UNIVERSITY OF SOUTHAMPTON

FUNCTIONAL FORM MISSPECIFICATION IN REGRESSIONS WITH INTEGRATED TIME SERIES

Ioannis Kasparis

Doctor of Philosophy

Department of Economics

May, 2004

"When logic and proportion
have fallen sloppy dead,
and the White Knight is talking backwards
and the Red Queen's off with her head!
Remember what the dormouse said:
Feed your head."
[White Rabbit]

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCES

ECONOMICS

Doctor of Philosophy

FUNCTIONAL FORM MISSPECIFICATION IN REGRESSIONS WITH INTEGRATED TIME SERIES

by

Ioannis Kasparis

This thesis addresses the issue of functional form misspecification in models with nonstationary covariates. In particular we assume that the variables of the model are unit root processes. First we examine the asymptotic behaviour of the least squares estimator, under functional form misspecification, in regression models like those analysed by Park and Phillips (2001) in their recent paper in *Econometrica*. In contrast to the stationary case, we find that convergence to some pseudo-true value does not always hold. In some cases the estimator diverges. Whenever the estimator converges the order of consistency is usually slower and the limit distribution theory different than the one under correct specification. Moreover a conditional moment test for functional form is considered within the theoretical framework of Park and Phillips (2001). In contrast to the stationary case, under nonstationarity the test may be two-sided. The asymptotic power of the test is derived against a set of alternatives where each alternative is characterised by the asymptotic order of the true specification. Moreover it is shown that the use of integrable weighting functions in the construction of the test statistic improves asymptotic power against a set of alternatives. Next the test for functional form is extended to cointegrating relationships. Our framework allows for a fitted model that is possibly nonlinear in variables and in view of this the linear specifications commonly used in practice constitute a special case. The test is consistent under both functional form misspecification and no cointegration. So the functional form test can be also used as a cointegration test. In both cases the divergence rate attained equals n/M, with n and M being the sample size and the bandwidth used in the estimation of long-run covariance matrices respectively.

Contents

Acknowledgements

I would like to thank my parents and my brother for their support and encouragement. I am greatly indebted to Grant Hillier for his support and for suggesting to me this area of research. I would also like to thank Peter Phillips for his invaluable advice. Many thanks to Tassos Magdalinos, John Aldrich, Ray O' Brien, Jean-Yves Pitarakis and Jan Podivinsky. Financial support from the Fuculty of Social Sciences of Southampton University is gratefully acknowledged.

Chapter 1
Introduction

The issue of functional form misspecification in statistics and econometrics has been addressed as early as the 1960's (Ramsey (1969)). The properties of estimators under functional form misspecification (FFM) have been studied in detail (e.g. White (1981)) and several functional form tests have been proposed over time. These testing procedures have been developed for independent and identically data (i.i.d.) but can be extended to weakly dependent data (see Bierens, (1990)). The purpose of this dissertation will be to extend some of these results to models with nonstationary regressors. In particular we will assume that covariates are unit root processes.

In order to address the issue of FFM, we need to depart from the standard linear specifications. Our theoretical framework is therefore naturally nonlinear, closely related to that of Park and Phillips (1999, 2001). However although the asymptotic properties of estimators for nonlinear models with stationary and weakly dependent data have been explored almost twenty years ago (e.g. Hansen (1981), White and Domowitz (1984)), no well developed limit distribution theory had been developed for nonlinear models with strongly dependent regressors until the recent development of Park and Phillips (1999, 2001). Our aim is to use these theoretical developments to analyse misspecified, nonlinear models with nonstationary covariates.

The thesis consists of three main chapters. Although they are related the focus of each chapter is different. The theoretical framework of the first two main chapters is the same. In the first two main chapters we consider single covariate models nonlinear in parameters and in variable. In the third main chapter we consider multi-factor models linear in parameters but nonlinear in variables. The focus and findings of each chapter are summarised below.

In the first main chapter (Chapter 2) we provide results for the asymptotic behaviour of the Nonlinear Least Squares (NLS) estimator in models with a single covariate. The theoretical framework is the same as that of Park and Phillips (2001), P&P hereafter. It is well known that for stationary models, under FFM, the NLS estimator has a well defined limit, referred to in the econometric literature as "pseudo-true" value. Moreover the NLS estimator about the pseudo-true value and scaled by $\sqrt{n}$ (n is the sample size) has a Gaussian limit distribution. So, in the stationary case under FFM, the NLS estimator has a well defined limit, the limit distribution is still Gaussian (although the variance is larger), and the convergence rate is unaffected.

We show that when the covariate is a unit root process, things may be quite different. First, when the parameter space is unbounded, convergence to some pseudo-true value does not always hold. In some cases the estimator is unbounded in probability. When the parameter space is bounded and the fitted model is of different asymptotic order than the true one, the estimator converges to a boundary point of the parameter space. In this case the limit objective function is not minimised at a turning point and therefore techniques used to obtain limit distribution results and convergence rates, when the limit is on a boundary (Andrews (1999), Phillips and Moon (2003)), are not applicable. When the true model is of the same order of magnitude as the fitted one, limit distribution results and convergence rates are obtained. When confined to the *I-regular* family, the limit distribution and the convergence rates under FFM are the same as those under correct functional form. The limit distribution under FFM is still mixed normal but with larger variance. So this is analogous to the stationary case. If both the true and fitted models are H-regular, convergence rates are slower and the limit distribution is different than the one under correct specification. The limit distribution involves only functionals of Brownian motion and not stochastic integrals.

The second main chapter (Chapter 3) focuses on testing. A conditional moment test for functional form is considered. We derive the limit distribution of the test under the null hypothesis and the asymptotic power rates under the alternative hypothesis. The theoretical framework again is the same as the one of P&P. Three estimation procedures are considered: NLS, an Instrumental Variables (IV) kind of estimator and the Efficient Nonstationary NLS (EN-NLS) proposed by Chang, Park and Phillips (2001). The asymptotic power of the test is obtained against a set of alternatives with each alternative charecterised by the asymptotic order of the true specification. Moreover we show that the use of *I-regular* weighting functions in the construction of the test statistic improves asymtotic power in some cases. A potential application within this framework is Park's (2002) nonlinear nonstationary stochastic volatility models.

In the third main chapter (Chapter 4), the conditional moment test is considered when the fitted model is linear in parameters and involves more than one covariate. The framework of Park and Phillips postulates that the regressors are exogenous. In order to make the test applicable to cointegrating relationships, the exogeneity assumption about the regressors is relaxed. Endogeneity and dependence structure is introduced by assuming that the errors of the model and the errors that drive the unit root variables is a vector linear process. We obtain limit distribution theory for sample moments under these conditions. From P&P we know that sample moment asymptotics (covariance) are determined by stochastic integrals. In our case extra terms appear that involve long-run covariance matrices and functionals of Brownian motion. To induce a chi-squared limit distribution under the null, a correction term is introduced in the test statistic and the parameters are estimated by a Fully Modified Least Squares type of estimator. The test is consistent under FFM and no cointegration. So the functional form test can be also used as a cointegration test. We show that in both cases the test diverges at a rate of n/M, with M being the bandwidth of the kernel used for the estimation of long-run covariances. The rate attained under the alternative is the same with the one of the CUSUM test for cointegration proposed by Xiao and Phillips (2002).

Before we proceed to the next chapter, we will clarify some issues regarding the presentation of definitions, assumptions and results. Our technical results will be referred to as "Propositions". Results from other authors are clearly indicated as such. The proofs of all Propositions are given in the Appendix to the chapter where the Proposition is stated. Each main chapter is followed by an Appendix relating to the particular chapter. Moreover definitions, assumptions and results are numbered with respect to the section in which they are stated but not with respect to the relevant chapter. In each chapter numbering restarts.

Chapter 2
Consequences of Functional Form Misspecification in Regressions with Integrated Time Series

2.1 Introduction

In this chapter we examine the behaviour of the NLS estimator under functional form misspecification, when both the true and the fitted models involve nonstationary covariates. In particular we will assume that the covariates are unit root processes. This work relies heavily on the developments made recently in a sequence of papers by Park and Phillips. Park and Phillips (1999) provide limit distribution theory for nonlinear transformations of a unit root process. Park and Phillips (2001) extend their earlier results to models which are nonlinear in parameters and Chang, Park and Phillips (2001) provide limit distribution theory for multiple regression models.

Our theoretical framework is the same as the one of P&P. We assume that the functional form of both the true and fitted models belongs to either the *I-regular* or *H-regular* family of transformations defined in P&P. One of our main findings is that convergence to a pseudo-true value does not always hold. In some cases the estimators diverge. We know from P&P that under correct functional form specification, the limit distribution theory is mixed normal for *I-regular* models and involves stochastic integrals for *H-regular* models. We will show that under functional form misspecification the limit distribution theory may be different. The convergence rates usually depend on the asymptotic order of the true model relative to that of the fitted model, and can be slower than the convergence rates attained under correct functional form specification.

The results provided in this chapter are useful for the development of specification tests. The alternative hypothesis of testing procedures like those proposed by Ramsey (1969), White (1981) and Bierens (1990) is that the fitted model in incorrectly specified. So our analysis is useful in determining the asymptotic power properties of these testing procedures.

The rest of this chapter is organised as follows: Section 2.2 briefly reviews the properties of the NLS estimator under functional form misspecification when the regressors are stationary. In Section 2.3 the theoretical framework is specified in detail and some useful results due to Park and Phillips (1999, 2001) and de Jong (2002) are provided. Section 2.4 provides consistency and limit distribution results and Section 2.5 concludes.

Before we proceed to the next section we introduce some notation. For a vector $x = (x_i)$ or a matrix $A = (a_{ij})$, $|x|$ and $|A|$ denotes the vector and matrix respectively of the moduli of their elements. The maximum of the moduli is denoted as $||.||$. Moreover $||.||$ may denote the supremum of function (possibly vector-valued or matrix-valued) over some set K , say i.e. $\sup_K ||f|| = ||f||_K$. For a function $g : \mathbf{R}^p \to \mathbf{R}$ define the arrays

$$\dot{g} = \left(\frac{\partial g}{\partial a_i}\right), \ddot{g} = \left(\frac{\partial^2 g}{\partial a_i \partial a_j}\right), \dddot{g} = \left(\frac{\partial^3 g}{\partial a_i \partial a_j \partial a_k}\right)$$

which will be assumed to be vectors arranged by the lexicographic ordering of their indices. Sometimes is more convenient to express the second derivatives of g in matrix form i.e. $\ddot{G} = \partial^2 g / \partial a \partial a'$. Moreover $1\{A\}$ will denote the indicator function of a set A.

2.2 Functional Form Misspecification in the Stationary Framework

As mentioned above, under incorrect functional form the NLS estimator converges to a well defined limit referred to in the literature as the "pseudo-true value". Moreover the NLS estimator has a Gaussian limit distribution around the pseudo-true value. In this section we briefly show how these results can be obtained. For easier exposition, in this section only we will confine ourselves to independent and identically distributed (*i.i.d.*) data. The same type of results can be obtained assuming weakly dependent data (see White (1984) and Bierens (1990)).

Let the true model be

$$y_t = f(x_t) + u_t \tag{1}$$

where $\{x_t\}_{t=1}^{n}$ and $\{u_t\}_{t=1}^{n}$ are *i.i.d.* sets of random variables and the distribution function of $\{x_t\}_{t=1}^{n}$ is $F(x)$. The function $f(.) : \mathbf{R} \to \mathbf{R}$. Assume that the fitted model is:

$$\hat{y}_t = g(x_t, \hat{a}) + \hat{u}_t \tag{2}$$

where $g : \mathbf{R} \times \mathbf{A} \to \mathbf{R}$ with A being a compact subset of $\mathbf{R}^m$. Define the NLS estimator $\hat{a}$ as

$$\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2 \tag{3}$$

Moreover let $c(x_t, a) = (y_t - g(x_t, a))^2$ and define $Q_n(a)$ as:

$$Q_n(a) = \sum_{t=1}^{n} c(x_t, a)$$

Convergence to pseudo-true value can be established from the following theorem due to Jennrich (1969):

THEOREM 2.1: (Jennrich (1969))

If $n^{-1}Q_n(a) \rightarrow Q(a)$ *in probability* $(a.s.)$ *uniformly in* a *as* $n \rightarrow \infty$ *, and* $Q(a)$ *is continuous in* a *and has a unique minimum at* a^* $a.s$*, then* $\hat{a} \rightarrow a^*$ *in probability* $(a.s.)$.

Theorem 2.1. is used by Jennrich (1969) to prove that the NLS estimator converges to the parameter of interest under correct functional form specification. The following result due White (1981) exploits Theorem 2.1 to establish convergence of the NLS estimator under incorrect functional form to a well defined limit:

THEOREM 2.2: (White (1981))

Let the true and the fitted model be given by (1) and (2) respectively. Moreover assume that:

(a) $g(x, a)$ *is a continuous function of* a *for each* x *in* X *and* a *in* A.

(b) $(f(x) - g(x, a))^2 \leq m(x)$ *for all* x *and* a *in* A*, where* $m(x)$ *is integrable with respect to* $F(x)$.

(c) a^* *uniquely minimises* $Q(a) = \int_{-\infty}^{\infty} (f(x) - g(x, a))^2 \, dF(x)$.

Then $a \overset{a.s.}{\rightarrow} a^*$*, as* $n \rightarrow \infty$.

It is obvious from a simple application of the law of large numbers that

$$\frac{1}{n}\sum_{t=1}^{n}(y_t - g(x_t, a))^2 \overset{a.s.}{\rightarrow} \int_{-\infty}^{\infty} (f(x) - g(x, a))^2 \, dF(x)$$

pointwise in a. Conditions (a) and (b) in Theorem 2.2 together with the compactness assumption ensure that the convergence is uniform over A. So under these conditions $\hat{a}$ is consistent for a^*, the value that minimises the mean square error.

Given the convergence of the NLS estimator, the limit distribution result can be obtained by a mean value expansion of the first derivative of the objective function around the pseudo-true value. $\dot{Q}_n(\hat{a})$ can be written as:

$$\dot{Q}_n(\hat{a}) = \dot{Q}_n(a^*) + \ddot{Q}_n(\bar{a})(\hat{a} - a^*),$$

where $\|\bar{a} - a^*\| \leq \|\hat{a} - a^*\|$. Provided that $\ddot{Q}_n(\bar{a})$ is invertible the expression above can be written as:

$$\sqrt{n}(\hat{a} - a^*) = \left[\frac{1}{n}\ddot{Q}_n(\bar{a})\right]^{-1} \frac{1}{\sqrt{n}}\dot{Q}_n(a^*) \tag{4}$$

Under conditions similar to the conditions (a) and (b) of Theorem 2.2 it follows that the term:

$$\frac{1}{n}\ddot{Q}_n(\bar{a}) \xrightarrow{a.s.} \ddot{Q}(a^*) = \mathbf{E}(\ddot{c}(x_t, a^*)). \tag{5}$$

Because a^* minimises $\mathbf{E}(c(x_t, a))$, under the extra assumptions that a^* is an interior point of A and $|\dot{c}(x, a^*)| \leq d(x)$ with $d(x)$ being integrable with respect to $F(x)$, it follows that $\mathbf{E}(\dot{c}(x_t, a^*)) = 0$. Hence from the Lindeberg-Lèvy central limit theorem:

$$\frac{1}{\sqrt{n}}\dot{Q}_n(a^*) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\dot{c}(x_t, a^*) \xrightarrow{d} N(0, \mathbf{E}(\dot{c}(x_t, a^*)\dot{c}(x_t, a^*)') \tag{6}$$

From (4), (5) and (6) follows that $\sqrt{n}(\hat{a} - a^*) \xrightarrow{d} N(0, \mathbf{\Sigma})$, with

$$\mathbf{\Sigma} = \mathbf{E}[\ddot{c}(x_t, a^*)]^{-1}\mathbf{E}[\dot{c}(x_t, a^*)\dot{c}(x_t, a^*)']\mathbf{E}[\ddot{c}(x_t, a^*)]^{-1}.$$

which establishes the aforementioned result.

2.3 Definitions, Assumptions and Preliminary Results

The models we consider are the same as those discussed in P&P. We assume that the series $\{y_t\}_{t=1}^n$ is generated by a model of the following form:

$$y_t = f(x_t) + u_t \tag{7}$$

where x_t is a unit root process, the function $f(.) : \mathbf{R} \to \mathbf{R}$ and u_t is a martingale difference. Assume that the fitted model is:

$$\hat{y}_t = g(x_t, \hat{a}) + \hat{u}_t \tag{8}$$

where $g(., a)$ is a transformation of the data "different from $f(.)$". This is defined precisely in the next section.

This section provides a set of definitions that specify the kind of models under consideration. First we specify the process that generates the covariates. Secondly, we specify the kinds of functions that define the functional form of the model, and finally we give a precise definition for the error term. Our theoretical framework is very similar to the one of P&P so most of the definitions outlined here are the same as those in P&P. Some useful results due to Park and Phillips (1999, 2001), de Jong (2002a) and Jeganathan (2003) are also provided.

We assume throughout that the sequence $\{x_t\}_{t=1}^n$ is a unit root process generated by $x_t = x_{t-1} + v_t$, with $x_0 = 0$[1]. v_t is assumed to have an infinite moving average representation:

---

[1] This initial condition can be replaced by $x_o = O_p(1)$ (see P&P).

$$v_t = \psi(L)\eta_t = \sum_{k=0}^{\infty} \psi_k \eta_{t-k},$$

with $\psi(1) \neq 0$ and $\{\eta_t\}_{t=1}^{n}$ a sequence of independent and identically distributed random variables with zero mean. The following assumption will be made in the rest of this chapter:

ASSUMPTION 3.1:

(a) $\sum_{k=0}^{\infty} k\,|\psi_k| < \infty$ *and* $\mathbf{E}(\eta_t)^p < \infty$ *for some* $p > 2$.

(b) η_t *has absolutely continuous distribution with respect to Lebesgue measure and its characteristic function* $\phi(s)$ *is such that* $\lim_{s\to\infty} s^r\phi(s) = 0$ *for some* $r > 0$.

Define the stochastic process:

$$V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} v_t$$

$V_n(r)$ takes values in the set of cadlag functions on the interval [0,1]. Moreover define the partial sum process of the errors u_t of the model:

$$U_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t$$

We assume that the processes $U_n(r)$ and $V_n(r)$ satisfy the following assumption:

ASSUMPTION 3.2:

(a) $(U_n(r), V_n(r)) \xrightarrow{d} (U, V)$, *where* (U, V) *is a vector Brownian motion,*

(b) $(u_t, \mathcal{F}_{n.t-1})$ *is a martingale difference sequence with* $\mathbf{E}(u_t^2|\mathcal{F}_{n.t-1}) = \sigma^2$ *a.s. for every* $t = 1, ...n$ *and* $\sup_{1\leq t\leq n} \mathbf{E}(|u_t|^q\,|\mathcal{F}_{n.t-1}) < \infty$ *a.s. for* $q > 2$,

(c) x_t *is adapted to* $\mathcal{F}_{n.t-1}$ *for every* $t = 1, ...n$.

Note that condition (c) implies that $\mathbf{E}(y_t|\mathcal{F}_{n.t-1}) = f(x_t)\ a.s.$ Under Assumptions 3.1 and 3.2, a strong approximation result holds for $(U_n(r), V_n(r))$ (see Park and Phillips (1999) Lemmas 2.3 and 6.2). Our asymptotic results will involve embedding arguments, but to avoid the repetition of such arguments in the derivation of our results, convergence in probability will be interpreted as convergence in distribution unless the limit is non stochastic.

Following P&P we restrict g and f to be members of two families of functions: *I-regular* and *H-regular* functions. The *I-regular* family defined in P&P involves integrable functions. The transformations we consider are typically functions of two arguments. The first argument corresponds to some economic variable and the second to some parameter(s). In particular we say that $f : \mathbf{R} \times \Pi \to \mathbf{R}$ is *I-regular* on the parameter space $\Pi \subset \mathbf{R}^m$ if the following condition is satisfied:

DEFINITION 3.1: *I-regular functions*

(a) *For each* $\pi_o \in \Pi$, *there exists a neighborhood* N_o *of* π_o *and* $T : \mathbf{R} \to \mathbf{R}$ *that is bounded and integrable such that* $\|f(s,\pi) - f(s,\pi_o)\| \leq |\pi - \pi_o| T(s)$ *for all* $\pi \in N_o$

(b) *for some constants* $c > 0$ *and* $k > 6/(p-2)$ *where* $p > 4$ *is as given in Assumption 3.1(a)* $\|f(x,\pi) - f(y,\pi)\| \leq c|x-y|^k$ *for all* $\pi \in \Pi$, *on each piece* S_i *of their common support* $S = \bigcup S_i \subset \mathbf{R}$.

Conditions (a) and (b) are smoothness conditions imposed on the function. The first one implies that for compact Π, $\sup_{\pi\in\Pi} |f(.,\pi)|$ is bounded and integrable, whilst the second one requires sufficient smoothness between the family members on each piece of their common support. P&P provide the following asymptotic results *for I-regular* functions:

THEOREM 3.1: (Park and Phillips (2001))

Let f be I-regular on the compact space Π. *Then, as* $n \to \infty$

(i) $1/\sqrt{n}\sum_{t=1}^{n} f(x_t,\pi) \xrightarrow{p} \int_{-\infty}^{\infty} f(s,\pi)dsL(1,0)$ *uniformly in* π;

(ii) $1/\sqrt[4]{n}\sum_{t=1}^{n} f(x_t,\pi)u_t \xrightarrow{d} \left[\int_{-\infty}^{\infty} f(s,\pi)f(s,\pi)'dsL(1,0)\right]^{1/2} W(1)$,

as $n \to \infty$, *where* $W(1)$ *is a vector Brownian motion at point 1 and* $L(1,0)$ *is the Brownian motion (V) local time at the origin up to time 1.*

The (variance rescaled) local time of the Brownian motion V up to time 1 at the vicinity of the point s is defined as

$$L(1,s) = \lim_{\epsilon \searrow 0} 1/\left(2\epsilon\right) \int_0^1 1\{|V(r) - s| < \epsilon\}dr.$$

The reader is directed to Phillips and Park (1998) and P&P for further discussion about local times and their use in econometrics. Moreover, we will use the following result for "zero energy" transformations of unit root processes due to Jeganathan (2003):

THEOREM 3.2: (Jeganathan (2003))

Let $f : \mathbf{R} \to \mathbf{R}$ be *such that* f, f^2 *are integrable,* $\int_{-\infty}^{\infty} f(s)ds = 0$ *and* $\int_{-\infty}^{\infty} |sf(s)|\, ds < \infty$.

Then, as $n \to \infty$

$$1/\sqrt[4]{n}\sum_{t=1}^{n} f(x_t) \xrightarrow{d} \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\left|\int_{-\infty}^{\infty} e^{isr}f(r)dr\right|^2 \frac{1+\phi(s)}{1-\phi(s)}\right) dsL(1,0)\right)^{1/2} \bar{W}(1),$$

where $\bar{W}(1)$ *is a standard normal variate independent of* $L(1,0)$.

Before we define the *H-regular* family we need to introduce the concept of regularity proposed by P&P.

DEFINITION 3.2: *Regular functions*

A transformation T(.) is said to be regular if and only if

(a) $T(.)$ *is continuous in a neighborhood of infinity, and*

(b) *for every compact set* $K \subset \mathbf{R}$, *there exist for each* $\epsilon > 0$ *continuous functions* $\underline{T}_\epsilon$, $\bar{T}_\epsilon$ *and* $\delta_\epsilon > 0$ *such that* $\underline{T}_\epsilon(x) \leq T(y) \leq \bar{T}_\epsilon(x)$ *for all* $|x - y| < \delta_\epsilon$ *on* K *and* $\int_K \left(\underline{T}_\epsilon(x) - \bar{T}_\epsilon(x)\right) dx \to 0$ *as* $\epsilon \to 0$.

DEFINITION 3.3:

A transformation T(.,π) is said to be regular on Π *if*

(a) $T(.,\pi)$ *is regular for every* $\pi \in \Pi$ *and*

(b) *for all* $x \in \mathbf{R}$ $T(x,.)$ *is equicontinuous in a neighborhood of* x.

P&P provide the following asymptotics for sums of regular transformations of normalised unit root processes:

THEOREM 3.3: (Park and Phillips (2001))

If T is regular on a compact set Π *and Assumption 2.1 holds then:*

(i) $\frac{1}{n}\sum_{t=1}^{n} T(\frac{x_t}{\sqrt{n}}, \pi) \xrightarrow{p} \int_0^1 T(V(r), \pi)dr$ *uniformly in* π,

and if T is regular, then

(ii) $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} T(\frac{x_t}{\sqrt{n}}, \pi)u_t \xrightarrow{d} \int_0^1 T(V(r), \pi)dU(r)$

as $n \to \infty$.

The conditions in Definition 3.2 are required to establish pointwise convergence of the sum in Theorem 3.3, while the equicontinuity requirement of Definition 3.3 establishes uniform convergence over the parameter space Π.

The family of *H-regular* transformations is defined as follows:

DEFINITION 3.4:

The transformation $f(s,\pi)$ *is H-regular on* Π *if and only if*

$$f(\lambda s,\pi) = k_f(\lambda,\pi)h_f(s,\pi) + R_f(\lambda,s,\pi)$$

where $h_f(s,\pi)$ *is regular on* Π, *with* $\pi \in \Pi$, *and* $R_f(\lambda,s,\pi)$ *is such that:*

(i) $|R_f(\lambda,s,\pi)| \leq a(\lambda,\pi)P(s,\pi)$ *with* $\limsup_{\lambda\to\infty}\sup_{\pi\in\Pi}\|a(\lambda,\pi)k_f(\lambda,\pi)^{-1}\| = 0$ *and* $P(s,\pi)$ *locally integrable*

or

(ii) $|R_f(\lambda,s,\pi)| \leq b(\lambda,\pi)Q(\lambda s,\pi)$ *with* $\limsup_{\lambda\to\infty}\sup_{\pi\in\Pi}\|b(\lambda,\pi)k_f(\lambda,\pi)^{-1}\| = O(1)$ *and* $Q(\lambda s,\pi)$ *bounded and vanishing at infinity.*

The functions $k_f(\lambda,\pi)$ and $h_f(s,\pi)$ are called the asymptotic order and the limit homogeneous function of f respectively. For notational brevity we may write the asymptotic order of f as $k_f(\lambda) = k_f$ and when that depends on some parameter e.g. a^* we will write $k_f(\lambda,a^*) = k_f^*$. If the asymptotic order of f does not depend on a parameter, then f will be referred to as H_o*-regular*. Moreover for *H-regular* $\dot{f}$, $\ddot{f}$ the limit homogeneous functions and asymptotic orders of $\dot{f}$, $\ddot{f}$ will be written as $\dot{h}_f$, $\ddot{h}_f$ and $\dot{k}_f$, $\ddot{k}_f$ respectively. Examples of *I-regular* and *H-regular* transformations will be provided later.

The following asymptotic result for sums of regular transformations of unnormalised unit root processes is due to P&P:

THEOREM 3.4: (Park and Phillips (2001))

If f is H-regular on a compact set Π *then:*

(i) $\frac{1}{n}k_f(\sqrt{n},\pi)^{-1}\sum_{t=1}^{n} f(x_t,\pi) \xrightarrow{p} \int_0^1 h_f(V(r),\pi)dr$ *uniformly in* π,

and

(ii) $\frac{1}{\sqrt{n}}k_f(\sqrt{n},\pi)^{-1}\sum_{t=1}^{n} f(x_t,\pi)u_t \xrightarrow{d} \int_0^1 h_f(V(r),\pi)dU(r)$,

as $n \to \infty$.

The leading term in the sum of Theorem 3.4 (i) is regular and its limit behavior is given by Theorem 3.3. The regular family is a subset of the set of locally integrable functions, i.e. functions integrable on any compact subset of **R**. Theorems 3.1 and 3.4 provide the essential asymptotic theory required for the derivation of limit distribution results for the NLS estimator under correct specification. The results of Theorem 3.1(i) and Theorem 3.3(i) will be referred as *sample mean asymptotics*, while the results of Theorem 3.1(ii) and Theorem 3.3(ii) will be referred as *sample covariance asymptotics*. Under functional form misspecification, the sample covariance asymptotics of Theorem 3.1(ii) are not relevant. In this case the sample covariance asymptotics of Theorem 3.2 are relevant.

Continuous and locally bounded monotone functions are regular. Functions with poles like:

$$f(s) = \log|s| \text{ and } f(s) = |s|^c \text{ with } -1 < c < 0,$$

which are appealing for econometric modelling, are not regular. de Jong (2002a) however have shown, that under Assumption 3.1 the result of Theorem 3.2 holds for another family of functions which are not *regular*. This family is comprised of *locally integrable* functions that have finitely many poles and are continuous and monotone between the poles. Pötscher

(2002) extends the result of Theorem 3.3(i) to all locally integrable transformations, under the condition that the normalised unit root process has bounded densities. For the purpose of this paper will follow the P&P and de Jong (2002a) exposition. Hereafter we will use the term *regular* to refer to functions that satisfy Definition 3.2 or belong to the family considered by de Jong (2002a).

Transformations with poles such as

$$f(s) = |s|^{-m} \; with \, m \geq 1, \tag{9}$$

are not locally integrable. Although our theoretical framework is confined to transformations that are integrable or locally integrable (i.e. *I-regular* and *H-regular*), there are occasions in which the product of a locally integrable transformation with an integrable one involves components that are not locally integrable. The asymptotic behaviour of non-locally integrable transformations is as yet unknown (see de Jong and Wang (2002)).

P&P provide several examples of *I-regular* and *H-regular* regression functions. Some examples are given below:

EXAMPLE 2.1:

(a) *I-regular* functions: i) $f(x, a, b) = ae^{-bx^2}$, with $\theta = (a, b) \in \Theta \subset \mathbf{R} \times \mathbf{R}_+$. ii) $f(x, a, b) = a(1 + bx^2)^{-1}$, with $\theta = (a, b) \in \Theta \subset \mathbf{R} \times \mathbf{R}_+$

(b) *H-regular* functions: $f(x, a, b) = ax^b$, with $\theta = (a, b) \in \Theta \subset \mathbf{R} \times \mathbf{R}_+$ is *H-regular* with limit homogenous function and asymptotic order: $h_f(x, a, b) = ax^b$ and $k(a, b, \lambda) = \lambda^b$.

(c) *H_o-regular* functions: The asymptotic order of *H_o-regular* functions does not depend on the parameters. Hence *H_o-regular* functions are usually linear in parameter e.g. $f(x,\theta) = \theta x$, $\theta \ln |x|$, $\theta |x|^c$ with $c > -1$ are *H_o-regular* with limit homogeneous function $h_f(x,\theta) = \theta x$, θ, $\theta |x|^c$ and asymptotic order $k(\lambda) = \lambda$, 1 and λ^c. $f(x,\theta) = x/(1+\theta x)1\{x \geq 0\}$, $x^2/(1+\theta x)1\{x \geq 0\}$, $\theta \in \Theta \subset \mathbf{R}_+$ are examples of *H_o-regular* functions non linear in parameter. The corresponding limit homogenous functions and the asymptotic orders are $h_f(x,\theta) = \theta^{-1}$, $x\theta^{-1}$ and $k(\lambda) = 1$, λ respectively.

(d) *Non-locally integrable* functions: Consider $f(x) = (1+|x|^2)^{-1}$ and $g(x) = |x|^c$, $0 < c < 1$ which are integrable. Now the product $fg = |x|^{-(2-c)} - |x|^{-(2-c)}(1+|x|^2)^{-1}$ involves non-locally integrable terms.

Park and Phillips (1999) have developed asymptotic theory for another class of functions. This class comprises functions that grow with exponential rate (*E-regular*). Moreover de Jong and Wang (2002) provide asymptotic theory for nearly non-locally integrable transformations. Some of our results could be extended also to these two families of functions, however such development will not be attempted in this study.

Before we consider the misspecified case we briefly consider the limit distribution of the NLS estimator under correct specification. Assume that the series, $\{y_t\}_{t=1}^n$, is generated by the model

$$y_t = f(x_t, \theta_o) + u_t$$

with $f : \mathbf{R} \times \mathbf{\Theta} \rightarrow \mathbf{R}$ with Θ being a compact subset of $\mathbf{R}^p$and $\theta_o \in \Theta$. Define the NLS estimator

$$\hat{\theta} = \arg\min_{\theta\in\Theta} \sum_{t=1}^{n} (y_t - f(x_t, \theta))^2$$

If $f(x, \theta)$ is *I-regular* or *H-regular* and under some regularity conditions P&P shown that $\hat{\theta}$ is consistent for θ_o. For *I-regular* $f(x, \theta)$ the following limit distribution result holds:

$$\sqrt[4]{n}(\hat{\theta} - \theta_o) \xrightarrow{d} \left(L(1,0) \int_{-\infty}^{\infty} \dot{f}(s, \theta_o)\dot{f}(s, \theta_o)' ds \right)^{-1/2} W(1)$$

where $\dot{f}(x, \theta)$ is the first derivative of the regression function with respect to the argument θ. We notice that the limit distribution is mixed normal and the convergence rate is $\sqrt[4]{n}$, which is smaller than the one obtained under stationary $(\sqrt{n})$. For *H-regular* $f(x, \theta)$, P&P provide the following limit distribution result:

$$\sqrt{n}\dot{k}_f'(\sqrt{n}, \theta_o)(\hat{\theta} - \theta_o) \xrightarrow{d} \left(\int_0^1 \dot{h}_f(V(r), \theta_o)\dot{h}_f(V(r), \theta_o)' dr \right)^{-1} \int_0^1 \dot{h}_f(V(r), \theta_o) dU(r)$$

where $\dot{k}_f(\sqrt{n}, \theta_o)$ is the asymptotic order (matrix) of $\dot{f}(s, \theta_o)$. The limit distribution involves a stochastic integral and functionals of Brownian motion. Mixed normal is obtained as special case when $E(v_t u_{t-1}) = 0$. The convergence rate is faster than $\sqrt{n}$ whenever $\dot{k}_f(\sqrt{n}, \theta_o)$ diverges.

2.4 Consequences of Functional Form Misspecification

In this section we investigate the behaviour of estimators when the fitted model is of wrong functional form. We assume that the time series $\{y_t\}_{t=1}^n$ is generated by the model:

$$y_t = f(x_t, \theta_o) + u_t \tag{10}$$

where $f : \mathbf{R} \to \mathbf{R}$ is either *I-regular* or *H-regular*[2], x_t and u_t are the unit-root and martingale difference processes defined in section 2.2. Consider the fitted model:

$$\hat{y}_t = g(x_t, \hat{a}) + \hat{u}_t \tag{11}$$

where $g : \mathbf{R} \times \mathbf{A} \to \mathbf{R}$ is either *I-regular* or *H-regular* with A a compact subset of $\mathbf{R}^m$. We will refer to the fitted model of being of correct functional form if

$$f(., \theta_o) = g(., a_o), \textit{for a unique } a_o \in A$$

Similarly we will say that the fitted model is of false functional form if

$$f(., \theta_o) \neq g(., a) \textit{ for every } a \in A.$$

The NLS estimator of a fitted model (equation (11)) is defined as:

$$\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2. \tag{12}$$

When the fitted model is linear in parameter, the relevant estimation procedure is Ordinary Least Squares (OLS) and the "pseudo-true parameter space" A is not required to be compact. The estimator for the fitted model in this instance is:

$$\hat{a} = \arg\min \sum_{t=1}^{n} (y_t - a g(x_t))^2. \tag{13}$$

2.4.1 Misspecification under Stationarity VS Misspecification under Nonstationarity: Some Theoretical Considerations

It is apparent from Theorems 2.1 and 2.3 that the asymptotics for unit root processes are quite different from the asymptotics for stationary data. The sample averages of unit root

[2] The parameter θ_o in f could be suppressed; however we will keep it to highlight what the parameter of interest might be in some examples later.

processes converge to random quantities rather than fixed numbers, which is the case under stationarity. Moreover the limits of the sample covariance expressions are different. Under stationarity, the limit distribution is Gaussian whilst under nonstationarity it is either mixed normal or involves stochastic integrals. The most important difference though - when it comes to functional form misspecification - is that different transformations of unit processes are of different asymptotic order, while transformations of stationary data are of the same asymptotic order.

