Inference of the Trend in a Partially Linear Model

Kun Ho Kim*

August 3, 2011

Abstract

In this paper, we construct the uniform confidence band (UCB) of a time-varying trend in a partially linear model. A two-stage local linear regression is proposed to estimate the time-varying trend. Based on this estimate, we develop an invariance principle to construct the UCB of the trend function. The proposed methodology is used to estimate the Non-Accelerating Inflation Rate of Unemployment (NAIRU) in the Phillips curve and to perform inference of the parameter based on its UCB. The empirical results strongly suggest that the U.S. NAIRU is time-varying.

Key words: Uniform confidence band (UCB), Time-varying trend, Partially linear model, Two-stage local linear regression, Strong invariane principle, Phillips curve, Time-varying NAIRU

JEL Classifications: C12, C13, C14

*Department of Economics, Michigan State University, Tel: 312-351-2942, email: kunhokim8@gmail.com

1 Introduction

The purpose of this paper is to construct the *uniform confidence band* (UCB) of a time-varying trend in a partially linear model. A partially linear model has been very popular in the semiparametric econometrics literature due to its flexibility to incorporate parametric and nonparametric components together. For an excellent treatment of this literature, we refer to Härdle, Liang & Gao (2000). The partial linear model that we are interested in here has the following framework:

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \mu(i/n) + \epsilon_i \tag{1}$$

where $i = 1, 2, \cdots, n$.[1] Note that y_i is the observed response variable and that ϵ_i is a stationary random variable with bounded fouth moment. Here $\mathbf{x}_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^\top$ is a $(p\times 1)$ vector of observed regressors. We allow $\mathbf{x}_i$ to be *non-stationary.* A $(p \times 1)$ vector $\boldsymbol{\beta}$ and a scalar $\mu(\cdot)$ are *unobserved.* The difference is that $\boldsymbol{\beta}$ is a vector of fixed parameters, while $\mu(\cdot)$ is a time-varying deterministic trend. A similar model is studied in Gao & Hawthorne (2006). For the assumptions on (1), we refer to Section 2. In this paper, our interest is in trend $\mu(\cdot)$. We estimate and carry out inference of parameter $\mu(\cdot)$.

Traditionally, there has been more interest in estimating unknown $\boldsymbol{\beta}$ rather than estimating unknown $\mu(\cdot)$. Asymptotically consistent estimators of parameter $\boldsymbol{\beta}$ have been proposed by Robinson (1988) and Hall, Kay & Titterington (1990), among others. For an excellent review on these estimators, we refer to Yatchew (1998). Despite this tendency, uniform confidence bands of $\mu(\cdot)$ based on its accurate estimate could be useful and valuable, given the importance of the time trend in an economic variable. One useful application of the framework (1) is to the *Phillips curve*, where the long-term trend in unemployment rate is usually called the Non-Accelerating Inflation Rate of Unemployment (NAIRU). In macroe-

[1] We use index i for time. The traditional time index t is reserved for continuous time in $[0, 1]$.

conomics, NAIRU plays an important role because this structural parameter allows us to determine what the current status of the economy is in business cycles. If the current unemployment rate is below the NAIRU, the economy is believed to be in an economic expansion. Otherwise, it is in a recession. In framework (1), $\mu(\cdot)$ can represent this NAIRU parameter (see Section 5.1). By constructing the UCB of the parameter, we can test any parametric specification of $\mu(\cdot)$. That is, we can perform *model validation* for trend $\mu(\cdot)$.

To construct the UCBs of the time-varying trend $\mu(\cdot)$ in model (1), we first propose a *two-stage local linear* estimate of $\mu(\cdot)$. We assume that $\mu(\cdot)$ is a nonparametric function that changes *smoothly* in continuous time. Specifically, we assume that the second derivative of $\mu(\cdot)$ is continuous over a fixed time domain $[0, 1]$. That is, $\mu(\cdot) \in \mathcal{C}^2[0, 1]$. A combination of the differencing estimation (Hall, Kay & Titterington, 1990) and the local-linear regression (Fan & Gijbels, 1996) is utilized to estimate $\mu(\cdot)$ in model (1). Then, we apply the *strong invariance principle* (Wu, 2007) to this two-stage local linear estimate of $\mu(\cdot)$ to construct the UCB of the time-varying trend. A detailed description of these procedures will be provided in Section 3 and Section 4.

In fact, it is common in economics to assume a smoothly time-varying process on structural parameters such as $\mu(\cdot)$ in (1). For example, many economists believe that the movement in the NAIRU of the Phillips curve is *smooth* (Ball & Mankiw, 2002; Gordon, 1997, 1998; Staiger, Stock & Watson, 1996, 1997). They impose parametric structures on time-varying NAIRU and estimate it. Cai (2007) introduces the slowly time-varying *beta-coefficient* to the traditional Capital Asset Pricing Model (CAPM) and estimates the coefficient based on the local-linear approach. González & Teräsvirta (2008) and González, Hubrich & Teräsvirta (2009) assume a smoothly time-varying parametric function for *core inflation* and use their parametric models to forecast the U.S. inflation series. Kim (2010) considers the Consumption-based CAPM (C-CAPM) where the *risk aversion* parameter is assumed to evolve slowly in time. He proposes a consistent estimator of time-varying risk aversion and studies the parameter's cyclical behavior based on the proposed estimate. Kim, Zhou & Wu (2010) introduce a partial equilibrium model with smoothly time-varying price/income *elasticities* for the U.S. durable goods market. They provide the empirical evidence that the demand elasticities in their model indeed change slowly in time. Given these implications, we assume that trend $\mu(\cdot)$ in (1) evolves smoothly over time.

The main contribution of this work is that we show how to construct the uniform confidence band (UCB) of $\mu(\cdot)$ in (1). The previous literature on time-varying trend mainly considers *point-wise* confidence intervals (Staiger, Stock & Watson, 1996, 1997). Since we model $\mu(\cdot)$ as a nonparametric function in continuous time, we need to construct confidence bands, not intervals. It is not only more challenging, but also more interesting to construct uniform bands than point-wise ones. In many occasions, uniform confidence bands allow us to perform important statistical inference that point-wise counterparts cannot. For example, the UCBs can be used to test whether $\mu(\cdot)$ takes any parametric form on it while the point-wise confidence bands cannot. For more detail on the construction of UCB, we refer to Section 4. To the best of our knowledge, this is the first time that one constructs the UCB of a time-varying trend in a partially linear model and carries out statitical inference about the parameter based on the constructed UCB. We provide an empirical application of this methodology in Section 5. That is, we construct the UCB of a time-varying NAIRU in the Phillips curve for the U.S. economy.

The second contribution is that we estimate the time-varying NAIRU in the Phillips curve without assuming any parametric structure on it. Staiger, Stock & Watson (1996) and Gordon (1998) posit a stochastic process on the NAIRU, namely, a random walk, to estimate the parameter. Due to this rather arbitrary choice on $\mu(\cdot)$, this approach is clearly prone to *mis-specification.* In contrast, we eliminate this type of arbitrary assumptions on the NAIRU, and model the parameter as a non-parametric function in continuous time. Estimation of the function is based on the two-stage local linear regression method. As a result, we reduce the possibility of mis-specification on the NAIRU. The minimum requirement for the unknown NAIRU is that the second derivative of the function is continuous over $t \in [0, 1]$.

The organization of the paper is the following: Section 2 discusses the assumptions needed for our asymptotic result. Section 3 introduces a *two-stage local linear* regression to estimate the time-varying trend $\mu(\cdot)$ in (1) on a fixed time domain. The first-stage parametric and the second-stage nonparametric estimations are carefully explained. The smoothness condition on $\mu(\cdot)$ is also discussed. Section 4 describes how to construct *uniform confidence band* (UCB) of $\mu(\cdot)$. To that end, we show how to apply the strong invariance principle (Wu, 2007) to our two-stage local linear estimation. The detailed steps to construct the UCB based on the strong invariance principle are also provided. As an empirical application, we estimate and perform inference of the NAIRU in the Phillips curve in Section 5. The results on the constructed UCB of NAIRU are also explained in detail. Section 6 concludes the paper and discusses related future research. The asymptotic results regarding the construction of UCB are given in Appendix.

2 Assumptions

We first introduce some notations that will be used throughout this study. For any vector $\mathbf{v} = (v_1, v_2, \ldots, v_p) \in \mathbb{R}^p$, we let $|\mathbf{v}| = (\sum_{i=1}^{p} v_i^2)^{1/2}$. For any random vector $\mathbf{V}$, we write $\mathbf{V} \in \mathcal{L}^q$ $(q > 0)$ if $||\mathbf{V}||_q = [\mathbb{E}(|\mathbf{V}|^q)]^{\frac{1}{q}} < \infty$. In particular, $||\mathbf{V}|| = ||\mathbf{V}||_2$. We denote

$\mathbf{L} : [0,1] \times \mathbb{R}^{\infty} \mapsto \mathbb{R}^p$ as a measurable function such that $\mathbf{L}(t, \mathcal{F}_i)$ is a properly defined $(p \times 1)$ random vector for all $t \in [0,1]$, where $\mathcal{F}_i = (\cdots, \eta_{i-1}, \eta_i)$ with independent and identically distributed (IID) random errors $\{\eta_i\}_{i \in \mathbb{Z}}$. Define the *physical dependence measure* (Wu, 2005) for $\mathbf{L}(t, \mathcal{F}_i)$ as the following:

$$\delta_q(\mathbf{L}, k) = \sup_{t \in [0,1]} \|\mathbf{L}(t, \mathcal{F}_k) - \mathbf{L}(t, \mathcal{F}_k^*)\|_q \tag{2}$$

where $\mathcal{F}_i^* = (\cdots, \eta_0^*, \cdots, \eta_{i-1}, \eta_i)$ is a coupled process with η_0^* being an IID copy of η_0. For a class of stochastic processes $\{\mathbf{L}(t, \mathcal{F}_i)\}_{i \in \mathbb{Z}}$, we say that the process is $\mathcal{L}^q$ *stochastic Lipschitz-continuous* over $[0,1]$ if:

$$\sup_{0 \le t_1 < t_2 \le 1} \frac{||\mathbf{L}(t_2, \mathcal{F}_0) - \mathbf{L}(t_1, \mathcal{F}_0)||_q}{|t_2 - t_1|} < \infty \tag{3}$$

