Bayesian Semiparametric Regression for Autoregressive Models with Possible Unit Roots

Ricardo Gonçalves Silva*
Instituto de Ciências Matemáticas e de Computação - ICMC
Departamento de Ciências de Computação e Estatística - SCE
Universidade de São Paulo - USP
Ph.D. Student

26th March 2004

Abstract

In this paper we consider bayesian semiparametric regression within the generalized linear model framework. Specifically, we study a class of autoregressive time series where the time trend is incorporated in a nonparametrically way. Estimation and inference where performed through Markov Chain Monte Carlo simulation techniques. Main results show that treating the time trend nonparametrically possible model misspecification and biased results from structural break issues are solved. Empirical applications are conducted using the extended Nelson and Plosser benchmark time series.
Keywords: Bayesian Inference, Unit Root, Structural Break, MCMC, Semiparametric Regression, Nonlinear Time Trend, Random Walk Prior, Macroeconomic Time Series.
J.E.L. Classification: C11, C13, C14, C22 and C44.

1 Introduction

The major issue on estimating and testing autoregressive time series are related to *specification*. The autoregressive order, the constancy of mean (unit root) and variance (heteroscedasticity), the inclusion of fitted constant and time trend and possible structural break are all related to *model's specification*. In each case a different methodology should be carefully chosen and applied in order to get consistent estimates of model's parameters.

*Corresponding Address: Rua Major Jose Ignacio 3985, São Carlos-SP. Zip Code: 13569-010, Brazil. E-mail: ricardo@icmc.usp.br

The problems point above should be considered from two point of views: the economic and the statistical. From the last, as pointed by White (1995), we need to try to ask under what conditions standard parametric statistical techniques such as maximum likelihood can be meaningfully applied to estimate the parameters of a misspecified model. In time series analysis, the answer for this question is directly linked to the concept of *Dynamic Misspecification*. A simple example used by (White, 1995, p. 57) concerns on adjusting a first order autoregressive model when the true process is generated by second order. Since the dependence on the second lag has been ignored, a dynamic misspecification thereby committed. There are circumstances in which dynamic misspecification not need destroy a possible economic interpretation for particular parameters, as in Robinson (1982), Gourieroux, Monfort, and Trognon (1985) and Poirier and Rudd (1988). Their results, show that consistent estimation of parameters of interest is possible, despite such dynamic misspecification.

More specifically in time series context, many other examples arises. First, the so-called block bootstrap, widely used recently in both stationary and non-stationary time series, e. g. Paparoditis and Politis (2003) and Lahiri (1999), easily generates dynamic misspecification, as showed by Corradi and Swanson (2002). Second, in performing classical hypothesis testing Perron and Campbell (1991) pointed out that a non rejection of the unit root hypothesis may be due to misspecification of the deterministic components included as regressors. The authors ague that *...care must be exercised in choosing the appropriate determinist regressor to include to have reasonable power properties.* (Perron and Campbell, 1991, p. 12).

In this paper we consider a possible solution for the problem point out above. The key is to observe that polynomials time trends of any order, even in the presence of structural break, constitutes a very strong assumption for long macroeconomic time series. We propose to estimate the standard regression model imposing a nonlinear time trend specification. As empirical application we apply the proposed nonlinear approach for the extended Nelson and Plosser dataset using a bayesian technique to overcome main difficulties in models estimation and inference.

The rest of the paper is organized as follows. Section 2 presents and discuss the Bayesian Structured Additive Regression with nonlinear components. Section 3 presents the priors densities we shall use throughout the paper. Section 4 concerns on the inference and implementation of Markov Chain Monte Carlo simulation techniques and the final section concludes. All graphics as well as the computational results were obtained using the program *Bayes X*; for details, see Brezger, Kneib, and Lang (2003).

2 Bayesian Structured Additive Regression

Generalized linear models (e.g. Fahrmeir and Tutz (2001)) assume that, given covariates $\mathbf{z}$ and unknown parameters γ, the distribution of the response variable y belongs to an exponential family, i.e.

$$p(y|\mathbf{z}) = \exp\left(\frac{y\theta - b(\theta)}{\phi}\right) c(y, \phi) \tag{1}$$

where $b(\cdot)$, $c(\cdot)$, θ and ϕ determines the respective distribution. A list of the most common distributions and their specific parameters can be found, e.g. in (Fahrmeir and Tutz, 2001, p. 21). The mean $\mu = E(y|\mathbf{z}, \gamma)$ is linked to a linear predictor η by

$$\mu = h(\eta) \qquad \eta = \mathbf{z}'\gamma \tag{2}$$

Here h is a known response function, and γ are usual unknown regression parameters. In the remainder of this paper we will consider that the response variable belongs to a Gaussian family, given by

$$p(y|\mathbf{z}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{1}{2\sigma^2}(\epsilon_t)^2\right\} \tag{3}$$