A transformation, say $T(x)$, applied to a unit-root process may strengthen or undermine the signal produced by the process itself. An *I-regular* transformation undermines the signal of the process. An *H-regular* transformation undermines the signal of the process if the asymptotic order $k_T(\lambda)$ of $T(x)$ converges. The signal of the process is strengthened if $k_T(\lambda)$ diverges. This effect is apparent from Theorems 3.1 and 3.4. The sample average of a unit root transformation is of order $O_p(\sqrt{n})$, if the transformation is *I-regular*, and of order $O_p(k_T(\sqrt{n})n)$ if the transformation is *H-regular*. By allowing the functional form of the true and the fitted model to be either in the *I-regular* or the *H-regular* family we effectively allow the true specification to be of different order than the fitted specification. As mentioned above, this kind of complication does not arise in the stationary framework. The relative asymptotic order of the fitted model to the one of the true model is of central importance for the subsequent analysis. For convenience we introduce the following notation:

(a) If the asymptotic order of the fitted model $g(a)$ is the same as that of the true model f, for some $a \in A$ we denote this by $g \approx f$,

(b) If the fitted model $g(a)$ is of higher (smaller) asymptotic order than the true model f, or all $a \in A$ that will be denoted by $g \succ f$ ($g \prec f$).

Members of the *H-regular* family are of higher asymptotic order than members of the *I-regular* family. Moreover, members of the *I-regular* family are of the same asymptotic order, while members of the *H-regular* family may be of different order,

When the fitted model is nonlinear in parameters the estimation procedure under consideration is NLS. The NLS estimator, $\hat{a}$, is defined as the minimiser of the objective function over a compact set, A (equation (12)). The compactness assumption is standard in problems that are non-linear in parameters for theoretical convenience. For (scalar parameter) models that are linear in parameter the relevant procedure is OLS. The OLS estimator is defined as in (13). The OLS problem has a closed form solution so the compactness assumption is redundant. One may consider a "restricted" version of the OLS estimator defined as the minimiser of the objective function over a compact set:

$$\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - ag(x_t))^2 \tag{14}$$

Under correct functional form specification, there is no difference asymptotically between the estimator of (13) and the one of (14). Both estimators will be consistent for the parameter of interest, provided the latter is contained within the specified parameter space. This is not the case though under functional form misspecification. Under functional form misspecification the two estimators may exhibit different asymptotic behaviour. For instance, as we are going to show later in this section, if the fitted model is of smaller asymptotic order than the true one, there are occasions in which the OLS estimator as defined by (13) diverges in probability. This divergent behavior is not exhibited by the estimator of (14) as in this case the estimator is defined over bounded interval. Actually in this instance, the "restricted" OLS estimator converges to a boundary point of the pseudo-true parameter space.

2.4.2 Models Linear in Parameter

We begin our analysis with models that are linear in parameter. The analysis of models that are nonlinear in parameter is left for section 2.4.3. The reason we treat the two types of models separately is twofold: first the techniques for the asymptotic analysis for models that are nonlinear in parameter are more involved, second, as was mentioned in the previous section, under functional form misspecification the OLS procedure is not directly comparable to the NLS procedure. Under the former procedure the interval on which the estimator is defined is unbounded, whilst under the latter procedure the estimator is defined on bounded set. So in principle the OLS procedure -in contrast to the NLS- can deliver estimators that diverge in probability.

Our findings can be summarised as follows. The OLS estimator converges to a pseudo-true value whenever the fitted model is of asymptotic order at least as large as that of the true model. In particular the pseudo-true value is zero when $g \succ f$ and non zero when $g \approx f$. The OLS estimator diverges in probability whenever $g \prec f$. An exception to the last case occurs when g is *I-regular* and f is *H-regular* but the product gf is *I-regular*. In this case although $g \prec f$, the signal produced by f is neutralised and as result the OLS estimator converges to some pseudo-true value.

We first consider the case when the fitted model, g is *I-regular* We distinguish three cases: (a) the true model f is *I-regular,* (b) the true model f is *H-regular* and the product fg is *I-regular* and (c) the true model f is *H-regular* and the product fg is *H-regular.* The properties of the OLS estimator for *I-regular* g are shown in Proposition 4.1:

PROPOSITION 4.1: (g *I-regular*)

Let $\hat{a} = \arg\min \sum_{t=1}^{n} \left(y_t - ag(x_t)\right)^2$ *and* g *be I-regular.*

(i) *If* f *is I-regular or H-regular such that* gf is *I-regular, then*

$$\hat{a} \xrightarrow{p} a^* = \frac{\int_{-\infty}^{\infty} f(s,\theta_o)g(s)ds}{\int_{-\infty}^{\infty} g(s)^2 ds}$$

and if $\int_{-\infty}^{\infty} |sz(s,a^*)|\, ds < \infty$, *where* $z(s,a^*) = f(s,\theta_o)g(s)ds - a^* g(s)^2$ *then*

$$n^{1/4}\left(\hat{a} - a^*\right) \xrightarrow{d} \left[\int_{-\infty}^{\infty} g(s)^2 ds L(1,0)^{1/2}\right]^{-1} \times$$
$$\left[\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\left|\int_{-\infty}^{\infty} e^{isr} z(r,a^*)dr\right|^2 \frac{1+\phi(s)}{1-\phi(s)}\right) ds\right)^{1/2} \bar{W}(1) + \left(\int_{-\infty}^{\infty} g(s)^2 ds\right)^{1/2} W(1)\right]$$

with $\bar{W}(1)$ *standard normal independent of* $L(1,0)$ *and* $W(1)$.[3]

(ii) *If* f *is H-regular such that* gf *is H-regular with limit homogeneous function* h_{gf} *and asymptotic order* k_{gf}, *then*

$$\frac{1}{\sqrt{n}k_{gf}(\sqrt{n})}\hat{a} \xrightarrow{d} \frac{\int_0^1 h_{gf}(V(r),\theta_o)dr}{\int_{-\infty}^{\infty} g(s)^2 ds L(1,0)}$$

and

$$\hat{a} = O_p(\sqrt{n}k_{gf}(\sqrt{n})) \xrightarrow{p} \infty$$

as $n \longrightarrow \infty$.

---

[3] This result has been suggested to me by Professor P. C. B. Phillips.

We can see from Proposition 4.1 that the OLS estimator converges to a pseudo-true value in two cases: when the true model is *I-regular* and when the true model is *H-regular* but the product of f and g give an *I-regular* transformation. If the product f and g give an *H-regular* transformation, then the OLS estimator will diverge in probability. The divergence rate depends on the asymptotic order of fg. We note for the kind of misspecification in (i), the limit distribution is still mixed normal, but with larger variance than one attained under correct specification.

Now we turn to the case when the fitted model is *H-regular*. The true model may belong to the *I-regular* or *H-regular* families. Although the members of the *I-regular* family are of the same asymptotic order, members of the *H-regular* family may be of different order. So whenever the true model is *H-regular* we distinguish three cases: (a) $g \approx f$, (b) $g \succ f$ and (c) $g \prec f$. The asymptotic properties of the OLS estimator for g *H-regular* are shown in Proposition 4.2:

PROPOSITION 4.2: (g *H-regular*)

Let $\hat{a} = \arg\min \sum_{t=1}^{n} \left(y_t - ag(x_t)\right)^2$ *and* g *be H-regular.*

(i) *If* f *is H-regular such that* $k_g = k_f$ *and* $f(x, \theta_o) - a^* g(x) = q(x, \theta_o)$, q *H-regular such that* $\left(k_q k_f^{-1}\right)(\lambda) \overset{\lambda \to \infty}{\to} \infty$. *Then*

$$\hat{a} \xrightarrow{p} a^*$$

and

$$\frac{k_g(\sqrt{n})}{k_q(\sqrt{n})}(\hat{a} - a^*) \xrightarrow{d} \frac{\int_0^1 h_g(V(r)) h_q(V(r), \theta_o) dr}{\int_0^1 h_g(V(r))^2 ds}$$

(ii) *If f is H-regular such that $(k_f k_g^{-1})(\lambda) \overset{\lambda\to\infty}{\to} 0$, then*

$$\hat{a} = O_p\left(\frac{k_f(\sqrt{n})}{k_g(\sqrt{n})}\right) = o_p(1)$$

and

$$\frac{k_g(\sqrt{n})}{k_f(\sqrt{n})}\hat{a} \overset{d}{\to} \frac{\int_0^1 h_f(V(r),\theta_o)h_g(V(r))dr}{\int_{-\infty}^{\infty} h_g(V(r))^2 ds}$$

(iii) *If f is H-regular such that $(k_f k_g^{-1})(\lambda) \overset{\lambda\to\infty}{\to} \infty$, then*

$$\hat{a} = O_p\left(\frac{k_f(\sqrt{n})}{k_g(\sqrt{n})}\right) \overset{p}{\to} \infty$$

and

$$\frac{k_g(\sqrt{n})}{k_f(\sqrt{n})}\hat{a} \overset{d}{\to} \frac{\int_0^1 h_f(V(r),\theta_o)h_g(V(r))dr}{\int_0^1 h_g(V(r))^2 dr}$$

(iv) *If f is I-regular such that fg* is *I-regular, then*

$$\hat{a} = o_p(1)$$

and if $k_g(\lambda) \overset{\lambda\to\infty}{\to} 0$

$$\sqrt{n}k_g(\sqrt{n})^2\hat{a} \overset{d}{\to} \frac{\int_{-\infty}^{\infty} f(\theta_o,s)g(s)dsL(1,0)}{\int_0^1 h_g(V(r))^2 dr},$$

while, if $k_g(\lambda) \overset{\lambda\to\infty}{\to} \infty$

$$\sqrt{n}k_g(\sqrt{n})\hat{a} \overset{d}{\to} \frac{\int_0^1 h_g(V(r))dU(r)}{\int_0^1 h_g(V(r))^2 dr},$$

(v) *If f is I-regular such that fg is H-regular with limit homogeneous function h_{gf} and asymptotic order k_{gf}, then*

$$\hat{a} = O_p\left(\frac{k_{gf}(\sqrt{n})}{k_g^2(\sqrt{n})}\right) = o_p(1)$$

$$\frac{k_g^2(\sqrt{n})}{k_{gf}(\sqrt{n})}\hat{a} \overset{d}{\to} \frac{\int_0^1 h_{gf}(V(r),\theta_o)dr}{\int_0^1 h_g(V(r))^2 dr}.$$

as $n \to \infty$.

Proposition 4.2 (i) considers the case when the true model, f can be represented as a sum of homogeneous components of different asymptotic orders and the fitted model agrees with the leading term of f. So effectively the fitted model is correctly specified up to a term that is of smaller order than k_f. The limit distribution is quite different to the one under correct specification. It involves functionals of Brownian motion but not a stochastic integral as opposed to the correctly specified case. The convergence rate is determined by the relative asymptotic order of the leading term of f to the one of the second leading term and is slower than the rate attained under correct specification. Whenever $g \succ f$, (Proposition 4.2 (ii), (iv) and (v)) the OLS estimator converges to zero. The limit distribution is comparable to the one under correct specification only in the second case of Proposition 4.2 (iv). In all other cases the limit distribution is different than the one under correct specification and the convergence rates are slower. The OLS estimator diverges when $g \prec f$ (Proposition 4.2 (iii)). The divergence rate is determined by the relative asymptotic order of the true model to the one of the fitted model.

2.4.3 Models Nonlinear in Parameter

Convergence to Pseudo-true Value

We have seen that the OLS estimator may exhibit divergent behaviour when the fitted model is of smaller asymptotic order than the true one. This is not the case though when the relevant estimation procedure is NLS. An obvious explanation for this is that the NLS estimator is defined over a compact set. The compactness assumption about the parameter space may appear restrictive, as it implicitly requires that there are known bounds for the parameter of interest. Nonetheless it is standard assumption for models nonlinear in parameter for theoretical convenience (see for example White (1981), Hansen (1982), Newey and McFadden (1994)). Moreover in practice, when it comes to fitting a model that is nonlinear in the parameter using a computer, one may effectively need to define an interval over which the computer minimises the objective function. As we are going to show later in this section, the NLS estimator converges to a *boundary* point of the pseudo parameter space A, whenever the true model is of different asymptotic order than the true specification. The NLS estimator may converge to an interior point of A only when the true and the fitted models are of the same asymptotic order.

Most of the consistency results provided follow from a Jennrich (1969) type of result. The Jennrich Theorem requires that the objective function, $Q_n(a)$, appropriately rescaled, converges uniformly to a function $Q(a)$ that is continuous and has a unique minimum with respect to a. The objective function for most of the models we are dealing with, involves components of different orders. In particular it is often the case that the order of magnitude of components that depend on a is dominated by the order of components that do not depend on a. For this reason it is more convenient to consider a shifted version of the objective function:

$$D_n(a, a^*) = Q_n(a) - Q_n(a^*), \text{ with } a^* \in A$$

rather the objective function itself. Following P.&P, we will establish consistency by verifying the following condition (CN1):

CN1:

Let v_n be a normalising sequence of real numbers. If $v_n^{-1}D_n(a, a^*) \to D(a, a^*)$ in probability uniformly in a as $n \to \infty$, and $D(a, a^*)$ is continuous in a and has a unique minimum at a^* $a.s.$ Then $\hat{a} \to a^*$ in probability.

Although CN1 is applicable to *I-regular* and *H_o-regular* functions, is not applicable to general *H-regular* functions (see P&P), as these functions have different rates for different values of a. The following condition (CN2) due to Wu (1981) is more relevant when the model is given by a general *H-regular* function:

CN2:

If for any $\delta > 0$, $\liminf_{n\to\infty} \inf_{|a-a^*|\geq\delta} D_n(a, a^*) > 0$ in probability then $\hat{a} \to a^*$ in probability.

The consistency results provided for general *H-regular* models follow from CN2.

The asymptotic behaviour of the NLS estimator for *I-regular* g is given in Propositions 4.3-4.5. and for g *H_o-regular* in Propositions 4.6-4.8. Assumption (c) in all propositions below is an identification condition. It ensures that the limit objective function has a unique minimum and hence in view of CN1 is sufficient for the convergence of the NLS estimator to some pseudo-true value. Some examples of f and g functions for which condition (c) is satisfied are also provided.

PROPOSITION 4.3: (*g I-regular, f I-regular*)

Let

(a) $\hat{a} = \arg\min_{a\in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2$ *and g be I-regular on A,*

(b) *f be I-regular,*

(c) $\int_{-\infty}^{\infty}[f(s,\theta_o)-g(s,a)]^2ds > \int_{-\infty}^{\infty}[f(s,\theta_o)-g(s,a^*)]^2ds$ *for all* $a \in A : a \neq a^*$.

Then

$$\hat{a} \xrightarrow{p} a^*.$$

In particular we have

$$D(a,a^*) = \left(\int_{-\infty}^{\infty}[f(s,\theta_o)-g(s,a)]^2ds - \int_{-\infty}^{\infty}[f(s,\theta_o)-g(s,a^*)]^2ds\right)L(1,0)$$

with $v_n = \sqrt[4]{n}$.

EXAMPLE 4.1:

Let $f(s,\theta_o) = (1+s^2)^{-1}$ and $g(s,a) = e^{-as^2}$ with $A \subset \mathbf{R}_+$. f is *H-regular*, g is *I-regular* on A and gf is *I-regular* on A. Now condition (c) of Proposition 4.3 requires that $\int_{-\infty}^{\infty}[g(s,a)-f(s,\theta_o)]^2ds$ has a unique minimum with respect to a. The integral does not have an analytical solution. From numerical integration we find that integral is minimised at a point close to zero.

PROPOSITION 4.4: (g *I-regular*, f *H-regular*, gf *I-regular*)

Let

(a) $\hat{a} = \arg\min_{a\in A}\sum_{t=1}^{n}(y_t - g(x_t,a))^2$ *and g be I-regular on A,*

(b) *f be H-regular and gf be I-regular on A,*

(c) $\int_{-\infty}^{\infty}[g(s,a)^2 - 2f(s,\theta_o)g(s,a)]ds > \int_{-\infty}^{\infty}[g(s,a^*)^2 - 2f(s,\theta_o)g(s,a^*)]ds$

for all $a \in A : a \neq a^*$.

Then

$$\hat{a} \xrightarrow{p} a^*.$$

In particular we have

$$D(a,a^*) = \left(\int_{-\infty}^{\infty} [g(s,a)^2 - 2f(s,\theta_o)g(s,a)]ds - \int_{-\infty}^{\infty} [g(s,a^*)^2 - 2f(s,\theta_o)g(s,a^*)]ds\right) L(1,0)$$

with $v_n = \sqrt[4]{n}$.

EXAMPLE 4.2:

Let $f(s,\theta_o) = \theta_o s$ and $g(s,a) = e^{-as^2}$ with $A \subset \mathbf{R}_+$. f is *H-regular*, g is *I-regular* on A and gf is *I-regular* on A. Now $\int_{-\infty}^{\infty}[g(s,a)^2 - 2f(s,\theta_o)g(s,a)]ds = \sqrt{\frac{\pi}{2a}}$ so condition (c) of Proposition 4.4 is satisfied with a^* being the upper boundary point of A.

PROPOSITION 4.5: (*g I-regular, f H-regular, gf H-regular*)

Let

(a) $\hat{a} = \arg\min_{a\in A} \sum_{t=1}^{n} (y_t - g(x_t,a))^2$ *and g be I-regular on A,*

(b) *f be H_o-regular and gf be H-regular on A with positive limit homogeneous function h_{gf} and asymptotic order k_{gf},*

(c) $\int_{-\infty}^{\infty} h_{gf}(s,\theta_o,a)L(1,s)ds < \int_{-\infty}^{\infty} h_{gf}(s,\theta_o,a^*)L(1,s)ds$ *a.s.* for all $a \in A : a \neq a^*$.

Then

$$\hat{a} \xrightarrow{p} a^*.$$

In particular we have

$$D(a,a^*) = \left(\int_{-\infty}^{\infty} h_{gf}(s,\theta_o,a^*)L(1,s)ds - \int_{-\infty}^{\infty} h_{gf}(s,\theta_o,a)L(1,s)ds\right)$$

with $v_n = nk_{gf}(\sqrt{n})$.

From the Propositions above we see that when f is of different order than g, the limit objective function $D(a, a^*)$ is not a "complete" quadratic form as in Proposition 4.3. In this case the limit objective function is minimised at a boundary point of the parameter space A.

Next we consider the case when f is *H-regular.*

PROPOSITION 4.6: (*g H-regular, f H-regular*)

Let

(a) $\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2$ *and g be H_o-regular on A,*

(b) *f be H-regular and with* $k_f = k_g$,

(c) $\int_{-\infty}^{\infty} [h_f(s, \theta_o) - h_g(s, a)]^2 L(1, s) ds > \int_{-\infty}^{\infty} [h_f(s, \theta_o) - h_g(s, a^*)]^2 L(1, s) ds \; a.s.$

for all $a \in A : a \neq a^*$.

Then

$$\hat{a} \xrightarrow{p} a^*.$$

In particular we have

$$D(a, a^*) = \left(\int_{-\infty}^{\infty} [h_f(s, \theta_o) - h_g(s, a)]^2 L(1, s) ds - \int_{-\infty}^{\infty} [h_f(s, \theta_o) - h_g(s, a^*)]^2 L(1, s) ds \right)$$

with $v_n = n k_f(\sqrt{n})^2$.

EXAMPLE 4.4:

Let $f(s, \theta_o) = \theta_o s^2 1\{s > 0\}$ and $g(s, (a) = s^3 (1 + as)^{-1} 1\{s > 0\}$ with $A \subset \mathbf{R}_+$. f is *H-regular and* g is *H_o-regular* on A. Condition (c) of Proposition 4.6 requires $\left(\theta_o - \frac{1}{a}\right)^2 \int_0^\infty s^2 L(1, s) ds > \left(\theta_o - \frac{1}{a^*}\right)^2 \int_0^\infty s^2 L(1, s) ds \; a.s.$ which holds whenever $a^* \neq \theta_o^{-1}$

PROPOSITION 4.7: (*g H-regular, f I-regular or H-regular with $g \succ f$*)

Let

(a) $\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2$ *and g be H_o-regular on A,,*

(b) *f be I-regular or H-regular with $(k_f k_g^{-1})(\lambda) \to 0$ as $\lambda \to \infty$,*

(c) $\int_{-\infty}^{\infty} h_g(s,a)^2 L(1,s)ds > \int_{-\infty}^{\infty} h_g(s,a^*)^2 L(1,s)ds \; a.s.$ for all $a \in A : a \neq a^*$.

Then

$$\hat{a} \xrightarrow{p} a^*.$$

In particular we have

$$D(a,a^*) = \left(\int_{-\infty}^{\infty} h_g(s,a)^2 L(1,s)ds - \int_{-\infty}^{\infty} h_g(s,a^*)^2 L(1,s)ds \right)$$

with $v_n = n k_g(\sqrt{n})^2$.

EXAMPLE 4.5:

Let $f(s,\theta_o) = \theta_o s$ and $g(s,a) = s^3(1+as)^{-1}1\{s > 0\}$ with $A \subset \mathbf{R}_+$. f is *H-regular and* g is *H_o-regular* on A. Condition (c) of Proposition 4.6 requires $\left(\frac{1}{a}\right)^2 \int_0^{\infty} s^2 L(1,s)ds > \left(\frac{1}{a^*}\right)^2 \int_0^{\infty} s^2 L(1,s)ds \; a.s.$ which holds whenever a^* is the upper boundary point of A.

PROPOSITION 4.8: (*g H-regular, f H-regular with $g \prec f$*)

Let

(a) $\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2$ *and g be H_o-regular on A,,*

(b) *f be H-regular with $(k_f k_g^{-1})(\lambda) \to \infty$ as $\lambda \to \infty$*

(c) $\int_{-\infty}^{\infty} h_f(s,\theta_o) h_g(s,a) L(1,s)ds < \int_{-\infty}^{\infty} h_f(s,\theta_o) h_g(s,a^*) L(1,s)ds \; a.s.$ *for all* $a \in A : a \neq a^*$.

Then

$$\hat{a} \xrightarrow{p} a^*.$$

In particular we have

$$\begin{aligned} D(a,a^*) &= 2\int_{-\infty}^{\infty} h_f(s,\theta_o)h_g(s,a^*)L(1,s)ds \\ &\quad -2\int_{-\infty}^{\infty} h_f(s,\theta_o)h_g(s,a)L(1,s)ds \end{aligned}$$

with $v_n = nk_f(\sqrt{n})k_g(\sqrt{n})$.

EXAMPLE 4.6:

Let $f(s,\theta_o) = \theta_o s^2$ and $g(s,a) = s(1+as)^{-1}1\{s>0\}$ with $A \subset \mathbf{R}_+$. f is *H-regular and* g is H_o*-regular* on A. Condition (c) of Proposition 4.7 requires $\frac{1}{a}\int_0^{\infty} s^3 L(1,s)ds < \frac{1}{a^*}\int_0^{\infty} s^3 L(1,s)ds$ $a.s.$ which holds whenever $a^*$is the lower boundary point of A.

We notice that in all examples above the limit objective function $D(a,a^*)$ is strictly monotonic in a whenever f and g are of different orders. In this case the NLS estimator converges to boundary point of the pseudo-true parameter space. The NLS estimator may converge to value that is interior in A when f and g are of the same order. For instance in Example 4.3 the NLS estimator converges to point a^* that is interior in A when θ_o^{-1} is interior in A.

Propositions 4.9 and 4.10 consider the case when the fitted model g is a general *H-regular* function. In this case the convergence of the NLS estimator is established by providing sufficient conditions for CN2. In Proposition 4.9 the fitted model is assumed to be correctly specified up some lower order *H-regular* component. This can be seen as a case of missing or redundant lower order components.

PROPOSITION 4.9: (*g H-regular, f H-regular with $g \approx f$*))

Let

(a) $\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2$ *and g be H-regular on A, with limit homogeneous function and asymptotic order h_g and $k_g(\lambda, a)$ respectively,*

(b) *f be H-regular with asymptotic order k_f such that* $f(x, \theta_o) - g(x, a^*) = q(x, a^*)$ *with* $k_f(\lambda) = k_g(\lambda, a^*)$ *and $q(x, a)$ H-regular on A with asymptotic order $k_q(\lambda, a)$ such that* $k_q(\lambda, a^*) k_g(\lambda, a^*)^{-1} \overset{\lambda \to \infty}{\to} 0$.

Then CN2 *holds if:*

(i) *for any $\bar{a} \neq a^*$ and $\bar{p}, \bar{c} > 0$, there exist $\varepsilon > 0$ and a neighborhood N of $\bar{a}$ such that as $\lambda \to \infty$*

$$\inf_{\substack{|c-\bar{c}|<\varepsilon \\ |p-\bar{p}|<\varepsilon}} \inf_{a \in N} |pk_g(\lambda, a) - ck_g(\lambda, a^*)| \times k_q(\lambda, a^*)^{-1} \to \infty;$$

(ii) *for all $a \in A$ and $\delta > 0$,* $\int_{|s| \leq \delta} h_g(s, a)^2 ds > 0$.

EXAMPLE 4.7:

The conditions of Proposition 4.9 are satisfied for $f(s, \theta_o) = s^{\theta_o}(1 + s)^{-1} 1\{s > 0\}$ and $g(s, a) = s^a 1\{s > 0\}$ with $\theta_o \in A \subset \mathbf{R}_+$ and $a^* = \theta_o - 1$.

Proposition 4.10 considers the case when $f \prec g$ for all $a \in A$. In this case the NLS estimator converges to a^*, the lower boundary point of A.

PROPOSITION 4.10: *(g H-regular, f H-regular with $g \succ f$))*

Let

(a) $\hat{a} = \arg\min_{a \in A} \sum_{t=1}^{n} (y_t - g(x_t, a))^2$ *and g be H-regular on $A \subset \mathbf{R}_+$ with limit homogeneous function h_g and asymptotic order $k_g(\lambda, a)$,*

(b) *f be I-regular or H-regular with asymptotic order $k_f(\lambda)$*

(c) $k_f(\lambda) k_g(\lambda, a)^{-1} \overset{\lambda \to \infty}{\to} 0$ *for all $a \in A$,*

(d) *a^* be the lower boundary point of A i.e. $k_g(\lambda, a^*) k_g(\lambda, a)^{-1} \overset{\lambda \to \infty}{\to} 0$ for all $a \in A : a \neq a^*$.*

Then CN2 holds if:

(i) *for any $\bar{a} \neq a^*$ and $\bar{p}, \bar{c} > 0$, there exist $\varepsilon > 0$ and a neighborhood N of $\bar{a}$ such that as $\lambda \to \infty$*

$$\inf_{\substack{|c - \bar{c}| < \varepsilon \\ |p - \bar{p}| < \varepsilon}} \inf_{a \in N} |p k_g(\lambda, a) - c k_g(\lambda, a^*)| \times k_g(\lambda, a^*)^{-1} \to \infty;$$

(ii) *for all $a \in A$ and $\delta > 0$, $\int_{|s| \leq \delta} h_g(s, a)^2 ds > 0$.*

EXAMPLE 4.8:

The conditions of Proposition 4.10 are satisfied for $f(s, \theta_o) = \theta_o \ln s 1\{s > 0\}$ and $g(s, a) = s^a 1\{s > 0\}$ with $a \in A \subset \mathbf{R}_+$.

Limit distribution

The limit distribution results provided cover the case when the true and fitted models are of the same asymptotic order. In particular when both models are *H-regular* we assume that the fitted model is correctly specified up to some lower order term, i.e.:

$$f(s,\theta_o) - g(s,a^*) = q(s,a^*) \text{ for some } a^* \in A$$

where q is a lower order (lower than f and g) component. This is the case considered in Propositions 4.6 and 4.9. Under this kind of misspecification the true and fitted models agree in order. Moreover we will assume that the pseudo-true value is an interior point in A. The limit distribution theory can be obtained in a similar fashion to the one in P&P. The examples in section 4.3 suggest that when the true and fitted models are of different orders the NLS estimator converges to boundary points in A. Andrews (1999) provides a general approach for obtaining limit distribution results for problems in which the parameter is a boundary point in the parameter space (see also Phillips and Moon (2003)). This approach is not applicable for the kind of problem under consideration, for the following reason. Andrews (1999) considers only the case in which the minimum of the limit objective function $D(a,a^*)$ is a turning point. For the kind of misspecification under consideration, $D(a^*,a^*)$ in CN1 does not correspond to a turning point.

The most common way of deriving a limit distribution result is based on an application of the mean value theorem on the derivative of the objective function $\dot{Q}_n(\hat{a})$ around a^*, the probability limit of $\hat{a}$. This is the method used in section 2. We recall that

$$\dot{Q}_n(\hat{a}) = \dot{Q}_n(a^*) + \ddot{Q}_n(\bar{a})(\hat{a} - a^*), \tag{15}$$

where $\|\bar{a} - a^*\| \leq \|\hat{a} - a^*\|$, $A \subset \mathbf{R}$ and a^* interior in A. Now given (15) suppose the following hold:

$$s_n^{-1} v_n' (\hat{a} - a^*) = - \left[v_n^{-1} \ddot{Q}_n(\bar{a}) v_n'^{-1} \right]^{-1} (s_n v_n)^{-1} \dot{Q}_n(a^*) \tag{16}$$
$$= - \left[v_n^{-1} \ddot{Q}_n(a^*) v_n'^{-1} \right]^{-1} (s_n v_n)^{-1} \dot{Q}_n(a^*) \xrightarrow{d}$$
$$- \ddot{Q}(a^*)^{-1} \dot{Q}(a^*)$$

where, v_n is sequence of normalising matrices and s_n a normalising sequence of real numbers. Consider the following conditions:

A1 : $(s_n v_n)^{-1} \dot{Q}_n(a^*) \xrightarrow{d} \dot{Q}(a^*)$, *as* $n \to \infty$.
A2 : $v_n^{-1} \ddot{Q}_n(a^*) v_n'^{-1} = v_n^{-1} \ddot{Q}_n^o(a^*) v_n'^{-1} + o_p(1)$, *for n large enough.*
A3 : $v_n^{-1} \ddot{Q}_n(a^*) v_n'^{-1} \xrightarrow{p} \ddot{Q}(a^*)$, *as* $n \to \infty$.
A4 : $\ddot{Q}(a^*) > \mathbf{0}$ *a.s.*
A5 : $\dot{Q}_n(\hat{a}) = \mathbf{0}$, *with probability approaching one, as* $n \to \infty$.
A6 : $v_n^{-1} \left(\ddot{Q}_n(\bar{a}) - \ddot{Q}_n(a^*) \right) v_n'^{-1} = o_p(1)$, *as* $n \to \infty$.
A7a : $s_n = 1$.
A7b : $s_n \to \infty$, *with* $\|s_n v_n^{-1}\| \to 0$, *as* $n \to \infty$.

Conditions $A1 - A7a$ are considered by P&P. They are standard assumptions for nonlinear models and are sufficient for (16). Given that $\hat{a} = a^* + o_p(1)$, these conditions can be easily verified for a variety of nonlinear problems. Condition $A5$ requires that the objective function satisfies a first order condition in the limit. Condition $A7b$ requires that the "Score" be of smaller order than the "Hessian". For most problems s_n is equal to one, leading to the familiar v_n-*consistency* result for extremum estimators. For the kind of misspecification under consideration, s_n diverges. The score is of higher order than is under correct specification, and as result the order of consistency is compromised.

The conditions above can be easily checked under the kind of misspecification considered in Propositions 4.3 and 4.6 provided a^* is interior in A. We recall that in Proposition 4.3 the true and fitted models are *I-regular.*

PROPOSITION 4.11: (*g I-regular, f I-regular*)

Let f be I-regular and g I-regular on $A \subset \mathbf{R}^m$ with

(a) a^* *as in Proposition 4.3,*

(b) $\int_{-\infty}^{\infty} |sz(s,a^*)|\, ds < \infty$ *where* $z(s,a^*) = \dot{g}(s,a^*)\left(f(s,\theta_o) - g(s,a^*)\right)$,

(c) $\int_{-\infty}^{\infty} \dot{g}(s,a^*)\dot{g}(s,a^*)'ds - \int_{-\infty}^{\infty} \ddot{G}(s,a^*)\left(f(s,\theta_o) - g(s,a^*)\right) > 0$.

Then

$$n^{1/4}\left(\hat{a} - a^*\right) \xrightarrow{d}$$

$$L(1,0)^{-1/2}\left[\int_{-\infty}^{\infty} \dot{g}(s,a^*)\dot{g}(s,a^*)'ds - \int_{-\infty}^{\infty} \ddot{G}(s,a^*)\left(f(s,\theta_o) - g(s,a^*)\right)ds\right]^{-1} \times$$

$$\left[\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\left|\int_{-\infty}^{\infty} e^{isr}z(r,a^*)dr\right|^2 \frac{1+\phi(s)}{1-\phi(s)}\right)ds\right)^{1/2}\bar{W}(1) + \left(\int_{-\infty}^{\infty} g(s)g(s)'ds\right)^{1/2} W(1)\right]$$

as $n \to \infty$.

In Proposition 4.6 the fitted model is *H_o-regular* and is of the same order as the true model.

PROPOSITION 4.12: (*g H-regular, f H-regular*)

Let f be H-regular and g H-regular on $A \subset \mathbf{R}^m$ with asymptotic order $k_f(\lambda)$ and $k_g(\lambda)$ respectively. Moreover assume

(a) $k_f(\lambda) = k_g(\lambda)$,

(b) $f(s,\theta_o) - g(s,a^*) = q(s,a^*)$ *with q, $\dot{g}$ and $\ddot{g}$ H_o-regular on A such that*

$$\sup_{\lambda}\left\|\left(\dot{k}_g \otimes \dot{k}_g\right)^{-1} k_g\ddot{k}_g\right\| < \infty,$$

(c) $T(s,a) = \dot{g}(s,a)\,(f(s,\theta_o) - g(s,a))$ *such that*

$$T(\lambda s, a) = k_g(\lambda)\dot{k}_g'(\lambda)H_1(s,a) + k_q(\lambda)\dot{k}_g'(\lambda)H_2(s,a) + R(\lambda, s, a)$$

with $k_q k_g(\lambda)^{-1} \overset{\lambda\to\infty}{\longrightarrow} 0$ *and* R *of smaller order than* $k_q\dot{k}_g'(\lambda)$,

(d) a^* *is as in Proposition 4.6 such that* $H_1(s,a^*) = \mathbf{0}$,

(e) $\int_{|s|\le\delta} \dot{h}_g(s,a^*)\dot{h}_g(s,a^*)'ds > \mathbf{0}$ *for all* $\delta > 0$.

Then

$$\frac{\dot{k}_g'(\sqrt{n})}{k_q(\sqrt{n})}(\hat{a} - a^*) \overset{d}{\longrightarrow} \left[\int_0^1 \dot{h}_g(V(r),a^*)\dot{h}_g(V(r),a^*)'dr\right]^{-1}\int_0^1 H_2(V(r),a^*)dr$$

as $n \longrightarrow \infty$.

The function T of condition (c) resembles the functional form of the "score", $\dot{Q}_n(a)$. We assume that the $\dot{Q}_n(a)$ can be represented as a sum of H_o*-regular* components of different orders with the leading term equal to zero at a^* (condition (d)). So we are considering the case when the fitted model is correctly specified up to some lower order H_o*-regular* term (see also Proposition 4.2). The limit distribution is comparable with that of Proposition 4.2. Under misspecification the order of the score increases by $\sqrt{n}k_q(\sqrt{n})$. This translates to a direct reduction in the estimator's convergence rate by the same amount.