We denote a collection of such systems by Lip_q. In addition, we write $a_n \asymp b_n$ if $|a_n/b_n|$ is bounded away from 0 and ∞ for all large n. Given these notations, we introduce the following assumptions that will be used throughout the paper. Given model (1):

Assumption 1. *The regressors are non-stationary such that*

$$\mathbf{x}_i = \mathbf{G}(i/n,\ \mathcal{U}_i), \quad i = 1, 2, \cdots, n, \tag{4}$$

where $\mathcal{U}_i = (\cdots, u_{i-1}, u_i)$ with $\{u_i\}_{i \in \mathbb{Z}}$ being a set of independent and identically distributed (IID) random elements. Here $\mathbf{G} := (G_1, G_2, \cdots, G_p)^\top : [0,1] \times \mathbb{R}^{\infty} \to \mathbb{R}^p$ is a measurable function such that $\mathbf{G}(t,\ \mathcal{U}_i)$ is well defined for each $t \in [0,1]$.

Assumption 2. *Let $M(t)$ be the $p \times p$ matrix with ijth entry $m_{ij}(t) = \mathbb{E}[G_i(t,\ \mathcal{U}_0) G_j(t,\ \mathcal{U}_0)]$. We assume that the smallest eigenvalue of $M(t)$ is bounded away from 0 on $t \in [0,1]$.*

Assumption 3. *Let $\mathbf{G}(t,\ \mathcal{U}_i) \in Lip_2$ and $\sup_{0 \le t \le 1} ||\mathbf{G}(t,\ \mathcal{U}_i)||_4 < \infty$.*

Assumption 4. *Let* ϵ_i *be a stationary random variable with* $\mathbb{E}\epsilon_i = 0$ *and* $\mathbb{E}|\epsilon_i|^4 < \infty$ *such that:*

$$\epsilon_i = H(\mathcal{V}_i), \quad i = 1, 2, \cdots, n, \tag{5}$$

where $\mathcal{V}_i = (\cdots, v_{i-1}, v_i)$ *with* $\{v_i\}_{i\in\mathbb{Z}}$ *being a set of IID random elements. Here* $H : \mathbb{R}^\infty \to \mathbb{R}$ *is a well-defined measurable function. Here* $\{v_i\}_{i\in\mathbb{Z}}$ *in (5) and* $\{u_j\}_{j\in\mathbb{Z}}$ *in (4) are independent of each other.*

Assumption 5. *Let* $\zeta_i = (u_i, v_i)^\top$ *and* $\mathcal{I}_i = (\cdots, \zeta_{i-2}, \zeta_{i-1}, \zeta_i)$. *Define* $\mathbf{U}(t, \mathcal{I}_i) = \mathbf{G}(t, \mathcal{U}_i)H(\mathcal{V}_i)$. *Then,* $\sum_{k=0}^{\infty}[\delta_4(\mathbf{G}, k) + \delta_2(\mathbf{U}, k)] < \infty$

Assumption 6. *Let* $\mu(\cdot)$ *be second-order continuous over* $[0, 1]$. *In other words,* $\mu(\cdot) \in \mathcal{C}^2[0, 1]$.

Assumption 7. *A kernel* $K(\cdot)$ *is positive and symmetric, and has a compact support* $[-1,\ 1]$*, and is Lipschitz continuously differentiable.*

Assumption 8. *A bandwidth* b_n *satisfies the following condition:*

$$nb_n^3 + \frac{1}{nb_n^2} = o(1) \tag{6}$$

We employ these assumptions to prove the lemmas and theorems in our paper. For detail on the proofs, we refer to Appendix. These assumptions are noteworthy and we have the following important remarks:

Remark 1. Assumption 1 allows regressors $\mathbf{x}_i$ in model (1) to be *non-stationary* since their moments are time-varying. It is often reasonable to assume that the characteristics of model regressors vary in time. In particular, we let the time variation in these characteristics is *smooth*, rather than abrupt. Note also that $\mathbf{x}_i$ is *dependent* due to the cumulative IID random elements. Assumption 2 prevents asymptotic *multicollinearity*.

Remark 2. Assumption 3 ensures that model regressors $\mathbf{x}_i$ are *local-stationary*, which is a mild form of non-stationarity. That is, if one observes local-stationary variables in a relatively short time span, they will be *approximately* stationary. Since many economic variables including unemployment rates can be modelled as local-stationary processes (see Figures 1 and 2), we can make our model specification more general by introducing this assumption. For more on this local-stationarity, we refer to Priestley (1965), Dahlhaus (1997), Mallat, Papanicolaou & Zhang (1998), Ombao, Von Sachs & Guo (2005) and Kim, Zhou & Wu (2010).

Remark 3. Assumption 4 allows error term ϵ_i to be *correlated* and thus *dependent* with bounded fourth-moments. The dependence structure for ϵ_i defined at (5) is flexible and general in that function $H(\cdot)$ is not specified. Note that the cumulative dependnece structure in ϵ_i allows us to apply the strong invariance principle (Wu, 2007) to our model framework. Examples of random process ϵ_i include a stationary ARCH process (Engle, 1982).

Remark 4. Assumption 5 ensures *short-range dependence* among the variables in our model. The interpretation is that the cumulative effect of a single error on all future values is *bounded*. The measure of dependence used here is the *physical dependence measure* (Wu, 2005) based on stationary causal processes. This measure is known to be particularly useful for characterizing dependence in nonlinear time series models (Wu, 2005, 2007; Zhao & Wu, 2008; Kim, 2010; Kim, Zhou & Wu, 2010).

Remark 5. Assumption 6 guarantees that trned $\mu(\cdot)$ changes *smoothly* in time. In particular, the second-order continuity of these parameters is required for the asymptotic consistency of the *local-linear estimates* (see Section 3) of their first-order derivatives.

Remark 6. Assumption 7 allows popular kernel functions such as the Epanechnikov kernel, which we will use in our semiparametric estimation of $\mu(\cdot)$. Assumption 8 is regarding the convergence order of the bandwidths for our local-linear regression.

3 Two-stage estimation of time-varying trend

As mentioned in the Introduction, the time-varying trend $\mu(\cdot)$ in model (1) has a fixed-time domain (Cressie, 1993) of $[0, 1]$, such that the period starts at 0 and ends at 1. Unlike other parametric approaches, the only restriction that we place on the trend $\mu(\cdot)$ in (1) is that $\mu(\cdot)$ changes *smoothly* over its time domain $[0, 1]$. Specifically, we assume that the time-varying trend $\mu(\cdot)$ satisfies $\mu(\cdot) \in \mathcal{C}^2[0, 1]$ (*i.e.* Assumption 6).

3.1 First-stage parametric estimation

Since the partially linear model (1) has both parameter $\boldsymbol{\beta}$ and nonparametric function $\mu(\cdot)$, it is natural to employ a *two-step* method to estimate both unknowns. A similar two-step method is considered in Kim (2010) among others. We first estimate β based on the first-differenced version of model (1):

$$\Delta y_i = \Delta \mathbf{x}_i^\top \boldsymbol{\beta} + \left[\mu\Big(\frac{i}{n}\Big) - \mu\Big(\frac{i-1}{n}\Big)\right] + \Delta\epsilon_i \tag{7}$$

where $\Delta y_i = y_i - y_{i-1}$, $\Delta \mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_{i-1}$ and $\Delta\epsilon_i = \epsilon_i - \epsilon_{i-1}$, respectively. Since $\mu(\cdot) \in \mathcal{C}^2[0, 1]$, we have the following result:

$$\left|\mu\Big(\frac{i}{n}\Big) - \mu\Big(\frac{i-1}{n}\Big)\right| = O\Big(\frac{1}{n}\Big) \tag{8}$$

which means that the left-hand side of the above equation disappears quickly as sample size n increases. Then, we can rewrite (7) by the following:

$$\Delta y_i = \Delta \mathbf{x}_i^\top \boldsymbol{\beta} + O(1/n) + \Delta\epsilon_i \tag{9}$$

The first-stage estimation of unknown parameter $\boldsymbol{\beta}$ in (1) is based on equation (9). This idea of differencing was first introduced by Hall, Kay & Titterington (1990). For an excellent summary on this differencing method, we refer to Yatchew (1997, 1998). In this study, we employ the *least squares* estimation method to estimate $\boldsymbol{\beta}$ in (7), treating $\Delta \mathbf{x}_i$ as the regressor and Δy_i as the response variable. That is, we propose the following *differencing estimator*:

$$\hat{\boldsymbol{\beta}}_D = \Big(\sum_{i=1}^{n} \Delta \mathbf{x}_i \Delta \mathbf{x}_i^\top\Big)^{-1} \sum_{i=1}^{n} \Delta \mathbf{x}_i \Delta y_i \tag{10}$$

where $\hat{\boldsymbol{\beta}}_D$ is based on the differenced data. The asymptotic consistency for this differencing estimator is given by the following theorem:

Theorem 1. *Under Assumption 1 – Assumption 8,*

$$|\hat{\boldsymbol{\beta}}_D - \boldsymbol{\beta}| = O_{\mathbb{P}}\Big(\frac{1}{\sqrt{n}}\Big) \tag{11}$$

where $\hat{\boldsymbol{\beta}}_D$ is the differencing estimator of $\boldsymbol{\beta}$.