Considering a Gaussian time series, y_t, observed for $t = 1, 2, \ldots, T$ characterized by a first order autoregressive (AR(1)) process with fitted constant and time trend

$$y_t = c + \beta t + \rho y_{t-1} + \epsilon_t \tag{4}$$

the density function model (5) takes the form

$$p(y|\mathbf{z}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{1}{2\sigma^2}\left(y_t - c - \beta t - \rho y_{t-1}\right)^2\right\} \tag{5}$$

In this way, the regression parameters γ becomes

$$\gamma = (c, \beta, \rho)' \tag{6}$$

and the vector of covariates $\mathbf{z}$ is then given by

$$\mathbf{z} = (1, t, y_{t-1}) \tag{7}$$

In time series context we should consider the temporal dependency among observations. To overcome the difficulties, we replace the strictly linear predictor in (2) by a structured additive predictor

$$\eta_t = f_{Trend}(t) + \mathbf{z}'\boldsymbol{\gamma} \tag{8}$$

where the function $f_{Trend}(t)$ is a nonlinear function for the time trend.

At first sight it may look strange to use one general notation for nonlinear functions of continuous covariates and temporally correlated effects as in (8). However, the unified treatment of the different components in the model has several advantages. First, since we adopt a Bayesian perspective it is generally not necessary impose any restrictions on parameters space, because in the Bayesian approach all unknown parameters are assumed to be random. Second, as we will see in next section, the priors for smooth functions and serially correlated effects can be cast into a general form. Finally, the general form of the priors also allows rather general and unified estimation procedures, named, Markov Chain Monte Carlo technique. As a side effect the implementation and description of these procedures is considerably facilitated.

3 Prior Assumptions

For Bayesian inference, the unknown function $f_{Trend}(t)$ in (8), more exactly corresponding to a function evaluation, and the fixed parameters γ are considered as random variables and must be supplemented by appropriate prior assumptions. In the absence of any prior knowledge diffuse priors are the appropriate choice for fixed effects parameters, i.e.

$$\gamma_j \propto constant \tag{9}$$

Priors for the unknown function $f_{Trend}(t)$ depends on the type of the covariates and on prior beliefs about the smoothness of $f_{Trend}(t)$. In the following we will be able to express the function evaluation $f_{Trend}(t)$ as the vector product of a design vector X and a unknown parameter δ, i.e.

$$f = XB \tag{10}$$

Then, we obtain the predictor (8) in matrix notation as

$$\eta = \delta X + \mathbf{U}\gamma \tag{11}$$

where U corresponds to the usual design matrix for model's parameters.

A prior for the function $f_{Trend}(t)$ is now defined by specifying a suitable design vector X and a prior distribution for the vector δ of unknown parameters. The general form of the prior for δ is given by

$$p(\delta|\sigma^2) \propto \frac{1}{(\sigma^2)^{rank(K)/2}} \exp\left(-\frac{1}{2\sigma^2}\delta' K \delta\right) \tag{12}$$

where K is a *penalty vector* that shrinks parameters towards zero or penalizes too abrupt jumps between neighboring parameters. In most cases K will be rank deficient and therefore the prior for δ is partially improper. The variance parameter σ^2 is equivalent to the inverse smoothing parameter in a frequentist approach and controls the trade off between flexibility and smoothness. For more detail on this type of prior distributions see Brezger, Kneib, and Lang (2003) and Lang, Fronk, and Fahrmeir (2002).

3.1 Priors for the Time Trend

Several alternatives have been proposed for specifying smoothness priors for continuous covariates or time trends. These are random walk priors or more generally, autoregressive priors (see Fahrmeir and Lang (2001a) and Fahrmeir and Lang (2001b)), Bayesian P-splines (Lang and Brezger (2003)) and Bayesian smoothing splines (Hastie and Tibshirani (2000)). In this paper we will use the class of random walk priors to model the time trend.

A common approach in dynamic or state space models is to estimate one parameter δ_m for each distinct regressor (m) and penalize too abrupt jumps between successive parameters using random walk priors. Most commonly used are first or second order random walk models

$$\delta_m = \delta_{m-1} + u_m \quad \text{or} \quad \delta_m = 2\delta_{m-1} - \delta_{m-2} + u_m \tag{13}$$

with Gaussian errors $u_m \sim N(0, \sigma^2)$ and diffuse priors for the $\delta_i \propto constant$, $i = 1, 2$ for initial values, respectively.

