DISENTANGLING SYSTEMATIC AND IDIOSYNCRATIC DYNAMICS IN PANELS OF VOLATILITY MEASURES

Matteo BARIGOZZI[1], Christian BROWNLEES[2], Giampiero M. GALLO[3] and David VEREDAS[4]

Abstract

Realized volatilities observed across several assets show a common secular trend and some idiosyncratic pattern which we accommodate by extending the class of Multiplicative Error Models (MEMs). In our model, the common trend is estimated nonparametrically, while the idiosyncratic dynamics are assumed to follow univariate MEMs. Estimation theory based on seminonparametric methods is developed for this class of models for large cross-sections and large time dimensions. The methodology is illustrated using two panels of realized volatility measures between 2001 and 2008: the SPDR Sectoral Indices of the S&P500 and the constituents of the S&P100. Results show that the shape of the common volatility trend captures the overall level of risk in the market and that the idiosyncratic dynamics have an heterogeneous degree of persistence around the trend. Out–of–sample forecasting shows that the proposed methodology improves volatility prediction over several benchmark specifications.

Keywords: Vector Multiplicative Error Model, Seminonparametric Estimation, Volatility.

JEL classification: C32, C51, G01.

[1]London School of Economics and Political Science – Department of Statistics; e-mail: m.barigozzi@lse.ac.uk.

[2]Universitat Pompeu Fabra – Department of Economics and Business & Barcelona GSE; e-mail: christian.brownlees@upf.edu.

[3]Università di Firenze – Dipartimento di Statistica, Informatica, Applicazioni; e-mail: gallog@disia.unifi.it.

[4]ECARES – Solvay Brussels School of Economics and Management – Université libre de Bruxelles; e-mail: david.veredas@ulb.ac.be.

Corresponding address: David Veredas, ECARES – Solvay Brussels School of Economics and Management – Université libre de Bruxelles, 50 Av F.D. Roosevelt CP114/04, B1050 Brussels, Belgium. Phone: +32(0)26504218. Fax: +32(0)26504475.

1 Introduction

The model introduced in this paper is motivated by the analysis of panels of realized volatility measures (RV). Visual inspection of the data shows that RVs tend to oscillate around a common average level that can be associated with the overall level of risk in the market. Disentangling systematic and idiosyncratic components allows us to understand which movements are due to common and individual sources. The memory of the idiosyncratic processes gives also interesting insights on how the persistence of individual shocks changes, once the presence of a common long run variance is accounted for.

From a statistical perspective, we extend the class of Multiplicative Error Models (MEM) for nonnegative multivariate time series (Engle (2002), Engle and Gallo (2006)), and dynamic models with slowly moving components as in, *inter alia*, Engle and Rangel (2008). Our model could also be applied to identify systematic and idiosyncratic dynamics in panels of market activity, risk or liquidity measures (e.g. traded volumes, spreads, trading intensities) exhibiting similar empirical regularities.

We introduce a vector MEM that decomposes the conditional expectation of each series as the product of a systematic trend (modeled as a nonparametric curve) and an idiosyncratic dynamic component (modeled as a univariate MEM). A simple estimation approach makes the model appealing even when the number of series in the panel is large.

We choose to model interdependence across realized variances, not resorting to the use of realized covariances. For several applications, as a matter of fact, such as variance derivatives trading, volatilities are the main object of interest. The construction and modelling of large realized covariance matrices still poses some theoretical and practical challenges, and is a topic of active research (Chiriac and Voev (2011), Noureldin *et al.* (2012a)). Asset pricing models developed in finance also motivate decomposing the risk of an asset in a systematic and idiosyncratic components. However, in these models systematic and idiosyncratic components are additive, while here we adopt a multiplicative framework motivated by the stylized facts of the data.

The estimation approach developed for this class of models combines ideas from the literature on profile likelihood and copulas. First, building up on the inference from the marginals framework of Joe (2005), the joint conditional likelihood of the model is decomposed in the contribution of the marginal densities and joint copula dependence. Second, the marginal densities are used to estimate both the nonparametric common trend and the parametric idiosyncratic dynamics. We justify this approach using results from profile likelihood maximization Staniswalis (1989), Severini and Wong (1992) and Veredas *et al.* (2007)). This technique allows to establish efficiency bounds in a seminonparametric setting. The estimation procedure boils down to a univariate estimation of the systematic component and the univariate estimation of the idiosyncratic marginal dynamics of each series. The large sample properties of the estimators are derived and, in particular, we show that the asymptotic variance of the estimated parameters is the smallest possible, given the seminonparametric and two–step nature of the procedure. The theory is developed in a setting that allows both large cross-sections and large time dimensions. A Monte Carlo study shows that in finite samples the estimators perform adequately and that standard inferential procedures behave satisfactorily.

We apply the model to two panels of daily realized volatilities spanning from January 2, 2001 to December 31, 2008. The first panel consists of the nine sectoral indices of the SPDR S&P500 index, while the second contains the ninety constituents of the S&P100 that have continuously been traded in the sample period. The datasets are related to each other in that the constituents of the S&P100 are also some of the main underlying assets of the SPDR sectoral indices. The empirical results of the two applications are consistent with one another. The estimated shape of the systematic risk is essentially the same in the two panels, and its level can be associated with the global level of uncertainty in the economy, which exhibits clear peaks at the beginning and end of the 2000s in correspondence to the dot–com bubble burst and the financial crisis. Once the systematic trend is accounted for, idiosyncratic dynamics are mean reverting. Interestingly, the speed of reversion is rather heterogeneous across assets. For instance, in the SPDR panel, mean reversion is more pronounced for Consumer Discretionary and Materials, while it is much slower for Technology and Energy. Moreover, the S&P100 panel exhibits on average more idiosyncratic dynamics (in the sense of slower mean reversion) than the SPDR sectors. By the same token, inspection of the idiosyncratic dynamics reveals interesting patterns: Technology appears to have been more volatile during the dot–com bubble burst, the Energy sector experienced turmoil during the energy crisis in 2005–2006, and Financials were under distress during later years when the crisis hit this sector the most. Finally, an out–of–sample forecasting exercise is used to assess the predictive ability of the specification. We forecast realized volatility from 2007 to the end of the sample using a number of MEM–based specifications, and showing that forecasts based on the our model are almost always able to improve the out–of–sample predictive ability.

Different strands of literature relate to our work. Starting from the contribution of Engle and Rangel (2008), there has been interest in capturing secular trends in financial volatility. Among others, the list of contributions in a univariate setting includes Amado and Teräsvirta (2008), Engle *et al.* (2009) and Brownlees and Gallo (2010). Rangel and Engle (2012), Hafner and Linton (2010), Long *et al.* (2011) and Colacito *et al.* (2011) extend these ideas in a multivariate setting. The paper relates also to the literature on multivariate extensions of the MEM model, like the works of Cipollini *et al.* (2006) and Hautsch (2008). Moreover, there is a long tradition of decomposing panels of financial time series into a common and an idiosyncratic component in econometrics, namely in additive conditional heteroskedastic factor models (see Alessi *et al.* (2009) and citations within). The paper contributes to the growing literature on modeling daily volatility using intra-daily information. Research in this area includes Andersen *et al.* (2007), Patton and Sheppard (2009), Shephard and Sheppard (2010), Hansen *et al.* (2012) and Chen *et al.* (2011). Chiriac and Voev (2011) and Noureldin *et al.* (2012a) explore models for realized covariance matrices. This work also fits with the larger segment of the literature that finds evidence of long range dependence in volatility and have proposed ways to capture it. Significant contributions include long memory models (Andersen *et al.* (2003), Deo *et al.* (2006), Andersen *et al.* (2007), Corsi (2010), and Luciani and Veredas (2011)). Finally, this paper relates to the recent strand of literature on panel volatility modeling. Contributions in this area include Bauwens and Rombouts (2007), Engle *et al.* (2008), Pakel *et al.* (2011), Hautsch *et al.* (2011), Wang and Zou (2010) and Noureldin *et al.* (2012b).

The paper is structured as follows. Section 2 describes the panels of realized volatility measures that we use in the empirical application. Section 3 describes our specification and Section 4 details the estimation strategy. The asymptotic properties of the estimator are given in Section 5. In Section 6 we carry out a Monte Carlo exercise to assess the reliability of the proposed estimation approach. Section 7 presents the estimation results for the SPDR sectoral indices and the constituents of the S&P100. We conclude in Section 8. Assumptions, proofs and additional empirical results are gathered in Appendix A and B.

2 Stylized Facts for Panels of Volatility Measures

We study two panels of realized volatility measures from January 2, 2001 to December 31, 2008. The first, referred to as SPDR, consists of the nine Select Sector SPDRs Exchange Traded Funds (ETF) that divide the S&P500 index into sector index funds. The sectors (with the abbreviations we use and the original ticker names) are Materials (Mat, XLB), Energy (Ener, XLE), Financial (Fin, XLF), Industrial (Ind, XLI), Technology (Tech, XLK), Consumer Staples (Stap, XLP), Utilities (Util, XLU), Health Care (Heal, XLV), and Consumer Discretionary (Disc, XLY). The second panel, named S&P100, consists of U.S. equity companies that are part of the S&P100 index. It contains all the constituents of the S&P100 index as of December 2008 that have been trading in the full sample period (90 in total).[1]

Among the available estimators of the daily integrated volatility based on intraday returns, we adopt the realized kernels (Barndorff-Nielsen *et al.* (2008)).[2] They are a family of heteroskedastic and autocorrelation consistent volatility estimators, robust to various forms of market microstructure noise present in high frequency data. Our choice is motivated by the appealing theoretical properties of this family of estimators, as well as their good forecasting performance (e.g. for predicting Value at Risk, Brownlees and Gallo (2010)). Parallel analysis using alternative estimators (not reported in the paper) suggests that our results do not hinge on the specific measure of volatility chosen.

We compute optimal realized kernels following the procedure detailed in Barndorff-Nielsen *et al.* (2009). Our primary source of data are tick–by–tick intra-daily quotes from the TAQ database. Data are extracted and filtered using the methods described in Brownlees and Gallo (2006) and Barndorff-Nielsen *et al.* (2009). Let $r_{i\,t\,j}$ denote the 1–minute frequency returns (sampled in tick time) at minute j on day t for ticker i. The realized Parzen kernel estimator is defined as

$$x_{i\,t} = \sum_{h=-H}^{H} \mathrm{K}_p\left(\frac{h}{H+1}\right)\gamma_h, \quad \text{with} \quad \gamma_h = \sum_{j=|h|+1}^{J} r_{i\,t\,j} r_{i\,t\,j-|h|},$$

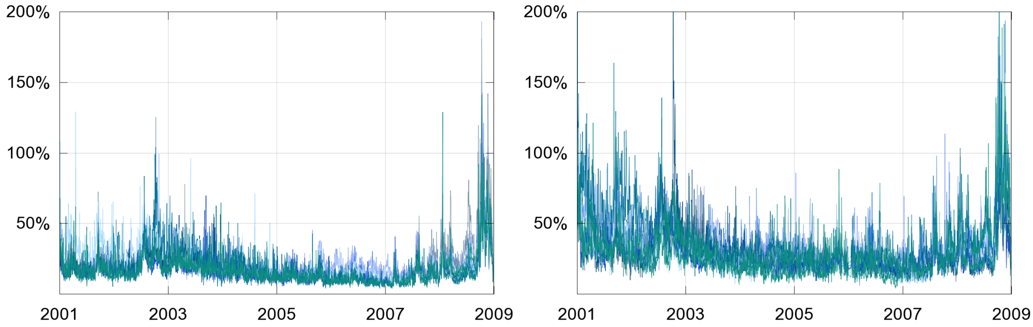
and where H is both the bandwidth of the kernel and the maximum order of the autocovariance, J is the number of 1–minute frequency returns within the day, and $\mathrm{K}_p(\cdot)$ denotes the Parzen kernel. Under appropriate conditions, Barndorff-Nielsen *et al.* (2008) show that the realized kernel estimator converges to the integrated variance of returns. The computation of the estimator and optimal choice of the bandwidth parameter for each series closely follows the guidelines described by Barndorff-Nielsen *et al.* (2009).

Figure 1 shows plots of the two panels of percent annualized volatility $\sqrt{252 x_{i\,t}}$; left for SPDR and right for S&P100. The plots suggest that the series cluster around a common time–varying average level that can be interpreted as *systematic volatility*. Statistical tests for the selection of the number of common factors in the panels strongly support the evidence of a one factor structure (see Luciani and Veredas (2011) for an exhaustive analysis, and Andersen *et al.* (2001) for stylized facts on a similar panel). The secular movements of systematic volatility can be attached to well known economic events or system wide innovations. The high level of volatility in the beginning of the 2000s is related to the aftermath of dot–com bubble burst and the recession. The period that goes from 2004 to July 2007 is characterized by a low level of uncertainty that corresponds to the market rally following the recession. Finally, volatility rises with the advent of the financial crisis

[1] The complete list of S&P100 tickers, company names and industry sectors is reported in Appendix B.

[2] The volatility measurement literature is indeed large: see, among others, Andersen *et al.* (2003), Aït-Sahalia *et al.* (2005), Bandi and Russell (2006), Alizadeh *et al.* (2002) for alternative estimators.

Figure 1: Realized volatilities



(a) SPDR

(b) S&P100

Annualized realized volatilities for the SPDR select sectors (left) and the 90 S&P100 constituents (right) from January 2, 2001 to December 31, 2008.

and it skyrockets to the highest level reached over the last 20 years in the fall of 2008, following the demise of Lehman Brothers.

Table 1 displays descriptive statistics. The table reports average percent annualized volatility, standard deviation of volatility (volatility of volatility), daily, weekly (5 days) and monthly (22 days) autocorrelations, average correlation at lag 0 with the other series, and the percentage of variance explained by the first principal component. For SPDR we report statistics for each sectoral index while for S&P100 we report the 25%, 50%, and 75% quantiles of the statistics across industry sectors. Mean and variability levels of the series are higher for the S&P100 panel rather then the SPDR, due to the fact the sectoral aggregation decreases the average and dispersion of volatility. Autocorrelations decay slowly, consistently with the evidence of long range dependence widely documented in volatility studies. The average cross-correlation with the other series and the proportion of variance explained by the first principal components are always above 0.50 and 50% respectively, which confirms the existence of strong co–movements in volatility.

3 A Seminonparametric Vector MEM

The empirical evidence of the previous section suggests that the dynamics of the volatility measures in the panel can be described by a common secular trend and residual idiosyncratic short run components. In this section we introduce a novel SeminonParametric Vector Multiplicative Error Model (SPvMEM) that captures these empirical regularities.

Let $x_{i\,t}$ be the value of the realized volatility measure for the i^{th} asset in period t, with $i = 1, ..., N$ and $t = 1, ..., T$, and let $z_t = t/T$ denote the (rescaled) time index. The realized measure $x_{i\,t}$ is modelled in a multiplicative specification of the form

$$x_{i\,t} = a_i\, \phi(z_t)\, \mu_{i\,t}\, \epsilon_{i\,t}, \quad \epsilon_{i\,t}|\mathcal{F}_{t-1} \sim D(1), \tag{1}$$

where $\mathcal{F}_{t-1}$ is the information set up to time $t-1$, a_i is the scale factor of the i^{th} series, $\phi(z_t)$ is a deterministic time trend, $\mu_{i\,t}$ is an idiosyncratic short term dynamic component, and $\epsilon_{i\,t}$ is a conditionally independent error term with positive support and unit expectation. More detailed assumptions on the process are given in the Appendix.

Table 1: Descriptive statistics

		vol	vov	$\hat{\rho}_{\mathsf{day}}$	$\hat{\rho}_{\mathsf{week}}$	$\hat{\rho}_{\mathsf{month}}$	$\bar{\rho}$	PC_1
				SPDR				
Mat		23.34	11.83	0.72	0.63	0.39	0.82	0.91
Ener		25.68	12.78	0.65	0.61	0.35	0.79	0.87
Fin		25.85	15.71	0.69	0.50	0.35	0.77	0.84
Ind		21.85	11.29	0.66	0.57	0.37	0.81	0.88
Tech		26.82	13.34	0.49	0.41	0.27	0.66	0.58
Stap		16.93	8.21	0.45	0.36	0.19	0.77	0.77
Util		24.12	12.73	0.65	0.55	0.35	0.72	0.70
Heal		17.27	8.61	0.34	0.27	0.17	0.73	0.68
Disc		20.56	10.57	0.64	0.55	0.35	0.82	0.90
				S&P100				
Mat	$q_{0.25}$	33.80	14.49	0.66	0.55	0.31	0.63	0.62
	$q_{0.50}$	34.73	15.17	0.69	0.56	0.34	0.67	0.72
	$q_{0.75}$	36.42	16.59	0.73	0.60	0.37	0.72	0.83
Ener	$q_{0.25}$	34.00	15.04	0.57	0.50	0.26	0.61	0.61
	$q_{0.50}$	37.98	16.36	0.66	0.57	0.32	0.68	0.72
	$q_{0.75}$	44.28	18.50	0.71	0.66	0.38	0.69	0.74
Fin	$q_{0.25}$	36.45	21.01	0.63	0.27	0.17	0.52	0.49
	$q_{0.50}$	39.77	23.77	0.66	0.47	0.28	0.64	0.70
	$q_{0.75}$	42.93	26.59	0.74	0.56	0.34	0.68	0.76
Ind	$q_{0.25}$	29.25	12.42	0.61	0.49	0.31	0.65	0.64
	$q_{0.50}$	30.61	13.44	0.64	0.54	0.33	0.69	0.73
	$q_{0.75}$	33.19	17.33	0.70	0.56	0.36	0.71	0.80
Tech	$q_{0.25}$	34.96	16.10	0.64	0.52	0.34	0.56	0.49
	$q_{0.50}$	38.53	17.85	0.67	0.58	0.37	0.62	0.58
	$q_{0.75}$	46.20	23.26	0.72	0.61	0.40	0.72	0.78
Util	$q_{0.25}$	28.24	13.58	0.67	0.45	0.25	0.60	0.54
	$q_{0.50}$	29.61	14.09	0.70	0.52	0.30	0.67	0.66
	$q_{0.75}$	31.42	15.20	0.72	0.60	0.34	0.70	0.73
Stap	$q_{0.25}$	25.46	11.43	0.46	0.36	0.22	0.60	0.52
	$q_{0.50}$	28.32	12.45	0.53	0.43	0.24	0.68	0.69
	$q_{0.75}$	29.80	13.30	0.58	0.47	0.30	0.71	0.76
Heal	$q_{0.25}$	29.19	13.02	0.48	0.34	0.22	0.64	0.62
	$q_{0.50}$	30.06	13.55	0.59	0.46	0.29	0.65	0.63
	$q_{0.75}$	32.95	15.00	0.61	0.50	0.34	0.66	0.67
Disc	$q_{0.25}$	32.98	15.49	0.58	0.47	0.30	0.63	0.60
	$q_{0.50}$	35.76	17.41	0.64	0.54	0.36	0.68	0.68
	$q_{0.75}$	41.13	21.01	0.67	0.58	0.41	0.70	0.76

The top part shows descriptive statistics for SPDR. The bottom part shows the same descriptive statistics for S&P100 with the assets grouped according to the same sectors as SPDR. For each group the table shows the 25, 50 and 75 quantiles. The columns report the average annualized volatility (vol), standard deviation of volatility, volatility of volatility (vov), the autocorrelations of order 1, 5 and 22 ($\hat{\rho}_{\mathsf{day}}$, $\hat{\rho}_{\mathsf{week}}$ and $\hat{\rho}_{\mathsf{month}}$), the average cross–correlation with the other series in the dataset ($\bar{\rho}$), and the percentage of the variance explained by the first principal component (PC_1).

The component $\phi(z_t)$ is a scalar smooth function capturing the low frequency common trend. It is assumed that $\phi : [0,1] \to \mathcal{P} \subset \mathbb{R}_+$ and that ϕ belongs to the set $\Gamma = \{p \in C^\infty[0,1] : p(z_t) \in \mathcal{P}$ for all $z_t \in [0,1]\}$. The use of $z_t = t/T$ as a regressor is a common assumption in models with time–varying parameters. The trend is normalized in a way such that $\int_0^1 \phi(u)\mathrm{d}u$ is equal to unity. In what follows, we denote ϕ without any reference to z_t as an infinite dimensional nuisance parameter belonging to Γ in the sense of Severini and Wong (1992). Our specification choice of the trend has important implications for estimation. In a rescaled time framework, i.e.

when $z_t = t/T$, as T increases the number of observations in a neighborhood of each point of the trend increases as well, and this allows to carry out pointwise inference on the trend.

The idiosyncratic component $\mu_{i\,t}$ is a nonnegative conditionally predictable scalar process with unit mean. It is defined as $\mu_{i\,t} = \mu_i(\mathcal{F}_{t-1}, \boldsymbol{\delta}_i)$, where $\boldsymbol{\delta}_i \in \mathcal{D}_i \subset \mathbb{R}^{p_\delta}$, is a vector of parameters characterizing the dynamics of the process. Several functional forms for $\mu_{i\,t}$ have been proposed in the MEM literature. Here, we opt for an asymmetric dynamics reminiscent of GARCH, that is

$$\mu_{i\,t} = \left(1 - \alpha_i - \beta_i - \frac{\gamma_i}{2}\right) + \alpha_i \frac{x_{i\,t-1}}{a_i \phi(z_{t-1})} + \beta_i \mu_{i\,t-1} + \gamma_i \frac{x_{i\,t-1}}{a_i \phi(z_{t-1})} \mathbf{1}_{\{r_{i\,t-1}<0\}}, \tag{2}$$

where $\alpha_i > 0$, $\gamma_i \geq 0$ and $\beta_i \geq 0$, and $r_{i\,t-1}$ denotes the return on day $t-1$ of asset i. Thus, in this case $\boldsymbol{\delta}_i = (\alpha_i, \beta_i, \gamma_i)^{\mathsf{T}}$ and $p_\delta = 3$. This type of functional form is often used in the literature (see Engle and Rangel, 2008) and it is typically found to be successful for prediction. We assume that the probability of $\{r_{i\,t-1} < 0\}$ is 1/2 and that $\mu_{i\,t}$ is stationary, that is $\alpha_i + \gamma_i/2 + \beta_i < 1$, so that the unconditional mean of (2) is one. Idiosyncratic dynamics can be equivalently parameterized in an alternative way for estimation convenience, as we will see later. Let $\tilde{\mu}_{i\,t} = a_i \mu_{i\,t}$, then

$$\tilde{\mu}_{i\,t} = \omega_i + \alpha_i \frac{x_{i\,t-1}}{\phi(z_{t-1})} + \beta_i \tilde{\mu}_{i\,t-1} + \gamma_i \frac{x_{i\,t-1}}{\phi(z_{t-1})} \mathbf{1}_{\{r_{i\,t-1}<0\}},$$

with $\omega_i = a_i \left(1 - \alpha_i - \beta_i - \frac{\gamma_i}{2}\right)$.

The conditional moments of the process are $\mathrm{E}[x_{i\,t}|\mathcal{F}_{t-1}] = a_i \phi(z_t) \mu_{i\,t}$ and $\mathrm{Var}[x_{i\,t}|\mathcal{F}_{t-1}] = a_i^2 \phi^2(z_t) \mu_{i\,t}^2 \mathrm{Var}(\epsilon_{i\,t})$. Thus, the SPvMEM is a conditionally heteroskedastic process where the conditional mean and variance change over time and are driven by the level of the common trend and of the idiosyncratic dynamics. On the other hand, the trend adjusted process $x_{i\,t}/\phi(z_t)$ follows a standard stationary MEM.

The model can be compactly expressed using vector notation. Let $\mathbf{x}_t = (x_{1\,t}, \ldots, x_{N\,t})^{\mathsf{T}}$ be the $N \times 1$ dimensional vector of volatility measures. Let $\mathbf{a} = (a_1, \ldots, a_N)^{\mathsf{T}} \in \mathcal{A} \subset \mathbb{R}_+^N$ and let the idiosyncratic dynamics be represented as a $N \times 1$ vector $\boldsymbol{\mu}(\mathcal{F}_{t-1}, \boldsymbol{\delta}) \subset \mathbb{R}_+^N$:

$$\boldsymbol{\mu}(\mathcal{F}_{t-1}, \boldsymbol{\delta}) = \boldsymbol{\mu}_t = \begin{pmatrix} \mu_1(\mathcal{F}_{t-1}, \boldsymbol{\delta}_1) \\ \mu_2(\mathcal{F}_{t-1}, \boldsymbol{\delta}_2) \\ \vdots \\ \mu_N(\mathcal{F}_{t-1}, \boldsymbol{\delta}_N) \end{pmatrix},$$

with $\boldsymbol{\delta} = (\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_N) \in \mathcal{D} \subset \mathbb{R}^{N p_\delta}$. The SPvMEM can be written as

$$\mathbf{x}_t = \phi(z_t) \cdot \mathbf{a} \odot \boldsymbol{\mu}_t \odot \boldsymbol{\epsilon}_t, \tag{3}$$

where $\boldsymbol{\epsilon}_t = (\epsilon_{1\,t}, ..., \epsilon_{N\,t})'$ and $\odot$ denotes the Hadamard component–wise product operator.

