Testing for Structural Change in the Presence of Auxiliary Models*

Eric Ghysels†

Alain Guay‡

This version: April 15, 2004

(First draft: February 1999)

Abstract

Several estimation procedures such as the Efficient Method of Moments (EMM) of Gallant and Tauchen (1996) and Indirect Inference procedure of Gouriéroux, Monfort and Renault (1993) involve two models, an auxiliary one and a model of interest. The role played by both models poses challenges and provides new opportunities for hypothesis testing beyond the usual Wald, LM and LR-type tests. In this paper we present and derive the asymptotic distribution theory for various classes of tests for structural change. Some procedures are extensions of standard tests while others are specific to the dual model setup and exploit its unique features.

*The first author gratefully acknowledges financial support of Fonds pour la Formation de Chercheurs et l'aide à la Recherche (FCAR). The second author would like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada through a grant to NCM2 [Network for Computing and Mathematical Modeling]. We would also like to thank Alastair Hall and Éric Renault for comments on an earlier draft of the paper.

†University of North Carolina and CIRANO. *e-mail*: eghysels@unc.edu.

‡Université du Québec à Montréal, CIRPÉE and CIREQ. *e-mail*: guay.alain@uqam.ca.

1 Introduction

There is now a fully developed asymptotic distribution theory for various types of test statistics associated with Generalized Method of Moments (henceforth GMM) and Simulated Method of Moments (henceforth SMM) estimators. The seminal paper by Hansen (1982) on GMM proposed a widely used test for overidentifying restrictions, while Gallant (1987) and Newey and West (1988) presented generic Wald, LM and LR-type tests. Andrews and Ploberger (1994) deal with optimal tests when a nuisance parameter is present only under the alternative. One of the most prominent applications of such test statistics involves the hypothesis of structural change with unknown breakpoint.[1] McFadden (1989), Pakes and Pollard (1989) and Duffie and Singleton (1993) extended the GMM framework to estimation methods involving simulated moments. A comprehensive treatment of Wald, LM and LR-type tests for SMM can be found in Gouriéroux, Monfort and Renault (1993). **Ghysels and Guay (2003) derive optimal structural change tests with unknown breakpoint for simulated method of moments.**

In recent years a number of estimation procedures have been proposed which involve a dual model setup. Examples include Asymptotic Least Squares (henceforth ALS) of Gouriéroux, Monfort and Trognon (1985), the indirect inference method of Gouriéroux et al. (1993) (henceforth I.I.) and the Efficient Method of Moments (called EMM) procedure of Gallant and Tauchen (1996). Estimation procedures involving auxiliary models are more commonly used, particularly in situations where likelihood-based estimation or method of moments are infeasible. Many empirical examples can be found in macroeconomic and financial econometrics literature. These procedures are driven by the fundamental distinction between an auxiliary model, parameterized by a vector θ, and a model of interest, which is parameterized by ρ. The distinct role played by both models can be viewed as adding complications to the formulation of traditional tests and can also be viewed as the basis for formulating new classes of tests. The purpose of our paper is to examine both issues.

We present several classes of tests for structural change, some are extensions of tests proposed for GMM and SMM while others genuinely exploit features unique to the dual model setup. We proceed in two steps. First we ignore the simulation uncertainty and deal with tests for structural change in a GMM-type setup involving an auxiliary model. Such tests are based on the ALS principle. Next we add the simulation uncertainty and present a generic class of tests for structural change with unknown breakpoints for EMM and I.I. estimators. **While Ghysels and Guay (2003) deal with structural change tests for simulated method of moments they do not consider estimation methods involving an auxiliary model as**

[1] It should be noted that various tests for the structural change hypothesis were developed for the GMM estimator; see for instance Andrews and Fair (1988), Dufour, Ghysels and Hall (1994), Ghysels, Guay and Hall (1997), Ghysels and Hall (1990), Guay (2003), Hall and Sen (1999), Hoffman and Pagan (1989), Sowell (1996a), among others.

ALS, I.I. and E.M.M.

Among the tests for structural change specifically tailored for EMM and I.I. figures a class of tests based on a principle of simulated scores which is specific to the combination of an auxiliary model and simulation based estimation. The simulated score tests we propose use simulated series from a restricted null model of interest. Using the reprojection arguments of Gallant and Tauchen (1998) we can fit a sieve seminonparametric SNP density to the simulated data. Under the null the simulated data should yield a reprojection score generator which is a martingale difference sequence when applied to the actual sample data.

Our analysis also relates to a EMM diagnostic test proposed by Liu and Zhang (1998). Their test, while meant to be a simulated score test, is closely related to one of the structural change tests we propose. We generalize and extend the test Liu and Zhang (1998) suggested. Recent work by van der Sluis (1998) also proposes structural change tests for EMM. We show that the asymptotic derivations in van der Sluis are invalid for the proposed statistics and compare our tests with the Hansen J-type and Hall-Sen type tests discussed in van der Sluis (1998).

The paper is organized as follows: In section 2 we discuss tests for structural change with unknown breakpoint. Section 3 deals with simulated score tests. Section 4 covers non-nested hypothesis testing while section 5 concludes.

2 Models and Parameter Estimators

In this section we describe the data generating processes as well as the various classes of estimators we will consider. A first subsection is devoted to the description of the data generating processes. The second subsection covers the parameter estimators.

2.1 The Data Generating Processes

The data generating process is described by a parametric nonlinear simultaneous equations model, namely:

$$r(y_t, y_{t-1}, x_t, u_t, \rho) = 0 \tag{2.1}$$

$$q(u_t, u_{t-1}, \varepsilon_t, \rho) = 0 \tag{2.2}$$

where $\rho \in \Re \subset R^p$, $\{y_t\}$ corresponds to the vector of dependent variables whereas $\{x_t\}$ is the vector of exogenous variables. Both vector processes are stationary and observable, in addition $\{x_t\}$ is a homogeneous

Markov process independent of $\{\varepsilon_t\}$ and $\{u_t\}$. The latter two are latent processes with ε_t white noise with known distribution G_0.[2] The fact that only one lag is considered in (2.1) and (2.2) is not essential and can easily be relaxed.[3] It will also be convenient to define the vector $Z_{t-1} \equiv (y_{t-1}, x_t)$. Equations (2.1) and (2.2) correspond to the data generating processes considered by Gouriéroux et al. (1993), Gallant and Tauchen (1996) and Dridi, Guay and Renault (2003). Since we will be dealing with simulation-based estimators we assume that samples of simulated $\{y_t^s(\rho)\}_{t=1}^T$ can be generated uniquely through (2.1) and (2.2), given ρ and conditional on initial values u_0 and y_0 as well as the observed path of exogenous variables $\{x_t\}_{t=1}^T$. Throughout the paper we assume that all processes are stationary **under the null of no structural change**. Following Andrews and McDermott (1995), one can extend the asymptotic distribution of the estimators and corresponding test statistics to nonlinear models with deterministically trending variables. As Andrews and McDermott show, this would only involve some straightforward modifications to the estimation of covariance matrices.

The indirect inference method of Gouriéroux et al. (1993) and the efficient method of moments of Gallant and Tauchen (1996) are estimation procedures designed for situations where the log-likelihood function of the structural model:

$$\zeta_T(\rho) = \sum_{t=1}^{T} \log p(y_t | Z_{t-1}, \rho) \tag{2.3}$$

is computationally intractable **and where $\{p(y_t|Z_{t-1},\rho)\}_{t=1}^T$ is a sequence of time-invariant conditional densities**. The likelihood-based method is therefore replaced by an instrumental criterion which involves a vector of parameters $\theta \in \Theta \subset R^q$, namely:

$$\Psi_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \psi_t(y_t | Z_{t-1}, \theta). \tag{2.4}$$

Minimizing (2.4) yields an M-estimator $\widehat{\theta}_T$ for θ. The auxiliary model parameters θ and those of the structural model are related through **a system of G-equations:**

$$g(\theta^*, \rho^0) = 0 \tag{2.5}$$

where ρ^0 is the true value of ρ defined for the structural model (2.1) and (2.2) and θ^* called the pseudo-true value is the value which minimizes the limit (as $T \to \infty$) of the M-estimation criterion (2.4). Equation (2.5) yields a so-called binding function $\theta^* = b(\rho^0)$. The I.I. and EMM procedures provide, in different ways, simulation-based approximations to the binding function. Moreover, the function in (2.5) must satisfy:

[2] The assumption of white noise can be relaxed, see Gouriéroux et al. (1993) for further discussion.

[3] In principle an infinite number of lags can be considered, though at a cost of additional regularity conditions, as discussed by Gallant and Tauchen (1996).

Assumption 2.1 *For the purpose of identification, $p \leq G \leq q$ in (2.5), where $\rho \in \Re \subset R^p$, $\theta \in \Theta \subset R^q$, $G_\theta = \partial g(\theta^*, \rho^0)/\partial \theta'$ and $G_\rho = \partial g(\theta^*, \rho^0)/\partial \rho'$ are both of full column rank.*

Finally, it will be useful to split the parameter vector ρ into two subvectors $\rho = (\rho^1,\ \rho^2)$. There are at least two motivating reasons for this. First, following Andrews (1993) one can consider tests for partial structural change where only a subvector ρ^1 of the parameter vector of interest ρ is tested for structural change. Second, following Dridi and Renault (2001) and Dridi et al. (2003) one can also consider situations where only a subvector ρ^1 of ρ is of direct interest while ρ^2 consists of nuisance parameters, such as parameters pertaining to distributional assumptions. Such a situation, which Dridi and Renault (2001) and Dridi and al. (2003) label semiparametric indirect inference, also suggests tests for structural change for subvectors corresponding to parameters of economic interest. Throughout the remainder of this paper we will discuss the implications of partial structural change and semiparametric indirect inference. To keep the notational complexity minimal, we avoid spliting the parameter vector ρ in subvectors. All the results we present can easily be modified to take into account the special cases of testing the null hypothesis of structural change for subvectors.

2.2 Parameter estimators

The Asymptotic Least Squares estimator of Gouriéroux, Monfort and Trognon (1985) is a procedure for estimating ρ through an auxiliary model parameterized by θ. Its main advantage, which we exploit here for expository purpose, is that it does not involve simulation uncertainty. Sidestepping this source of uncertainty, at least at a first stage, allows us to focus first and foremost on the key issue of testing for structural change when an auxiliary model is present.

2.2.1 The Asymptotic Least Squares estimator

We will consider several ALS estimators. In particular, we define the estimator for the entire sample of the parameter vector of the auxiliary model as the following M-estimator:

$$\hat{\theta}_T = \arg\min_{\theta\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \psi_t(y_t|Z_{t-1};\theta) \tag{2.6}$$

where $\theta \in \Theta \subset R^q$. Some tests for structural change involve parameter estimators over subsamples. We will call full sample estimators, like (2.6), as restricted estimators since the parameters are assumed identical across subsamples. To define an unrestricted estimator we consider explicitly two subsamples, the first is based on observations $t = 1, \cdots, [T\pi]$ while the second subsample covers $t = [T\pi] + 1, \cdots, T$ where $\pi \in \Pi \subset (0, 1)$. The separation $[T\pi]$ represents a possible breakpoint and $[\cdot]$ denotes the greatest integer function. The unrestricted asymptotic least squares estimators for the first and the second subsamples are,

$$\hat{\theta}_{1T}(\pi) = \arg\min_{\theta\in\Theta} \frac{1}{[T\pi]} \sum_{t=1}^{[T\pi]} \psi_t(y_t|Z_{t-1};\theta) \tag{2.7}$$

$$\hat{\theta}_{2T}(\pi) = \arg\min_{\theta\in\Theta} \frac{1}{T-[T\pi]} \sum_{t=[T\pi]+1}^{T} \psi_t(y_t|Z_{t-1};\theta). \tag{2.8}$$

The unrestricted least square estimators of the parameter vector ρ for the first and the second subsamples are obtained by:

$$\hat{\rho}_{iT}(\pi) = \arg\min_{\rho\in\Re} g(\hat{\theta}_{iT}(\pi),\rho)' W_{iT}(\pi) g(\hat{\theta}_{iT}(\pi),\rho) \tag{2.9}$$

where $W_{iT}(\pi)$ are positive definite matrices and g is defined in (2.5) and $i = 1, 2$ corresponding to the appropriate subsample. The restricted asymptotic least squares estimator for ρ is obtained via a function relating θ^* to the parameter of interest. This function is defined as $g(\theta^*, \rho^0) = 0$ where $\hat{\theta}$ replaces θ^*, i.e. the estimator which minimizes as $T \to \infty$ the limit of the M-criterion. Hence, the restricted (i.e. full sample) estimator is:

$$\hat{\rho}_T = \arg\min_{\rho\in\Re} g(\hat{\theta}_T,\rho)' W_T g(\hat{\theta}_T,\rho) \tag{2.10}$$

where W_T is a $G \times G$ positive definite matrix.

2.2.2 The Indirect Inference estimator

The indirect inference method of Gouriéroux et al. (1993) also involves the binding function which relates the estimator for the auxiliary model to the estimator of the structural model $\theta^* = b(\rho^0)$. The binding function is unknown, however, and therefore is approximated by simulation. Assume one selects a value of ρ and, using equations (2.1) and (2.2), one simulates the process $\{y_t^s(\rho)\}_{t=1}^T$. The estimator of the auxiliary model is then defined as:

$$\hat{\theta}_T^s(\rho) = \arg\min_{\theta\in\Theta} \frac{1}{T} \sum_{t=1}^{T} \psi_t(y_t^s(\rho)|Z_{t-1}^s(\rho),\theta)) \tag{2.11}$$

where Z_0^s are the initial values (y_{-1}, x_0) for the s simulated path. Note also that $Z_{t-1}^s(\rho) \equiv (y_{t-1}^s(\rho), x_t)$. For S simulated paths, we construct $\frac{1}{S}\sum_{s=1}^S \hat{\theta}_T^s(\rho)$, where $\hat{\theta}_T^s$ is a consistent estimator of the binding function. The indirect estimator of ρ is obtained as the solution of the following minimum distance problem

$$\hat{\rho}_T^S = \arg\min_{\rho\in\Re} \left[\hat{\theta}_T - \frac{1}{S}\sum_{s=1}^S \hat{\theta}_T^s(\rho)\right]' W_T \left[\hat{\theta}_T - \frac{1}{S}\sum_{s=1}^S \hat{\theta}_T^s(\rho)\right]. \tag{2.12}$$

where W_T is a $q \times q$ positive definite matrix.

