)	(0.948, 1.021)
Manitoba	1	0.959	0.026	(0.936, 1.030)	(0.941, 1.022)
NB	1	0.938	0.034	(0.905, 1.029)	(0.910, 1.016)
NF	4	0.885	0.044	(0.832, 1.017)	(0.838, 0.981)
NS	1	0.953	0.027	(0.925, 1.028)	(0.930, 1.017)
Ontario	2	0.941	0.021	(0.916, 1.004)	(0.919, 0.986)
PEI	1	0.927	0.033	(0.892, 1.022)	(0.896, 1.001)
Quebec	2	0.945	0.026	(0.917, 1.019)	(0.920, 1.006)
Saskatchewan	1	0.949	0.026	(0.923, 1.023)	(0.926, 1.009)
Canada	2	0.955	0.018	(0.935, 1.008)	(0.937, 0.994)

Only an intercept is included in the model.

Table 9: Results from applying Bai and Perron approach

Statistics	Alberta	BC	Manitoba	NB	NF	NS	Ontario	PEI	Quebec	Saskatchewan	Canada
UD_{MAX}	30.840^a	37.138^a	18.399^a	15.939^b	32.441^a	35.406^a	25.801^a	35.124^a	34.491^a	35.730^a	38.750^a
WD_{MAX} 5%	38.840*	37.138*	27.616*	26.357*	33.718*	45.733*	38.441*	35.124*	48.731*	45.249*	45.560*
$supF_T(1)$	38.840^a	37.138^a	16.233^b	13.787^d	32.441^a	35.406^a	19.557^b	35.124^a	8.056	35.730^a	18.619^b
$supF_T(2)$	21.175^a	30.141^a	16.005^a	15.939^a	25.225^a	30.296^a	25.801^a	26.389^a	34.491^a	24.258^a	38.750^a
$supF_T(3)$	27.310^a	27.683^a	18.398^a	15.304^a	19.814^a	33.989^a	22.136^a	23.005^a	25.366^a	28.081^a	30.712^a
$supF_T(4)$	22.400^a	22.044^a	17.877^a	14.602^a	23.341^a	29.037^a	24.554^a	19.494^a	21.451^a	23.859^a	26.591^a
$supF_T(5)$	18.609^a	19.538^a	12.935^a	14.065^a	19.549^a	22.191^a	21.583^a	15.212^a	27.360^a	24.146^a	23.138^a
$supF_T(2\|1)$	10.074	16.210^c	15.408^c	16.827^b	16.954	18.096^b	30.548^a	12.139^d	20.172^b	10.167	55.667^a
$supF_T(3\|2)$	10.169	18.699^c	13.431	13.81	8.424	39.681^a	25.527^a	11.498	12.742	10.618	14.753
$supF_T(4\|3)$	10.791	9.514	15.408^d	12.75	8.424	10.775	22.118^b	6.409	12.056	8.187	11.871
$supF_T(5\|4)$	5.819	9.514		11.537	8.678		2.573		0		0
BIC	1	1	0	0	0	3	0	1	0	1	2
LWZ	0	0	0	0	0	0	0	0	0	0	0
Sequential	$1^{a,b,c,d}$	$1^{a,b,c}\ 3^d$	$0^a,\ 1^{b,c}\ 4^d$	$0^{a,b,c}2^d$	$1^{a,b,c,d}$	$1^a,\ 3^{b,c,d}$	$0^a,\ 3^{b,c,d}$	$1^{a,b,c,d}$	$0^{a,b,c,d}$	$1^{a,b,c,d}$	$0^a,\ 2^{b,c,d}$

^{a,b,c,}and ^{d} shows significant at 1%, 2.5%, 5%, and 10%.

* significant at 5% level of significance.