EXAMPLE 4.9:

Let $f(s,\theta_o) = \theta_o s^2 1\{s > 0\}$ and $g(s,a) = s^3(1+as)^{-1}1\{s > 0\}$ with $A \subset \mathbf{R}_+$. f is *H-regular* with asymptotic order $k_f(\sqrt{n}) = n$ and g is H_o*-regular* on A with asymptotic order $k_g(\sqrt{n}) = n$. It follows from Proposition 4.6 that $\hat{a} \overset{p}{\longrightarrow} a^* = \theta_o^{-1}$. which is interior in A provided θ_o^{-1} is interior in A. Now

$\dot{g}(s,a) = -\frac{s^4}{(1+as)^2}1\{s>0\}$ is H_o-*regular* on A with $\dot{k}_g(\sqrt{n}) = n$,

$\ddot{g}(s,a) = \frac{2s^5}{(1+as)^3}1\{s>0\}$ is H_o-*regular* on A with $\ddot{k}_g(\sqrt{n}) = n$,

$H_1(s,a) = -a^{-1}s^2(\theta_o s^2 - a^{-1}s^2)1\{s>0\}$ with $H_1(s,a^*) = 0$,

$H_2(s,a) = -a^{-3}s^3 1\{s>0\}$ and $k_q(\sqrt{n}) = \sqrt{n}$,

$\dot{h}_g^2(s,a^*) = (a^{*-1}s)^4\, 1\{s>0\}$ which satisfies (d).

It follows from Proposition 4.12 that:

$$\sqrt{n}(\hat{a} - \theta_o^{-1}) \xrightarrow{d} \left[\theta_o^2 \int_0^\infty s^4 L(1,s)ds\right]^{-1} \int_0^\infty s^3 L(1,s)ds$$

It is difficult to establish $A6$ for general *H-regular* functions as the order of these functions depends on a. Alternatively consider the following assumption:

A8 : *There is a sequence μ_n such that $\mu_n v_n^{-1} \to 0$ as $n \to \infty$ and*

$$\sup_{a \in N_n} \left\| \mu_n^{-1} \left(\ddot{Q}_n(a) - \ddot{Q}_n(a^*) \right) \mu_n'^{-1} \right\| = o_p(1),$$

where $N_n = \{a : \|\mu_n'(a - a^*)\| \le 1\}$.

Assumption 8 requires the "Hessian" $\ddot{Q}_n(a)$ to converge to $\ddot{Q}_n(a^*)$ unifromly over a shrinking neibourhood of the value a^*. The following result due to Wooldridge (1994). It is utilised by P&P to obtain limit distribution results for general *H-regular* models under correct specification.

THEOREM 4.1: (Wooldridge (1994))

Assume $A1 - A4$, $A7a$ and $A8$ hold. Then as $n \to \infty$

$$v_n'(\hat{a} - a^*) \xrightarrow{d} -\ddot{Q}(a^*)^{-1}\dot{Q}(a^*).$$

Note that this approach of obtaining limit distribution results does not require $\hat{a} - a^* = o_p(1)$. Actually the convergence of $\hat{a}$ to a^* follows from the limit distribution result itself. Under functional form misspecification $A7b$ holds instead of $A7a$. Theorem 4.1 can be extended under $A7b$. Redefine the set N_n of $A8$ as $N_n^* = \{a : \|s_n^{-1}\mu_n'(a - a^*)\| \leq 1\}$ and consider the following modification of $A8$.

A8* : *There is a sequence* μ_n *such that* $\mu_n v_n^{-1} \to 0$ *as* $n \to \infty$ *and*

$$\sup_{a \in N_n^*} \left\| \mu_n^{-1} \left(\ddot{Q}_n(a) - \ddot{Q}_n(a^*) \right) \mu_n'^{-1} \right\| = o_p(1)$$

Then Theorem 4.1 can be restated as follows.

PROPOSITION 4.13:

Assume $A1 - A4$, $A7b$ *and* $A8^*$ *hold. Then*

$$s_n^{-1} v_n'(\hat{a} - a^*) \xrightarrow{d} -\ddot{Q}(a^*)^{-1}\dot{Q}(a^*)$$

as $n \to \infty$.

Before we present the limit distribution result for $\hat{a}$ we need to introduce the following notation. Define a neighborhood of a^* by

$$N(\varepsilon, \lambda) = \left\{ a : \left\| k_q(\lambda, a^*)^{-1} \dot{k}_g(\lambda, a^*)' (a - a^*) \right\| \leq \lambda^{\varepsilon} \right\}$$

Then the following result holds.

PROPOSITION 4.14: (*g H-regular, f H-regular*)

Let f be H-regular and g H-regular on $A \subset \mathbf{R}^m$ with asymptotic order $k_f(\lambda)$ and $k_g(\lambda, a)$ respectively. Moreover assume

(a) *there exists an interior point of A, a^* such that $k_f(\lambda) = k_g(\lambda, a^*)$,*

(b) $f(s, \theta_o) - g(s, a^*) = q(s, a^*)$ *with q, $\dot{g}$ H-regular on A,*

(c) $T(s, a) = \dot{g}(s, a)\,(f(s, \theta_o) - g(s, a))$ *such that*

$$T(\lambda s, a) = k_g(\lambda, a)\dot{k}_g(\lambda, a)' H_1(s, a) + k_q(\lambda, a)\dot{k}_g(\lambda, a)' H_2(s, a) + R(\lambda, s, a)$$

with H_1, H_2 regular, $k_q(\lambda, a) \in \mathbf{R}$, $k_q(\lambda, a)k_g(\lambda, a)^{-1} \overset{\lambda\to\infty}{\rightarrow} 0$ for all $a \in A$ and R of smaller order than k_q,

(d) $H_1(s, a^*) = \mathbf{0}$,

(e) $\int_{|s|\leq\delta} \dot{h}_g(s, a^*)\dot{h}_g(s, a^*)' ds > \mathbf{0}$ *for all $\delta > 0$,*

(f) *for any $\bar{s} > 0$, there exists $\varepsilon > 0$ such that as $\lambda \to \infty$,*

$$\left\| k_q(\lambda, a^*) \left(\dot{k}_g(a^*) \otimes \dot{k}_g(a^*)\right) (\lambda)^{-1} \left(\sup_{|s|\leq\bar{s}} |\ddot{g}(\lambda s, a^*)|\right) \right\| \to 0, \tag{17}$$

$$\lambda^{\varepsilon} \left\| k_q(\lambda, a^*) \left(\dot{k}_g(a^*) \otimes \dot{k}_g(a^*)\right) (\lambda)^{-1} \left(\sup_{|s|\leq\bar{s}} \sup_{a\in N(\varepsilon,\lambda)} |\ddot{g}(\lambda s, a)|\right) \right\| \to 0, \tag{18}$$

$$\lambda^{\varepsilon} \left\| k_q(\lambda, a^*) \left(\dot{k}_g(a^*) \otimes \dot{k}_g(a^*) \otimes \dot{k}_g(a^*)\right) (\lambda)^{-1} \left(\sup_{|s|\leq\bar{s}} \sup_{a\in N(\varepsilon,\lambda)} |\dddot{g}(\lambda s, a)|\right) \right\| \to 0; \tag{19}$$

Then

$$\frac{\dot{k}_g(\sqrt{n}, a^*)'}{k_q(\sqrt{n}, a^*)}(\hat{a} - a^*) \xrightarrow{d} \left[\int_0^1 \dot{h}_g(V(r), a^*)\dot{h}_g(V(r), a^*)' dr\right]^{-1} \int_0^1 H_2(V(r), a^*) dr$$

as $n \to \infty$.

Proposition 4.14 provides sufficient conditions for Proposition 4.13 to hold. Conditions (a)-(e) in Proposition 4.14 are similar to the conditions in Proposition 4.12. Condition (f) is sufficient for condition $A8^*$. The limit distribution result is comparable with the one of Proposition 4.12. In this case the speed of convergence depends on the pseudo-true value a^*.

2.5 Conclusion

The purpose of this chapter has been to examine the consequences of functional form misspecification when the data are strongly dependent. For this reason the theoretical framework was kept simple. In particular we have considered regressions with just a single covariate. Some of the results provided here can be easily extended to multiple regressions with additively separable components along the lines of Chang, Park and Phillips (2001). A comprehensive generalisation of our results however, to a multiple regression context is a quite challenging task. The econometric techniques utilised by Chang, Park and Phillips (2001) to obtain consistency and limit distribution results for multivariate models, were developed for correctly specified models and unfortunately their applicability is limited in our framework.

We have seen that in contrast to the stationary and the weakly dependent case, convergence to a pseudo-true value does not always hold when the covariates are unit root processes. In particular when OLS is the relevant estimation procedure the estimator may diverge in probability when the true model is of different order than the fitted model. This is not the case though when the NLS procedure is under consideration. The NLS estimator is defined on a compact set and as result it converges to boundary a point in the parameter space when f and g are of different orders. When the pseudo-true value is a boundary point the standard techniques used to obtain limit distribution results (e.g. Andrews (1999), Phillips and Moon (2003)) are not applicable as the limit objective function does not attain its minimum at a turning point. When the pseudo-true value is an interior point, the rate of convergence is the same as that under correct specification ($\sqrt[4]{n}$) when both f and g are *I-regular*. In this case the limit distribution is mixed normal but with larger variance than the one attained under correct functional form. In almost any other case the convergence rate is slower. Moreover the limit distribution theory is different than the one under correct specification.

The results provided here are not only interesting from a theoretical point of view. They are useful for the development of specification tests and model selection procedures. A convergence to pseudo-true value result is required to determine the asymptotic power of functional form testing procedures like the ones proposed by White (1981) and Newey (1985) (tests without specific alternative). Moreover knowledge about the estimator's convergence rate under functional form misspecification is required for obtaining limit distribution as well as asymptotic power results for testing procedures like the ones proposed by Cox (1961, 1962), Davidson and McKinnon (1981) and Voung (1989) (tests with specific alternative). The theory developed by Park and Phillips provides the applied worker with numerous specifications at his disposal. Given this wide range of models, the applied worker is faced with problem of choosing the appropriate one. The next two chapters address this problem. We develop a conditional test for functional form for regression models with unit roots.

2.6 Appendix to Chapter 2

PROOF OF PROPOSITION 4.1: First note that the OLS estimator is $\hat{a}$ is

$$\hat{a} = \frac{\sum_{t=1}^n f(x_t)g(x_t)}{\sum_{t=1}^n g(x_t)^2} + \frac{\sum_{t=1}^n g(x_t)u_t}{\sum_{t=1}^n g(x_t)^2}.$$

(i) We have from Theorem 3.1

$$\hat{a} = \frac{\sum_{t=1}^n f(x_t,\theta_o)g(x_t)}{\sum_{t=1}^n g(x_t)^2} + O_p(n^{-\frac{1}{4}}) = \frac{\int_{-\infty}^{\infty} f(s,\theta_o)g(s)ds}{\int_{-\infty}^{\infty} g(s)^2 ds} + o_p(1).$$

Also

$$\sqrt[4]{n}\,(\hat{a} - a^*) = \frac{\frac{1}{\sqrt[4]{n}}\left[\sum_{t=1}^n f(x_t,\theta_o)g(x_t) - a^* \sum_{t=1}^n g(x_t)^2\right]}{\frac{1}{\sqrt{n}}\sum_{t=1}^n g(x_t)^2} + \frac{\frac{1}{\sqrt[4]{n}}\sum_{t=1}^n g(x_t)u_t}{\frac{1}{\sqrt{n}}\sum_{t=1}^n g(x_t)^2}.$$

In view of the fact that $f(.,\theta_o)g(.) - a^*g(.)^2$ is of zero energy, the result follows by Theorems 3.1 and 3.2.

(ii) We have from Theorems 3.1 and 3.3 that

$$(\sqrt{n}k_{gf}(\sqrt{n}))^{-1}\hat{a} = \frac{\frac{1}{nk_{gf}}\sum_{t=1}^n f(x_t,\theta_o)g(x_t)}{\frac{1}{\sqrt{n}}\sum_{t=1}^n g(x_t)^2} + O_p(n^{-\frac{3}{4}}k_{gf}(\sqrt{n})^{-1}) \xrightarrow{d} \frac{\int_0^1 h_{gf}(V(r),\theta_o)dr}{\int_{-\infty}^{\infty} g(s)^2 ds L(1,0)} +$$

$o_p(1)$.

■

PROOF OF PROPOSITION 4.2:

(i) We have

$$\frac{k_g(\sqrt{n})}{k_q(\sqrt{n})}(\hat{a} - a^*) = \frac{\frac{1}{k_g n k_t}\sum_{t=1}^n g(x_t)q(x_t,\theta_o)}{\frac{1}{nk_g^2}\sum_{t=1}^n g(x_t)^2} + o_p(1) = \frac{\int_0^1 h_g(V(r))h_q(V(r),\theta_o)dr}{\int_0^1 h_g(V(r))^2 ds} + o_p(1).$$

where the first equality holds from the assumption that $f - a^*g = q$ and the limit distribution result from Theorem 3.4.

(ii) We have

$$\frac{k_g(\sqrt{n})}{k_f(\sqrt{n})}\hat{a} \;=\; \frac{\frac{1}{nk_gk_f}\sum_{t=1}^{n} f(x_t,\theta_o)g(x_t)}{\frac{1}{nk_g^2}\sum_{t=1}^{n} g(x_t)^2} + O_p(n^{-\frac{1}{2}}k_f(\sqrt{n})^{-1}) \;=\; \frac{\int_0^1 h_f(V(r),\theta_o)h_g(V(r))dr}{\int_0^1 h_g(V(r))^2dr} + o_p(1).$$

(iii) The proof is the same as (ii) and therefore omitted.

(iv) For $k_g(\lambda) \overset{\lambda\to\infty}{\longrightarrow} 0$ we have

$$\sqrt{n}k_g(\sqrt{n})^2\hat{a} \;=\; \frac{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} f(x_t,\theta_o)g(x_t)}{\frac{1}{nk_g^2}\sum_{t=1}^{n} g(x_t)^2} + O_p(n^{-\frac{1}{2}}k_g(\sqrt{n})^2) \;=\; \frac{\int_{-\infty}^{\infty} f(s,\theta_o)g(s)dsL(1,0)}{\int_0^1 h_g(V(r))^2ds} + o_p(1)$$

where the last equality holds from Theorems 3.1 and 3.4.

For $k_g(\lambda) \overset{\lambda\to\infty}{\longrightarrow} \infty$ from Theorem 3.4 we have

$$\sqrt{n}k_g(\sqrt{n})\hat{a} = \frac{\frac{1}{\sqrt{n}k_g}\sum_{t=1}^{n} g(x_t)u_t}{\frac{1}{nk_g^2}\sum_{t=1}^{n} g(x_t)^2} + O_p(k_g(\sqrt{n})^{-1}) \overset{d}{\to} \frac{\int_0^1 h_g(V(r))dU(r)}{\int_0^1 h_g(V(r))^2ds}.$$

(v) Under the assumption $O_p(n^{-\frac{1}{2}}k_{gf}(\sqrt{n})^{-1}k_g(\sqrt{n})) = o_p(1)$ and Theorem 3.4 we have

$$\frac{k_g^2(\sqrt{n})}{k_{gf}(\sqrt{n})}\hat{a} = \frac{\frac{1}{nk_{gf}}\sum_{t=1}^{n} f(x_t,\theta_o)g(x_t)}{\frac{1}{nk_g^2}\sum_{t=1}^{n} g(x_t)^2} + O_p(n^{-\frac{1}{2}}k_{gf}(\sqrt{n})^{-1}k_g(\sqrt{n})) \overset{d}{\to} \frac{\int_0^1 h_{gf}(V(r),\theta_o)dr}{\int_0^1 h_g(V(r))^2dr}.$$

■

For the proof of Propositions 4.3-4.8 note that

$$\begin{aligned} D_n(a,a^*) &= \sum_{t=1}^{n}\{[(f(x_t,\theta_o)-g(x_t,a)]^2+[(f(x_t,\theta_o)-g(x_t,a)]\,u_t)\} \\ &\quad -\sum_{t=1}^{n}\{[(f(x_t,\theta_o)-g(x_t,a^*)]^2+[(f(x_t,\theta_o)-g(x_t,a^*)]\,u_t)\} \end{aligned}$$

PROOF OF PROPOSITION 4.3: The objective function can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{n}}D_n(a,a^*) \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}[(f(x_t,\theta_o)-g(x_t,a)]^2-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}[(f(x_t,\theta_o)-g(x_t,a^*)]^2+o_p(1) \\
&= \left(\int_{-\infty}^{\infty}[f(s,\theta_o)-g(s,a)]^2ds-\int_{-\infty}^{\infty}[f(s,\theta_o)-g(s,a^*)]^2ds\right)L(1,0)+o_p(1)
\end{aligned}$$

In view of the I-regularity of f and g the first equality holds by Lemma A7(b) in P&P and the second by Theorem 3.1 and Lemma A6(b) in P&P. The convergence holds uniformly in $a \in A$ and in view of condition (c) and Lemma A8(b) in P&P the requisite result follows.■

PROOF OF PROPOSITION 4.4: The objective function can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{n}}D_n(a,a^*) \\
&= \frac{1}{\sqrt{n}}\left[\sum_{t=1}^{n}(g(x_t,a)^2-2f(x_t,\theta_o)g(x_t,a)-\sum_{t=1}^{n}g(x_t,a^*)^2-2f(x_t,\theta_o)g(x_t,a^*)\right] \\
&\quad +o_p(1) \\
&= \left(\int_{-\infty}^{\infty}g(s,a)^2-2f(s,\theta_o)g(s,a)ds-\int_{-\infty}^{\infty}g(s,a^*)^2-2f(s,\theta_o)g(s,a^*)ds\right)L(1,0) \\
&\quad +o_p(1)
\end{aligned}$$

In view of the I-regularity of g and gf the first equality holds by Lemma A7(b) in P&P and the second by Theorem 3.1 and Lemma A6(b) in P&P. The convergence holds uniformly in $a \in A$ and in view of condition (c) and Lemma A8(b) in P&P the requisite result follows.■

PROOF OF PROPOSITION 4.5: The objective function can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{n}} D_n(a, a^*) \\
= \;& \frac{1}{nk_{gf}} \sum_{t=1}^{n} 2f(x_t, \theta_o) g(x_t, a^*) - \frac{1}{nk_{gf}} \sum_{t=1}^{n} 2f(x_t, \theta_o) g(x_t, a) + o_p(1) \\
= \;& \left(\int_{-\infty}^{\infty} 2h_{gf}(s, \theta_o, a^*) L(1, s) ds - \int_{-\infty}^{\infty} 2h_{gf}(s, \theta_o, a) L(1, s) ds \right) + o_p(1)
\end{aligned}$$

The first equality holds by conditions (a), (b) and Lemma A7(c) in P&P and the second by Theorem 3.3. The convergence holds uniformly in $a \in A$ and in view of condition (c) and Lemma A8(a) in P&P the requisite result follows.■

PROOF OF PROPOSITION 4.6: The objective function can be written as

$$\begin{aligned}
&\frac{1}{nk_g} D_n(a, a^*) \\
= \;& \frac{1}{nk_g} \sum_{t=1}^{n} [(f(x_t, \theta_o) - g(x_t, a)]^2 - \frac{1}{nk_g} \sum_{t=1}^{n} [(f(x_t, \theta_o) - g(x_t, a^*)]^2 + o_p(1) \\
= \;& \int_{-\infty}^{\infty} [h_f(s, \theta_o) - h_g(s, a)]^2 L(1, s) ds - \int_{-\infty}^{\infty} [h_f(s, \theta_o) - h_g(s, a^*)]^2 L(1, s) ds \\
&+ o_p(1)
\end{aligned}$$

Given conditions (a) and (b) the first equality holds by Lemma A7(c) in P&P and the second by Theorem 3.3 and Lemma A6(c) in P&P. The convergence holds uniformly in $a \in A$. In view of condition (c) and Lemma A8(a) in P&P the requisite result follows.■

PROOF OF PROPOSITION 4.7: The objective function can be written as

$$
\begin{aligned}
&\frac{1}{nk_g}D_n(a,a^*)\\
&= \frac{1}{nk_g}\left[\sum_{t=1}^{n}(g(x_t,a)^2-2f(x_t,\theta_o)g(x_t,(a)-\sum_{t=1}^{n}g(x_t,a^*)^2-2f(x_t,\theta_o)g(x_t,a^*)\right]\\
&+O_p(\frac{k_f}{k_g})\\
&= \left(\int_{-\infty}^{\infty}h_g(s,a)^2L(1,s)ds-\int_{-\infty}^{\infty}h_g(s,a^*)^2L(1,s)ds\right)\\
&+o_p(1)
\end{aligned}
$$

The first equality holds by conditions (a), (b) and Lemma A7(c) in P&P and the second by Theorem 3.3 and Lemma A6(c) in P&P. The convergence holds uniformly in $a \in A$. In view of condition (c) and Lemma A8(a) in P&P the requisite result follows.■

PROOF OF PROPOSITION 4.8: The objective function can be written as

$$
\begin{aligned}
&\frac{1}{nk_gk_f}D_n(a,a^*)\\
&= \frac{1}{nk_gk_f}\sum_{t=1}^{n}2f(x_t,\theta_o)g(x_t,a^*)-\frac{1}{nk_gk_f}\sum_{t=1}^{n}2f(x_t,\theta_o)g(x_t,a)+O_p\left(\frac{k_g}{k_f}\right)\\
&= \left(\int_{-\infty}^{\infty}2h_f(s,\theta_o)h_g(s,a^*)L(1,s)ds-\int_{-\infty}^{\infty}2h_f(s,\theta_o)h_g(s,a)L(1,s)ds\right)+o_p(1)
\end{aligned}
$$

The first equality holds by conditions (a), (b) and Lemma A7(c) in P&P and the second by Theorem 3.3. The convergence holds uniformly in $a \in A$. In view of condition (c) and Lemma A8(a) in P&P the requisite result follows.■

PROOF OF PROPOSITION 4.9: Fix $\delta > 0$ and define $A_o = \{|a - a^*| \geq \delta\} \subset A$. Let $\bar{a}$ be an arbitrary point in A_o and let N be a neighborhood of $\bar{a}$ given in condition (i).

Define

$$A_n(a, a^*) = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t, a) - g(x_t, a^*)\right)^2,$$
$$B_n(a, a^*) = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t, a) - g(x_t, a^*\right) u_t,$$
$$C_n(a, a^*) = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t, a) - g(x_t, a^*)\right)q(x_t, a^*).$$

Given conditions (a) and (b) the objective function $D_n(a, a^*)$ can be expressed as

$$n^{-1}D_n(a, a^*) = A_n - 2B_n - 2C_n.$$

Now it follows from condition (i) and inequality (49) in P&P that

$$k_q(\sqrt{n}, a^*)^{-2}A_n \overset{a.s.}{\to} \infty \tag{20}$$

uniformly in N. Moreover since

$$\frac{1}{n}\sum_{t=1}^{n} u_t^2 = O_p(1),$$
$$\frac{1}{n}\sum_{t=1}^{n} q(x_t, a^*)^2 = O_p\left(k_q(\sqrt{n}, a^*)^2\right),$$

it follows from the Cauchy-Schwarz inequality and (20) that

$$A_n^{-1}|B_n| \leq A_n(a, a^*)^{-1/2}O_p(1) = o_p(1), \tag{21}$$
$$A_n^{-1}|C_n| \leq k_q(\sqrt{n}, a^*)A_n(a, a^*)^{-1/2}O_p(1) = o_p(1),$$

uniformly in $a \in A$. Now from (21) we have

$$\begin{aligned} n^{-1}D_n(a,a^*) &\geq A_n\left(1 - 2A_n^{-1}\left|B_n\right| - 2A_n^{-1}\left|C_n\right|\right) \\ &= A_n(1+o_p(1)) \xrightarrow{p} \infty, \end{aligned}$$

uniformly in N. Since A_o is compact and $\bar{a}$ an arbitrary point we have

$$n^{-1} \inf_{a \in A_o} D_n(a,a^*) \xrightarrow{p} \infty$$

and the result follows.

■

PROOF OF PROPOSITION 4.10: Fix $\delta > 0$ and define $A_o = \{|a - a^*| \geq \delta\} \subset A$. Let $\bar{a}$ be an arbitrary point in A_o and let N be a neighborhood of $\bar{a}$ given in condition (i). We first show that condition (i) is not satisfied for a^* that is not the lower boundary point in A.

Suppose that a^* is not the lower boundary point in A. Then we can choose $\bar{a} < a^*$. In that case the term in condition (i) is

$$\inf_{\substack{|p-\bar{p}|<\varepsilon \\ |c-\bar{c}|<\varepsilon}} \inf_{a \in N} \left|pk_g(\lambda,a)k_g(\lambda,a^*)^{-1} - c\right| \rightarrow \inf_{|c-\bar{c}|<\varepsilon} |c| < \infty$$

as $\lambda \rightarrow \infty$.

Now we prove that the condition in Theorem 4.1 holds.

Define

$A_n = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t,a) - g(x_t,a^*)\right)^2$,

$B_n = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t,a) - g(x_t,a^*\right) u_t$,

$C_n(a,a^*,\theta_o) = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t,a) - g(x_t,a^*)\right) f(x_t,\theta_o)$,

$E_n = \frac{1}{n}\sum_{t=1}^{n}\left(g(x_t,a) - g(x_t,a^*)\right)g(x_t,a^*).$

and note that the objective function $D_n(a, a^*)$ is

$$n^{-1}D_n(a,a^*) = A_n - 2B_n - 2C_n - 2E_n$$

Now it follows from condition (i) and inequality (49) in P&P that

$$k_g(\sqrt{n},a^*)^{-2}A_n \overset{a.s.}{\to} \infty \tag{22}$$

uniformly in N. Moreover since

$$\frac{1}{n}\sum_{t=1}^{n}u_t^2 = O_p(1),$$
$$\frac{1}{n}\sum_{t=1}^{n}f(x_t,\theta_o)^2 = O_p\left(k_f(\sqrt{n},\theta_o)^2\right),$$
$$\frac{1}{n}\sum_{t=1}^{n}g(x_t,a^*)^2 = O_p\left(k_g(\sqrt{n},a^*)^2\right),$$

it follows from the Cauchy-Schwarz inequality and (22) that

$$\begin{aligned} A_n^{-1}\left|B_n\right| &\le A_n(a,a^*)^{-1/2}O_p(1) = o_p(1), \\ A_n^{-1}\left|C_n\right| &\le k_f(\sqrt{n},\theta_o)A_n(a,a^*)^{-1/2}O_p(1) = o_p(1), \\ A_n^{-1}\left|E_n\right| &\le k_g(\sqrt{n},a^*)A_n(a,a^*)^{-1/2}O_p(1) = o_p(1), \end{aligned} \tag{23}$$

uniformly in $a \in A$. Now from (23) we have that

$$\begin{aligned} n^{-1}D_n(a,a^*) &\ge A_n\left(1 - 2A_n^{-1}\left|B_n\right| - 2A_n^{-1}\left|C_n\right| - 2A_n^{-1}\left|E_n\right|\right) \\ &= A_n(1 + o_p(1)) \overset{p}{\to} \infty, \end{aligned}$$

uniformly in N. Since A_o is compact and $\bar{a}$ an arbitrary point we have

$$n^{-1}\inf_{a\in A_o} D_n(a,a^*) \overset{p}{\to} \infty$$

and the result follows.

■

PROOF OF PROPOSITION 4.11: We establish conditions $A1 - A6$. We start with $A1$. Note that $z(s,a^*)$ is of zero energy as a^* is interior in A. Hence $A1$ follows from Therems 3.1 and 3.2. $A2$ and $A3$ follow directly from Theorem 3.1 and Lemma A7(b) of P&P. Moreover

$$\ddot{Q}(a^*) = \int_{-\infty}^{\infty} \dot{g}(s,a^*)\dot{g}(s,a^*)'ds - \int_{-\infty}^{\infty} \ddot{G}(s,a^*)\left(f(s,\theta_o) - g(s,a^*)\right)$$

and therefore $A4$ follows from condition (c). $A5$ follows trivially from the assumption that $\hat{a} = a^* + o_p(1)$ with a^* interior in A. Finally for $A6$ note that from Theorem 3.1 and Lemma A7(b) in P&P

$$\ddot{Q}_n(a) \xrightarrow{p} \int_{-\infty}^{\infty} \dot{g}(s,a)\dot{g}(s,a)'ds - \int_{-\infty}^{\infty} \ddot{G}(s,a)\left(f(s,\theta_o) - g(s,a)\right)$$

uniformly in A and in view of the convergence of $\hat{a}$ to a^* the result follows.■

PROOF OF PROPOSITION 4.12: We establish conditions $A1 - A6$. Under condition (c) $A1 - A3$ follow directly by Theorem 3.4 and Lemma A7(c) in P&P. Also

$$\ddot{Q}(a^*) = \int_0^1 \dot{h}_g(V(r),a^*)\dot{h}_g(V(r),a^*)'dr = \int_{-\infty}^{\infty} \dot{h}_g(s,a^*)\dot{h}_g(s,a^*)'L(1,s)ds,$$

and therefore $A4$ follows from condition (e). Moreover $A5$ holds trivially under the assumption that $\hat{a} = a^* + o_p(1)$. Now we check $A6$. Set $v_g = n^{1/2}\dot{k}_g(\sqrt{n})$. It follows from Theorem 3.3 and Lemma A7(c) in P&P that

$$v_g^{-1}\sum_{t=1}^{n} \dot{g}(x_t,a)\dot{g}(x_t,a)'v_g'^{-1} \xrightarrow{p} \int_0^1 \dot{h}_g(V(r),a)\dot{h}_g(V(r),a)'dr,$$

uniformly in A. Also from condition (b) and Lemma A7(c) in P&P we have

$$(v_g \otimes v_g)^{-1} \sum_{t=1}^{n} \ddot{g}(x_t, a) u_t = \left(\dot{k}_g \otimes \dot{k}_g\right)^{-1} \ddot{k}_g \left(\frac{\ddot{k}_g^{-1}}{n} \sum_{t=1}^{n} \ddot{g}(x_t, a) u_t \right) \xrightarrow{p} 0,$$

uniformly in $a \in A$. Hence it will suffice to show that

$$(v_g \otimes v_g)^{-1} \sum_{t=1}^{n} \ddot{g}(x_t, \bar{a}) \left(f(x_t, \theta_o) - g(x_t, \bar{a})\right) = o_p(1)$$

which is what we will set out to do. Following P&P, let $K = [s_{\min} - 1, s_{\max} + 1] \times A$ with $s_{\min} = \min_{0 \leq r \leq 1} V(r)$, $s_{\max} = \max_{0 \leq r \leq 1} V(r)$ and consider

$$\begin{aligned} &\left\| \left(\left(\dot{k}_g \otimes \dot{k}_g\right)^{-1} k_g \ddot{k}_g \right) \frac{1}{n} \sum_{t=1}^{n} \ddot{k}_g^{-1} \ddot{g}(x_t, a) \dot{k}_g^{-1} \left(f(x_t, \theta_o) - g(x_t, a)\right) \right\| \\ \leq\ & \left\| \ddot{h}_g \right\|_K \frac{1}{n} \sum_{t=1}^{n} \dot{k}_g^{-1} \left| g(x_t, a^*) - g(x_t, a) + q(x_t, a^*) \right| \\ \xrightarrow{p}\ & \left\| \ddot{h}_g \right\|_K \int_0^1 \left| h_g(V(r), a^*) - h_g(V(r), a) \right| dr, \end{aligned}$$

where the second line is due to the local boundeness of $\ddot{h}_g$ (Lemma A3(b) of P&P). Now because the limit function is continuous in a (Lemma A8(a) of P&P) the result follows.■

PROOF OF PROPOSITION 4.13: The proof is the same with the proof of Theorem 8.1 in Wooldridge (1994) and therefore omitted.■

PROOF OF PROPOSITION 4.14: We will show that $A1 - A4$ and $A8^*$ hold. $A1$ holds from Theorem 3.3 and conditions (c) and (d). To prove that $A2$ holds first define $v_g = n^{1/2} \dot{k}_g^*$. Now it follows from (19) that

$$\left\| (v_g \otimes v_g)^{-1} \sum_{t=1}^{n} \ddot{g}(x_t, a^*) u_t \right\| \leq \left\| \left(\dot{k}_g^* \otimes \dot{k}_g^*\right)^{-1} \left(\sup_{s \leq \bar{s}} \left| \ddot{g}(\sqrt{n} s, a^*) \right| \right) \right\| \frac{1}{n} \sum_{t=1}^{n} |u_t| \xrightarrow{p} 0,$$

and from (19) together with conditions (b), (c) and (d) it follows that

$$\begin{aligned}
&\left\| (v_g \otimes v_g)^{-1} \sum_{t=1}^{n} \ddot{g}(x_t, a^*) \left[q(x_t, a^*)\right] \right\| \\
\leq\ & \left\| k_q(\sqrt{n}, a^*) \left(\dot{k}_g^* \otimes \dot{k}_g^*\right)^{-1} \left(\sup_{s \leq \bar{s}} \left|\ddot{g}(\sqrt{n}s, a^*)\right|\right) \right\| \\
& \left[\frac{1}{n k_q(\sqrt{n}, a^*)} \sum_{t=1}^{n} |q(x_t, a^*)|\right] \\
& \xrightarrow{p} 0.
\end{aligned}$$

which establishes $A2$. Given $A2$, $A3$ follows from Theorem 3.3 and Lemma $A6(b)$ in P&P.