We complete the definition of the model described in (3) with the specification of the distribution of the vector of innovations $\boldsymbol{\epsilon}_t$. Its probability density function (pdf) and cumulative density function (cdf) are denoted by $f_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}_t; \boldsymbol{\theta})$ and $F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}_t; \boldsymbol{\theta})$ respectively, where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^{p_\theta}$.

Choosing an appropriate multivariate specification for the innovation term can be challenging in this setting as there are few multivariate distributions for positive real–valued random vectors. The multivariate exponential and Gamma are the two most prominent, but they are cumbersome and their properties may not always dovetail with those of the volatility measures. Building up on Cipollini *et al.* (2006), we follow a modelling strategy based on copulas. Let $F_{\epsilon_i}(\epsilon_{i\,t})$ denote the marginal cdf of $\epsilon_{i\,t}$ and define $u_{i\,t} = F_{\epsilon_i}(\epsilon_{i\,t})$ for $i = 1, \ldots, N$. Then, by Sklar's theorem, the joint cdf and pdf of $\boldsymbol{\epsilon}_t$ can be written respectively as

$$F_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}_t) = C(u_{1\,t}, \ldots, u_{N\,t}) \quad \text{and} \quad f_{\boldsymbol{\epsilon}}(\epsilon_t) = \prod_{i=1}^{N} f_{\epsilon_i}(\epsilon_{i\,t}) \cdot c(u_{1\,t}, \ldots, u_{N\,t}), \tag{4}$$

where $C(\cdot)$ is the copula function and $c(\cdot)$ is its derivative with respect to $(u_{1\,t}, \ldots, u_{N\,t})$. The advantage of copulas is that they allow to decompose the specification of the distribution in the choice of the marginal pdfs $f_{\epsilon_i}(\cdot)$ and the copula density $c(\cdot)$.

As far as the choice of the marginal distribution is concerned, we opt for a Gamma distribution. It is a flexible distribution, it belongs to the exponential family, and it nests the exponential, chi–square, Erlang, and Maxwell–Boltzmann distributions. If the marginal distribution of $\epsilon_{i\,t}$ is a Gamma distribution with parameters (k_i, ν_i), then the marginal conditional distribution of $x_{i\,t}$ is also Gamma $x_{i\,t}|\mathcal{F}_{t-1} \sim Gamma\left((a_i \phi(z_t)\mu_{i\,t})^{-1} k_i, \nu_i \right)$, with conditional pdf

$$f_{x_i}\left(x_{i\,t}; |\mathcal{F}_{t-1}\right) = \frac{k_i}{\Gamma(\nu_i) a_i \phi(z_t) \mu_{i\,t}} \left(\frac{x_{i\,t} k_i}{a_i \phi(z_t) \mu_{i\,t}} \right)^{\nu_i - 1} \exp\left(-\frac{x_{i\,t} k_i}{a_i \phi(z_t) \mu_{i\,t}} \right). \tag{5}$$

In what follows, we fix $k_i = \nu_i$ to ensure that $\epsilon_{i\,t}$ has unit mean, while $\text{Var}[\epsilon_{i\,t}] = \nu_i^{-1}$.

As far as the choice of the copula function is concerned, we adopt a Gaussian meta–copula, as delivers estimators that are easy to compute numerically in large dimensions:

$$c_\Phi(u_{1\,t}, \ldots, u_{N\,t}; \mathbf{R}) = |\mathbf{R}|^{-1/2} \text{exp} \left(-\frac{1}{2} \left(\Phi^{-1}(\mathbf{u}_t) \right)^{\mathsf{T}} (\mathbf{R} - \mathbf{I}) \left(\Phi^{-1}(\mathbf{u}_t) \right) \right) , \tag{6}$$

where $\Phi(\cdot)$ is the Gaussian cdf, $\mathbf{R}$ is the correlation matrix and $\mathbf{I}$ is the identity matrix. The marginal and copula parameters are collected in the vector $\boldsymbol{\theta}$ defined as $(\nu_1 \ldots \nu_N, \text{vech}(\mathbf{R}))^{\mathsf{T}}$, which has a dimension of $N + N(N-1)/2$.

Our choices of the marginal and copula function are primarily driven on the grounds of simplicity (Song (2000), Cipollini *et al.* (2006)). We acknowledge that the Gamma and Gaussian meta–copula have some limitations in fitting the moments of the data in empirical applications, and that it might be of interest adopting more sophisticated distributions. For instance, in some applications (e.g. financial derivatives based on volatilities), measuring tail dependences may be of interest.

Let us emphasize nevertheless that the theory developed in this paper carries through for different choices of the marginal densities and for any copula function. As for the marginals, any distribution belonging to the exponential family delivers simple closed form estimators for ϕ. Moreover, by Quasi Maximum Likelihood arguments (see Engle and Gallo, 2006), if the conditional moments are correctly specified the marginals deliver consistent estimates of the parameters. Concerning the copula choice, the theory developed in this work allows for general types of copulas, and inference under possible copula misspecification can be addressed by adapting Chen and Fan (2006) to our theory.

4 Estimation Procedure

The specification introduced in the previous section is a nonlinear model containing both parametric and nonparametric elements. Since joint estimation of both components is cumbersome, we propose a three step estimation procedure that stems from profile likelihood estimation (Severini

Table 2: SPvMEM Estimation Algorithm

Iteration 0

0.1 For each $t \in [0,T]$ define $z_t = t/T$ and compute the initial estimate:

$$\widehat{\phi}^0(z_\tau) = \frac{\sum_{t=1}^T \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right) \sum_{i=1}^N \frac{x_{i\,t}}{\bar{x}_i} \frac{1/s_i^2}{\sum_{i=1}^N 1/s_i^2}}{\sum_{t=1}^T \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)},$$

where $\bar{x}_i$ and s_i^2 are, respectively, sample mean and variance of the i^{th} series.

0.2 For each $i = 1, \ldots, N$ set $(\boldsymbol{\xi}_1^0, \ldots, \boldsymbol{\xi}_N^0)$ where $\boldsymbol{\xi}_i^0 = (a_i^0, \alpha_i^0, \beta_i^0, \gamma_i^0, \nu_i^0)$

Iteration $q > 0$

1 For each $i = 1, \ldots, N$ maximize the N log–likelihoods

$$\widehat{\boldsymbol{\xi}}_i^q = \arg\max_{\boldsymbol{\xi}_i} \sum_{t=1}^T \log f_{x_i}(x_{i\,t}; \boldsymbol{\xi}_i, \widehat{\phi}^{q-1} | \mathcal{F}_{t-1})$$

2 Compute the N idiosyncratic components evaluated at $\widehat{\phi}^{q-1}$

$$\widehat{\mu}_{i\,t}^q = \left(1 - \widehat{\alpha}_i^q - \widehat{\beta}_i^q - \frac{\widehat{\gamma}_i^q}{2}\right) + \widehat{\alpha}_i^q \frac{x_{i\,t-1}}{\widehat{a}_i^q \widehat{\phi}^{q-1}(z_{t-1})} + \widehat{\beta}_i^q \hat{\mu}_{i\,t-1}^q + \widehat{\gamma}_i^q \frac{x_{i\,t-1}}{\widehat{a}_i^q \widehat{\phi}^{q-1}(z_{t-1})} \mathbf{1}_{r_{i\,t-1}<0}.$$

For each $z_t \in [0,1]$, given $\hat{\mu}_{i\,t}^q$, compute:

$$\widehat{\phi}^q(z_t) = \frac{\sum_{t=1}^T \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right) \sum_{i=1}^N \frac{\widehat{\nu}_i}{\sum_{i=1}^N \widehat{\nu}_i} \frac{x_{i\,t}}{\widehat{a}_i^q \widehat{\mu}_{i\,t}^q}}{\sum_{t=1}^T \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)}.$$

3 Check for convergence otherwise go back to 1

Copula Estimation

1 The Gaussian meta-copula is estimated with the sample covariance matrix of $\Phi^{-1}(\hat{u}_{i\,t})$ where $\hat{u}_{i\,t} = \widehat{F}_{\epsilon_i}(\hat{\epsilon}_{i\,t})$

and Wong, 1992) and inference from the marginals (Joe, 2005): i) nonparametric estimation of the common trend ϕ; ii) estimation of the marginal parameters $\boldsymbol{\xi}_i = (a_i, \boldsymbol{\delta}_i^{\mathrm{T}}, \nu_i)^{\mathrm{T}}$ of each series; and iii) estimation of the copula parameter $\boldsymbol{\psi} = \mathrm{vech}(\mathbf{R})$. We know explain these steps in detail.

i) Conditionally on the idiosyncratic dynamics $\mu_{i\,t}$ and the marginal parameters $\boldsymbol{\xi}_i$, a natural estimator of the common trend is a Nadaraya–Watson type estimator applied to the weighted average of the rescaled series. That is, for any $z_\tau \in [0,1]$,

$$\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) = \frac{\sum_{t=1}^T \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right) \sum_{i=1}^N \frac{x_{i\,t}}{a_i \mu_{i\,t}} \frac{\nu_i}{\sum_{i=1}^N \nu_i}}{\sum_{t=1}^T \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)}, \tag{7}$$

where $\mathrm{K}(\cdot)$ is a suitable kernel function with a bandwidth h_{NT} that can vary both with N and T (see the next section for details). This estimator has an intuitive meaning: the common trend at z_τ is estimated as a nonparametric regression of a weighted sum (across N) of $x_{i\,t}$ adjusted by the idiosyncratic component $a_i \mu_{i\,t}$, where the weights are $\nu_i / \sum_{i=1}^N \nu_i$. Since ν_i is the reciprocal of the variance of the i^{th} innovation, the weights also have an intuitive interpretation: the least erratic series $x_{i\,t}$ (denoted by larger ν's) receives more weight in the estimation of the common trend. The technical justification for this weighting scheme is given in the next section when the estimator is derived. Other weighting schemes are also possible, e.g. equal weights $1/N$. As far as the choice of the kernel function is concerned, it is important to stress that such a choice ought to be based on the purpose of the application of the SPvMEM. If the objective of the application is in–sample estimation then a two–sided kernel is more suitable, and the next section develops the related asymptotic properties of the estimated curve in this setting. On the other hand, if the objective of the application is out–of–sample forecasting, then one should resort to a one–sided kernel (see Gijbels *et al.*, 1999).

$ii)$ Conditionally on the common component ϕ, we can adjust each series for the level of the common trend. From the properties of the Gamma distribution, it follows that

$$\left.\frac{x_{i\,t}}{\phi(z_t)}\right| \mathcal{F}_{t-1} \sim Gamma\left((a_i\mu_{i\,t})^{-1}\nu_i, \nu_i\right). \tag{8}$$

Thus, given an estimator of the common trend, we can estimate the marginal parameter vector $\boldsymbol{\xi}_i = (a_i, \boldsymbol{\delta}_i^{\mathrm{T}}, \nu_i)^{\mathrm{T}}$ by maximising the marginal log–likelihood function of the trend adjusted series associated to (8), for each series in the panel. In this way we obtain the estimator $\widehat{\boldsymbol{\xi}}_{i\,T}$.

$iii)$ Finally, conditionally on the first two steps, the correlation $\mathbf{R}$ across the innovations can be estimated as the sample covariance matrix of the transformed marginal cdfs $\Phi^{-1}(u_{i\,t}) = \Phi^{-1}(F_{x_i}(x_{i\,t}; \boldsymbol{\xi}_i, \phi|\mathcal{F}_{t-1}))$.

These considerations motivate us to propose an iterative procedure to estimate the SPvMEM applying iteratively the first and second steps until convergence, and deriving the copula parameters after that.

As detailed in Table 2, the procedure is initialized as follows. The initial estimate of the trend requires an educated guess of $a_i\mu_{i\,t}$ and ν_i, for all i (step 0.1). As for $a_i\mu_{i\,t}$, since it is the conditional mean of the i^{th} series adjusted for the common component, we consider the sample means $\bar{x}_i$. The values of the marginal parameters are also set to an educated initial guess (step 0.2). The parameters α_i^0, β_i^0 and γ_i^0 are set respectively to 0.1, 0.8 and 0.1, while a_i is set to $(1 - \alpha_i^0 - \beta_i^0 - \gamma_i^0)\bar{x}_i$.As for ν_i, since it is the inverse of the variance of a Gamma distribution, we consider the inverse of the sample variance s_i^2 (see Hafner and Linton, 2010). After these initializations, the estimation algorithm proceeds as follows. The estimated parameters at iteration q are obtained by maximizing the N log–likelihood functions using an estimate of the curve obtained at iteration $q-1$ (step 1). Next, using the updated estimates of $\mu_{i\,t}$, we update the common trend estimation (step 2). Steps 1 and 2 are repeated until convergence of the estimated parameters (typically achieved within a few iterations). Finally, the Gaussian meta–copula is estimated.

5 Asymptotic Theory

This section establishes the asymptotic properties of our estimators. We establish consistency and asymptotic Gaussianity of the common trend and the parameters. The asymptotic framework considered in this section is developed for $T \to \infty$, and both N fixed and $N \to \infty$. All the assumptions and proofs are in the Appendix. An interesting feature of our setting is that consistent estimation of the common trend does not necessarily require a large cross–section. As the sample size T increases, the number of observations in the neighborhood of each point of the trend increases as well, and this allows for consistent inference. However, we show in this section that, when also the cross–sectional dimension N is allowed to increase to infinity, consistency is still achieved and some of the inferential procedures simplify. This result is obtained if we add two additional conditions, on the choice of the kernel bandwidth and on the dependence captured by the copula. In particular, we impose a smaller bandwidth and assume weak cross–sectional dependence among the innovations. This latter condition is the analog of the condition imposed on the dependence between idiosyncratic components in linear models (see e.g. Bai and Ng (2002)).

We adopt the following notation for the log–likelihood function of the SPvMEM. The model contains an infinite dimensional nuisance parameter ϕ and a $5N + N(N-1)/2$–dimensional vector of parameters $\boldsymbol{\eta} = (\mathbf{a}^{\mathrm{T}}, \boldsymbol{\delta}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}}$. The log–likelihood of observation $\mathbf{x}_t$ at time t conditional on the information set $\mathcal{F}_{t-1}$ is defined as

$$\boldsymbol{\ell}_t(\boldsymbol{\eta}, \phi) = \log f_{\mathbf{x}}(\mathbf{x}_t; \boldsymbol{\eta}, \phi|\mathcal{F}_{t-1}), \tag{9}$$

using the notation $\boldsymbol{\ell}_t(\boldsymbol{\eta}, \phi(z_t))$ when necessary. It is convenient to rearrange the parameter vector $\boldsymbol{\eta}$. We collect the parameters of the marginals in a $5N$–dimensional vector $\boldsymbol{\xi} = (\boldsymbol{\xi}_1^{\mathrm{T}} \dots \boldsymbol{\xi}_N^{\mathrm{T}})^{\mathrm{T}}$, where $\boldsymbol{\xi}_i = (a_i, \boldsymbol{\delta}_i^{\mathrm{T}}, \nu_i)^{\mathrm{T}} \in \boldsymbol{\Xi}_i \subset \mathbb{R}^{5N}$ and such that $\boldsymbol{\Xi}_i \cap \boldsymbol{\Xi}_j = \emptyset$ for $i \neq j$. The parameters of the copula are collected in a vector $\boldsymbol{\psi} \in \boldsymbol{\Psi} \subset \mathbb{R}^{p_\psi}$. It follows from the joint pdf of the innovations in equation (4) and Sklar's theorem that the conditional log–likelihood can be expressed as

$$\mathcal{L}_{NT}(\boldsymbol{\eta}, \phi) = \sum_{t=1}^{T} \sum_{i=1}^{N} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi) + \sum_{t=1}^{T} \ell_t^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi), \tag{10}$$

where

$$\begin{aligned} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi) &= \log f_{x_i}(x_{it}; \boldsymbol{\xi}_i, \phi|\mathcal{F}_{t-1}), \quad \text{and} \\ \boldsymbol{\ell}_t^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi) &= \log c(u_{1t}, \dots, u_{Nt}; \boldsymbol{\psi}|\mathcal{F}_{t-1}). \end{aligned}$$

Again, we also use the notation $\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi(z_t))$ and $\boldsymbol{\ell}_t^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi(z_t))$ when necessary.

In the sequel we use a more detailed notation. We denote by $\boldsymbol{\eta}_0 = (\boldsymbol{\xi}_{10}, \dots, \boldsymbol{\xi}_{N0}, \boldsymbol{\psi}_0)$ the true values of the parameters, by ϕ_0 we indicate the true curve, and we denote by E_0 the expectation taken under the true model.

In an $i.i.d.$ univariate context, and based on Staniswalis (1989), Severini and Wong (1992) propose and prove the asymptotic properties of an estimator of ϕ_0 based on smoothed profile log–likelihood maximization. This estimator has been generalized to the univariate dependent case by Veredas *et al.* (2007). In the multivariate context of the SPvMEM, this technique requires an estimator of $\boldsymbol{\eta}_0$, which in principle we could obtain by maximizing the joint log–likelihood (9). The procedure to estimate ϕ_0 and $\boldsymbol{\eta}_0$ would be based on iterating between two optimizations: on the one hand the global and parametric optimization with respect to the $5N$ parameters in $\boldsymbol{\xi}_0 = (\boldsymbol{\xi}_{10}, \dots, \boldsymbol{\xi}_{N0})$ plus the p_ψ parameters in $\boldsymbol{\psi}_0$ and, on the other hand, and at each iteration, the optimization of the localized (or smoothed) log–likelihood (to be performed T times). This procedure has two main drawbacks. First, it is computationally intensive, and, second, it requires an expression for the joint log–likelihood. Even for small values of N, this approach seems unfeasible, but it can be simplified by means of inference from the marginals, which leads to the estimator introduced in the previous section.

We first show how to derive the estimator of the common trend ϕ_0, followed by the analysis of the effect of the nonparametric component on the inference on $\boldsymbol{\xi}_0$. In Theorems 1–2 we show the asymptotic properties of $\widehat{\phi}_{\boldsymbol{\xi}_o NT}$, defined in (7), under the true parameters of the marginals $\boldsymbol{\xi}_0$. In

Theorem 3 and Corollary 1 we study the relation between the estimator of the curve and a generic value of the parameters $\boldsymbol{\xi}$. In Theorem 4 we show consistency and the asymptotic distribution of $\widehat{\boldsymbol{\xi}}_T$. Finally, in Theorem 5 we prove consistency of $\widehat{\boldsymbol{\psi}}_T$. Theorems 1–4 are proved in the cases of i) fixed N and large T and ii) both N and T large.

5.1 Estimation of the common trend

Given the marginal likelihoods, for any $z_\tau \in [0,1]$ and for a given value of the parameters $\boldsymbol{\xi}$, the estimator of the curve is such that

$$\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) = \arg\sup_{\phi\in\Gamma} \sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right) \sum_{i=1}^{N} \boldsymbol{\ell}_{i\,t}^{m}(\boldsymbol{\xi}_i, \phi(z_t)). \tag{11}$$

Here $\mathrm{K}(\cdot)$ is a suitable kernel function satisfying assumption K in the Appendix, and h_{NT} is a bandwidth parameter that can depend on both N and T, as explained in Theorem 2 below. Given the choice of Gamma distributions for the marginals, this optimization has a closed form solution:

$$\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) = \frac{\sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right) \sum_{i=1}^{N} \frac{x_{i\,t}}{a_i \mu_{i\,t}} \frac{\nu_i}{\sum_{i=1}^{N} \nu_i}}{\sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)}. \tag{12}$$

This is the estimator proposed in (7) and this derivation shows that the weighting scheme based on the innovation variances is actually a direct consequence of the maximization in (11).

The following Theorems show consistency and asymptotic Gaussianity of $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)$, i.e. of the estimator in (12) when $\boldsymbol{\xi} = \boldsymbol{\xi}_0$.

Theorem 1 – Nonparametric component – Consistency *Given the estimator $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)$ and under assumptions A, B, C.1, I, K, L in the Appendix, and if $NTh_{NT} \to \infty$ and $h_{NT} \to 0$ as $NT \to \infty$ we have:*

a) if $T \to \infty$ and N is finite, then, for any $z_\tau \in [0,1]$, $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau) \xrightarrow{P} \phi_0(z_\tau)$;

b) if $T \to \infty$ and $N \to \infty$ and assumption C.2 in the Appendix holds, the result in part a) still holds.

Proof. The proof of the theorem is in the Appendix.

Theorem 2 – Nonparametric component – Asymptotic Gaussianity *Under assumptions A, B, C.1, I, K, L, P in the Appendix, and if $NTh_{NT} \to \infty$ as $NT \to \infty$ and one of the following conditions holds true:*

a) $T \to \infty$, N finite, and $h_{NT} \to 0$;

b) $T \to \infty$, $N \to \infty$, $Nh_{NT}^2 \to 0$, and assumptions C.2 and D in the Appendix hold;

then, the estimator $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)$ in (12) is such that

$$\sqrt{NTh_{NT}}\left(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau) - \phi_0(z_\tau)\right) \xrightarrow{d} \mathcal{N}\left(0, V_{\boldsymbol{\xi}_o}(z_\tau)\right),$$

where, in case a)

$$
\begin{aligned}
V_{\boldsymbol{\xi}_o}(z_\tau) &= \left(\int_{-1}^{1}\mathrm{K}^2(u)\mathrm{d}u\right)\frac{\bar{\imath}_{N\,\boldsymbol{\xi}_o}(z_\tau)}{\bar{\jmath}^2_{N\,\boldsymbol{\xi}_o}(z_\tau)},\\
\bar{\imath}_{N\,\boldsymbol{\xi}_o}(z_\tau) &= \frac{1}{N}\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,o},\phi_0(z_\tau))\right)^2\right],\\
\bar{\jmath}_{N\,\boldsymbol{\xi}_o}(z_\tau) &= -\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,o},\phi_0(z_\tau))\right],
\end{aligned}
$$

and in case b)

$$
\begin{aligned}
V_{\boldsymbol{\xi}_o}(z_\tau) &= \left(\int_{-1}^{1}\mathrm{K}^2(u)\mathrm{d}u\right)\frac{i_{\boldsymbol{\xi}_o}(z_\tau)}{j^2_{\boldsymbol{\xi}_o}(z_\tau)},\\
i_{\boldsymbol{\xi}_o}(z_\tau) &= \lim_{N\to\infty}\bar{\imath}_{N\,\boldsymbol{\xi}_o}(z_\tau),\quad j_{\boldsymbol{\xi}_o}(z_\tau)=\lim_{N\to\infty}\bar{\jmath}_{N\,\boldsymbol{\xi}_o}(z_\tau).
\end{aligned}
$$

Proof. The proof of the theorem is in the Appendix.

Some remarks on these theorems are in order. First, the estimator $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}$ converges to the true curve ϕ_0 by virtue of Lemma 1 in the Appendix, which proves that ϕ_0 is a maximizer not only of the localized version of the joint log–likelihood (10), but also of each localized marginal log–likelihood. Second, with respect to Veredas *et al.* (2007), asymptotic efficiency is lost since we are neglecting the copula part of the log–likelihood (10). As for the effect of large N, we emphasize two results. In the limit $N,T\to\infty$, the bandwidth has to decrease at a faster rate, i.e. the local averages have to be computed using a smaller window of observations when the cross–section is large in comparison to when it is finite. Moreover the term $i_{\boldsymbol{\xi}_o}(z_\tau)$ in the asymptotic variance contains not only the variances of the marginal scores but also their covariances. Thus the asymptotic variance is well defined also in the limit $N\to\infty$, provided that we assume these covariances to be bounded (see assumption D). In particular, we have[3]

$$
\begin{aligned}
\bar{\imath}_{N\,\boldsymbol{\xi}_o}(z_\tau) &= \frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\left(\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,o},\phi_0(z_\tau))\right)^2\right]+\\
&\quad+\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N}\left(\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,o},\phi_0(z_\tau))\right)\left(\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{j\,t}(\boldsymbol{\xi}_{j\,o},\phi_0(z_\tau))\right)\right]=\\
&= \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}_0\left[\frac{\nu^2_{i0}(\epsilon_{i\,t}-1)^2}{\phi_0^2(z_\tau)}\right]+\frac{1}{N}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N}\mathrm{E}_0\left[\frac{\nu_{i0}\nu_{j0}(\epsilon_{i\,t}-1)(\epsilon_{j\,t}-1)}{\phi_0^2(z_\tau)}\right].\quad(13)
\end{aligned}
$$

By recalling that $\mathrm{E}_0[\epsilon_{i\,t}]=1$, we see that the first term of (13) is proportional to the variances of the innovations, while the second is proportional to their covariance. Under assumption C.2, the first term is bounded for any N, while the second one diverges with N as there are $N(N-1)/2$ covariance terms. By assumption D, the second term of (13) is bounded for any N. Therefore, in large panels we require the cross–sectional dependence among the innovations to be weak. Similar conditions for the idiosyncratic components are made in the context of approximate factor models

[3] We denote by $\partial/\partial\phi$ the Fréchet functional derivative.