For certain structural change tests we will need again to define subsample estimators. They are obtained with the auxiliary model for the first and the second subsamples, namely:

$$\hat{\theta}_{1T}^s(\rho,\pi) = \arg\min_{\theta\in\Theta} \frac{1}{[T\pi]} \sum_{t=1}^{[T\pi]} \psi_t(y_t^s(\rho)|Z_{t-1}^s(\rho),\theta) \tag{2.13}$$

and

$$\hat{\theta}^s_{2T}(\rho,\pi) = \arg\min_{\theta\in\Theta} \frac{1}{T-[T\pi]} \sum_{t=[T\pi]+1}^{T} \psi_t(y^s_t(\rho)|Z^s_{t-1}(\rho),\theta) \tag{2.14}$$

Therefore the indirect estimators for the first and the second subsamples are obtained by:

$$\hat{\rho}^S_{iT}(\pi) = \arg\min_{\rho\in\Re} \left[\hat{\theta}_{iT}(\pi) - \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}^s_{iT}(\rho,\pi)\right]' W_{iT}(\pi) \left[\hat{\theta}_{iT}(\pi) - \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}^s_{iT}(\rho,\pi)\right] \tag{2.15}$$

for $i = 1,2$ and $W_{iT}(\pi)$ are positive definite matrices.

To conclude this section we elaborate on the simulation of processes with structural breaks. Suppose the parameters of interest for the two subsamples are ρ_i for $i = 1,2$. Then for the first subsample one generates data based on (2.1) and (2.2), modified accordingly, namely $r(y^s_t, y^s_{t-1}, x_t, u_t, \rho_1) = 0$ and $q(u^s_t, u^s_{t-1}, \varepsilon^s_t, \rho_1) = 0$ for $t = 1,\cdots,[T\pi]$. This is repeated for the second subsample which covers $t = [T\pi]+1,\cdots,T$ with ρ_2 as parameter. Hence, one creates a series $\{y^s_t(\rho_1,\rho_2)\}^T_{t=1} \equiv (\{y^s_t(\rho_1)\}^{[T\pi]}_{t=1}, y^s_t(\rho_2)\}^T_{t=[T\pi]+1})$.

2.2.3 The Efficient Method of Moments estimator

According to the D.G.P. (2.1) and (2.2), Gallant and Tauchen (1996) define what they call the maintained model via the corresponding sequence of time-invariant densities $\{p_1(Z_0|\rho), p(y_t|Z_{t-1},\rho)\}^\infty_{t=1}$, $\rho \in \Re \subset R^p$, whereas the auxiliary model is represented by a sequence of time-invariant densities $\{f_1(Z_0|\theta)$, $\{f(y_t|Z_{t-1},\theta)\}^\infty_{t=1}$, $\theta \in \Theta \subset R^q$. It should be noted that we continue to use Z_{t-1} as the conditional information set. Typically, Gallant and Tauchen consider densities conditional on y_{t-1}. However, in some circumstances Z_{t-1} contains only x_t, as for instance is the case with reprojection schemes, see Gallant and Tauchen (1998). For the sake of simplicity we will keep the conditioning set as Z_{t-1} and it will be obvious from the context what the conditional information set is. The following assumption introduced by Gallant and Tauchen (1996) is used for the validity of the EMM criterion as a specification test for the maintained model.

Assumption 2.2 *The maintained model* $\{p(Z_0|\rho), p(y_t|Z_{t-1},\rho)\}^\infty_{t=1}$ $\rho \in \Re$ *is smoothly embedded within the auxiliary model* $\{f(Z_0|\theta), f(y_t|Z_{t-1},\theta)\}^\infty_{t=1}$ $\theta \in \Theta$, *i.e. for some open neighborhood* $\Re^0 \to \Theta$, *it is such that:* $p\left(y_t|\, Z_{t-1},\rho\right) = f\left[y_t|\, Z_{t-1}, b(\rho)\right], t = 1,2,\ldots$ *for every* $\rho \in \Re^0$ *and* $p\left(Z_0|\,\rho\right) = f(Z_0|b(\rho))$ *for every* $\rho \in \Re^0$.

Under this embedding assumption, the parameters of the auxiliary model (θ^*) are related to the parameters of the maintained model (ρ^0) according to $\theta^* = b(\rho^0)$. Assumption 2.2 is comparable to Assumption 2.1, both play the same role guaranteeing identification of ρ via the auxiliary model. However, Assumption 2.2 is stronger than Assumption 2.1. Indeed, under Assumption 2.2 the E.M.M. estimator is fully efficient.

The EMM estimator is obtained in two steps. The first step is to compute the (pseudo) maximum likelihood estimate of the auxiliary model:

$$\hat{\theta}_T = \arg\max_{\theta\in\Theta} \frac{1}{T}\sum_{t=1}^{T} log[f\left(y_t|\, Z_{t-1},\theta\right)], \tag{2.16}$$

and the corresponding estimate of the information matrix:

$$I_T = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{\partial}{\partial\theta}\log f\left(y_t|\, Z_{t-1},\hat{\theta}_T\right)\right]\left[\frac{\partial}{\partial\theta}\log f\left(y_t|\, Z_{t-1},\hat{\theta}_T\right)\right]'.^4 \tag{2.17}$$

In the second step, a vector of moment conditions is constructed using the expectation under the maintained model of the scores from the auxiliary model. The EMM estimator is obtained by minimizing a GMM criterion function formed by the above moment conditions, i.e.,

$$\hat{\rho}_T^S = \arg\min_{\rho\in\Re} m_T^S(\rho,\hat{\theta}_T)'(I_T)^{-1}m_T^S(\rho,\hat{\theta}_T) \tag{2.18}$$

where

$$m_T^S(\rho,\theta) = \frac{1}{TS}\sum_{t=1}^{TS}\frac{\partial}{\partial\theta}\log[f(y_t^s(\rho)|Z_{t-1}^s(\rho),\theta)] \tag{2.19}$$

and $y_t^s(\rho), Z_{t-1}^s(\rho)_{t=1}^{TS}$ is a long series of realizations simulated from the maintained model with the parameter vector ρ. Under suitable regularity conditions discussed in Gallant and Tauchen (1996) and Assumption 2.2, we have $\sqrt{T}(\hat{\theta}_T-\theta^*)\xrightarrow{d} N\{0,I^{-1}\}$ and, $\sqrt{T}(\hat{\rho_T}-\rho^0)\xrightarrow{d} N\{0,[M_\rho' I^{-1}M_\rho]^{-1}\}$ where $M_\rho = (\partial/\,\partial\rho')m(\rho^0,\theta^*)$ and I is the outer product of scores, as suggested by the estimator in (2.17). All these results apply to the case where the number of simulations goes to infinity. In the case of possible structural changes with unknown breakpoint, theoretical results based of the number of simulations equal to infinity are not so appealing as the computational cost involved can be prohibitively high. For this reason, the asymptotic results need to be modified to account for a finite number of simulations. When S is finite, the randomness of the EMM estimator $\hat{\rho}_T$ will not only depend on the randomness of $\hat{\theta}_T$ but also on the randomness of the moment conditions due to a finite length of series simulated from the structural model. Therefore the asymptotic variance-covariance matrix in equation (2.17) is scaled by $(1+1/S)$ using arguments similar to Duffie and Singleton (1993).

To conclude this section we present partial sample estimators which appear in certain tests for structural change. The unrestricted EMM estimator for the subsamples are defined as:

$$\hat{\rho}_{iT}^S(\pi) = \arg\min_{\rho\in\Re} m_i'(\rho,\hat{\theta}_{iT}(\pi))(I_{iT})^{-1}m_i(\rho,\hat{\theta}_{iT}(\pi)) \tag{2.20}$$

[4]See Gallant and Tauchen (1996) for alternative consistent estimators of I.

where $\rho_i \in \Re \subset R^p$, I_{iT} is the estimator of the matrix I for the i^{th} subsample and:

$$\hat{\theta}_{1T}(\pi) = \arg\max_{\theta\in\Theta} \frac{1}{[T\pi]} \sum_{t=1}^{[T\pi]} log[f\left(y_t|\, Z_{t-1},\theta\right)], \tag{2.21}$$

$$\hat{\theta}_{2T}(\pi) = \arg\max_{\theta\in\Theta} \frac{1}{T-[T\pi]} \sum_{t=[T\pi]+1}^{T} log[f\left(y_t|\, Z_{t-1},\theta\right)], \tag{2.22}$$

$$m_1(\rho,\hat{\theta}_{1T}(\pi)) = \frac{1}{[TS\pi]} \sum_{t=1}^{[TS\pi]} \frac{\partial}{\partial\theta} \log f[y_t^s(\rho)|Z_{t-1}^s(\rho),\hat{\theta}_1(\pi))]. \tag{2.23}$$

$$m_2(\rho,\hat{\theta}_{2T}(\pi)) = \frac{1}{TS-[TS\pi]} \sum_{t=[TS\pi]+1}^{TS} \frac{\partial}{\partial\theta} \log f[y_t^s(\rho)|Z_{t-1}^s(\rho),\hat{\theta}_2(\pi)]. \tag{2.24}$$

The simulation of processes when a break is present can be characterized by the following sequence of densities $\{p(Z_0|\rho_1), p(y_t|Z_{t-1},\rho_1)\}_{t=1}^{[TS\pi]}$, for the first subsample and $\{p(y_t|Z_{t-1},\rho_2)\}_{t=[TS\pi]+1}^{TS}$, for the second. It is important to note that the simulated path length is function of the fraction of the sample (π). This point is crucial. Indeed, The asymptotic distribution of several structural change tests could depend on the nuisance parameter S and hence the critical values depend on S, in the case where the simulated path length is not split according to the presumed breakpoint π. Section 3.2 will examine this problem.

3 GMM-like Tests for Structural change with unknown breakpoint

The purpose of this section is to generalize GMM-based tests for structural change presented by Andrews (1993), Andrews and Ploberger (1994), Sowell (1996a,b) and Ghysels, Guay and Hall (1997). A variety of tests were proposed ranging from (optimal) Wald, LM and LR-type tests to predictive tests with unknown breakpoint. In this section we deal with the issues posed by procedures involving two models, an auxiliary one and a model of interest. We noted that the role played by both models poses challenges and provides new opportunities for hypothesis testing. Here we only deal with the usual Wald, LM and LR-type and predictive tests. In the next section we cover tests which are specifically designed for the dual model setup. We cover tests based on ALS, Indirect Inference and EMM estimators. One of the first issues to resolve is to clearly define the null hypothesis of interest in tests for structural change analysis. The analysis in the first subsection involves the ALS since it allows us again to focus directly on the key issues of hypotheses and test statistics. The added complication of simulation uncertainty is considered in the second subsection.

3.1 Tests for Asymptotic Least Squares

The purpose of this section is twofold: (1) clearly spell out the null hypotheses involved in tests for structural change when an auxiliary model is present and (2) adapt the usual Wald, LM and LR-type and predictive tests for such situations. A subsection is devoted to each of the two issues.

3.1.1 The Null Hypotheses

The null hypothesis of interest for structural parameters is:

$$H_0^\rho : \rho_t = \rho^0 \qquad \forall t = 1, ..., T. \tag{3.1}$$

The fact that we estimate the parameter vector ρ indirectly via an auxiliary model implies that we also should consider the null hypothesis:

$$H_0^\theta : \theta_t = \theta^* \qquad \forall t = 1, ..., T. \tag{3.2}$$

The null hypotheses (3.1) and (3.2), while related, are obviously not identical. Accepting H_0^θ implies that there is no structural change for ρ because of the identification Assumption 2.1 for the binding function. Rejecting H_0^θ does not necessarily imply that H_0^ρ is violated since the dimension of θ is equal or greater than the dimension of ρ. To unravel whether the rejection of H_0^θ is due to a structural change of the overidentifying restrictions, one can follow the approach of Sowell (1996b) and characterize via projection the subspace which identifies ρ. Such projection can distinguish structural change of the structural parameters from breaks in the overidentifying restrictions. This distinction becomes even more interesting when we allow for partial structural change, i.e. consider subvectors of ρ. In particular, in the context of semiparametric indirect inference, following Dridi et al. (2003), this may involve a subvector of nuissance parameters ρ^2 for which structural change may be more tolerated.