Now we show $A8^*$. Fix δ such that $0 < \delta < \varepsilon/3$, and define $\mu_g = n^{1/2-\delta}\dot{k}_g^*$ and N_n as in $A8^*$. Following P&P we write

$$\ddot{Q}_n(a) - \ddot{Q}_n(a^*) = \left(\ddot{D}_{1n}(a) + \ddot{D}_{1n}(a)'\right) + \ddot{D}_{2n}(a) + \ddot{D}_{3n}(a) + \ddot{D}_{4n}(a) + \ddot{D}_{5n}(a^*) + \ddot{D}_{6n}(a),$$

where

$$\begin{aligned}
\ddot{D}_{1n}(a) &= \sum_{t=1}^{n} \dot{g}(x_t, a^*)(\dot{g}(x_t, (a) - \dot{g}(x_t, a^*))', \\
\ddot{D}_{2n}(a) &= \sum_{t=1}^{n} (\dot{g}(x_t, a) - \dot{g}(x_t, a^*))(\dot{g}(x_t, a) - \dot{g}(x_t, a^*))', \\
\ddot{D}_{3n}(a) &= \sum_{t=1}^{n} \ddot{G}(x_t, a)(g(x_t, a) - g(x_t, a^*)) \\
\ddot{D}_{4n}(a) &= -\sum_{t=1}^{n} \ddot{G}(x_t, a) q(x_t, a^*) \\
\ddot{D}_{5n}(a^*) &= \sum_{t=1}^{n} \ddot{G}(x_t, a^*) q(x_t, a^*) \\
\ddot{D}_{6n}(a) &= -\sum_{t=1}^{n} \left(\ddot{G}(x_t, a) - \ddot{G}(x_t, a^*)\right) u_t
\end{aligned}$$

and define

$$\bar{\omega}_{in}^2(a) = \left\| \mu_g^{-1} \ddot{D}_{in} \mu_g^{-1\prime} \right\|,$$

For $i = 1, ..., 6$. For all $a \in N_n$, we have

$$\bar{\omega}_{1n}^2(a) \leq \sum_{t=1}^{n} \left\|\mu_g^{-1}\dot{g}(x_t, a^*)\right\| \left\|k_q^* \left(\mu_g \otimes \mu_g\right)^{-1} \ddot{g}(x_t, \bar{a})\right\|, \tag{24}$$

$$\bar{\omega}_{2n}^2(a) \leq \sum_{t=1}^{n} \left\|k_q^* \left(\mu_g \otimes \mu_g\right)^{-1} \ddot{g}(x_t, \bar{a})\right\|^2, \tag{25}$$

$$\begin{aligned}\bar{\omega}_{3n}^2(a) \leq & \sum_{t=1}^{n} \left\|\mu_g^{-1}\dot{g}(x_t, a^*)\right\| \left\|k_q^* \left(\mu_g \otimes \mu_g\right)^{-1} \ddot{g}(x_t, \bar{a})\right\| \\ & + \sum_{t=1}^{n} \left\|\left(\mu_g \otimes \mu_g\right)^{-1} \ddot{g}(x_t, \bar{a})\right\| \left\|k_q^* \left(\mu_g \otimes \mu_g\right)^{-1} \ddot{g}(x_t, a)\right\|,\end{aligned} \tag{26}$$

$$\bar{\omega}_{4n}^2(a) \leq \sum_{t=1}^{n} \left\|\left(\mu_g \otimes \mu_g\right)^{-1} \ddot{G}(x_t, a) q(x_t, a))\right\|, \tag{27}$$

$$\bar{\omega}_{5n}^2(a^*) \leq \sum_{t=1}^{n} \left\|\left(\mu_g \otimes \mu_g\right)^{-1} \ddot{G}(x_t, a^*) q(x_t, a^*)\right\|, \tag{28}$$

$$\bar{\omega}_{6n}^2(a) \leq \sum_{t=1}^{n} \left\|k_q^* \left(\mu_g \otimes \mu_g \otimes \mu_g\right)^{-1} \dddot{g}(x_t, \bar{a})\right\| |u_t|, \tag{29}$$

where $\|a^* - \bar{a}\| \leq \|a^* - a\|$. Let $\bar{s} = \max(s_{\max}, -s_{\min}) + 1$. For n large enough we have

$$\sup_{a \in N_n} |\ddot{g}(x_t, a)| \leq \sup_{|s| \leq \bar{s}} \sup_{a \in N_n} \left|\ddot{g}(\sqrt{n}s, a)\right|,$$

for all $t = 1, ..., n$. Now from (24)-(29) we have

$$\begin{aligned}\bar{\omega}_{1n}^2(a) \leq & n^{3\delta} \left\|k_q^* \left(\dot{k}_g^* \otimes \dot{k}_g^*\right)^{-1} \sup_{|s| \leq \bar{s}} \sup_{a \in N_n} \left|\ddot{g}(\sqrt{n}s, a)\right|\right\| \\ & \times \frac{1}{n} \sum_{t=1}^{n} \left\|\dot{k}_g^{*-1}\dot{g}(x_t, a^*)\right\|,\end{aligned} \tag{30}$$

$$\bar{\omega}_{2n}^2(a) \leq n^{4\delta} \left\|k_q^* \left(\dot{k}_g^* \otimes \dot{k}_g^*\right)^{-1} \sup_{|s| \leq \bar{s}} \sup_{a \in N_n} \left|\ddot{g}(\sqrt{n}s, a)\right|\right\|^2, \tag{31}$$

$$\bar{\omega}_{3n}^2(a) \leq n^{3\delta}\left\| k_q^*\left(\dot{k}_g^*\otimes\dot{k}_g^*\right)^{-1}\sup_{|s|\leq\bar{s}}\sup_{a\in N_n}\left|\ddot{g}(\sqrt{n}s,a)\right|\right\| \times\frac{1}{n}\sum_{t=1}^{n}\left\|\dot{k}_g^{*-1}\dot{g}(x_t,a^*)\right\|, +\frac{n^{4\delta}}{2}\left\| k_q^*\left(\dot{k}_g^*\otimes\dot{k}_g^*\right)^{-1}\sup_{|s|\leq\bar{s}}\sup_{a\in N_n}\left|\ddot{g}(\sqrt{n}s,a)\right|\right\|^2, \quad (32)$$

$$\bar{\omega}_{5n}^2(a) \leq n^{2\delta}\left\|\left(\dot{k}_g^*\otimes\dot{k}_g^*\right)^{-1}\sup_{|s|\leq\bar{s}}\sup_{a\in N_n}\left|\ddot{g}(\sqrt{n}s,a)q(\sqrt{n}s,a)\right|\right\|, \quad (33)$$

$$\bar{\omega}_{6n}^2(a^*) \leq n^{2\delta}\left\|\left(\dot{k}_g^*\otimes\dot{k}_g^*\right)^{-1}\sup_{|s|\leq\bar{s}}\left|\ddot{g}(\sqrt{n}s,a^*)\right|\right\| \times\frac{1}{n}\sum_{t=1}^{n}k_q^*\left|q(x_t,a^*)\right|, \quad (34)$$

$$\bar{\omega}_{6n}^2(a) \leq n^{3\delta}\left\| k_q^*\left(\dot{k}_g^*\otimes\dot{k}_g^*\otimes\dot{k}_g^*\right)^{-1}\sup_{|s|\leq\bar{s}}\sup_{a\in N_n}\left|\dddot{g}(\sqrt{n}s,a)\right|\right\|\frac{1}{n}\sum_{t=1}^{n}|u_t|. \quad (35)$$

Now from (30)-(35) together with (17)-(19) we have $\bar{\omega}_{in}^2(a) = o_{a.s.}(1)$, $i = 1, ..., 5$, uniformly in N_n which establishes $A8^*$ and the proof is completed.■

Chapter 3
Testing for Functional Form Under Nonstationarity

3.1 Introduction

Several specification tests for functional form have been developed over time (see for example Ramsey (1969), White (1981), Bierens (1990)). The basis for these tests is to exploit moment conditions that hold when, and only when, the fitted model is of correct functional form (see Newey (1985)). They were originally developed for models with independent and identically distributed data and can be extended to models with weakly dependent data (see Bierens (1990)). In this chapter we examine how a conditional moment test for functional form performs, when the data are are nonstationary. In particular we will assume as in Chapter 2 that the covariates are unit root processes.

The theoretical framework in this chapter is exactly the same with one of Chapter 2. We assume that the functional form of both the true and fitted models belongs to either the *I-regular* or *H-regular* family of transformations defined there. As we have seen, the kinds of models treated in P&P involve completely different limit distribution theory from that involving stationary covariates considered in the literature on functional form testing. The most important difference though - when it comes to functional form misspecification - is that different transformations of unit root processes are of different asymptotic order, while transformations of stationary data are of the same asymptotic order. By allowing the functional form of the true and the fitted model to be either in the *I-regular* or the *H-regular* family we effectively allow the true specification to be of different order than the fitted specification.

As mentioned above this kind of complication does not arise in the stationary framework, and it has an impact on the behaviour of the test under the alternative (incorrect specification). One such effect is that whereas for stationary data, the functional form tests mentioned above are one-sided, in our case they can be two sided. For instance a typical statistic for the testing problem under consideration is of the form:

$$CM_n = \frac{(SM_n)^2}{VN_n}$$

where SM_n is some sample moment and VN_n is a variance normalisation term. For stationary data, under misspecification we have

$$CM_n \sim \frac{O_p(n)}{O_p(1)} \xrightarrow{p} \infty, \ \mathit{as}\ n \to \infty$$

So we reject the null for large values of CM_n. For unit root data the test can be two-sided. VN_n is not necessarily bounded under misspecification. VN_n may diverge, sometimes even faster than $(SM_n)^2$. When this is the case the statistic converges to zero. If $(SM_n)^2 \sim VN_n$ the test is inconsistent.

The asymptotic power of the test is derived for a set of alternatives where each alternative is characterised by the asymptotic order of the true specification. We show that in contrast to the stationary case, under nonstationarity there is not a single divergence rate under FFM. The divergence rate depends on the true model, the fitted model and the nature of any weighting functions employed. Moreover it shown that when integrable weighting functions are used in the construction of the test statistic, better asymptotic power is achieved against a set of alternatives.

We use integrable weightings in two ways. First an integrable weighting may be used in the construction of the sample moment used in the test statistic. This is particularly beneficial when the true model is *I-regular* as it makes the test consistent. Secondly we use integrable weightings for the construction of bounded -under the alternative- variance estimators. The benefit of such variance estimators is more apparent if one is confined to the *H-regular* framework (i.e. both f and g are *H-regular*). We show that in this instance if a standard variance estimator is employed, the test under FFM diverges with the same rate attained under stationarity (n). On the other hand if a bounded variance estimator is employed instead, the test may attain higher divergence rates under incorrect specification. Finally we show that when an *I-regular* weighting is used in the sample moment of the statistic the test may become two sided. In this instance the divergence rate of the test under misspecification can be very large for some alternatives but very poor for some others. This problem is avoided if a bounded variance estimator is employed together with the *I-regular* weighting. The use of a bounded variance estimator in this case makes the test one-sided and results in a single divergence rate ($\sqrt{n}$) under FFM.

It is apparent from what is mentioned above, that we are not considering a single test but a family of tests. Each member of the family is characterised by the nature of the weighting function used in the construction of the sample moment of the statistic, by the nature of variance estimator employed and as we are going to see later by the method used to estimate the fitted model.

The rest of the chapter is organised as follows: In Section 2 the theoretical framework is specified and some preliminary results are given. The main results are provided in Section 3 and Section 4 concludes.

3.2 Theoretical Framework and Preliminary Results

As in Chapter 2, the true model will be

$$y_t = f(x_t, \theta_o) + u_t, \tag{1}$$

where, $f(.) : \mathbf{R} \to \mathbf{R}$ belongs either to the *I-regular or H-regular* family of transformations defined in P&P. Assume that the fitted model is:

$$\hat{y}_t = g(x_t, \hat{a}) + \hat{u}_t, \tag{2}$$

where $g(., a) : \mathbf{R} \times \mathbf{A} \to \mathbf{R}$ is either *I-regular* or *H-regular* with A being a compact subset of $\mathbf{R}^q$ and $\hat{a}$ is some estimator. The variables x_t and u_t will be assumed to be as in Chapter 2. Moreover "correct" and "incorrect" functional form should be understood as in Chapter 2.

REMARK 2.1:

Given the $\mathcal{F}_{t-1}$ measurability of x_t the following conditional moment conditions hold:

Under correct functional form

$$\mathbf{E}\left\{(y_t - g(x_t, a_o)) \,|\mathcal{F}_{t-1}\right\} = 0 \; a.s., \textit{ for a unique } a_o \in A. \tag{3}$$

Under incorrect functional form

$$\mathbf{E}\left\{(y_t - g(x_t, a)) \,|\mathcal{F}_{t-1}\right\} \neq 0 \; a.s., \; for \; all \; a \in A. \tag{4}$$

For the construction of the test statistics three estimators will be considered: NLS, an Instrumental Variables (IV) type of estimator and the Efficient Nonstationary-Nonlinear Least Squares (EN-NLS) proposed by Chang, Park and Phillips (2001) (CPP hereafter). Define the NLS estimator $\hat{a}$ as

$$\hat{a} = \arg\min_{a \in A} Q_n(a), \tag{5}$$

where $Q_n(a) = \sum_{t=1}^{n} (y_t - g(x_t, a))^2$. The Instrumental Variables (IV) estimator with an integrable instrument will be employed as well. Define the IV estimator $\hat{a}$ with instrument $r(.)$ as

$$\hat{a} = \arg\min_{a \in A} Q_n^r(a), \tag{6}$$

where $Q_n^r(a) = \sum_{t=1}^{n} r(x_t) (y_t - g(x_t, a))^2$ and $r(.)$ is *I-regular*. This a special kind of IV estimator as the instrument used is the covariate itself. The IV estimator is usually employed in situations where there is endogeneity bias. In such circumstances, the instrument is usually different than the regressors. Here we consider the IV estimator for a different reason. The choice of the particular instrument makes all the components of the objective function of the same order. As we will see later this is particularly beneficial when the true model is *I-regular*. In this instance the use of this particular IV estimator together with an integrable weighting in the sample moment of our test statistic makes the test consistent.

The asymptotic properties of the IV estimator with an *I-regular* instrument are shown in Proposition 2.1 for correct functional form and Proposition 2.2 for incorrect functional form.

PROPOSITION 2.1:

Let

(a) g *be of correct functional form,*

(b) g, $\dot{g}$, $\ddot{g}$ be *I-regular or H-regular on A,*

(c) $\hat{a} = \arg\min_{a\in A} Q_n^r(a)$ *with* $r(.)$ *I-regular such that* $r(.)g(.,a)$, $r(.)f(.)$ are *I-regular.*

(d) $Q(a) = L(1,0)\int_{-\infty}^{\infty} r(s)\left(g(s,a_o) - g(s,a)\right)^2 ds > 0$ $a.s.$ *for all* $a \neq a_o$ *in* A.

(e) $\int_{-\infty}^{\infty} r(s)\dot{g}(s,a_o)\dot{g}(s,a_o)'ds > 0$ $a.s.$

Then as $n \to \infty$,

$$\sqrt[4]{n}(\hat{a} - a_o) \xrightarrow{d}$$
$$\left(L(1,0)^{1/2}\int_{-\infty}^{\infty} r(s)\dot{g}(s,a_o)\dot{g}(s,a_o)'ds\right)^{-1}\left(\int_{-\infty}^{\infty} r(s)^2\dot{g}(s,a_o)\dot{g}(s,a_o)'ds\right)^{1/2} W(1).$$

REMARK 2.2:

The convergence rate of the NLS estimator is $n^{1/4}$ and $n^{1/2}\dot{k}(\sqrt{n})$ for *I-regular* and *H-regular* g respectively. The convergence rate of the IV estimator under consideration is $n^{1/4}$ irrespective of g. Clearly this is due to the use of an *I-regular* instrument.

PROPOSITION 2.2:

Let

(a) g *be of incorrect functional form,*

(b) g, $\dot{g}$ *and* $\ddot{g}$ be *I-regular or H-regular on A,*

(c) $\hat{a} = \arg\min_{a\in A} Q_r(a)$ *with* $r(.)$ *I-regular such that* $r(.)g(.,a)$, $r(.)f(.)$ are *I-regular.*

(d) a^* *interior in* A :

$L(1,0)\left[\int_{-\infty}^{\infty} r(s)\left(f(s)-g(s,a)\right)^2 ds - \int_{-\infty}^{\infty} r(s)\left(f(s)-g(s,a^*)\right)^2 ds\right] > 0$ *a.s. for all* $a \neq a^*$ *in* A,

(e) $\int_{-\infty}^{\infty} |sz(s,a^*)|\, ds < \infty$, *where* $z(s,a^*) = r(s)\dot{g}(s,a^*)\left(f(s,\theta_o) - g(s,a^*)\right)$,

(f) $\int_{-\infty}^{\infty} \dot{g}(s,a^*)\dot{g}(s,a^*)'ds - \int_{-\infty}^{\infty} \ddot{G}(s,a^*)\left(f(s,\theta_o)-g(s,a^*)\right) ds > \mathbf{0}$.

Then as $n \to \infty$,

$$n^{1/4}\left(\hat{a}-a^*\right) \xrightarrow{d}$$
$$\left[\int_{-\infty}^{\infty} \dot{g}(s,a^*)\dot{g}(s,a^*)'ds - \int_{-\infty}^{\infty} \ddot{G}(s,a^*)\left(f(s,\theta_o)-g(s,a^*)\right) ds\right]^{-1} \times \left[I_1\bar{W}(1) + I_2 W(1)\right],$$

where $I_1 = \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\left|\int_{-\infty}^{\infty} e^{isr} z(r,a^*)dr\right|^2 \frac{1+\phi(s)}{1-\phi(s)}\right) ds\right)^{1/2}$, $I_2 = \left(\int_{-\infty}^{\infty} \dot{g}(s,a^*)\dot{g}(s,a^*)'ds\right)^{1/2}$, ϕ *and* $\bar{W}(1)$ *are the characteristic function and the Gaussian variable respectively defined in Chapter 2.*

REMARK 2.3:

In the previous chapter we have seen that, in some cases the NLS estimator converges to a boundary point of A. Under such circumstances standard techniques (e.g. Andrews (1999), Phillips and Moon (2003)) cannot provide any limit distribution or convergence rate results. This problem is avoided when the IV procedure is employed with I-regular instrument $r(.)$, say. If the tails of $r(.)$ converge faster than polynomial rates, then all the terms in the objective function, $Q_n(a)$, are of the same order and as a result the limit objective function,

$Q(a)$ is minimised at a turning point. The "score" is a zero energy function in the limit i.e.

$$n^{1/2}\dot{Q}_n^r(a^*) \xrightarrow{p} \int_{-\infty}^{\infty} r(s)\dot{g}(s,a^*)\left(f(s)-g(s,a^*)\right)dsL(1,0)$$
$$= \int_{-\infty}^{\infty} z(s,a^*)dsL(1,0)$$

where $\int_{-\infty}^{\infty} z(s,a^*)ds = \mathbf{0}$.

As is mentioned in P&P as long as $\sigma_{\varepsilon u} = \mathbf{E}\left(\varepsilon_{t+1}u_t\right) \neq 0$, the covariance asymptotics for *H-regular* transformations are not mixed normal. As result the estimators are inefficient and the usual t- and chi-squared tests do not have standard limits. This problem is avoided when EN-NLS procedure is employed instead of NLS. The procedure was developed by CPP and is in the spirit of the Fully Modified Least Squares (FM-LS) of Phillips and Hansen (1990). The EN-NLS estimator is defined as

$$\hat{a} = \arg\min_{a\in A} Q_n^+(a) \tag{7}$$

where

$$\begin{aligned} Q_n^+(a) &= \sum_{t=1}^{n}\left(y_t^+ - g(x_t,a)\right)^2,\ y_t^+ = y_t - \hat{\sigma}_{u\varepsilon}\hat{\sigma}_{\varepsilon}^{-2}\hat{\varepsilon}_{l,t+1}, \\ \hat{\sigma}_{u\varepsilon} &= n^{-1}\sum_{t=1}^{n}\hat{u}_t\hat{\varepsilon}_{l,t+1},\ \hat{\sigma}_{\varepsilon}^2 = n^{-1}\sum_{t=1}^{n}\hat{\varepsilon}_{l,t}^2, \end{aligned}$$

$\hat{u}_t$ comes from first step NLS estimation and $\hat{\varepsilon}_{l,t}$ from the regression

$$v_t = \hat{\pi}_1 v_{t-1} + \hat{\pi}_2 v_{t-2} +, \ldots, + \hat{\pi}_l v_{t-l} + \hat{\varepsilon}_{l,t}$$

where l is allowed to to increase as $n \to \infty$ at certain rates. For more details see CPP.

Moreover for the construction of the test statistic one needs an estimator for σ^2, the variance of the error term u_t in (1). P&P have shown that under correct specification

$$\hat{\sigma}^2 = n^{-1}Q_n(\hat{a}) \xrightarrow{P} \sigma^2.$$

Alternatively consider

$$\breve{\sigma}^2 = n^{-1/2}Q_n^w(\hat{a})/w_n$$

where

(i) $Q_n^w(a) = \sum_{t=1}^n w(x_t)\left(y_t - g(x_t, a)\right)^2$,

(ii) $w(.)$ *is I-regular such that* $w(.)f(.)$ *is I-regular and* $w(.)g(., a)$ *is I-regular on* A,

(iii) $w_n = n^{-1/2}\sum_{t=1}^n w(x_t)$.

As shown in Proposition 2.3, $\breve{\sigma}_n^2$ can be used for consistent estimation of σ^2.

PROPOSITION 2.3:

Let Assumptions 2.1 and 2.2 hold and g be of correct functional form.

Then as $n \to \infty$

$$\breve{\sigma}^2 \xrightarrow{P} \sigma^2.$$

The crucial difference between $\hat{\sigma}^2$ and $\breve{\sigma}^2$ is that under incorrect functional form the former may diverge while the latter is bounded in probability. As we are going to see later the use of $\breve{\sigma}^2$ instead of $\hat{\sigma}^2$ may result in better asymptotic power over certain alternatives. A bounded estimator will also be used in conjuction with EN-NLS. Define

$$\breve{\sigma}_+^2 = \breve{\sigma}^2 - \breve{\sigma}_{\varepsilon u}^2\hat{\sigma}_\varepsilon^{-2}$$

where $\breve{\sigma}_{\varepsilon u} = n^{-1/2}\left(\sum_{t=1}^n w(x_t)\hat{u}_t\hat{\varepsilon}_{l,t+1}\right)/w_n$. Then we have the following result:

PROPOSITION 2.4:

Let Assumptions 2.1, 2.2 and Assumption 5.1 in CPP hold. Also let g be of correct functional form.

Then as $n \to \infty$

$$\check{\sigma}_{+}^{2} \xrightarrow{p} \sigma^{2} - \sigma_{\varepsilon u}^{2}\sigma_{\varepsilon}^{-2},$$

where $\sigma_{\varepsilon}^{2} = \mathbf{E}(\varepsilon_{t}^{2})$.

3.3 Detection of Functional Form Misspecification

The functional form test we will consider is based on conditional moment conditions which are valid under correct functional form (Bierens (1990)). Under the null hypothesis we have:

$$H_o : E[(y_t - g(x_t, \tilde{a})r(x_t)|\mathcal{F}_t] = 0,$$

where $\tilde{a} = p\lim \hat{a}$ and $r(x_t)$ is a weighing function. A weighting function is employed in the Bierens test as well. Bierens (1990) suggests that the use of some weighting function may improve power under specific alternatives and in small samples. As we will explain later, in our framework the employment of some weighing function is necessary to make our test consistent when the true model is *I-regular*. Under the alternative hypothesis:

$$H_1 : E[(y_t - g(x_t, \tilde{a})r(x_t)|\mathcal{F}_t] \neq 0.$$

The test is based on the following sample moment expression:

$$\left[\frac{1}{v_n}\sum_{t=1}^{n}(y_t - g(x_t, \hat{a}))\, r(x_t)\right]$$

where v_n is some normalising sequence. For stationary data $v_n = n^{1/2}$. In our framework, the normalising sequence v_n depends on the choice of the weighting function and the fitted model. In particular $v_n = n^{1/4}$ when r is *I-regular*. When both r and g are *H-regular* $v_n = n^{1/2}k_r(n^{1/2})$, with k_r the asymptotic order of r. When r is *I-regular* it will be chosen in a way such that $f(.)r(.)$ and $g(.)r(.)$ are *I-regular* irrespective of f and g. As we are going to see later, the use of an *I-regular* weighting function may improve the asymptotic power of the test in some cases.

Five results are provided in this section. In terms of presentation the first three consider the case where an integrable weighting function is used in the sample moment of the test, while the last two involve a homogeneous weighting functions. Moreover we will assume that the estimator converges to some pseudo-true value under the alternative. When IV is used, this is established by Proposition 2.2. When NLS is the relevant procedure, sufficient conditions for convergence to some pseudo-true value are given in the previous chapter. We assume that the parameter space is compact and convergence to some pseudo-true value will be taken for granted here. In the previous chapter we mentioned that when the fitted model is linear in parameter, the compactness assumption can be dropped. When the parameter space in not bounded and there is FFM, the estimator may not converge to a pseudo-true value. It turns out that most of the results provided below do not differ when the estimator is unbounded in probability. Whenever they do differ, it will be pointed out. Before we present the first result we introduce the following assumption:

ASSUMPTION 3.1:

For $\dot{g}$ and r I-regular or H-regular define

$$\hat{A}^r_n = \sum_{i=1}^n \dot{g}'(x_i, \hat{a}) r(x_i),$$

$$\hat{B}^r_n = \sum_{i=1}^n \dot{g}(x_i, \hat{a}) \dot{g}'(x_i, \hat{a}) r(x_i) \text{ and its inverse when it exists,}$$

$$\hat{B}_n = \sum_{i=1}^n \dot{g}(x_i, \hat{a}) \dot{g}'(x_i, \hat{a}) \text{ and its inverse when it exists.}$$

Then for some normalising matrices v_{1n}, v_{2n}, v_{3n},

$$v_{1n}^{-1} \hat{A}^r_n \xrightarrow{p} A^r,$$

$$v_{2n}^{-1} \hat{B}^r_n v_{2n}'^{-1} \xrightarrow{p} B^r > 0,$$

$$v_{3n}^{-1} \hat{B}_n v_{3n}'^{-1} \xrightarrow{p} B > 0.$$

The test statistic when (2) is the fitted model is:

$$CM_n = \frac{\left[\sum_{t=1}^n (y_t - g(x_t, \hat{a}) - c_n) r(x_t)\right]^2}{\hat{\varsigma}^2 \sum_{t=1}^n \left[\hat{A}^{r\prime}_n \hat{C}_n^{-1} \dot{g}(x_t, \hat{a}) - r(x_t)\right]^2},$$

where $\hat{\alpha}$ is some least squares estimator, $\hat{\varsigma}^2$ is some variance estimator (e.g. $\hat{\sigma}^2$, $\breve{\sigma}^2$), $\hat{C}_n$ will be set equal to $\hat{B}^r_n$ or $\hat{B}_n$ and c_n is a correction term used to induce standard limit distribution. c_n will be set equal to zero unless otherwise specified. Moreover when $r(.)$ is *I-regular* we will need the following regularity condition (RC) that ensures that the test diverges under the alternative.

RC:

$$\int_{-\infty}^{\infty} (f(s, \theta_o) - g(s, a^*)) r(s) ds \neq 0.$$

If RC does not hold, our statistic will be negligible in probability up to some order, rather than bounded up to some order. Clearly the choice of the weighting function $r(.)$ is important. For inctance consider the following weihting function $r(x) = \exp\left[-(x)^2\right]$ that resembles the Gaussian density. This function is centred around zero and threrefore emphasizes differences between f and g at the origin. Bierens (1990) develops a modified version of the conditional moment test we consider in order to ensure that the test is consistent -in the stationary framework- when conditions like RC fail to hold (see also de Jong (1996) for an extension of Bierens test to weakly dependent data). This kind of development will not be attempted here.

We proceed to our first result. First we consider the case where the fitted model is *I-regular*. The asymptotic properties of the test are given in Proposition 3.1:

PROPOSITION 3.1:

Let

(a) $r_j(.)$, $g(., a)$ *and* $\dot{g}(., a)$ *be I-regular with* $a \in A \subset \mathbf{R}^p$,

(b) $\hat{a} = \arg\min_{a \in A} Q_n(a)$, $\hat{\varsigma}^2 = \hat{\sigma}^2$, $\hat{C}_n = \hat{B}_n$,

(c) Assumption 3.1 *holds,*

Then under the null hypothesis,

$$CM_n \xrightarrow{d} \chi_1^2$$

Under the alternative hypothesis provided that RC *holds,*

$$CM_n = \left\{ \begin{array}{l} O_p(\sqrt{n}), \textit{for I-regular } f; \\ O_p(\sqrt{n}), \textit{ if } k_f(\sqrt{n}) \overset{n\to\infty}{\to} 0, \\ O_p(\sqrt{n} k_f(\sqrt{n})^{-2}), \textit{ if } k_f(\sqrt{n}) \overset{n\to\infty}{\to} \infty, \end{array} \right\} \textit{for H-regular } f.$$

The statistic diverges with rate $n^{1/2}$ when the true model is *I-regular*. When the true model is *H-regular* the statistic diverges as long as the true model defines a transformation of asymptotic order, $k_f(\sqrt{n})$, smaller than $n^{1/4}$. For $k_f(\sqrt{n}) \sim n^{1/4}$, CM_{nj} is bounded in probability and for $k_f(\sqrt{n})$ greater than $n^{1/4}$ converges to zero.

In contrast to the stationary case, the estimator $\hat{\sigma}^2$ may exhibit divergent behaviour under incorrect functional form. Whenever $\hat{\sigma}^2$ diverges fast enough (i.e. at rate greater than $n^{1/2}$) the statistic converges to zero. The order of the test consistency may be quite small for a range of alternatives. For instance if the f is *H-regular* with $k_f(\sqrt{n}) \sim n^b$, $b \in (0, 1/2)$ the order of consistency is smaller that $n^{1/2}$ that is attained for *I-regular* f. As we have seen earlier the test is inconsistent, when $b = 1/4$.

REMARK 3.1:

Clearly the choice of the weighting function can improve the power of the test against specific alternatives. In our case an integrable weighting function is necessary for the test to be consistent. If a weighting is not included, the statistic would be bounded in probability under the alternative. To see this consider the following sample moment mentioned above without any weighting:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (y_t - g(x_t, \hat{a}))$$

Under H_o this statistic has a Gaussian limit distribution. Under H_1 the statistic is bounded in probability:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (y_t - g(x_t, \hat{a})) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (f(x_t, \theta_o) - g(x_t, \hat{a})) + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t = O_p(1)$$

for f and g *I-regular*. The use of an integrable weighting function reduces the normalisation rate required, for the sample moment to converge to a standard distribution, under the null, from $\sqrt{n}$ to $\sqrt[4]{n}$, and effectively induces a divergence rate of order $\sqrt[4]{n}$ under the alternative hypothesis.

Now we will consider the case when the fitted model is *H-regular* and the weighting function *I-regular*. Under the alternative hypothesis, we distinguish two cases for *H-regular* f. First we consider the case the true and fitted models are of different asymptotic orders. Secondly we consider the case the fitted model is correctly specified up to some smaller order component $q(x_t, a^*)$. The asymptotic properties of the test are given in Proposition 3.2:

PROPOSITION 3.2:

Let

(a) g, $\dot{g}$ *and* $\ddot{g}$ *be H-regular on* $A \subset \mathbf{R}^p$, r *I-regular such that* $r\dot{g}\dot{g}'$ *is I-regular,*

(b) $\hat{a} = \arg\min_{a \in A} Q_n^r(a)$, $\hat{\varsigma}^2 = \hat{\sigma}^2$, $\hat{C}_n = \hat{B}_n^r$,

(c) Assumption 3.1 *holds,*

Then under the null hypothesis,

$$CM_n \xrightarrow{d} \chi_1^2.$$

Under the alternative hypothesis provided that RC *holds,*

$$CM_n = \begin{cases} \left.\begin{array}{l} O_p(\sqrt{n}),\ \textit{if}\ k_g(\sqrt{n}, a^*) \overset{n\to\infty}{\to} 0, \\ O_p(\sqrt{n}k_g^{-2}(\sqrt{n}, a^*)),\ \textit{if}\ k_g(\sqrt{n}, a^*) \overset{n\to\infty}{\to} \infty, \end{array}\right\} ,\textit{for } f \textit{ I-regular,} \\ \left.\begin{array}{l} O_p(\sqrt{n}/\max\left\{k_f^2(\sqrt{n}), k_g^2(\sqrt{n}, a^*)\right\}),\ \textit{if}\ k_f \neq k_g(a^*) \\ O_p(\sqrt{n}k_q^{-2}(\sqrt{n}, a^*)),\ \textit{if}\ k_f = k_g(a^*)\ \textit{and}\ f - g(a^*) = q(a^*) \end{array}\right\} ,\textit{for } f \textit{ H-regular.} \end{cases}$$

We notice that the test is two sided for *H-regular* g as well. This occurs when $\hat{\sigma}^2$ diverges. The order of consistency varies for different alternatives. For instance if the f is *H-regular* with $\max\{k_f(\sqrt{n}), k_g(\sqrt{n}, a^*)\}$ $(k_f \neq k_g(a^*))$ or $k_q^{-2}(\sqrt{n}, a^*))$,$(k_f \neq k_g(a^*)) \sim n^b$, $b \in (0, 1/2)$ the order of consistency is smaller that $n^{1/2}$ and larger than $n^{1/2}$ otherwise.

REMARK 3.2:

The use of an integrable weighting function together with IV enables us to use the covariance asymptotics of P&P for integrable transformations. So although the fitted model is H-regular, the limit distribution is mixed normal, and the use of a more involved procedure like EN-NLS is avoided.

We have seen that the variance estimator $\hat{\sigma}^2$ may diverge under incorrect specification. As result higher consistency rates are attained for models of higher order. This however comes at the cost of poorer rates for alternatives of lower order and in particular for alternatives of order n^b, $b \in (0, 1/2)$. For these particular alternatives better performance could be expected if the estimator $\breve{\sigma}^2$ is employed instead. Actually if $\breve{\sigma}^2$ were employed, then under incorrect functional form the test is of order $\sqrt{n}$ for all alternatives under consideration.

Now we will consider the same scenario as in of Proposition 3.2. with a bounded estimator for σ^2. The following result holds:

PROPOSITION 3.3:

Let

(a) g, $\dot{g}$ *and* $\ddot{g}$ *be either I-regular or H-regular on* $A \subset \mathbf{R}^p$, r *I-regular such that* $r\dot{g}\dot{g}'$ *is I-regular,*

(b) $\hat{a} = \arg\min_{a \in A} Q_n^r(a)$, $\hat{\varsigma}^2 = \breve{\sigma}^2$, $\hat{C}_n = \hat{B}_n^r$,

(c) Assumption 3.1 *holds,*

Then under the null hypothesis,

$$CM_n \xrightarrow{d} \chi_1^2.$$

Under the alternative hypothesis provided that RC *holds,*

$$CM_n = O_p(\sqrt{n}).$$

We will now consider the case where the weighting function used is *H-regular* when the fitted model is *H-regular* as well. Two versions of the test will be examined. The first involves an unbounded (under the alternative) variance estimator and the second a bounded variance estimator. In both cases the limit distribution under the null is not mixed normal. To resolve this problem two actions are taken. First the fitted model is estimated by EN-NLS . Secondly the sample moment of the test is corrected by the term $\hat{\sigma}_{u\varepsilon}\hat{\sigma}_{\varepsilon}^{-2}\hat{\varepsilon}_{l,t+1}$ ($\breve{\sigma}_{u\varepsilon}\hat{\sigma}_{\varepsilon}^{-2}\hat{\varepsilon}_{l,t+1}$ for the second version of the test). These modifications induce chi-squared limit distribution under the null.

PROPOSITION 3.4:

Let

(a) g, $\dot{g}$ *and* $\ddot{g}$ *be H-regular on* $A \subset \mathbf{R}^p$ *and* r_j *H-regular,*

(b) $\hat{a} = \arg\min_{a \in A} Q_n^+(a)$, $\hat{\varsigma}^2 = \hat{\sigma}_+^2$, $\hat{C}_n = \hat{B}_n$, $c_n = \hat{\sigma}_{u\varepsilon}\hat{\sigma}_\varepsilon^{-2}\hat{\varepsilon}_{l,t+1}$,

(c) Assumption 3.1 *holds,*

Then under the null hypothesis,

$$CM_n \xrightarrow{d} \chi_1^2.$$

Under the alternative hypothesis,

$$CM_n = \begin{cases} \left.\begin{array}{l} O_p(nk_g^2(a^*)), \text{ if } k_g(\sqrt{n}, a^*) \overset{n\to\infty}{\to} 0, \\ O_p(n), \text{ if } k_g(\sqrt{n}, a^*) \overset{n\to\infty}{\to} \infty, \end{array}\right\}, \text{for } f \text{ I-regular and } h_g(x_t, a^*) \neq 0, \\ O_p(n), \text{for } f \text{ H-regular.} \end{cases}$$

First we notice that the use of an *H-regular* weighting in the construction of the variance estimator makes the test one sided. Under the alternative, the divergence rate is smaller than n, when the true model is *I-regular* and the asymptotic order the fitted model is vanishing. In any other case the divergence rate of the test equals n which is the rate attained under stationarity.

REMARK 3.3:

*I*n Proposition 3.4. we assume that $h_g(x_t, a^*)$ is bounded away from zero, when the true model is *I-regular*. When this does not hold and a^* is on the boundary of A, then the power rate cannot be obtained as we do not have any result for the convergence rate of the estimator. When the fitted model is linear in parameter and there is no compactness restriction on the parameter space (or at least the estimator is required to be bounded away from zero), then $h_g(x_t, a^*) = 0$ and the test is inconsistent. This holds for the statistic considered in Proposition 3.5 below as well. This inconsistency problem however is avoided, when the weighting function $r(.)$ is chosen to be *I-regular*.

The explosive behaviour of the variance estimator, when the models are *H-regular*, prevents any realisation of faster rates than those attained under stationarity. It turns out that faster rates can be attained when the bounded variance estimator $\breve{\sigma}_{+}^{2}$ is used instead. This is shown in Proposition 3.5.