(see Bai and Ng (2002)). Assumption D is also supported by the empirical evidence (see Section 7). We also note that analogous results are established by Li and Racine (2006) in additive panel models with a nonparametric component.

In order to carry out inference, the expressions in $V_{\boldsymbol{\xi}_o}(z_\tau)$ have to be replaced by their sample analogues. The integral of the squared kernel is a kernel–specific constant (e.g. it is 1 if we use a Gaussian kernel, $5/7$ for a quartic kernel). Then, using the estimators of the parameters $\widehat{a}_{iT}, \widehat{\boldsymbol{\delta}}_{iT}, \widehat{\nu}_{iT}$, given in Theorem 4 below, we define the estimated residuals

$$\widehat{\epsilon}_{it} = \frac{x_{it}}{\widehat{a}_{iT}\mu_{it}(\mathcal{F}_{t-1}, \widehat{\boldsymbol{\delta}}_{iT})\widehat{\phi}_{NT\,\widehat{\boldsymbol{\xi}}_T}(z_t)}.$$

The localized Fisher information $\bar{i}_{N\,\boldsymbol{\xi}_o}(z_\tau)$ can be estimated, for any $z_\tau \in [0,1]$, by using the sample counterpart of (13):

$$\widehat{\bar{i}}_{NT\,\widehat{\boldsymbol{\xi}}_T}(z_\tau) \;=\; \frac{\sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\left[\sum_{i=1}^{N} \frac{\widehat{\nu}_{iT}}{\widehat{\phi}_{NT\,\widehat{\boldsymbol{\xi}}_T}(z_t)}\left(\widehat{\epsilon}_{it} - 1\right)\right]^2}{N\sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)}. \tag{14}$$

Analogously, $\bar{j}_{N\,\boldsymbol{\xi}_o}(z_\tau)$, is estimated by

$$\widehat{\bar{j}}_{NT\,\widehat{\boldsymbol{\xi}}_T}(z_\tau) \;=\; \frac{\sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\sum_{i=1}^{N} \frac{\widehat{\nu}_{iT}}{\widehat{\phi}^2_{NT\,\widehat{\boldsymbol{\xi}}_T}(z_t)}\left(2\,\widehat{\epsilon}_{it} - 1\right)}{N\sum_{t=1}^{T} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)}. \tag{15}$$

Alternatively, the denominators in (14) and (15) can be substituted by NTh_{NT}. Proofs of the consistency of these estimators are in Müller (1984).

Prior to turning to the estimation of the parameter vector $\boldsymbol{\xi}$ in the next section, we study its relationship with the curve ϕ. Let us first consider the case $N = 1$, and focus on $\boldsymbol{\xi}_{1\,0}$. In the seminonparametric setting, the Fisher information matrix for $\boldsymbol{\xi}_{1\,0}$ is the sum of a parametric component minus a correction due to the presence of the curve:

$$\mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{1t}}{\partial \boldsymbol{\xi}_1}(\boldsymbol{\xi}_{1\,0}, \phi_0)\frac{\partial \boldsymbol{\ell}^m_{1t}}{\partial \boldsymbol{\xi}_1^{\mathrm{T}}}(\boldsymbol{\xi}_{1\,0}, \phi_0)\right] - \mathbf{v}\,\mathrm{E}_0\left[\left(\frac{\partial \boldsymbol{\ell}^m_{1t}}{\partial \phi}(\boldsymbol{\xi}_{1\,0}, \phi_0)\right)^2\right]\mathbf{v}^{\mathrm{T}}, \tag{16}$$

where $\mathbf{v}$ is a generic vector of the same size as $\boldsymbol{\xi}_{1\,0}$. Note that, since the second term in (16) is positive definite, we have a smaller Fisher information with respect to the fully parametric case. We define a *least favorable direction* $\mathbf{v}^*$ as the minimizer of the seminonparametric Fisher information matrix (16) over all possible directions $\mathbf{v}$. Severini and Wong (1992) prove that an estimator of the curve having as tangent vector the least favorable direction, and called *least favorable curve*, delivers an unbiased estimator of the parameter $\boldsymbol{\xi}_{1\,0}$. It can be shown that the explicit form of the least favorable direction is

$$\mathbf{v}^* = -\frac{\mathrm{E}_0\left[\frac{\partial^2 \boldsymbol{\ell}^m_{1t}}{\partial \boldsymbol{\xi}_1 \partial \phi}(\boldsymbol{\xi}_{1\,0}, \phi_0)\right]}{\mathrm{E}_0\left[\frac{\partial^2 \boldsymbol{\ell}^m_{1t}}{\partial \phi^2}(\boldsymbol{\xi}_{1\,0}, \phi_0)\right]}. \tag{17}$$

Since all marginals depend on the curve, when $N > 1$, in (17) we have to substitute $\boldsymbol{\ell}^m_{1t}$ with $\sum_i \boldsymbol{\ell}^m_{it}/N$. We denote the least favorable curve in a generic value of the parameters as $\phi_{\boldsymbol{\xi}}$ and we require it to satisfy the regularity assumption L in the Appendix. The tangent vector to this curve computed at the true value of the parameters (i.e., the first derivative of the least favorable curve with respect to the parameters) is then defined as

$$\phi'_{\boldsymbol{\xi}_o} = -\frac{\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^N \frac{\partial^2 \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}\partial\phi}(\boldsymbol{\xi}_{i\,0},\phi_0)\right]}{\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^N \frac{\partial^2 \boldsymbol{\ell}^m_{it}}{\partial \phi^2}(\boldsymbol{\xi}_{i\,0},\phi_0)\right]} \equiv -\frac{\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}}{\bar{j}_{N\,\boldsymbol{\xi}_o}}. \tag{18}$$

In the following Theorem we show the asymptotic properties of the vector tangent to the estimated curve $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}$.

Theorem 3 – Least Favorable Direction *Given the estimator $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)$ and under assumptions A, B, C.1, I, K, L, S in the Appendix, and if $NTh_{NT} \to \infty$ and $h_{NT} \to 0$ as $NT \to \infty$ we have:*

a) if $T \to \infty$ and N is finite, then, for any $z_\tau \in [0,1]$, $\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}(z_\tau) \xrightarrow{P} \phi'_{\boldsymbol{\xi}_o}$, where $\phi'_{\boldsymbol{\xi}_o}$ is defined in (18)*;*

b) if $T \to \infty$ and $N \to \infty$ and assumption C.2 in the Appendix holds, then, for any $z_\tau \in [0,1]$, $\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}(z_\tau) \xrightarrow{P} \mathbf{0}$.

Proof. The proof of the theorem is in the Appendix.

When N is large the tangent vector to the curve becomes increasingly smaller. Intuitively, this is a consequence of the fact that when $N \to \infty$ the nonparametric and the parametric components are asymptotically uncorrelated.

Last, we show the asymptotic properties of the estimated curve for any value for the parameters. Indeed, $\widehat{\boldsymbol{\xi}}_T$ and $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$ depend on each other. To break this feedback loop, in the following Corollary we show that, for any given value of the parameters $\boldsymbol{\xi}$, the estimator defined in (12) is an estimator of a least favorable curve, a result which is necessary to prove consistency of the estimated parameters in Theorem 4. We prove the Corollary only for finite N as, when $N \to \infty$, the effect of the nuisance parameter ϕ becomes negligible (see part b of Theorem 3).

Corollary 1 – Least Favorable Curve *Under the same Assumptions of Theorem 3.a, for any $\boldsymbol{\xi} \in \Xi$ and any $z_\tau \in [0,1]$, $\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) \xrightarrow{P} \phi_{\boldsymbol{\xi}}(z_\tau)$.*

Proof. The proof of the corollary is in the Appendix.

The first part of the proof of the Corollary relies on assumption N in the Appendix about the existence of a limiting curve (and its derivatives) for any value of the parameters. In a second part, we show that for any $\boldsymbol{\xi}$, the nonparametric estimator $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$ converges to the least favorable curve $\phi_{\boldsymbol{\xi}}$, which is done in Lemma 4.

5.2 Estimation of the parameters

We now turn to the estimation of the marginals and copula parameters, $(\boldsymbol{\xi}_{1\,0},\ldots,\boldsymbol{\xi}_{N\,0})$ and $\boldsymbol{\psi}$. Since $\Xi_i \cap \Xi_j = \emptyset$ for $i \neq j$, estimation of $(\boldsymbol{\xi}_{1\,0},\ldots,\boldsymbol{\xi}_{N\,0})$ boils down to N independent optimizations of the marginal log–likelihoods,

$$\widehat{\boldsymbol{\xi}}_{i\,T} = \arg\max_{\boldsymbol{\xi}_i \in \Xi_i} \sum_{t=1}^T \boldsymbol{\ell}^m_{it}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}), \quad i = 1,\ldots,N. \tag{19}$$

The asymptotic properties are proved in the following Theorem.

Theorem 4 – Parameters of the Marginals *Consider the estimator of a least favorable curve in* (12) *and for any* $i = 1, \ldots, N$, *let* $\widehat{\boldsymbol{\xi}}_{iT}$ *be the vector of parametric estimates in* (19)*, then under assumptions A, B, C.1, I, L, P, S, in the Appendix, we have*

a) if $T \to \infty$, *then* $\widehat{\boldsymbol{\xi}}_{iT} \xrightarrow{P} \boldsymbol{\xi}_{i0}$;

b) if $T \to \infty$ *and* N *is finite, then*

$$\sqrt{T}\left(\widehat{\boldsymbol{\xi}}_T - \boldsymbol{\xi}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, (\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}\, \mathbf{I}^*_{\boldsymbol{\xi}_o}\, (\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}\,\right),$$

where

$$\mathbf{I}^*_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \ldots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} - \left(\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}\bar{\boldsymbol{d}}^T_{N\boldsymbol{\xi}_o}\right) \otimes \frac{\bar{\imath}_{N\boldsymbol{\xi}_o}}{\bar{\jmath}^2_{N\boldsymbol{\xi}_o}},$$

$$\mathbf{H}^*_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{H}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \ldots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathcal{H}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} - \left(\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}\bar{\boldsymbol{d}}^T_{N\boldsymbol{\xi}_o}\right) \otimes \frac{1}{\bar{\jmath}_{N\boldsymbol{\xi}_o}},$$

where $\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}$ *is the numerator of* (18), $\bar{\jmath}_{N\boldsymbol{\xi}_o}$ *is defined in Theorem 2, and, for any* $i = 1, \ldots, N$,

$$\mathcal{I}_{\boldsymbol{\xi}_{io}\boldsymbol{\xi}_{io}} = \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i0}, \phi_0)\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}^T_i}(\boldsymbol{\xi}_{i0}, \phi_0)\right], \quad \mathcal{H}_{\boldsymbol{\xi}_{io}\boldsymbol{\xi}_{io}} = -\mathrm{E}_0\left[\frac{\partial^2 \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}^T_i}(\boldsymbol{\xi}_{i0}, \phi_0)\right].$$

c) if $T \to \infty$ *and* $N \to \infty$ *and under assumptions C.2 and D in the Appendix, then*

$$\sqrt{T}\left(\widehat{\boldsymbol{\xi}}_T - \boldsymbol{\xi}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{H}^{-1}_{\boldsymbol{\xi}_o}\, \mathbf{I}_{\boldsymbol{\xi}_o}\, \mathbf{H}^{-1}_{\boldsymbol{\xi}_o}\,\right),$$

where

$$\mathbf{I}_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \ldots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix}, \quad \mathbf{H}_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{H}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \ldots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathcal{H}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix}.$$

Proof. The proof of the theorem is in the Appendix.

For N fixed, consistency and asymptotic Gaussianity of the estimated parameters hold, provided we have a consistent estimator of ϕ_0 satisfying assumption N. Theorem 3 guarantees that this is the case for $\widehat{\phi}_{\boldsymbol{\xi}NT}$, even when $N \to \infty$. Notice that the parameter estimation of the marginals is based on N univariate maximizations, thus N plays no role in computing the first term of $\mathbf{I}^*_{\boldsymbol{\xi}_o}$ and $\mathbf{H}^*_{\boldsymbol{\xi}_o}$. The asymptotic covariance matrix is larger than the parametric lower bound. This is due to two reasons. First, since we use the marginals to estimate $\boldsymbol{\xi}_0$, we obtain the usual sandwich form as in Quasi Maximum Likelihood estimation. Second, as explained in the previous section, the presence of the curve modifies both the information matrix and the Hessian, and it is straightforward to see that the asymptotic covariance matrix is greater than or equal to the seminonparametric lower bound, i.e. $(\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}\mathbf{I}^*_{\boldsymbol{\xi}_o}(\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1} \succeq (\mathbf{I}^*_{\boldsymbol{\xi}_o})^{-1}$. The correction term $\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_o}$ is not block diagonal, as the curve is contained in all the marginal distributions. Also, by Theorem 3, this correction term is $O(N^{-2})$. Thus, if we let $N \to \infty$ this term becomes negligible and the seminonparametric asymptotic covariance converges to its parametric counterpart, i.e. to $\mathbf{H}^{-1}_{\boldsymbol{\xi}_o}\, \mathbf{I}_{\boldsymbol{\xi}_o}\, \mathbf{H}^{-1}_{\boldsymbol{\xi}_o}$.

In short, we see the advantage of inference from the marginals: estimation reduces to a simple iterative process between a closed form estimator, $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$, and N univariate optimizations with respect to five parameters each ($\boldsymbol{\xi}_i = (a_i, \alpha_i, \beta_i, \gamma_i, \nu_i)$).

Finally, given the estimators $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$ and $\widehat{\boldsymbol{\xi}}_T = (\widehat{\boldsymbol{\xi}}^{\mathrm{T}}_{1\,T}, \ldots, \widehat{\boldsymbol{\xi}}^{\mathrm{T}}_{N\,T})^{\mathrm{T}}$ obtained in (12) and (19), we estimate $\boldsymbol{\psi}_0$ by maximizing the copula log–likelihood:

$$\widehat{\boldsymbol{\psi}}_T = \arg\max_{\boldsymbol{\psi}\in\,\boldsymbol{\Psi}} \sum_{t=1}^{T} \boldsymbol{\ell}^c_t(\widehat{\boldsymbol{\xi}}_T, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}). \tag{20}$$

Consistency of this estimator is proved in the following Theorem.

Theorem 5 – Parameters of the Copula *Consider the estimator of a least favorable curve in* (12) *and the estimators of the marginals' parameters* $\widehat{\boldsymbol{\xi}}_{i\,T}$*, for* $i = 1, \ldots, N$*, in* (19)*, let* $\widehat{\boldsymbol{\psi}}_T$ *be the vector of estimators in* (20)*, then under assumptions A, B, C.1, I, L, P in Appendix A, if* $T \to \infty$*,* $\widehat{\boldsymbol{\psi}}_T \overset{P}{\to} \boldsymbol{\psi}_0$*.*

Proof. The proof of the theorem is in the appendix.

6 Monte Carlo Study

We carry out a simulation study to investigate the finite sample properties of estimators, and to assess the reliability of the asymptotic standard errors estimators for inference. We chose to simulate $x_{i\,t}$ directly from the SPvMEM model, using univariate MEMs for the idiosyncratic components and then linking them through the systematic component. To do so, we also specify a DGP for the signs of the daily returns of each assets, deriving them from iid Bernoulli random deviates to reproduce equi–probability of negative and positive signs. The alternative approach, not pursued here, would be to simulate a panel of high frequency prices from a continuous time model with stochastic volatility and leverage, constructing the high frequency returns $r_{i\,tj}$ by discretely sampling such process, and then finally constructing the realized measures $x_{i\,t}$, which would then need to be linked together via the systematic component.

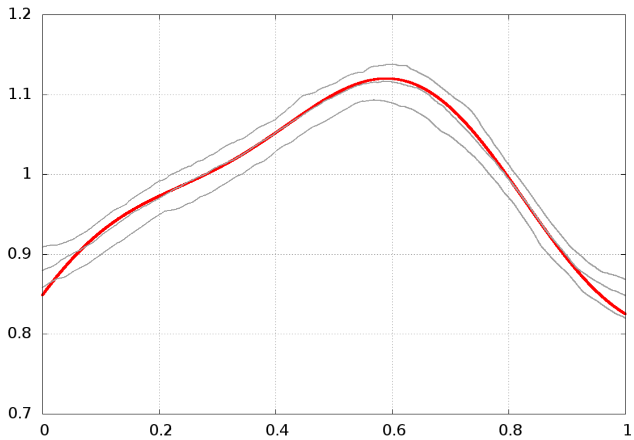
The coefficients of the idiosyncratic components are chosen to reproduce approximately the empirical characteristics of the data: α_i, γ_i and β_i are set respectively to 0.05, 0.06 and 0.90 for each series. The scale factors a_i and the marginal variances $1/\nu_i$ are drawn from an Exponential distribution. The plot of the common trend component $\phi(z_t)$ used in the simulations is displayed in Figure 2.

In order to induce cross–sectional dependence in the innovations $\epsilon_{i\,t}$, we proceed as follows: for each replication of the Monte Carlo, we simulate a covariance matrix from a Wishart distribution of order N. The parameters of the Wishart are chosen in a way such that its expectation is the identity and the standard deviation of the off diagonal elements is 0.03. We then construct the correlation matrix by appropriately standardizing the covariance matrix and simulate a panel of T random numbers from a multivariate normal with mean zero and covariance matrix equal to the generated correlation matrix. The normal random numbers are then mapped into uniforms random variables through the normal cdf. Finally, we plug each of the N uniform random numbers series into the inverse cdf of a Gamma distribution with dispersion equal to $1/\nu_i$. The simulation procedure ensures that marginally the innovations $\epsilon_{i\,t}$ have a Gamma distribution and exhibit cross–sectional dependence. The Monte Carlo experiment is replicated for 1000 times. The dimensions of the panel are equal to $N = 100$ and $T = 5000$, which mimics the dimensions of the S&P100 panel analyzed in this paper.

Table 3 reports results of the experiment in the form of summary averages across all Monte Carlo replications and all series in the panel. The left panel shows the squared bias and variance of the trend estimator evaluated in five different points of the support of the curve. The table also reports the average of the squared estimated asymptotic standard error, and the coverage rate of the asymptotic 90% confidence interval constructed as the trend point estimate plus or minus the normal quantile times the estimated asymptotic standard error. We also report in Figure 2 the 95% quantile range and the median of the simulated distribution of the nonparametric trend estimator. Results show that the nonparametric estimator behaves satisfactorily and that the estimator is rather precise. The average estimated standard error is close to its target, albeit slightly downward biased. The downward bias of the standard error also affects inference: in fact, the coverage rate of the confidence intervals is slightly smaller than the nominal level of 90%. The right panel of Table 3 reports similar summary statistics for the parametric part of the model. Again, the estimation procedure delivers satisfactory estimates. The average estimated standard errors closely track their population analogs and the coverage rate of the 90% confidence interval closely matches the desired level.

Overall, results show that the proposed estimation procedure performs well in moderately large panels, and that the large sample standard error estimators provide adequate inference.

Figure 2: Simulation Study: Common Trend.



The figure shows the profile of the common trend (thick line) used in the Monte Carlo simulations together with the 95% quantile range and the median of the simulated distribution of the trend estimator (thin lines).

7 Empirical Analysis

In this section we apply the SPvMEM to the SPDR and S&P100 panels. Section 7.1 presents the estimation results over the full sample, while Section 7.2 presents the results of a forecasting exercise where the model is compared against a number of alternative specifications.

Table 3: Simulation Study

Nonparametric Component					Parametric Component				
z_τ	squared bias (× 100)	variance (× 100)	average est. variance (× 100)	90% CI coverage		squared bias (× 100)	variance (× 100)	average est. variance (× 100)	90% CI coverage
0.17	0.0005	0.0123	0.0087	0.8322	a_i	0.0028	0.0914	0.0927	0.9079
0.33	0.0015	0.0131	0.0112	0.8260	α_i	0.0000	0.0055	0.0069	0.9384
0.50	0.0008	0.0159	0.0128	0.8406	γ_i	0.0001	0.0143	0.0171	0.9056
0.67	0.0021	0.0138	0.0121	0.8340	β_i	0.0010	0.0461	0.0476	0.8958
0.83	0.0007	0.0117	0.0087	0.7521	ν_i	0.0002	0.0293	0.0216	0.8945

The table reports the squared bias, variance, average estimated variance and coverage of the 90% confidence interval of the nonparametric common trend estimator (in different points of the support) and of the marginals' parameter estimators. 1000 replications for $N = 100$ and $T = 5000$.

7.1 In–Sample Estimation Results

Table 4: SPDR Estimated Parameters

	SPvMEM						Univariate MEM					
	a_i	α_i	γ_i	β_i	ν_i	π_i	a_i	α_i	γ_i	β_i	ν_i	π_i
XLB	0.32 (0.033)	0.22 (0.018)	0.13 (0.012)	0.57 (0.026)	0.33 (0.000)	0.86 (0.032)	0.05 (0.012)	0.25 (0.037)	0.13 (0.027)	0.66 (0.035)	0.33 (0.031)	0.97 (0.053)
XLE	0.13 (0.018)	0.26 (0.015)	0.09 (0.013)	0.66 (0.016)	0.34 (0.000)	0.96 (0.023)	0.10 (0.023)	0.25 (0.032)	0.12 (0.027)	0.64 (0.036)	0.34 (0.030)	0.95 (0.050)
XLF	0.22 (0.017)	0.29 (0.011)	0.12 (0.012)	0.54 (0.015)	0.60 (0.000)	0.90 (0.019)	0.02 (0.012)	0.27 (0.057)	0.13 (0.062)	0.67 (0.045)	0.61 (0.066)	1.00 (0.079)
XLI	0.12 (0.013)	0.27 (0.012)	0.08 (0.009)	0.64 (0.015)	0.47 (0.000)	0.94 (0.019)	0.02 (0.009)	0.29 (0.033)	0.11 (0.023)	0.66 (0.032)	0.48 (0.032)	1.00 (0.047)
XLK	0.09 (0.010)	0.24 (0.010)	0.10 (0.006)	0.69 (0.010)	0.69 (0.000)	0.98 (0.014)	0.04 (0.018)	0.27 (0.045)	0.13 (0.027)	0.66 (0.045)	0.70 (0.044)	0.99 (0.065)
XLP	0.04 (0.004)	0.16 (0.005)	0.07 (0.006)	0.78 (0.005)	0.77 (0.000)	0.98 (0.008)	0.02 (0.011)	0.19 (0.045)	0.11 (0.044)	0.74 (0.039)	0.75 (0.065)	0.99 (0.063)
XLU	0.10 (0.006)	0.25 (0.011)	0.05 (0.009)	0.69 (0.010)	0.59 (0.000)	0.97 (0.016)	0.02 (0.012)	0.27 (0.052)	0.06 (0.042)	0.70 (0.040)	0.65 (0.041)	1.00 (0.069)
XLV	0.11 (0.008)	0.37 (0.013)	0.10 (0.009)	0.52 (0.012)	1.15 (0.000)	0.94 (0.018)	0.02 (0.014)	0.39 (0.064)	0.12 (0.043)	0.55 (0.052)	1.22 (0.048)	1.00 (0.085)
XLY	0.24 (0.019)	0.19 (0.011)	0.12 (0.011)	0.61 (0.019)	0.46 (0.000)	0.86 (0.022)	0.02 (0.008)	0.25 (0.036)	0.13 (0.034)	0.68 (0.035)	0.48 (0.038)	0.99 (0.053)

Estimated parameters and standard errors (in parenthesis) for the SPvMEM (left) and the univariate MEM (right).

Estimation is carried out using a quartic kernel with a bandwidth resulting in a trend computed over a three month window. For the SPDR panel estimation details of the individual series are reported on the left side of Table 4; for the S&P100 panel the results are conveniently aggregated by quantiles across industry groups in Table 11. Detailed estimation results of the S&P100 panel are provided in the Appendix in Table . The univariate asymmetric MEM(1,1)

$$x_{it} = a_i \mu_{it} \epsilon_{it} = \tilde{\mu}_{it} \epsilon_{it} = \left(\omega_i + (\alpha_i + \gamma_i \mathbf{1}_{\{r_{i\,t-1}<0\}}) x_{it-1} + \beta_i \tilde{\mu}_{it-1}\right) \epsilon_{it}$$

provide estimation benchmark results on the right side of the same tables.