The proof of Theorem 1 is provided in Appendix. The similar result under a different set of assumptions is obrained in Yatchew (1997). Robinson (1988) introduced his kernel-based estimation of $\boldsymbol{\beta}$ in (1). In this work, we prefer to employ the differece-based method in Hall, Kay & Titterington (1990) to the Robinson's estimator, because the latter involves quite tedious separate nonparametric regressions for each parametric variable and for the dependent variable. On the other hand, the difference-based estimation here avoids those preliminary nonparametric regressions required for the Robinson's method. A simple smoothness condition on the trend function $\mu(\cdot)$ allows us to employ this difference-based estimation method for our study.

3.2 Second-stage nonparametric estimation

Given the difference-based estimate of $\boldsymbol{\beta}$ in (1), we are ready to estimate the time-varying trend, $\mu(t)$. We first rewrite the partially linear model (1), and obtain the following:

$$y_i - \mathbf{x}_i^\top \boldsymbol{\beta} = \mu(i/n) + \epsilon_i \tag{12}$$

where ϵ_i is a stationary martingale difference sequence satisfying $\mathbb{E}|\epsilon_i|^4 < \infty$. Given that $\mu(t)$ changes *smoothly* in time $t \in [0,1]$, we employ the *local linear regression* method to estimate the trend function $\mu(\cdot)$. Among many available nonparametric methods, local polynomial regression (Fan & Gijbels, 1996) is frequently used due to its simple form, ease of computation and analytical tractability. In contrast to other popular nonparametric methods such as the Nadaraya-Watson estimator ($i.e.$ local constant estimator), a local linear estimator does suppress the well-known *boundary problem* and achieve nearly optimal statistical efficiency (Fan & Gijbels, 1996). Thus, we shall use a local linear estimate of $\mu(\cdot)$ in this study. The local linear regression gives:

$$(\hat{\mu}(t), \hat{\mu}'(t)) = arg \min_{(\eta_0, \eta_1)} \sum_{i=1}^{n} [y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \eta_0 - \eta_1(t_i - t)]^2 K\Big(\frac{t_i - t}{b_n}\Big) \tag{13}$$

where $t_i = \frac{i}{n}$ and $i = 1, \cdots, n$. Here $K(\cdot)$ is a kernel function and b_n is a bandwidth. We let $K(\cdot)$ be a symmetric probability density function with support $[-1, 1]$. A popular choice is the *Epanechnikov* kernel $K(x) = 3\max(1 - x^2, 0)/4$. The bandwidth b_n can be viewed as the size of the neighborhood on which $\mu(\cdot)$ is estimated. If b_n is large, then the estimation is based on a large neighborhood and the estimated $\hat{\mu}(t)$ will have a small variance and a large bias. Analogously, if b_n is small, $\hat{\mu}(t)$ will have a large variance and a small bias. In this study, we select our bandwidth based on the *generalized cross-validation* (GCV) method (Wahba, 1977; Craven & Wahba, 1979). Here $\hat{\mu}(\cdot)$ is called the *local linear estimate* of $\mu(\cdot)$. As a by-product, (13) also gives an estimate of its derivative $\mu'(\cdot)$. Since (13) is essentially a weighted least squares estimate, we can write the solution of (13) for $\mu(\cdot)$ as the following:

$$\hat{\mu}(t) = \sum_{i=1}^{n} w_n(t,i)(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}) \tag{14}$$

where $w_n(t,i) = K\Big(\frac{t-i/n}{b_n}\Big)\frac{S_2(t)-(t-i/n)S_1(t)}{S_2(t)S_0(t)-S_1^2(t)}$ with $S_j(t) = \sum_{i=1}^{n}(t-i/n)^j K\Big(\frac{t-i/n}{b_n}\Big)$. The time domain of t is fixed over $t \in [0,1]$ and $w_n(t,i)$ is the weight given to each observation. Note here that estimator (14) is *infeasible* because $\boldsymbol{\beta}$ is *unknown*. Hence, we propose the following *feasible* estimator of the time-varying trend $\mu(t)$ in (12):

$$\tilde{\mu}(t) = \sum_{i=1}^{n} w_n(t,i)(y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}_D) \tag{15}$$

where $\hat{\boldsymbol{\beta}}_D$ is the differencing estimate of $\boldsymbol{\beta}$. We call $\tilde{\mu}(t)$ a *two-stage* local linear estimator of $\mu(\cdot)$ since the first-stage differencing estimate $\hat{\boldsymbol{\beta}}_D$ is embedded in the second-stage local linear estimate of time-varying $\mu(\cdot)$. The asymptotic consistency of this two-stage local linear estimator (15) is given by the following Theorem:

Theorem 2. *Under Assumption 1 – Assumption 8, for any fixed $t \in (0,1)$,*

$$\Big|\tilde{\mu}(t) - \mu(t)\Big| = O_{\mathbb{P}}\Big(b_n + \frac{1}{\sqrt{nb_n}} + \frac{1}{\sqrt{n}}\Big) \tag{16}$$

where $\tilde{\mu}(t)$ is the two-stage local linear estimator for $\mu(t)$ defined by (15).

The proof of Theorem 2 is provided in Appendix. By Assumption 8, the asymptotic consistency of $\tilde{\mu}(t)$ to $\mu(t)$ follows easily.

4 Uniform confidence bands (UCB) of time-varying trend

Until recently, the empirical literature on the partially linear model rarely provided any statistical inference based on the estimates of either fixed or time-varying trend. A handful of works that pioneered in this direction include Gordon (1997) and Staiger, Stock & Watson (1997), where the authors estimate a time-varying NAIRU in the Phillips curve and construct its confidence intervals. Unfortunately, these previous works on the statistical inference about time-varying trend are only regarding the construction of *point-wise* confidence intervals of the trend. In treatment of dynamic models such as the partially linear model with a *time-varying* trend, it is much more appropriate and more useful to construct *uniform confidence bands* (UCB) than their point-wise counterparts, becaue the UCBs allow us to perform interesting statistical inference about the trend, such as a test on its *uniform constancy* over a certain time period. In principle, point-wise confidence intervals/bands are not appropriate for testing the uniform constancy of a time-varying parameter because any inference results based on them are regarding one specific point only.

In order to construct the asymptotic UCB of trend $\mu(t)$ over the fixed domain $t \in [0,1]$ with the confidence level $100(1-\alpha)\%$, $\alpha \in (0,1)$, we have to find the following two functions $f_n(\cdot)$ and $g_n(\cdot)$ based on the data that we have:

$$\lim_{n\to\infty} \mathbb{P}\{f_n(t) \le \mu(t) \le g_n(t) \text{ for all } t \in \mathcal{T}\} = 1-\alpha \tag{17}$$

where $\mathcal{T} = [0,1]$. The purpose of constructing the above UCB is to test whether the volatility $\mu(\cdot)$ takes a certain parametric form. That is, using the UCB of $\mu(\cdot)$, we are able to test the null hypothesis $H_0 : \mu(\cdot) = \mu_\theta(\cdot)$, where $\theta \in \Theta$ and Θ is a parameter space. For example, in order to test $H_0 : \mu(t) = \theta_0 + \theta_1 t + \theta_2 t^2$, one can simply check whether

$f_n(t) \leq \hat{\theta}_0 + \hat{\theta}_1 t + \hat{\theta}_2 t^2 \leq g_n(t)$ holds for *all* $t \in \mathcal{T}$. Here $\hat{\theta}_i$ is the least squares estimate of θ_i, $i = 0, 1, 2$. If the condition does hold for *all* $t \in \mathcal{T}$, then we *fail to reject* the null hypothesis at significance level α.

In general, the UCB is a more *conservative* confidence band than the traditional point-wise confidence band in the sense that the UCB is usually *wider* than its point-wise copunterpart. Thus, any test results based on the UCB would be more *robust* than those under the point-wise ones. For these reasons, the UCB has recently gained more attention in econometrics and statistics literature. For example, Härdle & Song (2009) construct the UCBs of quantiles for the age-earning relation in the U.S. labor market. Liu & Wu (2010) construct the UCBs of the mean and volatility of the U.S. Treasury yield curve rates. Given these interesting results, we shall construct the UCBs of time-varying trend in the partially linear model (12), and carry out statistical inference about it. To our knowledge, this is the first time that one performs a formal test on parametric specifications of the trend in model (12) based on its UCB.

4.1 Strong invariance principle

In constructing the UCBs for time-varying trend $\mu(\cdot)$ in model (12), we shall employ a very useful tool in asypmtotic theories, called the *strong invariance principle.* This principle basically allows us to access asymptotic properties of a partial sum of stationary and ergodic random variables. Komlós, Major & Tusnády (1975) introduce their celebrated strong invariance principle for IID random variables. Wu (2007) extends the idea by developing his invariance principle for *dependent* random variables. Given the dependence among random error ϵ_i of the partially linear model (12) (see Assumption 3), we employ the strong invariance principle in Wu (2007) in this study. By the strong invariance principle in Wu (2007), we have the following:

$$\max_{i\le n}\Big|\sum_{k=1}^{i}\epsilon_k-\sigma_\epsilon\mathbb{B}(i)\Big|=o_{\mathbb{AS}}\Big(n^{1/4}\log(n)\Big) \tag{18}$$

where ϵ_i is the error term in the partially linear model (12). Here $\sigma_\epsilon^2=\sum_{k\in\mathbb{Z}}\mathbb{E}(\epsilon_0\epsilon_k)$ is the *long-run variance* of ϵ_i, and $\mathbb{B}(\cdot)$ is the standard Brownian motion. By applying the strong invariance principle (18) and the *summation-by-parts* formula, we can show the following:

$$\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n}w_n(t,i)\epsilon_i-\sigma_\epsilon\sum_{i=1}^{n}w_n(t,i)[\mathbb{B}(i)-\mathbb{B}(i-1)]\Big|=o_{\mathbb{AS}}\Big(\frac{\log(n)}{n^{3/4}b_n}\Big) \tag{19}$$

where $w_n(t,i)$ is the weight for the local-linear regression defined in (14). For the proof of (19), we refer to Lemma 3 in Appendix. Given (19), we can obtain the following Theorem:

Theorem 3. (Invariance Principle) *Recall that $\tilde{\mu}(t)$ is the two-stage local linear regression estimator of $\mu(t)$ from (15). Under Assumption 1 – Assumption 8,*

$$\sqrt{nb_n}\sup_{0\le t\le 1}\Big|\tilde{\mu}(t)-\mu(t)-\sigma_\epsilon\sum_{i=1}^{n}w_n(t,i)Z_i\Big|=o_{\mathbb{P}}(1) \tag{20}$$

where $\sigma_\epsilon^2=\sum_{k\in\mathbb{Z}}\mathbb{E}(\epsilon_0\epsilon_k)$ is the long-run variance of ϵ_i, and Z_i is an IID standard normal random variable.