To elaborate further on the distinction between the null hypotheses (3.1) and (3.2), and in particular the interpretation of rejecting the null hypotheses, we consider a sequence of local of alternatives:

$$\theta_{t,T} = \theta^* + h(\eta, \nu, \frac{t}{T})/\sqrt{T} \tag{3.3}$$

where $h(\eta, \nu, \pi)$, for $\pi \in [0, 1]$, is a q-dimensional function which can be expressed as the uniform limit of step functions, $\eta \in R^i$, $\nu \in R^j$ such that $0 < \nu_1 < \nu_2 < \ldots < \nu_j < 1$ and θ^* is in the interior of Θ. The function $h(\cdot)$ allows for a wide range of alternative hypotheses (see Sowell (1996b)). The parameter ν locates structural changes as a fraction of the sample size and the vector η defines the local alternatives. To simplify the notation $h(\eta, \nu, \frac{t}{T})$ will be denoted $h(\nu)$. The following theorem provides the asymptotic distribution for the optimally weighted $g(\cdot)$ for both subsamples, using $W_T = \Omega_T^{-1}$[5], where Ω_T is the full sample estimator of the optimal weighting matrix Ω which is defined in Appendix E:

Theorem 3.1 *Under Assumptions 2.1, A.1, B.1 and sequence of local alternatives (3.3), we have:*

$$\begin{aligned}\pi\sqrt{T}\Omega_T^{-1/2}g(\hat{\theta}_{1T}(\pi),\hat{\rho}_T) \quad &\Rightarrow \quad -\left[B(\pi) - \Omega^{-1/2}G_\theta H(\pi)\right] + \\ &\pi\Omega^{-1/2}G_\rho(G_\rho'\Omega^{-1}G_\rho)^{-1}G_\rho'\Omega^{-1/2}\left[B(1) - \Omega^{-1/2}G_\theta H(1)\right].\end{aligned}$$

$$\begin{aligned}(1-\pi)\sqrt{T}\Omega_T^{-1/2}g(\hat{\theta}_{2T}(\pi),\hat{\rho}_T) \quad &\Rightarrow \quad -\left[B(1) - B(\pi) - \Omega^{-1/2}G_\theta\left(H(1) - H(\pi)\right)\right] + \\ &(1-\pi)\Omega^{-1/2}G_\rho(G_\rho'\Omega^{-1}G_\rho)^{-1}G_\rho'\Omega^{-1/2}\left[B(1) - \Omega^{-1/2}G_\theta H(1)\right].\end{aligned}$$

where $H(\pi) = \int_0^\pi h(\eta,\nu,r)dr$*,* $B(\pi)$ *is a G-dimensional vectors of independent Brownian motions.*

Proof: See Appendix E

Under the null hypothesis (3.2), a version of Corollary 1 of Sowell(1996a) holds, namely there exists an orthonormal matrix C such that

$$\pi C\sqrt{T}\Omega_T^{-1/2}g(\hat{\theta}_{1T}(\pi),\hat{\rho}_T) \quad \Rightarrow \quad \begin{bmatrix} -BB_p(\pi) \\ -B_{G-p}(\pi) \end{bmatrix}. \tag{3.4}$$

where $BB_p(\pi)$ is a p-dimensional Brownian bridge, $B_{G-p}(\pi)$ is a $G-p$-dimensional Brownian motion, C is such that $\Omega^{-1/2}G_\rho(G_\rho'\Omega^{-1}G_\rho)^{-1}G_\rho'\Omega^{-1/2} = C'\Lambda C, CC' = Id$ where Id is the identity matrix and

$$\Lambda = \begin{bmatrix} Id_p & 0_{p\times(G-p)} \\ 0_{(G-p)\times p} & 0_{(G-p)\times(G-p)} \end{bmatrix}.$$

For the function $g(\cdot)$ evaluated at the estimator obtained from the second subsample, we have

$$(1-\pi)C\sqrt{T}\Omega_T^{-1/2}g(\hat{\theta}_{2T}(\pi),\hat{\rho}_T) \quad \Rightarrow \quad \begin{bmatrix} BB_p(\pi) \\ -B^*_{G-p}(\pi) \end{bmatrix} \tag{3.5}$$

where $BB_G(\pi)$ is defined above and $B^*_{G-p}(\pi) = B_{G-p}(1) - B_{G-p}(\pi)$.

As shown by Sowell (1996b), structural change tests can be constructed in projecting on the appropriate subspace. The limiting stochastic processes in (3.4) and (3.5) are equivalent to the limiting stochastic processes for the GMM estimator in Sowell or those obtained for the Simulated Method of Moments estimator in Ghysels and Guay (2003). Under the null hypotheses (3.1) and (3.2), the results in (3.4) and (3.5) show that the limiting continuous stochastic processes are linear combinations of p Brownian bridges, one for each

[5] This expression may exist only with probability going to one. When this expression is singular, a g-inverse can be used in place of the inverse (see Andrews (1993)). Similar comments apply elsewhere below

parameter estimated, and $G-p$ Brownian motions, spanning the space of *overidentifying restrictions*, where G is the dimension of $g(\cdot)$.

We can refine now the null hypothesis (3.1). In particular, following Hall and Sen (1999) we consider the generic null, for the case of a single breakpoint, which seperates the identifying restrictions across the two subsamples:

$$H_0^{I\rho}(\pi) = \begin{cases} P_G' \Omega^{-1/2} g(\theta^*, \rho^0) = 0 & \forall t = 1, \ldots, [\pi T] \\ P_G \Omega^{-1/2} g(\theta^*, \rho^0) = 0 & \forall t = [\pi T] + 1, \ldots, T \end{cases}$$

where $P_G = \Omega^{-1/2} G_\rho (G_\rho' \Omega^{-1} G_\rho)^{-1} G_\rho' \Omega^{-1/2}$. Moreover, the overidentifying restrictions are stable if they hold before and after the breakpoint. This is formally stated as $H_0^{O-g}(\pi) = H_0^{O-g1}(\pi) \cap H_0^{O-g2}(\pi)$ with:

$$\begin{aligned} H_0^{O-g1}(\pi) : (Id_G - P_G)\Omega^{-1/2} g(\theta^*, \rho^0) &= 0 \qquad \forall t = 1, \ldots, [\pi T] \\ H_0^{O-g2}(\pi) : (Id_G - P_G)\Omega^{-1/2} g(\theta^*, \rho^0) &= 0 \qquad \forall t = [\pi T] + 1, \ldots, T \end{aligned}$$

Using the projection applied to the decomposition appearing in (3.4) and (3.5), it is clear that instability must be reflected in a violation of at least one of the three hypotheses: $H_0^{I\rho}(\pi)$, $H_0^{O-g1}(\pi)$, or $H_0^{O-g2}(\pi)$. It is only the former of those three which corresponds to the null hypothesis (3.1). Violation of $H_0^{O-g1}(\pi)$, or $H_0^{O-g2}(\pi)$ mean that there are reasons to reject the null hypothesis H_0^θ in (3.2), but still accept H_0^ρ in (3.1). Various tests can be constructed with local power properties against any particular one of these null hypotheses (and typically no power against the other two).

To conclude we need to discuss the implication of various structural change tests in presence of auxiliary models. The decomposition of the hypothesis (and associated tests) into $H_0^{I\rho}(\pi)$, $H_0^{O-g1}(\pi)$, or $H_0^{O-g2}(\pi)$ has different implications for the structural model. The auxiliary model can be viewed as a window through which information is obtained about the structural model. Consequently, structural change can only be assessed via the information about the structural model revealed by the auxiliary model. For example, Guay and Renault (2003) examine indirect encompassing when both models are misspecified and estimated by auxiliary models. In the first step of their proposed procedure, the auxiliary model is used only to obtain consistent estimators of structural model parameters. Structural parameter instability detected through the intermediary of the auxiliary model is crucial for the consistency of the procedure. However, instability of the overidentifying restrictions (of the auxiliary model) without change of the structural parameters is innocuous.

3.1.2 Test statistics

A structural change test is obtained for the vector of parameters ρ when the function $\Omega_T^{-1/2} g(\cdot)$ is projected on the subspace identifying the parameters with the first subsample estimator $\hat{\theta}_{1T}$. This statistic is

$$\begin{aligned} Q_{1T}(\pi) &= T\pi^2 g(\hat{\theta}_{1T}(\pi), \hat{\rho}_T)' \Omega_T^{-1/2} \left(\Omega_T^{-1/2} G_{\rho,T} (G'_{\rho,T} \Omega_T^{-1} G_{\rho,T})^{-1} G'_{\rho,T} \Omega_T^{-1/2} \right) \\ &\times \Omega_T^{-1/2} g(\hat{\theta}_{1T}(\pi), \hat{\rho}_T). \end{aligned} \tag{3.6}$$

where $G_{\rho,T}$ is a consistent estimator of G_ρ. The statistic with the estimator of the second subsample is:

$$\begin{aligned} Q_{2T}(\pi) &= T(1-\pi)^2 g(\hat{\theta}_{2T}(\pi), \hat{\rho}_T)' \Omega_T^{-1/2} \left(\Omega_T^{-1/2} G_{\rho,T} (G'_{\rho,T} \Omega_T^{-1} G_{\rho,T})^{-1} G'_{\rho,T} \Omega_T^{-1/2} \right) \\ &\times \Omega_T^{-1/2} g(\hat{\theta}_{2T}(\pi), \hat{\rho}_T). \end{aligned} \tag{3.7}$$

A structural change test for overidentifying restrictions is obtained when projecting the function $\Omega_T^{-1/2} g(\cdot)$ on the subspace orthogonal to the subspace identifying the parameters. For example, the statistic with the first subsample estimator is

$$\begin{aligned} Q^0_{1T}(\pi) &= T\pi^2 g(\hat{\theta}_{1T}(\pi), \hat{\rho}_T)' \Omega_T^{-1/2} \left(Id - \Omega_T^{-1/2} G_{\rho,T} (G'_{\rho,T} \Omega_T^{-1} G_{\rho,T})^{-1} G'_{\rho,T} \Omega_T^{-1/2} \right) \\ &\times \Omega_T^{-1/2} g(\hat{\theta}_{1T}(\pi), \hat{\rho}_T) \end{aligned} \tag{3.8}$$

In the case of unknown breakpoint, statistics can be constructed by mapping on $\pi \in \Pi$. Andrews and Ploberger (1994) in the context of maximum likelihood estimation and Sowell (1996a,b) for GMM estimation derive optimal tests which are characterized by an average exponential mapping. In the case of a one time structural break alternative and a particular integral weight functions for $h(\cdot)$ in (3.3), the tests with the greatest weighted average asymptotic power have the following form for structural parameter instability

$$Exp - Q_{iT} = (1+c)^{(-p/2)} \int exp\left(\frac{1}{2}\frac{c}{1+c} Q_{iT}(\pi)\right) dR(\pi)$$

where $R(\pi)$ is the weight function over the set of possible breakpoints Π. The parameter c controls the distance of the alternative. For close alternatives $c \to 0$, the asymptotic test with the greatest weighted average power is an average (ave) over $\pi \in \Pi$ and has the form: $\int Q_{iT}(\pi) dR(\pi)$. For a distant alternative $c \to \infty$, the functional is: $log \int exp\left(\frac{1}{2} Q_{iT}(\pi)\right) dR(\pi)$. The supremum form $\sup_{\pi \in \Pi} Q_{it}(\pi)$ often used in the litterature corresponds to the case where $c/(1+c) \to \infty$. The LM (or $LM_T(\pi)$ for given π) test statistic of structural change corresponds to the case where $R(\pi) = 1/\left(\pi(1-\pi)\right) d\pi$. A Wald ($Wald_T(\pi)$) and a LR-type ($LR_T(\pi)$) test statistics can be constructed as usual with the restricted and unrestricted ALS estimators. Following Andrews (1993), we can show that $Wald_T(\pi) = LM_T(\pi) + o_p(1)$ and $LR_T(\pi) = LM_T(\pi) + o_p(1)$.

The following proposition gives the asymptotic distribution for the exponential mapping for Q_{iT} when Q_{iT} corresponds to the Wald, LM and LR ratio-type tests.

Proposition 3.1 *Under the null hypothesis H_0 in (3.1) and Assumptions 2.1, A.1, B.1, the following processes indexed by π for a given set Π whose closure lies in (0,1) satisfy:*

$$\sup Q_{iT} \Rightarrow \sup_{\pi\in\Pi} Q_p(\pi), \quad aveQ_{iT} \Rightarrow \int_\Pi Q_p(\pi)dR(\pi), \quad expQ_{iT} \Rightarrow log\left(\int_\Pi \exp[\frac{1}{2}Q_p(\pi)]dR(\pi)\right),$$

with

$$Q_p(\pi) = BB_p(\pi)'BB_p(\pi)$$

for $i = 1, 2$.

This result is obtained through the application of the continuous mapping theorem (see Pollard (1984)). The asymptotic distribution is a quadratic form of weighted Brownian bridge such as when the breakpoint is known the asymptotic distribution is a chi-square with a degree of freedom equal to the dimension of the structural vector parameters. **In the case of an unknown breakpoint, critical values are given in Andrews (1993) for a weighting equals to $1/(\pi(1-\pi))$.**

The next proposition gives the asymptotic distribution for the exponential mapping for Q^0_{iT} when Q^0_{iT} is the statistic for the structural change in overidentifying restrictions corresponding to the null $H_0^{O-gi}(\pi)$.

Proposition 3.2 *Under the null hypothesis of no structural change for the over-identifying restrictions and Assumptions 2.1, A.1, B.1, the following processes indexed by π for a given set Π whose closure lies in (0,1) satisfy:*

$$\sup Q^0_{iT} \Rightarrow \sup_{\pi\in\Pi} Q_{i,G-p}(\pi),$$

$$aveQ^0_{iT} \Rightarrow \int_\Pi Q_{i,G-p}(\pi)dR(\pi),$$

$$expQ^0_{iT} \Rightarrow log\left(\int_\Pi \exp[\frac{1}{2}Q_{i,G-p}(\pi)]dR(\pi)\right),$$

with $Q_{1,G-p}(\pi) = B_{G-p}(\pi)'B_{G-p}(\pi)$ and $Q_{2,G-p}(\pi) = B^*_{G-p}(\pi)'B^*_{G-p}(\pi)$, where $B_{G-p}(\pi)$ is a $G-p$-dimensional vector of independent Brownian motion, $B^*_{G-p}(\pi) = B_{G-p}(1) - B_{G-p}(\pi)$ and $i = 1, 2$.

The asymptotic distribution is a quadratic form of Brownian motion such as when the breakpoint is known the asymptotic distribution is a chi-square with a degree of freedom equal to $G-p$. **Critical values for $Q_{1,G-p}(\pi)$ and $Q_{2,G-p}(\pi)$ with respective weighting equals to $1/\pi$ and $1/(\pi(1-\pi))$ are tabulated in Guay (2003).** Predictive tests, discussed in Ghysels, Guay and Hall (1997) and Guay (2003), can also be constructed and the asymptotic distribution of those tests can be easily obtained from Theorem 3.1.

3.2 Structural Change for Indirect Inference and EMM

The analysis in Section 3.1 involves the Asymptotic Least Squares estimator of Gouriéroux, Monfort and Trognon (1985), a procedure which we have chosen to discuss first as it features the estimation of ρ through an auxiliary model parameterized by θ. We now turn our attention to procedures with similar features, but which require simulations to obtain the binding function appearing in (2.5). Sidestepping simulation uncertainty allowed us to focus exclusively on the key issue of testing for structural change when an auxiliary model is present. The results in Ghysels and Guay (2003) may help us to understand the effect of simulation uncertainty on tests for structural change. They propose a set of tests for structural change in models estimates via Simulated Method of Moments (see Duffie and Singleton (1993)) and show that the number of simulations does not affect the asymptotic distribution nor the asymptotic local power of tests for structural change. Hence, the asymptotic results obtained for GMM-based tests are also valid for SMM-based procedures. The intuition for this result is that in the case of tests for structural change one compares parameter estimates that are subject to the same simulation uncertainty (unlike tests of a fixed hypothesis where the distance of the estimates to the null depends on the simulation uncertainty). **Ghysels and Guay (2003) show that the asymptotic distribution under the local alternatives depends on the number of simulations.** Nontheless, they also shown by a Monte Carlo investigation that a relatively small number of simulations suffices to obtain tests with desirable small sample size and power properties.