Table 10: Estimation results from Bai and Perron approach

	Alberta	BC	Manitoba	NB	NF	NS	Ontario	PEI	Quebec	SAS	Canada
μ_1	1.474	-1.746	0.708	3.888	2.258	5.203	3.280	3.910	13.588	1.598	2.286
	(1.143)	(-1.347)	(2.492)	(2.739)	(2.436)	(3.823)	(2.785)	(3.693)	(8.214)	(2.171)	(2.613)
β_1	-0.022	0.049	0.006	-0.030	0.005	-0.028	0.006	0.032	-0.040	0.003	0.001
	(-1.235)	(2.383)	(1.642)	(-1.580)	(0.689)	(-1.773)	(0.566)	(3.018)	(-1.909)	(0.312)	(0.167)
ρ_{11}	0.519	1.671	0.886	0.698	0.697	0.516	1.146	0.597	1.165	0.628	1.342
	(1.446)	(8.409)	(18.500)	(6.096)	(6.057)	(3.727)	(4.958)	(5.459)	(12.758)	(3.612)	(7.621)
ρ_{21}	0.180	-0.514			0.267		-0.635		-0.238		-0.634
	(0.498)	(-1.801)			(1.905)		(-2.800)		(-2.640)		(-3.611)
ρ_{31}					0.192						
					(1.349)						
ρ_{41}					-0.300						
					(-2.606)						
μ_2	2.109	1.547	2.387	3.388	20.744	2.755	5.584	12.176		2.775	6.528
	(5.667)	(2.082)	(1.672)	(3.595)	(4.691)	(3.567)	(5.473)	(5.205)		(6.098)	(6.858)
β_2	-0.013	-0.008	-0.011	-0.014	-0.076	-0.041	-0.073	-0.063		-0.011	-0.069
	(-5.292)	(-1.945)	(-1.217)	(-3.512)	(-4.541)	(-4.673)	(-5.182)	(-5.032)		(-5.598)	(-6.468)
ρ_{12}	1.073	0.958	0.764	0.799	0.312	0.897	1.241	0.528		0.710	1.059
	(11.572)	(10.008)	(7.747)	(13.528)	(2.136)	(17.063)	(10.638)	(5.749)		(14.28)	(8.479)
ρ_{22}	-0.233	-0.063			0.063		-0.602				-0.446
	(2.751)	(-0.673)			(0.422)		(-6.256)				(-4.619)
ρ_{32}					0.197						
					(1.311)						
ρ_{42}					-0.365						
					(-2.673)						
μ_3				20.359		-4.723	4.014				2.637
				(3.495)		(-1.461)	(6.040)				(4.620)
β_3				-0.092		0.236	-0.030				-0.014
				(-2.904)		(1.59)	(-5.548)				(-4.465)
ρ_{13}				-0.013		0.199	1.030				1.186
				(-0.048)		(0.391)	(8.510)				(10.331)
ρ_{23}							-0.224				-0.345
							(-2.087)				(-3.305)
μ_4						5.002	5.801				
						(2.796)	(2.446)				
β_4						-0.025	-0.018				
						(-2.405)	(-1.079)				
ρ_{14}						0.741	0.550				
						(9.411)	(1.385)				
ρ_{24}							-0.108				
							(-0.303)				
$\bar{R}^2$	0.965	0.952	0.926	0.891	0.815	0.943	0.967	0.897	0.922	0.937	0.976
F	403.420^a	285.613^a	243.32^a	107.569^a	42.728^a	161.275^a	213.197^a	170.776^a	458.4422^a	289.125^a	390.785^a

^{a,b,c,}and ^{d} shows significant at 1%, 2.5%, 5%, and 10%.

t-statistics are shown in paranthesis.

Table 11: The estimated break points and their 90% confidence intervals

	Alberta	BC	Manitoba	NB	NF	NS	Ontario	PEI	SAS	Canada
T_{B1}	1981:1	1982:1	1993:4	1981:1	1990:4	1981:1	1981:3	1990:2	1981:4	1981:3
C.I.	1980:3-1981:2	1981:3-1982:4	1993:2-1995:4	1980:4-1982:3	1990:3-1991:1	1980:3-1981:2	1981:1-1981:4	1989:3-1990:3	1981:1-1982:1	1981:1-198
T_{B2}				2000:3		1988:4	1989:4			1989:4
C.I.				1998:4-2000:4		1988:2-1989:1	1989:2-1990:1			1989:2-199
T_{B3}						1993:3	2000:4			
C.I.						1993:1-1993:4	2000:2-2001:2			