The proof of Theorem 3 is given in Appendix. The invariance principle in Theorem 3 states that we can *approximate* the quantiles of $\sup_{0\le t\le 1}\Big|\frac{\tilde{\mu}(t)-\mu(t)}{\sigma_\epsilon}\Big|$ by the quantiles of the sampling distribution of $\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n}w(t,i)Z_i\Big|$. This is an extremely important and useful result, because it means that we can easily approximate the quantiles of the proposed two-stage estimator by using IID standard normals instead. Without this result, we have to use the asymptotic distribution of $\sup_{0\le t\le 1}\Big|\frac{\tilde{\mu}(t)-\mu(t)}{\sigma_\epsilon}\Big|$ in order to construct the uniform confidence bands of $\mu(t)$. However, this approach should be taken with a great caution because the asymptotic distribution of $\sup_{0\le t\le 1}\Big|\frac{\tilde{\mu}(t)-\mu(t)}{\sigma_\epsilon}\Big|$ is the extreme-value (or Gumbel) distribution

(Bickel & Rosenblatt, 1973; Johnston, 1982; Härdle, 1989; Fan & Zhang, 2000; Wu & Zhao, 2007; Liu & Wu, 2010). It is well-known that the convergence to this distribution is *extremely slow* (Wu & Zhao, 2007; Liu & Wu, 2010), and the confidence bands based directly on this distribution could be very inaccurate if the sample size is not large enough.

4.2 Construction of uniform confidence bands (UCB)

Given this useful result of Theorem 3, we can overcome the potential problem of directly using the asymptotic distribution of $\sup_{0\le t\le 1}\left|\frac{\tilde{\mu}(t)-\mu(t)}{\sigma_\epsilon}\right|$ to construct the uniform confidence bands (UCB) of the time-varying trend in our model. This idea of constructing uniform confidence bands using IID standard normals was utilized by Eubank & Speckman (1993), where they considered the case of IID random errors. Wu & Zhao (2007) extended the idea to the case of dependent errors. In this work, we apply the idea to a partially lnear model. Based on Theorem 3, we suggest the following procedures to construct the UCBs of time-varying trend $\mu(t)$ in (12):

(i) Obtain the differencing estimate $\hat{\boldsymbol{\beta}}_D$ in (10) by running the least squares on (9).

(ii) Select the optimal bandwidth b_n for our local linear regression (15) based on the generalized cross-validation (GCV) method (Wahba, 1977; Craven & Wahba, 1979).

(iii) Obtain the two-stage local linear estimate $\tilde{\mu}(t)$ proposed in (15).

(iv) Compute $\sup_{0\le t\le 1}\left|\sum_{i=1}^{n} w_n(t,i)Z_i\right|$, where $w_n(t,i)$ is the weight for local linear regression in (14), and $\{Z_i\}$ are generated IID standard normals.

(v) Repeat (iv), say $1,000$ times. We obtain the 95th quantile of this sampling distribution of $\sup_{0\le t\le 1}\left|\sum_{i=1}^{n} w(t,i)Z_i\right|$, and denote it as $\hat{q}_{0.95}$.

(vi) Estimate σ_ϵ using the following *variant* of the subseries variance estimator proposed by Carlstein (1986) and extended by Wu & Zhao (2007):

$$\hat{\sigma}_\epsilon^2 = \frac{1}{2(m-1)k_n} \sum_{i=1}^{m-1} \Big[\sum_{j=1}^{k_n} \Big(y_{j+ik_n} - y_{j+(i-1)k_n} - (\mathbf{x}_{j+ik_n} - \mathbf{x}_{j+(i-1)k_n})^\top \hat{\boldsymbol{\beta}}_D \Big) \Big]^2 \tag{21}$$

where k_n is the length of subseries and $m = [n/k_n]$ is the largest integer not exceeding n/k_n. Carlstein (1986) shows that the optimal length of subseries is $k_n \asymp n^{1/3}$. Hence, we let $k_n \asymp n^{1/3}$ here. For a finite sample, we choose $k_n = [n^{1/3}]$. The asymptotic consistency of $\hat{\sigma}_\epsilon^2$ to the long-run variance σ_ϵ^2 is given by Lemma 4 in Appendix.

(vii) The 95% UCB of $\mu(t)$ is $[\tilde{\mu}(t) \pm \hat{\sigma}_\epsilon \hat{q}_{0.95}]$.

As an application of these proposed steps to construct uniform bands of $\mu(\cdot)$, we shall construct the UCBs of a time-varying NAIRU in the Phillips curve. Detailed descriptions of the model, the data and the empirical result will be provided in the following section.

5 Application

5.1 Phillips curve with time-varying NAIRU

One direct and useful application of the framework (1) is to the inflation–unemployment tradeoff, which is known as the Phillips curve in the macroeconomics literature. Specifically, we intend to construct the UCB of the Non-Accelerating Inflation Rate of Unemployment (NAIRU) in the traditional Phillips curve. One of the simple versions of the NAIRU equation in the literature is the following random-walk-type Phillips curve (Staiger, Stock & Watson, 1996, 1997; Gordon, 1997, 1998; Fair, 2000; Ball & Mankiw, 2002):

$$\pi_i = \pi_{i-1} + \alpha(U_i - U_N) + \epsilon_i \tag{22}$$

where π_i and U_i are inflation and unemployment rates at time $i = 1, \cdots, n$, respectively. Here ϵ_i is a mean zero random error at time i, often dubbed the *supply shock*. This equation has two unknown parameters: α and U_N. Here U_N is the NAIRU that is also called the natural rate of unemployment. The parameter U_N embeds all shifts in the inflation-unemployment tradeoff. In that sense, the natural rate can be viewed as the unemployment rate that the economy reaches in the long run.

In principle, U_N can exhibit substantial *variation over time* (Gordon, 1997; Staiger, Stock & Watson, 1996, 1997; Ball & Mankiw, 2002). To implement this idea, we build on the traditional random-walk-type model (22) and propose the following Phillips curve with *time-varying* NAIRU $U_N(\cdot)$:

$$\pi_i = \pi_{i-1} + \alpha\Big[U_i - U_N\Big(\frac{i}{n}\Big)\Big] + \epsilon_i \tag{23}$$

where $U_N(\cdot)$ is time-varying over its domain $[0, 1]$. Obviously, this model is an extension of the general framework (1). One of the differences between our model (23) and the previous works on time-varying NAIRU in the literature (Ball & Mankiw, 2002; Gordon, 1997; Staiger, Stock & Watson, 1996, 1997) is that the domain of $U_N(\cdot)$ in (23) is bounded strictly between zero and one because of $i = 1, \cdots, n$. By using $\frac{i}{n}$ instead of i for $U_N(\cdot)$, we can naturally restrict the domain of this nonparametric function to $[0, 1]$, which can help construct a consistent semiparametric estimator of $U_N(\cdot)$. As discussed previously, we assume that $U_N(\cdot)$ is second-order continuous over its fixed time domain $[0, 1]$. Hence, the time-varying NAIRU $U_N(t)$ can be estimated by the two-stage local linear estimator $\tilde{U}_N(t)$ proposed in (15)[2]:

$$\tilde{U}_N(t) = \sum_{i=1}^{n} w_n(t, i)\Big(U_i - \frac{\Delta\pi_i}{\hat{\alpha}_D}\Big) \tag{24}$$

where $\Delta\pi_i = \pi_i - \pi_{i-1}$ and $\hat{\alpha}_D$ is the differencing estimate of α in (23)[3]. Then, by following

[2] We use an Epanechnikov kernel. The GCV chooses $b_n = 0.10$ for (24).

[3] Here $\hat{\alpha}_D = -0.2832$.

the procedures given in Section 4.2, we can construct the UCBs of time-varying NAIRU $U_N(t)$.

5.2 Empirical Results

To construct the UCB of NAIRU for the U.S. economy, we use *monthly* unemployment rates and *monthly* Consumer Price Index (CPI) for the U.S. economy from March 1950 to September 2007. The data on both unemployment rate and CPI were obtained from the Global Insight Basic Economics database at the Wharton Research Data Services (WRDS) of the Wharton School, the University of Pennsylvania (http://wrds.wharton.upenn.edu). The estimated time-varying NAIRU and its 95% uniform confidence bands (UCB) are plotted along with the monthly unemployment data in Figure 1. Similarly, we plot the 99% UCB in Figure 2. Note that the 99% UCB is naturally wider than the 95% UCB[4].

As we can observe from Figures 1 and 2, there has been a substantial amount of fluctuation in the U.S. NAIRU between 1950 and 2007. Throughout the 1950s and the 1960s, the estimated NAIRU stays below 6%, although it rises during the 1950s and falls during the 1960s. It steadily increases during the 1970s. The peak in the estimated NAIRU comes in the early 1980s. Since then, it has declined gradually, which becomes most notable during the late 1990s. It reaches its trough around 2000, and then slightly *increases* after that. Ball & Mankiw (2002) and Staiger, Stock & Watson (1997, 2001) report very similar findings to ours. One difference between our result and theirs is that we estimate NAIRU before 1960 and after 2000 as well, which means that our result is more extensive.