The purpose of this section is to extend the results of Ghysels and Guay (2003). In particular, it will be shown that for both the I.I. and EMM estimators, the simulation uncertainty does not affect the asymptotic distribution of tests for structural stability. Hence, there is an asymptotic analogue between ALS-based tests and I.I. or EMM-based procedures. We begin with the Indirect Inference procedure to show that this is indeed the case.

Theorem 3.2 *For the full and the partial sample indirect inference estimators appearing in (2.13), (2.14), under Assumptions 2.1, A.1 and C.1 and sequence of local alternatives (3.3), we have*

$$\begin{aligned}\pi\sqrt{T}\Omega_T^{-1/2}\left[\hat{\theta}_{1T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_{1T}^{s}(\hat{\rho}_T^s,\pi)\right] \;\Rightarrow\; & -\left[B(\pi)-\frac{1}{S}\sum_{s=1}^{S}B^s(\pi)-\Omega^{-1/2}H(\pi)\right]\\ & +\pi\Omega^{-1/2}b_\rho(\rho^0)\left[b_\rho'(\rho^0)\Omega^{-1}b_\rho(\rho^0)\right]^{-1}b_\rho(\rho^0)'\Omega^{-1/2}\times\\ & \left[B(1)-\frac{1}{S}\sum_{s=1}^{S}B^s(1)-\Omega^{-1/2}H(1)\right]\end{aligned}$$

and for the second subsample

$$(1-\pi)\sqrt{T}\Omega_T^{-1/2}\left[\hat{\theta}_{2T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_{2T}^{s}(\hat{\rho}_T^s,\pi)\right]\Rightarrow-\left[B(1)-B(\pi)-\frac{1}{S}\sum_{s=1}^{S}(B(1)^s-B(\pi)^s)\right]$$

$$+\Omega^{-1/2}\left(H(1)-H(\pi)\right)+(1-\pi)\Omega^{-1/2}b_\rho(\rho^0)\left[b_\rho(\rho^0)'\Omega^{-1}b_\rho(\rho^0)\right]^{-1}\times$$
$$b'_\rho(\rho^0)\Omega^{-1/2}\left[B(1)-\frac{1}{S}\sum_{s=1}^{S}B^s(1)-\Omega^{-1/2}H(1)\right]$$

where $H(\pi)=\int_0^\pi h(\eta,\nu,r)dr$ and $B(\pi)$ and $B^s(\pi)$ are two q-dimensional vectors of mutually independent Brownian motions and Ω_T is a consistent estimator of $\Omega = J^{-1}IJ^{-1}$.

Proof: See Appendix E

Under the null hypothesis, by replacing Ω_T by $\tilde{\Omega}_T=(1+\frac{1}{S})\Omega_T$ in the expressions of Theorem 3.2, a version of Corollary 1 of Sowell(1996a) can be easily shown such as

$$\pi C\sqrt{T}\tilde{\Omega}_T^{-1/2}\left[\hat{\theta}_{1T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\theta_{1T}^s(\rho_T^s,\pi)\right]\Rightarrow\begin{bmatrix}-BB_p(\pi)\\-B_{q-p}(\pi)\end{bmatrix}.$$

For the second subsample, we have

$$(1-\pi)C\sqrt{T}\tilde{\Omega}_T^{-1/2}\left[\hat{\theta}_{2T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\theta_{2T}^s(\rho_T^s,\pi)\right]\Rightarrow\begin{bmatrix}BB_p(\pi)\\-B^*_{q-p}(\pi)\end{bmatrix}.$$

where $B^*_{q-p}(\pi)=B_{q-p}(1)-B_{q-p}(\pi)$. To obtained these results, we note that

$$\left(1+\frac{1}{S}\right)^{-1/2}\left[B(\pi)-\frac{1}{S}\sum_{s=1}^{S}B^s(\pi)\right]$$

is a q-dimensional vector of standard Brownian motion. As shown in Section 3.1.2, structural change tests can be constructed by projection on the appropriate subspace. A structural change test is obtained for the vector of parameters ρ when the difference between the estimator obtained with the auxiliary model for the data and the average estimators obtained with simulated paths is projected on the subspace identifying the parameters for the first or the second subsample. The statistic based on the first subsample estimator is

$$\begin{aligned}Q_{1,T}^S(\pi) &= T\pi^2\left[\hat{\theta}_{1T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\theta_{1T}^s(\rho_T^s)\right]'\tilde{\Omega}_T^{-1/2}\left(\tilde{\Omega}_T^{-1/2}b_{\rho,T}\left[b'_{\rho,T}\tilde{\Omega}_T^{-1}b_{\rho,T}\right]^{-1}b'_{\rho,T}\tilde{\Omega}_T^{-1/2}\right)\times\\ &\quad\tilde{\Omega}_T^{-1/2}\left[\hat{\theta}_{1T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\theta_{1T}^s(\rho_T^s)\right]\end{aligned}$$

where $b_{\rho,T}$ is a consistent estimator of $b_\rho(\rho^0)$[6]. A structural change tests for overidentifying restrictions is obtained by projecting the same function on the subspace orthogonal to the subspace identifying the parameters. The resulting statistic based on the first subsample estimator is:

$$\begin{aligned}Q_{1,T}^{S,0}(\pi) &= T\pi^2\left[\hat{\theta}_{1T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\theta_{1T}^s(\rho_T^s)\right]'\tilde{\Omega}_T^{-1/2}\left(Id-\tilde{\Omega}_T^{-1/2}b_{\rho,T}\left[b'_{\rho,T}\tilde{\Omega}_T^{-1}b_{\rho,T}\right]^{-1}b'_{\rho,T}\tilde{\Omega}_T^{-1/2}\right)\times\\ &\quad\tilde{\Omega}_T^{-1/2}\left[\hat{\theta}_{1T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}\theta_{1T}^s(\rho_T^s)\right].\end{aligned}$$

[6]See Gouriéroux et al. (1993).

The asymptotic distribution under the null of the exponential mappings of these statistics is given in Proposition 3.1 for the parameter stability and in Proposition 3.2 for stability of overidentifying restrictions.

Using Theorem 3.2, we show that simulation uncertainty does not affect the asymptotic distribution of tests for structural stability under the null. Hence, the implication of structural change detected in the auxiliary model has the same interpretation as in the ALS case. In Section 3.1.1 it was noted that the importance of instability in the auxiliary model for the structural model depends on which hypotheses is violated, namely $H_0^{I\rho}(\pi)$, $H_0^{O-g1}(\pi)$, or $H_0^{O-g2}(\pi)$. An interesting case to examine can be found in Dridi and Renault (2001) who develop a generalization of Indirect Inference to semi-parametric settings. Their approach produces a theory of robust estimation despite misspecifications of the structural model. Suppose economic theory only provides information about a subvector of parameters of interest ρ^1 (in our notation) and direct estimation of the structural model can not be performed so that the econometrician relies on Indirect Inference. To simulated the structural model, an additional nuisance parameter vector ρ^2 is required. Since we are only interested in a consistent estimator for ρ^1, the importance of finding structural change in the auxiliary model depends on the impact of instability on the parameter vector of interest. In particular, instability for the nuisance parameters ρ^2 or for overidentifying restrictions without affecting stability of the parameter vector of interest have no impact on the consistency of the semi-parametric Indirect Inference estimator. Structural change tests must therefore focus on the parameter vector of interest ρ^1. With our results, partial structural stability tests for this parameter vector of interest can be constructed and the asymptotic distribution of exponential mappings is given in Section 3.1.2.

Next, we turn to the EMM estimator, in particular:

Theorem 3.3 *For the partial sample Efficient Method of Moments estimators appearing in (2.18), (2.23), (2.24), under Assumptions A.1,C.1, and D.1 and sequence of the local alternatives (3.3), we have*

$$\pi\sqrt{T}I_T^{-1/2}m_1(\hat{\rho}_T^S,\hat{\theta}_{1T}(\pi)) \Rightarrow \left[B(\pi)-\frac{1}{\sqrt{S}}B^s(\pi)-I^{-1/2}JH(\pi)\right]$$
$$-\pi I^{-1/2}M_\rho\left[M_\rho' I^{-1}M_\rho\right]^{-1}M_\rho' I^{-1/2}\left[B(1)-\frac{1}{\sqrt{S}}B^s(1)-I^{-1/2}JH(1)\right]$$

and for the second subsample

$$(1-\pi)\sqrt{T}I_T^{-1/2}m_2(\hat{\rho}_T^S,\hat{\theta}_{2T}(\pi)) \Rightarrow \left[B(1)-B(\pi)-\frac{1}{\sqrt{S}}\left(B(1)^s-B(\pi)^s\right)-I^{-1/2}J\left(H(1)-H(\pi)\right)\right]$$
$$-\pi I^{-1/2}M_\rho\left[M_\rho' I^{-1}M_\rho\right]^{-1}M_\rho' I^{-1/2}\left[B(1)-\frac{1}{\sqrt{S}}B^s(1)-I^{-1/2}JH(1)\right]$$

where $H(\pi)=\int_0^\pi h(\eta,\nu,r)dr$ *and* $B(\pi)$ *and* $B^s(\pi)$ *are two q-dimensional vectors of mutually independent Brownian motions.*

Proof: See Appendix E

Structural change tests can be constructed by replacing I_T by $\tilde{I}_T = (1+\frac{1}{S})I_T$ as shown in the previous sections. These statistics based on the first subsample estimator are:

$$Q^S_{1,T}(\pi) = T\pi^2 m_1(\hat{\rho}^S_T, \hat{\theta}_{1T}(\pi))'\tilde{I}_T^{-1/2}\left(\tilde{I}^{-1/2}M_{\rho,T}\left[M'_{\rho,T}\tilde{I}^{-1}M_{\rho,T}\right]^{-1}M'_{\rho,T}\tilde{I}^{-1/2}\right)\tilde{I}_T^{-1/2}m_1(\hat{\rho}^S_T, \hat{\theta}_{1T}(\pi))$$

and

$$Q^{S,0}_{1,T}(\pi) = T\pi^2 m_1(\hat{\rho}^S_T, \hat{\theta}_{1T}(\pi))'\tilde{I}_T^{-1/2}\left(Id - \tilde{I}^{-1/2}M_{\rho,T}\left[M'_{\rho,T}\tilde{I}^{-1}M_{\rho,T}\right]^{-1}M'_{\rho,T}\tilde{I}^{-1/2}\right)\tilde{I}_T^{-1/2}m_1(\hat{\rho}^S_T, \hat{\theta}_{1T}(\pi))$$

where $M_{\rho,T}$ is a consistent estimator of M_ρ.

The asymptotic distribution of the resulting tests are the same than the one derived above by noting that $(1+1/S)^{-1/2}\left[B(\pi) - \frac{1}{\sqrt{S}}B^s(\pi)\right]$ is a q-dimensional vector of Brownian motions. The asymptotic distributions under the null are then given in Proposition 3.1 for the parameter stability and in Proposition 3.2 for stability of overidentifying restrictions.

van der Sluis (1998) proposes similar structural change tests for EMM. However, in contrast to our strategy, the length of the simulated series used to construct the structural change tests in van der Sluis (1998) is the same for the estimation of the full sample estimator of ρ and for the evaluation of the moment restrictions with the unrestricted estimators $\hat{\theta}_{iT}(\pi)$. Such a strategy has an important impact on the asymptotic distribution as will be shown in the remaining of this section. Suppose that the length of the simulated series is equal to TS. The statistic proposes by van der Sluis is based on the following moment restrictions:

$$\frac{1}{TS}\sum_{t=1}^{TS}\frac{\partial}{\partial\theta}\log f(y^s_t(\hat{\rho}^S_T)|Z^s_{t-1}(\hat{\rho}^S_T), \hat{\theta}_{iT}(\pi))].$$

for $i = 1, 2$. For the case where the moment restrictions are evaluated at $\hat{\theta}_{1T}(\pi)$ which is also obtained with a simulated path equals to TS, we can show the following result under the null:

$$\sqrt{T}I_T^{-1/2}\frac{1}{TS}\sum_{t=1}^{TS}\frac{\partial}{\partial\theta}\log f(y^s_t(\hat{\rho}^S_T)|Z^s_{t-1}(\hat{\rho}^S_T), \hat{\theta}_{1T}(\pi))] \Rightarrow \left[\frac{B(\pi)}{\pi} - \frac{1}{\sqrt{S}}B^s(1)\right]$$
$$-I^{-1/2}M_\rho\left[M'_\rho I^{-1}M_\rho\right]^{-1}M'_\rho I^{-1/2}\left[B(1) - \frac{1}{\sqrt{S}}B^s(1)\right].$$

The LM structural change statistic is constructed by projecting the above moment restrictions on the subspace identifying the parameters. Such a statistic has the usual asymptotic distribution (see Proposition 3.1). This result holds because the nuisance term introduced by simulation $(1/\sqrt{S})B^s(1)$ cancels out. However, the asymptotic distribution of a structural change test for overidentifying restrictions constructed by projection on the subspace orthogonal to the subspace identifying the parameters is not the same as given in Proposition 3.2. In this case, one can show that the asymptotic distribution is function of the following process:

$$\frac{\pi^2 B^s_{q-p}(1)' B^s_{q-p}(1)}{S} + B_{q-p}(\pi)' B_{q-p}(\pi).$$

The quadratic form of Brownian motion given in Proposition 3.2 is augmented by an extra term. This extra term contains a nuisance parameter that depends on the length of the simulated path, and hence the critical values depend on S. Consequently, the statistics in van der Sluis (1998) are valid only in the case where S equals infinity, or would require critical values that need to be computed for various values of S (the applications in van der Sluis (1998) use $S = 3$ with $S = \infty$ critical values).

4 Tests Exploiting Auxiliary Models

Thus far we examined a set of tests which were introduced to the literature in the context of GMM and SMM estimation and have their roots in the earlier literature (see e.g. Andrews (1993) for references). We studied the consequences of having estimation and inference via auxiliary models. The purpose of this section is to present statistics which are designed to tests for structural breaks and take advantage of the dual model setup. We cover two types of tests, a first class relates to recent work of Liu and Zhang (1998) on diagnostic testing of EMM score generators which we show are implicitly tests for structural change. The second class is based on the simulated score principle.