When contrasted to typical GARCH estimates, the values of the α's and γ's are higher, while they are lower for the β's. This is the result of a better informative content of realized measures as estimates of volatility (see Brownlees and Gallo (2010) and Shephard and Sheppard (2010) for similar evidence). As customary, we have positive asymmetric reaction to negative news. The estimated persistence $\alpha_i + \gamma_i/2 + \beta_i$ reveals important differences across assets. For SPDR the sectors with higher persistence are Consumer Staples (XLP) and Technology (XLK), meaning that these are the sectors with longer lasting idiosyncratic departures from the systematic component.

Table 5: S&P100 Estimated Parameters

		SPvMEM						Univariate MEM					
		a_i	α_i	γ_i	β_i	ν_i	π_i	a_i	α_i	γ_i	β_i	ν_i	π_i
Disc	$q_{0.25}$	0.11	0.27	0.07	0.60	0.26	0.95	0.06	0.27	0.07	0.62	0.27	0.98
	$q_{0.50}$	0.16	0.30	0.09	0.62	0.31	0.97	0.07	0.29	0.09	0.66	0.32	0.98
	$q_{0.75}$	0.24	0.31	0.10	0.66	0.37	0.98	0.09	0.31	0.10	0.67	0.37	0.99
Ener	$q_{0.25}$	0.10	0.25	0.06	0.65	0.23	0.97	0.09	0.25	0.07	0.65	0.24	0.97
	$q_{0.50}$	0.12	0.26	0.08	0.68	0.32	0.98	0.12	0.26	0.08	0.68	0.34	0.98
	$q_{0.75}$	0.21	0.28	0.09	0.70	0.53	0.98	0.14	0.28	0.09	0.70	0.54	0.99
Fin	$q_{0.25}$	0.10	0.30	0.09	0.53	0.25	0.97	0.05	0.30	0.08	0.56	0.26	0.99
	$q_{0.50}$	0.12	0.35	0.11	0.57	0.30	0.98	0.05	0.35	0.11	0.58	0.31	0.99
	$q_{0.75}$	0.16	0.37	0.13	0.64	0.43	0.98	0.07	0.38	0.13	0.64	0.45	1.00
Heal	$q_{0.25}$	0.09	0.27	0.06	0.58	0.35	0.97	0.07	0.27	0.06	0.56	0.35	0.97
	$q_{0.50}$	0.14	0.31	0.07	0.63	0.46	0.98	0.08	0.31	0.09	0.63	0.46	0.98
	$q_{0.75}$	0.17	0.37	0.09	0.66	0.69	0.98	0.13	0.37	0.10	0.66	0.66	0.99
Ind	$q_{0.25}$	0.10	0.28	0.08	0.58	0.25	0.96	0.06	0.28	0.08	0.58	0.26	0.98
	$q_{0.50}$	0.12	0.33	0.09	0.60	0.28	0.97	0.08	0.33	0.09	0.61	0.29	0.98
	$q_{0.75}$	0.16	0.34	0.12	0.63	0.36	0.98	0.10	0.34	0.11	0.63	0.37	0.99
Mat	$q_{0.25}$	0.17	0.26	0.09	0.59	0.24	0.95	0.09	0.27	0.09	0.59	0.24	0.97
	$q_{0.50}$	0.18	0.29	0.10	0.61	0.26	0.96	0.13	0.29	0.10	0.63	0.27	0.97
	$q_{0.75}$	0.27	0.32	0.10	0.65	0.32	0.96	0.17	0.32	0.10	0.66	0.33	0.98
Stap	$q_{0.25}$	0.07	0.24	0.05	0.61	0.31	0.97	0.05	0.24	0.05	0.59	0.31	0.97
	$q_{0.50}$	0.10	0.30	0.06	0.62	0.39	0.97	0.07	0.31	0.07	0.63	0.39	0.98
	$q_{0.75}$	0.14	0.33	0.07	0.70	0.56	0.98	0.11	0.34	0.08	0.71	0.56	0.98
Tech	$q_{0.25}$	0.11	0.26	0.08	0.49	0.22	0.96	0.06	0.27	0.08	0.51	0.24	0.97
	$q_{0.50}$	0.21	0.29	0.10	0.61	0.27	0.97	0.09	0.31	0.10	0.60	0.28	0.99
	$q_{0.75}$	0.28	0.39	0.12	0.68	0.37	0.99	0.14	0.40	0.12	0.68	0.38	0.99
Util	$q_{0.25}$	0.09	0.27	0.05	0.62	0.30	0.97	0.07	0.28	0.05	0.62	0.31	0.98
	$q_{0.50}$	0.11	0.29	0.09	0.65	0.36	0.98	0.08	0.30	0.09	0.65	0.36	0.98
	$q_{0.75}$	0.13	0.31	0.10	0.68	0.56	0.98	0.10	0.31	0.10	0.67	0.56	0.98

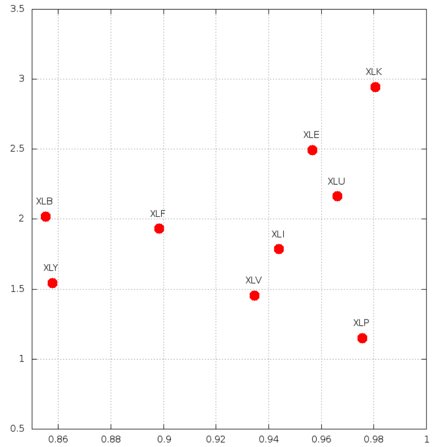
Estimated parameters for the SPvMEM (left) and the univariate MEM (right). The table reports the median, 1^{st} quartile and 3^{rd} quartile of the parameter estimates across the each industry group.

For S&P100 the differences between the least and the most persistent idiosyncratic components are wide and, in general, they appear to be higher than the ones in SPDR. We interpret this as the result that individual assets have a higher level of idiosyncrasy in comparison to sectoral ETFs. The systematic scale factors are fairly close to the sample average volatilities (see Table 1). Last, the estimates of the Gamma distribution also differ substantially across sectors and assets, implying differences in the marginal distributions. By contrast, in the vast majority of cases, the persistence implied by the MEMs are essentially equal to one, hinting to the violation of the stationarity condition, and the unconditional volatilities (denoted by a_i for comparison purposes with our model) are often far (especially for SPDR) from the sample average volatilities.

Indeed, Figure 3 displays the scatter plot of the estimated (log) unconditional mean versus persistence for the SPvMEM in the SPDR panel. Technology (XLK) is the sector with the highest unconditional volatility (and high persistence), followed by Energy (XLE) and Utilities (XLU). The Industrial (XLI), Health Care (XLV) and Consumer Staple (XLP) sectors have lower levels of volatility but still have persistent dynamics while Financials (XLF), Materials (XLM), and Consumer Discretionary (XLY) have lower levels of volatility and lower persistence. By contrast, and as a consequence of the misspecification, the estimated $\alpha_i + \gamma_i/2 + \beta_i$ in the univariate MEM (not reported in the plot), collapse to unity in all sectors regardless of the unconditional level of volatility, forcing the estimates of unconditional volatility to be often excessively large.

Figure 4 shows (from top to bottom) the estimated mean $\sqrt{252\ a_i\phi(z_t)\mu_{i\,t}}$, the systematic component $\sqrt{\phi(z_t)}$, and the idiosyncratic components $\sqrt{252\ a_i\mu_{i\,t}}$ for SPDR (left) and S&P100 (right). The fitted mean series accurately track the movements of realized volatilities. The sys-

Figure 3: SPDR Persistence versus volatility



Scatter plots of persistence (X-axis) versus (log) unconditional volatility (Y-axis) for the SPvMEM.

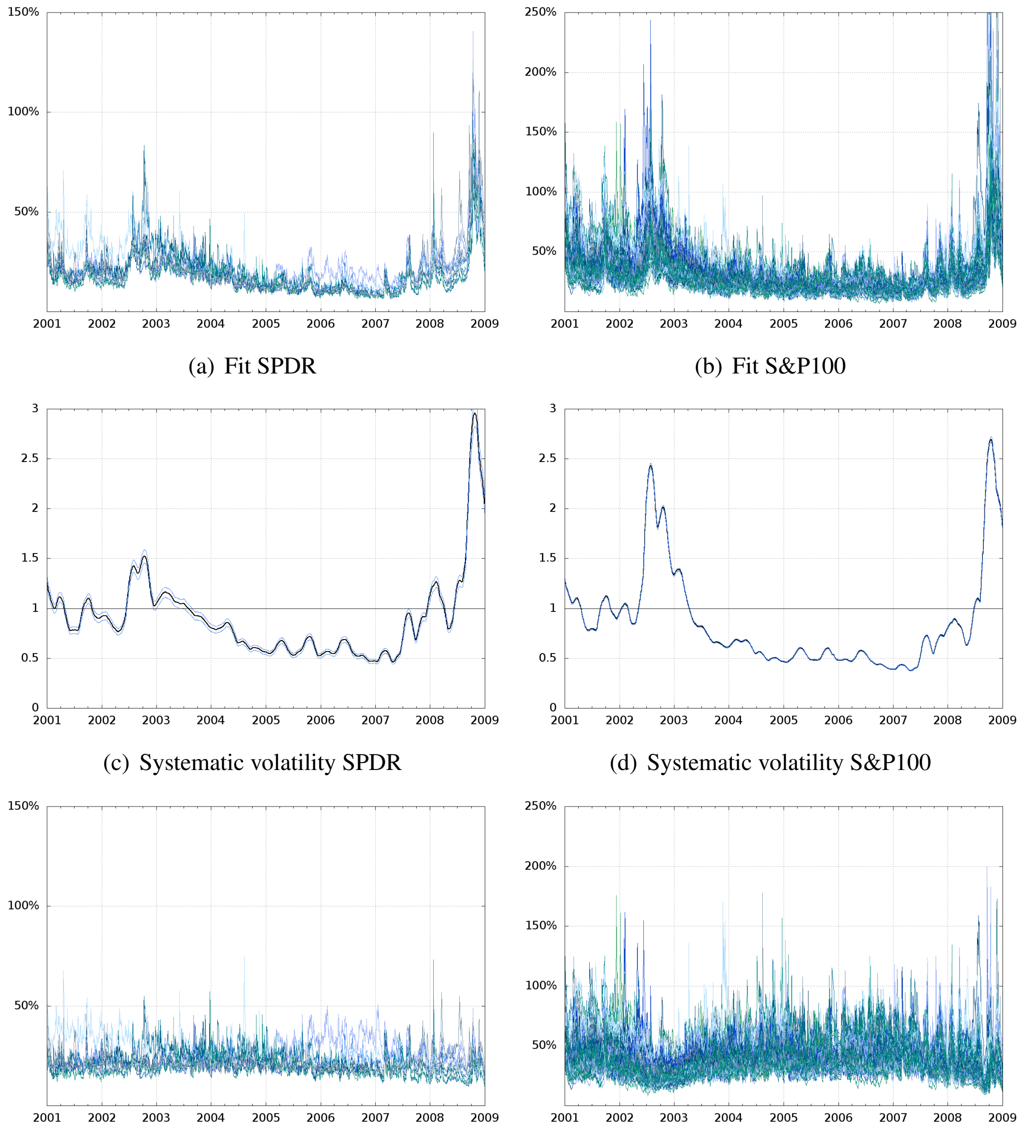
tematic volatility for both panels are essentially close with minor differences for the 2002 burst in volatility recorded differently in the two sets of data (averages vs individual stocks). The plot of the systematic volatility components also displays the (pointwise) 95% confidence bands. Note that, especially in the S&P100 panel, the width of interval is rather tight and the bands are hard to see. The horizontal line at 1 is used as a benchmark to identify periods of systematic risk amplification and contraction. The systematic trend can be interpreted in terms of the underlying movements in the business cycle and financial markets. We trace the increase in volatility following the market drop in mid 2002 with a cut to a half to its unconditional value during the years of volatility moderation, and the sudden increase to three times its normal value toward the last quarter of 2008. The idiosyncratic volatilities are stationary and vary around the unconditional means.

To get deeper insights, Figure 5 displays the idiosyncratic volatilities of the Energy, Financial and Technology sectors only, which allows to visually identify periods of distress around the systematic trend. At the beginning of 2001, Technology was the most volatile sector due to the aftermath of the burst of the dot com bubble. Between 2005 and 2007, concerns about oil prices generated an increased level of uncertainty in the Energy sector. Finally, the Financial sector had a surge in volatility starting from July 2007 with the beginning of the credit crunch.

Table 6 reports the sample correlation matrix $\mathbf{R}$ as defined in Section 3 for the SPDR while Figure 6 provides analog estimate for the S&P100 in a heat map. The average correlations are around 0.40 and 0.25 in the SDPR and S&P100 panels respectively. In comparison to the original data, we see that the common trend captures a significant amount of cross–sectional dependence. The heat map of the S&P100 unveils clustering among stocks that belong to the same sector that are not visible from the raw data. This is the case for Technology, Financials, Energy and Utilities.

Last, we find that residuals of the model do not exhibit significant serial dependence. Inspection of the lag one autocorrelation matrices (not reported in the paper) for both models reveals that the autocorrelations are negligible and there is virtually no significant evidence. Also, the cross autocorrelations are exiguous.

Figure 4: Volatility decomposition



(a) Fit SPDR

(b) Fit S&P100

(c) Systematic volatility SPDR

(d) Systematic volatility S&P100

(e) Idiosyncratic volatilities SPDR

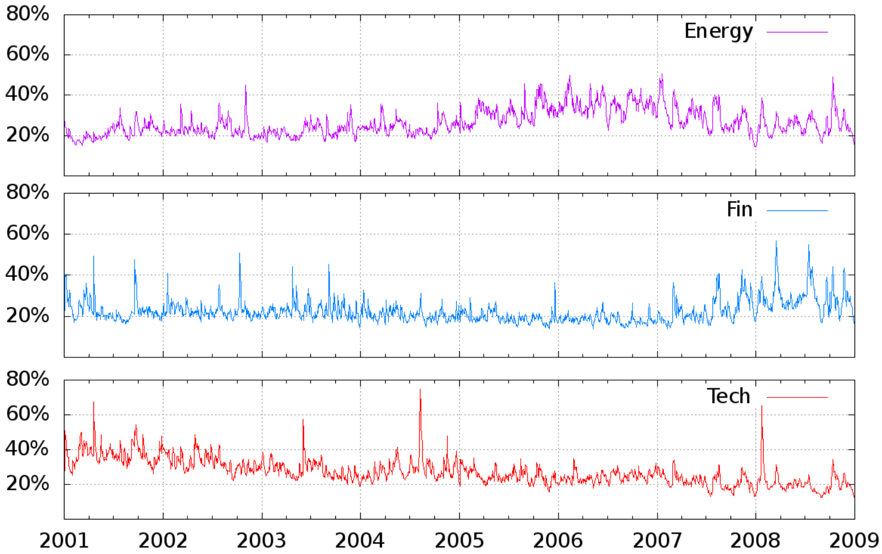
(f) Idiosyncratic volatilities S&P100

The top row shows the estimated fit of the annualized volatilities entailed by the model: $\sqrt{252\ a_i\ \phi(z_t)\ \mu_{i\,t}}$. The middle row shows the systematic volatility $\sqrt{\phi(z_t)}$, and the bottom row shows the annualized idiosyncratic volatilities $\sqrt{252\ a_i\ \mu_{i\,t}}$. Left column is for SPDR and right column for S&P100.

7.2 Out–of–Sample Forecasting Analysis

To assess the predictive ability of our proposed specification we carry out a forecasting exercise. We compare the SPvMEM against four alternative specifications: a MEM(1,1), a MEM(2,2), a Composite Likelihood MEM(1,1) (CL–MEM(1,1), cf. Engle *et al.* (2008)) and a (univariate) Seminonparametric MEM (SPMEM). All the MEM specifications considered allow for asymmetric reactions to past negative returns. In detail:

Figure 5: Idiosyncratic volatilities for Energy, Financial and Technology



The figure shows the idiosyncratic volatility for Energy (top), Financials (middle) and Technology (bottom).

Table 6: SPDR Residual Copula Dependence

	XLB	XLE	XLF	XLI	XLK	XLP	XLU	XLV	XLY
XLB	1								
XLE	0.37	1							
XLF	0.29	0.25	1						
XLI	0.44	0.36	0.44	1					
XLK	0.27	0.27	0.36	0.30	1				
XLP	0.25	0.21	0.21	0.23	0.28	1			
XLU	0.33	0.30	0.20	0.28	0.22	0.19	1		
XLV	0.26	0.22	0.26	0.26	0.51	0.42	0.22	1	
XLY	0.37	0.28	0.45	0.41	0.36	0.26	0.25	0.29	1

The table reports the estimated Gaussian meta-copula correlation matrix for the SPDR panel.

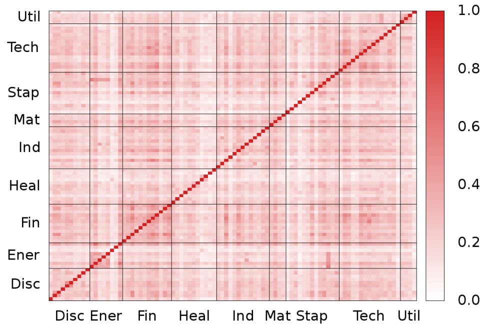
Table 7: SPDR Residual Lag 1 Autocorrelation

	XLB	XLE	XLF	XLI	XLK	XLP	XLU	XLV	XLY
XLB	-0.00	0.03	0.03	0.03	0.04	0.02	0.09	0.02	0.00
XLE	0.03	-0.02	0.02	0.03	0.02	0.07	0.04	0.02	0.00
XLF	-0.01	0.03	-0.01	-0.00	0.05	0.03	0.06	0.01	0.00
XLI	0.04	0.03	0.01	-0.02	0.03	0.03	0.06	0.03	-0.00
XLK	0.04	-0.02	0.01	0.02	-0.02	0.00	0.04	-0.01	0.00
XLP	0.03	0.03	0.01	0.04	0.03	-0.01	0.06	-0.01	0.02
XLU	0.03	0.01	0.02	0.04	0.02	0.05	0.01	0.03	0.03
XLV	0.01	0.00	0.01	0.02	0.01	0.04	0.04	-0.03	-0.00
XLY	0.04	0.04	0.02	0.00	0.05	0.03	0.09	0.03	-0.02

The table reports the residual lag one autocorrelation matrix for the SPDR panel.

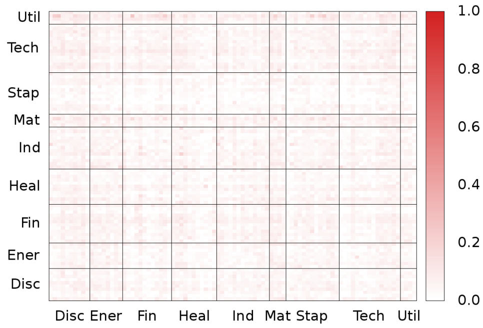
1. MEM without trend: $x_{it} = a_i \mu_{it} \epsilon_{it} = \tilde{\mu}_{it} \epsilon_{it}$, with $\tilde{\mu}_{it}$ specified as

Figure 6: S&P100 Residual Copula Dependence



The heatmap displays the estimated Gaussian meta-copula correlation matrix for the S&P100 panel.

Figure 7: S&P100 Residual Lag 1 Autocorrelation



The heatmap displays the residual lag one autocorrelation matrix for the S&P100 panel.

Table 8: SPDR Out–of–sample Forecasting

	SPVMEM	MEM(1,1)	MEM(2,2)	CL-MEM(1,1)	SPMEM
XLB	69.18	67.79	69.05	86.06	73.81
XLE	62.78	64.05	64.04	75.27	68.26
XLF	89.53	86.84	86.10	121.47	86.55
XLI	75.45	77.37	168.06	131.10	81.88
XLK	100.84	101.87	226.81	182.50	102.30
XLP	116.90	120.10	123.27	152.77	126.64
XLU	86.41	92.91	184.00	146.47	94.28
XLV	130.73	142.53	148.27	175.00	145.37
XLY	85.49	88.75	89.09	125.27	93.17

Out–of–sample QL losses for the SPvMEM, MEM(1,1), MEM(2,2), Composite Likelihood MEM(1,1) and (univariate) SPMEM.

(a) MEM(1,1): $\omega_i + (\alpha_i + \gamma_i \mathbf{1}_{\{r_{i\,t-1}<0\}})x_{it-1} + \beta_i \tilde{\mu}_{it-1}$

(b) MEM(2,2): $\omega_i + (\alpha_{1i} + \gamma_i \mathbf{1}_{\{r_{i\,t-1}<0\}})x_{it-1} + \beta_{1i}\tilde{\mu}_{it-1} + \alpha_{2i}x_{it-2} + \beta_{2i}\tilde{\mu}_{it-2}$

(c) CL–MEM(1,1): $\omega + (\alpha + \gamma \mathbf{1}_{\{r_{i\,t-1}<0\}})x_{it-1} + \beta \tilde{\mu}_{it-1}$

2. SPMEM(1,1) (individual trend): $x_{it} = a_i \phi_i(z_t)\mu_{it}\epsilon_{it} = \phi_i(z_t)\tilde{\mu}_{it}\epsilon_{it}$, with $\tilde{\mu}_{it}$ specified as in MEM(1,1) and $\phi_i(z_t)$ an asset specific, univariate version of expression (7).

Table 9: S&P100 Out–of–sample Forecasting

		SPVMEM	MEM(1,1)	MEM(2,2)	CL-MEM(1,1)	SPMEM
Disc	Mean	65.47	66.11	79.56	105.96	66.70
	q_{10}	56.12	56.79	56.78	78.37	55.95
	q_{90}	75.84	76.39	132.10	143.99	78.39
Ener	Mean	67.81	73.22	135.90	116.34	69.57
	q_{10}	47.61	47.24	45.51	61.28	47.72
	q_{90}	117.24	150.49	346.55	259.27	124.07
Fin	Mean	75.98	77.45	107.68	108.34	79.24
	q_{10}	59.25	58.97	58.36	83.35	61.40
	q_{90}	97.28	101.22	181.95	138.90	102.77
Heal	Mean	82.37	91.09	97.88	122.59	84.18
	q_{10}	57.07	68.79	75.65	97.19	58.11
	q_{90}	100.57	119.03	136.95	152.94	102.12
Ind	Mean	67.48	70.31	90.76	108.16	68.44
	q_{10}	57.63	57.01	55.55	82.67	58.80
	q_{90}	79.73	97.67	173.82	135.09	81.51
Mat	Mean	68.07	68.76	64.54	91.43	68.91
	q_{10}	57.55	58.54	54.90	75.37	58.43
	q_{90}	73.79	76.91	70.83	112.36	75.86
Stap	Mean	80.51	81.25	95.05	114.22	81.81
	q_{10}	62.16	62.48	58.28	74.34	60.61
	q_{90}	94.81	94.46	141.41	134.90	95.66
Tech	Mean	72.44	71.94	79.27	127.07	72.51
	q_{10}	58.33	57.68	54.12	95.50	55.80
	q_{90}	93.34	93.13	128.50	165.02	92.98
Util	Mean	70.63	70.47	91.54	107.64	72.27
	q_{10}	57.86	56.27	52.07	82.07	59.17
	q_{90}	89.70	91.28	129.84	144.28	88.45

Out–of–sample QL losses for the SPvMEM, MEM(1,1), MEM(2,2), Composite Likelihood MEM(1,1) and (univariate) SPMEM. The table reports the average loss, the 10% quantile and 90% quantile of each industry group.

Models without trend are the MEM(1,1) (seen above), possibly extended in the specification to accommodate a second lag (MEM(2,2)). Since parameter coefficients estimated over a panel of financial time series may cluster, the CL–MEM(1,1) assumes parameters to be constant across series. The last specification forces volatility trends to be series specific, and hence allows for assessing the benefits (if any) of estimating the common trend by pooling series.

The forecasting exercise is designed as follows. Starting from the beginning of 2007 we produce the series of one step ahead forecasts for each model using parameter estimates updated on the last weekday of each week. The prediction of the nonparametric trend is constructed by keeping constant the last estimate for the forecast horizon. The series of forecasts produced by the five approaches are evaluated using the QL loss function (Patton (2010)).

The choice of the kernel and the bandwidth for the out–of–sample exercise require further details. First, building on Gijbels *et al.* (1999) among others, we fit the SPvMEM using a one sided quartic kernel for the forecasting application. Second, the bandwidth is chosen using an out–of–sample cross–validation criterion. We consider different bandwidths corresponding to different window lengths varying from one to six months. We then estimate the models from the beginning of the sample until June 2006 using the different bandwidths and choose the bandwidth that delivers the one step ahead best forecasts over the last six months of 2007.