An important contribution of our work is that we can statistically test the *uniform constancy* of NAIRU based on our empirical results. Based on the estimated 95% UCB of NAIRU

[4]The width of 95% UCB of time-varying NAIRU is 2.10% while that of 99% UCB is 2.50%.

(Figure 1), we *reject*, at 5% significance level, the null hypothesis that the U.S. NAIRU is constant between 1950 and 2007. This is because no horizontal line can be contained by the constructed 95% UCB of the parameter (see Figure 1). We *reject* the null hypothesis even at 1% significance level since we cannot place any horizontal line inside the 99% UCB in Figure 2. Based on these findings, there seems to be a strong evidence against the hypothesis of time-invariant NAIRU. In order to test the uniform constancy of a parameter, one needs to construct $uniform$ confidence bands of the parameter, the U.S. NAIRU in this example. To the best of our knowledge, this is the first paper that constructs the UCB of NAIRU and rejects its uniform constancy based on the UCB.

Typically, statistical inference about structural parameters has been based mostly on their *point-wise* confidence intervals. Staiger, Stock & Watson (1997) employ this approach and report the 95% confidence intervals of U.S. NAIRU at three specific dates (first quarters of 1984, 1989 and 1994). From Table 1 (page 39) of Staiger, Stock & Watson (1997), the widths of CPI-based 95% confidence intervals of NAIRU at these three dates are 3.7%, 3.1% and 3.7%, respectively. Since their results are point-wise, the width of interval varies from point to point. On the other hand, the width of our CPI-based 95% UCB of NAIRU is only **2.10**%, while the width of 99% UCB is **2.50**%[5]. This implies that our two-stage estimate of U.S. NAIRU along with its constructed UCB are *more informative* than the estimates reported in Staiger, Stock & Watson (1997), because our UCBs have a *smaller* width than the point-wise ones reported in Table 1 of their paper. Considering that UCBs are wider than point-wise bands in general, this is a strong evidence that our estimates of NAIRU are *more precise* than those in Staiger, Stock & Watson (1997). From the same Table in Staiger, Stock & Watson (1997), our UCBs are either comparable to or smaller than the other confidence intervals based on different measures of U.S. inflation. Our UCB with a

[5]In our case, the width of confidence band is constant because the bands are $uniform$.

small width is an important contribution to the literature because the major drawback with the NAIRU estimates from previous studies (Staiger, Stock & Watson, 1996, 1997) was the large confidence interval.

Furthermore, the estimation results in this study confirm that the Phillips curve has a *negative* slope. In other words, the difference-based estimate of α in the equation (23) is $\hat{\alpha}_D = -0.2832$. This result confirms the findings from previous works; there seems to be a clear *negative* relationship between unemployment and inflation. The literature on the Phillips curve with the traditional *constant* NAIRU reports similar findings on this relationship. Here we instead work on the Phillips curve with a *time-varying* NAIRU, and find out that the negative relationship between unemployment and inflation still remains unchanged.

6 Conclusion

In this paper, we illiustrate how to construct the uniform confidence bands (UCB) of a time-vaying trend in a partially linear model. The main idea is to apply the strong invariance principle to a partial sum of the model error terms and to estimate quantiles of the proposed test statistic involving the two-stage local linear estimator by using IID standard normals. The asymptotic justification of this approach is provided by Theorem 3. As an application of this methodology, we employ the random-walk-type Phillips curve (Staiger, Stock & Watson, 1996, 1997; Fair, 2000; Ball & Mankiw, 2002) and construct the UCB of the time-varying NAIRU in the model. The empirical results suggest that there has been a substantial amount of time variation in the U.S. NAIRU during the $1950-2007$ period. The constructed UCB of time-varying NAIRU with both 95% and 99% confidence levels supports the time variation (Figures 1 and 2). In other words, the hypothesis of time-invariant NAIRU is rejected at both 5% and 1% significance levels by the test based on the UCB of NAIRU.

Moreover, the widths of our UCBs of NAIRU are *smaller* than those of the confidence intervals reported in Staiger, Stock & Watson (1996, 1997), which means that our two-stage local linear estimates of NAIRU appear *more informative* than those reported in Staiger, Stock & Watson (1996, 1997). This is a clear progress in that the major drawback of the results in Staiger, Stock & Watson (1996, 1997) is the *wide* confidence interval. Furthermore, the results in this study also confirm the gradual increase in the U.S. NAIRU during the 1970s and early 1980s, and the steady decrease during the mid/late 1980s and 1990s, which was also reported in Staiger, Stock & Watson (1996, 1997) and Ball & Mankiw (2002). Our results go beyond this, and further reveal that the NAIRU slightly increases after 2000 (Figures 1 and 2).

Regarding future research, this project suggests a couple of interesting issues and topics for consideration. First, we can extend the current work to the case of *conditionally heteroskedastic* errors. Thus, we can generalize the result found here. A rather straightforward extension of the current work will suffice. Second, we can investigate the *forecasting* ability of the proposed model with a time-varying trend. One of the main purposes of using the Phillips curve in macroeconomics is to forecast inflation series. We can hopefully improve the forecasting ability of a partially linear model by adopting a time-varying trend in it. Further insight can be gained by extending the current work in these and other directions.

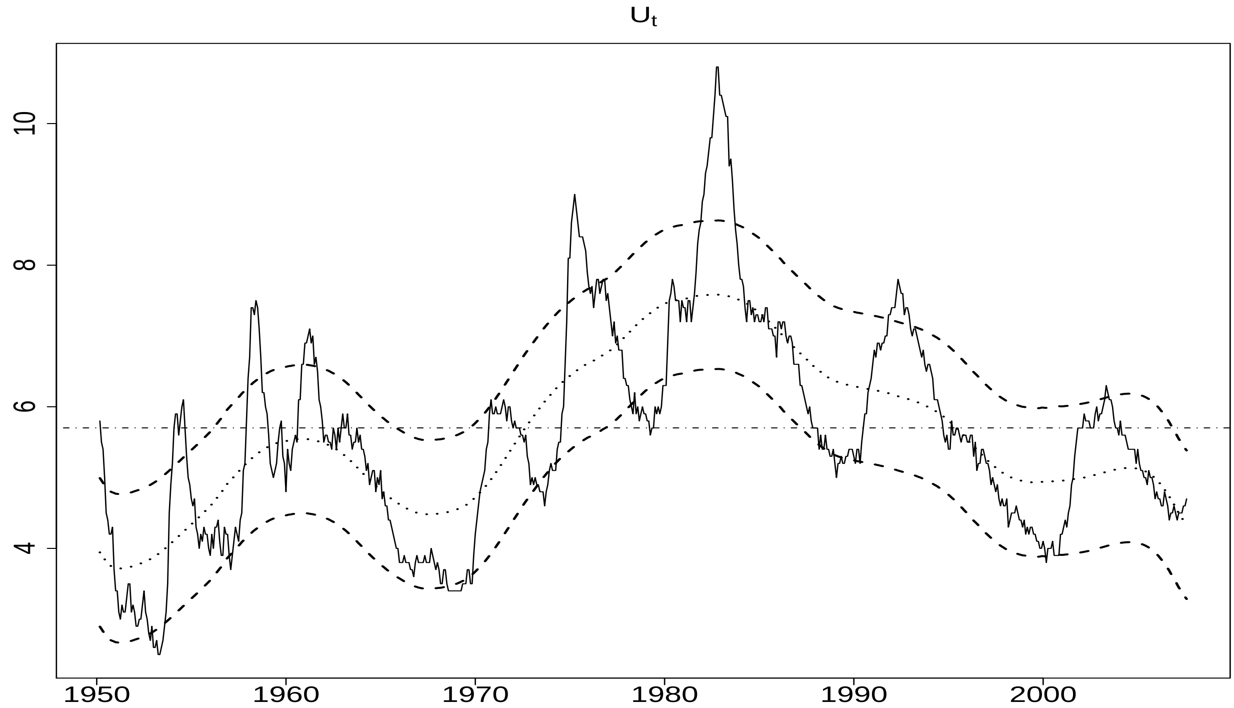

Figure 1: Time-varying NAIRU for the U.S. economy (1950-2007): The dashed curves are **95 % uniform confidence band** (UCB) of NAIRU. The width of 95% UCB is **2.10%**. The dotted curve in the middle of the UCB is the local linear estimate of time-varying NAIRU. For the local linear regression, we use an Epanechnikov kernel. The GCV chooses $b_n = 0.10$. The estimate of NAIRU and the UCB are placed over the U.S. monthly unemployment rates (solid). The dot-dashed horizontal line is the estimate of *fixed* NAIRU, 5.70%.