PROPOSITION 3.5:

Let

(a) *g, $\dot{g}$ and $\ddot{g}$ be H-regular on $A \subset \mathbf{R}^p$ and r_j H-regular,*

(b) $\hat{a} = \arg\min_{a \in A} Q_n^{+}(a)$, $\hat{\varsigma}^2 = \breve{\sigma}_{+}^{2}$, $\hat{C}_n = \hat{B}_n$, $c_n = \breve{\sigma}_{u\varepsilon}\hat{\sigma}_{\varepsilon}^{-2}\hat{\varepsilon}_{l,t+1}$,

(c) Assumption 3.1 *holds,*

Then under the null hypothesis,

$$CM_n \xrightarrow{d} \chi_1^2.$$

Under the alternative hypothesis,

$$CM_n = \left\{ \begin{array}{l} O_p(n), \textit{for } f \textit{ I-regular and } h_g(x_t, a^*) \neq 0, \\ \left. \begin{array}{l} O_p(n \max\left\{k_f^2(\sqrt{n}), k_g^2(\sqrt{n}, a^*)\right\}), \textit{ if } k_f \neq k_g(a^*) \\ O_p(n k_q^2(\sqrt{n}, a^*)), \textit{ if } k_f = k_g(a^*) \textit{ and } f - g(a^*) = q(a^*) \end{array} \right\}, \textit{for } f \textit{ H-regular.} \end{array} \right.$$

From Proposition 3.5 we see that when a bounded estimator is employed, the divergence rate improves, as long as $k_g(a^*)$ and $k_q(a^*)$ increase as $n \to \infty$.

3.4 Conclusion

A conditional moment test for functional form was considered in regressions with a unit root covariate. We have shown that under nonstationarity, the power properties of the test are quite different from those under stationarity. In contrast to the stationary case, under nonstationarity there is no single divergence rate under the alternative. The divergence rate depends on the true model, the fitted model and the nature of any weighting functions employed. The use of an *I-regular* weighting in the sample moment of the statistic results in a consistent test. Nonetheless an *I-regular* weighting may result in a two-sided test. When the test is two sided, the statistic has good asymptotic power against higher order alternatives, but performs poorly against lower order alternatives. If a bounded variance estimator is employed together with an *I-regular* weighting the test becomes one-sided and attains a single divergence rate ($\sqrt{n}$) under misspecification. The benefit from using a bounded variance estimator is more apparent when, both the true and fitted models are *H-regular*. The use of bounded estimator in this case may result in higher divergence rates than the one attained under stationarity (n).

The bounded variance estimator was constructed in a very simple way. We weight the squares of the regression residuals $\hat{u}(x_t)^2$ say, by an integrable function, $w(x_t)$, of the regression covariate. The term $\hat{u}(x_t)^2 w(x_t)$ is *I-regular* even if $\hat{u}(x_t)^2$ contains *H-regular* components. As result the variance estimator $\sum \left(\hat{u}(x_t)^2 w(x_t)\right) / \sum w(x_t)$ is bounded under FFM. This approach will not work in multiple regression models. Suppose for instance that our model involves *H-regular* functions of two unit root covariates x_{1t} and x_{2t} say. Then under FFM, $\sum \hat{u}(x_t)^2 \left(w(x_{1t}) + w(x_{2t})\right) / \sum \left(w(x_{1t}) + w(x_{2t})\right) \sim \sum \left(\hat{u}_1(x_{1t})^2 w(x_{2t}) + \hat{u}_2(x_{2t})^2 w(x_{1t})\right) / \sum \left(w(x_{1t}) + w(x_{2t})\right) \sim k_{u_1}(\sqrt{n}) + k_{u_2}(\sqrt{n})$ where the last step is due to Lemma 2 of Hu and Phillips (2002). Whether a bounded variance estimator can be obtained in a multi-factor setting remains an open question.

The purpose of this chapter was to examine to what extent functional form testing is different under nonstationarity and to propose an effective testing procedure. For this reason our theoretical framework was kept simple. We have assumed that the regression model involves a single exogenous covariate. Although single factor models are used in financial theory e.g. Park's (2002) nonlinear nonstationary stochastic volatility models, most econometric applications, such as cointegration analysis, would require multi-factor endogenous models. An extension of the current framework in this direction is attempted in Chapter 4.

3.5 Appendix to Chapter 3

We introduce some notation

$G_{qq} = \int_0^1 h_q^2(V(s), a^*)ds$, $G_{rq} = \int_0^1 h_r(V(s))h_q(V(s), a^*)ds$

$G_{q\dot{g}} = \int_0^1 h_q(V(s), a^*)\dot{h}_g(V(s), a^*)ds$, $G_{r\dot{g}} = \int_0^1 h_r(V(s))\dot{h}_g(V(s), a^*)ds$

$G_{\dot{g}} = \int_0^1 \dot{h}_g(V(s), a^*)ds$, $G_{\dot{g}\dot{g}} = \int_0^1 \dot{h}_g(V(s), a^*)\dot{h}_g(V(s), a^*)'ds$;

PROOF OF PROPOSITION 2.1: We first show that $\hat{a} \xrightarrow{p} a_o$. From Theorem 3.1(i) in P&P we have $n^{-1/2}Q_n(a) \xrightarrow{p} Q(a)$ uniformly in a, where

$$Q(a) = L(1,0)\int_{-\infty}^{\infty} r(s)\left(g(s, a_o) - g(s, a)\right)^2 ds$$

Now $Q(.)$ is continuous (Lemma 8(b) of P&P) $a.s.$ and in view of condition (d) is uniquely minimised at a_o. Hence $\hat{a} \xrightarrow{p} a_o$. Now we show the limit distribution result. Following the same arguments as in P&P, from the first order condition and Mean Value Theorem (MVT) we have

$$\dot{Q}_n^r(\hat{a}) = o_p(1) \Rightarrow$$

$$\sqrt[4]{n}\,(\hat{a} - a_o) = \left[\tfrac{1}{\sqrt{n}}\ddot{Q}_n^r(\bar{a})\right]^{-1}\tfrac{1}{\sqrt[4]{n}}\dot{Q}_n^r(a_o) + o_p(1) \text{ (where } \|\hat{a} - \bar{a}\| \leq \|\hat{a} - a_o\|)$$

Now in view of Theorem 3.1 in P&P is straightforward to show that

$$\tfrac{1}{\sqrt[4]{n}}\dot{Q}_n^r(a_o) \xrightarrow{d} \left[L(1,0)\int_{-\infty}^{\infty} r^2(s)\dot{g}(s, a_o)\dot{g}'(s, a_o)ds\right]^{1/2} W(1)$$

$$\tfrac{1}{\sqrt{n}}\ddot{Q}_n^r(\bar{a}) \xrightarrow{p} L(1,0)\int_{-\infty}^{\infty} r(s)\dot{g}(s, a_o)\dot{g}'(s, a_o)ds$$

which establishes the result.■

PROOF OF PROPOSITION 2.2:

We first show that $\hat{a} \xrightarrow{p} a^*$. The same arguments as those in the proof above give $n^{-1/2}Q_n(a) \xrightarrow{p} Q(a)$ uniformly in a, where

$$Q(a) = L(1,0) \int_{-\infty}^{\infty} r(s) \left(f(s) - g(s,a)\right)^2 ds$$

$Q(.)$ is continuous (Lemma 8(b) of P&P) $a.s.$ and in view of condition (d) is uniquely minimised at a^* Hence $\hat{a} \xrightarrow{p} a^*$. From first order condition and Mean Value Theorem (MVT) we have

$$\dot{Q}_n^r(\hat{a}) = o_p(1) \Rightarrow$$
$$\sqrt[4]{n}\,(\hat{a} - a^*) = \left[\tfrac{1}{\sqrt{n}}\ddot{Q}_n^r(\bar{a})\right]^{-1} \tfrac{1}{\sqrt[4]{n}}\dot{Q}_n^r(a^*) + o_p(1)$$

(where $\|\hat{a} - \bar{a}\| \leq \|\hat{a} - a^*\|$)

Using Theorem 3.1 in P&P the term

$$\tfrac{1}{\sqrt{n}}\ddot{Q}_n^r(\bar{a}) \xrightarrow{p} L(1,0) \int_{-\infty}^{\infty} r(s)\dot{g}(s,a^*)\dot{g}'(s,a^*)ds.$$

Now the term

$$\tfrac{1}{\sqrt[4]{n}}\dot{Q}_n^r(a^*) = \tfrac{1}{\sqrt[4]{n}} \sum_{t=1}^{n} r(x_t)\dot{g}(x_t,a^*) \left(f(x_t) - g(x_t,a^*)\right) + \tfrac{1}{\sqrt[4]{n}} \sum_{t=1}^{n} r(x_t)\dot{g}(x_t,a^*)u_t$$

Write the first term in the expression above

$$\tfrac{1}{\sqrt[4]{n}} \sum_{t=1}^{n} r(x_t)\dot{g}(x_t,a^*) \left(f(x_t) - g(x_t,a^*)\right) = \tfrac{1}{\sqrt[4]{n}} \sum_{t=1}^{n} z(x_t,a^*)$$
$$\xrightarrow{d} \left(\tfrac{1}{2\pi} \int_{-\infty}^{\infty} \left(\left|\int_{-\infty}^{\infty} e^{isr} z(r,a^*)dr\right|^2 \tfrac{1+\psi(s)}{1-\psi(s)}\right) dsL(1,0)\right)^{1/2} \bar{W}(1)$$

(from Theorem 3.2 in Chapter 2)

The second term

$$\tfrac{1}{\sqrt[4]{n}} \sum_{t=1}^{n} r(x_t)\dot{g}(x_t,a^*)u_t \xrightarrow{d} \left[L(1,0) \int_{-\infty}^{\infty} r^2(s)\dot{g}(s,a_o)\dot{g}'(s,a_o)ds\right]^{1/2} W(1)$$

and the result follows.■

PROOF OF PROPOSITION 2.3:

$$\breve{\sigma}^2 = n^{-1/2}\left[\sum_{t=1}^{n}(y_t - g(x_t,\hat{a}))^2 w(x_t)\right]/w_n$$

$$= n^{-1/2}\left[\sum_{t=1}^{n}(g(x_t,a_o) - g(x_t,\hat{a}))^2 w(x_t)\right]/w_n$$

$$+2\left[\sum_{t=1}^{n}(g(x_t,a_o) - g(x_t,\hat{a}))\,w(x_t)u_t\right]/w_n + \left[\sum_{t=1}^{n}w(x_t)u_t^2\right]/w_n$$

$$= n^{-1/2}\left[\sum_{t=1}^{n}w(x_t)u_t^2\right]/w_n + o_p(1)$$

(from Lemma 7(a) and Theorem 3.2 in P&P, given the consistency of $\hat{a}$)

$$= n^{-1/2}\left[\sum_{t=1}^{n}w(x_t)\sigma^2\right]/w_n + n^{-1/2}\left[\sum_{t=1}^{n}w(x_t)\left(u_t^2 - \sigma^2\right)\right]/w_n$$

$$= n^{-1/2}\left[\sum_{t=1}^{n}w(x_t)\sigma^2\right]/w_n + o_p(1)$$

(Lemma 7(a) in P&P)

$$= \frac{\sigma^2 L(1,0)\int_0^1 w(s)ds}{L(1,0)\int_0^1 w(s)ds} + o_p(1) = \sigma^2 \text{ as required.}\blacksquare$$

PROOF OF PROPOSITION 2.4:

$$\breve{\sigma}_+^2 = \breve{\sigma}^2 - \left(\breve{\sigma}_{u\varepsilon}\right)^2\sigma_\varepsilon^{-2} = \sigma^2 - \left(\breve{\sigma}_{u\varepsilon}\right)^2\sigma_\varepsilon^{-2} + o_p(1)$$

(from the previous result).

Now $\breve{\sigma}_{u\varepsilon} = n^{-1/2}\left[\sum_{t=1}^{n}w(x_t)\hat{u}_t\hat{\varepsilon}_{l,t+1}\right]/w_n$

$$= n^{-1/2}\left[\sum_{t=1}^{n}(g(x_t,a_o) - g(x_t,\hat{a}))\,w(x_t)\hat{\varepsilon}_{l,t+1} + \sum_{t=1}^{n}w(x_t)u_t\hat{\varepsilon}_{l,t+1}\right]/w_n$$

$$= n^{-1/2}\left[\sum_{t=1}^{n}(g(x_t,a_o) - g(x_t,\hat{a}))\,w(x_t)\varepsilon_{t+1} + \sum_{t=1}^{n}w(x_t)u_t\varepsilon_{t+1}\right]/w_n$$

(CPP proof of Theorem 5.2)

$$= n^{-1/2}\left[\sum_{t=1}^{n}w(x_t)u_t\varepsilon_{t+1}\right]/w_n + o_p(1)$$

(due to Theorem 3.2 in P&P, given the consistency of $\hat{a}$)

$$= n^{-1/2}\left[\sum_{t=1}^{n} w(x_t)\sigma_{\varepsilon u} + \sum_{t=1}^{n} w(x_t)\left(u_t\varepsilon_{t+1} - \sigma_{\varepsilon u}\right)\right]/w_n = n^{-1/2}\sum_{t=1}^{n} w(x_t)\sigma_{\varepsilon u}/w_n +$$

$o_p(1)$

(Lemma 7(a) in P&P)

$$= \frac{\sigma_{u\varepsilon}L(1,0)\int_0^1 w(s)ds}{L(1,0)\int_0^1 w(s)ds} + o_p(1) = \sigma_{u\varepsilon}.$$

Hence

$\check{\sigma}_+^2 \xrightarrow{p} \sigma_+^2$ as required.■

PROOF OF PROPOSITION 3.1:

A. We first show the result under the null hypothesis:

From MVT with $\bar{a}$ being an intermediate point we have

$$CM_n =$$

$$= \frac{n^{-1/2}\left[(\hat{a}-a_o)'\sum_{t=1}^{n}\dot{g}(x_t,\bar{a})r(x_t)-\sum_{t=1}^{n}r(x_t)u_t\right]^2}{\hat{\sigma}^2 n^{-1/2}\sum_{t=1}^{n}\left[\hat{A}_n^{r\prime}\hat{B}_n^{-1}\dot{g}(x_t,\hat{a})+r_j(x_t)\right]^2}$$

$$= \frac{n^{-1/2}\left[\left(\sum_{t=1}^{n}\dot{g}(x_t,\bar{a})'r(x_t)\right)\left(\sum_{t=1}^{n}\dot{g}(x_t,\bar{a})\dot{g}(x_t,\bar{a})'\right)^{-1}\sum_{t=1}^{n}\dot{g}(x_t,a_o)u_t-\sum_{t=1}^{n}r(x_t)u_t\right]^2}{\hat{\sigma}^2 n^{-1/2}\sum_{t=1}^{n}\left[\hat{A}_n^{r\prime}\hat{B}_n^{-1}\dot{g}(x_t,\hat{a})-r_j(x_t)\right]^2}$$

$$= \frac{\left[n^{-1/4}\sum_{t=1}^{n}\left(\hat{A}_n^{r\prime}\bar{B}_n^{-1}\dot{g}(x_t,a_o)-r(x_t)\right)u_t\right]^2}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_o^{r\prime}B_o^{-1}\dot{g}(s,a_o)-r(s)\right)^2 ds+o_p(1)} = \frac{\left[n^{-1/2}\sum_{t=1}^{n}\left(A^{r\prime}B^{-1}\dot{g}(x_t,a_o)-r(x_t)\right)u_t\right]^2+o_p(1)}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_o^{r\prime}B_o^{-1}g(s,a_o)-r(s)\right)^2 ds}$$

(due to Theorem 2.2(i) and Corollary 5.4 in P&P)

$$\xrightarrow{d} \frac{L(1,0)\int_{-\infty}^{\infty}\left(A_o^{r\prime}B_o^{-1}\dot{g}(s,a_o)-r(s)\right)^2 ds W(1)^2}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_o^{r\prime}B_o^{-1}g(s,a_o)-r(s)\right)^2 ds} = (W(1)/\sigma)^2\ (W(1)\sim N(0,\sigma^2))$$

and the last result due to Theorem 2.2(ii) in P&P.

B. Now we show the result under the alternative hypothesis:

$$CM_n = \frac{n^{-1/2}\left[\sum_{t=1}^{n}f(x_t)-g(x_t,\hat{a})+\sum_{t=1}^{n}r(x_t)u_t\right]^2}{\hat{\sigma}_n^2 n^{-1/2}\sum_{t=1}^{n}\left[\hat{A}_n^{r\prime}\hat{B}_n^{-1}\dot{g}(x_t,\hat{a})-r_j(x_t)\right]^2} = \frac{n^{-1/2}\left[\sum_{t=1}^{n}f(x_t)-g(x_t,\hat{a})+\sum_{t=1}^{n}r(x_t)u_t\right]^2}{\hat{\sigma}_n^2 L(1,0)\int_{-\infty}^{\infty}(A^{r\prime}B^{-1}g(s,a^*)-r(s))^2 ds+o_p(1)}$$

$$= \frac{n^{1/2}\left[n^{-1/2}\sum_{t=1}^{n}(f(x_t)-g(x_t,\hat{a}))r(x_t)+n^{-1/4}\sum_{t=1}^{n}r_j(x_t)u_t\right]^2}{\hat{\sigma}^2 L(1,0)\int_{-\infty}^{\infty}\left(A_*^{r\prime}B_*^{-1}g(s,a^*)-r(s)\right)^2 ds+o_p(1)}$$

$$= \frac{n^{1/2}\left[L(1,0)\int_{-\infty}^{\infty}(f(s)-g(s,a^*))r(s)ds+o_p(1)+O_p(n^{-1/4})\right]^2}{\hat{\sigma}^2 L(1,0)\int_{-\infty}^{\infty}\left(A_*^{r\prime}B_*^{-1}g(s,a^*)-r(s)\right)^2 ds} + o_p(1)$$

in view of RC this gives

$$\frac{n^{1/2}\left[L(1,0)\int_{-\infty}^{\infty}(f(s)-g(s,a^*))r_j(s)ds+o_p(1)+O_p(n^{-1/4})\right]^2}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_*'B_*^{-1}g(s,a^*)-r_j(s)\right)^2 ds} = O_p(n^{1/2})$$

for f I-regular or H-regular with $k_f(\lambda) \overset{\lambda\to\infty}{\to} 0$

and

$$\frac{n^{1/2}\left[L(1,0)\int_{-\infty}^{\infty}(f(s)-g(s,a^*))r_j(s)ds+O_p(n^{-1/4})\right]^2}{k_f(n^{1/2})L(1,0)\int_{-\infty}^{\infty}h_f^2(s)L(1,s)ds\int_{-\infty}^{\infty}\left(A_*^{r\prime}B_*^{-1}g(s,a^*)-r_j(s)\right)^2 ds+o_p(k_f(n^{1/2}))} = O_p(n^{1/2}k_f^{-1}(n^{1/2}))$$

for f H-regular with $k_f(\lambda) \overset{\lambda\to\infty}{\to} \infty$.■

PROOF OF PROPOSITION 3.2:

A. Under the null hypothesis:

$$CM_n = \frac{n^{-1/2}\left[(\hat{a}-a_o)'\sum_{t=1}^{n}\dot{g}(x_t,\bar{a})r(x_t)-\sum_{t=1}^{n}r(x_t)u_t\right]^2}{\hat{\sigma}_n^2 n^{-1/2}\sum_{t=1}^{n}\left[\tilde{A}_n'\tilde{B}_n^{r-1}\dot{g}(x_t,\bar{a})r(x_t)-r(x_t)\right]^2}$$

(MVT)

$$= \frac{\left[n^{-1/4}\sum_{t=1}^{n}\left(\hat{A}_n'\bar{B}_n^{-1}\dot{g}(x_t,a_o)r_j(x_t)-r_j(x_t)\right)u_t\right]^2}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_o'B_o^{r-1}g(s,a_o)r(s)-r(s)\right)^2 ds+o_p(1)}$$

$$= \frac{\left[n^{-1/2}\sum_{t=1}^{n}\left(A'B^{-1}\dot{g}(x_t,a_o)r(x_t)-r(x_t)\right)u_t\right]^2+o_p(1)}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_o'B_o^{r-1}g(s,a_o)r(s)-r(s)\right)^2 ds}$$

Hence $CM_n \overset{d}{\to} \frac{L(1,0)\int_{-\infty}^{\infty}\left(A_o'B_o^{-1}\dot{g}(s,a)-r(s)\right)^2 ds W(1)^2}{\sigma^2 L(1,0)\int_{-\infty}^{\infty}\left(A_o'B_o^{r-1}g(s,a_o)r(s)-r(s)\right)^2 ds} = (W(1)/\sigma)^2$.

B. Under the alternative hypothesis: Similar arguments with ones of Proof 3.1 lead

to

$$CM_n = \frac{n^{1/2}\left[L(1,0)\int_{-\infty}^{\infty}(f(s)-g(s,a^*))r(s)ds+o_p(1)+O_p(n^{-1/4})\right]^2}{\hat{\sigma}^2 L(1,0)\int_{-\infty}^{\infty}\left(A_*'B_*^{r-1}g(s,a^*)r(s)-r(s)\right)^2 ds+o_p(1)}$$

hence the behaviour of the statistic is determined by the term $\hat{\sigma}_n^2$.

We consider the following cases for

$$\hat{\sigma}^2 = \sum_{t=1}^{n} u_t^2 + \tfrac{1}{n}\sum_{t=1}^{n}(f(x_t) - g(x_t,\hat{a}))^2 - \tfrac{2}{n}\sum_{t=1}^{n}(f(x_t) - g(x_t,\hat{a}))\,u_t :$$

(i) f I-regular with $k_g(\sqrt{n}, a^*) \overset{n\to\infty}{\to} 0$:

$$\hat{\sigma}^2 = \sigma^2 + o_p(1) + O_p(k_g(\sqrt{n}, a^*)^2) + o_p(k_g(\sqrt{n}, a^*)) = \sigma^2 + o_p(1)$$

(ii) f I-regular with $k_g(\sqrt{n}, a^*) \overset{n\to\infty}{\to} \infty$:

$$\hat{\sigma}^2 = \sigma^2 + o_p(1) + O_p(k_g(\sqrt{n}, a^*)^2) + o_p(k_g(\sqrt{n}, a^*)) = O_p(k_g(\sqrt{n}, a^*)^2)$$

(iii) f H-regular with $k_f \neq k_g(a^*)$:

$$\hat{\sigma}^2 = O_p(k_f(\sqrt{n})^2) + O_p(k_g(\sqrt{n}, a^*)^2) + O_p(k_f(\sqrt{n})k_g(\sqrt{n}, a^*)) + ..$$
$$= O_p(\max\{k_f(\sqrt{n})^2), O_p(k_g(\sqrt{n}, a^*)^2\})$$

where the results above are due to Proposition 4.10 of Chapter 2 and Lemma 7 of P&P.

(iv) f H-regular with $k_f = k_g(a^*)$, $f - g(a^*) = q(a^*)$:

Now from an application of mean value theorem on $(f(x_t) - g(x_t,\hat{a}))^2$ around a^* we have

$$\hat{\sigma}^2 = \sigma^2 + o_p(1) + \tfrac{1}{n}\sum_{t=1}^{n}(q(x_t,a^*) + (\hat{a}-a^*)'\dot{g}(x_t,\bar{a}))^2 + \tfrac{1}{n}\sum_{t=1}^{n}((\hat{a}-a^*)'\dot{g}(x_t,\bar{a})u_t)$$
$$= k_q^2(a^*)\left\{G_{qq} - 2G'_{q\dot{g}}G_{\dot{g}\dot{g}}^{-1}G_{q\dot{g}} + \left[G'_{\dot{g}}G_{\dot{g}\dot{g}}^{-1}G_{q\dot{g}}\right]^2\right\} + o_p\left(k_q^2(a^*)\right) + o_p\left(k_q(a^*)\right)$$
$$= O_p\left(k_q^2(a^*)\right)$$

where the second equality above is due to Proposition 4.13 of Chapter 2 and Lemma 7 of P&P and this completes the proof.■

PROOF OF PROPOSITION 3.3:

A. Under the null hypothesis similar arguments with the ones of Proposition 3.2 together with Proposition 2.3 give the result.

B. Under the alternative hypothesis along the same lines of Proposition 3.2 we have:

$$CM_n = \frac{n^{1/2}\left[L(1,0)\int_{-\infty}^{\infty}(f(s)-g(s,a^*))r(s)ds+o_p(1)+O_p(n^{-1/4})\right]^2}{\breve{\sigma}^2 L(1,0)\int_{-\infty}^{\infty}\left(A_*' B_*^{r-1} g(s,a^*)r(s)-r(s)\right)^2 ds+o_p(1)}$$

Now from Proposition 2.2 and using the same arguments as those in the proof of 2.3 we have

$$\breve{\sigma}^2 =$$

$$= \frac{L(1,0)\int_{-\infty}^{\infty}(f(s)-g(s,a^*))^2 w(s)ds+L(1,0)\sigma^2\int_{-\infty}^{\infty}w(s)ds+o_p(1)}{L(1,0)\int_{-\infty}^{\infty}w(s)ds+o_p(1)}$$

$$= \frac{\int_{-\infty}^{\infty}(f(s)-g(s,a^*))^2 w(s)ds}{\int_{-\infty}^{\infty}w(s)ds} + \sigma^2 + o_p(1)$$

which completes the proof.■

PROOF OF PROPOSITION 3.4:

A. Under the null hypothesis

$$CM_n = \frac{\left[\sum_{t=1}^{n}(y_t-g(x_t,\hat{a})-c_n)r(x_t)\right]^2}{\hat{\sigma}_+^2\sum_{t=1}^{n}\left[\hat{A}_n\hat{B}_n^{-1}g(x_t,\hat{a})+r(x_t)\right]^2} = \frac{\frac{1}{nk_r^2}\left[(\hat{a}-a_o)'\sum_{t=1}^{n}\dot{g}(x_t,\bar{a})r(x_t)-\sum_{t=1}^{n}r(x_t)u_t^+\right]^2}{\hat{\sigma}_+^2\frac{1}{nk_r^2}\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}\dot{g}(x_t,\hat{a})+r(x_t)\right]^2} \text{ (MVT)}$$

$$= \frac{\left[\left(\frac{1}{nk_r}\sum_{t=1}^{n}r(x_t)\dot{g}(x_t,\bar{a})'\dot{k}_g'^{-1}\right)\left(\frac{1}{n}\sum_{t=1}^{n}\dot{k}_g^{-1}\dot{g}(x_t,\bar{a})\dot{g}(x_t,\bar{a})'\dot{k}_g'^{-1}\right)^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\dot{k}_g^{-1}\dot{g}(x_t,a_o)u_t^+-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}k_r^{-1}r(x_t)u_t^+\right]^2}{\hat{\sigma}_+^2\frac{1}{n}\sum_{t=1}^{n}\left[\left(k_r^{-1}\hat{A}_n'\dot{k}_g'^{-1}\right)\left(\dot{k}_g^{-1}\hat{B}_n\dot{k}_g'^{-1}\right)^{-1}\dot{k}_g^{-1}\dot{g}(x_t,\hat{a})-k_r^{-1}r(x_t)\right]^2}$$

$$= \frac{\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(A_o'B_o^{-1}\dot{k}_g'^{-1}\dot{g}(x_t,a_o)-k_r^{-1}r(x_t)\right)u_t^+\right]^2}{\sigma_+^2\frac{1}{n}\sum_{t=1}^{n}\left[A_o'B_o^{-1}\dot{k}_g'^{-1}\dot{g}(x_t,\hat{a})-k_r^{-1}r(x_t)\right]^2} + o_p(1)$$

$$= \frac{\left[\int_0^1\left(A_o'B_o^{-1}\dot{h}_g(V(r),a_o)-h_r(V(r))\right)dU^+(r)\right]^2}{\sigma_+^2\int_0^1\left(A_o'B_o^{-1}\dot{h}_g(V(r),a_o)-h_r(V(r))\right)^2 dr} + o_p(1)$$

Hence $CM_n \xrightarrow{d} \left(\frac{W^+(1)}{\sigma_+}\right)^2$,

where U^+ (& W^+) is a Brownian motion independent of V.

B. Under the alternative hypothesis:

First we will determine the behaviour of $\hat{\sigma}_+^2 = \hat{\sigma}^2 - \hat{\sigma}_{u\varepsilon}^2 \sigma_\varepsilon^{-2}$. The behaviour of $\hat{\sigma}^2$ is already given in the proof of 3.2. Now under misspecification σ_ε^{-2} is not affected so we will concentrate on $\hat{\sigma}_{u\varepsilon}$. The arguments below show that $\hat{\sigma}_+^2 \sim \hat{\sigma}^2$.

(i) f is I-regular:

$$\hat{\sigma}_{u\varepsilon} = \tfrac{1}{n}\sum_{t=1}^{n} \hat{u}_t \hat{\varepsilon}_{l,t+1}$$
$$= \tfrac{1}{n}\sum_{t=1}^{n} f(x_t)\hat{\varepsilon}_{l,t+1} + \tfrac{1}{n}\sum_{t=1}^{n} g(x_t,\hat{a})\hat{\varepsilon}_{l,t+1} + \tfrac{1}{n}\sum_{t=1}^{n} u_t\hat{\varepsilon}_{l,t+1}$$
$$= \tfrac{1}{n}\sum_{t=1}^{n} f(x_t)\varepsilon_{t+1} + \tfrac{1}{n}\sum_{t=1}^{n} g(x_t,\hat{a})\varepsilon_{t+1} + \tfrac{1}{n}\sum_{t=1}^{n} u_t\varepsilon_{t+1} + o_p(1)$$

(CPP, proof of Theorem 5.2)

$$= O_p(n^{-3/4}) + o_p(k_g(a^*)) + \sigma_{u\varepsilon}$$

(from P&P Lemma 7(c) and Proposition 4.10 of Chapter 2).

In view of the results in proof of 3.2(i) $\hat{\sigma}_+^2 \sim \begin{cases} \sigma^2 - \sigma_{u\varepsilon}^2\sigma_\varepsilon^{-2}, & \text{for } k_g(a^*) \to 0 \\ O_p(k_g^2(a^*)), & \text{for } k_g(a^*) \to \infty \end{cases}$

Now the statistic

$$CM_n = \frac{\left[\sum_{t=1}^{n}(y-g(x_t,\hat{a})-c_n)r(x_t)\right]^2}{\hat{\sigma}_+^2\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n g(x_t,\hat{a})+r(x_t)\right]^2} = \frac{\frac{1}{nk_r^2}\left[\sum_{t=1}^{n}(f(x_t)-g(x_t,\hat{a}))r(x_t)+\sum_{t=1}^{n}r(x_t)u_t+\hat{\sigma}_{u\varepsilon}\hat{\sigma}_\varepsilon^{-2}\sum_{t=1}^{n}r(x_t)\hat{\varepsilon}_{l,t+1}\right]^2}{\hat{\sigma}_+^2\frac{1}{n}\sum_{t=1}^{n}\left[\left(k_r^{-1}\hat{A}_n'\dot{k}_g'^{-1}\right)\left(\dot{k}_g^{-1}\hat{B}_n\dot{k}_g'^{-1}\right)^{-1}\dot{k}_g^{-1}\dot{g}(x_t,\hat{a})+k_r^{-1}r(x_t)\right]^2}$$
$$= \frac{\left[\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n}(f(x_t)-g(x_t,\hat{a}))r(x_t)+\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n}r(x_t)u_t+\hat{\sigma}_{u\varepsilon}\hat{\sigma}_\varepsilon^{-2}\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n}r(x_t)\varepsilon_{t+1}\right]^2}{\hat{\sigma}_+^2\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V(r),a^*)-h_r(V(r))\right)^2dr} + o_p(1)$$

For $k_g(a^*) \to 0$,

$$CM_n = \frac{\left[\sqrt{n}k_g\int_0^1 h_g(V(r),a^*)h_r(V(r))dr+o_p(\sqrt{n}k_g)+\int_0^1 h_r(V(r))dU^+(r)\right]^2}{\sigma_+^2\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V(r),a_o)+h_r(V(r))\right)^2dr} + o_p(1) = O_p(nk_g^2)$$

For $k_g(a^*) \to \infty$,

$$CM_n =$$
$$= \frac{\left[\sqrt{n}k_g^*G_{rg}+o_p(\sqrt{n}k_g^*)+\int_0^1 h_r(V(r))dU^+(r)+o_p(k_g^*)\sigma_\varepsilon^{-2}\int_0^1 h_r(V(r))d\left(\psi^{-1}(1)V(r)\right)\right]^2}{k_g^2G_{gg}\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V(r),a_o)-h_r(V(r))\right)^2dr+o_p(k_g^2)}$$
$$+o_p(1)$$

$$= \frac{\left[\sqrt{n}G_{rg}+o_p(\sqrt{n})\right]^2}{G_{gg}\int_0^1\left(A'_*B_*^{-1}\dot{h}_g(V(r),a_o)-h_r(V(r))\right)^2dr} = O_p(n)$$

(ii) f is H-regular with $k_f \neq k_g(a^*)$:

$$\hat{\sigma}_{u\varepsilon} =$$

$$= \tfrac{1}{n}\sum_{t=1}^{n} f(x_t)\varepsilon_{t+1} + \tfrac{1}{n}\sum_{t=1}^{n} g(x_t,\hat{a})\varepsilon_{t+1} + \tfrac{1}{n}\sum_{t=1}^{n} u_t\varepsilon_{t+1} + o_p(1)$$

$$= O_p(n^{-1/2}k_f) + o_p(k_g(a^*)) + \sigma_{u\varepsilon}$$

Hence $\hat{\sigma}_+^2 \sim O_p(\max\left\{k_g^2(a^*), k_f^2\right\})$.

Now the statistic

$$CM_n = \frac{\left[\sum_{t=1}^{n}(y-g(x_t\hat{a})-c_n)r(x_t)\right]^2}{\hat{\sigma}_+^2\sum_{t=1}^{n}\left[\hat{A}'_n\hat{B}_ng(x_t,\hat{a})-r(x_t)\right]^2} = \frac{\frac{1}{nk_r^2}\left[\sum_{t=1}^{n}(f(x_t)-g(x_t,\hat{a}))r(x_t)+\sum_{t=1}^{n}r(x_t)u_t+\hat{\sigma}_{u\varepsilon}\hat{\sigma}_\varepsilon^{-2}\sum_{t=1}^{n}r(x_t)\hat{\varepsilon}_{l,t+1}\right]^2}{\hat{\sigma}_+^2\frac{1}{n}\sum_{t=1}^{n}\left[\left(k_r^{-1}\hat{A}'_n\dot{k}_g'^{-1}\right)\left(\dot{k}_g^{-1}\hat{B}_n\dot{k}_g'^{-1}\right)^{-1}\dot{k}_g^{-1}\dot{g}(x_t,\hat{a})-k_r^{-1}r(x_t)\right]^2}$$

$$\frac{\left[\sqrt{n}k_fG_{rf}-\sqrt{n}k_g^*G_{rg}+o_p(\sqrt{n}\left(k_f+k_g^*\right))+R_n\right]^2}{\left\{\left(k_g^{*2}G_{gg}\right)+\left(k_f^2G_{ff}\right)\right\}\int_0^1\left(A'_*B_*^{-1}\dot{h}_g(V(r),a^*)-h_r(V(r))\right)^2dr+o_p(k_g^{*2}+k_f^2)} = O_p(n)$$

where

$$R_n = \int_0^1 h_r(V(r))dU^+(r)+\left\{O_p(n^{-1/2}k_f)+o_p(k_g(a^*))\right\}\sigma_\varepsilon^{-2}\int_0^1 h_r(V(r))d\left(\psi^{-1}(1)V(r)\right)$$

(iii) f is H-regular with $k_f = k_g(a^*)$ and $f - g(a^*) = q(a^*)$:

$$\hat{\sigma}_{u\varepsilon} =$$

$$\tfrac{1}{n}\sum_{t=1}^{n} q(x_t,a^*)\varepsilon_{t+1} + \tfrac{1}{n}(\hat{a}-a^*)'\sum_{t=1}^{n}\dot{g}(x_t,\bar{a})\varepsilon_{t+1} + \tfrac{1}{n}\sum_{t=1}^{n}u_t\varepsilon_{t+1} + o_p(1)$$

(from the mean value Theorem)

$$= O_p(n^{-1/2}k_q(a^*)) + o_p(k_q(a^*)) + \sigma_{u\varepsilon}$$

(from P&P Lemma 7(c) and Proposition 4.13 of Chapter 2).