Results are shown in Tables 9 and 10. We report out–of–sample losses for each of the series of the SPDR, and cross sectional means and quantiles across industry groups for the S&P100. The evidence from the two forecasting exercises is similar. The SPvMEM delivers the best out-of-sample performance in the majority of cases. The second best performing model is the baseline

MEM(1,1) which closely tracks the SPvMEM. The performance of the MEM(2,2), on the other hand, varies substantially. In a number of cases the QL loss of this specification is much larger than the baseline MEM(1,1). The CL-MEM(1,1) does not appear to improve the predictions of the MEM(1,1). As it can be observed in Tables 5 and 6 parameters may vary quite considerably in the cross section so that, contrary to the encouraging results in GARCH panels, pooling in the MEM case may not be the best estimation strategy. The SPMEM generally performs closely to the SPvMEM but the latter generally performs better. Interestingly, in the SPDR dataset the discrepancy between the two models is larger than for the S&P100. Overall, the SPvMEM improves predictive ability in the majority of cases. An interesting outcome of the forecast comparison is that the SPvMEM delivers better results for the SPDR and S&P100 panels relative to the SPMEM. In the former model, jointly estimating the systematic risk component truly exploits the properties of interdependence across assets.

8 Conclusions

Modeling large panels of volatilities may prove a formidable task if one allows for dynamic interdependence. We follow parsimony in parametric specification, exploiting the stylized fact that volatilities appear to be driven by an underlying factor that captures the secular systematic trend. We propose a novel SeminonParametric vector MEM (SPvMEM) specification that decomposes risk measures in a systematic and idiosyncratic components, and it allows for cross–section dependence in the innovations. The systematic component is a secular trend and the idiosyncratic components are parametric. We develop an estimation technique that is based on profile likelihood estimation and inference from the marginals. Regardless of the dimension of the panel, estimation boils down to univariate likelihood maximization and the computation of a sample correlation matrix. The ease of estimation of our model makes it appealing in large dimensional applications.

We analyse two panels of daily realized volatility measures between 2001 and 2008. The first panel consists of the nine SPDR Sectoral Indices of the S&P500, and the second panel contains the ninety constituents of the S&P100 that have been continuously trading in the sample period. There is evidence of a common trend in both panels. Once the common component is accounted for all series exhibit mean reversion around it. The model also unveils dependencies in volatility innovations across assets and sectoral clusters for Technology, Financial, Energy and Utilities companies. A forecasting horse race against a set of competing specifications shows that the SPvMEM delivers the best out–of–sample performance for the majority of series in both panels.

Further refinements or uses of the model can be envisaged. As realized measures are estimators, we would need to investigate the benefits of taking the measurement error in the volatility estimation into explicit account in the modelling step (see Hansen and Lunde (2010)). As in other contributions, a relationship between macroeconomic variables and the common component can help highlight some determinants of the changes in risk levels or, reverting the perspective, the spillover effects of market volatility onto the real economy.

Acknowledgements

We are grateful to Yacine Aït-Sahalia, Robert F. Engle, Jianqing Fan, Christian Hafner, Marc Hallin, Eric Jondeau, Charles Mathias, Oliver Linton, Alexei Onatsky, Davy Paindaveine, Eric Renault, Antonio Rubia, Song Song, Valeri Voev and Liuren Wu for insightful remarks. We are also grateful to the participants of the 4th Brussels-Waseda Seminar in Time Series and Financial

Statistics (June 2009), the Stanford SITE summer 2009 workshop (June 2009), the 3rd International Conference on Computational and Financial Econometrics (October, 2009), the 2009 EC^2 conference in Aarhus (December 2009), the Oberwolfach Workshop of Semiparametric Modeling of Multivariate Time Series with Changing Dynamics (January, 2010), the NYU-Stern conference on Volatility and Systemic Risk (April 2010), the EUI Econometrics Workshop on Recent Advances in Time Series Econometrics (June 2010), the CEQURA Conference on Advances in Financial and Insurance Risk Management (September 2010), the XVIII Finance Forum (November 2010), the SETA conference (April 2011), and the seminar participants at Carlos III, CORE, ECARES, LSE, Lugano, Princeton, Queen Mary College, Rutgers, and Université catholique de Louvain. Any error and inaccuracy are ours.

Christian Brownlees acknowledges financial support from Spanish Ministry of Science and Technology (Grant MTM2012-37195) and from Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075). Giampiero M. Gallo acknowledges financial support from the italian MIUR under the PRIN 2011 grant "MISURA". David Veredas acknowledges financial support from the Belgian National Bank and the IAP P6/07 contract, from the IAP program (Belgian Scientific Policy), 'Economic policy and finance in the global economy'. Any errors and inaccuracies are ours.

References

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous–time process in the presence of market microstructure noise. *The Review of Financial Studies*, **28**, 351–416.

Alessi, L., Barigozzi, M., and Capasso, M. (2009). Estimation and forecasting in large datasets with conditionally heteroskedastic dynamic common factors. Technical Report 09–1115, European Central Bank.

Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, **57**, 1047–1091.

Amado, C. and Teräsvirta, T. (2008). Modelling conditional and unconditional heteroskedasticity with smoothly time-varying structure. Technical Report 691, SSE/EFI Working Paper Series in Economics and Finance.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, **71**, 579–625.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*, **89**, 701–720.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, **70**, 191–221.

Bandi, F. M. and Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, **79**, 655–692.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, **76**, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realised kernels in practice: Trades and quotes. *Econometrics Journal*, **12**, 1–32.

Bauwens, L. and Rombouts, J. V. K. (2007). Bayesian clustering of many GARCH models. *Econometric Reviews*, **26**, 365–386.

Brownlees, C. T. and Gallo, G. M. (2006). Financial econometric analysis at ultra–high frequency: Data handling concerns. *Computational Statistics and Data Analysis*, **51**, 2232–2245.

Brownlees, C. T. and Gallo, G. M. (2010). Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics*, **8**, 29–56.

Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula–based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, **135**, 125–154.

Chen, X., Ghysels, E., and Wang, F. (2011). HYBRID GARCH models and intra-daily return periodicity. *Journal of Time Series Econometrics*, **3**.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, **26**, 922–947.

Cipollini, F., Engle, R. F., and Gallo, G. M. (2006). Vector multiplicative error models: Representation and inference. Technical Report 12690, National Bureau of Economic Research.

Colacito, R., Engle, R. F., and Ghysels, E. (2011). A component model for dynamic correlations. *Journal of Econometrics*, **164**, 45–59.

Corsi, F. (2010). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, **7**, 174–196.

Deo, R., Hurvich, C., and Lu, Y. (2006). Forecasting realized volatility using a long-memory stochastic volatility model: Estimation, prediction and seasonal adjustment. *Journal of Econometrics*, **131**, 29–58.

Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, **17**, 425–446.

Engle, R. F. and Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, **131**, 3–27.

Engle, R. F. and Rangel, J. G. (2008). The spline-GARCH model for low frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, **21**, 1187–1222.

Engle, R. F., Shephard, N., and Sheppard, K. (2008). Fitting vast dimensional time-varying covariance models. Technical report.

Engle, R. F., Ghysels, E., and Sohn, B. (2009). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics*. available online.

Gijbels, I., Pope, A., and Wand, M. (1999). Understanding exponential smoothing via kernel regression. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **61**, 39–50.

Hafner, C. M. and Linton, O. (2010). Efficient estimation of a multivariate multiplicative volatility model. *Journal of Econometrics*, **159**, 55–73.

Hansen, P. R. and Lunde, A. (2010). Estimating the persistence and the autocorrelation function of a time series that is measured with error. Technical Report 2010-8, CREATES.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: A joint model of returns and realized measures of volatility. *Journal of Applied Econometrics*, **27**, 877–906.

Hautsch, N. (2008). Capturing common components in high-frequency financial time series: A multivariate stochastic multiplicative error model. *Journal of Economic Dynamics and Control*, **32**, 3978 – 4015.

Hautsch, N., Hess, D., and Veredas, D. (2011). The impact of macroeconomic news on quote adjustments, noise, and informational volatility. *Journal of Banking and Finance*, **35**, 2733 – 2746.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, **94**, 401–419.

Li, Q. and Racine, J. S. (2006). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Princeton.

Long, X., Su, L., and Ullah, A. (2011). Estimation and forecasting of dynamic conditional covariance: A semiparametric multivariate model. *Journal of Business & Economic Statistics*, **29**, 109–125.

Luciani, M. and Veredas, D. (2011). A simple model for vast panels of volatilities. Technical Report 2011/28, ECARES.

McLeish, D. (1975). A maximal inequality and dependent strong laws. *Annals of Probability*, **3**, 826–836.

Müller, H. (1984). Smooth optimum kernel estimators of densities, regression curves, and modes. *Annals of Statistics*, **12**, 766–744.

Noureldin, D., Shephard, N., and Sheppard, K. (2012a). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, **27**, 907–933.

Noureldin, D., Shephard, N., and Sheppard, K. (2012b). Multivariate rotated ARCH models. Technical report.

Pakel, C., Shephard, N., and Sheppard, K. (2011). Nuisance parameters, composite likelihoods and a panel of GARCH models. *Statistica Sinica*, **21**, 307–329.

Patton, A. (2010). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, **160**, 246–256.

Patton, A. J. and Sheppard, K. (2009). Good volatility, bad volatility: Signed jumps and the persistence of volatility. Technical report.

Rangel, J. G. and Engle, R. F. (2012). The factor–spline–GARCH model for high and low frequency correlations. *Journal of Business & Economic Statistics*, **30**, 109–124.

Severini, T. A. and Wong, W. H. (1992). Profile likelihood and conditionally parametric models. *Annals of Statistics*, **20**, 1768–1802.

Shephard, N. and Sheppard, K. (2010). Realising the future: forecasting with high frequency based volatility (HEAVY) models. *Journal of Applied Econometrics*, **25**, 197–231.

Song, P. X. (2000). Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics*, **27**, 305–320.

Staniswalis, J. G. (1989). On the kernel estimate of a regression function in likelihood based models. *Journal of the American Statistical Association*, **84**, 273–283.

Veredas, D., Rodriguez-Poo, J., and Espasa, A. (2007). Semiparametric estimation for financial durations. In L. Bauwens, W. Pohlmeier, and D. Veredas, editors, *High Frequency Financial Econometrics. Recent Developments*, pages 204–208. Springer Verlag.

Wang, Y. and Zou, J. (2010). Vast volatility matrix estimation for high-frequency financial data. *Annals of Statistics*, **38**, 943–978.

White, H. and Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, **52**, 143–162.

Wooldridge, J. and White, H. (1988). Some invariance principles and central limit theorems for dependent and heterogeneous processes. *Econometric Theory*, **4**, 210–230.

Appendix A

Assumptions

Assumption A states the properties of the vector of innovations. Assumptions B and P are standard in Maximum Likelihood estimation. Assumption C1 contains differentiability conditions (i.e. bounded derivatives) on the log–likelihoods, C.2 allows for the large N setting. Assumption D assumes bounded cross–sectional covariance among innovations as $N \to \infty$ (and is new to this paper). Assumption I is necessary for the identification assumption of both the parameters and the curve. Assumption K states the properties of the kernel. Assumption L contains conditions on the smoothness of the curve necessary to define the least favorable direction. Assumption N contains regularity conditions for the estimated curve and its derivatives (and is also proved in Lemma 4). Assumption S assumes bounded mixed derivatives needed to show the convergence of the estimated tangent vector to the least favorable direction.

A.1 The detrended process $x_{i\,t}/\phi_0(z_t) = a_{i0}\mu_{i\,t}(\boldsymbol{\delta}_{i0})\epsilon_{i\,t}$ is a strong mixing process for any $i = 1, \dots N$, where, for some $p > 2$ and $r \in \mathbb{N}$, the mixing coefficients $\{\alpha_j\}$ must satisfy

$$\sum_{j=1}^{\infty} j^{r-1}\alpha_j^{1-2/p} < \infty.$$

Furthermore, for some even integer $q \leq 2r$ $\mathrm{E}_0[|\mu_{i\,t}(\boldsymbol{\delta}_{i0})\epsilon_{i\,t}|^q] < m$, where m is a constant not depending on z_t.

A.2 $\boldsymbol{\epsilon}_t$ is a conditionally independent random vector process such that, for any $i = 1, \dots N$, $\epsilon_{i\,t}|\mathcal{F}_{t-1} \sim \mathit{Gamma}(\nu_{i0}, \nu_{i0})$, and $\mathrm{E}_0[\epsilon_{i\,t}] = 1$, $\mathrm{Var}_0[\epsilon_{i\,t}] = \nu_{i0}^{-1}$.

A.3 For any $i = 1, \dots, N$, we have $0 < \min_i(\nu_{i0}) \leq \nu_{i0} \leq \max_i(\nu_{i0}) < \infty$.

B The true values of the parameters are such that $\boldsymbol{\xi}_0 \in int(\boldsymbol{\Xi})$ and $\boldsymbol{\psi}_0 \in int(\boldsymbol{\Psi})$, with $\boldsymbol{\Xi} \subset \mathbb{R}^{5N}$, $\boldsymbol{\Psi} \subset \mathbb{R}^{p_\psi}$ and $\boldsymbol{\Xi}$, $\boldsymbol{\Psi}$ both compact sets. We also use the notation $\boldsymbol{\eta}_0 = (\boldsymbol{\xi}_0^{\mathrm{T}}, \boldsymbol{\psi}_0^{\mathrm{T}})^{\mathrm{T}} \in int(\boldsymbol{\Lambda}) \subset \mathbb{R}^{5N+\psi}$, with $\boldsymbol{\Lambda}$ compact. The least favorable curve is such that, for any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ and any $z_t \in [0,1]$, $\phi_{\boldsymbol{\xi}}(z_t) \in int(\mathcal{P})$, with $\mathcal{P} \subset \mathbb{R}_+$ compact.

C.1 We assume that for each $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$, $i = 1, \dots, N$ and $z_t \in [0,1]$,

$$\sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \sup_{\phi \in \Gamma} \sup_{z_t \in [0,1]} \mathrm{E}_0\left[\left|\frac{\partial^k}{\partial \boldsymbol{\xi}_i^k}\frac{\partial^l}{\partial z_t^l}\frac{\partial^j}{\partial \phi^j}\boldsymbol{\ell}_{i\,t}^m(\boldsymbol{\xi}_i, \phi(z_t))\right|^q\right] < \infty,$$

for $j = 0,1,2,3$, $k = 0,1,2$, $l = 0,1,2$ and $q = 2$. The same holds also for $\boldsymbol{\ell}_t^c(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi(z_t))$. We use the notation $\partial/\partial\phi$ to indicate the Fréchet functional derivative.

C.2 We also assume that:

$$\sup_{N \in \mathbb{N}} \sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} \sup_{\phi \in \Gamma} \sup_{z_t \in [0,1]} \mathrm{E}_0\left[\left|\frac{1}{N}\sum_{i=1}^{N}\frac{\partial^k}{\partial \boldsymbol{\xi}^k}\frac{\partial^l}{\partial z_t^l}\frac{\partial^j}{\partial \phi^j}\boldsymbol{\ell}_{i\,t}^m(\boldsymbol{\xi}_i, \phi(z_t))\right|^q\right] < \infty,$$

for $j = 0,1,2,3$, $k = 0,1,2$, $l = 0,1,2$, and $q = 2$.

D Define $\mathrm{Cov}_0[\epsilon_{i\,t}, \epsilon_{j\,t}] = \tau_{ij}$, then there exists a positive real constant M such that

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}|\tau_{ij}| < M < \infty.$$

I.1 For any $i = 1, \ldots, N$ and for fixed but arbitrary $\tilde{\boldsymbol{\eta}} \in \boldsymbol{\Lambda}$, $\tilde{\boldsymbol{\xi}}_i \in \boldsymbol{\Xi}_i$, $\tilde{\boldsymbol{\psi}} \in \boldsymbol{\Psi}$ and $\tilde{\phi} \in \mathcal{P}$, let

$$\rho(\boldsymbol{\eta}, \phi) = \widetilde{\mathrm{E}}\left[\boldsymbol{\ell}_t(\boldsymbol{\eta}, \phi)\right], \quad \rho_i^m(\boldsymbol{\xi}_i, \phi) = \widetilde{\mathrm{E}}\left[\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_i, \phi)\right], \quad \rho^c(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi) = \widetilde{\mathrm{E}}\left[\boldsymbol{\ell}_t^c(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi)\right]$$

where expectation is taken with respect to the distribution of $\mathbf{x}_t$ with parameters $\tilde{\phi}$ and $\tilde{\boldsymbol{\eta}}$, $\tilde{\boldsymbol{\xi}}_i$, or $\tilde{\boldsymbol{\psi}}$ respectively. Then, if $\phi \neq \tilde{\phi}$, we have

$$\rho(\boldsymbol{\eta}, \phi) < \rho(\tilde{\boldsymbol{\eta}}, \tilde{\phi}), \quad \rho_i^m(\boldsymbol{\xi}_i, \phi) < \rho_i^m(\tilde{\boldsymbol{\xi}}_i, \tilde{\phi}), \quad \rho^c(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi) < \rho^c(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{\psi}}, \tilde{\phi}).$$

I.2 Let $\phi_{\boldsymbol{\xi}}(z_t)$ be such that

$$\frac{\partial}{\partial \phi} \mathrm{E}_0\left[\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}}(z_t))\right] = 0,$$

for any $z_t \in [0,1]$ and for each fixed $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$, $i = 1, \ldots, N$. Then, we assume that $\phi_{\boldsymbol{\xi}}(z_t)$ is unique and that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that, if

$$\sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \sup_{z_t \in [0,1]} \left| \frac{\partial}{\partial \phi} \mathrm{E}_0\left[\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_i, \bar{\phi}(z_t))\right] \right| \leq \delta,$$

then

$$\sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} \sup_{z_t \in [0,1]} \left| \bar{\phi}(z_t) - \phi_{\boldsymbol{\xi}}(z_t) \right| \leq \varepsilon.$$

K Assume that the kernel function $\mathrm{K}(\cdot)$ is of order $k > 3/2$ with support $[-1,1]$ and it is such that

$$\int_{-1}^{1} \mathrm{K}(u) \mathrm{d}u = 1, \quad \int_{-1}^{1} u \mathrm{K}(u) \mathrm{d}u = 0, \quad \int_{-1}^{1} u^p \mathrm{K}(u) \mathrm{d}u < \infty, \quad \int_{-1}^{1} u^q \mathrm{K}^2(u) \mathrm{d}u < \infty,$$

for $p = 0, \ldots, 3$ and $q = 0, \ldots, 6$. Assume also that

$$\sup_{u \in [-1,1]} \left| \frac{\partial^r \mathrm{K}(u)}{\partial u^r} \right| < \infty, \qquad r = 0, \ldots, 4.$$

The conditions on the bandwidth vary and are stated in Theorem 2.

L Given the least favorable curve $\phi_{\boldsymbol{\xi}}$, then, for any $z_t \in [0,1]$ and any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, define

$$\phi'_{\boldsymbol{\xi}}(z_t) \equiv \frac{\partial \phi_{\boldsymbol{\xi}}}{\partial \boldsymbol{\xi}}(z_t) \text{ and } \phi''_{\boldsymbol{\xi}}(z_t) \equiv \frac{\partial^2 \phi_{\boldsymbol{\xi}}}{\partial \boldsymbol{\xi}^2}(z_t),$$

and define the norm of a vector $\mathbf{w}$ as $||\mathbf{w}|| = \sup_{z_t \in [0,1]} |\mathbf{w}(z_t)|$. Then, we assume that $\phi'_{\boldsymbol{\xi}}(z_t)$ and $\phi''_{\boldsymbol{\xi}}(z_t)$ exist with $||\phi'_{\boldsymbol{\xi}}|| < \infty$ and $||\phi''_{\boldsymbol{\xi}}|| < \infty$.

N For any $z_t \in [0,1]$ and $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, the estimated curve $\widehat{\phi}_{\boldsymbol{\xi} NT}(z_t)$ converges in probability to some constant both if $T \to \infty$ and N is small and if both $N, T \to \infty$. Denote that constant as $\widetilde{\phi}_{\boldsymbol{\xi}}(z_t)$. For any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, we require that $\widetilde{\phi}_{\boldsymbol{\xi}} \in \Gamma$ and, and for all $r, s = 0, 1, 2$ such that $r + 2 \leq 2$, that

$$\frac{\partial^{r+s}}{\partial z_t^r \partial \boldsymbol{\xi}^s} \widetilde{\phi}_{\boldsymbol{\xi}}(z_t) \text{ and } \frac{\partial^{r+s}}{\partial z_t^r \partial \boldsymbol{\xi}^s} \widehat{\phi}_{\boldsymbol{\xi} NT}(z_t),$$

exist. We require that

$$\sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} ||\widehat{\phi}_{\boldsymbol{\xi} NT} - \widetilde{\phi}_{\boldsymbol{\xi}}|| = o_P(1), \quad \sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} ||\widehat{\phi}'_{\boldsymbol{\xi} NT} - \widetilde{\phi}'_{\boldsymbol{\xi}}|| = o_P(1), \quad \sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} ||\widehat{\phi}''_{\boldsymbol{\xi} NT} - \widetilde{\phi}''_{\boldsymbol{\xi}}|| = o_P(1),$$

where the norm is defined in Assumption L. Finally, for some $\delta > 0$, we require that

$$\left\|\frac{\partial}{\partial z_t}\widehat{\phi}_{\boldsymbol{\xi}_o NT} - \frac{\partial}{\partial z_t}\widetilde{\phi}_0\right\| = o_P(T^{-\delta}), \qquad \left\|\frac{\partial}{\partial z_t}\widehat{\phi}'_{\boldsymbol{\xi}_o NT} - \frac{\partial}{\partial z_t}\widetilde{\phi}'_0\right\| = o_P(T^{-\delta}).$$

P The following matrices are positive definite for $i = 1, \ldots, N$:

$$\mathcal{I}_{\boldsymbol{\xi}_{io}\boldsymbol{\xi}_{io}} = \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i0}, \phi_0)\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i0}, \phi_0)\right], \quad \mathcal{H}_{\boldsymbol{\xi}_{io}\boldsymbol{\xi}_{io}} = -\mathrm{E}_0\left[\frac{\partial^2 \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right],$$
$$\mathbf{I}_{\boldsymbol{\psi}_o} = \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^c_t}{\partial \boldsymbol{\psi}}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\frac{\partial \boldsymbol{\ell}^c_t}{\partial \boldsymbol{\psi}^{\mathrm{T}}}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\right], \quad \mathbf{H}_{\boldsymbol{\psi}_0} = -\mathrm{E}_0\left[\frac{\partial^2 \boldsymbol{\ell}^c_t}{\partial \boldsymbol{\psi}\partial \boldsymbol{\psi}^{\mathrm{T}}}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\right].$$

Moreover, the matrices $\mathbf{I}^*_{\boldsymbol{\xi}_0}$ and $\mathbf{H}^*_{\boldsymbol{\xi}_o}$ defined in Theorem 4 are positive definite. We also assume

$$\bar{\jmath}_{N\boldsymbol{\xi}_o}(z_\tau) = -\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^N \frac{\partial^2 \boldsymbol{\ell}^m_{it}}{\partial \phi^2}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_\tau))\right] > 0, \; z_\tau \in [0,1].$$

S Assume that for all $r, s = 0, \ldots, 4$, $r+s \leq 4$, and any $i = 1, \ldots, N$, the derivative $\frac{\partial^{r+s}\boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i^r \partial \phi^s}(\boldsymbol{\xi}_i, \phi)$ exist for almost all $\mathbf{x}_t$ and assume that

$$\mathrm{E}_0\left[\sup_{\boldsymbol{\eta}\in\Lambda}\sup_{\phi\in\Gamma}\left\|\frac{\partial^{r+s}\boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i^r\,\partial \phi^s}(\boldsymbol{\xi}_i, \phi)\frac{\partial^{r+s}\boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i^{r\mathrm{T}}\partial \phi^s}(\boldsymbol{\xi}_i, \phi)\right\|\right] < \infty.$$

Proof of Theorem 1

a) Define for any $z_\tau \in [0,1]$ the smoothed likelihood

$$\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \phi(z_\tau)) \equiv \sum_{t=1}^T \sum_{i=1}^N \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\boldsymbol{\ell}^m_{it}(\boldsymbol{\xi}_{i\,0}, \phi(z_t)). \tag{A-1}$$

Then the estimated curve is such that

$$\widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau) = \arg\sup_{\phi\in\Gamma}\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \phi(z_\tau)), \tag{A-2}$$

and, from Lemma 1 and assumption P, the true value of the curve for N fixed is such that:

$$\phi_0(z_\tau) = \arg\sup_{\phi\in\Gamma}\bar{\lambda}_N(\boldsymbol{\xi}_0, \phi(z_\tau)), \tag{A-3}$$

where

$$\bar{\lambda}_N(\boldsymbol{\xi}_0, \phi(z_\tau)) = \frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^N \boldsymbol{\ell}^m_{it}(\boldsymbol{\xi}_{i\,0}, \phi(z_\tau))\right]. \tag{A-4}$$