4.1 Liu and Zhang Tests

The results obtained in Appendix E allow us to examine the specification test in the Efficient Method of Moments framework proposes by Liu and Zhang (1998). This test is a measure of the overall goodness of fit of the auxiliary model. The zeta statistic introduced by Liu and Zhang is defined as follows:

$$\zeta_T = \frac{12}{T^3} \left(\sum_{i=1}^{T} \sum_{t=1}^{i} s_t'(\hat\theta_T) \right) (I_T)^{-1} \left(\sum_{i=1}^{T} \sum_{t=1}^{i} s_t(\hat\theta_T) \right)$$

where $s_t(\hat\theta_T) = \partial/\partial\theta \log f\left(y_t \middle| Z_{t-1}, \hat\theta_T\right)$. We will see that the zeta statistic is in fact a structural change statistic test for the parameters of the auxiliary model. Using results derived in this paper, we can show that the asymptotic distribution of the zeta statistic under the alternative is given by:

$$12 \left[\int_0^1 BB_q(\pi) - I^{-1/2} J \left(H(\pi) - \pi H(1) \right) \right]' \left[\int_0^1 BB_q(\pi) - I^{-1/2} J \left(H(\pi) - \pi H(1) \right) \right]$$

where $BB(\pi)$ is a vector of independent Brownian Bridge of dimension q. The second term in the bracket shows that the zeta statistic is powerful against a structural change alternative for the parameter vector θ.

However, this test is not an optimal test of structural change as defined by Andrews and Ploberger (1994) and Sowell (1996a,b).

4.2 Simulated Score Tests

We introduce a specific structural change test for the Efficient Method of Moments called simulated score tests. The tests rely on simulated series from a restricted null model of interest. Using the reprojection arguments of Gallant and Tauchen (1998), we can fit a sieve seminonparametric SNP density to the simulated data. Under the null, the simulated data should yield a reprojection score generator which is a martingale difference sequence when applied to the actual sample data.

In the case of structural change tests, the simulated score test consists of evaluating the score for the actual sample data for a possible breakpoint using the estimator of the auxiliary model for the simulated data. The first step is to simulate series with the restricted estimator defined in equation (2.18). The second step is to obtain the estimator of θ of the auxiliary model with the simulated series. The score for this second step is:

$$s_N(\hat{\rho}_T^S, \hat{\theta}_N) = \frac{1}{N}\sum_{t=1}^{N} \frac{\partial}{\partial \theta} \log f(y_t^s(\hat{\rho}_T^S) | Z_{t-1}^s(\hat{\rho}_T^S), \hat{\theta}_N).$$

where N is the length of the simulated series. The third step is to evaluate the score with the data for a possible breakpoint at the estimator obtained in the second step. The simulated score structural change test is then based on the following statistic:

$$m_1(\hat{\theta}_N(\hat{\rho}_T^S), \pi) = \frac{1}{T\pi}\sum_{t=1}^{T\pi} \frac{\partial}{\partial \theta} \log f\,(y_t |\, Z_{t-1}, \hat{\theta}_N(\hat{\rho}_T^S)). \tag{4.9}$$

where $\hat{\theta}_N(\hat{\rho}_T^S)$ is the estimator of the auxiliary model obtained with simulated series. The next theorem gives the asymptotic distribution of the statistic defined above.

Theorem 4.1 *Under Assumptions A.1, C.1 and D.1 and the alternative (3.3)*

$$\begin{aligned}\pi\sqrt{T} I_T^{-1/2} m_1(\hat{\theta}_N(\hat{\rho}_T^S), \pi) \quad \Rightarrow \quad & -\left[B(\pi) - I_1^{-1/2} J H(\pi)\right] + \pi \frac{\sqrt{T}}{\sqrt{N}} B^N(1) \\ & + \pi I^{-1/2} M_\rho \left[M_\rho' I^{-1} M_\rho\right]^{-1} M_\rho' I^{-1/2} \left[B(1) - \frac{1}{\sqrt{S}} B^s(1) + I^{-1/2} J H(1)\right]\end{aligned}$$

where $H(\pi) = \int_0^\pi h(\eta, \nu, r) dr$ and $B(\pi)$ and $B^N(\pi)$ are two q-dimensional vectors of mutually independent Brownian motions.

Proof: See Appendix E

Under the null, the asymptotic distribution differs from the one obtained in Theorem 3.3. In particular, it depends on the length of the simulated series N. However, replacing N by $TS\pi$ for $\pi \in \Pi$ as the length of the simulated series results in the asymptotic distribution appearing in Theorem 3.3. This is the same argument as the one developed in the discussion of the van der Sluis statistic. Using $TS\pi$ as the length for the simulated series yields asymptotic distributions under the null and under the local alternative that are identical to the distributions of the test proposes in Section 3.2. Structural change tests can then be constructed by replacing I_T by $\tilde{I}_T = (1 + \frac{1}{S})I_T$ as shown in the previous sections. The asymptotic distributions of these structural change tests are given in Proposition 3.1 for the parameter stability and in Proposition 3.2 for the stability of overidentifying restrictions. Hence, the simulated score tests have the same asymptotic distribution under the null and the alternative as the the tests appearing in Section 3.2. However, the small sample properties can differ. In particular, the usual statistics proposed in Section 3.2 are based on the unrestricted estimators θ_{iT} for $i = 1, 2$. For small or large value of $\pi \in \Pi$, the properties of the unrestricted estimators $\hat{\theta}_{iT}$ could be poor since the partial samples used to obtain these estimators are relatively small. This problem does not occur for the computation of the simulated score test.

5 Conclusions

Estimation procedures involving auxiliary models are more commonly used, particularly in situations where likelihood-based estimation is infeasible. Many empirical examples can be found in the financial econometrics literature, particularly pertaining to the estimation of continuous time processes. Financial markets experience regular disruptions, sometimes modeled as so called jumps. There may be more fundamental shifts at work and the tests proposed here would be applicable.

Besides generalizing existing test procedures we also introduced new ones which rely on the dual model setup. The simulated score tests introduced in the paper can easily be extended to hypotheses other than structural breaks. As a by product of the paper, we also showed that some recently proposed diagnostic tests for auxiliary models are de facto tests for structural change, albeit suboptimal ones.

Appendices

A Technical Assumptions

To simplify the notation, $\psi_t(y_t|Z_{t-1}, x_t; \theta)$ will be noted $\psi_t(\theta)$.

A.1 General Assumptions

Assumption A.1 *Let us suppose the following assumptions:*

- $\hat{\theta}_T - \theta^* \xrightarrow{p} 0$ *under* θ^* *and* θ^* *is an interior point of* Θ.[7]
- $\sup_{\pi\in\Pi} \|\hat{\theta}_{iT}(\pi) - \theta^*\| \xrightarrow{p} 0$ *under* θ^* *for* $i = 1, 2$ *and* θ^* *is an interior point of* Θ.[8]
- $\psi(\theta)$ *is twice continuously partially differential in* θ *for all* $\theta \in \Theta^*$ *with probability one under* θ^* *where* Θ^* *is some neighborhood of* θ^*.
- *The matrix* $\frac{1}{T}\sum_{t=1}^{T\pi} \frac{\partial^2\psi_t}{\partial\theta\partial\theta'}(\theta)$ *converges in probability to* πJ *uniformly over* $\theta \in \Theta^*$ *under* θ^*, $\forall\pi \in [0, 1]$ *for the positive definite matrix* $J = \lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T} \frac{\partial^2\psi_t}{\partial\theta\partial\theta'}(\theta^*)\right]$.
- *The conditional variance matrix* $var\left(\frac{1}{\sqrt{T\pi}}\sum_{t=1}^{T} \frac{\partial\psi_t}{\partial\theta}(\theta)|x_t\right)$ *converges in probability to* πI *uniformly over* $\theta \in \Theta^*$ *under* θ^*, $\forall\pi \in [0, 1]$ *for the positive definite matrix* $I = \lim_{T\to\infty} var\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \frac{\partial\psi_t}{\partial\theta}(\theta^*)|x_t\right)$ *which is the conditional variance according to the DGP.*
- $\|W_T - W_0\| \xrightarrow{p} 0$ *and* $sup_{\pi\in\Pi}\|W_{iT}(\pi) - W_0\| \xrightarrow{p} 0$ *for* $i = 1, 2$.
- $\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\pi]} \frac{\partial\psi_t}{\partial\theta}(\theta^*) \Rightarrow I^{1/2}B(\pi) - JH(\pi)$ *under the alternative (3.3) where* $B(\pi)$ *is a q-dimensional vector of standard Brownian motions and* $H(\pi) = \int_0^\pi h(r)dr$.

B Asymptotic Least Squares Regularity Conditions

Assumption B.1 *The following are assumed to hold:*

- $g(\theta, \rho)$ *is continuous in* θ *and* ρ *for all* $\theta \in \Theta^*$ *and all* $\rho \in \Re^0$.
- $g(\theta, \rho)$ *is continuously partially differential in* θ *for all* $\theta \in \Theta^*$.

[7]This holds, for example by Theorem 6 of Gallant (1987, Ch. 7) under Assumptions 1-6 of Gallant (1987, Ch. 7) on $\psi_t(y_t|Z_{t-1}, x_t; \theta)$. See also Pötscher and Prucha (1989), Newey (1991) or Andrews (1992) and Davidson (1994) for alternative derivations.

[8]This uniform convergence can be obtained by imposing, for example, Assumptions A in Andrews (1993, p. 845) or Assumptions in Ghysels, Guay and Hall (1997) on the score of the objective function.

- *$g(\theta, \rho)$ is continuously partially differential in ρ for all $\rho \in \Re^0$.*
- *$\hat{\rho}_T - \rho^0 \stackrel{p}{\to} 0$ under ρ^0 and ρ^0 is an interior point of $\Re$.*[9]
- *$\sup_{\pi \in \Pi} \|\hat{\rho}_{iT}(\pi) - \rho^0\| \stackrel{p}{\to} 0$ under ρ^0 for $i = 1, 2$ and ρ^0 is an interior point of $\Re$.*

C Indirect Inference Regularity Conditions

To simplify the notation, $\psi\left(y_t^s(\rho)|Z_{t-1}(\rho), \theta\right)$ will be noted $\psi_t^s\left(\theta(\rho)\right)$.

Assumption C.1 *The following are assumed to hold:*

- *$\hat{\rho}_T^S - \rho^0 \stackrel{p}{\to} 0$ under ρ_0 and ρ^0 is an interior point of $\Re$.*[10]
- *$sup_{\pi \in \Pi} \|\hat{\rho}_{iT}^S(\pi) - \rho^0\| \stackrel{p}{\to} 0$ under ρ^0 for $i = 1, 2$ and ρ^0 is an interior point of $\Re$.*
- *$\sup_{\rho \in \Re} \|\hat{\theta}_T^s(\rho) - b(\rho)\| \stackrel{p}{\to} 0$ for $s = 1, \ldots, S$.*
- *$\hat{\theta}_{iT}^s(\rho, \pi)$ converges in probability to $b(\rho)$ uniformly over $\pi \in \Pi$ and $\rho \in \Re$ for $s = 1, \ldots, S$ and $i = 1, 2$.*
- *$\psi_t^s(\theta(\rho))$ is twice continuously partially differential in θ for all $\theta \in \Theta^*$ with probability one under $\theta^* = b(\rho^0)$.*
- *The matrix $\frac{1}{T}\sum_{t=1}^{T\pi} \frac{\partial^2 \psi_t^s}{\partial\theta\partial\theta'}(\theta(\rho))$ converges in probability to πJ uniformly over $\theta \in \Theta^*$ under $\theta^* = b(\rho^0)$, $\forall \pi \in [0, 1]$ and $s = 1, \ldots, S$.*
- *The conditional variance matrix $var\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\pi]} \frac{\partial \psi_t^s}{\partial\theta}\left(\theta(\rho)\right)|x_t\right)$ converges in probability to πI uniformly over $\theta \in \Theta^*$ under $\theta^* = b(\rho^0)$, $\forall \pi \in [0, 1]$ and $s = 1, \ldots, S$.*
- *$\hat{\theta}_T^s(\rho)$ is continuously partially differentiable in ρ for all $\rho \in \Re^0$ with probability one under ρ^0 for $s = 1, \ldots, S$.*
- *$\hat{\theta}_{iT}^s(\rho, \pi)$ is continuously partially differentiable in ρ for all $\rho \in \Re^0$ and $\pi \in \Pi$ with probability one under ρ^0 for $i = 1, 2$ and $s = 1, \ldots, S$.*
- *$b(\rho)$ is one-to-one and is continuously partially differentiable in ρ for all $\rho \in \Re^0$ and is noted $b_\rho(\rho)$.*
- *The matrix $\frac{\partial \hat{\theta}_T^s}{\partial \rho'}(\rho)$ has full column rank and converges in probability to the full column rank matrix $b_\rho(\rho)$ uniformly over $\rho \in \Re$, for $s = 1, \ldots, S$.*

[9] The uniform convergence of $\hat{\rho}_T$ and $\hat{\rho}_{iT}(\pi)$ **can be obtained by assumptions imposed above and under an assumption of identification and imposing that $\Re$ is compact (see Gouriéroux, Monfort and Trognon (1985)).**

[10] Uniform convergence for simulated estimators of ρ and θ can be obtained under regularity conditions considered in Duffie and Singleton (1993) and Ghysels and Guay (2003).

- *The matrices* $\frac{\partial \hat{\theta}^s_{iT}}{\partial \rho'}(\rho,\pi)$ *have full rank column and converge in probability to the full column rank matrix* $b_\rho(\rho)$ *uniformly over* $\rho \in \Re$ *and* $\pi \in \Pi$, $i = 1, 2$ *and* $s = 1, \ldots, S$.

- $\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\pi]} \frac{\partial \psi^s_t}{\partial \theta}(b(\rho^0)) \Rightarrow I^{1/2} B(\pi)^s$ *where* $B(\pi)^s$ *is a q-dimensional vector of Brownian motions for* $s = 1, \ldots, S$ *with* $b(\rho^0) = \theta^*$

D Efficient Method of Moments Regularity Conditions

To simplify the notation $\frac{\partial}{\partial \theta} \log[f(y^s_t(\rho)|Z^s_{t-1}(\rho), \theta)]$ will be noted $\frac{\partial}{\partial \theta} \log[f^s_t(\theta(\rho))]$.