Hence in view of proof 3.2(iv) we have $\hat{\sigma}_+^2 \sim O_p(k_q(a^*))$.

Now the test statistic

$$CM_n = \frac{\left[\sum_{t=1}^{n}(y-g(x_t,\hat{a})-c_n)r(x_t)\right]^2}{\hat{\sigma}_+^2\sum_{t=1}^{n}\left[\hat{A}'_n\hat{B}_ng(x_t,\hat{a})+r(x_t)\right]^2}$$

$$= \frac{\left[\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n} q(x_t,a^*)r(x_t)+\sqrt{n}k_q\left(\hat{a}^{+}-a^*\right)'\frac{\dot{k}_g^*}{k_q}\sum_{t=1}^{n}\frac{\dot{k}_g^{*-1}}{nk_r}\dot{g}(x_t,\bar{a})r(x_t)+\hat{\sigma}_{u\varepsilon}\hat{\sigma}_{\varepsilon}^{-2}\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n} r(x_t)\hat{\varepsilon}_{l,t+1}\right]^2}{\hat{\sigma}_{+}^2\frac{1}{n}\sum_{t=1}^{n}\left[\left(k_r^{-1}\hat{A}_n'\dot{k}_g'^{-1}\right)\left(\dot{k}_g^{-1}\hat{B}_n\dot{k}_g'^{-1}\right)^{-1}\dot{k}_g^{-1}\dot{g}(x_t,\hat{a})-k_r^{-1}r(x_t)\right]^2}+$$

$o_p(1)$ (MVT)

$$= \frac{\left[\sqrt{n}k_q^*\left(G_{rq}-G_{r\dot{g}}'G_{\dot{g}\dot{g}}^{-1}G_{\dot{g}q}\right)+R_n\right]^2}{\left\{k_q^{*2}\left(G_{qq}-2G_{qg}'G_{gg}^{-1}G_{gg}+\left[G_g'G_{gg}^{-1}G_{qg}\right]^2\right)+o_p(k_q^{*2})+O_p(n^{-1}k_q^{*2})\right\}\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V,a^*)-h_r(V)\right)^2 dr} =$$

$O_p(n)$

where

$$R_n = \int_0^1 \left(G_{r\dot{g}}'G_{\dot{g}\dot{g}}^{-1}h_{\dot{g}}(V,a^*)+h_q(V,a^*)\right)dU(r)^{+}$$

$$+\left\{O_p(n^{-1/2}k_q)+o_p(k_q)\right\}\sigma_{\varepsilon}^{-2}\left[\int_0^1\left(G_{r\dot{g}}'G_{\dot{g}\dot{g}}^{-1}\dot{h}_g(V(r),a^*)-h_r(V(r))\right)d\left(\psi^{-1}(1)V(r)\right)\right].\blacksquare$$

PROOF OF PROPOSITION 3.5:

A. For the null the result follows from Proposition 2.4 using the same arguments as in Proposition 3.4.

B. First we will consider the behaviour of $\breve{\sigma}_{u\varepsilon}$, $\breve{\sigma}_{+}^2$ under misspecification.

$$\breve{\sigma}_{u\varepsilon} = n^{-1/2}\left[\sum_{t=1}^{n} w(x_t)\hat{u}_t\hat{\varepsilon}_{l,t+1}\right]/w_n$$

$$= n^{-1/2}\left[\sum_{t=1}^{n}\left(f(x_t)-g(x_t,\hat{a})\right)w(x_t)\varepsilon_{t+1}+\sum_{t=1}^{n} w(x_t)u_t\varepsilon_{t+1}\right]/w_n + o_p(1)$$

$$= n^{-1/2}\left[\sum_{t=1}^{n} w(x_t)u_t\varepsilon_{t+1}\right]/w_n + o_p(1)$$

(P&P Lemma 7(a))

$$= \sigma_{u\varepsilon} + o_p(1)$$

Hence given the proof of Proposition 3.3 we have

$$\breve{\sigma}_{+}^2 = \frac{\int_{-\infty}^{\infty}(f(s)-g(s,a^*))^2 w(s)ds}{\int_{-\infty}^{\infty} w(s)ds} + \sigma_{+}^2 + o_p(1)$$

Now the statistic

$$CM_n = \frac{\left[\sum_{t=1}^{n}(y_t - g(x_t,\check{a}) - c_n)r(x_t)\right]^2}{\hat{\sigma}_+^2 \sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n g(x_t,\hat{a}) - r(x_t)\right]^2}$$

$$= \frac{\left[\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n}(f(x_t)-g(x_t,\check{a}))r(x_t) + \frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n} r(x_t)u_t + \check{\sigma}_{u\varepsilon}\hat{\sigma}_\varepsilon^{-2}\frac{1}{\sqrt{n}k_r}\sum_{t=1}^{n} r(x_t)\varepsilon_{t+1}\right]^2}{\check{\sigma}_+^2\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V(r),a^*) - h_r(V(r))\right)^2 dr} + o_p(1)$$

(i) for H-regular f with $k_f \neq k_g(a^*)$:

Using similar arguments as those in the proof of 3.4(ii) we have

$$CM_n = \frac{\left[\sqrt{n}k_f G_{rf} - \sqrt{n}k_g^* G_{rg} + o_p(\sqrt{n}(k_f + k_g^*)) + R_n\right]^2}{\left(\frac{\int_{-\infty}^{\infty}(f(s)-g(s,a^*))^2 w(s)ds}{\int_{-\infty}^{\infty} w(s)ds} + \sigma_+^2\right)\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V(r),a^*) - h_r(V(r))\right)^2 dr + o_p(1)}$$
$$= O_p(nk_f^2) + O_p(nk_g^{*2})$$

where

$$R_n = \int_0^1 h_r(V(r))dU^+(r) + \left\{O_p(n^{-1/2}k_f) + o_p(k_g(a^*))\right\}\sigma_\varepsilon^{-2}\int_0^1 h_r(V(r))d\left(\psi^{-1}(1)V(r)\right)$$

(ii) for H-regular f with $k_f = k_g(a^*)$ and $f - g(a^*) = q(a^*)$:

Using similar arguments as those in the proof of 3.4(iii) we have

$$CM_n = \frac{\left[\sqrt{n}k_q^*\left(G_{rq} - G_{\dot{g}}'G_{\dot{g}\dot{g}}^{-1}G_{r\dot{g}}\right) + o_p(\sqrt{n}k_q^*) + R_n\right]^2}{\left(\frac{\int_{-\infty}^{\infty}(f(s)-g(s,a^*))^2 w(s)ds}{\int_{-\infty}^{\infty} w(s)ds} + \sigma_+^2\right)\int_0^1\left(A_*'B_*^{-1}\dot{h}_g(V(r),a^*) - h_r(V(r))\right)^2 dr + o_p(1)}$$
$$= O_p(nk_q^{*2})$$

where

$$R_n = \int_0^1\left(G_{r\dot{g}}'G_{\dot{g}\dot{g}}^{-1}h_{\dot{g}}(V(r),a^*) + h_q(V(r),a^*)\right)dU(r)^+$$
$$+\left\{O_p(n^{-1/2}k_q) + o_p(k_q)\right\}\sigma_\varepsilon^{-2}\left[\int_0^1\left(G_{r\dot{g}}'G_{\dot{g}\dot{g}}^{-1}\dot{h}_g(V(r),a^*) + h_r(V(r))\right)d\left(\psi^{-1}(1)V(r)\right)\right]. \blacksquare$$

Chapter 4
Detection of Functional Form Misspecification in Cointegrating Relationships

4.1 Introduction

In this chapter the conditional moment test for functional form is extended to cointegrating relationships. The only work that is close to ours, that we are aware of, is a paper by Hong and Phillips (2004). Hong and Phillips (2004) adapt the RESET test to cointgrating relationships. Their framework allows for fitted models that are linear with scalar covariates. Our framework includes multiple regression models nonlinear in variables. The theoretical framework of this chapter is confined to the *H-regular* family of transformations. The exogeneity assumption about the regressors is dropped. Moreover we introduce dependence in the error of the model. We assume that the error of the model and the errors that drive the unit root processes are a vector linear process. To induce a standard limit distribution under the null, a semiparametric approach is followed similar to the one of Xiao and Phillips (2002). The fitted model is estimated by a Fully Modified-Least Squares (FM-LS) type of estimator and the sample moment of the test is corrected for endogeneity. We show that under the null hypothesis of correct functional form the test has a chi-squared distribution. Moreover, we examine the properties of the test under two kinds of misspecification. First we consider the case where cointegration exists (possible nonlinear) but the fitted model is of incorrect functional form. Secondly we examine the case where the dependent variable of our fitted model does not cointegrate with independent variables of our model, i.e. when there is no cointegration at all (neither linear nor nonlinear). We show that the test is consistent under both functional form misspecification and no cointegration. The same divergence rate is attained under both kinds of misspecification and it depends on the bandwidth term used in the estimation of long-run covariance matrices.

The divergence rate of our statistic under misspecification is the same as the one attained by the CUSUM, the KPSS tests (see Xiao and Phillips (2002)) and the modified RESET test of Hong and Phillips (2004). As mentioned in Chapter 3 the statistic for the testing problem under consideration is of the form:

$$CM_n = \frac{(SM_n)^2}{VN_n}$$

where SM_n is some sample moment and VN_n is a variance normalisation term. Under FFM or no cointegration the residuals of the fitted model will be dominated by some *H-regular* term. Denote by $\bar{k}$ the asymptotic order of that term. Then our statistic under misspecification:

$$CM_n \sim \frac{O_p(n\bar{k}^2)}{O_p(M\bar{k}^2)} = O_p(n/M) \xrightarrow{p} \infty,$$

provided the bandwidth parameter M is such that $n/M \to \infty$ as $n \to \infty$. If M is set equal to a constant, the divergence rate of our test is the same (n) with that attained under stationarity (see also Proposition 3.4 of Chapter 3). The variance normalisation term we are using in the particular test statistic is unbounded, under misspecification, and therefore the divergence rate of our statistic cannot be better than that achieved under stationarity.

The test we propose is a Bierens (1990) type of test for functional form. Other tests for functional form developed for stationary data, White's (1981) Hauseman-type test, for example, could also be extended to our framework. Xiao and Phillips (2002) point out that the CUSUM test for structural change can be used as an alternative way of testing for cointegration. It turns out that testing procedures originally developed for testing for functional form can be used as tests for cointegration as well. Traditional residual-based tests for cointegration examine whether the regression residuals contain a unit root. Our procedure can detect absence of cointegration in a different way. Under correct specification, the sum of regression residuals order of magnitude is $\sqrt{n}$. Under the alternative (no cointegration or incorrect functional form), the sum of residuals will be of higher order than $\sqrt{n}$ hence the test statistic will diverge. As Xiao and Phillips suggest, this is also the rational behind the CUSUM test. In view of this it will not be surprising if the CUSUM test can be used as a test for functional form as well.

The rest of this chapter is organised as follows. In Section 2 our theoretical framework is specified and some preliminary results are provided. In Section 3 our testing procedure is presented and its properties are developed. Section 4 concludes. As usual for a function $f : \mathbf{R} \rightarrow \mathbf{R}$, $\dot{f}$ will denote its first derivative with respect to its argument Finally for a vector or a matrix A say, A' will denote its transpose.

4.2 Theoretical Framework and Preliminary Results

We assume that the series $\{y_t\}_{t=1}^{n}$ is generated by

$$\begin{aligned} y_t &= \theta_{o1} f_1(x_{1t})+, ..., +\theta_{op} f_p(x_{pt}) + u_t \qquad (1) \\ &= f'(x_t)\theta_o + u_t \end{aligned}$$

or by

$$y_t = s(z_t), \tag{2}$$

where $f(.)$ and $s(.)$ belong to the H-regular family, the x_{it}'s and z_t are unit root processes, and u_t is an error term that will be specified in detail later. Our purpose is to examine the case of Functional Form Misspecification (FFM) and the case of no cointegration. For this purpose we will consider two possible data generating mechanisms for our dependent variable. The model in (1) will be the true specification when there is cointegration (possibly nonlinear) between y_t and the variables of interest x_{it}. The specification (2) will be the data generating mechanism when there is no cointegrating relationship between y_t and the variables of interest. When the latter is the case, it is usually assumed in the literature (e.g. Xiao and Phillips (2002)) that y_t is a unit root process, z_t say, that is unrelated to the regressors (x_{it}'s). Here we will assume that y_t is a possibly a nonlinear function of such a process. In this way we allow y_t to be of different order of magnitude than the z_t. Clearly when $s(.)$ is linear, y_t is a unit root process. The fitted model will be given by

$$\begin{aligned} \hat{y}_t &= \hat{a}_1 g_1(x_{1t})+, ..., +\hat{a}_p g_p(x_{pt}) + \hat{u}_t \\ &= g'(x_t)\hat{a} + \hat{u}_t \end{aligned} \tag{3}$$

For notational brevity, the vectors $f(x_t)$ and $g(x_t)$ in (1) and (3) may be written as f_t and g_t respectively.

DEFINITION 2.1:

(i) *We will say that the fitted model (3) is of correct functional form, when* $g_i(.) = f_i(.)$ *for all* $i = \{1, .., p\}$ *and (1) holds.*

(ii) *We will say that the fitted model (3) is of incorrect functional form, when the true model is given by (1) and* $g_i(.) \neq f_i(.)$ *for some* $i = \{1, .., p\}$.

(ii) *We will say that there is no cointegration, when the fitted model is given by (3) and the true model by (2).*

Moreover when some component g_i of our fitted model is of incorrect functional form, we will assume that one of the following holds:

(i) $g_i \succ f_i$,

(ii) $g_i \prec f_i$ or

(iii) $g_i \approx f_i(.)$,with $g_i - f_i = q_i$ and $q_i \prec g_i, f_i$.

(The notation: $\succ$, $\prec$ and $\approx$ and is defined in Chapter 2)

We will rule out the possibility of having a second cointegrating raltionship between $g_1(x_{1t})$,. .., $g_p(x_{pt})$. It is obvious from Definition 2.1 that our framework does not allow for omitted or redundant variables. An extension of our results in that direction is possible but will not be attempted here, as it would result in more complexity in our presentation.

Next we will specify in detail the variables and the functions that appear in (1), (2) and (3). The variables $x_t' = (x_{1t}, ..., x_{pt})$ and z_t are a unit root processes given by:

$$x_t = x_{t-1} + v_t \textit{ and } z_t = z_{t-1} + w_t$$

We will assume that $e_t' = (u_t, v_t', w_t)$ are linear processes given by:

$$\begin{aligned} u_t &= \sum_{j=1}^{\infty} \Phi_j \varepsilon_{t-j} = \Phi(L)\varepsilon_t, \\ v_t &= \sum_{j=1}^{\infty} \Psi_j \eta_{t-j} = \Psi(L)\eta_t, \end{aligned}$$

and

$$w_t = \sum_{j=1}^{\infty} \Gamma_j \omega_{t-j} = \Gamma(L)\omega_t,$$

with $\Phi(1), \Psi(1), \Gamma(1) \neq 0$ and $\sum_{j=1}^{\infty} j^{\alpha} |\Phi_j|$, $\sum_{j=1}^{\infty} j^{\alpha} \|\Psi_j\|$, $\sum_{j=1}^{\infty} j^{\alpha} |\Gamma_j| < \infty$ with $\alpha > 1$. Moreover v_t, u_t and w_t satisfy the following assumption:

ASSUMPTION 2.1:

(i) $\{\xi_t' = (\varepsilon_t, \eta_{t+1}', \omega_{t+1}), \mathcal{F}_t = \sigma(\xi_s, -\infty \leq s \leq t)\}$ *is a stationary and ergodic martingale difference sequence with* $\mathbf{E}[\xi_t \xi_t' \mid \mathcal{F}_{t-1}] = \Sigma$.

(ii) *The sequence* ξ_t *is i.i.d. with* $\mathbf{E}\|\xi_t\|^r < \infty$ *for some* $r > 4$ *and its distribution is absolutely continuous with respect to Lebesgue measure and has characteristic function* $\varphi(\lambda) = o(\|\lambda\|^{-\delta})$ *as* $\lambda \to \infty$.

For the purpose of our analysis we will need to comformably partition the covariance matrix Σ as follows:

$$\Sigma = \begin{pmatrix} \Sigma_{\varepsilon\varepsilon} & \Sigma_{\varepsilon\eta} & \Sigma_{\varepsilon\omega} \\ \Sigma_{\eta\varepsilon} & \Sigma_{\eta\eta} & \Sigma_{\eta\omega} \\ \Sigma_{\omega\varepsilon} & \Sigma_{\omega\eta} & \Sigma_{\omega\omega} \end{pmatrix}$$

For v_t, u_t and w_t define the usual partial sum processes

$$U_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t,$$

$$V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} v_t,$$

and

$$W_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t,$$

with $0 \leq r \leq 1$. Under assumption 2.1 it follows from a multivariate extension of the results of Phillips and Solo (1992) that

$$(U_n(r), V_n'(r), W_n(r)) \xrightarrow{d} (U(r), V'(r), W(r))$$

with $(U(r), V'(r), W(r))$ being an $(p+2)$-dimensional Brownian motion with covariance matrix Ω conformably partitioned as

$$\Omega = \begin{pmatrix} \Omega_{uu} & \Omega_{uv} & \Omega_{uw} \\ \Omega_{vu} & \Omega_{vv} & \Omega_{vw} \\ \Omega_{wu} & \Omega_{wv} & \Omega_{ww} \end{pmatrix}$$

Under our assumptions, strong approximations hold for the vector $(U_n(r), V_n'(r), W_n(r))$ that allow the use of embedding arguments. Such embedding arguments are extensively utilised by P&P and will be used here as well. So when we use convergence in probability arguments those should be interpreted as convergence in distribution unless the limit is nonstochastic. Moreover for our purposes we need to introduce the usual one-sided long-run covariance matrices. Note that Ω can be expressed as:

$$\Omega = \sum_{k=-\infty}^{\infty} \mathbf{E}\left(e_t e_{t+k}'\right)$$

and the one sided long-run covariance matrix, say Λ is:

$$\Lambda = \sum_{k=0}^{\infty} \mathbf{E}\left(e_t e_{t+k}'\right) = \begin{pmatrix} \Lambda_{uu} & \Lambda_{uv} & \Lambda_{uw} \\ \Lambda_{vu} & \Lambda_{vv} & \Lambda_{vw} \\ \Lambda_{wu} & \Lambda_{wv} & \Lambda_{ww} \end{pmatrix}$$

Next we specify the functions that appear in (1), (2) and (3). As mentioned earlier we will work within the H-regular family but due to the introduction of weak dependence in the error structure of the model, we will need to impose some smoothness on the first derivatives of our functions. We will call our functions H_1-regular with H_1-regularity defined as follows:

DEFINITION 2.2:

The transformation $f : \mathbf{R}^p \to \mathbf{R}^p$, *such that* $f'(x) = (f_1(x_1), .., f_p(x_p))$ *will be called* H_1*-regular if:*

(i) $f(\lambda x) = k_f(\lambda)h_f(x) + R_f(x,\lambda)$ *with* $h_f(.)$ *regular and*

(a) $|R_f(x,\lambda)| \leq a_f(\lambda)P_f(x)$, *with* $\limsup_{\lambda\to\infty} \left\|a_f(\lambda)k_f^{-1}(\lambda)\right\| = 0$ *and* $P_f(.)$ *locally integrable, or*

(b) $|R_f(x,\lambda)| \leq b_f(\lambda)Q_f(\lambda x)$, *with* $\limsup_{\lambda\to\infty} \left\|b_f(\lambda)k_f^{-1}(\lambda)\right\| < \infty$ *and* $Q_f(.)$ *locally integrable and vanishing at infinity.*

(ii) $\lambda\dot{f}(\lambda x) = k_f(\lambda)\dot{h}_f(x) + \dot{R}_f(x,\lambda)$ *with* $\dot{h}_f(.)$ *regular and*

(a) $\left|\dot{R}_f(x,\lambda)\right| \leq \dot{a}_f(\lambda)\dot{P}_f(x)$, *with* $\limsup_{\lambda\to\infty} \left\|\lambda\dot{a}_f(\lambda)k_f^{-1}(\lambda)\right\| = 0$ *and* $\dot{P}_f(.)$ *locally integrable, or*

(b) $\left|\dot{R}_f(x,\lambda)\right| \leq \dot{b}_f(\lambda)\dot{Q}_f(\lambda x)$, *with* $\limsup_{\lambda\to\infty} \left\|\lambda\dot{b}_f(\lambda)k_f^{-1}(\lambda)\right\| < \infty$ *and* $\dot{Q}_f(.)$ *locally integrable and vanishing at infinity.*

(iii) *For any constant* K *and two sequences* s_n *and* m_n *such that* $s_n \downarrow 0$ *and* $m_n \to \infty$, *as* $n \to \infty$,

$$\limsup_{n\to\infty} \left\|m_n\sqrt{n}k_f(\sqrt{n})^{-1}\right\| \sup_{\|x_1\|\leq K} \sup_{\|x_1-x_2\|\leq s_n} \left\|\dot{f}(\sqrt{n}x_1) - \dot{f}(\sqrt{n}x_2)\right\| = 0.$$

As usual h_f and k_f will be called the limit homogenous functions and asymptotic order of f respectively. Moreover note that when f is a p-dimensional vector, k_f and $\dot{f}$ will be $(p \times p)$ diagonal metrices. Condition (iii) in the definition above is the same smoothness condition employed by de Jong (2002b)

For linear models it is well known (e.g. Phillips (1986, 1988)) that when the errors of the model are weakly dependent, the covariance asymptotics involve extra terms. In the limit, apart from the stochastic integral, a long-run covariance matrix appears. For nonlinear models the long-run covariance matrix is weighted by functionals of Brownian motion. This result was originally shown by de Jong (2002b), when the errors satisfy mixing conditions. Below we provide the same result for errors that are linear processes.

PROPOSITION 2.1:

Let $f'(x_t) = (f_1(x_{1t}), ..., f(x_{pt}))$ be H_1-regular. Under Assumption 2.1

$$\text{(i)}\ \frac{1}{\sqrt{n}} k_f^{-1}(\sqrt{n}) \sum_{t=1}^{n} f(x_t) u_t \xrightarrow{d} \int_0^1 h_f(V(r)) dU(r) + \int_0^1 \dot{h}_f(V(r)) dr \Lambda_{vu}$$

and

$$\text{(ii)}\ \frac{1}{\sqrt{n}} k_f^{-1}(\sqrt{n}) \sum_{t=1}^{n} f(x_t) v_t' \xrightarrow{d} \int_0^1 h_f(V(r)) dV'(r) + \int_0^1 \dot{h}_f(V(r)) dr \Lambda_{vv}$$

as $n \to \infty$.

Notice that the linear model that is commonly used in cointegrating relationships is H_1-regular. Below we provide some examples for f and g which are covered by our theoretical framework.

EXAMPLE 4.1:

(i) Let the true model be $y_t = \theta_{o1} x_{1t} + \theta_{o2} x_{2t} \left| x_{2t} \right|^{1/2} + u_t$ and the fitted $\hat{y}_t = \hat{a}_1 x_{1t} + \hat{a}_2 x_{2t} + \hat{u}_t$.

(ii) Let the true model be $y_t = \theta_{o1} x_{1t} + \theta_{o2} x_{2t}^2 + u_t$ and the fitted $\hat{y}_t = \hat{a}_1 x_{1t} \left| x_{2t} \right|^{1/4} + \hat{a}_2 x_{2t} + \hat{u}_t$.

(iii) Let the true model be $y_t = \theta_o x_t \exp(x_t)(1+\exp(x_t))^{-1} + u_t$ and the fitted $\hat{y}_t = \hat{a}x_t + \hat{u}_t$.

In practice functional form misspecification could arise from neglected lower order components (lower order than the linear specification). Consider for instance the case where $f(x) = x + |x|^{1/2}$ and $g(x) = x$. Under this kind of misspecification we have shown in Chapter 3, that the conditional moment test will be consistent when the regression errors are martingale differences. Under the current framework however we are unable to obtain expilicit power rate results as the component $|x|^{1/2}$ is not H_1*-regular*.

Saikkonen and Choi (2004) have recently analysed cointegrating Smooth Transition Regression (STR) models. Their specification is comprised by a linear component multiplied by a transition function. They explicitly consider a logistic function. Because the logistic function lacks identification when the covariates are unit root processes (see P&P), Saikkonen and Choi (2004) actually consider a model where the covariates are normalised by the square root of the sample size. In practice one might want to have a constistent test for functional form, when the transition function used is misspecified. Our framework does not cover fitted models of this kind because they are nonlinear in parameters. Moreover the fact that their model involves normalised variables creates an extra complication. We have assumed that if the fiitted model g, is of incorrect functional form and of the same order as the true model f, then $g(a^*)$ and f agree up to some lower order component $q(a^*)$, say. Actually to the best of our knowledge, this has to be the case (when both f and g are *H-regular*). Now if the model involves normalised variables one can find examples of f, and g for which f and $g(a^*)$ do not agree at all, for example let $f(x) = 1\{x/\sqrt{n} > c\}$ and

$g(x,a) = \exp(ax/\sqrt{n})\,(1+\exp(ax/\sqrt{n}))^{-1}$. An extension of our theoretical results to this kind of models could be possible given our results in Chapter 2, however a development of a limit distribution result for locally integrable functions of zero energy would be required.

4.3 A Conditional Moment Test for Functional Form

The asymptotic behaviour of the LS estimator and the usual likelihood based tests and t-tests is determined by sample covariances like those in Proposition 2.1. Because the limit distribution theory is not mixed normal, the aforementioned statistical tests do not have standard distributions under the null. In our case to induce a chi-squared distribution for our test statistic, the model is fitted by a FM-LS type of estimator and an endogeneity correction term is introduced in the statistic. To obtain our estimator and the correction term, kernel estimators for Ω_{uu}, Ω_{vv}, Ω_{vu}, Λ_{vu} and Λ_{vv} are used:

$$\hat{\Omega}_{uu} = \sum_{h=-M}^{M} \kappa\left(\tfrac{h}{M}\right) C_{uu}(h), \quad \hat{\Omega}_{vv} = \sum_{h=-M}^{M} \kappa\left(\tfrac{h}{M}\right) C_{vv}(h),$$
$$\hat{\Omega}_{vu} = \sum_{h=-M}^{M} \kappa\left(\tfrac{h}{M}\right) C_{vu}(h), \quad \hat{\Lambda}_{vv} = \sum_{h=0}^{M} \kappa\left(\tfrac{h}{M}\right) C_{vv}(h),$$
$$\hat{\Lambda}_{vu} = \sum_{h=0}^{M} \kappa\left(\tfrac{h}{M}\right) C_{vu}(h),$$

where $\kappa(.)$ is the lag window defined on $[-1,1]$ such that $\kappa(0) = 1$ and M is a bandwidth such that $M \to \infty$, $n/M \to 0$ as $n \to \infty$. Moreover $C_{uu}(h)$, $C_{vv}(h)$, and $C_{vu}(h)$ are sample covariances defined by $C_{uu}(h) = n^{-1}\sum_t' \hat{u}_t\hat{u}_{t+h}$, $C_{vv}(h) = n^{-1}\sum_t' v_t v_{t+h}'$ and $C_{vu}(h) = n^{-1}\sum_t' v_t u_{t+h}$ where $\sum_t'$ is summation over $1 \leq t,\, t+h \leq n$. Consistency results for this kind of kernel estimators are provided by Andrews (1991), when the processes satisfy mixing conditions. Under the current framework consistency results are provided by Phillips (1995).

Our estimator closely resembles the original FM-LS estimator introduced by Phillips and Hansen (1990). Before we present the estimator we need to define the following quantities:

$$y_t^+ = y_t - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu} \text{ and } \hat{\Lambda}_{vu}^+ = \hat{\Lambda}_{vu} - \hat{\Lambda}_{vv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}$$

The FM-LS estimator under consideration is:

$$\hat{a} = \left[\sum_{t=1}^{n} g(x_t)g'(x_t)\right]^{-1}\left[\sum_{t=1}^{n} g(x_t)y_t^+ - \dot{g}_n\hat{\Lambda}_{vu}^+\right]$$

with $\dot{g}_n = \sum_{t=1}^{n}\dot{g}(x_t)$. The following result holds:

PROPOSITION 3.1:

Under correct functional form

$$\sqrt{n}k_g\left(\hat{a} - \theta_o\right) \xrightarrow{d} \left[\int_0^1 h_g\left(V(r)\right)h_g'\left(V(r)\right)dr\right]^{-1}\int_0^1 h_g\left(V(r)\right)dU^+(r),$$

as $n \to \infty$, *where* $U^+(r) = U(r) - V'(r)\Omega_{vv}^{-1}\Omega_{vu}$.

Notice that the limit distribution of the estimator is mixed normal as V and U^+ are independent.

Now we can present the test statistic. First define the matrices $\hat{A}_n$, $\hat{B}_n$, A and B as follows:

$$\frac{1}{n}k_g^{-1}\hat{A}_n = \frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g(x_t) \xrightarrow{p} A,$$

and

$$\frac{1}{n}k_g^{-1}\hat{B}_nk_g^{-1} = \frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g(x_t)g'(x_t)k_g^{-1} \xrightarrow{p} B > \mathbf{0}.$$

The test statistic is:

$$CM_n = \frac{\left[\sum_{t=1}^{n}\left(y_t^{+} - g(x_t)'\hat{a} - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\right]^2}{\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g(x_t) - 1\right]^2},$$

PROPOSITION 3.2:

Under the null hypothesis

$$CM_n \xrightarrow{d} \chi_1^2,$$

as $n \to \infty$.

We will now examine the asymptotic power of the test. Note that under the alternative some of the kernel estimators mentioned earlier will be inconsistent. Before we consider their limit behaviour we need to introduce some notation. Define $d_t = f_t - g_t$ with f_t, g_t as in (1) and (3). Moreover let k_{g^*} and k_{d^*} be the asymptotic orders of the leading elements of g_t and d_t respectively. For the purpose of the subsequent analysis we will distinguish two cases, when the fitted model is of incorrect functional form. First $k_{d^*} \neq k_{g^*}$ and secondly $k_{d^*} = k_{g^*}$. In the second case the leading term of the fitted model is misspecified and dominates all the components of the true model. The first case takes into account all the other scenarios. Denote by $\hat{a}_{LS}$ the least squares estimator corresponding to the fitted model. Moreover we will introduce some notation.

DEFINITION 3.1:

Define:

(i) *The p-dimensional vectors* $h_{f^*}(.)$, $h_{g^*}(.)$, $h_{d^*}(.)$ *and the* $(p \times p)$*-dimensional matrices* $\dot{h}_{f^*}(.)$, $\dot{h}_{g^*}(.)$, $\dot{h}_{d^*}(.)$ *by*

$$\begin{array}{ll} (nk_{f^*})^{-1}\sum_{t=1}^{n} f_t \xrightarrow{p} \int_0^1 h_{f^*}(V(r))dr, & (\sqrt{n}k_{f^*})^{-1}\sum_{t=1}^{n} \dot{f}_t \xrightarrow{p} \int_0^1 \dot{h}_{f^*}(V(r))dr, \\ (nk_{g^*})^{-1}\sum_{t=1}^{n} g_t \xrightarrow{p} \int_0^1 h_{g^*}(V(r))dr, & (\sqrt{n}k_{f^*})^{-1}\sum_{t=1}^{n} \dot{g}_t \xrightarrow{p} \int_0^1 \dot{h}_{g^*}(V(r))dr, \\ (nk_{d^*})^{-1}\sum_{t=1}^{n} d_t \xrightarrow{p} \int_0^1 h_{d^*}(V(r))dr, & (\sqrt{n}k_{f^*})^{-1}\sum_{t=1}^{n} \dot{d}_t \xrightarrow{p} \int_0^1 \dot{h}_{d^*}(V(r))dr, \end{array}$$

(ii) *The vectors* ζ_1, ζ_2 *and* ζ_3 *are the following limits*

$$\begin{array}{l} \frac{k_g}{k_{d^*}}\left(\hat{a}_{LS} - \theta_o\right) \xrightarrow{p} \zeta_1, \text{ under FFM when } k_{d^*} \neq k_{g^*}, \\ \frac{k_g}{k_{f^*}}\hat{a}_{LS} \xrightarrow{p} \zeta_2, \text{ under FFM when } k_{d^*} = k_{g^*}, \\ \frac{k_g}{k_s}\hat{a}_{LS} \xrightarrow{p} \zeta_3, \text{ under no cointegration.} \end{array}$$

(iii) *The vectors* $\bar{\zeta}_1$, $\bar{\zeta}_2$, $\bar{\zeta}_3$, $\bar{h}_1$, $\bar{h}_2$, $\bar{h}_3$ *and the matrices* $\dot{H}_1$, $\dot{H}_2$, $\dot{H}_3$, $\bar{\Omega}$, $\bar{\Lambda}$

$$\begin{array}{l} \bar{\zeta}_1' = (\theta_o', -\zeta_1'), \ \bar{\zeta}_2' = (\theta_o', -\zeta_2'), \ \bar{\zeta}_3' = (1, -\zeta_3'), \\ \bar{h}_1' = \left(h_{d^*}', h_g'\right), \ \bar{h}_2' = \left(h_{f^*}', h_g'\right), \ \bar{h}_3' = \left(h_s, h_g'\right), \\ \dot{H}_1' = \left(\dot{h}_{d^*}', \dot{h}_g'\right), \ \dot{H}_2' = \left(\dot{h}_f', \dot{h}_g'\right), \ \dot{H}_3' = \left(\dot{h}_s, \dot{h}_g'\right), \\ \bar{\Omega} = \left(\Omega_{vw}, \Omega_{vv}\right), \ \bar{\Lambda} = \left(\Lambda_{vw}, \Lambda_{vv}\right). \end{array}$$

Let $K(s) = \lim_{n\to\infty}(2\pi M)^{-1}\sum_{h=-M}^{M}\kappa(h/M)e^{ihs}$ and $K_1(s)$ is its one-sided version. The limit behaviour of the kernel estimators under the alternative as $n \to \infty$, is given in the following proposition.