Then, following the same argument as in the proof of Lemma 4, we can use the Weak Law of Large Numbers by McLeish (1975) and Lemma 2.1 of White and Domowitz (1984). Thus, we have, for any $z_\tau \in [0,1]$ and $\phi \in \mathcal{P}$,

$$\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \phi(z_\tau))}{NTh_{NT}} \xrightarrow{P} \bar{\lambda}_N(\boldsymbol{\xi}_0, \phi(z_\tau)), \;\text{ as }\; T \to \infty,\, Th_{NT} \to \infty,\, h_{NT} \to 0. \tag{A-5}$$

Furthermore, since, for any $z_\tau \in [0,1]$,

$$\sup_{\phi\in\Gamma} \frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \phi(z_\tau))}{NTh_{NT}} \xrightarrow{P} \sup_{\phi\in\Gamma} \bar{\lambda}_N(\boldsymbol{\xi}_0, \phi(z_\tau)), \text{ as } T\to\infty,\, Th_{NT}\to\infty,\, h_{NT}\to 0,$$

given Lemma 4, we have, for any $z_\tau \in [0,1]$,

$$\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau))}{NTh_{NT}} \xrightarrow{P} \bar{\lambda}_N(\boldsymbol{\xi}_0, \phi_0(z_\tau)), \text{ as } T\to\infty,\, Th_{NT}\to\infty,\, h_{NT}\to 0. \tag{A-6}$$

By applying (A-5) to the left hand side of (A-6), we have, for any $z_\tau \in [0,1]$, $\bar{\lambda}_N(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau)) \xrightarrow{P} \bar{\lambda}_N(\boldsymbol{\xi}_0, \phi_0(z_\tau))$ as $T\to\infty$,$Th_{NT}\to\infty$ and $h_{NT}\to 0$. Assumptions I and S imply that, for any $z_\tau \in [0,1]$, and for any N fixed, we have $\widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau) \xrightarrow{P} \phi_0(z_\tau)$ as $T\to\infty$, $Th_{NT}\to\infty$ and $h_{NT}\to 0$.

b) Now let us consider the case $N\to\infty$. By assumption C.2 we know that

$$\sup_{N\in\mathbb{N}} \left|\bar{\lambda}_N(\boldsymbol{\xi}_0, \phi(z_\tau))\right| < \infty.$$

Therefore, the following limit exists, and we define

$$\lambda(\boldsymbol{\xi}_0, \phi(z_\tau)) = \lim_{N\to\infty} \bar{\lambda}_N(\boldsymbol{\xi}_0, \phi(z_\tau)).$$

Summing up we have that for any $N \in \mathbb{N}$ $\phi_0(z_\tau) = \arg\sup_{\phi\in\Gamma} \lambda(\boldsymbol{\xi}_0, \phi(z_\tau))$. Moreover,

$$\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \phi(z_\tau))}{NTh_{NT}} \xrightarrow{P} \lambda(\boldsymbol{\xi}_0, \phi(z_\tau)), \text{ as } N,T\to\infty,\, Th_{NT}\to\infty,\, h_{NT}\to 0. \tag{A-7}$$

By the same arguments as before we have, for any $z_\tau \in [0,1]$,

$$\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau))}{NTh_{NT}} \xrightarrow{P} \lambda(\boldsymbol{\xi}_0, \phi_0(z_\tau)), \text{ as } N,T\to\infty,\, Th_{NT}\to\infty,\, h_{NT}\to 0,$$

and $\lambda(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau)) \xrightarrow{P} \lambda(\boldsymbol{\xi}_0, \phi_0(z_\tau))$, as $N,T\to\infty$, $Th_{NT}\to\infty$ and $h_{NT}\to 0$, which imply that, for any $z_\tau \in [0,1]$, we have $\widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau) \xrightarrow{P} \phi_0(z_\tau)$ as $N,T\to\infty$, $Th_{NT}\to\infty$ and $h_{NT}\to 0$. □

Proof of Theorem 2

Define $u_t = (z_\tau - z_t)/h_{NT}$, then, for $z_\tau \in (0,1)$, as $T\to\infty$ and $Th_{NT}\to\infty$, we have

$$\frac{1}{Th_{NT}} \sum_{t=1}^{T} g(u_t)\mathrm{K}(u_t) \to \int_{-1}^{1} g(u)\mathrm{K}(u)\mathrm{d}u, \tag{A-8}$$

for $g(u) = u^p$, with $p = 0,\ldots,3$ and $g(u) = u^q\mathrm{K}(u)$ with $q = 0,\ldots,6$. All integrals in (A-8) are finite because of assumption K.

Given the estimated curve $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}$, we have, for any $z_\tau \in [0,1]$ and for $i = 1, \dots, N$,

$$\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)) = \frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_0(z_\tau)) + \frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \bar{\phi}(z_\tau))(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau) - \phi_0(z_\tau)), \quad \text{(A-9)}$$

where $\bar{\phi}(z_\tau)$ lies between $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)$ and $\phi_0(z_\tau)$. Then, taking the first order conditions of (A-2), from (A-9), for any $z_\tau \in [0,1]$, we have

$$\begin{aligned}
0 \;=\; & \frac{1}{NTh_{NT}}\frac{\partial}{\partial\phi}\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)) = \underbrace{\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_t)) +}_{A_{NT}} \qquad \text{(A-10)}\\
+ & \underbrace{\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \bar{\phi}(z_t))\right](\phi_0(z_\tau) - \phi_0(z_t)) +}_{D_{NT}}\\
+ & \underbrace{\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_t))\right]}_{B_{NT}}(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau) - \phi_0(z_\tau)) +\\
+ & \underbrace{\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \bar{\phi}(z_t)) - \frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_t))\right]}_{C_{NT}}(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau) - \phi_0(z_\tau)).
\end{aligned}$$

By re-arranging (A-10) we obtain

$$\sqrt{NTh_{NT}}(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau) - \phi_0(z_\tau)) = -\frac{\sqrt{NTh_{NT}}(A_{NT} + D_{NT})}{(B_{NT} + C_{NT})}. \quad \text{(A-11)}$$

Let us consider each term on the right hand side of (A-11) separately.

$\mathbf{A_{NT}}$. We have

$$\begin{aligned}
\mathrm{E}_0[A_{NT}] \;=\; & \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_\tau + h_{NT}u_t))\mathrm{K}(u_t)\right] =\\
=\; & \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_\tau))\right] +\\
& + \frac{1}{NTh_{NT}}\sum_{t=1}^{T}h_{NT}u_t\phi_0'(z_\tau)\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_\tau))\right].
\end{aligned}$$

If we use (A-8) and assumption K we have, as $T \to \infty$ and $Th_{NT} \to \infty$,

$$\mathrm{E}_0[A_{NT}] \to \frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_\tau))\right] = 0. \quad \text{(A-12)}$$

If N is fixed the expectation in (A-12) is zero by Lemma 1. When $N \to \infty$, by assumption C.2, (A-12) is bounded for any N, and, again by Lemma 1, we have, as $T \to \infty$, $Th_{NT} \to \infty$, and

$N \to \infty$, that $\mathrm{E}_0[A_{NT}] \to 0$. Then,

$$\begin{aligned}\mathrm{E}_0\left[A_{NT}^2\right] &= \frac{1}{N^2T^2h_{NT}^2}\sum_{t=1}^{T}\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau+h_{NT}u_t))\right)^2\mathrm{K}^2(u_t)\right]+\\ &+\frac{1}{N^2T^2h_{NT}^2}\sum_{\substack{t,s=1\\t\neq s}}^{T}\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau+h_{NT}u_t)\right)\mathrm{K}(u_t)\right.\\ &\left.\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{is}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau+h_{NT}u_s))\right)\mathrm{K}(u_s)\right]=\mathrm{E}_0[A_{1NT}^2]+\mathrm{E}_0[A_{2NT}^2].\end{aligned}$$

By using (A-8) and assumption K, we have, as $T\to\infty$ and $Th_{NT}\to\infty$,

$$\mathrm{E}_0[A_{2NT}^2]\to\frac{1}{N^2}\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau)\right)\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{is}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)\right]=0, \quad \text{(A-13)}$$

by independence, as we are computing likelihoods in the true value of parameters, and Lemma 1. Moreover

$$\begin{aligned}\mathrm{E}_0\left[A_{1NT}^2\right] &= \frac{1}{N^2T^2h_{NT}^2}\sum_{t=1}^{T}\mathrm{K}^2(u_t)\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]+\\ &+\frac{1}{N^2T^2h_{NT}^2}\sum_{t=1}^{T}h_{NT}^2u_t^2\phi_0'^2(z_\tau)\mathrm{K}^2(u_t)\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right].\end{aligned}$$

If we use (A-13), (A-8), and assumption K, we have, as $T\to\infty$ and $Th_{NT}\to\infty$,

$$\begin{aligned}\mathrm{E}_0\left[A_{NT}^2\right] &\to \frac{1}{N^2Th_{NT}}\kappa_1\,\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]+\\ &+\frac{h_{NT}}{N^2T}\phi_0'^2(z_\tau)\kappa_2\,\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]=\\ &=\frac{1}{NTh_{NT}}\bar{\imath}_{N\boldsymbol{\xi}_o}(z_\tau)\left[\kappa_1+h_{NT}^2\phi_0'^2(z_\tau)\kappa_2\right], \qquad \text{(A-14)}\end{aligned}$$

where we have defined $\kappa_1=\int_{-1}^{1}\mathrm{K}^2(u)\mathrm{d}u$ and $\kappa_2=\int_{-1}^{1}u^2\mathrm{K}^2(u)\mathrm{d}u$. Therefore, since $\mathrm{E}_0[A_{NT}]\to 0$, as $T\to\infty$ and $Th_{NT}\to\infty$, we have

$$\mathrm{Var}_0\left[A_{NT}\right]\to\frac{1}{NTh_{NT}}\bar{\imath}_{N\boldsymbol{\xi}_o}(z_\tau)\left[\kappa_1+h_{NT}^2\phi_0'^2(z_\tau)\kappa_2\right]. \qquad \text{(A-15)}$$

We can then apply the Weak Law of Large Numbers to A_{NT} which implies $A_{NT}\xrightarrow{P}0,\;$ as $T\to\infty$, $Th_{NT}\to\infty$. If assumptions C.2 and D hold and we define $i_{\boldsymbol{\xi}_o}(z_\tau)=\lim_{N\to\infty}\bar{\imath}_{N\boldsymbol{\xi}_o}(z_\tau)$, this limit exists and is bounded. Thus when $T\to\infty$, $Th_{NT}\to\infty$, and $N\to\infty$

$$\mathrm{Var}_0\left[A_{NT}\right]\to\frac{1}{NTh_{NT}}i_{\boldsymbol{\xi}_o}(z_\tau)\left[\kappa_1+h_{NT}^2\phi_0'^2(z_\tau)\kappa_2\right]. \qquad \text{(A-16)}$$

Analogously we have $A_{NT} \xrightarrow{P} 0$, as $N, T \to \infty$, $Th_{NT} \to \infty$.

$\mathbf{B_{NT}}$. Following the same reasoning we have that

$$\begin{aligned}
\mathrm{E}_0[B_{NT}] &= \frac{1}{NTh_{NT}} \sum_{t=1}^{T} \mathrm{K}(u_t) \mathrm{E}_0 \left[\sum_{i=1}^{N} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right] + \\
&+ \frac{1}{NTh_{NT}} \sum_{t=1}^{T} h_{NT} u_t \phi_0'(z_\tau) \mathrm{K}(u_t) \mathrm{E}_0 \left[\sum_{i=1}^{N} \frac{\partial^3}{\partial \phi^3} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right].
\end{aligned}$$

If we use (A-8) and assumption K we have, as $T \to \infty$ and $Th_{NT} \to \infty$,

$$\mathrm{E}_0[B_{NT}] \to \frac{1}{N} \mathrm{E}_0 \left[\sum_{i=1}^{N} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right] = -\bar{\jmath}_{N\boldsymbol{\xi}_o}(z_\tau). \tag{A-17}$$

If assumption C.2 holds then $\lim_{N\to\infty} \bar{\jmath}_{N\boldsymbol{\xi}_o}(z_\tau) = j_{\boldsymbol{\xi}_o}(z_\tau)$ exists and is finite and, as $T \to \infty$, $Th_{NT} \to \infty$, and $N \to \infty$, $\mathrm{E}_0[B_{NT}] \to = -j_{\boldsymbol{\xi}_o}(z_\tau)$. Then let us compute the variance. We have

$$\begin{aligned}
\mathrm{E}_0[B_{NT}^2] &= \frac{1}{N^2T^2h_{NT}^2} \sum_{t=1}^{T} \mathrm{E}_0 \left[\left(\sum_{i=1}^{N} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau + h_{NT}u_t)) \right)^2 \mathrm{K}^2(u_t) \right] + \\
&+ \frac{1}{N^2T^2h_{NT}^2} \sum_{\substack{t,s=1 \\ t\neq s}}^{T} \mathrm{E}_0 \left[\left(\sum_{i=1}^{N} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau + h_{NT}u_t)) \right) \mathrm{K}(u_t) \right. \\
&\left. \left(\sum_{i=1}^{N} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{is}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau + h_{NT}u_s)) \right) \mathrm{K}(u_s) \right] = \mathrm{E}_0[B_{1NT}^2] + \mathrm{E}_0[B_{2NT}^2].
\end{aligned}$$

By using (A-8) and assumption K, we can prove that, as $T \to \infty$ and $Th_{NT} \to \infty$, we have $\mathrm{E}_0[B_{2NT}^2] \to \bar{\jmath}_{N\boldsymbol{\xi}_o}^2(z_\tau)$. the other term in (**??**) becomes

$$\begin{aligned}
\mathrm{e}_0[b_{1nt}^2] &= \frac{1}{n^2t^2h_{nt}^2} \sum_{t=1}^{t} \mathrm{k}^2(u_t) \mathrm{e}_0 \left[\left(\sum_{i=1}^{n} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right)^2 \right] + \\
&+ \frac{1}{n^2t^2h_{nt}^2} \sum_{t=1}^{t} h_{nt}^2 u_t^2 \phi_0'^2(z_\tau) \mathrm{k}^2(u_t) \mathrm{e}_0 \left[\left(\sum_{i=1}^{n} \frac{\partial^3}{\partial \phi^3} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right)^2 \right] + \\
&+ \frac{2}{n^2t^2h_{nt}^2} \sum_{t=1}^{t} h_{nt} u_t \phi_0'(z_\tau) \mathrm{k}^2(u_t) \mathrm{e}_0 \left[\left(\sum_{i=1}^{n} \frac{\partial^2}{\partial \phi^2} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right) \right. \\
&\left. \left(\sum_{i=1}^{n} \frac{\partial^3}{\partial \phi^3} \boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0}, \phi_0(z_\tau)) \right) \right].
\end{aligned}$$

if we use (**??**) and assumption k we have, as $t \to \infty$ and $th_{nt} \to \infty$,

$$\begin{aligned}
\mathrm{e}_0\left[b_{1nt}^2\right] &\to \frac{1}{n^2 th_{nt}}\kappa_1\,\mathrm{e}_0\left[\left(\sum_{i=1}^{n}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]+\\
&+ \frac{h_{nt}}{n^2 t}\phi_0'^2(z_\tau)\kappa_2\,\mathrm{e}_0\left[\left(\sum_{i=1}^{n}\frac{\partial^3}{\partial\phi^3}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]=\\
&= \frac{1}{th_{nt}}\kappa_1\bar{\jmath}_n(z_\tau)+\frac{h_{nt}}{t}\phi_0'^2(z_\tau)\kappa_2\bar{\amalg}_n(z_\tau), \qquad \text{(A-18)}
\end{aligned}$$

where κ_1 and κ_2 are defined above, $\kappa_3 = \int_{-1}^{1} u^2 \mathrm{k}(u)\mathrm{d}u$ and

$$\begin{aligned}
\bar{\jmath}_n(z_\tau) &= \frac{1}{n^2}\mathrm{e}_0\left[\left(\sum_{i=1}^{n}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right],\\
\bar{\amalg}_n(z_\tau) &= \frac{1}{n^2}\mathrm{e}_0\left[\left(\sum_{i=1}^{n}\frac{\partial^3}{\partial\phi^3}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]. \qquad \text{(A-19)}
\end{aligned}$$

Then we have, as $T \to \infty$ and $Th_{NT} \to \infty$,

$$\mathrm{Var}_0[B_{NT}] \to \frac{1}{Th_{NT}}\kappa_1\bar{\mathcal{S}}_N(z_\tau)+\frac{h_{NT}}{T}\phi_0'^2(z_\tau)\kappa_2\bar{\mathcal{Q}}_N(z_\tau).$$

We can then apply the Weak Law of Large Numbers to B_{NT} which implies $B_{NT} \xrightarrow{P} -\bar{\jmath}_{N\boldsymbol{\xi}_o}(z_\tau)$ as $T \to \infty$ and $Th_{NT} \to \infty$. Moreover, if assumption C.2 holds, then $j_{\boldsymbol{\xi}_o}(z_\tau) = \lim_{N\to\infty}\bar{\jmath}_{N\boldsymbol{\xi}_o}(z_\tau)$, $\mathcal{S}(z_\tau) = \lim_{N\to\infty}\bar{\mathcal{S}}_N(z_\tau)$ and $\mathcal{Q}(z_\tau) = \lim_{N\to\infty}\bar{\mathcal{Q}}_N(z_\tau)$ exist and are finite. Therefore, $B_{NT} \xrightarrow{P} -j_{\boldsymbol{\xi}_o}(z_\tau)$ as $N, T \to \infty$ and $Th_{NT} \to \infty$.

$\mathbf{C_{NT}}$. By Theorem 1 we know that, when N is fixed, $\widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau) \xrightarrow{P} \phi_0(z_\tau)$, as $T \to \infty$, $Th_{NT} \to \infty$. Since for any $z_\tau \in [0,1]$, we have $|\bar{\phi}(z_\tau) - \phi_0(z_\tau)| \leq |\widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau) - \phi_0(z_\tau)|$, we also have $\bar{\phi}(z_\tau) \xrightarrow{P} \phi_0(z_\tau)$, as $T \to \infty$, $Th_{NT} \to \infty$. which implies

$$|C_{NT}| \leq \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\left|\frac{\partial^3}{\partial\phi^3}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_t))\right|\left|\widehat{\phi}_{\boldsymbol{\xi}_o NT}(z_\tau)-\phi_0(z_\tau)\right|.$$

Therefore, $|C_{NT}| \xrightarrow{P} 0$ as $T \to \infty$, $Th_{NT} \to \infty$. If assumption C.2 holds then all terms in the previous inequality are bounded even when we let $N \to \infty$. Then, by using Theorem 1, we have $|C_{NT}| \xrightarrow{P} 0$ as $N, T \to \infty$, $Th_{NT} \to \infty$.

$\mathbf{D_{NT}}$. This is the bias term and it can be decomposed as

$$
\begin{aligned}
D_{NT} &= \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_t))(\phi_0(z_\tau)-\phi_0(z_t)) + \\
&+\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^3}{\partial\phi^3}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_t))(\bar{\phi}(z_t)-\phi_0(z_t))(\phi_0(z_\tau)-\phi_0(z_t)) = \\
&= D_{1NT} + D_{2NT}.
\end{aligned}
$$

Since for any $z_\tau \in [0,1]$, we have $|\bar{\phi}(z_\tau)-\phi_0(z_\tau)| \leq |\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau)-\phi_0(z_\tau)|$, we have $|D_{2NT}| \xrightarrow{P} 0$ as $T\to\infty$, $Th_{NT}\to\infty$, by Theorem 1. If assumption C.2 holds this is true even when $N\to\infty$. Let us consider D_{1NT}, we have:

$$
\begin{aligned}
\mathrm{E}_0[D_{1NT}] &= \frac{1}{NTh_{NT}}\mathrm{E}_0\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_t))(\phi_0(z_\tau)-\phi_0(z_t))\right] = \\
&= \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau+h_{NT}u_t))(\phi_0(z_\tau)-\phi_0(z_\tau+h_{NT}u_t))\mathrm{K}(u_t)\right] = \\
&= \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\left(h_{NT}u_t\phi_0'(z_\tau)+\frac{h_{NT}^2u_t^2}{2}\phi_0''(z_\tau)\right)\mathrm{K}(u_t)\left\{\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right]+\right. \\
&\left.+\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^3}{\partial\phi^3}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))h_{NT}u_t\phi_0'(z_\tau))\right]\right\} = \\
&= \frac{1}{NTh_{NT}}\sum_{t=1}^{T}h_{NT}u_t\phi_0'(z_\tau)\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right]+ \\
&+\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\frac{h_{NT}^2u_t^2}{2}\phi_0''(z_\tau)\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right]+ \\
&+\frac{1}{NTh_{NT}}\sum_{t=1}^{T}h_{NT}^2u_t^2\phi_0'^2(z_\tau)\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^3}{\partial\phi^3}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right]+ \\
&+\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\frac{h_{NT}^3u_t^3}{2}\phi_0''(z_\tau)\phi_0'(z_\tau)\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right].
\end{aligned}
$$

If we use (A-8) and assumption K we have, as $T\to\infty$ and $Th_{NT}\to\infty$,

$$
\begin{aligned}
\mathrm{E}_0[D_{1NT}] &\to h_{NT}^2\kappa_3\frac{\phi_0''(z_\tau)}{2}\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right]+ \\
&+ h_{NT}^2\kappa_3\phi_0'^2(z_\tau)\frac{\partial}{\partial\phi}\frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_{i\,0},\phi_0(z_\tau))\right] = \\
&= h_{NT}^2\kappa_3\left[-\frac{\phi_0''(z_\tau)}{2}\bar{j}_{N\boldsymbol{\xi}_o}(z_\tau)-\phi_0'^2(z_\tau)\frac{\partial}{\partial\phi}\bar{j}_{N\boldsymbol{\xi}_o}(z_\tau)\right] = \\
&= h_{NT}^2\kappa_3\bar{\mathcal{B}}_N(z_\tau), \qquad \text{(A-20)}
\end{aligned}
$$

where $\bar{j}_{N\boldsymbol{\xi}_o}(z_\tau)$ is defined in (A-17), κ_3 is defined in (**??**) and

$$\bar{\mathcal{B}}_N(z_\tau) = -\bar{j}_{N\boldsymbol{\xi}_o}(z_\tau)\frac{\phi_0''(z_\tau)}{2} - \frac{\partial}{\partial\phi}\bar{j}_{N\boldsymbol{\xi}_o}(z_\tau)\phi_0'^2(z_\tau).$$

Now, let us compute the variance.

$$\begin{aligned} \mathrm{E}_0\left[D_{1NT}^2\right] &= \frac{1}{N^2T^2h_{NT}^2}\sum_{t=1}^{T}\left(\phi_0(z_\tau)-\phi_0(z_\tau+h_{NT}u_t)\right)^2\mathrm{K}^2(u_t) \\ &\quad \mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau+h_{NT}u_t))\right)^2\right]+ \\ &\quad +\frac{1}{N^2T^2h_{NT}^2}\sum_{\substack{t,s=1\\t\neq s}}^{T}\left(\phi_0(z_\tau)-\phi_0(z_\tau+h_{NT}u_t)\right)\left(\phi_0(z_\tau)-\phi_0(z_\tau+h_{NT}u_s)\right)\mathrm{K}(u_t)\mathrm{K}(u_s) \\ &\quad \mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau+h_{NT}u_t))\right)\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau+h_{NT}u_s))\right)\right]= \\ &= \mathrm{E}_0[D_{1,1NT}^2]+\mathrm{E}_0[D_{1,2NT}^2]. \end{aligned} \tag{A-21}$$

By using (A-8) and assumption K, we can prove that, as $T\to\infty$ and $Th_{NT}\to\infty$,

$$\mathrm{E}_0[D_{1,2NT}^2]\to O(h_{NT}^3). \tag{A-22}$$

The other term in (A-21) becomes

$$\mathrm{E}_0[D_{1,1NT}^2] = \frac{1}{N^2T^2h_{NT}^2}\sum_{t=1}^{T}h_{NT}^2u_t^2\phi_0'^2(z_\tau)\mathrm{K}^2(u_t)\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]+O(h_{NT}^3).$$

If we use (A-8) and assumption K we have, as $T\to\infty$ and $Th_{NT}\to\infty$,

$$\begin{aligned} \mathrm{E}_0[D_{1,1NT}^2] &\to \frac{h_{NT}}{N^2T}\phi_0'^2(z_\tau)\kappa_2\,\mathrm{E}_0\left[\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\phi^2}\boldsymbol{\ell}_{it}^m(\boldsymbol{\xi}_{i0},\phi_0(z_\tau))\right)^2\right]+O(h_{NT}^3)= \\ &= \frac{h_{NT}}{T}\phi_0'^2(z_\tau)\kappa_2\bar{\mathcal{S}}_N(z_\tau)+O(h_{NT}^3), \end{aligned} \tag{A-23}$$

where κ_2 is defined in (**??**) and $\bar{\mathcal{S}}_N(z_\tau)$ is defined in (**??**). Therefore keeping only terms up to $O(h_{NT}^2)$ we have, as $T\to\infty$ and $Th_{NT}\to\infty$,

$$\mathrm{Var}_0\left[D_{1NT}\right]\to\frac{h_{NT}}{T}\phi_0'^2(z_\tau)\kappa_2\bar{\mathcal{S}}_N(z_\tau). \tag{A-24}$$

Since by $D_{NT}=D_{1NT}+o_P(1)$, as $T\to\infty$ and $Th_{NT}\to\infty$, we can then apply the Weak Law of Large Numbers to D_{NT} which implies $D_{NT}\xrightarrow{P}h_{NT}^2\kappa_3\bar{\mathcal{B}}_N(z_\tau)$ as $T\to\infty$ and $Th_{NT}\to\infty$. Moreover, if assumption C.2 holds, then $\mathcal{B}(z_\tau)=\lim_{N\to\infty}\bar{\mathcal{B}}_N(z_\tau)$ and $\mathcal{S}(z_\tau)=\lim_{N\to\infty}\bar{\mathcal{S}}_N(z_\tau)$ exist and are finite. Thus, as $T\to\infty$, $Th_{NT}\to\infty$, and $N\to\infty$,

$$\mathrm{Var}_0\left[D_{1NT}\right]\to\frac{h_{NT}}{T}\phi_0'^2(z_\tau)\kappa_2\mathcal{S}(z_\tau), \tag{A-25}$$

and $D_{NT} \xrightarrow{P} h_{NT}^2 \kappa_3 \mathcal{B}(z_\tau)$.