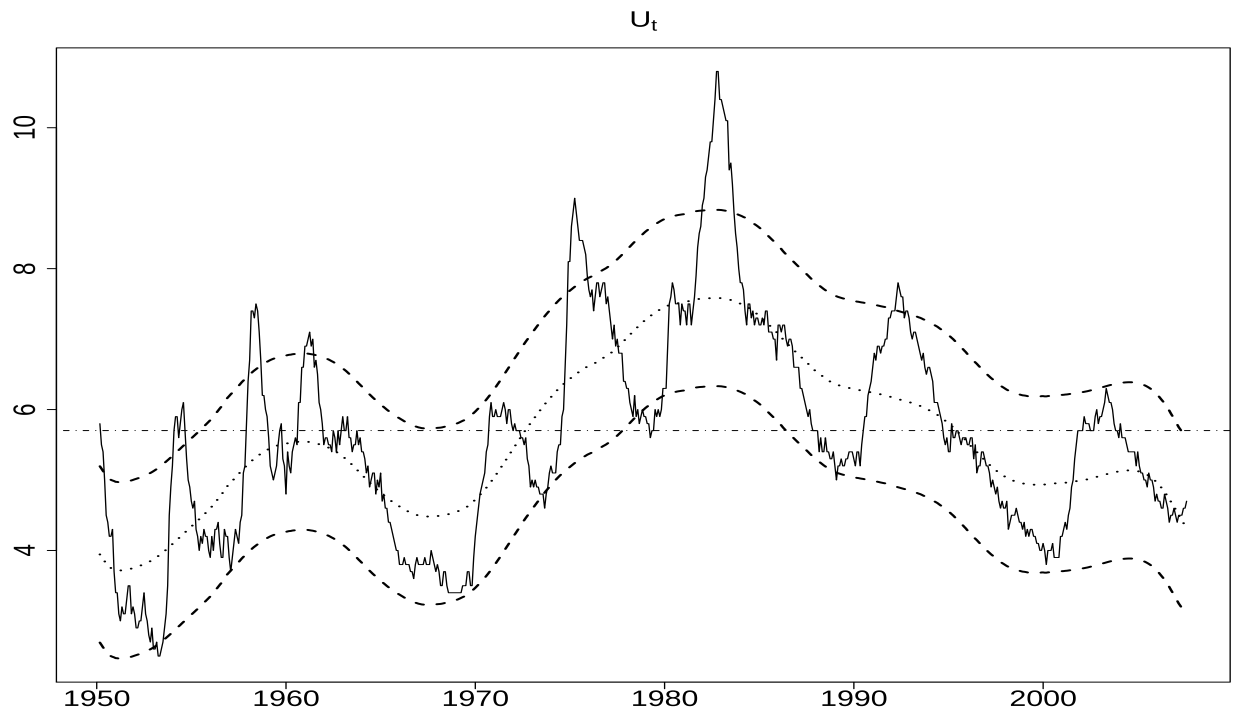

Figure 2: Time-varying NAIRU for the U.S. economy (1950-2007): The dashed curves are **99 % uniform confidence band** (UCB) of NAIRU. The width of 95% UCB is **2.50%**. The dotted curve in the middle of the UCB is the local linear estimate of time-varying NAIRU. For the local linear regression, we use an Epanechnikov kernel. The GCV chooses $b_n = 0.10$. The estimate of NAIRU and the UCB are placed over the U.S. monthly unemployment rates (solid). The dot-dashed horizontal line is the estimate of *fixed* NAIRU, 5.70%.

7 Appendix

We define the first-differenced regressors as the following:

$$\Delta \mathbf{x}_i = \mathbf{G}(t_i,\ \mathcal{U}_i) - \mathbf{G}(t_{i-1},\ \mathcal{U}_{i-1}) := \mathbf{G}_d(t_i,\ \mathcal{U}_i) \tag{25}$$

where $t_i = i/n$. Similarly, $\Delta\epsilon_i = H_d(\mathcal{V}_i)$. Then, we can define

$$\mathbf{L}_d(t, \mathcal{I}_i) := \mathbf{G}_d(t,\ \mathcal{U}_i) H_d(\mathcal{V}_i) \tag{26}$$

Lemma 1. *Under Assumption 1 – Assumption 8,*

$$\mathbf{G}_d(t,\ \mathcal{U}_i) \in Lip_2 \quad \textit{and} \quad \sup_{0\le t\le 1} ||\mathbf{G}_d(t,\ \mathcal{U}_i)||_4 < \infty$$

proof) Note the following:

$$\begin{aligned}
& \sup_{0\le t_i < t_j \le 1} \frac{||\mathbf{G}_d(t_j, \mathcal{U}_0) - \mathbf{G}_d(t_i, \mathcal{U}_0)||_2}{|t_j - t_i|} \\
\le & \sup_{0\le t_i < t_j \le 1} \frac{||\mathbf{G}(t_j, \mathcal{U}_0) - \mathbf{G}(t_{j-1}, \mathcal{U}_0) - [\mathbf{G}(t_i, \mathcal{U}_0) - \mathbf{G}(t_{i-1}, \mathcal{U}_0)]||_2}{|t_j - t_i|} \\
\le & \sup_{0\le t_i < t_j \le 1} \frac{||\mathbf{G}(t_j, \mathcal{U}_0) - \mathbf{G}(t_i, \mathcal{U}_0)||_2}{|t_j - t_i|} + \sup_{0\le t_{i-1} < t_{j-1} \le 1} \frac{||\mathbf{G}(t_{j-1}, \mathcal{U}_0) - \mathbf{G}(t_{i-1}, \mathcal{U}_0)||_2}{|t_{j-1} - t_{i-1}|}
\end{aligned}$$

where $t_j - t_i = t_{j-1} - t_{i-1}$ by definition. By **Assumption 3**,

$$\sup_{0\le t_i < t_j \le 1} \frac{||\mathbf{G}(t_j, \mathcal{U}_0) - \mathbf{G}(t_i, \mathcal{U}_0)||_2}{|t_j - t_i|} < \infty$$

and

$$\sup_{0\le t_{i-1} < t_{j-1} \le 1} \frac{||\mathbf{G}(t_{j-1}, \mathcal{U}_0) - \mathbf{G}(t_{i-1}, \mathcal{U}_0)||_2}{|t_{j-1} - t_{i-1}|} < \infty$$

Thus,

$$\sup_{0\le t_i < t_j \le 1} \frac{||\mathbf{G}_d(t_j, \mathcal{U}_0) - \mathbf{G}_d(t_i, \mathcal{U}_0)||_2}{|t_j - t_i|} < \infty$$

which means $\mathbf{G}_d(t,\, \mathcal{U}_i) \in \text{Lip}_2$. Moreover, by **Assumption 3**,

$$\sup_{0\le t_i\le 1} ||\mathbf{G}_d(t_i,\, \mathcal{U}_i)||_4 \le \sup_{0\le t_i\le 1} ||\mathbf{G}(t_i,\, \mathcal{U}_i)||_4 + \sup_{0\le t_{i-1}\le 1} ||\mathbf{G}(t_{i-1},\, \mathcal{U}_{i-1})||_4 < \infty$$

Therefore, the lemma easily follows.

$\square$

Lemma 2. *Under Assumption 1 – Assumption 8,*

$$\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\epsilon_i\Big| = O_{\mathbb{P}}(1)$$

proof) Let $\psi_d = \frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i\Delta\epsilon_i$. Define

$$\psi_{d,k} = \frac{1}{n}\sum_{i=1}^{n} \mathcal{P}_{i-k}\Big(\Delta\mathbf{x}_i\Delta\epsilon_i\Big)$$

where the projection operator is $\mathcal{P}_k(\cdot) := \mathbb{E}[\cdot|\mathcal{I}_k] - \mathbb{E}[\cdot|\mathcal{I}_{k-1}]$. Then,

$$\begin{aligned}
||\psi_{d,k}||^2 &= \Big|\Big|\frac{1}{n}\sum_{i=1}^{n} \mathcal{P}_{i-k}\Big(\Delta\mathbf{x}_i\Delta\epsilon_i\Big)\Big|\Big|^2 \\
&\le \frac{1}{n^2}\sum_{i=1}^{n} \Big|\Big|\mathbb{E}(\Delta\mathbf{x}_i\Delta\epsilon_i|\mathcal{I}_{i-k}) - \mathbb{E}(\Delta\mathbf{x}_i\Delta\epsilon_i|\mathcal{I}_{i-k-1})\Big|\Big|^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \Big|\Big|\mathbb{E}(\Delta\mathbf{x}_i\Delta\epsilon_i|\mathcal{I}_{i-k}) - \mathbb{E}(\Delta\mathbf{x}^*_{i,k}\Delta\epsilon^*_{i,k}|\mathcal{I}_{i-k})\Big|\Big|^2 \\
&\le \frac{1}{n^2}\sum_{i=1}^{n} \Big|\Big|\Delta\mathbf{x}_i\Delta\epsilon_i - \Delta\mathbf{x}^*_{i,k}\Delta\epsilon^*_{i,k}\Big|\Big|^2 \qquad (27)
\end{aligned}$$

where $\Delta\mathbf{x}^*_{i,k}\Delta\epsilon^*_{i,k}$ denotes $\Delta\mathbf{x}_i\Delta\epsilon_i$ with ζ_{i-k} replaced by its IID copy ζ^*_{i-k}. The last inequality (27) is due to the Jensen's Inequality. Thus, by (27),

$$||\psi_{d,k}||^2 \le \frac{1}{n^2}\sum_{i=1}^{n} \delta_2^2(\mathbf{L}_d, k) = \frac{1}{n}\delta_2^2(\mathbf{L}_d, k) \qquad (28)$$

Moreover,

$$\psi_d - \mathbb{E}\psi_d = \sum_{k=0}^{\infty} \mathcal{P}_{i-k}(\psi_d) = \sum_{k=0}^{\infty} \psi_{d,k} \qquad (29)$$

Thus, by (28) and **Assumption 5**,

$$||\psi_d - \mathbb{E}\psi_d|| \le \sum_{k=0}^{\infty} ||\psi_{d,k}|| = \frac{1}{\sqrt{n}} \sum_{k=0}^{\infty} \delta_2(\mathbf{L}_d, k) = O\Big(\frac{1}{\sqrt{n}}\Big) \tag{30}$$

By $\psi_d = \frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\epsilon_i$ and $\mathbb{E}\psi_d = 0$, (30) leads to

$$\Big|\Big|\frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\epsilon_i\Big|\Big| = O\Big(\frac{1}{\sqrt{n}}\Big)$$

Thus, the lemma follows.