PROPOSITION 3.3:

Let Assumption 2.1 hold. Under incorrect functional form as $n \to \infty$ *we have*

$$\begin{array}{l} \frac{n^{1/2}}{Mk_{d^*}}\hat{\Omega}_{vu} \xrightarrow{p} 2\pi K(0)\int_0^1 dV(r)\bar{h}_1'(V(r))\bar{\zeta}_1 + \Omega_{vv}\int_0^1 \dot{H}_1'(V(r))\bar{\zeta}_1 dr, \\ \frac{n^{1/2}}{Mk_{d^*}}\hat{\Lambda}_{vu} \xrightarrow{p} 2\pi K_1(0)\int_0^1 dV(r)\bar{h}_1'(V(r))\bar{\zeta}_1 + \Lambda_{vv}\int_0^1 \dot{H}_1'(V(r))\bar{\zeta}_2 dr, \\ \frac{1}{Mk_{d^*}^2}\hat{\Omega}_{uu} \xrightarrow{p} 2\pi K(0)\int_0^1 \bar{\zeta}_1'\bar{h}_1(V(r))\bar{h}_1'(V(r))\bar{\zeta}_1 dr, \end{array}$$

when $k_{d^*} \neq k_{g^*}$ *and*

$$\begin{array}{l} \frac{n^{1/2}}{Mk_{f^*}}\hat{\Omega}_{vu} \xrightarrow{p} 2\pi K(0)\int_0^1 dV(r)\bar{h}_2'(V(r))\bar{\zeta}_2 + \Omega_{vv}\int_0^1 \dot{H}_2'(V(r))\bar{\zeta}_2 dr, \\ \frac{n^{1/2}}{Mk_{f^*}}\hat{\Lambda}_{vu} \xrightarrow{p} 2\pi K_1(0)\int_0^1 dV(r)\bar{h}_2'(V(r))\bar{\zeta}_2 + \Lambda_{vv}\int_0^1 \dot{H}_2'(V(r))\bar{\zeta}_2 dr, \\ \frac{1}{Mk_{f^*}^2}\hat{\Omega}_{uu} \xrightarrow{p} 2\pi K(0)\int_0^1 \bar{\zeta}_2'\bar{h}_2(V(r))\bar{h}_2'(V(r))\bar{\zeta}_2 dr, \end{array}$$

when $k_{d^*} \neq k_{g^*}$. Under no cointegration

$$\frac{n^{1/2}}{Mk_s}\hat{\Omega}_{vu} \xrightarrow{p} 2\pi K(0)\int_0^1 dV(r)\bar{h}_3'(W(r),V(r))\bar{\zeta}_3 + \bar{\Omega}\int_0^1 \dot{H}_3'\left(W(r),V(r)\right)\bar{\zeta}_3 dr$$
$$\frac{n^{1/2}}{Mk_s}\hat{\Lambda}_{vu} \xrightarrow{p} 2\pi K_1(0)\int_0^1 dV(r)\bar{h}_3'(W(r),V(r))\bar{\zeta}_3 + \bar{\Lambda}\int_0^1 \dot{H}_3'\left(W(r),V(r)\right)\bar{\zeta}_3 dr$$
$$\frac{1}{Mk_s^2}\hat{\Omega}_{uu} \xrightarrow{p} 2\pi K(0)\int_0^1 \bar{\zeta}_3'\dot{H}_3\left(W(r),V(r)\right)\dot{H}_3'\left(W(r),V(r)\right)\bar{\zeta}_3 dr$$

as $n \longrightarrow \infty$.

The behaviour of the statistic under the alternative is given by the following result:

PROPOSITION 3.4:

Under FFM or no cointegration we have

$$CM_n = O_p\left(n/M\right).$$

This result is consistent with the one in Proposition 3.4 of Chapter 3, where the same test was considered but with exogenous regressors. It was reported there that when both f and g are *H-regular,* the asymptotic power rate of the test under FFM is n, the same attained under stationarity. When the conditions of Proposition 3.4 hold, the power rate is smaller than n. Clearly the reduction in the divergence rate of the test in this case is due to the semiparametric approach followed. Moreover in Chapter 3 it was shown that the test will not be consistent, when the fitted model is *H-regular* and the true one *I-regular* unless the estimator is bounded away from zero. It turns out that is the case under the current framework as well.

From the simulation study of Xiao and Phillips (2002) it is obvious that when it comes to the choice of the bandwidth parameter there is a trade-off between size and power. Their simulation results seem to suggest that a choice of $M \sim n^{1/3}$ is a reasonable one. Andrews (1991) proposes an automatic bandwidth method where $M = 1.447\left(\hat{\delta}n\right)^{1/3}$, with $\hat{\delta} = 4\hat{\rho}/\left(1-\hat{\rho}^2\right)^2$ and $\hat{\rho}$ is the estimator from the residuals autoregression. Methods like this one are inappropriate in our case. As Xiao and Phillips (2002) point out these kind of procedures were developed for stationary processes. In our case the residuals are stationary only under the null. Under the alternative they are not stationary. When this type of bandwidth method is used the, CUSUM test has no power. Xiao and Phillips (2002) suggest that under the alternative of their test $M \sim n$. As the following proposition shows, this is true in our case as well.

PROPOSITION 3.5:

Let Assumption 2.1 hold and the derivatives of f, g and s be H_1-regular. Then under FFM or no cointegration

$$\left(\hat{\delta} n\right)^{1/3} = O_p(n).$$

4.4 Conclusion

A conditional moment test for functional form was developed for cointegrating relationships. The standard linear models used in practice are a special case in our theoretical framework. The test has a chi-squared limit distribution under correct functional form and is consistent under FFM or in the case where there is no cointegration. Following Xiao and Phillips (2002), a semiparametric approach was used to induce a standard limit distribution under the null. Under the alternative, the divergence rate is adversely affected by the magnitude of the bandwidth parameter. The divergence rate is the same as the one of the CUSUM test for cointegration and in the best case is as good as the rate attained under stationarity. In Chapter 3 we have seen that for models with single and exogenous covariate, integrable weighting functions can improve the power rate of the test. We expect that this result can be easily extended to the current framework as long as the model has a single regressor.

4.5 Appendix to Chapter 4

LEMMA A:

Let Assumption 2.1 hold and f H_1-regular. Then for $0 \leq h \leq M$

$$\frac{1}{n}k_f^{-1}\sum_{t=1}^{n} f(x_t)f'(x_{t+h})k_f^{-1} \xrightarrow{p} \int_0^1 h_f(V(r))h_f'(V(r))dr,\ 1 \leq t,\ t+h \leq n,$$

as $n \longrightarrow \infty$.

PROOF OF LEMMA A:

Let $\bar{n} = n - h$ and $V_{\bar{n}}(r) = \frac{1}{\sqrt{\bar{n}}}\sum_{t=1}^{[r\bar{n}]} v_t$. Now

$$\begin{aligned}
& \frac{1}{\bar{n}}k_f^{-1}\sum_{t=1}^{\bar{n}} f(x_t)f'(x_{t+h})k_f^{-1} \\
\sim\ & \frac{1}{\bar{n}}\sum_{t=1}^{\bar{n}} h_f\left(\frac{x_t}{\sqrt{n}}\right) h_f'\left(\frac{x_{t+h}}{\sqrt{n}}\right) \\
& \text{(from the definition of } H_1\textit{-regularity}) \\
=\ & \int_0^1 h_f(V_{\bar{n}}(r))h_f'\left(V_{\bar{n}}\left(r+\frac{h}{\bar{n}}\right)\right) dr \\
& \xrightarrow{p} \int_0^1 h_f(V_{\bar{n}}(r))h_f'\left(V_{\bar{n}}\left(r+\frac{h}{\bar{n}}\right)\right) dr,
\end{aligned}$$

as $n \longrightarrow \infty$, from an application of the continuous mapping theorem since

$$V_{\bar{n}}\left(r+\frac{h}{\bar{n}}\right) - V_{\bar{n}}(r) = O_p\left(\frac{h^{1/2}}{\bar{n}^{1/2}}\right) = o_p(1).$$

■

PROOF OF PROPOSITION 2.1:

We start with the proof of (i). From the Beveridge-Nelson (BN) decomposition we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)u_t\\
&= \frac{1}{\sqrt{n}}k_f(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Phi(1)\varepsilon_t - \frac{1}{\sqrt{n}}k_f(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Delta\tilde{\varepsilon}_t,
\end{aligned}$$

where $\tilde{\varepsilon}_t = \sum_{j=1}^{\infty}\left(\sum_{i=j+1}^{\infty}\Phi_i\right)\varepsilon_{t-j}$. Now $\frac{1}{\sqrt{n}}k_f(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Phi(1)\varepsilon_t \xrightarrow{p} \int_0^1 h_f(V(r))dU(r)$ from Theorem 3.3 in P&P. The term

$$\begin{aligned}
\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Delta\tilde{\varepsilon}_t &= \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})f(x_n)\tilde{\varepsilon}_n - \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta f(x_t)\tilde{\varepsilon}_{t-1}\\
&= O_p(n^{-1/2}) - \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1}+\gamma_t v_t)v_t\tilde{\varepsilon}_{t-1},
\end{aligned}$$

where $\gamma_t = diag\left(\gamma_{1t},..,\gamma_{pt}\right)$ with $\gamma_{it}\in[-1,1]$.

Set $\bar{x}_{t-1} = x_{t-1}+\gamma_t v_t$, and $\bar{V}_n(r) = \frac{1}{\sqrt{n}}\sum_{t=1}^{[rn]}v_{t-1}$. First we will show that

$$\frac{1}{\sqrt{n}}\left\|k_f^{-1}(\sqrt{n})\left(\sum_{t=1}^{n}\dot{f}(\bar{x}_{t-1})v_t\tilde{\varepsilon}_{t-1} - \sum_{t=1}^{n}\dot{f}(x_{t-1})v_t\tilde{\varepsilon}_{t-1}\right)\right\| \xrightarrow{p} 0.$$

Note that $\sup_{r\in[0,1]}\left\|\bar{V}_n(r)-V(r)\right\| \le \sup_{r\in[0,1]}\left\|V_n(r)-V(r)\right\| + \sup_{r\in[0,1]}\left\|\frac{v_{[nr]}}{\sqrt{n}}\right\|$ with the first term $o_{a.s.}(1)$ from Lemma 2.3 in Park and Phillips (1999). Next, the second term is $o_{a.s.}(1)$ as well, because for any $\delta>0$

$$\begin{aligned}
&\sum_{n=1}^{\infty}P\left(\max_{1\le t\le n}\|v_t\|\ge\delta\sqrt{n}\right)\\
&\le \sum_{n=1}^{\infty}\sum_{t=1}^{n}P\left(\|v_t\|\ge\delta\sqrt{n}\right)\le\sum_{n=1}^{\infty}\mathbf{E}\left(\frac{\|v_t\|^r}{\delta n^{r/2-1}}\right)<\infty,\ \text{for } r>4,
\end{aligned}$$

where the last inequality can be easily checked given Assumption 2.1. Hence $\sup_{r\in[0,1]}\left\|\bar{V}_n(r)-V(r\right.$ $= o_{a.s.}(1)$. Similarly we can show that $\sup_{r\in[0,1]}\left\|\bar{x}_{[nr]}-V(r)\right\| = o_{a.s.}(1)$. Also note that

$\sup_{r\in[0,1]} \|V(r)\| \leq K$ $a.s.$ for some $K > 0$. Therefore for some sequences s_{1n} and s_{2n} that decrease to zero we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\left\|k_f^{-1}(\sqrt{n})\left(\sum_{t=1}^{n}\dot{f}(\bar{x}_{t-1})v_t\tilde{\varepsilon}_{t-1}-\sum_{t=1}^{n}\dot{f}(x_{t-1})v_t\tilde{\varepsilon}_{t-1}\right)\right\| \\
\leq{}& \left\|\sqrt{n}k_f^{-1}\right\|\left(\sup_{r\in[0,1]}\left\|\dot{f}(\sqrt{n}\bar{x}_{[nr]})-\dot{f}(\sqrt{n}V(r))\right\|+\sup_{r\in[0,1]}\left\|\dot{f}(\sqrt{n}\bar{V}_n(r))-\dot{f}(\sqrt{n}V(r))\right\|\right) \\
&\times\frac{1}{n}\sum_{t=1}^{n}\|v_t\tilde{\varepsilon}_{t-1}\| \\
\leq{}& m_n\sqrt{n}\left\|k_f^{-1}\right\|\sup_{\|x_1\|\leq K}\sup_{\|x_1-x_2\|\leq s_{1n}}\left\|\dot{f}(\sqrt{n}x_1)-\dot{f}(\sqrt{n}x_2)\right\|\times\frac{1}{nm_n}\sum_{t=1}^{n}\|v_t\tilde{\varepsilon}_{t-1}\| \\
&+m_n\sqrt{n}\left\|k_f^{-1}\right\|\sup_{\|x_1\|\leq K}\sup_{\|x_1-x_2\|\leq s_{2n}}\left\|\dot{f}(\sqrt{n}x_1)-\dot{f}(\sqrt{n}x_2)\right\|\times\frac{1}{nm_n}\sum_{t=1}^{n}\|v_t\tilde{\varepsilon}_{t-1}\| \\
&\xrightarrow{p} 0 \text{ as } n\to\infty,
\end{aligned}$$

given condition (iii) of Definition 2.2 and the fact that $\mathbf{E}\|v_t\tilde{\varepsilon}_{t-1}\| < \infty$ that can be easily checked.

Therefore we have shown that

$$\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Delta\tilde{\varepsilon}_t = o_p(1)-\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})v_t\tilde{\varepsilon}_{t-1}.$$

But

$$\begin{aligned}
&\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})v_t\tilde{\varepsilon}_{t-1} \\
={}& \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\mathbf{E}\left(v_t\tilde{\varepsilon}_{t-1}\right)-\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\left(v_t\tilde{\varepsilon}_{t-1}-\mathbf{E}\left(v_t\tilde{\varepsilon}_{t-1}\right)\right) \\
={}& \int_0^1\dot{h}_f(V(r))dr\mathbf{E}\left(v_t\tilde{\varepsilon}_{t-1}\right)-\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\left(v_t\tilde{\varepsilon}_{t-1}-\mathbf{E}\left(v_t\tilde{\varepsilon}_{t-1}\right)\right)+o_p(1).
\end{aligned}$$

In view of the fact $\mathbf{E}\left(v_t\tilde{\varepsilon}_{t-1}\right) = \Lambda_{vu}$ it would suffice to show that the last term above is asymptotically negligible, to complete the proof of part (i). This is what we set out to do next. Consider the term

$$\begin{aligned}
&\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\left(v_t\tilde{\varepsilon}_{t-1}-\mathbf{E}\left(v_t\tilde{\varepsilon}_{t-1}\right)\right)\\
=\ &\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\sum_{j=0}^{\infty}\Psi_j\tilde{\Phi}_{j+1}\left(\eta_{t-j}\varepsilon_{t-j-1}-\Sigma_{\eta\varepsilon}\right)\\
&+\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\sum_{r=1}^{\infty}\sum_{j=0}^{\infty}\Psi_j\tilde{\Phi}_{j+r+1}\eta_{t-j}\varepsilon_{t-j-r-1}\\
&+\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\sum_{r=1}^{\infty}\sum_{j=0}^{\infty}\Psi_{j+r-1}\tilde{\Phi}_j\eta_{t-j-r}\varepsilon_{t-j-1}\\
:\ &=I_{1n}+I_{2n}+I_{3n}.
\end{aligned}$$

We will show that I_{1n}, I_{2n}, and I_{3n} are asymptotically negligible. First define the lag polynomials $A(L)$, $\tilde{A}(L)$, $B_r(L)$, $\tilde{B}_r(L)$, $C_r(L)$, $\tilde{C}_r(L)$ with $r\in\mathbb{N}$. by:

$$\begin{aligned}
A(L) &= \sum_{j=0}^{\infty}A_jL^j,\ \tilde{A}(L)=\sum_{j=0}^{\infty}\tilde{A}_jL^j\\
\text{with } A_j &= \Psi_j\tilde{\Phi}_{j+1}\text{ and }\tilde{A}_j=\sum_{s=j+1}^{\infty}A_s,
\end{aligned}$$

$$\begin{aligned}
B_r(L) &= \sum_{j=0}^{\infty}B_{rj}L^j,\ \tilde{B}_r(L)=\sum_{j=0}^{\infty}\tilde{B}_{rj}L^j\\
\text{with } B_{rj} &= \Psi_j\tilde{\Phi}_{j+1+r}\text{ and }\tilde{B}_{rj}=\sum_{s=j+1}^{\infty}B_{rs},
\end{aligned}$$

and

$$\begin{aligned}
C_r(L) &= \sum_{j=0}^{\infty}C_{rj}L^j,\ \tilde{C}_r(L)=\sum_{j=0}^{\infty}\tilde{C}_{rj}L^j\\
\text{with } C_{rj} &= \Psi_{j+r-1}\tilde{\Phi}_j\text{ and }\tilde{C}_{rj}=\sum_{s=j+1}^{\infty}C_{rs}.
\end{aligned}$$

Also define ζ_{tj}, ζ_{tj}^r and $\bar{\zeta}_{tj}^r$ by

$$\begin{aligned}\zeta_{tj} &= \eta_{t-j}\varepsilon_{t-j-1} - \Sigma_{\eta\varepsilon},\\ \zeta_{tj}^r &= \eta_{t-j}\varepsilon_{t-j-r-1},\\ \bar{\zeta}_{tj}^r &= \eta_{t-j-r+1}\varepsilon_{t-j-1}.\end{aligned}$$

We will apply BN decomposition to the lag polynomials $A(L)$, $B_r(L)$, $C_r(L)$:

$$A(L) = A(1) - (1-L)\,\tilde{A}(L),$$

$$B_r(L) = B_r(1) - (1-L)\,\tilde{B}_r(L)$$

and

$$C_r(L) = C_r(1) - (1-L)\,\tilde{C}_r(L).$$

We start with I_{1n}. Define $\tilde{\zeta}_t = \sum_{j=0}^{\infty} \tilde{A}_j \zeta_{tj}$. Then using BN on I_{1n}

$$\begin{aligned}I_{1n} &= \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n} \dot{f}(x_{t-1})\sum_{j=0}^{\infty} A_j\zeta_{tj}\\ &= \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n} \dot{f}(x_{t-1})A(1)\zeta_{t0}\\ &\quad -\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n} \dot{f}(x_{t-1})\Delta\tilde{\zeta}_t.\end{aligned}$$

Note that $\{\zeta_{t0}, \mathcal{F}_{t-1}\}_{t=1}^{n}$ is a martingale difference sequence. Hence $\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n} \dot{f}(x_{t-1})A(1)\zeta_{t0}$ $O_p(1/\sqrt{n})$ by Theorem 3.3 in P&P. The second term above

$$\begin{aligned}&\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n} \dot{f}(x_{t-1})\Delta\tilde{\zeta}_t\\ =\;& \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\dot{f}(x_{n-1})\tilde{\zeta}_n - \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta\dot{f}(x_{t-1})\tilde{\zeta}_{t-1}\\ =\;& O_p(1/n) - \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta\dot{f}(x_{t-1})\tilde{\zeta}_{t-1}.\end{aligned}$$

Now given the H_1*-regularity* of f and the fact that $\mathbf{E}\left\|\tilde{\zeta}_{t-1}\right\| < \infty$, using the same arguments as above we can show that $\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta\dot{f}(x_{t-1})\tilde{\zeta}_{t-1} = o_p(1)$. Hence $I_{1n} = o_p(1)$.

Next will show the result for I_{2n}. First define $\zeta_t^B = \eta_t\sum_{r=1}^{\infty}B_r(1)\varepsilon_{t-r-1}$ and $\tilde{\zeta}_t^B = \sum_{r=1}^{\infty}\sum_{j=0}^{\infty}\tilde{B}_{jr}\zeta_{tj}^r$. Hence an application of the BN decomposition on I_{2n} gives

$$\begin{aligned} I_{2n} &= \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\sum_{r=1}^{\infty}\sum_{j=0}^{\infty}B_{jr}\zeta_{tj}^r \\ &= \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\zeta_t^B - \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\Delta\tilde{\zeta}_t^B. \end{aligned}$$

Note that $\{\zeta_t^B, \mathcal{F}_{t-1}\}_{t=1}^n$ is a martingale difference sequence, therefore $\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\zeta_{t-1}^B$ $= O_p(1/\sqrt{n})$ by Theorem 3.3 in P&P. The second term above

$$\begin{aligned} &\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\Delta\tilde{\zeta}_t^B \\ &= \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\dot{f}(x_{n-1})\tilde{\zeta}_n^B - \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta\dot{f}(x_{t-1})\tilde{\zeta}_{t-1}^B \\ &= O_p(1/n) - \frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta\dot{f}(x_{t-1})\tilde{\zeta}_{t-1}^B. \end{aligned}$$

Now note that $\mathbf{E}\left\|\tilde{\zeta}_{t-1}^B\right\| < \infty$ because,

$$\begin{aligned} &\mathbf{E}\left\|\tilde{\zeta}_{t-1}^B\right\| \\ &= \mathbf{E}\left\|\sum_{r=1}^{\infty}\sum_{j=0}^{\infty}\tilde{B}_{jr}\zeta_{tj}^r\right\| \leq \lim\inf_{s_1\to\infty}\lim\inf_{s_2\to\infty}\sum_{r=1}^{s_1}\sum_{j=0}^{s_2}\mathbf{E}\left\|\tilde{B}_{jr}\zeta_{tj}^r\right\| \\ &\leq \mathbf{E}\left\|\zeta_{tj}^r\right\|\sum_{r=1}^{\infty}\sum_{j=0}^{\infty}\sum_{s=j+1}^{\infty}\left\|\Psi_s\tilde{\Phi}_{s+r}\right\| = \mathbf{E}\left\|\zeta_{tj}^r\right\|\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\sum_{j=0}^{s-1}\left\|\Psi_s\tilde{\Phi}_{s+r}\right\| \\ &\leq \mathbf{E}\left\|\zeta_{tj}^r\right\|\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\left\|s\Psi_s\right\|\left\|\tilde{\Phi}_{s+r}\right\| \leq \mathbf{E}\left\|\zeta_{tj}^r\right\|\left(\sum_{s=1}^{\infty}\left\|s\Psi_s\right\|\right)\left(\sum_{r=1}^{\infty}\left\|\tilde{\Phi}_r\right\|\right) \\ &< \infty. \end{aligned}$$

Therefore using the same arguments as before it follows that $\frac{1}{n}\dot{k}_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\Delta\dot{f}(x_{t-1})\tilde{\zeta}_{t-1}^{B} = o_p(1)$ as well, which establishes that $I_{2n} = o_p(1)$. The proof that I_{3n} is negligible is the same with the proof for I_{2n} and therefore omitted.

For part (ii) applying BN decomposition we have

$$\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)v_t' = \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\eta_t'\Psi'(1) - \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Delta\tilde{\eta}_t',$$

with $\tilde{\eta}_t = \sum_{j=1}^{\infty}\left(\sum_{i=j+1}^{\infty}\Psi_i\right)\eta_{t-j}$ Now note that note that using similar arguments as above one can show that

$$\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\eta_t'\Psi'(1) = \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_{t-1})\eta_t'\Psi'(1) + o_p(1)$$

and hence $\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\eta_t'\Psi'(1) \xrightarrow{p} \int_0^1 h_f(V(r))dV'(r)$ from Theorem 3.3 in P&P. Using the same arguments as those in part (i) we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}f(x_t)\Delta\tilde{\eta}_t' \\
= \;& o_p(1) - \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})v_t\tilde{\eta}_t' - \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\left(v_t\tilde{\eta}_t' - \mathbf{E}\left(v_t\tilde{\eta}_{t-1}'\right)\right) \\
= \;& o_p(1) - \int_0^1 \dot{h}_f(V(r))dr\mathbf{E}\left(v_t\tilde{\eta}_{t-1}'\right) - \frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\left(v_t\tilde{\eta}_t' - \mathbf{E}\left(v_t\tilde{\eta}_{t-1}'\right)\right)
\end{aligned}$$

Using the same line of argument with part (i) it can be shown that

$$\frac{1}{\sqrt{n}}k_f^{-1}(\sqrt{n})\sum_{t=1}^{n}\dot{f}(x_{t-1})\left(v_t\tilde{\eta}_t' - \mathbf{E}\left(v_t\tilde{\eta}_{t-1}'\right)\right) = o_p(1)$$

and in view of the fact that $\mathbf{E}\left(v_t\tilde{\eta}_{t-1}'\right) = \Lambda_{vv}$ the result follows.

■

PROOF OF PROPOSITION 3.1:

Note that

$$\sqrt{n}k_g\left(\hat{a}-a^*\right)=\left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n}g(x_t)g'(x_t)k_g^{-1}\right]^{-1}\frac{k_g^{-1}}{\sqrt{n}}\left[\sum_{t=1}^{n}g(x_t)u_t^+-\dot{g}_n\hat{\Lambda}_{vu}^+\right].$$

Now

$$\frac{1}{n}k_g^{-1}\sum_{t=1}^{n}g(x_t)g'(x_t)\dot{k}_g^{-1}\xrightarrow{p}\int_0^1 h_g\left(V(r)\right)h_g'\left(V(r)\right)dr \tag{1}$$

$$\frac{k_g^{-1}}{\sqrt{n}}\dot{g}_n\hat{\Lambda}_{vu}^+\xrightarrow{p}\int_0^1\dot{h}_g(V(r))dr\Lambda_{vu}^+, \tag{2}$$

and

$$\frac{k_g^{-1}}{\sqrt{n}}\sum_{t=1}^{n}g(x_t)u_t^+=\frac{k_g^{-1}}{\sqrt{n}}\sum_{t=1}^{n}g(x_t)u_t-\frac{k_g^{-1}}{\sqrt{n}}\sum_{t=1}^{n}g(x_t)v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}.$$

The first term

$$\frac{k_g^{-1}}{\sqrt{n}}\sum_{t=1}^{n}g(x_t)u_t=\int_0^1 h_g\left(V(r)\right)dU(r)+\int_0^1\dot{h}_g(V(r))dr\Lambda_{vu}+o_p(1),$$

by Proposition 2.1 (i). The second term

$$\begin{aligned}&\frac{k_g^{-1}}{\sqrt{n}}\sum_{t=1}^{n}g(x_t)v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\\ =\;&\int_0^1 h_g\left(V(r)\right)d\left(V'(r)\Omega_{vv}^{-1}\Omega_{vu}\right)+\int_0^1\dot{h}_g(V(r))dr\Lambda_{vv}\Omega_{vv}^{-1}\Omega_{vu}+o_p(1),\end{aligned}$$

by Proposition 2.1 (ii). Hence

$$\frac{k_g^{-1}}{\sqrt{n}}\sum_{t=1}^{n}g(x_t)u_t^+\xrightarrow{p}\int_0^1 h_g\left(V(r)\right)dU^+(r)+\int_0^1\dot{h}_g(V(r))dr\left(\Lambda_{vu}-\Lambda_{vv}\Omega_{vv}^{-1}\Omega_{vu}\right) \tag{3}$$

In view of (1), (2) and (3) the result follows.■

PROOF OF PROPOSITION 3.2:

Note that

$$\begin{aligned}
CM_n &= \frac{\left[\sum_{t=1}^{n}\left(y_t^{+} - g_t{}'\hat{a} - v_t'\Omega_{vv}^{-1}\Omega_{vu}\right)\right]^2}{\left(\Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}\right)\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g_t - 1\right]^2} + o_p(1) \\
&= \frac{\left[\sum_{t=1}^{n}\left(g_t'a_o - g_t{}'\hat{a} + u_t - v_t'\Omega_{vv}^{-1}\Omega_{vu}\right)\right]^2}{\left(\Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}\right)\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g_t - 1\right]^2} \\
&= \frac{\left[\sum_{t=1}^{n}\left(g_t'\left(\hat{a} - a_o\right) - u_t + v_t'\Omega_{vv}^{-1}\Omega_{vu}\right)\right]^2}{\left(\Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}\right)\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g_t - 1\right]^2} \\
&= \frac{\left[\sum_{t=1}^{n} g_t'\left[\sum_{i=1}^{n} g_i g_i'\right]^{-1}\left[\sum_{j=1}^{n} g_j u_j^{+} - \dot{g}_n\hat{\Lambda}_{vu}^{+}\right] - \sum_{t=1}^{n}\left(u_t + v_t'\Omega_{vv}^{-1}\Omega_{vu}\right)\right]^2}{\left(\Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}\right)\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g_t - 1\right]^2}
\end{aligned}$$

Now consider the term

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g_t'\left[\sum_{i=1}^{n} g_i g_i'\right]^{-1}\left[\sum_{j=1}^{n} g_j u_j^{+} - \dot{g}_n\hat{\Lambda}_{vu}^{+}\right] \\
&\qquad - \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(u_t + v_t'\Omega_{vv}^{-1}\Omega_{vu}\right) \\
&= \frac{1}{n}\sum_{t=1}^{n} g_t' k_g^{-1}\left[\frac{1}{n}k_g^{-1}\sum_{i=1}^{n} g_i g_i' k_g^{-1}\right]^{-1}\frac{k_g^{-1}}{\sqrt{n}}\left[\sum_{j=1}^{n} g_j u_j^{+} - \dot{g}_n\hat{\Lambda}_{vu}^{+}\right] \\
&\qquad - \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(u_t + v_t'\Omega_{vv}^{-1}\Omega_{vu}\right) \\
&= \frac{1}{n}\sum_{t=1}^{n} g_t' k_g^{-1} B^{-1}\int_0^1 h_g(V(r))dU^{+}(r) - \int_0^1 dU^{+}(r) + o_p(1) \\
&= A'B^{-1}\int_0^1 h_g(V(r))dU^{+}(r) - \int_0^1 dU^{+}(r) + o_p(1) \\
&= \int_0^1\left[A'B^{-1}h_g(V(r)) - 1\right]dU^{+}(r) + o_p(1).
\end{aligned}$$

The term

$$\frac{1}{n}\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g(x_t) - 1\right]^2 = \int_0^1\left[A'B^{-1}h_g(V(r)) - 1\right]^2 dr + o_p(1)$$

Hence

$$CM_n \xrightarrow{d} \frac{\left[\int_0^1 \left[A'B^{-1}h_g(V(r)) - 1\right] dU^+(r)\right]^2}{\left(\Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}\right)\int_0^1 \left[A'B^{-1}h_g(V(r)) - 1\right]^2 dr}$$

and in view of the fact that V and U^+ are independent the result follows.■

PROOF OF PROPOSITION 3.3:

We start with the case of FFM and in particular when $k_{d^*} \neq k_{g^*}$. The arguments we use are similar to the ones of Phillips (1991). Under incorrect functional form

$$\hat{a}_{LS} = \left[\sum_{t=1}^{n} g_t g_t'\right]^{-1}\left[\sum_{t=1}^{n} g_t f_t'\theta_o + g_t u_t\right]$$

$$\frac{k_g}{k_{d^*}}\left(\hat{a}_{LS} - \theta_o\right) = \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t d_t'\theta_o + \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t u_t$$

Hence

$$\begin{aligned}\frac{k_g}{k_{d^*}}\left(\hat{a}_{LS} - \theta_o\right) = &\left[\int_0^1 h_g\left(V(r)\right)h_g'\left(V(r)\right)dr\right]^{-1}\int_0^1 h_g\left(V(r)\right)h_{d^*}'\left(V(r)\right)\theta_o dr + o_p(1)\\ &+\frac{1}{\sqrt{n}k_{d^*}}\left[\int_0^1 h_g\left(V(r)\right)h_g'\left(V(r)\right)dr\right]^{-1}\int_0^1 h_g\left(V(r)\right)dU(r)\\ &+\frac{1}{\sqrt{n}k_{d^*}}\left[\int_0^1 h_g\left(V(r)\right)h_g'\left(V(r)\right)dr\right]^{-1}\int_0^1 \dot{h}_g(V(r))dr\Lambda_{vu}\\ &= \zeta_1 + o_p(1).\end{aligned}$$

Define the normalising matrix $N_{d^*,n} = \begin{pmatrix} I_p & \mathbf{0} \\ \mathbf{0} & k_g/k_{d^*}\end{pmatrix}$. In what follows the regression residuals (from OLS estimation) will be written in the following form:

$$\begin{aligned}\hat{u}_t &= f_t'\theta_o - g_t'\hat{a} + u_t\\ &= d_t'\theta_o - g_t'\left(\hat{a} - \theta_o\right) + u_t \qquad (4)\end{aligned}$$

Hence

$$\begin{aligned}
\frac{n^{1/2}}{Mk_{d^*}}\hat{\Omega}_{vu} &= \frac{1}{Mk_{d^*}}\sum_{h=-M}^{M}\kappa\left(\frac{h}{M}\right)\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}v_t\begin{pmatrix} d'_{t+h} & g'_{t+h}\end{pmatrix}\begin{pmatrix}\theta_o \\ -(\hat{a}-\theta_o)\end{pmatrix}+\frac{1}{n}\sum_{t=1}^{n}v_t u_{t+h}\right) \\
&= \frac{1}{Mk_{d^*}}\sum_{h=-M}^{M}\kappa\left(\frac{h}{M}\right)\left(\frac{1}{n}\sum_{t=1}^{n}v_t\begin{pmatrix} d'_{t+h} & g'_{t+h}\end{pmatrix}N_{d^*,n}^{-1}N_{d^*,n}\begin{pmatrix}\theta_o \\ -(\hat{a}-\theta_o)\end{pmatrix}\right) \\
&\quad +\frac{n^{1/2}}{Mk_{d^*}}\Omega_{vu} \\
&= 2\pi K(0)\int_0^1 dV(r)\left(h'_{d^*}(V(r))\theta_o - h'_g(V(r))\zeta_1\right) \\
&\quad +\Omega_{vv}\int_0^1\left[\dot{h}'_{d^*}(V(r))\theta_o - \dot{h}'_g(V(r))\zeta_1\right]dr + o_p(1) \\
&\quad +\frac{n^{1/2}}{Mk_{d^*}}\Omega_{vu} + o_p(1) \\
&= 2\pi K(0)\int_0^1 dV(r)\bar{h}'_1(V(r))\bar{\zeta}_1 + \Omega_{vv}\int_0^1 \dot{H}'_1(V(r))\bar{\zeta}_1 dr + \frac{n^{1/2}}{Mk_{d^*}}\Omega_{vu} + o_p(1)
\end{aligned}$$

Hence

$$\hat{\Omega}_{vu} = O_p\left(\frac{Mk_{d^*}}{n^{1/2}}\right) + O_p(1) \tag{5}$$

Now using similar arguments as above it turns out that

$$\frac{n^{1/2}}{Mk_{d^*}}\hat{\Lambda}_{vu} = 2\pi K_1(0)\int_0^1 dV(r)\bar{h}'_1(V(r))\bar{\zeta}_1 + \Lambda_{vv}\int_0^1 \dot{H}'_1(V(r))\bar{\zeta} dr + \frac{n^{1/2}}{Mk_{d^*}}\Lambda_{vu} + o_p(1)$$

So

$$\hat{\Lambda}_{vu} = O_p\left(\frac{Mk_{d^*}}{n^{1/2}}\right) \tag{6}$$

Next we will find the order of $\hat{\Omega}_{uu}$. Note that

$$\hat{\Omega}_{uu} = \sum_{h=-M}^{M}\kappa\left(\frac{h}{M}\right)C_{uu}(h)$$

and

$$\frac{1}{k_{d^*}^2}C_{uu}(h) = \frac{1}{nk_{d^*}^2}\sum_{t=1}^{n}\begin{pmatrix}\theta_o & -(\hat{a}-\theta_o)\end{pmatrix}\begin{pmatrix} d_t \\ g_t\end{pmatrix}\begin{pmatrix} d'_{t+h} & g'_{t+h}\end{pmatrix}\begin{pmatrix}\theta_o \\ -(\hat{a}-\theta_o)\end{pmatrix}$$

$$
\begin{aligned}
& +\frac{1}{nk_{d^*}^2}\sum_{t=1}^{n}\left(\begin{array}{cc}\theta_o & -(\hat{a}-\theta_o)\end{array}\right)\left(\begin{array}{c}d_t\\ g_t\end{array}\right)u_{t+h} \\
& +\frac{1}{nk_{d^*}^2}\sum_{t=1}^{n}\left(\begin{array}{cc}d_{t+h}' & g_{t+h}'\end{array}\right)\left(\begin{array}{c}\theta_o\\ -(\hat{a}-\theta_o)\end{array}\right)u_t \\
= & \frac{1}{nk_{d^*}^2}\sum_{t=1}^{n}\left(\begin{array}{cc}\theta_o & -\frac{k_g}{k_{d^*}}(\hat{a}-\theta_o)\end{array}\right)N_{d^*,n}^{-1}\left(\begin{array}{cc}d_t d_{t+h}' & d_t g_{t+h}'\\ g_t d_{t+h}' & g_t g_{t+h}'\end{array}\right)N_{d^*,n}^{-1}\left(\begin{array}{c}\theta_o\\ -\frac{k_g}{k_{d^*}}(\hat{a}-\theta_o)\end{array}\right) \\
& +\frac{1}{nk_{d^*}^2}\sum_{t=1}^{n}\left(\begin{array}{cc}\theta_o & -\frac{k_g}{k_{d^*}}(\hat{a}-\theta_o)\end{array}\right)\left(\begin{array}{c}d_t\\ g_t\end{array}\right)u_{t+h} \\
& +\frac{1}{nk_{d^*}^2}\sum_{t=1}^{n}\left(\begin{array}{cc}d_{t+h}' & g_{t+h}'\end{array}\right)\left(\begin{array}{c}\theta_o\\ -(\hat{a}-\theta_o)\end{array}\right)u_t \\
= & \int_0^1 \bar{\zeta}_1'\bar{h}_1(V(r))\bar{h}_1'(V(r))\bar{\zeta}_1 dr
\end{aligned}
$$

(from Lemma A)

$$
\begin{aligned}
& +\frac{1}{\sqrt{n}k_{d^*}}\left(\int_0^1 \bar{\zeta}_1'\bar{h}_1(V(r))dU(r)+\int_0^1 \bar{\zeta}_1'\dot{H}_1(V(r))dr\Lambda_{vu}(h)\right) \\
& +\frac{1}{\sqrt{n}k_{d^*}}\left(\int_0^1 \bar{h}_1'\bar{\zeta}_1(V(r))dU(r)+\Lambda_{uv}(h)\int_0^1 \dot{H}_1'(V(r))\bar{\zeta}_1 dr\right).
\end{aligned}
$$

So,

$$
\begin{aligned}
\frac{1}{Mk_{d^*}^2}\hat{\Omega}_{uu} &= 2\pi K(0)\int_0^1 \bar{\zeta}_1'\bar{h}_1(V(r))\bar{h}_1'(V(r))\bar{\zeta}_1 dr \\
& +O_p\left(\frac{1}{M\sqrt{n}k_{d^*}}\right)
\end{aligned}
$$

Hence under the assumption that $\int_0^1 \bar{h}_1'\bar{\zeta}_1(V(r))dr \neq 0$[4] we have

$$
\hat{\Omega}_{uu} = O_p\left(Mk_{d^*}^2\right)+O_p\left(\frac{k_{d^*}}{\sqrt{n}}\right) = O_p\left(Mk_{d^*}^2\right) \tag{7}
$$

Consequently under FFM (4) and (5) give

$$
\begin{aligned}
\hat{\Omega}_{uu}-\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu} &= O_p\left(Mk_{d^*}^2\right)+O_p\left(\frac{M^2k_{d^*}^2}{n}\right) \\
&= O_p\left(Mk_{d^*}^2\right) \text{ when } M/n\to 0 \text{ as } n\to\infty. \qquad (8)
\end{aligned}
$$

[4] The case where $\int_0^1 \bar{h}_1'\bar{\zeta}_1(V(r))dr = 0$ will be discussed later.