Both specifications act as smoothness priors penalizing too rough functions $f_{Trend}(t)$. A first order random walk penalizes too abrupt jumps $\delta_m - \delta_{m-1}$ between successive states and a second order random walk penalizes deviations from the linear trend $2\delta_{m-1} - \delta_{m-2}$. In dealing with economic time series, we should consider both specifications, controlling then for unit roots and possible structural breaks.

The joint distribution of the regression parameters δ is easily computed as a product of conditional densities defined by (13) and can be brought into the general form (12). The penalty matrix is of the form $K = D'D$ where D is a first or second order difference matrix. For example, for a random walk of first order the penalty matrix is given by:

$$K = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix} \tag{14}$$

The design matrix X is a simple $0/1$ matrix where the number of columns is equal to the number of parameters, respectively the number of distinct covariate values.

Considering the prior densities setup presented above we are able now to conduct some empirical applications. Next section briefly review some basic concepts of Bayesian inference and present the main results.

4 Inference and Empirical Results

For full Bayesian inference, the unknown variance parameters σ^2 are also considered as random and estimated simultaneously with the unknown regression parameters δ . Therefore, hyperpriors are assigned to the variances σ^2 in a further stage of the hierarchy by highly dispersed (but proper) inverse Gamma priors $p(\sigma^2) \sim IG(a, b)$. The probability density is given by

$$\sigma^2 \propto (\sigma^2)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right) \tag{15}$$

The prior for σ^2 must not be diffuse in order to obtain a proper posterior for δ. A common choice for the hyperparameters are small values for a and b, e.g. $a = b = 0.001$ which is also the default values we shall use.

In some situations, the estimated nonlinear function f_{Trend} may considerably depend on the particular choice of hyperparameters a and b[1]. In that sense, the variation of hyperparameters can be used as a tool for model diagnostics.

Bayesian inference is based on the posterior of the model which is given by

$$p(\delta, \sigma^2, \gamma | y) \propto \mathcal{L}(y, \delta, \gamma) \left(p\left(\delta | \sigma^2\right) p(\sigma^2)\right) \tag{16}$$

where $\mathcal{L}(\cdot)$ denotes the likelihood which is the product of individual likelihood con-

[1]This may be the case for very low signal to noise ratios or/and small sample sizes. It is therefore highly recommended to estimate all models under consideration using a (small) number of different choices for a and b to assess the dependence of results on minor changes in the model assumptions.

tributions.

In many practical situations (as is the case here) the posterior distribution is numerically intractable. A common technique to overcome these problems are Markov Chain Monte Carlo (MCMC) simulation methods that have become very popular recently. MCMC methods allow the drawing of random numbers from the numerically intractable posterior distribution and in this way the estimation of characteristics of the posterior like means, standard deviations or quantiles via their empirical analogues.

The main idea is very simple. Instead of drawing directly from the posterior (which is impossible in most cases anyway) a Markov chain is created, whose iterations of the transition kernel converge to the posterior. In this way a sample of dependent random numbers of the posterior is obtained. As a rule, the first part of the sample is discarded to take into account the time the algorithm needs for convergence to the posterior. This part is known as burn-in period.

As we suppose at the beginning, the distribution of the response variable is Gaussian, i.e. $y_t|\eta_t, \sigma^2 \sim N(\eta_t, \tau^2/c_i), i = 1, \ldots, T$. In this case full conditionals for fixed parameters as well as nonlinear functions f_j are multivariate Gaussian. Thus a Gibbs sampler can be used where posterior samples are drawn directly from the multivariate Gaussian distributions. The full conditional $\gamma|\cdot$ for fixed parameters with diffuse priors is Gaussian with mean

$$E(\gamma|\cdot) = (\mathbf{U}'\mathbf{C}\mathbf{U})^{-1}\mathbf{U}'\mathbf{C}(y - \widetilde{\eta}) \tag{17}$$

and covariance matrix

$$Cov(\gamma|\cdot) = \tau^2(\mathbf{U}'\mathbf{C}\mathbf{U})^{-1} \tag{18}$$

where $\mathbf{U}$ is the design matrix of fixed parameters and $\widetilde{\eta}$ is the part of the additive predictor associated with the other effects in the model (for example nonparametric terms). Similarly, the full conditional for the regression coefficients δ of a function $f_{Trend}(t)$ is Gaussian with mean

$$\mu = E(\delta|\cdot) = \left(\frac{1}{\tau^2}\mathbf{X}'\mathbf{C}\mathbf{X} + \frac{1}{\sigma^2}K\right)^{-1} \frac{1}{\tau^2}\mathbf{X}'\mathbf{C}(y - \widetilde{\eta}) \tag{19}$$

and covariance matrix

$$Cov(\delta|\cdot) = P^{-1}\tau^2(\mathbf{U}'\mathbf{C}\mathbf{U})^{-1} \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{C}\mathbf{X} + \frac{1}{\tau^2}K\right)^{-1} \tag{20}$$