□

7.1 Proof of Theorem 1

Recall the differencing estimator $\hat{\boldsymbol{\beta}}_D$ given by (10).

$$\begin{aligned} \hat{\boldsymbol{\beta}}_D &= \Big(\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\mathbf{x}_i^\top\Big)^{-1} \sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta y_i \\ &= \Big(\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\mathbf{x}_i^\top\Big)^{-1} \sum_{i=1}^{n} \Delta\mathbf{x}_i \Big(\Delta\mathbf{x}_i^\top \boldsymbol{\beta} + O(1/n) + \Delta\epsilon_i\Big) \end{aligned}$$

which leads to

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_D - \boldsymbol{\beta}) = O\Big(\frac{1}{\sqrt{n}}\Big)\Big(\frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\mathbf{x}_i^\top\Big)^{-1} \frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i + \Big(\frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\mathbf{x}_i^\top\Big)^{-1} \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\epsilon_i \tag{31}$$

Moreover, by **Lemma 1**, $\Delta\mathbf{x}_i$ is *local-stationary*. Thus, for all $i = 1, \cdots, n$,

$$|\Delta\mathbf{x}_i| = O_{\mathbb{P}}(1)$$

which also leads to

$$\Big|\frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i\Big| = O_{\mathbb{P}}(1) \tag{32}$$

Similarly, by the local-stationarity of $\Delta\mathbf{x}_i$,

$$\Big|\frac{1}{n}\sum_{i=1}^{n} \Delta\mathbf{x}_i \Delta\mathbf{x}_i^\top\Big| = O_{\mathbb{P}}(1) \tag{33}$$

where $|\cdot|$ is a matrix norm induced by the Euclidean norm on a vector. By applying (32), (33), **Assumption 2** and **Lemma 2** to (31), the theorem follows.

$\square$

Lemma 3. *Under Assumption 1 – Assumption 8,*

$$\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)(\epsilon_i-\sigma_\epsilon Z_i)\Big| = o_{\mathbb{AS}}\Big(\frac{\log(n)}{n^{3/4}b_n}\Big) \tag{34}$$

where $\sigma_\epsilon^2 = \sum_{j\in\mathbb{Z}}\mathbb{E}(\epsilon_0\epsilon_j)$ *is the long-run variance of* ϵ_i *and* $w_n(t,i)$ *is the weight for the local-linear regression. Here* Z_i *is an IID standard normal random variable.*

proof) First, let $Z_i = \mathbb{B}(i)-\mathbb{B}(i-1)$. Then,

$$\begin{aligned}
&\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\epsilon_i - \sigma_\epsilon\sum_{i=1}^{n} w_n(t,i)Z_i\Big| \\
=\ & \sup_{0\le t\le 1}\Big|w_n(t,n)\sum_{j=1}^{n}(\epsilon_j-\sigma_\epsilon Z_j) - \sum_{j=2}^{n}[w_n(t,j)-w_n(t,j-1)]\sum_{i=1}^{j}(\epsilon_i-\sigma_\epsilon Z_i)\Big| \\
\le\ & \sup_{0\le t\le 1}\Big\{\Big|w_n(t,n)\Big|\Big|\sum_{j=1}^{n}(\epsilon_j-\sigma_\epsilon Z_j)\Big| + \sum_{j=2}^{n}\Big|w_n(t,j)-w_n(t,j-1)\Big|\Big|\sum_{i=1}^{j}(\epsilon_i-\sigma_\epsilon Z_i)\Big|\Big\} \\
\le\ & \sup_{0\le t\le 1}\Big\{\max_{j\le n}\Big|\sum_{i=1}^{j}\epsilon_i-\sigma_\epsilon\mathbb{B}(j)\Big|\cdot\Big[|w_n(t,n)|+\sum_{j=2}^{n}|w_n(t,j)-w_n(t,j-1)|\Big]\Big\} \\
\le\ & \max_{i\le n}\Big|\sum_{k=1}^{i}\epsilon_k-\sigma_\epsilon\mathbb{B}(i)\Big|\cdot\sup_{0\le t\le 1}\Big[|w_n(t,n)|+\sum_{j=2}^{n}|w_n(t,j)-w_n(t,j-1)|\Big]
\end{aligned} \tag{35}$$

where the first equality is due to the *summation-by-parts* formula. By **Assumption 4**, error ϵ_i is stationary with $\mathbb{E}\epsilon_i = 0$ and $\mathbb{E}|\epsilon_i|^4<\infty$. Thus, by the strong invariance principle in Wu (2007),

$$\max_{i\le n}\Big|\sum_{k=1}^{i}\epsilon_k-\sigma_\epsilon\mathbb{B}(i)\Big| = o_{\mathbb{AS}}\Big(n^{1/4}\log(n)\Big) \tag{36}$$

Moreover, a simple computation shows that the following equality holds for the local-linear regression weight $w_n(t,j)$:

$$\sup_{0\le t\le 1}\Big[|w_n(t,n)|+\sum_{j=2}^{n}|w_n(t,j)-w_n(t,j-1)|\Big] = O\Big(\frac{1}{nb_n}\Big) \tag{37}$$

Then, by applying (36) and (37) to (35), we have

$$\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\epsilon_i - \sigma_\epsilon \sum_{i=1}^{n} w_n(t,i)Z_i\Big| = o_{\mathbb{AS}}\Big(\frac{n^{1/4}\log(n)}{nb_n}\Big)$$

which leads to the lemma.

□

Lemma 4. *Let $k_n \asymp n^{1/3}$. Then, under Assumption 1 – Assumption 8,*

$$\hat{\sigma}_\epsilon^2 = \sigma_\epsilon^2 + O_{\mathbb{P}}(n^{-1/3}) \tag{38}$$

where $\hat{\sigma}_\epsilon^2$ is the subseries variance estimate of σ_ϵ^2 introduced by (21).

proof) For notational convenience, we let $y_i^* = y_i - \mathbf{x}_i^\top \boldsymbol{\beta}$. Then, a simple computation shows that we can rewrite $\hat{\sigma}_\epsilon^2$ as the following:

$$\begin{aligned}\hat{\sigma}_\epsilon^2 \;=\; & \frac{1}{2(m-1)k_n}\sum_{i=1}^{m-1}\Big[\sum_{j=1}^{k_n}\Big(y^*_{j+ik_n} - y^*_{j+(i-1)k_n}\Big)\Big]^2 \\ & + \frac{1}{2(m-1)k_n}\sum_{i=1}^{m-1}\Big[\sum_{j=1}^{k_n}\Big(\mathbf{x}_{j+ik_n} - \mathbf{x}_{j+(i-1)k_n}\Big)^\top(\hat{\boldsymbol{\beta}}_D - \boldsymbol{\beta})\Big]^2 \\ & - \frac{1}{(m-1)k_n}\sum_{i=1}^{m-1}\Big[\sum_{j=1}^{k_n}\Big(y^*_{j+ik_n} - y^*_{j+(i-1)k_n}\Big)\sum_{j=1}^{k_n}\Big(\mathbf{x}_{j+ik_n} - \mathbf{x}_{j+(i-1)k_n}\Big)^\top(\hat{\boldsymbol{\beta}}_D - \boldsymbol{\beta})\Big]\end{aligned} \tag{39}$$

By Carlstein (1986) and Wu & Zhao (2007), we have

$$\frac{1}{2(m-1)k_n}\sum_{i=1}^{m-1}\Big[\sum_{j=1}^{k_n}\Big(y^*_{j+ik_n} - y^*_{j+(i-1)k_n}\Big)\Big]^2 = \sigma_\epsilon^2 + O_{\mathbb{P}}(n^{-1/3}) \tag{40}$$

Moreover, by $k_n \asymp n^{1/3}$, **Assumption 2** and **Theorem 1**, we have

$$\frac{1}{2(m-1)k_n}\sum_{i=1}^{m-1}\Big[\sum_{j=1}^{k_n}\Big(\mathbf{x}_{j+ik_n} - \mathbf{x}_{j+(i-1)k_n}\Big)^\top(\hat{\boldsymbol{\beta}}_D - \boldsymbol{\beta})\Big]^2 = O_{\mathbb{P}}(n^{-2/3}) \tag{41}$$

Furthermore, by (40), (41) and the Cauchy-Schwarz inequality,

$$\frac{1}{(m-1)k_n}\sum_{i=1}^{m-1}\Big[\sum_{j=1}^{k_n}\Big(y^*_{j+ik_n} - y^*_{j+(i-1)k_n}\Big)\sum_{j=1}^{k_n}\Big(\mathbf{x}_{j+ik_n} - \mathbf{x}_{j+(i-1)k_n}\Big)^\top(\hat{\boldsymbol{\beta}}_D - \boldsymbol{\beta})\Big] = O_{\mathbb{P}}(n^{-1/3}) \tag{42}$$

Thus, by applying (40)–(42) to (39), the lemma easily follows.

□

7.2 Proof of Theorem 2

Recall the two-stage local linear regression estimator in (15). By a Taylor's expansion on $\mu(\cdot)$, we can show the following:

$$\begin{aligned}
\tilde{\mu}(t) &= \sum_{i=1}^{n} w_n(t,i)(y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}_D) \\
&= \sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_D) + \sum_{i=1}^{n} w_n(t,i)(\mu(i/n) + \epsilon_i) \\
&= \sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_D) + \sum_{i=1}^{n} w_n(t,i)\Big[\mu(t) + (i/n - t)\mu'(t_0) + \epsilon_i\Big]
\end{aligned}$$

where $t_0 \in (t,\ i/n)$. Thus, from $\sum_{i=1}^{n} w_n(t,i) = 1$, we have

$$\tilde{\mu}(t) - \mu(t) = \sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_D) + \sum_{i=1}^{n} w_n(t,i)(i/n - t)\mu'(t_0) + \sum_{i=1}^{n} w_n(t,i)\epsilon_i \tag{43}$$

By **Assumption 2** and **Theorem 1**, for any fixed $t \in (0,1)$,

$$\Big|\sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_D)\Big| = O_{\mathbb{P}}(n^{-1/2}) \tag{44}$$

Moreover, by **Assumption 6** and **Assumption 7**,

$$\Big|\sum_{i=1}^{n} w_n(t,i)(i/n - t)\mu'(t_0)\Big| = O(b_n) \tag{45}$$

for any fixed $t \in (0,1)$. Furthermore, by **Assumption 4**,

$$\Big|\sum_{i=1}^{n} w_n(t,i)\epsilon_i\Big| = O_{\mathbb{P}}\Big(\frac{1}{\sqrt{nb_n}}\Big) \tag{46}$$

for any fixed $t \in (0,1)$. Therefore, by applying (44)–(46) to (43), the theorem easily follows.