Assumption D.1 *The following are assumed to hold:*

- $\frac{\partial}{\partial \theta} \log[f^s_t(\theta(\rho))]$ *is continuously partially differential in* θ *for all* $\theta \in \Theta^*$ *with probability one under* $\theta^* = b(\rho^0)$.

- $\frac{\partial}{\partial \theta} \log[f^s_t(\theta(\rho))]$ *is continuously partially differential in* ρ *for all* $\rho \in \Re^0$ *with probability one under* ρ^0.

- *The matrix* $\left[-\frac{1}{TS} \sum_{t=1}^{TS\pi} \frac{\partial^2 \log f^s_t}{\partial \theta \partial \theta'}(\theta(\rho))\right]$ *converges in probability to* πJ *uniformly over* $\theta \in \Theta^*$ *under* $\theta^* = b(\rho^0)$, $\forall \pi \in [0, 1]$ *where* $J = \lim_{T\to\infty} \left[-\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \log f_t}{\partial \theta \partial \theta'}(\theta^*)\right]$.

- *The matrix* $\frac{\partial m^S}{\partial \rho'}(\rho, \hat{\theta}_T)$ *has full column rank and converges in probability to* M_ρ *uniformly over* $\rho \in \Re^0$ *under* ρ^0 *where* $M_\rho = \frac{\partial m^S}{\partial \rho'}(\rho^0, \theta^*)$

- *The matrices* $\frac{\partial m^S}{\partial \rho'}(\rho, \hat{\theta}_{iT}(\pi))$ *have full column rank and converge in probability to* M_ρ *for* $i = 1, 2$ *uniformly over* $\pi \in \Pi$ *and* $\rho \in \Re^0$ *under* ρ^0 *where* $M_\rho = \frac{\partial m^S}{\partial \rho'}(\rho^0, \theta^*)$

E Proof of Theorems

E.1 Proof of Theorem 3.1

We need to use the following Lemma to proof the Theorem.

Lemma E.1 *Under assumptions 2.1, A.1, B.1 and the alternative hypothesis (3.3), the asymptotic distribution of the full sample ALS estimator is*

$$\sqrt{T}\left(\hat{\rho}_T - \rho_0\right) \xrightarrow{d} \left[G_\rho^{'} W_0 G_\rho\right]^{-1} G_\rho^{'} W_0 G_\theta J^{-1} I^{1/2}\left[B(1) - I^{-1/2} J H(1)\right]$$

and the asymptotic distributions of the unrestricted M-estimators are:

$$\sqrt{T}\left(\hat{\theta}_{1T}(\pi) - \theta^*\right) \Rightarrow -J^{-1} I^{1/2}\left[\frac{B(\pi) - I^{-1/2} J H(\pi)}{\pi}\right]$$

and

$$\sqrt{T}\left(\hat{\theta}_{2T}(\pi) - \theta^*\right) \Rightarrow -J^{-1} I^{1/2}\left[\frac{B(1) - B(\pi) - I^{-1/2} J\left(H(1) - H(\pi)\right)}{(1-\pi)}\right].$$

Proof of Lemma E.1:

First, the asymptotic distribution for the restricted estimator is shown. By the mean value expansion for the F.O.C. evaluated at $\hat{\theta}_T$:

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial \psi_t(\hat{\theta}_T)}{\partial \theta} = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial \psi_t(\theta^*)}{\partial \theta} + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \psi_t(\tilde{\theta})}{\partial \theta \partial \theta'}(\hat{\theta}_T - \theta^*) + o_p(1)$$

where $\tilde{\theta}' = [\tilde{\theta}^{(1)} \ldots \tilde{\theta}^{(q)}]$ and $\tilde{\theta}^{(k)} = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}_T^{(k)}$ for some $0 \le \lambda^{(k)} \le 1$ and $k = 1, \ldots, q$. Since $\hat{\theta}_T$ is consistent for θ^*, $\tilde{\theta} \xrightarrow{p} \theta^*$ and by assumption A.1, the expression above yields:

$$\sqrt{T}(\hat{\theta}_T - \theta^*) = -J^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial \psi_t(\theta^*)}{\partial \theta} + o_p(1).$$

By assumption A.1, we also have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial \psi_t(\theta^*)}{\partial \theta} \xrightarrow{d} I^{1/2}B(1) - JH(1)$$

where $H(1) = \int_0^1 h(r)dr$. The asymptotic distribution of the full sample estimator $\hat{\theta}_T$ is then given by

$$\sqrt{T}(\hat{\theta}_T - \theta^*) \xrightarrow{d} -J^{-1} I^{1/2}\left[B(1) - I^{-1/2} J H(1)\right]. \tag{E.1}$$

The mean value expansion for the restricted ALS estimator is

$$\left(\frac{\partial g(\hat{\theta}_T, \hat{\rho}_T)}{\partial \rho^{'}}\right)^{'} W_T g(\hat{\theta}_T, \hat{\rho}_T) \quad = \quad \left(\frac{\partial g(\hat{\theta}_T, \hat{\rho}_T)}{\partial \rho^{'}}\right)^{'} W_T g(\theta^*, \rho^0)$$

$$+\left(\frac{\partial g(\hat{\theta}_T,\hat{\rho}_T)}{\partial\rho'}\right)' W_T\frac{\partial g}{\partial\theta'}(\tilde{\theta},\hat{\rho}_T)(\hat{\theta}_T-\theta^*)$$
$$+\left(\frac{\partial g(\hat{\theta}_T,\hat{\rho}_T)}{\partial\rho'}\right)' W_T\frac{\partial g}{\partial\rho'}(\hat{\theta},\tilde{\rho})(\hat{\rho}_T-\rho^0)+o_p(1)$$

where $\tilde{\rho}'=[\tilde{\rho}^{(1)}\dots\tilde{\rho}^{(p)}]$ and $\tilde{\rho}^{(k)}=\lambda^{(k)}\rho^{0,(k)}+(1-\lambda^{(k)})\hat{\rho}^{(k)}$ for some $0\le\lambda^{(k)}\le 1$ and $k=1,\dots,p$ and $\tilde{\theta}$ is defined above.

By the fact that the first term of the right hand side is equal to zero, the full sample estimator $\hat{\rho}_T$ is only function of the asymptotic distribution of $\hat{\theta}_T$. Since $\hat{\theta}_T$ is consistent for θ^*, then $\tilde{\theta}\xrightarrow{p}\theta^*$ and $\hat{\rho}_T$ is consistent for ρ^0 then $\tilde{\rho}\xrightarrow{p}\rho^0$. Under assumptions 2.1, A.1, B.1, the continuous mapping theorem, the consistency of $\tilde{\theta}$ and $\tilde{\rho}$, the expression above yields

$$\sqrt{T}\left(\hat{\rho}_T-\rho^0\right)=-\left[G_\rho' W_0 G_\rho\right]^{-1}G_\rho' W_0 G_\theta\sqrt{T}(\hat{\theta}_T-\theta^*)+o_p(1).$$

and by the result (E.1),

$$\sqrt{T}\left(\hat{\rho}_T-\rho^0\right)\xrightarrow{d}\left[G_\rho' W_0 G_\rho\right]^{-1}G_\rho' W_0 G_\theta J^{-1}I^{1/2}\left[B(1)-I^{-1/2}JH(1)\right].$$

The optimal estimator is obtained with the following weighting matrix

$$W_T=\Omega_T^{-1}=\left[\frac{\partial g(\hat{\theta}_T,\hat{\rho}_T)}{\partial\theta'}J_T^{-1}(\hat{\theta}_T)I_T(\hat{\theta}_T)J_T^{-1}(\hat{\theta}_T)\left(\frac{\partial g(\hat{\theta}_T,\hat{\rho}_T)}{\partial\theta'}\right)'\right]^{-1}$$ [11]

and

$$\Omega_T\xrightarrow{p}\Omega=\left[G_\theta J^{-1}IJ^{-1}G_\theta'\right].$$

where $J_T(\hat{\theta}_T)=\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2\psi_t}{\partial\theta\partial\theta'}(\hat{\theta}_T)$ and $I_T(\hat{\theta}_T)$ is consistent estimator of I.[12]

Now, we derive the asymptotic distribution for the unrestricted estimators. By the mean value expansion for the M-estimators $\hat{\theta}_{1T}(\pi)$ for the first subsample:

$$\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi_t(\hat{\theta}_{1T}(\pi))}{\partial\theta}=\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi_t(\theta^*)}{\partial\theta}+\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial^2\psi_t(\tilde{\theta_1}(\pi))}{\partial\theta\partial\theta'}(\hat{\theta}_{1T}(\pi)-\theta^*)+o_p(1)$$

where $\tilde{\theta_1}'(\pi)=[\tilde{\theta_1}^{(1)}(\pi)\dots\tilde{\theta_1}^{(q)}(\pi)]$ and $\tilde{\theta_1}^{(k)}(\pi)=\lambda^{(k)}\theta^{*(k)}+(1-\lambda^{(k)})\hat{\theta}_{1T}^{(k)}(\pi)$ for some $0\le\lambda^{(k)}\le 1$ and $k=1,\dots,q$. Since $\hat{\theta}_{1T}(\pi)$ is consistent for θ^*, $\tilde{\theta}_1(\pi)\xrightarrow{p}\theta^*$ and under assumption A.1, this yields

$$\sqrt{T}\left(\hat{\theta}_{1T}(\pi)-\theta^*\right)=-J^{-1}\sqrt{T}\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi_t(\theta^*)}{\partial\theta}+o_p(1).$$

[11] This expression may exist only with probability → 1. When this expression is singular, a g-inverse can be used in place of the inverse.

[12] In pure time series context, a consistent estimator can be obtained using methods developed by Gallant (1987), Andrews and Monahan (1992), Newey and West (1994) among others. In the general case, the same methods applied, but the score must be centered around its empirical mean (see Gouriéroux et al. (1993)).

By the mean value expansion for the M-estimators $\hat{\theta}_{2T}(\pi)$ for the second subsample:

$$\frac{1}{T-[T\pi]}\sum_{t=[T\pi]+1}^{T}\frac{\partial\psi_t(\hat{\theta}_T(\pi))}{\partial\theta} = \frac{1}{T-[T\pi]}\sum_{t=[T\pi]+1}^{T}\frac{\partial\psi_t(\theta^*)}{\partial\theta} + \frac{1}{T-[T\pi]}\sum_{t=[T\pi]+1}^{T}\frac{\partial^2\psi_t(\tilde{\theta_2}(\pi))}{\partial\theta\partial\theta'}(\hat{\theta}_{2T}(\pi)-\theta^*)+o_p(1)$$

where $\tilde{\theta_2}'(\pi) = [\tilde{\theta_2}^{(1)}(\pi)\ldots\tilde{\theta_2}^{(q)}(\pi)]$ and $\tilde{\theta_2}^{(k)}(\pi) = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}_{2T}^{(k)}(\pi)$ for some $0 \leq \lambda^{(k)} \leq 1$ and $k = 1,\ldots,q$. By Assumption A.1,

$$\frac{1}{T-[T\pi]}\sum_{t=[T\pi]+1}^{T}\frac{\partial^2\psi_t(\tilde{\theta_2}(\pi))}{\partial\theta\partial\theta'}\xrightarrow{p} J$$

using $\sum_{t=[T\pi]+1}^{T} = \sum_{t=1}^{T} - \sum_{t=1}^{[T\pi]}$. Since $\hat{\theta}_{2T}(\pi)$ is consistent for θ^*, $\tilde{\theta}_2(\pi)\xrightarrow{p}\theta^*$, we then obtain

$$\sqrt{T}\left(\hat{\theta}_{2T}(\pi)-\theta^*\right) = -J^{-1}\sqrt{T}\frac{1}{T-T[\pi]}\sum_{t=[T\pi]+1}^{T}\frac{\partial\psi_t(\theta^*)}{\partial\theta}+o_p(1).$$

By the weak convergence of the score (Assumption A.1), we get

$$\sqrt{T}\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi_t(\theta^*)}{\partial\theta}\Rightarrow I^{1/2}\left[\frac{B(\pi)-I^{-1/2}JH(\pi)}{\pi}\right]$$

and

$$\sqrt{T}\frac{1}{T-[T\pi]}\sum_{t=[T\pi]+1}^{T}\frac{\partial\psi_t(\theta^*)}{\partial\theta}\Rightarrow I^{1/2}\left[\frac{B(1)-B(\pi)-I^{-1/2}J\left(H(1)-H(\pi)\right)}{(1-\pi)}\right].$$

The asymptotic distribution of the unrestricted M-estimators are then given by:

$$\sqrt{T}\left(\hat{\theta}_{1T}(\pi)-\theta^*\right)\Rightarrow -J^{-1}I^{1/2}\left[\frac{B(\pi)-I^{-1/2}JH(\pi)}{\pi}\right] \tag{E.2}$$

$$\sqrt{T}\left(\hat{\theta}_{2T}(\pi)-\theta^*\right)\Rightarrow -J^{-1}I^{1/2}\left[\frac{B(1)-B(\pi)-I^{-1/2}J\left(H(1)-H(\pi)\right)}{(1-\pi)}\right].$$

Proof of Theorem 3.1:

First, we show the result for the first subsample. We do the mean value expansion for $\Omega_T^{-1/2}g(\hat{\theta}_{1T}(\pi),\hat{\rho}_T)$ which yields

$$\Omega_T^{-1/2}g(\hat{\theta}_{1T}(\pi),\hat{\rho}_T) = \Omega_T^{-1/2}g(\theta^*,\rho^0)+\Omega_T^{-1/2}\frac{\partial g}{\partial\theta'}(\tilde{\theta}_1(\pi),\hat{\rho}_T)(\hat{\theta}_{1T}(\pi)-\theta^*) +\Omega_T^{-1/2}\frac{\partial g}{\partial\rho'}(\hat{\theta}_{1T}(\pi),\tilde{\rho})(\hat{\rho}_T-\rho^0)+o_p(1)$$

where $\tilde{\theta_1}(\pi)$ and $\tilde{\rho}$ are defined above. By using Lemma E.1 and under assumptions 2.1, A.1 and B.1 with the convergence in probability of Ω_T to Ω and the continuous mapping theorem, we obtain that

$$\sqrt{T}\Omega_T^{-1/2}g(\hat{\theta}_{1T}(\pi),\hat{\rho}_T) \Rightarrow -\Omega^{-1/2}G_\theta J^{-1}I^{1/2}\left[\frac{B(\pi)-I^{-1/2}JH(\pi)}{\pi}\right] +\Omega^{-1/2}G_\rho\left(G_\rho'\Omega^{-1}G_\rho\right)^{-1}G_\rho'\Omega^{-1}G_\theta J^{-1}I^{1/2}\left[B(1)-I^{-1/2}JH(1)\right].$$

Since $\Omega^{-1/2}G_\theta J^{-1}I^{1/2}B(\pi)$ is a q-dimensional standard Brownian motions, the result follows. The asymptotic distribution for the second sample can be obtained in a similar way.