Although the full conditional is Gaussian, drawing random samples in an efficient way is not trivial, since linear equation systems with a high dimensional precision matrix P must be solved in every iteration of the MCMC scheme. Following Rue (2001), drawing random numbers from $p(\delta|\cdot)$ is as follows: We first compute the Cholesky decomposition $P = LL'$. We proceed by solving $L'\delta = z$, where z is a vector of independent standard Gaussians. It follows that $\delta \sim N(0, P^{-1})$. We then compute the mean μ by solving

$$P\mu = \frac{1}{\tau^2}\mathbf{X}'\mathbf{C}(y - \widetilde{\eta})$$

This is achieved by first solving

$$L\nu = \frac{1}{\tau^2}\mathbf{X}'\mathbf{C}(y - \widetilde{\eta})$$

by forward substitution followed by backward substitution $L'\mu = \nu$. Finally, adding μ to the previously simulated δ yields $\delta \sim N(\mu, P^{-1})$.

In all cases, the posterior precision matrix P can be brought into a band matrix like structure with bandsize depending on the prior. Random samples from the full conditional can now be drawn in a very efficient way using Cholesky decompositions for band matrices or band matrix like matrices. In our implementation we use the envelope method for band matrix like matrices as described in George and Liu (1981). The full conditionals for the variance parameters σ^2 and τ^2 are all inverse Gamma distributions with parameters

$$a' = a + \frac{rank(K)}{2} \quad \text{and} \quad b' = b + \frac{1}{2}\delta' K \delta \tag{21}$$

for σ^2 . For τ^2 we obtain

$$a'_\sigma = a_\sigma + \frac{T}{2} \quad \text{and} \quad b'_\sigma = b_\sigma + \frac{1}{2}\epsilon'\epsilon \tag{22}$$

where ϵ is the usual vector of residuals.

We note that the response variable is standardized prior to estimation in order to avoid numerical problems with too large or too small values of the response. All results are, however, retransformed into the original scale.

The sampling scheme used can be summarized in the following four steps:

1 Initialization: Compute the posterior mode for δ and γ given fixed (usually small) smoothing parameters $\lambda = \tau^2/\sigma^2$, using $\lambda = 0.1$. The mode is computed via backfitting. Use the posterior mode estimates as the current state δ^{mc},

$(\sigma^2)^{mc}$, $\gamma^{mc}$of the chain.

2 Update regression parameters γ: Update regression parameters γ by drawing from the Gaussian full conditional with mean and covariance in (17) and (18).

3 Update regression parameter δ: Update δ by drawing from the Gaussian full conditionals with mean and covariance matrix given in (19) and (20).

4 Update variance parameter σ^2 and τ^2: Update variance parameters by drawing from inverse gamma full conditionals with parameters given in (21) and (22).

Now we are able to present the an empirical application. Results presented in table (1) are a comparison of full bayesian estimations, using a flat priori for all regressor and a semiparametric bayesian approach, where a random walk prior of first or second order (RW1 or RW2) is used for the time trend. The estimation used a Markov Chain with 330.000 steps, with 30.000 discarded at the burn-in period. All chains converged and show none degree of correlation[2]

Let's now consider the unit root issue. We should note that, the bayesian criteria to test the presence of a unit root in a time series is the inclusion of the unity inside the posterior credible interval[3]. The first point we observe is that when time trend is treated parametrically, the results are very close to that obtained using classical procedures, rejecting the null hypotheses of unit root only for two time series, unemployment rate and velocity of money. Nelson and Plosser (1982), in their seminal work, found that all of the fourteen macroeconomic time series, except the unemployment rate, were characterized by stochastic nonstationarity. Using a different testing procedure, Perron (1988), found that his results strongly support the conclusion reached by Nelson and Plosser. Even when compared with another bayesian procedures, like the ones presented by Phillips (1991a), Phillips (1991b) and Schotman and van Dijk (1991), our first results are not an exception.

However, when time trend is treated nonparametrically, the results are striking. The unit root hypothesis is now maintained for eight time series: real and nominal GNP, consumer price and production indexes, GNP deflator, money stock, nominal wages and S&P 500. All other time series remain stationary and, most important, the autoregressive coefficient are lower in all cases.

[2]Given the high number of Markov Chains used, these results are available upon request.

[3]Exists a growing literature about Bayesian unit root tests. The discussion was started by Sims (1988) and Sims and Uhlig (1988/1991) and become very controversial before the publication of Phillips (1991a) and Phillips (1991b). Discuss this literature is beyond the scope of this paper, however, the procedure used here is also considered as a possible alternative for unit root testing in all the paper cited above.