We now consider the limiting distribution of (A-11) up to terms of order $O(h_{NT}^2)$. First consider the case $i)$ in Theorem 2, i.e. when N is fixed and T is large. Then, from (A-11) and using (**??**) and (**??**) we have

$$\left(\widehat{\phi}_{\boldsymbol{\xi}_o T}(z_\tau) - \phi_0(z_\tau) - h_{NT}^2 \bar{\mathcal{B}}_N(z_\tau)\kappa_3\right) \xrightarrow{P} 0 \text{ as } T \to \infty, \, Th_{NT} \to \infty.$$

By applying the Central Limit Theorem by Wooldridge and White (1988) and Slutsky's Theorem, we have convergence in distribution, as $T \to \infty$ and $Th_{NT} \to \infty$:

$$\sqrt{NTh_{NT}} \left(\widehat{\phi}_{\boldsymbol{\xi}_o T}(z_\tau) - \phi_0(z_\tau) - h_{NT}^2 \bar{\mathcal{B}}_N(z_\tau)\kappa_3\right) \xrightarrow{d} \mathcal{N}\left(0, \widetilde{V}_{\boldsymbol{\xi}_o}(z_\tau) + W_{\boldsymbol{\xi}_o}(z_\tau) + U_{\boldsymbol{\xi}_o}(z_\tau)\right).$$

The asymptotic variance is

$$\begin{aligned} \widetilde{V}_{\boldsymbol{\xi}_o}(z_\tau) &= NTh_{NT} \frac{\mathrm{Var}_0\left[A_{NT}\right]}{\mathrm{E}_0\left[B_{NT}\right]^2} = \frac{\bar{i}_{N\,\boldsymbol{\xi}_o}(z_\tau)}{\bar{j}^2_{N\,\boldsymbol{\xi}_o}(z_\tau)}\left[\kappa_1 + h_{NT}^2 \phi_0'^2(z_\tau)\kappa_2\right], \qquad \text{(A-26)} \\ W_{\boldsymbol{\xi}_o}(z_\tau) &= NTh_{NT} \frac{\mathrm{Var}_0\left[D_{NT}\right]}{\mathrm{E}_0\left[B_{NT}\right]^2} = Nh_{NT}^2 \frac{\phi_0'^2(z_\tau)\bar{\mathcal{S}}_N(z_\tau)}{\bar{j}^2_{N\,\boldsymbol{\xi}_o}(z_\tau)}, \\ U_{\boldsymbol{\xi}_o}(z_\tau) &= NTh_{NT} \frac{2\,\mathrm{Cov}_0\left[A_{NT}, D_{NT}\right]}{\mathrm{E}_0\left[B_{NT}\right]^2}. \end{aligned}$$

If we also let $h_{NT} \to 0$, and by the Cauchy–Schwarz inequality, we have $\widetilde{V}_{\boldsymbol{\xi}_o}(z_\tau) \to V_{\boldsymbol{\xi}_o}(z_\tau)$, $W_{\boldsymbol{\xi}_o}(z_\tau) \to 0$, $U_{\boldsymbol{\xi}_o}(z_\tau) \leq \sqrt{2V_{\boldsymbol{\xi}_o}(z_\tau)W_{\boldsymbol{\xi}_o}(z_\tau)} \to 0$. Moreover, also the bias term becomes negligible and we have the claim of the result of part *i*) of the Theorem.

Now let us consider case $ii)$ in Theorem 2, i.e. when both N and T are large. By assuming that assumptions C.2 and D hold

$$\left(\widehat{\phi}_{\boldsymbol{\xi}_o T}(z_\tau) - \phi_0(z_\tau) - h_{NT}^2 \mathcal{B}(z_\tau)\kappa_3\right) \xrightarrow{P} 0 \text{ as } N, T \to \infty, \, NTh_{NT} \to \infty.$$

Then, by applying the Central Limit Theorem by Wooldridge and White (1988) and Slutsky's Theorem, we have convergence in distribution, as $N, T \to \infty$ and $NTh_{NT} \to \infty$:

$$\sqrt{NTh_{NT}} \left(\widehat{\phi}_{\boldsymbol{\xi}_o T}(z_\tau) - \phi_0(z_\tau) - h_{NT}^2 \mathcal{B}(z_\tau)\kappa_3\right) \xrightarrow{d} \mathcal{N}\left(0, \widetilde{V}_{\boldsymbol{\xi}_o}(z_\tau) + W_{\boldsymbol{\xi}_o}(z_\tau) + U_{\boldsymbol{\xi}_o}(z_\tau)\right),$$

where

$$\widetilde{V}_{\boldsymbol{\xi}_o}(z_\tau) = \frac{i_{\boldsymbol{\xi}_o}(z_\tau)}{j^2_{\boldsymbol{\xi}_o}(z_\tau)}\left[\kappa_1 + h_{NT}^2 \phi_0'^2(z_\tau)\kappa_2\right],$$

and

$$W_{\boldsymbol{\xi}_o}(z_\tau) = Nh_{NT}^2 \frac{\phi_0'^2(z_\tau)\mathcal{S}(z_\tau)}{j^2_{\boldsymbol{\xi}_o}(z_\tau)}, \quad U_{\boldsymbol{\xi}_o}(z_\tau) = \lim_{N \to \infty} NTh_{NT} \frac{2\,\mathrm{Cov}_0\left[A_{NT}, D_{NT}\right]}{\mathrm{E}_0\left[B_{NT}\right]^2}.$$

If we also let $Nh_{NT}^2 \to 0$ which implies $h_{NT} \to 0$, and by the Cauchy–Schwarz inequality, we have $\widetilde{V}_{\boldsymbol{\xi}_o}(z_\tau) \to V_{\boldsymbol{\xi}_o}(z_\tau)$, $W_{\boldsymbol{\xi}_o}(z_\tau) \to 0$, $U_{\boldsymbol{\xi}_o}(z_\tau) \leq \sqrt{2V_{\boldsymbol{\xi}_o}(z_\tau)W_{\boldsymbol{\xi}_o}(z_\tau)} \to 0$. Moreover, also the bias term becomes negligible and we have the claim of the result of part *ii*) of the Theorem. □

Proof of Theorem 3

a) From the first order conditions of (A-2), for any $z_\tau \in [0,1]$ and any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, we have

$$\begin{aligned}\mathbf{0} = \frac{\partial}{\partial \boldsymbol{\xi}}\left(\frac{\partial \widetilde{\mathcal{L}}_{NT}}{\partial \phi}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_\tau))\right) \quad &= \quad \sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \boldsymbol{\xi}\partial \phi}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t)) + \\ &+ \sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \phi^2}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t))\, \widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}(z_t).\end{aligned}$$

Solving for $\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}$ we have, for any $z_\tau \in [0,1]$,

$$\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}(z_\tau) = -\frac{\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \boldsymbol{\xi}\partial \phi}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t))}{\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \phi^2}(\boldsymbol{\xi}_0, \widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t))} = -\frac{\boldsymbol{\alpha}_{NT}(z_\tau)}{\beta_{NT}(z_\tau)}. \quad \text{(A-27)}$$

Moreover, using a Taylor expansion in a neighborhood of ϕ_0, and for T sufficiently large,

$$\begin{aligned}&\left|\left|\boldsymbol{\alpha}_{NT}(z_\tau) - \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \boldsymbol{\xi}\partial \phi}(\boldsymbol{\xi}_0, \phi_0(z_t))\right|\right|_2 \leq \\ &\leq \left|\left|\frac{1}{NTh_{NT}}\frac{\partial}{\partial \phi}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \boldsymbol{\xi}\partial \phi}(\boldsymbol{\xi}_0, \bar{\phi}(z_t))\right|\right|_2 \sup_{z_t \in [0,1]}\left|\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t) - \phi_0(z_t)\right| = o_P(1),\end{aligned}$$

where $\bar{\phi}$ is between $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}$ and ϕ_0. The previous result is a consequence of Theorem 1 on the consistency of $\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}$ and assumption S. Similarly we can prove, for any $z_\tau \in [0,1]$, and for T sufficiently large,

$$\left|\beta_{NT}(z_\tau) - \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \phi^2}(\boldsymbol{\xi}_0, \phi_0(z_t))\right| = o_P(1).$$

Using calculations similar to those in the proof of Theorem 2 and by applying the Weak Law of Large Numbers, we have, for any $z_\tau \in [0,1]$, as $T \to \infty$, $Th_{NT} \to \infty$, and $h_{NT} \to 0$,

$$\begin{aligned}\boldsymbol{\alpha}^*_{NT}(z_\tau) &\equiv \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \boldsymbol{\xi}\partial \phi}(\boldsymbol{\xi}_0, \phi_0(z_t)) \quad \xrightarrow{P} \quad \bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_o}(z_\tau), \\ \beta^*_{NT}(z_\tau) &\equiv \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right)\frac{\partial^2 \boldsymbol{\ell}_{i\,t}^m}{\partial \phi^2}(\boldsymbol{\xi}_0, \phi_0(z_t)) \quad \xrightarrow{P} \quad -\bar{j}_{N\,\boldsymbol{\xi}_o}(z_\tau).\end{aligned}$$

We then have that as $T \to \infty$, $Th_{NT} \to \infty$, and $h_{NT} \to 0$,

$$\left|\left|\boldsymbol{\alpha}_{NT}(z_\tau) - \bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_o}(z_\tau)\right|\right|_2 \;\leq\; ||\boldsymbol{\alpha}_{NT}(z_\tau) - \boldsymbol{\alpha}^*_{NT}(z_\tau)||_2 + \left|\left|\boldsymbol{\alpha}^*_{NT}(z_\tau) - \bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_o}(z_\tau)\right|\right|_2 = o_P(1),$$

$$|\beta_{NT}(z_\tau) - (-\bar{j}_{N\,\boldsymbol{\xi}_o}(z_\tau))| \;\leq\; |\beta_{NT}(z_\tau) - \beta^*_{NT}(z_\tau)| + |\beta^*_{NT}(z_\tau) - (-\bar{j}_{N\,\boldsymbol{\xi}_o}(z_\tau))| = o_P(1),$$

which, substituted in (A-27), implies $\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}(z_\tau) \xrightarrow{P} \frac{\bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_o}(z_\tau)}{\bar{j}_{N\,\boldsymbol{\xi}_o}(z_\tau)}$ as $T \to \infty$, $Th_{NT} \to \infty$ and $h_{NT} \to 0$.

b) If assumption C.2 holds then we know that $\lim_{N\to\infty} \bar{j}_{N\,\boldsymbol{\xi}_o}(z_\tau) = j_{\boldsymbol{\xi}_o}(z_\tau)$ exists and is finite. Since each marginal depends only on its parameters $\boldsymbol{\xi}_i$, we have

$$\begin{aligned}\bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_o}(z_\tau) &= \frac{1}{N}\left\{ \mathrm{E}_0 \begin{bmatrix} \frac{\partial^2 \boldsymbol{\ell}^m_{1\,t}}{\partial \boldsymbol{\xi}_1 \partial \phi}(\boldsymbol{\xi}_{1\,0}, \phi_0(z_\tau)) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} + \ldots + \mathrm{E}_0 \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \frac{\partial^2 \boldsymbol{\ell}^m_{N\,t}}{\partial \boldsymbol{\xi}_N \partial \phi}(\boldsymbol{\xi}_{N\,0}, \phi_0(z_\tau)) \end{bmatrix} \right\} = \\ &= \frac{1}{N}\mathrm{E}_0 \begin{bmatrix} \frac{\partial^2 \boldsymbol{\ell}^m_{1\,t}}{\partial \boldsymbol{\xi}_1 \partial \phi}(\boldsymbol{\xi}_{1\,0}, \phi_0(z_\tau)) \\ \vdots \\ \frac{\partial^2 \boldsymbol{\ell}^m_{N\,t}}{\partial \boldsymbol{\xi}_N \partial \phi}(\boldsymbol{\xi}_{N\,0}, \phi_0(z_\tau)) \end{bmatrix}.\end{aligned}$$

Therefore, $\lim_{N\to\infty} \bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_o}(z_\tau) = \mathbf{0}$, for any $z_\tau \in [0,1]$. □

Proof of Corollary 1

We define $\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau)$ such that it solves (A-2) when computed in generic values of the parameters $\boldsymbol{\xi} \in \Xi$. According to assumption N, the limit in probability of $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$ exists and we denote it as $\widetilde{\phi}_{\boldsymbol{\xi}}$. Lemma 4 proofs that assumption N is always satisfied. In order to prove the Corollary, we have to show that $\widetilde{\phi}_{\boldsymbol{\xi}}$ is a least favorable curve, i.e. its derivative is equal to the least favorable direction, which for $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ is defined in (18).

First, notice that, if $\boldsymbol{\xi} = \boldsymbol{\xi}_0$, the result is a direct consequence of Theorems 1 and 3. Therefore, $\widetilde{\phi}_{\boldsymbol{\xi}_o} = \phi_0$. Moreover, from Theorem 3.a, we have that $\widetilde{\phi}'_{\boldsymbol{\xi}_o} = \phi'_0$. Thus, $\widetilde{\phi}_{\boldsymbol{\xi}_o} = \phi_0$ is a least favorable curve.

Now let us move to the case in $\boldsymbol{\xi} \neq \boldsymbol{\xi}_0$. From Lemma 4 we have that, for any $z_\tau \in [0,1]$ and any $\boldsymbol{\xi} \in \Xi$, as $T \to \infty$, $Th_{NT} \to \infty$, $h_{NT} \to 0$, $\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) \xrightarrow{P} \phi_{\boldsymbol{\xi}}(z_\tau)$, and $\widehat{\phi}'_{\boldsymbol{\xi}\,NT}(z_\tau) \xrightarrow{P} \phi'_{\boldsymbol{\xi}}(z_\tau)$, where $\phi_{\boldsymbol{\xi}}(z_\tau)$ is such that it solves (A-4). Therefore, assumption N is always satisfied with $\widetilde{\phi}_{\boldsymbol{\xi}} = \phi_{\boldsymbol{\xi}}$. Finally, by using the same arguments as in Theorem 3.a, and a Taylor series approximation of the first order conditions of (A-4) in a neighborhood of $\phi_{\boldsymbol{\xi}}$, we can prove that

$$\phi'_{\boldsymbol{\xi}} = \frac{\bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}}(z_\tau)}{\bar{j}_{N\,\boldsymbol{\xi}}(z_\tau)}, \quad \text{as } T \to \infty, Th_{NT} \to \infty, h_{NT} \to 0, \tag{A-28}$$

where $\bar{j}_{N\,\boldsymbol{\xi}}(z_\tau)$ and $\bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}}(z_\tau)$ are analogous to the ones defined in Theorem 3.a. Therefore, by comparing (A-28) with (18) in the paper, we recognize $\phi'_{\boldsymbol{\xi}}$ as a least favorable direction, which implies that $\phi_{\boldsymbol{\xi}}$ is a least favorable curve and $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$ is its estimator. To conclude also notice that, when considering $\boldsymbol{\xi} = \boldsymbol{\xi}_0$, we have $\phi_0 = \phi_{\boldsymbol{\xi}_o}$ and $\phi'_0 = \phi'_{\boldsymbol{\xi}_o}$ where the latter is defined in (18). □

Proof of Theorem 4

a) Given the estimator $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$ in (12), we define, for any $i = 1, \ldots, N$,

$$\mathcal{L}^m_{iT}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) = \sum_{t=1}^{T} \boldsymbol{\ell}^m_{it}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}).$$

We first prove consistency of $\widehat{\boldsymbol{\xi}}_{i\,T}$, such that (see (19)),

$$\widehat{\boldsymbol{\xi}}_{i\,T} = \arg\max_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}). \tag{A-29}$$

We also define the function

$$\gamma_i(\boldsymbol{\xi}_i) = \mathrm{E}_0\left[\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right],$$

where $\phi_{\boldsymbol{\xi}}$ is the least favorable curve computed in a generic value of the parameters $\boldsymbol{\xi}$. From Lemma 1 and assumption H the true value of the parameters is such that

$$\boldsymbol{\xi}_{i\,0} = \arg\max_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \gamma_i(\boldsymbol{\xi}_i). \tag{A-30}$$

Then, following the same argument as in the proof of Lemma 4 and using (A-51), we can use the Weak Law of Large Numbers by McLeish (1975) and Lemma 2.1 of White and Domowitz (1984), which imply that, for any $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$,

$$\frac{1}{T}\mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}}) \xrightarrow{P} \gamma_i(\boldsymbol{\xi}_i), \ \text{ as } \ T \to \infty. \tag{A-31}$$

Moreover,

$$\begin{aligned}
&\frac{1}{T}\left|\mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) - \mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right| \le \frac{1}{T}\sum_{t=1}^T \left|\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) - \boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right| = \\
&= \frac{1}{T}\sum_{t=1}^T \left|\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right| \sup_{z_\tau \in [0,1]} \left|\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)\right| \le \\
&\le \frac{1}{T}\sum_{t=1}^T \sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \left|\frac{\partial}{\partial\phi}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right| \sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \sup_{z_\tau \in [0,1]} \left|\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)\right|.
\end{aligned}$$

From Theorem 1 we have (notice that if assumption C.2 holds the following is true also when $N \to \infty$)

$$\sup_{z_\tau \in [0,1]} \left|\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)\right| = o_P(1), \ \text{ as } \ T \to \infty,$$

therefore, for any $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$,

$$\frac{1}{T}\left|\mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) - \mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right| \xrightarrow{P} 0, \ \text{ as } \ T \to \infty. \tag{A-32}$$

By combining (A-31) and (A-32), we have, for any $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$,

$$\frac{1}{T}\mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) \xrightarrow{P} \gamma_i(\boldsymbol{\xi}_i), \ \text{ as } \ T \to \infty. \tag{A-33}$$

Furthermore,

$$\sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \frac{1}{T}\mathcal{L}^m_{i\,T}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) \xrightarrow{P} \sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \gamma_i(\boldsymbol{\xi}_i), \ \text{ as } \ T \to \infty,$$

which, by means of (A-29) and (A-30), is equivalent to

$$\frac{1}{T}\mathcal{L}^m_{i\,T}(\widehat{\boldsymbol{\xi}}_{i\,T}, \widehat{\phi}_{\widehat{\boldsymbol{\xi}}_T\,NT}) \xrightarrow{P} \gamma_i(\boldsymbol{\xi}_{i\,0}), \ \text{ as } \ T \to \infty,$$

where the curve is now estimated in $\widehat{\boldsymbol{\xi}}_T = (\widehat{\boldsymbol{\xi}}_{1\,T}^{\mathrm{T}} \dots \widehat{\boldsymbol{\xi}}_{N\,T}^{\mathrm{T}})^{\mathrm{T}}$. From (A-33), the term on the left hand side is such that

$$\frac{1}{T}\mathcal{L}_{iT}^m(\widehat{\boldsymbol{\xi}}_{i\,T}, \widehat{\phi}_{\widehat{\boldsymbol{\xi}}_T\,NT}) \xrightarrow{P} \gamma_i(\widehat{\boldsymbol{\xi}}_{i\,T}), \text{ as } T \to \infty,$$

thus

$$\gamma_i(\widehat{\boldsymbol{\xi}}_{i\,T}) \xrightarrow{P} \gamma_i(\boldsymbol{\xi}_{i\,0}), \text{ as } T \to \infty.$$

Given assumptions I and S, we have consistency, for any $i = 1, \dots, N$,

$$\widehat{\boldsymbol{\xi}}_{i\,T} \xrightarrow{P} \boldsymbol{\xi}_{i\,0}, \text{ as } T \to \infty. \tag{A-34}$$

b) Given a consistent estimator of the curve $\widehat{\phi}_{\boldsymbol{\xi}\,NT}$, the estimated marginals' parameters, $\widehat{\boldsymbol{\xi}}_{1\,T}, \dots, \widehat{\boldsymbol{\xi}}_{N\,T}$, have to satisfy (19). First order conditions and a Taylor series expansion around the true values $\boldsymbol{\xi}_{i\,0}$ give, for any $i = 1, \dots, N$,

$$\mathbf{0} = \frac{1}{T}\frac{\partial \mathcal{L}_{iT}^m}{\partial \boldsymbol{\xi}_i}(\widehat{\boldsymbol{\xi}}_{i\,T}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) = \underbrace{\frac{1}{T}\frac{\partial \mathcal{L}_{iT}^m}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \widehat{\phi}_{\boldsymbol{\xi}\,NT})}_{\mathcal{A}_{i\boldsymbol{\xi}_i\,T}} + \underbrace{\frac{1}{T}\frac{\partial^2 \mathcal{L}_{iT}^m}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\bar{\boldsymbol{\xi}}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT})}_{-\mathcal{B}_{i\bar{\boldsymbol{\xi}}_i\,T}}(\widehat{\boldsymbol{\xi}}_{i\,T} - \boldsymbol{\xi}_{i\,0}), \tag{A-35}$$

where $\bar{\boldsymbol{\xi}}_i$ is between $\widehat{\boldsymbol{\xi}}_{i\,T}$ and $\boldsymbol{\xi}_{i\,0}$. By rearranging (A-35), we get

$$\sqrt{T}(\widehat{\boldsymbol{\xi}}_{i\,T} - \boldsymbol{\xi}_{i\,0}) = \sqrt{T}\mathcal{A}_{i\boldsymbol{\xi}_i\,T}\left(\mathcal{B}_{i\bar{\boldsymbol{\xi}}_i\,T}\right)^{-1}. \tag{A-36}$$

Since both terms depend on the estimated curve, we cannot apply directly the Law of Large Numbers or the Central Limit Theorem.