E.2 Proof of Theorem 3.2

Lemma E.2 *Under assumptions 2.1, A.1, C.1 and the alternative (3.3), the asymptotic distribution of the indirect inference estimator is*

$$\sqrt{T}(\hat{\rho}_T^S - \rho_0) \xrightarrow{d} -\left[b'_\rho(\rho^0)W_0 b_\rho(\rho^0)\right]^{-1} b'_\rho(\rho^0)W_0 J^{-1}I^{1/2}\left[B(1) - \frac{1}{S}\sum_{s=1}^{S} B^s(1) - I^{-1/2}JH(1)\right]$$

and the asymptotic distributions of the unrestricted simulated M-estimators are:

$$\sqrt{T}\left(\hat{\theta}_{1T}^s(\rho^0,\pi) - \theta^*\right) \Rightarrow -J^{-1}I^{1/2}\frac{B(\pi)^s}{\pi}$$

$$\sqrt{T}\left(\hat{\theta}_{2T}^s(\rho^0,\pi) - \theta^*\right) \Rightarrow -J^{-1}I^{1/2}\left[\frac{B(1)^s - B(\pi)^s}{(1-\pi)}\right]$$

Proof of Lemma E.2:

Now, we derive the asymptotic distribution of the restricted estimator $\hat{\rho}^S$. For the simulated path s, we have the following mean value expansion for the F.O.C. evaluated at $\hat{\theta}^s(\rho^0)$:

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\psi_t^s(\hat{\theta}_T^s(\rho^0))}{\partial\theta} = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial\psi_t^s(\theta^*)}{\partial\theta} + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2\psi_t^s(\tilde{\theta}^s(\rho^0))}{\partial\theta\partial\theta'}(\hat{\theta}_T^s(\rho^0) - \theta^*) + o_p(1)$$

where $\tilde{\theta}^s(\rho^0)' = [\tilde{\theta}^{s,(1)}(\rho^0)\ldots\tilde{\theta}^{s,(q)}(\rho^0)]$ and $\tilde{\theta}^{s,(k)}(\rho^0) = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}_T^{s,(k)}(\rho^0)^{(k)}$ for some $0 \le \lambda^{(k)} \le 1$ and $k = 1,\ldots,q$ and $\theta^* = b(\rho^0)$. Under Assumption C.1 and by the consistency of $\tilde{\theta}_T^s(\rho^0)$,

$$\sqrt{T}(\hat{\theta}_T^s(\rho^0) - \theta^*) = -J^{-1}\sqrt{T}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\psi_t^s(\theta^*)}{\partial\theta} + o_p(1).$$

The asymptotic distribution of $\hat{\theta}_T^s(\rho^0)$ is then given by:

$$\sqrt{T}(\hat{\theta}_T^s(\rho^0) - \theta^*) \xrightarrow{d} -J^{-1}I^{1/2}B(1)^s. \tag{E.3}$$

In contrast to the asymptotic distribution of the estimator $\hat{\theta}_T$ obtained with data, the asymptotic distribution of $\hat{\theta}_T^s(\rho^0)$ does not depend on the alternative for obvious reasons. The mean value expansion for the restricted indirect inference estimator is

$$\begin{aligned}\left[\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\hat{\theta}_T^s}{\partial\rho'}(\hat{\rho}_T^S)\right]' W_T\left[\hat{\theta}_T - \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_T^s(\hat{\rho}_T^S)\right] \quad = \quad & \left[\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\hat{\theta}_T^s}{\partial\rho'}(\hat{\rho}_T^S)\right]' W_T \\ & \left[(\hat{\theta}_T - \theta^*) - \frac{1}{S}\sum_{s=1}^{S}(\hat{\theta}_T^s(\hat{\rho}_T^S) - \theta^*)\right. \\ & \left. - \left(\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\hat{\theta}_T^s}{\partial\rho'}(\tilde{\rho}_T^S)\right)(\hat{\rho}_T^S - \rho_0)\right] + o_p(1)\end{aligned}$$

where $\tilde{\rho}^S$ is defined as in the proof of Lemma E.1 but for the estimator obtained with S simulated paths. By Assumptions C.1 and the consistency of $\tilde{\rho}^S_T$, this yields:

$$\sqrt{T}(\hat{\rho}^S_T - \rho_0) = \left[b'_\rho(\rho^0)W_0 b_\rho(\rho^0)\right]^{-1} b'_\rho(\rho^0)W_0\sqrt{T}\left[(\hat{\theta}_T - \theta^*) - \frac{1}{S}\sum_{s=1}^{S}(\hat{\theta}^s_T(\hat{\rho}^S_T) - \theta^*)\right] + o_p(1).$$

Using results (E.1) and (E.3), the asymptotic distribution of $\hat{\rho}^S_T$ is given by:

$$\sqrt{T}(\hat{\rho}^S_T - \rho_0) \xrightarrow{d} -\left[b'_\rho(\rho^0)W_0 b_\rho(\rho^0)\right]^{-1} b'_\rho(\rho^0)W_0 J^{-1}I^{1/2}\left[B(1) - \frac{1}{S}\sum_{s=1}^{S}B^s(1) - I^{-1/2}JH(1)\right]$$

The asymptotic distribution depends on the matrix W_0 and the number of simulations S. The restricted optimal Indirect Inference estimator is obtained with the following weighting matrix $W_T \xrightarrow{p} JI^{-1}J$.[13]

Now, we derive the asymptotic distribution for the unrestricted estimators. The mean value expansion of the M-estimators for the first subsample evaluated at $\hat{\theta}^s_{1T}(\rho^0,\pi)$ gives

$$\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi^s_t(\hat{\theta}^s_{1T}(\rho^0))}{\partial\theta} = \frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi^s_t(\theta^*)}{\partial\theta} + \frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial^2\psi^s_t(\tilde{\theta}^s(\rho^0),\pi)}{\partial\theta\partial\theta'}(\hat{\theta}^s_{1T}(\rho^0,\pi) - \theta^*) + o_p(1).$$

where $\tilde{\theta}^s(\rho^0,\pi)' = [\tilde{\theta}^{s,(1)}(\rho^0,\pi)\ldots\tilde{\theta}^{s,(q)}(\rho^0,\pi)]$ and $\tilde{\theta}^{s,(k)}(\rho^0,\pi) = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}^{s,(k)}_{1T}(\rho^0,\pi)$ for some $0 \le \lambda^{(k)} \le 1$ and $k = 1,\ldots,q$ and $\theta^* = b(\rho^0)$. This yields the following expression:

$$\sqrt{T}\left(\hat{\theta}^s_{1T}(\rho^0,\pi) - \theta^*\right) = -J^{-1}\frac{1}{\sqrt{[T\pi]}}\sum_{t=1}^{[T\pi]}\frac{\partial\psi^s_t(\theta^*)}{\partial\theta} + o_p(1).$$

We can obtain the equivalent expression for $\hat{\theta}_{2T}(\pi)$ by a similar mean value expansion.

By Assumption C.1, we have the following weak convergence of the score for the first and the second subsamples:

$$\sqrt{T}\frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\frac{\partial\psi^s_{1t}(\theta^*)}{\partial\theta} \Rightarrow I^{1/2}\frac{B(\pi)^s}{\pi}$$

and

$$\sqrt{T}\frac{1}{[T(1-\pi)]}\sum_{t=[T\pi]+1}^{T}\frac{\partial\psi^s_{2t}(\theta^*)}{\partial\theta} \Rightarrow I^{1/2}\frac{B(1)^s - B(\pi)^s}{(1-\pi)}.$$

Given the results above, the asymptotic distributions of the unrestricted simulated M-estimators are respectively:

$$\sqrt{T}\left(\hat{\theta}^s_{1T}(\rho^0,\pi) - \theta^*\right) \Rightarrow -J^{-1}I^{1/2}\frac{B(\pi)^s}{\pi}$$

$$\sqrt{T}\left(\hat{\theta}^s_{2T}(\rho^0,\pi) - \theta^*\right) \Rightarrow -J^{-1}I^{1/2}\left[\frac{B(1)^s - B(\pi)^s}{(1-\pi)}\right]$$

Proof of Theorem 3.2:

[13] see Gouriéroux et al. (1993).

By a mean value expansion for the first subsample, we have that:

$$\Omega_T^{-1/2}\left[\hat{\theta}_{1T}(\pi) - \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_{1T}^{s}(\hat{\rho}_T^S,\pi)\right] = \Omega_T^{-1/2}\left[(\hat{\theta}_{1T}(\pi)-\theta^*) - \frac{1}{S}\sum_{s=1}^{S}(\hat{\theta}_{1T}^{s}(\rho^0,\pi)-\theta^*)\right] - \Omega_T^{-1/2}\left[\left(\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\hat{\theta}_{1T}^{s}}{\partial\rho'}(\tilde{\rho}_T^S,\pi)\right)(\hat{\rho}_T^S-\rho_0)\right] + o_p(1)$$

By Lemma E.1 and E.2, Assumption C.1 and the consistency of $\tilde{\rho}_T^S$, we obtain that:

$$\sqrt{T}\Omega_T^{-1/2}\left[\hat{\theta}_{1T}^{s} - \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_{1T}^{s}(\hat{\rho}_T^S,\pi)\right] \Rightarrow -\Omega^{-1/2}J^{-1}I^{1/2}\left[\frac{B(\pi)}{\pi} - \frac{1}{S}\sum_{s=1}^{S}\frac{B(\pi)^s}{\pi} - \frac{I^{-1/2}JH(\pi)}{\pi}\right]$$
$$+\Omega^{-1/2}b_\rho(\rho^0)\left[b_\rho'(\rho^0)\Omega^{-1}b_\rho(\rho^0)\right]^{-1}b_\rho'(\rho^0)\Omega^{-1}J^{-1}I^{1/2}\left[B(1) - \frac{1}{S}\sum_{s=1}^{S}B(1)^s - I^{1/2}JH(1)\right]$$

Since $\Omega^{-1/2}J^{-1}I^{1/2}B(\pi)$ is a q-dimensional vector of standard Brownian motions, the result follows. The asymptotic distribution for the second sample is obtained similarly.

E.3 Proof of Theorems 3.3 and 4.1

Lemma E.3 *Under assumptions A.1, C.1, D.1 and the alternative hypothesis (3.3), the asymptotic distribution of the full sample EMM estimator is*

$$\sqrt{T}\left(\hat{\rho}_T^S - \rho_0\right) \xrightarrow{d} -\left[M_\rho' I^{-1} M_\rho\right]^{-1} M_\rho' I^{-1/2}\left[B(1) - \frac{1}{\sqrt{S}}B(1)^s - I^{-1/2}JH(1)\right].$$

Proof of Lemma E.3:

First, the asymptotic distribution for the restricted estimator is shown. By the mean value expansion for the F.O.C. evaluated at $\hat{\theta}_T$:

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial\log f_t(\hat{\theta}_T)}{\partial\theta} = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial\log f_t(\theta^*)}{\partial\theta} + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2\log f_t(\tilde{\theta})}{\partial\theta\partial\theta'}(\hat{\theta}_T-\theta^*) + o_p(1)$$

where $\tilde{\theta}' = [\tilde{\theta}^{(1)}\ldots\tilde{\theta}^{(q)}]$ and $\tilde{\theta}^{(k)} = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}_T^{(k)}$ for some $0 \le \lambda^{(k)} \le 1$ and $k = 1,\ldots,q$. By assumption A.1 and the consistency $\tilde{\theta}$, the expression above yields:

$$\sqrt{T}(\hat{\theta}_T - \theta^*) = J^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial\log f_t(\theta^*)}{\partial\theta} + o_p(1)$$

where $J = \lim_{T\to\infty}\left[-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2\log f_t}{\partial\theta\partial\theta'}(\theta^*)\right]$. By the last assumption of A.1 and noting that $\psi_t(\cdot) = -\log f_t(\cdot)$, the asymptotic distribution of the score at the pseudo-true value θ^* is:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial\log f_t(\theta^*)}{\partial\theta} \xrightarrow{d} -\left[I^{1/2}B(1) - JH(1)\right]. \quad \text{(E.4)}$$

The asymptotic distribution of the full sample estimator $\hat{\theta}_T$ is then given by

$$\sqrt{T}(\hat{\theta}_T - \theta^*) \xrightarrow{d} -J^{-1}I^{1/2}\left[B(1) - I^{-1/2}JH(1)\right] \tag{E.5}$$

The mean value expansion of the F.O.C. evaluated at the unrestricted EMM estimator is

$$\begin{aligned}\left(\frac{\partial m_T^S(\hat{\rho}_T^S,\hat{\theta}_T)}{\partial\rho'}\right)' I_T^{-1} m_T^S(\hat{\rho}_T^S,\hat{\theta}_T) &= \left(\frac{\partial m_T^S(\hat{\rho}_T^S,\hat{\theta}_T)}{\partial\rho'}\right)' I_T^{-1} m_T^S(\rho^0,\theta^*) \\ &+ \left(\frac{\partial m_T^S(\hat{\rho}_T^S,\hat{\theta}_T)}{\partial\rho'}\right)' I_T^{-1}\frac{\partial m_T^S}{\partial\theta'}(\tilde{\theta},\hat{\rho}_T^S)(\hat{\theta}_T - \theta^*) \\ &+ \left(\frac{\partial m_T^S(\hat{\rho}_T^S,\hat{\theta}_T))}{\partial\rho'}\right)' I_T^{-1}\frac{\partial m_T^S}{\partial\rho'}(\hat{\theta}_T,\tilde{\rho}^S)(\hat{\rho}_T^S - \rho^0) + o_p(1)\end{aligned}$$

where $\tilde{\rho}^{S\prime} = [\tilde{\rho}^{S,(1)}\dots\tilde{\rho}^{S,(p)}]$ and $\tilde{\rho}^{S,(k)} = \lambda^{(k)}\rho^{0,(k)} + (1-\lambda^{(k)})\hat{\rho}^{S,(k)}$ for some $0 \le \lambda^{(k)} \le 1$ and $k = 1,\dots,p$ and $\tilde{\theta}$ is defined above.