Considering the nonstationary time series above, a graphical analyses presented at the end of the paper, shows that the firsts three clearly presents a breaking trend around the Seventeen's, while the remaining ones shows a high degree of nonlinearity on the estimated time trend, indicating the presence of possible multiple breaks. A recent related work was presented by Gonzlez-Rivera and Dahl (2003) provides further evidence on the existence of a nonlinear component in the quarterly growth rate of the US real GNP. All these results are according with the standard findings presented in Perron (1988), Zivot and Andrews (1992) and Christiano (1992) and are related to the presence of one or more structural break on the data generating process.

Finally, we should note that some of the stationary time series, as Unemployment and Bonds Yield, presents the *crash pattern (Perron (1988))* around the 1929-1932 years, also pointed by the works cited above and in the references therein.

5 Concluding Comments

This paper presents a Bayesian approach for semiparametric modeling of time series. Our results demonstrate that the approach is a useful and flexible tool for estimating realistically complex models. Empirical findings presented here receive strong support on the ongoing literature and show that in treating the time trend nonparametrically, possible model's misspecification, concerning on unit roots and structural break among others, are solved. It is not the message invalidate any parametric model but instead, indicate that more research should be considered if the focus is determinate the autoregressive order of any particular time series of interest.

There are many directions under which the methodology used here should be extended in future research. The basic ones are include additional lags, quadratic and possible cubic nonlinear time trends and another exponential families, like Student-t and generalized error distribution (Exponential Power). All this features can be successful implemented without excessive effort within the Bayesian methodology through Markov Chain Monte Carlo (MCMC) techniques.

	Full Bayesian				Semiparametric Bayesian			
Variable	Trend Prior	ρ	*Std*	5% Credible Interval	Trend Prior	ρ	*Std*	5% Credible Interval
Real GNPPC	Diffuse	0.9454	0.0343	[0.8773 ; 1.0131]	RW1	0.5937	0.1217	[0.3300 ; 0.8290]
Real GNP	Diffuse	1.0080	0.0155	[0.9777 ; 1.0387]	RW1	0.8771	0.0760	[0.7157 ; 1.0101]
Nonimal GNP	Diffuse	1.0454	0.0040	[1.0376 ; 1.0535]	RW1	1.0306	0.0214	[0.9881 ; 1.0717]
CPI	Diffuse	1.0239	0.0075	[1.0090 ; 1.0389]	RW1	0.9590	0.0395	[0.8776 ; 1.0333]
Industrial Production	Diffuse	1.0016	0.0249	[0.9521 ; 1.0507]	RW1	0.7722	0.1659	[0.4096 ; 1.0235]
GNP Deflator	Diffuse	1.0142	0.0074	[0.9999 ; 1.0287]	RW1	0.9698	0.0329	[0.9043 ; 1.0327]
Interest Rate	Diffuse	0.9448	0.0356	[0.8733 ; 1.0156]	RW2	0.5072	0.1690	[0.1204 ; 0.7864]
Money Stock	Diffuse	1.0242	0.0069	[1.0102 ; 1.0376]	RW1	1.0089	0.0326	[0.9458 ; 1.0739]
Velocity of Money	Diffuse	0.9006	0.0442	[0.8142 ; 0.9882]	RW2	0.6215	0.0932	[0.4402 ; 0.8061]
Real Wages	Diffuse	0.9949	0.0275	[0.9407 ; 1.0504]	RW1	0.5725	0.1174	[0.3509 ; 0.8075]
Nominal Wages	Diffuse	1.0171	0.0066	[1.0040 ; 1.0296]	RW1	0.9933	0.0314	[0.9251 ; 1.0510]
Common Stock Prices	Diffuse	1.2199	0.0220	[1.1758 ; 1.2613]	RW1	1.2736	0.0538	[1.1710 ; 1.3819]
Unemployment Rate	Diffuse	0.8789	0.0524	[0.7767 ; 0.9822]	RW1	0.3556	0.1222	[0.1345 ; 0.6172]
Bond Yield	Diffuse	0.9745	0.0261	[0.9228 ; 1.0259]	RW2	0.4591	0.1632	[0.0813 ; 0.7290]

Table 1: Bayesian Semiparametric Estimation for Extended Nelson & Plosser Data Set.

References

BREZGER, A., T. KNEIB, AND S. LANG (2003): *BayesX: Analysing Bayesian structured additive regression models*.

CHRISTIANO, L. (1992): "Searching for a Break in GNP," *Journal of Business and Economic Statistics*, 10, 237–250.