Figures

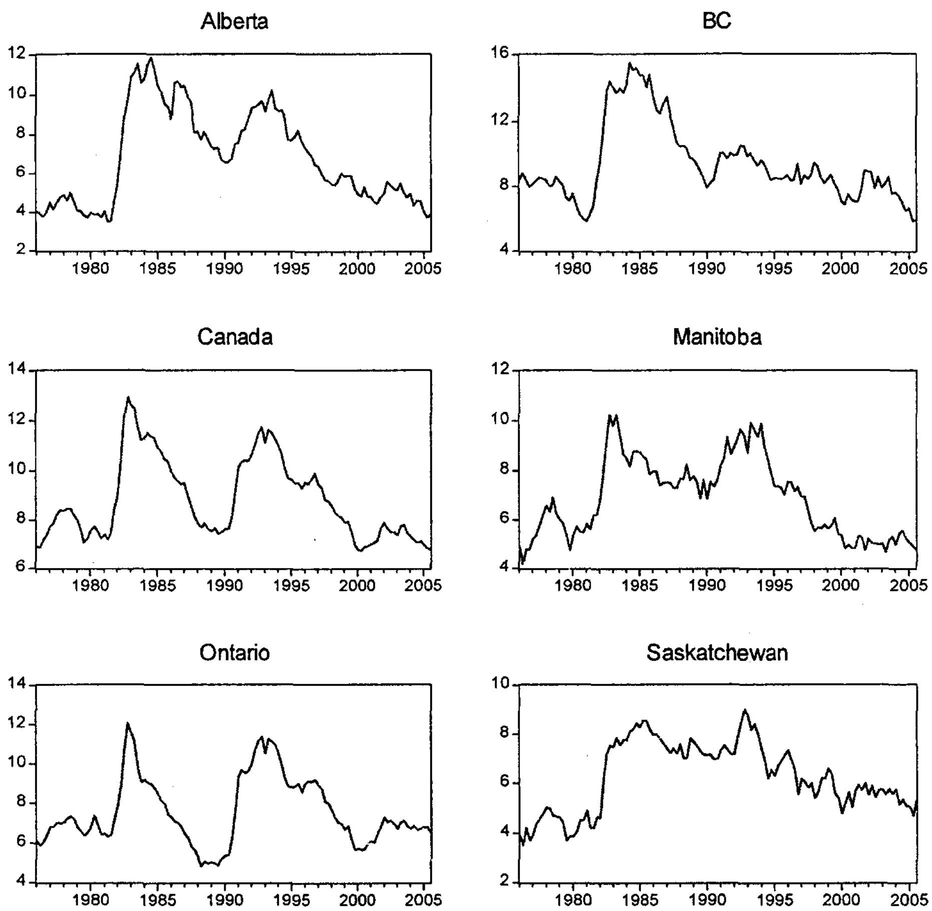

Figure 1: Unemployment rates of Canada and the Canadian provinces.

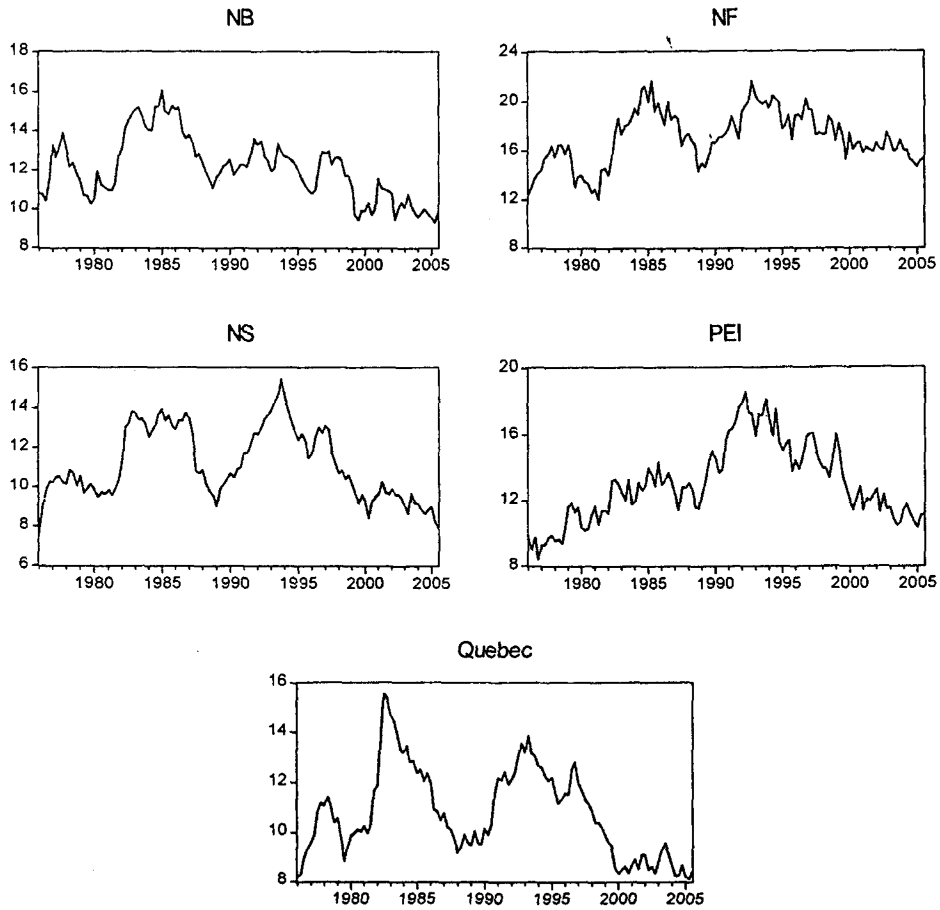

Figure 2: Unemployment rates of the Canadian provinces.