Under assumptions A.1 and D.1 and the consistency of $\tilde{\theta}$ and $\tilde{\rho}^S$, this yields

$$\sqrt{T}\left(\hat{\rho}_T^S - \rho_0\right) = \left[M_\rho' I^{-1} M_\rho\right]^{-1} M_\rho' I^{-1}\left[J(\hat{\theta}_T - \theta^*) - m_T^S(\rho^0,\theta^*)\right] + o_p(1).$$

since $\frac{\partial m_T^S}{\partial\theta'}(\tilde{\theta}_T,\hat{\rho}_T^S)$ converge in probability to $-J$. Now we examine the expression $m_T^S(\rho^0,\theta^*)$. By definition

$$\sqrt{T}m_T^S(\rho^0,\theta^*) = \sqrt{T}\frac{1}{TS}\sum_{t=1}^{TS}\frac{\partial}{\partial\theta}\log[f(y_t^s(\rho^0)|Z_{t-1}^s(\rho^0),\theta^*)].$$

Using the last assumption of C.1 with $\psi_t(\cdot) = -\log f_t(\cdot)$, we obtain

$$\sqrt{T}m_T^S(\rho^0,\theta^*) \xrightarrow{d} -\frac{1}{\sqrt{S}}I^{1/2}B(1)^s. \tag{E.6}$$

By E.5 and E.6, the asymptotic distribution of $\hat{\rho}_T^S$ is then given by:

$$\sqrt{T}\left(\hat{\rho}_T^S - \rho_0\right) \xrightarrow{d} -\left[M_\rho' I^{-1} M_\rho\right]^{-1} M_\rho' I^{-1/2}\left[B(1) - \frac{1}{\sqrt{S}}B(1)^s - I^{-1/2}JH(1)\right].$$

In the case where $S = \infty$, the term $B(1)^s$ disappears.

Proof of Theorem 3.3:

First, we show the result for the first subsample. We do the mean value expansion for $I_T^{-1/2}m_1^S(\hat{\rho}_T,\hat{\theta}_{1T}(\pi))$ which gives

$$\begin{aligned}I_T^{-1/2}m_1^S(\hat{\rho}_T^S,\hat{\theta}_{1T}(\pi)) &= I_T^{-1/2}m_1^S(\theta^*,\rho^0) + I_T^{-1/2}\frac{\partial m_1^S}{\partial\theta'}(\tilde{\theta}_1(\pi),\hat{\rho}_T^S)(\hat{\theta}_{1T}(\pi) - \theta^*) \\ &+ I_T^{-1/2}\frac{\partial m_1^S}{\partial\rho'}(\hat{\theta}_{1T}(\pi),\tilde{\rho}^S)(\hat{\rho}_T^S - \rho^0) + o_p(1)\end{aligned}$$

with $\tilde{\theta_1}'(\pi) = [\tilde{\theta_1}^{(1)}(\pi) \dots \tilde{\theta_1}^{(q)}(\pi)]$ and $\tilde{\theta_1}^{(k)}(\pi) = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}^{(k)}_{1T}(\pi)$ for some $0 \leq \lambda^{(k)} \leq 1$ and $k = 1, \dots, q$ and $\tilde{\rho}^S$ is defined above. Let us examine the expression $m_1^S(\rho^0, \theta^*)$. By definition

$$\sqrt{T} m_1^S(\rho^0, \theta^*) = \sqrt{T} \frac{1}{[TS\pi]} \sum_{t=1}^{[TS\pi]} \frac{\partial}{\partial \theta} \log[f(y_t^s(\rho^0)|Z_{t-1}^s(\rho^0), \theta^*)].$$

Using the last assumption of C.1 with $\psi_t(\cdot) = -\log f_t(\cdot)$, this yields

$$\sqrt{T} m_1^S(\rho^0, \theta^*) \xrightarrow{d} -\frac{1}{\sqrt{S}} I^{1/2} \left(\frac{B(\pi)^s}{\pi} \right). \tag{E.7}$$

By using Lemma E.1 and E.3, assumptions A.1, D.1, the result E.7 and the consistency of $\tilde{\theta}_1(\pi)$ and $\tilde{\rho}^S$, we get

$$\begin{aligned} \sqrt{T} I_T^{-1/2} m_1^S(\hat{\rho}_T^S, \hat{\theta}_{1T}(\pi)) \quad \Rightarrow \quad & \left[\frac{B(\pi)}{\pi} - \frac{1}{\sqrt{S}} \left(\frac{B(\pi)^s}{\pi} \right) - \frac{I^{-1/2} J H(\pi)}{\pi} \right] \\ & - I^{-1/2} M_\rho \left(M_\rho' I^{-1} M_\rho \right)^{-1} M_\rho I^{-1/2} \left[B(1) - \frac{1}{\sqrt{S}} B(1)^s - I^{-1/2} J H(1) \right]. \end{aligned}$$

The asymptotic distribution for the second sample can be obtained in a similar way.

Proof of Theorem 4.1:

First, we show the asymptotic distribution of the estimator $\hat{\theta}_N$ obtained with the simulated series for ρ fixed at the restricted estimator $\hat{\rho}_T^S$. The mean value expansion of the score of the auxiliary model evaluated at $\hat{\theta}_N(\hat{\rho}_T^S)$ is given by:

$$\begin{aligned} s_N(\hat{\rho}_T^S, \hat{\theta}_N) \quad = \quad & s_N(\theta^*, \rho^0) + \frac{\partial s_N}{\partial \theta'}(\tilde{\theta}(\rho^0), \hat{\rho}_T^S)(\hat{\theta}_N(\hat{\rho}_T^S) - \theta^*) \\ & + \frac{\partial s_N}{\partial \rho'}(\hat{\theta}_N, \tilde{\rho}^S)(\hat{\rho}_T^S - \rho^0) + o_p(1). \end{aligned}$$

with $\tilde{\theta}'(\rho^0) = [\tilde{\theta}^{(1)}(\rho^0) \dots \tilde{\theta}^{(q)}(\rho^0)]$ and $\tilde{\theta}^{(k)}(\rho^0) = \lambda^{(k)}\theta^{*(k)} + (1-\lambda^{(k)})\hat{\theta}_N^{(k)}(\rho^0)$ for some $0 \leq \lambda^{(k)} \leq 1$ and $k = 1, \dots, q$, $\theta^* = b(\rho^0)$ and $\tilde{\rho}^S$ defined above. By the asymptotic distribution of the restricted estimator $\hat{\rho}_T^S$ given in Lemma E.3, assumptions A.1, D.1, result E.6 and the consistency of $\tilde{\theta}$ and $\tilde{\rho}^S$, we obtain that

$$\begin{aligned} \sqrt{N}(\hat{\theta}_N(\hat{\rho}_T^S) - \theta^*) \quad \xrightarrow{d} \quad & -J^{-1} I^{1/2} B(1)^s - \frac{\sqrt{N}}{\sqrt{T}} J^{-1} M_\rho \left(M_\rho' I^{-1} M_\rho \right)^{-1} M_\rho I^{-1/2} \\ & \left[B(1) - \frac{1}{\sqrt{S}} B(1)^s - I^{-1/2} J H(1) \right]. \end{aligned} \tag{E.8}$$

The mean value expansion of the score evaluated at $\hat{\theta}_N$ for the data under the alternative for the first subsample is

$$I_T^{-1/2} m_1(\hat{\theta}_N(\hat{\rho}_T^S), \pi) = I_T^{-1/2} m_1^S(\theta^*, \pi) + I_T^{-1/2} \frac{\partial m_1}{\partial \theta'}(\tilde{\theta}, \pi)(\hat{\theta}_N(\hat{\rho}_T^S) - \theta^*) + o_p(1).$$

with obvious notation for $\tilde{\theta}$. By definition

$$\sqrt{T} m_1^S(\theta^*, \pi) = \sqrt{T} \frac{1}{[T\pi]} \sum_{t=1}^{[T\pi]} \frac{\partial}{\partial \theta} \log[f(y_t)|Z_{t-1}, \theta^*)]$$

and using the last assumption of A.1 with $\psi_t(\cdot) = -\log f_t(\cdot)$, this yields

$$\sqrt{T} m_1^S(\theta^*, \pi) \Rightarrow -\left(\frac{I^{1/2}B(\pi) - JH(\pi)}{\pi}\right). \tag{E.9}$$

By the asymptotic distribution of the $\hat{\theta}_N(\hat{\rho}_T^S)$ derived above, Assumptions A.1 and D.1, result E.9 and the consistency of $\tilde{\theta}$, we obtain under the alternative that

$$\begin{aligned} \pi\sqrt{T} I_T^{-1/2} m_1(\hat{\theta}_N(\hat{\rho}_T^S), \pi) \quad \Rightarrow \quad & -\left[B(\pi) - I^{-1/2}JH(\pi)\right] + \pi\frac{\sqrt{T}}{\sqrt{N}}B(1)^s \\ & + \pi I^{-1/2} M_\rho \left(M_\rho' I^{-1} M_\rho\right)^{-1} M_\rho I^{-1/2}\left[B(1) - \frac{1}{\sqrt{S}}B(1)^s - I^{-1/2}JH(1)\right]. \end{aligned}$$

References

[1] Andrews, D.W.K. (1992), "Generic Uniform Convergence," *Econometric Theory,* **8,** 241-257.

[2] Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica,* **61,** 821-856.

[3] Andrews, D.W.K. and R. Fair (1988), "Inference in Econometrics Models with Structural Change," *Review of Econmic Studies,* **55,** 615-640.

[4] Andrews, D.W.K. and C.J. McDermott (1995), "Nonlinear Econometric Models with Deterministically Trending Variables," *Review of Economic Studies,* **62,** 343-360.

[5] Andrews, D.W.K. and J.C. Monahan (1992), "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica,* **60,** 953-966.

[6] Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests when a Nuisance Parameter is Present Only under the Alternative," *Econometrica,* **62,** 1383-1414.

[7] Davidson, J. (1994), "Stochastic Limit Theory: An Introduction for Econometricians," Oxford University Press, Oxford.

[8] Dridi, R. L., A. Guay and E. Renault (2003), "Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models," *Discussion Paper.*

[9] Dridi and E. Renault (2001), "Semi-parametric Indirect Inference," *unpublished paper.*

[10] Duffie, D. and K. J. Singleton (1993), "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica,* **61,** 929-952.

[11] Dufour, J.M., E. Ghysels and A. Hall (1994), "Generalized Predictive Tests and Structural Change Analysis in Econometrics," *International Economic Review,* **35,** 199-229.

[12] Gallant, A.R. (1987), "Nonlinear Statistical Models," John Wiley, New York.

[13] Gallant, A.R. and G. Tauchen (1996), "Which Moments to Match?" *Econometric-Theory,* **12**, 657-681.

[14] Gallant, A.R. and G. Tauchen (1998), "Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions," *Journal of American Statistical Association,***93,** 10-24.

[15] Ghysels, E., A. Guay (2003), "Structural change tests for Simulated Method of Moments," *Journal of Econometrics,* **115,** 91-123.

[16] Ghysels, E., A. Guay and A. Hall (1997), "Predictive Test for Structural Change with Unknown Breakpoint," *Journal of Econometrics*, **82,** 209-233.

[17] Ghysels, E. and A. Hall (1990), "A Test for Structural Stability of Euler Conditions Parameters Estimated Via the Generalized Method of Moments Estimator," *International Economic Review,* **31,** 355-364.

[18] Gouriéroux, C., A. Monfort and E. Renault (1993), "Indirect Inference" *Journal of Applied Econometrics,* **8**, S85-S118.

[19] Gouriéroux, C., A. Monfort and A. Trognon (1985), "Moindres Carrés Asymptotiques," *Annales de l'INSEE*, **58,** 91-121.

[20] Guay, A. (2003), "Optimal Predictive Tests," *Econometric Reviews*, **22,** 379-410.

[21] Guay, A. and E. Renault (2003), "Indirect Encompassing with Misspecied Models," manuscript.

[22] Hall, A. and A. Sen (1999), "Structural Stability Testing in Models Estimated by Generalized Method of Moments," *Journal of Business and Economic Statistics* **17**, 335-348.

[23] Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* **50,** 1029-1054.

[24] Hoffman, D. and A. Pagan (1989), "Post-Sample Prediction Tests for Generalized Method of Moments Estimators," *Oxford Bulletin of Economics and Statistics,* **51,** 331-343.

[25] Liu, M. and H.H. Zhang (1998), "Specification Tests in the Efficient Method of Moments Framework with Application to the Stochastic Volatility Models," *Discussion Paper*, Carnegie Mellon University.

[26] McFadden (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica,* **57,** 995-1026.

[27] Newey, W.K. (1991), "Uniform Convergence in Probability and Stochastic Equicontinuity," *Econometrica,* **59,** 703-708.

[28] Newey, W.K. and K. West (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies,* **61,** 631-653.

[29] Pakes, A. and D. Pollard (1984), "Simulation and the Asymptotic of Optimization Estimators," *Econometrica,* **57,** 1027-1058.

[30] Pollard, D. (1984), "Convergence of Stochastic Processes," Springer-Verlag, New York.

[31] Potscher, B.M. and Prucha, I. R. (1989), "A Uniform Law of Large Numbers for Dependent and Heterogenous Processes," *Econometricea,* **57,** 675-683.

[32] Sowell, F. (1996a), "Optimal Tests for Parameter Instability in the Generalized Method of Moments Framework," *Econometrica,* **64,** 1085-1107.

[33] Sowell, F. (1996b), "Tests for Violations of Moment Conditions," *Manuscript*, Graduate School of Industrial Administration, Carnegie Mellon University.

[34] van der Sluis, P. J. (1998), "Structural Stability Tests with Unknown Breakpoint for the Efficient Method of Moments with Application to Stochastic Volatility Models," *Manuscript*, University of Amsterdam, department of Econometrics.