CORRADI, V., AND N. N. SWANSON (2002): "Bootstrap Conditional Distribution Tests Under Dynamic Misspecification," *Journal of Econometrics*, Under Revision.

FAHRMEIR, L., AND S. LANG (2001a): "Bayesian Inference for Generalized Additive Mixed Models Based on Markov Random Field Priors," *Journal of the Royal Statistical Society C*, 50, 201–220.

——— (2001b): "Bayesian Semiparametric Regression Analysis of Multicategorical Time-Space Data," *Annals of the Institute of Statistical Mathematics*, 53, 10–30.

FAHRMEIR, L., AND G. TUTZ (2001): *Multivariate Statistical Modelling based on Generalized Linear Models*. Springer-Verlag, New York.

GEORGE, A., AND J. W. LIU (1981): *Computer Solution of Large Sparse Positive Definite Systems*, Series in computational mathematics. PrenticeHall, New Jersey.

GONZLEZ-RIVERA, G., AND C. DAHL (2003): "Identifying Nonlinear Components by Random Fields in the US GNP growth. Implications for the Shape of the Business Cycle," *Studies on Nonlinear Dynamics and Econometrics*, 7(1).

GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1985): "A General Approach to Serial Correlation," *Econometric Theory*, 1, 315–340.

HASTIE, T., AND R. TIBSHIRANI (2000): "Bayesian Backfitting," *Statistical Science*, 15, 193–223.

LAHIRI, S. (1999): "Theoretical Comparisons of Block Bootstrap Methods," *Annals of Statistics*, 27, 386–404.

LANG, S., AND A. BREZGER (2003): "Bayesian Splines," *Journal of Computational and Graphical Statistics*, Forthcoming.

LANG, S., E. M. FRONK, AND L. FAHRMEIR (2002): "Function estimation with locally adaptive dynamic models," *Computational Statistics*, 17, 479–500.

NELSON, C., AND C. PLOSSER (1982): "Trends ans Random Walks in Macroeconomic Time series.," *Journal of Monetary economics*, 10, 139–162.

PAPARODITIS, E., AND D. N. POLITIS (2003): "Residual-Based Block Bootstrap for Unit Root Testing," *Econometrica*, 71(3), 813–855.

PERRON, P. (1988): "Trends and Random Walks in Macroeconomic Time Series Further Evidence from a New Approach," *Journal of Economic Dynamics and Control*, 12, 297–332.

PERRON, P., AND J. Y. CAMPBELL (1991): *Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots*vol. 6. MIT Press, Cambridge: MA.

PHILLIPS, P. C. B. (1991a): "Bayesian Rootes and Unit Roots: De Rebus Prioribus Semper est Disputandum," *Journal of Applied Econometrics*, 6, 435–473.

——— (1991b): "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Treds.," *Journal of Applied Econometrics*, 6, 333–364.

POIRIER, D. J., AND P. A. RUDD (1988): "Probit with Dependent Observations," *Review of Economic Studies*, 54, 593–614.

ROBINSON, P. M. (1982): "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables," *Econometrica*, 50, 27–40.

RUE, H. (2001): "Fast Sampling of Gaussian Markov Random Fields with Applications," *Journal of the Royal Statistical Society B*, 63, 325–338.

SCHOTMAN, P. C., AND H. K. VAN DIJK (1991): "On Bayesian Routes to Unit Roots," *Journal of Applied Econometrics*, 6, 387–401.

SIMS, C. (1988): "Bayesian Scepticism on Unit Root Econometrics," *Journal of Economics Dynamics and Control*, 12, 436–474.