□

7.3 Proof of Theorem 3

From equation (1), we have $y_i - \mathbf{x}_i^\top \boldsymbol{\beta} = \mu(i/n) + \epsilon_i$, where ϵ_i satisfies **Assumption 4**. Since $\mathbf{x}_i$ is local-stationary by **Assumption 3**, we have the following:

$$\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top\Big| = O_{\mathbb{P}}(1) \tag{47}$$

Then, we do a Taylor's expansion on $\mu(\cdot)$ to show the following:

$$\begin{aligned}
& \sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)[y_i - \mathbf{x}_i^\top\hat{\boldsymbol{\beta}}_D - \mu(t)] - \sigma_\epsilon\sum_{i=1}^{n} w_n(t,i)Z_i\Big| \\
\le\ & \sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_D) + \sum_{i=1}^{n} w_n(t,i)\epsilon_i - \sigma_\epsilon\sum_{i=1}^{n} w_n(t,i)Z_i \\
& \qquad + \sum_{i=1}^{n} w_n(t,i)[(t-i/n)\mu'(t) + C(t-i/n)^2\mu''(t_0)]\Big| \\
\le\ & \sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_D)\Big| + \sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\epsilon_i - \sigma_\epsilon\sum_{i=1}^{n} w_n(t,i)Z_i\Big| \\
& \quad + \sup_{0\le t\le 1}\Big|\sum_{i=1}^{n}(t-i/n)w_n(t,i)\mu'(t)\Big| + C\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n}(t-i/n)^2 w_n(t,i)\mu''(t_0)\Big| \\
\le\ & \Big|\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_D\Big|\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\mathbf{x}_i^\top\Big| + \sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\epsilon_i - \sigma_\epsilon\sum_{i=1}^{n} w_n(t,i)Z_i\Big| \\
& \quad + b_n\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\mu'(t)\Big| + Cb_n^2\sup_{0\le t\le 1}\Big|\sum_{i=1}^{n} w_n(t,i)\mu''(t_0)\Big| \\
=\ & O_{\mathbb{P}}(n^{-1/2}) + o_{\mathbb{AS}}\Big(\frac{\log(n)}{n^{3/4}b_n}\Big) + O(b_n) + O(b_n^2) \\
=\ & O_{\mathbb{P}}\Big(n^{-1/2} + \frac{\log(n)}{n^{3/4}b_n} + b_n\Big)
\end{aligned} \tag{48}$$

where $t_0 \in (t,\ i/n)$ and C is some constant. The third inequality comes from that the weight $w_n(t,i)$ involves $K\Big(\frac{t-i/n}{b_n}\Big)$ from (14) and that $K(\cdot)$ has a compact support $[-1,\ 1]$ by **Assumption 7**. The second last equality follows from (47), **Theorem 1**, **Lemma 3** and **Assumption 6**. Thus, by applying (15) and $\sum_{i=1}^{n} w_n(t,i) = 1$ to (48), we have the following:

$$\sqrt{nb_n} \sup_{0 \le t \le 1} \left| \tilde{\mu}(t) - \mu(t) - \sigma_\epsilon \sum_{i=1}^{n} w_n(t,i) Z_i \right| = O_{\mathbb{P}}\Big(b_n^{1/2} + \frac{\log(n)}{n^{1/4} b_n^{1/2}} + n^{1/2} b_n^{3/2} \Big)$$

Therefore, by **Assumption 8**,

$$\sqrt{nb_n} \sup_{0 \le t \le 1} \left| \tilde{\mu}(t) - \mu(t) - \sigma_\epsilon \sum_{i=1}^{n} w_n(t,i) Z_i \right| = o_{\mathbb{P}}(1)$$

□

References

[1] Ball, L. & Mankiw, N.G. (2002). The NAIRU in Theory and Practice. *Journal of Economic Perspectives* **16**, 115–136.

[2] Bickel, P.J. & Rosenblatt, M. (1973). On Some Global Measures of the Deviations of Density Function Estimates. *Annals of Statistics* **1**, 1071–1095.

[3] Cai, Z. (2007). Trending Time-varying Coefficient Time Series Models with Serially Correlated Errors. *Journal of Econometrics* **136**, 163–188.

[4] Carlstein, E. (1986). The Use of Subseries Values for Estimating the Variance of a General Statistic from A Stationary Sequence. *The Annals of Statistics* **14**, 1171–1179.

[5] Craven, P. & Wahba, G. (1979). Smoothing Noisy Data with Spline Functions. *Numerische Mathematik* **31**, 377–403.

[6] Cressie, N.A.C. (1993). *Statistics for Spatial Data*. Wiley, New York.

[7] Eubank, R.L. & Speckman, P.L. (1993). Confidence Bands in Nonparametric Regression. *Journal of the American Statistical Association* **88**, 1287–1301.

[8] Fair, R.C. (2000). Testing the NAIRU Model for the United States. *The Review of Economics and Statistics* **82**, 64–71.

[9] Fan, J. & Gijbels, I. (1996). *Local Polynomial Modelling and its Applications.* Chapman and Hall, London.

[10] Fan, J. & Zhang, W. (2000). Simultaneous Confidence Bands and Hypothesis Testing in Varying-Coefficient Models. *Scandinavian Journal of Statistics* **27**, 715–731.

[11] Gao, J. & Hawthorne, K. (2006). Semiparametric Estimation and Testing of the Trend of Temperature Series. *Econometrics Journal* **9**, 332–355.

[12] González, A., Hubrich, K. & Teräsvirta, T. (2009). Forecasting Inflation with Gradual Regime Shifts and Exogenous Information. manuscript.

[13] González, A. & Teräsvirta, T. (2008). Modelling Autoregressive Processes with a Shifting Mean. *Studies in Nonlinear Dynamics & Econometrics* **12**, 1–28.

[14] Gordon, R.J. (1997). The Time-varying NAIRU and its implications for Economic Policy. *Journal of Economic Perspectives* **11**, 11-32.

[15] Gordon, R.J. (1998). Foundations of the Goldilocks Economy: Supply Shocks and the Time-varying NAIRU. *Brookings Papers on Economic Activity* **2**, 197–333.

[16] Hall, P., Kay, J.W. & Titterington, D.M. (1990). Asymptotically Optimal Difference-Based Estimation of Variance in Nonparametric Regression. *Biometrika* **77**, 521-528.

[17] Härdle, W.K. (1989). Asymptotic Maximal Deviation of M-Smoothers. *Journal of Multivariate Analysis* **29**, 163–179.

[18] Härdle, W.K., Liang, H. & Gao, J. (2000). *Partially Linear Models.* Heidelberg, Physica-Verlag.

[19] Härdle, W.K. & Song, S. (2009). Confidence Bands in Quantile Regression. *Econometric Theory* **69**, 1–21.

[20] Johnston, G.J. (1982). Probabilities of Maximal Deviations for Nonparametric Regression Function Estimates. *Journal of Multivariate Analysis* **12**, 402–414.

[21] Kim, K.H. (2011). Counter-cyclical Risk Aversion. Michigan State University, manuscript.

[22] Kim, K.H., Zhou, Z. & Wu, W.B. (2010). Non-stationary Structural Model with Time-varying Demand Elasticities. *Journal of Statistical Planning and Inference* **140**, 3809–3819.

[23] Komlós, J., Major, P. & Tusnády, G. (1975). An Approximation of Partial Sums of Independent RV's and the Sample DF I. *Z. Wahrsch. Ver. Geb.* **32**, 111-131.

[24] Liu, W. & Wu, W.B. (2010). Simultaneous Nonparametric Inference of Time Series. *The Annals of Statistics* **38**, 2388–2421.

[25] Robinson, P.M. (1988). Root-N-Consistent Semiparametric Regression. *Econometrica* **56**, 931–954.

[26] Staiger, D., Stock, J.H. & Watson, M.W. (1996). How Precise Are Estimates of the Natural Rate of Unemployment? *NBER Working Paper No. 5477.*

[27] Staiger, D., Stock, J.H. & Watson, M.W. (1997). The NAIRU, Unemployment and Monetary Policy. *Journal of Economic Perspectives* **11**, 33-49.

[28] Staiger, D., Stock, J.H. & Watson, M.W. (2001). Prices, Wages, and the U.S. NAIRU in the late 1990s. *The Roaring Nineties: Can Full Employment Be Sustained?* Alan B. Krueger and Robert Solow, eds. New York.

[29] Wahba, G. (1977). A Survey of Some Smoothing Problems and the Method of Generalized Cross-Validation for Solving Them. In *Applications of Statistics* (P.R. Krishnaiah, ed.). North-Holland, Amsterdam, pp. 507–523.

[30] Wu, W.B. (2007). Strong Invariance Principles for Dependent Random Variables. *The Annals of Probability* **35**, 2294–2320.

[31] Wu, W.B. & Zhao, Z. (2007). Inference of Trends in Time Series. *Journal of the Royal Statistical Society, Series B.* **69**, 391–410.

[32] Yatchew, A. (1997). An Elementary Estimator of the Partial Linear Model. *Economics Letters* **57**, 135–143.

[33] Yatchew, A. (1998). Nonparametric Regression Techniques in Economics. *Journal of Economic Literature* **36**, 669–721.