To obtain the result we have assumed that the term $\int_0^1 \bar{h}_1'(V(r))\bar{\zeta}_1 dr \neq 0$. When this condition does not hold the line of argument used in the proof above shows that under FFM, the statistic is $o_p(n/M)$ rather than $O_p(n/M)$. The condition above does not hold when k_{d^*}, the order of the dominating element in d_t is determined by some element of the vector g_t the condition above does not hold. To see this suppose that $k_{d^*} = k_{g^*}$ and the first element of g_t dominates the rest. In this case the FM-LS estimator

$$\begin{aligned}
\frac{k_g}{k_{d^*}}(\hat{a}-\theta_o) &= \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}} k_g^{-1} \sum_{t=1}^{n} g_t d_t' \theta_o + o_p(1) \\
&= \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{g^*}} k_g^{-1} \sum_{t=1}^{n} g_t g_t' \theta_o \\
&= \left[\int_0^1 h_g(V(r))h_g'(V(r))dr\right]^{-1} \begin{pmatrix} \int_0^1 h_{g_1}^2(V)dr & 0 & \dots & 0 \\ \cdot & 0 & & 0 \\ \cdot & \cdot & \cdot & \\ \int_0^1 h_{g_p}(V)h_{g_1}(V)dr & 0 & \dots & 0 \end{pmatrix} \theta_o \\
&= \begin{pmatrix} \theta_{1o} \\ 0 \\ \cdot \\ 0 \end{pmatrix}. \qquad (9)
\end{aligned}$$

Now consider the sum of residuals that appear in (9):

$$\begin{aligned}
& \frac{1}{nk_{d^*}} \sum\nolimits_{t=1}^{n} (y_t - g(x_t)'\hat{a}) \\
&= \frac{1}{nk_{d^*}} \sum_t \left(d_t'\theta_o - g_t'(\hat{a}-\theta_o)\right) \\
&= \frac{1}{nk_{d^*}} \sum_t d_t'\theta_o - \frac{1}{n}\sum_t g_t' k_g^{-1} k_g \frac{k_g}{k_{d^*}}(\hat{a}-\theta_o) \\
&= \theta_{1o}\int_0^1 h_{g_1}(V(r))dr - \int_0^1 h_g'(V(r))dr \begin{pmatrix} \theta_{1o} \\ 0 \\ \cdot \\ 0 \end{pmatrix} + o_p(1) \\
&= o_p(1) \text{ (from (9))}.
\end{aligned}$$

Hence $\int_0^1 \bar{h}_1'(V(r))\bar{\zeta}_1 = 0$ and therefore the numerator of the test statistic is of order $o_p\left((nk_{d^*})^2\right)$ rather than $O_p\left((nk_{d^*})^2\right)$ that is stated in (9).

To obtain the requisite result in the case when $k_{d^*} = k_{g^*}$ we use different expression for the regression residuals than the one of (4). First the LS estimator

$$\begin{aligned}
\frac{k_g}{k_{f^*}}\hat{a}_{LS} &= \left[k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{k_g^{-1}}{k_{f^*}}\left[\sum_{t=1}^{n} g_t f_t'\theta_o + g_t u_t\right] \\
&= \left[\int_0^1 h_g\left(V(r)\right) h_g'\left(V(r)\right) dr\right]^{-1} \int_0^1 h_g\left(V(r)\right) h_{f^*}'\left(V(r)\right)\theta_o dr + o_p(1) \\
&= \zeta_2 + o_p(1)
\end{aligned}$$

Now the LS residuals

$$\begin{aligned}
& \frac{1}{nk_{f^*}}\sum\nolimits_{t=1}^{n}\left(y_t - g(x_t)'\hat{a}\right) \\
=\ & \frac{1}{nk_{f^*}}\sum\nolimits_{t=1}^{n} f_t'\theta_o - \frac{1}{n}\sum\nolimits_{t=1}^{n} g(x_t)'k_g^{-1}\frac{k_g}{k_{f^*}}\hat{a} + o_p(1) \\
=\ & \int_0^1 h_{f^*}'\left(V(r)\right)\theta_o dr - \int_0^1 h_g'\left(V(r)\right)\zeta_2 dr + o_p(1) \\
=\ & \int_0^1 \bar{h}_2'\left(V(r)\right)\bar{\zeta}_2 dr + o_p(1)
\end{aligned}$$

Now similar arguments as those above give

$$\frac{n^{1/2}}{Mk_{f^*}}\hat{\Omega}_{vu} = 2\pi K(0)\int_0^1 dV(r)\bar{h}_2'(V(r))\bar{\zeta}_2 + \Omega_{vv}\int_0^1 \dot{H}_2'(V(r))\bar{\zeta}_2 dr + \frac{n^{1/2}}{Mk_{2^*}}\Omega_{vu} + o_p(1),$$

$$\frac{n^{1/2}}{Mk_{f^*}}\hat{\Lambda}_{vu} = 2\pi K_1(0)\int_0^1 dV(r)\bar{h}_2'(V(r))\bar{\zeta}_2 + \Lambda_{vv}\int_0^1 \dot{H}_2'(V(r))\bar{\zeta}_2 dr + \frac{n^{1/2}}{Mk_{f^*}}\Lambda_{vu} + o_p(1)$$

and

$$\begin{aligned}
\frac{1}{Mk_{f^*}^2}\hat{\Omega}_{uu} &= 2\pi K(0)\int_0^1 \bar{\zeta}_2'\bar{h}_2(V(r))\bar{h}_2'(V(r))\bar{\zeta}_2 dr \\
&\quad + O_p\left(\frac{1}{M\sqrt{n}k_{f^*}}\right). \qquad (10)
\end{aligned}$$

When there is no cointegration we have

$$\begin{aligned}
\frac{n^{1/2}}{Mk_s}\hat{\Omega}_{vu} &= \frac{1}{Mk_s}\sum_{h=-M}^{M}\kappa\left(\frac{h}{M}\right)\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}v_t\begin{pmatrix} s_{t+h} & g'_{t+h}\end{pmatrix}\begin{pmatrix}1\\ -\hat{a}\end{pmatrix}\right)\\
&= \frac{1}{Mk_s}\sum_{h=-M}^{M}\kappa\left(\frac{h}{M}\right)\left(\frac{1}{n}\sum_{t=1}^{n}v_t\begin{pmatrix} s_{t+h} & g'_{t+h}k_g^{-1}k_s\end{pmatrix}\begin{pmatrix}1\\ -\frac{k_g}{k_s}\hat{a}\end{pmatrix}\right)\\
&= 2\pi K(0)\int_0^1 dV(r)\left(h_s(W(r)) - h'_g(V(r))\zeta_3\right)\\
&\quad +\Omega_{vw}\int_0^1 h_{\dot{s}}(W(r))dr - \Omega_{vv}\int_0^1 h'_{\dot{g}}(V(r))\zeta_3 dr + o_p(1)\\
&= 2\pi K(0)\int_0^1 dV(r)\bar{h}'_3(W(r),V(r))\bar{\zeta}_3\\
&\quad +\bar{\Omega}\int_0^1 \dot{H}'_3\left(W(r),V(r)\right)\bar{\zeta}_3 dr + o_p(1)
\end{aligned}$$

Now using the same arguments as above we have

$$\begin{aligned}
\frac{n^{1/2}}{Mk_s}\hat{\Lambda}_{vu} &= 2\pi K_1(0)\int_0^1 dV(r)\bar{h}'_3(W(r),V(r))\bar{\zeta}_3\\
&\quad +\bar{\Lambda}\int_0^1 \dot{H}'_3\left(W(r),V(r)\right)\bar{\zeta}_3 dr + o_p(1)
\end{aligned}$$

and

$$\frac{1}{Mk_s^2}\hat{\Omega}_{uu} = \int_0^1 \bar{\zeta}'_3\dot{H}_3\left(W(r),V(r)\right)\dot{H}'_3\left(W(r),V(r)\right)\bar{\zeta}_3 dr + o_p(1) \tag{11}$$

and this completes the proof.■

PROOF OF PROPOSITION 3.4:

We start with the case of FFM and in particular with $k_{d^*} \neq k_{g^*}$. Note that the FM-LS estimator

$$\hat{a} = \left[\sum_{t=1}^{n} g(x_t)g'(x_t)\right]^{-1}\left[\sum_{t=1}^{n} g(x_t)y_t^{+} - \dot{g}_n\hat{\Lambda}_{vu}^{+}\right],$$

rearranging

$$
\begin{aligned}
\frac{k_g}{k_{d^*}}(\hat{a}-\theta_o) &= \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t d_t'\theta_o \\
&+ \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t u_t \\
&- \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu} \\
&- \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\dot{g}_n\hat{\Lambda}_{vu}^{+} \\
&= \zeta_1 + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\right) + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\times\frac{Mk_{d^*}}{\sqrt{n}}\right) + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\times\frac{Mk_{d^*}}{\sqrt{n}}\right) + o_p(1) \\
&= \zeta_1 + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\right) + O_p\left(\frac{M}{n}\right) + o_p(1),
\end{aligned}
$$

where we have used the fact that $\hat{\Lambda}_{vu}, \hat{\Omega}_{vu} = O_p\left(\frac{Mk_{d^*}}{\sqrt{n}}\right)$ (equations (4),(5)). Recall that the test statistic is

$$
CM_n = \frac{\left[\sum_{t=1}^{n}\left(y_t^{+} - g(x_t)'\hat{a} - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\right]^2}{\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g(x_t) - 1\right]^2},
$$

Consider first the numerator rescaled by $(nk_{d^*})^2$:

$$
\begin{aligned}
&\left[\frac{1}{nk_{d^*}}\sum_{t=1}^{n}\left\{\left(y_t^{+} - g(x_t)'\hat{a}\right) - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right\}\right]^2 \\
=\ &\left[\frac{1}{nk_{d^*}}\sum_{t=1}^{n}\left\{\begin{pmatrix} d_t' & g_t' \end{pmatrix}\begin{pmatrix} \theta_o \\ -(\hat{a}-\theta_o) \end{pmatrix} + \left(u_t - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\right\}\right]^2 \\
=\ &\left[\frac{1}{nk_{d^*}}\sum_{t=1}^{n}\left\{\begin{pmatrix} d_t' & g_t' \end{pmatrix} N_{d^*,n}^{-1}\begin{pmatrix} \theta_o \\ -\frac{k_g}{k_{d^*}}(\hat{a}-\theta_o) \end{pmatrix} + \left(u_t - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\right\}\right]^2 \\
=\ &\left[\int_0^1 \bar{h}_1'(V(r))\bar{\zeta}_1 dr + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\right) + O_p\left(\frac{M}{n}\right)\right]^2 \qquad (12)
\end{aligned}
$$

with $N_{d^*,n} = \begin{pmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \frac{k_g}{k_{d^*}} \end{pmatrix}$. Now the the denominator rescaled by n

$$\begin{aligned}
&\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\frac{1}{n}\sum_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g(x_t) - 1\right]^2 \\
=& \left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\int_0^1 \left[A'B^{-1}h_g(V(r)) - 1\right]^2 dr + o_p(1) \\
=& O_p\left(Mk_{d^*}^2\right) \text{ (from equation (8))} \qquad (13)
\end{aligned}$$

In view of (12) and (13) the result follows.

When $k_{d^*} = k_{g^*}$ the FM-LS estimator

$$\begin{aligned}
\frac{k_g}{k_{f^*}}\hat{a} =& \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{f^*}}k_g^{-1}\sum_{t=1}^{n} g_t f_t'\theta_o \\
&+ \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t u_t \\
&- \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\sum_{t=1}^{n} g_t v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu} \\
&- \left[\frac{1}{n}k_g^{-1}\sum_{t=1}^{n} g_t g_t' k_g^{-1}\right]^{-1} \frac{1}{nk_{d^*}}k_g^{-1}\dot{g}_n\hat{\Lambda}_{vu}^{+} \\
=& \zeta_2 + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\right) + O_p\left(\frac{1}{\sqrt{n}k_{d^*}} \times \frac{Mk_{d^*}}{\sqrt{n}}\right) + O_p\left(\frac{1}{\sqrt{n}k_{d^*}} \times \frac{Mk_{d^*}}{\sqrt{n}}\right) + o_p(1) \\
=& \zeta_2 + O_p\left(\frac{1}{\sqrt{n}k_{d^*}}\right) + O_p\left(\frac{M}{n}\right) + o_p(1),
\end{aligned}$$

Consider again the numerator of the statistic this time rescaled by $(nk_{f^*})^2$:

$$\begin{aligned}
&\left[\frac{1}{nk_{f^*}}\sum_{t=1}^{n}\left\{\left(y_t^+ - g(x_t)'\hat{a}\right) - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right\}\right]^2 \\
=& \left[\frac{1}{nk_{f^*}}\sum_{t=1}^{n}\left\{\begin{pmatrix} f_t' & g_t' \end{pmatrix}\begin{pmatrix} \theta_o \\ -\hat{a} \end{pmatrix} + \left(u_t - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\right\}\right]^2 \\
=& \left[\frac{1}{nk_{f^*}}\sum_{t=1}^{n}\left\{\begin{pmatrix} f_t' & g_t' \end{pmatrix} N_{f^*,n}^{-1}\begin{pmatrix} \theta_o \\ -\frac{k_g}{k_{f^*}}\left(\hat{a} - \theta_o\right) \end{pmatrix} + \left(u_t - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\right\}\right]^2 \\
=& \left[\int_0^1 \bar{h}_2'(V(r))\bar{\zeta}_2 dr + O_p\left(\frac{1}{\sqrt{n}k_{f^*}}\right) + O_p\left(\frac{M}{n}\right)\right]^2 \qquad (14)
\end{aligned}$$

with $N_{f^*,n} = \begin{pmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \frac{k_g}{k_{f^*}} \end{pmatrix}$. Now the the denominator rescaled by n

$$
\begin{aligned}
&\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\frac{1}{n}\sum\nolimits_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g(x_t) - 1\right]^2 \\
= &\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\int_0^1\left[A'B^{-1}h_g(V(r)) - 1\right]^2 dr + o_p(1) \\
= &\; O_p\left(Mk_{f^*}^2\right) \text{ (from equation (10))} \qquad (15)
\end{aligned}
$$

in view of (14) and (15) this completes this part of the proof.

Under no cointegration using the same arguments as before the numerator of the test statistic rescaled by $(nk_s)^2$ is

$$
\begin{aligned}
&\left[\frac{1}{nk_s}\sum\nolimits_{t=1}^{n}\left\{\left(y_t^+ - g(x_t)'\hat{a}\right) - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right\}\right]^2 \\
= &\left[\frac{1}{nk_s}\sum\nolimits_{t=1}^{n}\left\{\begin{pmatrix} s_t & g_t' \end{pmatrix}\begin{pmatrix} 1 \\ -\hat{a} \end{pmatrix} - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right\}\right]^2 \\
= &\left[\frac{1}{nk_s}\sum\nolimits_{t=1}^{n}\left\{\begin{pmatrix} s_t & g_t' \end{pmatrix}N_{s,n}^{-1}\begin{pmatrix} 1 \\ -\frac{k_g}{k_s}\hat{a} \end{pmatrix} - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right\}\right]^2 \\
= &\left[\int_0^1 \bar{h}_3'(W(r), V(r))\bar{\zeta}_3 dr + O_p\left(\frac{1}{\sqrt{n}k_s}\right) + O_p\left(\frac{M}{n}\right)\right]^2 \qquad (15)
\end{aligned}
$$

with $N_{s,n} = \begin{pmatrix} I_p & \mathbf{0} \\ \mathbf{0} & \frac{k_g}{k_s} \end{pmatrix}$.

The denominator

$$
\begin{aligned}
&\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\frac{1}{n}\sum\nolimits_{t=1}^{n}\left[\hat{A}_n'\hat{B}_n^{-1}g(x_t) - 1\right]^2 \\
= &\left(\hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}\right)\int_0^1\left[A'B^{-1}h_g(V(r)) - 1\right]^2 dr + o_p(1) \\
= &\; O_p\left(Mk_s^2\right) \text{ (from equation (11))} \qquad (16)
\end{aligned}
$$

and in view of (16) and (17) the result follows.■

PROOF OF PROPOSITION 3.5:

We will show that the result under FFM when $k_{d^*} \neq k_{g^*}$. The proof for the other cases is similar and therefore omitted. Denote by $u(x_t)$ the regressions residuals from FM-LS estimation. From the proof of Proposition 3.4 (equation (12)) we have that $(nk_{d^*})^{-1}\sum_{t=1}^n u(x_t) = \int_0^1 \bar{h}_1'(V(r))\bar{\zeta}_1 dr + o_p(1) = \int_0^1 h_{u^*}(V(r))dr + o_p(1)$. and similarly define $\dot{u}(x_t)$, $\ddot{u}(x_t)$, $\dot{h}_u$ and $\ddot{h}_u$. First consider

$$\begin{aligned}
\hat{\rho}^2 - 1 &= \frac{\left(\sum_{t=2}^n u(x_t)u(x_{t-1})\right)^2 - \left(\sum_{t=2}^n u(x_{t-1})^2\right)^2}{\left\{\sum_{t=2}^n u(x_{t-1})^2\right\}^2} \\
&= \frac{\left\{\sum_{t=2}^n u(x_t)\left(u(x_t) + u(x_{t-1})\right)\right\}\left\{\sum_{t=2}^n u(x_t)\left(u(x_t) - u(x_{t-1})\right)\right\}}{\left\{\sum_{t=2}^n u(x_{t-1})^2\right\}^2} \\
&= \frac{\left\{\sum_{t=2}^n u(x_t)\left(u(x_t) + u(x_{t-1})\right)\right\}\left\{\sum_{t=2}^n u(x_t)\dot{u}(x_{t-1})'v_t\right\}}{\left\{\sum_{t=2}^n u(x_{t-1})^2\right\}^2} + o_p(1) \\
&= \frac{\left\{\sum_{t=2}^n u(x_t)\left(u(x_t) + u(x_{t-1})\right)\right\}\left\{\sum_{t=2}^n u(x_{t-1})\dot{u}(x_{t-1})'v_t\right\}}{\left\{\sum_{t=2}^n u(x_{t-1})^2\right\}^2} + o_p(1)
\end{aligned}$$

where the third line is due to the mean value Theorem and the H_1-*regularity* and the last one is due to Lemma A. Hence

$$\begin{aligned}
\frac{\sqrt{n}k_{d^*}}{k_{\dot{d}^*}}\left(\hat{\rho}^2 - 1\right) &= \left\{\left(nk_{d^*}^2\right)^{-1}\sum_{t=2}^n u(x_{t-1})^2\right\}^{-2} \\
&\quad \times \left\{\left(nk_{d^*}^2\right)^{-1}\sum_{t=2}^n u(x_t)\left(u(x_t) + u(x_{t-1})\right)\right\} \\
&\quad \times \left\{\left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^n u(x_{t-1})\dot{u}(x_{t-1})'v_t\right\}
\end{aligned}$$

Consider the term

$$\begin{aligned}
&\left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^n u(x_{t-1})\dot{u}(x_{t-1})'v_t \\
&= \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^n u(x_{t-1})\dot{u}(x_{t-1})'\Psi(1)\eta_t - \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=1}^n u(x_{t-1})\dot{u}(x_{t-1})'\Delta\tilde{\eta}_t.
\end{aligned}$$

The first term $\left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^n u(x_{t-1})\dot{u}(x_{t-1})'\Psi(1)\eta_t \xrightarrow{p} \int_0^1 h_{u^*}(V(r))\dot{h}_{u^*}'(V(r))dV(r)$.

The second

$$\left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}u(x_{t-1})\dot{u}(x_{t-1})'\Delta\tilde{\eta}_t$$

$$= \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}u(x_{n-1})\dot{u}(x_{n-1})'\tilde{\eta}_{n-1} - \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=1}^{n}\Delta\left(u(x_{t-1})\dot{u}(x_{t-1})'\right)\tilde{\eta}_{t-2}$$

$$= o_p(1) - \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}\dot{u}(x_{t-2})'v_{t-1}\dot{u}(x_{t-2})'\tilde{\eta}_{t-2}$$

$$- \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}u(x_{t-1})v_{t-1}'\ddot{u}(x_{t-2})\tilde{\eta}_{t-2} \text{ (with } \ddot{u} \text{ diagonal)}$$

$$= o_p(1) - \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}\dot{u}(x_{t-2})'v_{t-1}\tilde{\eta}_{t-2}'\dot{u}(x_{t-2})'$$

$$- \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}u(x_{t-1})\sum_{i=1}^{p}\ddot{u}_i(x_{t-2})v_{it-1}\tilde{\eta}_{it-2}$$

$$= o_p(1) - \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}\dot{u}(x_{t-2})'\mathbf{E}\left(v_{t-1}\tilde{\eta}_{t-2}'\right)\dot{u}(x_{t-2})$$

$$- \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}\dot{u}(x_{t-2})'\left(v_{t-1}\tilde{\eta}_{t-2}' - \mathbf{E}\left(v_{t-1}\tilde{\eta}_{t-2}'\right)\right)\dot{u}(x_{t-2})$$

$$- \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}u(x_{t-1})\sum_{i=1}^{p}\ddot{u}_i(x_{t-2})\mathbf{E}\left(v_{it-1}\tilde{\eta}_{it-2}\right)$$

$$- \left(\sqrt{n}k_{d^*}k_{\dot{d}^*}\right)^{-1}\sum_{t=2}^{n}u(x_{t-1})\sum_{i=1}^{p}\ddot{u}_i(x_{t-2})\left(v_{it-1}\tilde{\eta}_{it-2} - \mathbf{E}\left(v_{it-1}\tilde{\eta}_{it-2}\right)\right)$$

$$= o_p(1) - \int_0^1 \dot{h}_{u^*}'(V(r))\Lambda_{vv}\dot{h}_{u^*}(V(r))dr - \int_0^1 h_{u^*}(V(r))\sum_{i=1}^{p}\ddot{h}_{u_i^*}\left(V(r)\right)\Lambda_{vv_{ii}}dr,$$

where v_{it-1}, $\tilde{\eta}_{it-2}$ are the i^{th} elements of v_t, $\tilde{\eta}_t$ and $\ddot{u}_{it}$, $\Lambda_{vv_{ii}}$ the i^{th} diagonal element of $\ddot{u}_t$.

and Λ_{vv} respectively. Hence

$$\begin{aligned}\frac{\sqrt{n}k_{d^*}}{k_{\dot{d}^*}}\left(\hat{\rho}^2-1\right) &= 2\left\{\int_0^1 h_{u^*}\left(V(r)\right)^2 dr\right\}^{-1} \\ &\times\left\{\int_0^1 h_{u^*}(V(r))\dot{h}_{u^*}'(V(r))dV(r)\right\} \\ &\times\left\{\int_0^1 \dot{h}_{u^*}'(V(r))\Lambda_{vv}\dot{h}_{u^*}(V(r))dr + \int_0^1 h_{u^*}(V(r))\sum_{i=1}^{p}\ddot{h}_{u_i^*}\left(V(r)\right)\Lambda_{vv_{ii}}dr\right\}\end{aligned}$$

Since $\frac{k_{d^*}}{k_{\dot{d}^*}} = \sqrt{n}$

$$\hat{\rho}^2 - 1 = O_p(1/n^2)$$

Hence

$$\hat{\rho}^2 / \left(\hat{\rho}^2 - 1\right)^2 = O_p(n)$$

and therefore

$$\left(\hat{\delta} n\right)^{1/3} = O_p(n)$$

as required.■

Chapter 5
Conclusion

The aim of this thesis was to address the issue of functional form misspecification under nonstationarity. Much of our development relies on the work produced by Park and Phillips (1999, 2001). We tried to throw some light on two aspects of functional form misspecification. First we studied the large sample properties of the least squares estimator under nonstationarity (Chapter 2). Secondly we focused on specification testing (Chapter 3 & 4).

With respect to the first area of study we find that the properties of the least squares estimator under nonstationarity are analogous to those under stationarity as long as one is confined to the *I-regular* family of transformations proposed by Park and Phillips (1999, 2001). When the theoretical framework is restricted to the *I-regular* family, under functional form misspecification the least squares estimator converges to some pseudo-true value, while the limit distribution and the convergence rates are the same as those under correct specification. The behaviour of the least squares estimator is quite different when the theoretical framework is extended to the *H-regular* family of transformations. When the true model is of different asymptotic order than the fitted one and the parameter space unbounded, the estimator may be unbounded in probability, while for compact parameter space the estimator converges to a boundary point of the parameter space. In the latter case, techniques developed to obtain limit distribution results, when the parameter is on the boundary (e.g. Andrews (1999)), are not applicable as in our case the limit objective function is not minimised at a turning point. Finally, when the true and the fitted models are of the same asymptotic order, the estimator converges to some pseudo-true value. The limit distribution is usually different and the convergence rates slower than those under correct specification. The theoretical framework of Chapter 2 was kept simple. We assumed that the true and fitted models involved a single covariate. Extensions of our results to multivariate specifications is a quite challenging task. The applicability of econometric techniques for the asymptotic analysis of nonlinear models is limited in our case. These techniques were originally developed for correctly specified and stationary models. For instance as P&P point out the Jennrich (1969) approach for establishing consistency results is not applicable to general *H-regular* models even if they are correctly specified. The Wooldridge (1994) approach is utilised by Chang, Park and Phillips (2001) to establish asymptotic results for multivariate models under correct specification. This approach requires the parameter to be interior point in the parameter space. We have seen that in many cases this does not hold when there is functional form misspecification.

The second area of our study is specification testing. In particular we considered a Bierens (1990) type of conditional moment test for functional form in two different theoretical frameworks. First in Chapter 3 the test was considered within the theoretical framework of P&P. In contrast to the stationary case we find that there is not a single divergence rate under the alternative. The divergence rate depends on the nature of the fitted and the true models, the nature of any weighting functions used and the nature of the variance estimator. While under stationarity the test is one-sided, under non-stationarity it may become two-sided when an integrable weighting function is used in the sample moment of the test statistic. The use of an integrable weighting is particularly beneficial in the case the fitted model is *H-regular* and the true one *I-regular*, as it makes the test consistent. The use of a bounded variance estimator may result in hyper divergence rates under the alternative as long as we are confined within the *H-regular* family. In Chapter 4 we attempted to make the test applicable to cointegrating relationships. A multiple regression model linear in parameters was considered. The endogeneity assumption of Chapter 3 was dropped and dependence in the errors of the model was introduced. A semiparametric approach similar to the one of Xiao and Phillips (2002) was followed in order to induce standard limit distribution under the null. The test is consistent when there is functional form misspecification or no cointegration. The divergence rate attained under the alternative is the same as the one of the CUSUM test for cointegration proposed by Xiao and Phillips (2002).

The specification test considered here is one of the many tests for functional form proposed over time. We expect that other functional form tests like White's (1981) Hauseman type of test and information equality test can be extended to our framework. Apart from the specification tests mentioned above, model selection procedures have been developed over time e.g. Cox (1961, 1962), Davidson and McKinnon (1981), Voung (1989). In contrast to the usual specification tests, those procedures involve a specific alternative. Whether these model selection procedures can be extended to nonstationary models is an open question. We expect that it is easier to do so for the Davidson and McKinnon (1981) model selection procedure. This however may prove more difficult for the Voung (1989) procedure. Voung's procedure utilises the Kullback-Liebler distance to compare rival models. In the stationary framework, from an application of the law of large numbers, the

Kullback-Liebler distance appears as the limit of the likelihood ratio of two rival models. In our case the asymptotic theory is completely different and the Kullback-Liebler distance is not relevant.

References

ANDREWS, D.W.K. (1991): "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

ANDREWS, D.W.K. (1999): "Estimation when a Parameter is on a Boundary," *Econometrica,* 67, 1341-1383.

BIERENS, H. (1990): "A Consistent Test for Functional Form", *Econometrica* 58, 1443-1458.

CHANG, Y. PARK J.Y. AND P.C.B. PHILLIPS (2001): "Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors" . *Econometrics Journal*, 4.

COX, D.R. (1961): "Tests for Separate Families of Hypotheses," *Journal of the Royal Statistical Society,* Series B, 32, 323-344.

————— (1962): "Further Results on Tests for Separate Families of Hypotheses," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability,* Vol 1. Berkeley: University of California Press.

DAVIDSON, R. AND J.G. MCKINNON, (1981): "Several Tests for Model Specification in the Presence of Alternative Hypotheses," *Econometrica,* 49, 781-793.

DE JONG, R.M (1996): "The Bierens Test under data Dependance," *Journal of Econometrics*, 72, 1-32.

——————— (2002a): " Addendum to "Asymptotics for Nonlinear Transformations of Integrated Time Series"," Working Paper, Department of Economics, Michigan State University.

——————— (2002b): "Nonlinear Estimators with Integrated Regressors but without Exogeneity," Working Paper, Department of Economics, Michigan State University.

DE JONG, R.M. AND C-H. WANG (2002): "Further Results on the Asymptotics for Nonlinear Transformations of Integrated Time Series," Working Paper, Department of Economics, Michigan State University.

HANSEN, L.P. (1981): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* 50, 1029-1054.

HONG, S. H. AND P.C.B. PHILLIPS (2004): "Testing Linearity in Cointegrating Relations with an Application to PPP," , Working Paper,.Yale University.

HU, L AND P.C.B. PHILLIPS (2004): "Nonstationary Discrete Choice," Cowles Foundation Discussion Paper 1364.

JEGANATHAN, P. (2003): "Second Order Limits of Functionals of Sums of Linear Processes that Converge to Fractional Stable Motions," Working Paper, Indian Statistical Insitute.

JENNRICH, R.I. (1969): "Asymptotic Properties of Non-linear Least Squares Estimation," *Annals of Mathematical Statistics,* 40, 633-643.

NEWEY, W.K. (1985): "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica,* 53, 1047-1070.

NEWEY, W.K. AND D.M. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics,* Vol. IV, ed. by R.F. Engle and D.L. McFadden. Amsterdam: Elsevier, 2213-2145.

PARK, J.Y. (2002): "Nonlinear Nonstationary Heteroscedasticity," *Journal of Econometrics,* 110, 383-415.

PARK, J.Y. AND P.C.B. PHILLIPS (1999): "Asymptotics for Nonlinear Transformations of Integrated Time Series," *Econometric Theory,* 15, 269-298.

———— (2001): "Nonlinear Regressions with Integrated Time Series," *Econometrica,* 69, 117-161.

PHILLIPS, P.C.B. (1986): "Time Series Regression with a Unit Root," *Econometrica,* 55, 277-301.

PHILLIPS, P.C.B. (1988): "Weak Convergence to the Matrix Stochastic Integral $\int_0^1 BdB'$," *Journal of Multivariate Analysis*, 24, 252-264.

PHILLIPS, P.C.B. (1991): "Spectral Regression for Cointegrated Time Series," in *Non-parametric and semiparametric methods in econometrics and statistics*, edited by W. Barnett, J. Powell and G. Tauchen. Cambridge: Cambridge University Press.

PHILLIPS, P.C.B. (1995): "Fully Modified Least Squares and Vector Autoregression," *Econometrica*, 63, 1023-1078.

PHILLIPS, P.C.B. AND B. E. HANSEN (1990): "Statistical Inference in Instrumental Variables Regressions with I(1) Processes," *Review of Economic Studies*, 53, 473-496.

PHILLIPS, P.C.B. AND R.H. MOON (2003): "GMM Estimation of Autoregressive Roots Near Unity with Panel Data," working paper, Cowles Foundation Discussion Paper 1390.

PHILLIPS, P.C.B. AND J. Y. PARK (1998): "Nonstationary Density Estimation and Kernel Autoregression," Cowles Foundation Discussion Paper 1181.

PHILLIPS, P.C.B. AND V. SOLO (1992): "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971-1001.

PÖTSCHER B.M. (2002): "Nonlinear Functions and Convergence to Brownian Motion: Beyond the Continuous Mapping Theorem," Working Paper University of Vienna.

RAMSEY, J.B. (1969): "Classical Model Selection Through Specification Error Tests," *Journal of the Royal Statistical Society,* Series B, 32, 350-371.

SAIKKONEN, P. AND I. CHOI (2004): "Cointegrating Smooth Transition Regressions," *Econometric Theory*, 20, 301-340.

TAUCHEN, G. (1985): "Diagnostic Testing and Evaluation of Maximum Likelihood Models," *Journal of Econometrics,* 30, 415-443.

VOUNG, Q. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica,* 57, 307-333.

WHITE, H. (1981): "Consequences and Detection of Misspecified Nonlinear Regression Models," *Journal of the American Statistical Association,* 76, 419-433.

WHITE, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica,* 50, 1-25.

WHITE, H AND I. DOMOWITZ. (1984): "Nonlinear Regression with Dependent Observations" *Econometrica,* 52, 143-162.

WOOLDRIDGE, J.M. (1994): "Estimation and Inference for Dependent Processes," in *Handbook of Econometrics,* Vol. IV, ed. by R.F. Engle and D.L. McFadden. Amsterdam: Elsevier, 2639-2738.

WU, C.F. (1981): "Asymptotic Theory for Non-linear Least Squares Estimation," *Annals of Statistics,* 9, 501-513.

XIAO, Z. AND P.C.B. PHILLIPS (2002): "A CUSUM Test for Cointegration Using Regression Residuals," *Journal of Econometrics*, 108.