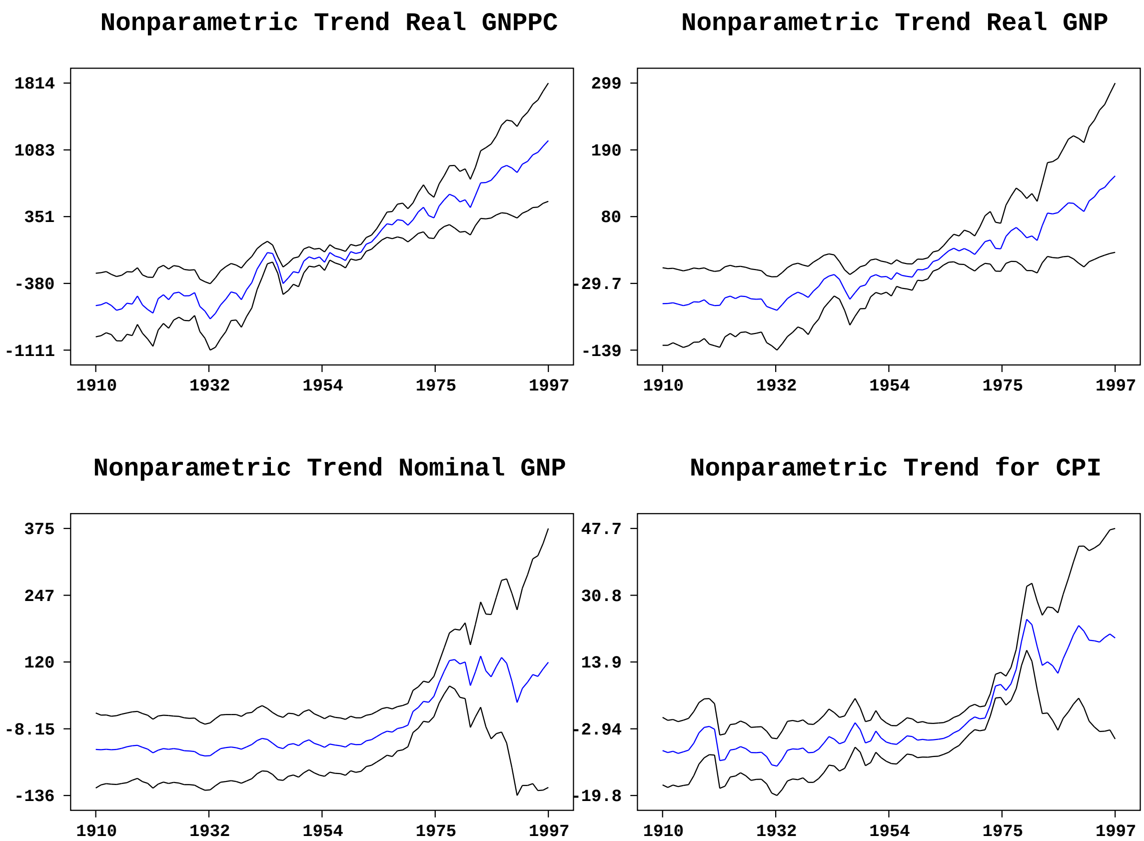
SIMS, C., AND H. UHLIG (1988/1991): "Understanding Unit Rooters: An Helicopter Tour," *Econometrica*, 59(6), 1591–1599.

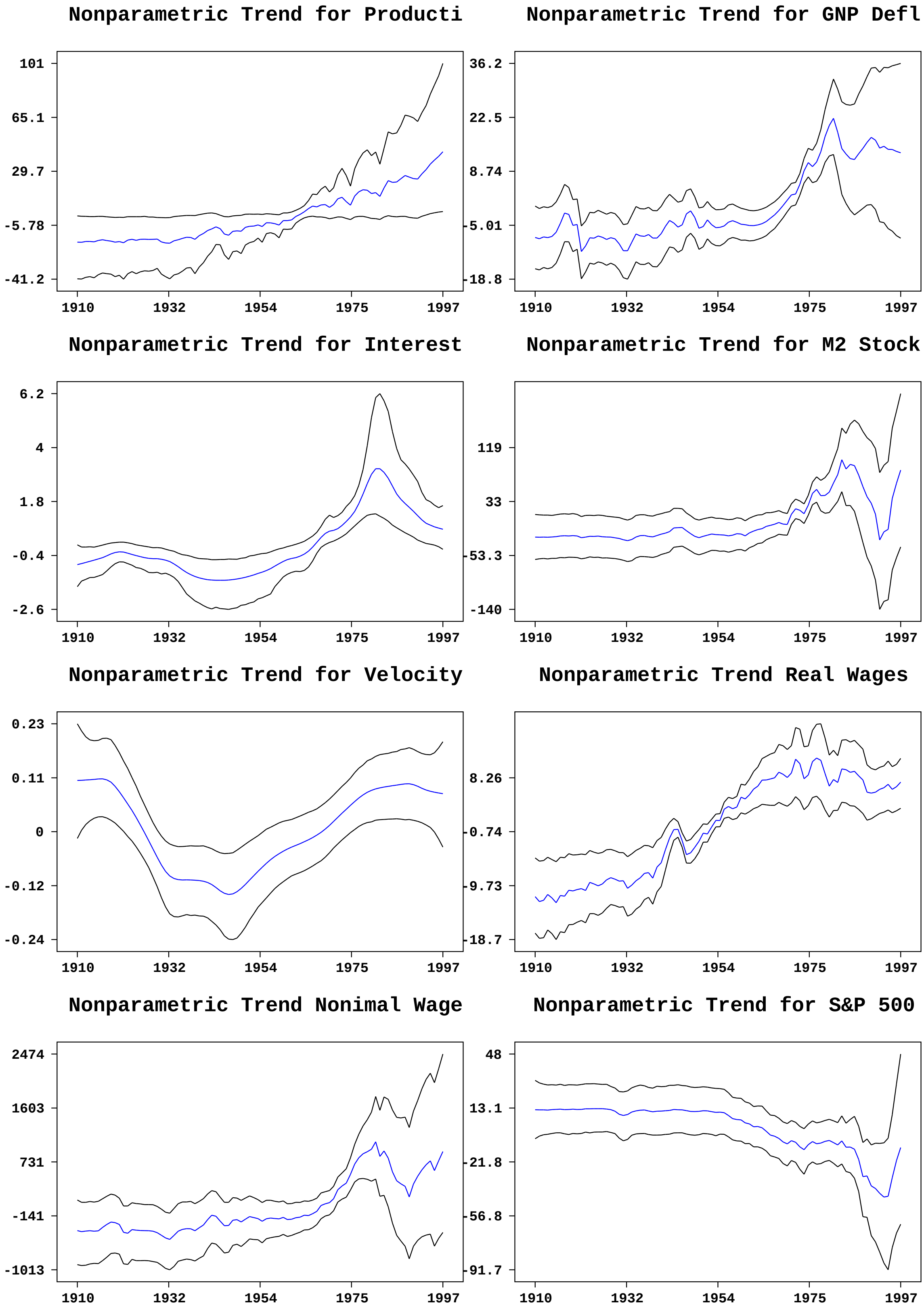
WHITE, H. (1995): *Estimation, Inference and Specification Analysis*, Econometric Society Monographs. Cambridge University Press, New York, NY.

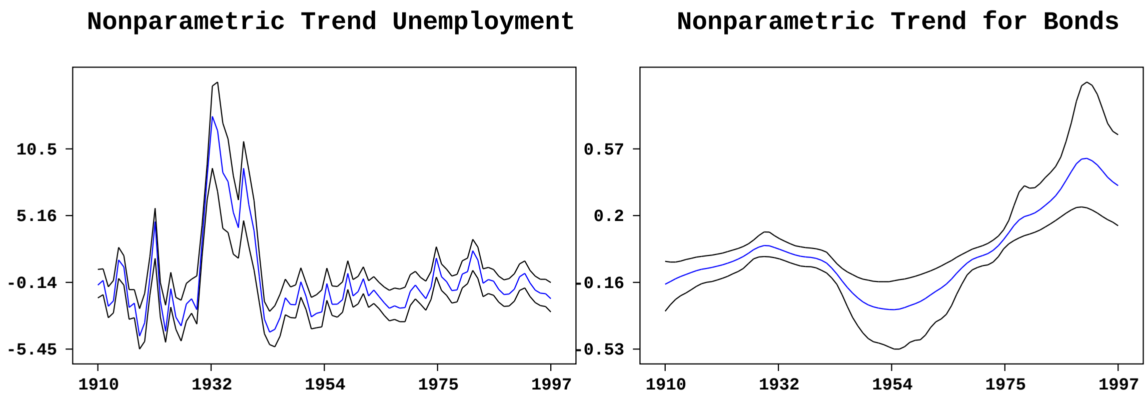
ZIVOT, E., AND D. ANDREWS (1992): "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis.," *Journal of Business and Economic Statistics*, 10(3), 251–270.

A Graphs of Nonparametric Trend Estimation

Nonlinear effect of *time trend* are presented for all the 14 time series on the graphics below. Shown are the posterior means(blue line), together with 95% pointwise credible intervals, (black lines). Note that all but three time series have it's time trend estimated using a first order random walk prior (RW1). The exception are interest rate, velocity of money and bond yield, which we estimated using a second order prior (RW2). As pointed in subsection 3.1, this was done because the deviation of the observed time realizations from the time trend was very rough.




Nonparametric Trend for Producti
101
65.1
29.7
-5.78
-41.2
1910
1932
1954
1975
1997
Nonparametric Trend for GNP Defl
36.2
22.5
8.74
-5.01
-18.8
Nonparametric Trend for Interest
6.2
4
1.8
-0.4
-2.6
Nonparametric Trend for M2 Stock
119
33
-53.3
-140
Nonparametric Trend for Velocity
0.23
0.11
0
-0.12
-0.24
Nonparametric Trend Real Wages
8.26
-0.74
-9.73
-18.7
Nonparametric Trend Nonimal Wage
2474
1603
731
-141
-1013
Nonparametric Trend for S&P 500
48
13.1
-21.8
-56.8
-91.7



Nonparametric Trend Unemployment
10.5
5.16
-0.14
-5.45
1910
1932
1954
1975
1997
Nonparametric Trend for Bonds
0.57
0.2
-0.16
-0.53
1910
1932
1954
1975
1997