By Lemma 3.d we have, for any $\boldsymbol{\xi}_i \in \Xi_i$, and any $i = 1, \dots, N$,

$$\left|\left|\frac{1}{T}\frac{\partial}{\partial \boldsymbol{\xi}_i}\left(\mathcal{L}_{iT}^m(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) - \mathcal{L}_{iT}^m(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right)\right|\right|_2 = \left|\left|\frac{1}{T}\frac{\partial}{\partial \boldsymbol{\xi}_i}\left(\frac{\partial \mathcal{L}_{iT}^m}{\partial \phi}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})(\widehat{\phi}_{\boldsymbol{\xi}\,NT} - \phi_{\boldsymbol{\xi}}) + r^{(2)}(\boldsymbol{\xi}_i)\right)\right|\right|_2,$$

where $\phi_{\boldsymbol{\xi}}$ is the least favorable curve. Using the last equality and a Taylor expansion around ϕ_0, we have

$$\begin{aligned}\left|\left|\mathcal{A}_{i\,\boldsymbol{\xi}_i\,T} - \frac{1}{T}\frac{\partial \mathcal{L}_{iT}^m}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right|\right|_2 \quad \leq \quad & \left|\left|\frac{1}{T}\frac{\partial}{\partial \boldsymbol{\xi}_i}\frac{\partial \mathcal{L}_{iT}^m}{\partial \phi}(\boldsymbol{\xi}_{i\,0}, \phi_0)(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT} - \phi_0)\right|\right|_2 + \\ & + \left|\left|\frac{1}{T}\frac{\partial \mathcal{L}_{iT}^m}{\partial \phi}(\boldsymbol{\xi}_{i\,0}, \phi_0)(\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT} - \phi'_0)\right|\right|_2 + \\ & + \left|\left|\frac{1}{T}\frac{\partial r^{(2)}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0})\right|\right|_2,\end{aligned}$$

where the vectors $\widehat{\phi}'_{\boldsymbol{\xi}_o\,NT}$ and ϕ'_0 are defined in Theorem 3 and assumption L, respectively. By Lemma 3.a,b,d, we have, for any $i = 1, \dots, N$,

$$\mathcal{A}_{i\,\boldsymbol{\xi}_i\,T} \xrightarrow{P} \frac{1}{T}\frac{\partial \mathcal{L}_{iT}^m}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0), \text{ as } T \to \infty.$$

By defining

$$\mathcal{A}_{i\,\boldsymbol{\xi}_i\,T}^* = \frac{1}{T}\frac{\partial \mathcal{L}_{NT}^m}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0),$$

the Weak Law of Large Numbers and Lemma 1, imply

$$\mathcal{A}^*_{i\,\boldsymbol{\xi}_i\,T} \xrightarrow{P} \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right] = \mathbf{0}, \;\; \text{as} \;\; T \to \infty. \tag{A-37}$$

A similar reasoning holds for $\mathcal{B}_{i\bar{\boldsymbol{\xi}}_i\,T}$. First, define

$$\mathcal{B}_{i\boldsymbol{\xi}_i\,T} = -\frac{1}{T}\frac{\partial^2 \mathcal{L}^m_{iT}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i\,0}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}).$$

Then, notice that, since $\left|\left|\bar{\boldsymbol{\xi}}_i - \boldsymbol{\xi}_{i\,0}\right|\right|_2 \leq \left|\left|\widehat{\boldsymbol{\xi}}_{i\,T} - \boldsymbol{\xi}_{i\,0}\right|\right|_2 = o_P(1)$ by part a) of this Theorem, and by using a Taylor series expansion in a neighborhood of $\boldsymbol{\xi}_{i\,0}$, we have

$$\left|\left|\mathcal{B}_{i\bar{\boldsymbol{\xi}}_i\,T} - \mathcal{B}_{i\boldsymbol{\xi}_i\,T}\right|\right|_2 = o_P(1).$$

From Lemma 3.c, we have, for any $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$ and any $i = 1, \ldots, N$,

$$\sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \left|\left|\frac{1}{T}\frac{\partial^2 \mathcal{L}^m_{iT}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) - \frac{1}{T}\frac{\partial^2 \mathcal{L}^m_{iT}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})\right|\right|_2 = o_P(1), \;\; \text{as} \;\; T \to \infty,$$

which, computed in $\boldsymbol{\xi}_{i\,0}$, implies

$$\mathcal{B}_{i\,\boldsymbol{\xi}_i\,T} \xrightarrow{P} -\frac{1}{T}\frac{\partial^2 \mathcal{L}^m_{iT}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i\,0}, \phi_0), \;\; \text{as} \;\; T \to \infty.$$

By defining

$$\mathcal{B}^*_{i\,\boldsymbol{\xi}_i\,T} = -\frac{1}{T}\frac{\partial^2 \mathcal{L}^m_{iT}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i\,0}, \phi_0),$$

the Weak Law of Large Numbers and Lemma 1, imply

$$\mathcal{B}^*_{i\,\boldsymbol{\xi}_i\,T} \xrightarrow{P} -\mathrm{E}_0\left[\frac{\partial^2 \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right] = \mathcal{H}_{\boldsymbol{\xi}_{io}\boldsymbol{\xi}_{io}}, \;\; \text{as} \;\; T \to \infty. \tag{A-38}$$

By combining (A-37) and (A-38) in (A-36), we have, for any $i = 1, \ldots, N$,

$$\left(\widehat{\boldsymbol{\xi}}_{i\,T} - \boldsymbol{\xi}_{i\,0}\right) \xrightarrow{P} \mathbf{0}, \;\; \text{as} \;\; T \to \infty.$$

In order to study the asymptotic covariance matrix of the parameters of the marginals, we have to take into account the presence of the nuisance parameter, which depends on all parameters $\boldsymbol{\xi} = (\boldsymbol{\xi}_1^{\mathrm{T}} \ldots \boldsymbol{\xi}_N^{\mathrm{T}})^{\mathrm{T}}$. For this reason, by using (A-36) jointly for all marginals and Lemma 3, we can write

$$\sqrt{T}\left(\widehat{\boldsymbol{\xi}}_T - \boldsymbol{\xi}_0\right) = \sqrt{T}\begin{pmatrix} \mathcal{A}^*_{1\,\boldsymbol{\xi}_1\,T} \\ \vdots \\ \mathcal{A}^*_{N\,\boldsymbol{\xi}_N\,T} \end{pmatrix}\begin{pmatrix} \left(\mathcal{B}^*_{1\,\boldsymbol{\xi}_1\,T}\right)^{-1} & \ldots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \ldots & \left(\mathcal{B}^*_{N\,\boldsymbol{\xi}_N\,T}\right)^{-1} \end{pmatrix} = \sqrt{T}\mathcal{A}^*_{\boldsymbol{\xi}\,T}\left(\mathcal{B}^*_{\boldsymbol{\xi}\,T}\right)^{-1}.$$

By the Central Limit Theorem by Wooldridge and White (1988) and Slutsky's Theorem, we have

$$\sqrt{T}\left(\widehat{\boldsymbol{\xi}}_T - \boldsymbol{\xi}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Omega}^*\right) \;\; \text{as} \;\; T \to \infty.$$

where the asymptotic covariance matrix is

$$\mathbf{\Omega}^* = \left(\mathrm{E}_0\left[\mathcal{B}^*_{\boldsymbol{\xi}T}\right]\right)^{-1}\mathrm{Var}_0\left[\mathcal{A}^*_{\boldsymbol{\xi}T}\right]\left(\mathrm{E}_0\left[\mathcal{B}^*_{\boldsymbol{\xi}T}\right]\right)^{-1}.$$

First consider the case in which there is no correction due to the presence of a curve. Then

$$\mathrm{Var}_0\left[\mathcal{A}^*_{\boldsymbol{\xi}T}\right] = \mathrm{E}_0\left[\mathcal{A}^*_{\boldsymbol{\xi}T}\mathcal{A}^{*\mathrm{T}}_{\boldsymbol{\xi}T}\right] = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix}, \tag{A-39}$$

where we used (A-37), Lemma 2, and we have defined

$$\mathcal{I}_{\boldsymbol{\xi}_{io}\boldsymbol{\xi}_{io}} = \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i0},\phi_0)\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i^{\mathrm{T}}}(\boldsymbol{\xi}_{i0},\phi_0)\right].$$

If we now correct for the presence of the curve we have to compute the correction using the least favorable direction. Thus, by adapting (16) in the paper to the multivariate case, (A-39) becomes

$$\mathbf{I}^*_{\boldsymbol{\xi}_o} = \mathrm{E}_0\left[\mathcal{A}^*_{\boldsymbol{\xi}T}\mathcal{A}^{*\mathrm{T}}_{\boldsymbol{\xi}T}\right] = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} - \phi_0'\frac{1}{N}\mathrm{E}_0\left[\left(\sum_{i=1}^N \frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \phi}(\boldsymbol{\xi}_0,\phi_0)\right)^2\right]\phi_0'^{\mathrm{T}}.$$

Now, by using ϕ_0' as defined in (18) in the paper and in Theorem 3, we have

$$\mathbf{I}^*_{\boldsymbol{\xi}_o} = \mathrm{E}_0\left[\mathcal{A}^*_{\boldsymbol{\xi}T}\mathcal{A}^{*\mathrm{T}}_{\boldsymbol{\xi}T}\right] = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} - \bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_o}\otimes\frac{\bar{\imath}_{N\boldsymbol{\xi}_o}}{\bar{\jmath}^2_{N\boldsymbol{\xi}_o}}, \tag{A-40}$$

where $\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}$ is defined in (A-28). More precisely, using (A-28) we can define

$$\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o} = \begin{pmatrix} \bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_{1\,o}}(z_\tau) \\ \vdots \\ \bar{\boldsymbol{d}}_{N\,\boldsymbol{\xi}_{N\,o}}(z_\tau) \end{pmatrix},$$

and we have

$$\mathbf{I}^*_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} - \begin{pmatrix} \bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_{1\,o}}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_{1\,o}} & \cdots & \bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_{1\,o}}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_{N,o}} \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_{N\,o}}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_{1\,o}} & \cdots & \bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_{N\,o}}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_{N,o}} \end{pmatrix}\otimes\frac{\bar{\imath}_{N\boldsymbol{\xi}_o}}{\bar{\jmath}^2_{N\boldsymbol{\xi}_o}},$$

Analogously, when correcting for the curve we can compute:

$$\mathbf{H}^*_{\boldsymbol{\xi}_o} = \mathrm{E}_0\left[\mathcal{B}^*_{\boldsymbol{\xi}T}\right] = \begin{pmatrix} \mathcal{H}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathcal{H}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} - \left(\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}\bar{\boldsymbol{d}}^{\mathrm{T}}_{N\boldsymbol{\xi}_o}\right)\otimes\frac{1}{\bar{\jmath}_{N\boldsymbol{\xi}_o}}. \tag{A-41}$$

By combining (A-40) and (A-41) we have $\mathbf{\Omega}^* = (\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}\,\mathbf{I}^*_{\boldsymbol{\xi}_o}\,(\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}$, and

$$\sqrt{T}\left(\widehat{\boldsymbol{\xi}}_T - \boldsymbol{\xi}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, (\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}\,\mathbf{I}^*_{\boldsymbol{\xi}_o}\,(\mathbf{H}^*_{\boldsymbol{\xi}_o})^{-1}\right) \quad \text{as } T\to\infty.$$

c) Let us consider each term of (A-40) separately. We see that the sums on the first term on the right hand side is $O(1)$. Moreover, the second term on the right hand side is the product of a term $\bar{i}_{N\boldsymbol{\xi}_o}/\bar{j}^2_{N\boldsymbol{\xi}_o}$ which is bounded for any N provided assumptions C.2 and D hold, times the mixed derivatives $\bar{\boldsymbol{d}}_{N\boldsymbol{\xi}_o}$ which, by Theorem 3.b, are of order $O(N^{-1})$. Therefore, we have

$$\lim_{N\to\infty} \mathbf{I}^*_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{I}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathcal{I}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} = \mathbf{I}_{\boldsymbol{\xi}_o}.$$

Analogously, provided assumptions C.2 holds, the same argument as before applies for (A-41), and we have

$$\lim_{N\to\infty} \mathbf{H}^*_{\boldsymbol{\xi}_o} = \begin{pmatrix} \mathcal{H}_{\boldsymbol{\xi}_{1o}\boldsymbol{\xi}_{1o}} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathcal{H}_{\boldsymbol{\xi}_{No}\boldsymbol{\xi}_{No}} \end{pmatrix} = \mathbf{H}_{\boldsymbol{\xi}_o}.$$

Thus, by applying the Central Limit Theorem by Wooldridge and White (1988) and Slutsky's Theorem, we have

$$\sqrt{T}\left(\widehat{\boldsymbol{\xi}}_T - \boldsymbol{\xi}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{H}^{-1}_{\boldsymbol{\xi}_o}\,\mathbf{I}_{\boldsymbol{\xi}_o}\,\mathbf{H}^{-1}_{\boldsymbol{\xi}_o}\right), \text{ as } N, T \to \infty.$$

This completes the proof. □

Proof of Theorem 5

We proceed as in Theorem 4.a. Define

$$\mathcal{L}^c_T(\widehat{\boldsymbol{\xi}}_T, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) = \sum_{t=1}^{T} \boldsymbol{\ell}^c_t(\widehat{\boldsymbol{\xi}}_T, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}).$$

We prove consistency of $\widehat{\boldsymbol{\psi}}_T$, which is such that (see (20) in the paper),

$$\widehat{\boldsymbol{\psi}}_T = \arg\max_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} \mathcal{L}^c_T(\widehat{\boldsymbol{\xi}}_T, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}), \tag{A-42}$$

We define the function

$$\mu(\boldsymbol{\psi}) = \mathrm{E}_0\left[\boldsymbol{\ell}^c_t(\boldsymbol{\xi}_0, \boldsymbol{\psi}, \phi_{\boldsymbol{\xi}})\right].$$

From Lemma 1 and assumption H the true value of the parameters is such that

$$\boldsymbol{\psi}_0 = \arg\max_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} \mu(\boldsymbol{\psi}). \tag{A-43}$$

Then, following the same argument as in the proof of Lemma 4 and using (A-51), we can use the Weak Law of Large Numbers by McLeish (1975) and Lemma 2.1 of White and Domowitz (1984), which imply that, for any $\boldsymbol{\psi} \in \boldsymbol{\Psi}$,

$$\frac{1}{T}\mathcal{L}^c_T(\boldsymbol{\xi}_0, \boldsymbol{\psi}, \phi_{\boldsymbol{\xi}}) \xrightarrow{P} \mu(\boldsymbol{\psi}), \text{ as } T \to \infty. \tag{A-44}$$

Similarly to what proved above we can use Corollary 1 and (A-34) to prove that, for any $\boldsymbol{\psi} \in \boldsymbol{\Psi}$,

$$\frac{1}{T}\left|\mathcal{L}^c_T(\widehat{\boldsymbol{\xi}}_T, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) - \mathcal{L}^c_T(\boldsymbol{\xi}_0, \boldsymbol{\psi}, \phi_{\boldsymbol{\xi}})\right| \xrightarrow{P} 0, \text{ as } T \to \infty. \tag{A-45}$$

Therefore, by combining (A-44) and (A-45), we have, for any $\boldsymbol{\psi} \in \boldsymbol{\Psi}$,

$$\frac{1}{T}\mathcal{L}_T^c(\widehat{\boldsymbol{\xi}}_T, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}) \xrightarrow{P} \mu(\boldsymbol{\psi}), \text{ as } T \to \infty. \tag{A-46}$$

Furthermore, by means of (A-42) and (A-43), we get

$$\frac{1}{T}\mathcal{L}_T^c(\widehat{\boldsymbol{\xi}}_T, \widehat{\boldsymbol{\psi}}_T, \widehat{\phi}_{\widehat{\boldsymbol{\xi}}_T\,NT}) \xrightarrow{P} \mu(\boldsymbol{\psi}_0), \text{ as } T \to \infty.$$

From (A-46), the term on the left hand side is such that

$$\frac{1}{T}\mathcal{L}_T^c(\widehat{\boldsymbol{\xi}}_T, \widehat{\boldsymbol{\psi}}_T, \widehat{\phi}_{\widehat{\boldsymbol{\xi}}_T\,NT}) \xrightarrow{P} \mu(\widehat{\boldsymbol{\psi}}_T), \text{ as } T \to \infty,$$

thus

$$\mu(\widehat{\boldsymbol{\psi}}_T) \xrightarrow{P} \mu(\boldsymbol{\psi}_0), \text{ as } T \to \infty.$$

Given assumptions I and S, we have consistency

$$\widehat{\boldsymbol{\psi}}_T \xrightarrow{P} \boldsymbol{\psi}_0, \text{ as } T \to \infty.$$

This completes the proof. □

Additional results

Notice that, although not explicitly indicated in the proofs, all density functions f_{x_i} and c are to be considered as conditional on $\mathcal{F}_{t-1}$.

Lemma 1 *Under assumptions C.1 and S for $i = 1, \dots, N$ and $z_\tau \in [0,1]$*

$$\mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}_{i\,t}^m}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right] = \mathbf{0}, \quad \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}_t^c}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\right] = \mathbf{0}, \quad \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}_t^c}{\partial \boldsymbol{\psi}}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\right] = \mathbf{0},$$

and

$$\mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}_{i\,t}^m}{\partial \phi}(\boldsymbol{\xi}_{i\,0}, \phi_0(z_\tau))\right] = 0, \quad \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}_t^c}{\partial \phi}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0(z_\tau))\right] = 0.$$

Proof. We prove just the first relation, the proof of the others being analogous:

$$\begin{aligned}
&\mathrm{E}_0\left[\frac{\partial}{\partial \boldsymbol{\xi}_i}\boldsymbol{\ell}_{i\,t}^m(\boldsymbol{\xi}_{i\,0}, \phi_0)\right] = \\
= \;& \int_{\mathbf{x}} \left(\frac{\partial}{\partial \boldsymbol{\xi}_i}\log f_{x_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right)\left(\prod_{k=1}^N f_{x_k}(\boldsymbol{\xi}_{k\,0}, \phi_0)\right) c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\mathrm{d}\mathbf{x}_t = \\
= \;& \int_{\mathbf{x}} \left(\frac{\partial}{\partial \boldsymbol{\xi}_i} f_{x_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\right)\frac{\left(\prod_{k=1}^N f_{x_k}(\boldsymbol{\xi}_{k\,0}, \phi_0)\right)}{f_{x_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)} c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\mathrm{d}\mathbf{x}_t = \\
= \;& \frac{\partial}{\partial \boldsymbol{\xi}_i}\left(\int_{\mathbf{x}}\left(\prod_{k=1}^N f_{x_k}(\boldsymbol{\xi}_{k\,0}, \phi_0)\right) c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\mathrm{d}\mathbf{x}_t\right) = \frac{\partial}{\partial \boldsymbol{\xi}_i} 1 = 0. \;\square
\end{aligned}$$

Lemma 2 *Under assumptions C.1, S and P, for any $i, j = 1, \dots, N$*

$$\mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\frac{\partial \boldsymbol{\ell}^c_t}{\partial \boldsymbol{\psi}^{\scriptscriptstyle T}}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)\right] = \mathbf{0}, \quad \textit{and} \quad \mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0}, \phi_0)\frac{\partial \boldsymbol{\ell}^m_{jt}}{\partial \boldsymbol{\xi}^{\scriptscriptstyle T}_j}(\boldsymbol{\xi}_{j\,0}, \phi_0)\right] = \mathbf{0}.$$

Proof. Equation (**??**) is in the appendix in Joe (2005). The proof of (**??**) is similar. Let $\mathbf{x}_{-i\,t}$ be the vector $\mathbf{x}_t$ when omitting the i-th component. We omit the dependence on ϕ_0 for simplicity. Then the expectation in (**??**) is equivalent to

$$\begin{aligned}
&\mathrm{E}_0\left[\frac{\partial \boldsymbol{\ell}^m_{it}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i\,0})\frac{\partial \boldsymbol{\ell}^m_{jt}}{\partial \boldsymbol{\xi}^{\mathrm{T}}_j}(\boldsymbol{\xi}_{j\,0})\right] = \\
= \ & \int_{\mathbf{x}}\left(\frac{\partial}{\partial \boldsymbol{\xi}_i}\log f_{x_i}(\boldsymbol{\xi}_{i\,0})\right)\left(\frac{\partial}{\partial \boldsymbol{\xi}^{\mathrm{T}}_j}\log f_{x_j}(\boldsymbol{\xi}_{j\,0})\right)\left(\prod_{k=1}^{N} f_{x_k}(\boldsymbol{\xi}_{k\,0})\right)c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0)\mathrm{d}\mathbf{x}_t = \\
= \ & \int_{x_i}\left(\frac{\partial}{\partial \boldsymbol{\xi}_i}\log f_{x_i}(\boldsymbol{\xi}_{i\,0})\right)\left[\int_{\mathbf{x}_{-i}}\left(\frac{\partial}{\partial \boldsymbol{\xi}^{\mathrm{T}}_j}\log f_{x_j}(\boldsymbol{\xi}_{j\,0})\right)\left(\prod_{k=1}^{N} f_{x_k}(\boldsymbol{\xi}_{k\,0})\right)c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0)\mathrm{d}\mathbf{x}_{-i\,t}\right]\mathrm{d}x_{i\,t} = \\
= \ & \int_{x_i}\left(\frac{\partial}{\partial \boldsymbol{\xi}_i}\log f_{x_i}(\boldsymbol{\xi}_{i\,0})\right)\left[\int_{\mathbf{x}_{-i}}\left(\frac{\partial}{\partial \boldsymbol{\xi}^{\mathrm{T}}_j}f_{x_j}(\boldsymbol{\xi}_{j\,0})\right)\left(\prod_{k\neq j;k=1}^{N} f_{x_k}(\boldsymbol{\xi}_{k\,0})\right)c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0)\mathrm{d}\mathbf{x}_{-i\,t}\right]\mathrm{d}x_{i\,t} = \\
= \ & \int_{x_i}\left(\frac{\partial}{\partial \boldsymbol{\xi}_i}\log f_{x_i}(\boldsymbol{\xi}_{i\,0})\right)\frac{\partial}{\partial \boldsymbol{\xi}^{\mathrm{T}}_j}\left[\int_{\mathbf{x}_{-i}}\left(\prod_{k=1}^{N} f_{x_k}(\boldsymbol{\xi}_{k\,0})\right)c(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0)\mathrm{d}\mathbf{x}_{-i\,t}\right]\mathrm{d}x_{i\,t} = \\
= \ & \int_{x_i}\left(\frac{\partial}{\partial \boldsymbol{\xi}_i}\log f_{x_i}(\boldsymbol{\xi}_{i\,0})\right)\frac{\partial}{\partial \boldsymbol{\xi}^{\mathrm{T}}_j}f_{x_i}(\boldsymbol{\xi}_{i\,0})\mathrm{d}x_{i\,t} = \mathbf{0}. \ \square
\end{aligned}$$

Lemma 3 *Under assumptions C.1, N and S*

a) for any $z_t \in [0,1]$ and for $i = 1, ..., N$, as $T \to \infty$,

$$\left\|\frac{1}{\sqrt{T}}\frac{\partial}{\partial \boldsymbol{\xi}_i}\frac{\partial \sum_{t=1}^{T}\boldsymbol{\ell}^m_{it}}{\partial \phi}(\boldsymbol{\xi}_{i\,0}, \phi_0)(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t) - \phi_0(z_t))\right\|_2 = o_P(1),$$
$$\left\|\frac{1}{\sqrt{T}}\frac{\partial}{\partial \boldsymbol{\psi}}\frac{\partial \sum_{t=1}^{T}\boldsymbol{\ell}^c_{t}}{\partial \phi}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)(\widehat{\phi}_{\boldsymbol{\xi}_o\,NT}(z_t) - \phi_0(z_t))\right\|_2 = o_P(1).$$

b) for any $z_t \in [0,1]$ and for $i = 1, ..., N$, as $T \to \infty$,

$$\left\|\frac{1}{\sqrt{T}}\frac{\partial \sum_{t=1}^{T}\boldsymbol{\ell}^m_{it}}{\partial \phi}(\boldsymbol{\xi}_{i\,0}, \phi_0)(\widehat{\phi'}_{\boldsymbol{\xi}_o\,NT}(z_t) - \phi'_0(z_t))\right\|_2 = o_P(1),$$
$$\left\|\frac{1}{\sqrt{T}}\frac{\partial \sum_{t=1}^{T}\boldsymbol{\ell}^c_{t}}{\partial \phi}(\boldsymbol{\xi}_0, \boldsymbol{\psi}_0, \phi_0)(\widehat{\phi'}_{\boldsymbol{\xi}_o\,NT}(z_t) - \phi'_0(z_t))\right\|_2 = o_P(1).$$

c) for $i = 1, ..., N$ and for any $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$ and $\boldsymbol{\psi} \in \boldsymbol{\Psi}$,

$$\sum_{t=1}^{T} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi} T}) - \sum_{t=1}^{T} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}}) = r^{(1)}(\boldsymbol{\xi}_i),$$
$$\sum_{t=1}^{T} \boldsymbol{\ell}_{t}^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi} T}) - \sum_{t=1}^{T} \boldsymbol{\ell}_{t}^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi_{\boldsymbol{\xi}}) = r^{(3)}(\boldsymbol{\psi}),$$

such that, as $T \to \infty$

$$\sup_{\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i} \left|\left| \frac{1}{T} \frac{\partial^2 r^{(1)}}{\partial \boldsymbol{\xi}_i \partial \boldsymbol{\xi}_i^{T}}(\boldsymbol{\xi}_i) \right|\right|_2 = o_P(1), \qquad \sup_{\boldsymbol{\psi} \in \boldsymbol{\Psi}} \left|\left| \frac{1}{T} \frac{\partial^2 r^{(3)}}{\partial \boldsymbol{\psi} \partial \boldsymbol{\psi}^{T}}(\boldsymbol{\psi}) \right|\right|_2 = o_P(1).$$

d) for $i = 1, \dots, N$ and for any $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$ and $\boldsymbol{\psi} \in \boldsymbol{\Psi}$

$$\begin{aligned}
\sum_{t=1}^{T} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \widehat{\phi}_{\boldsymbol{\xi} T}) &= \sum_{t=1}^{T} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}}) + \frac{\partial}{\partial \phi} \sum_{t=1}^{T} \boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi_{\boldsymbol{\xi}})(\widehat{\phi}_{\boldsymbol{\xi}}(z_t) - \phi_{\boldsymbol{\xi}}(z_t)) + r^{(2)}(\boldsymbol{\xi}_i), \\
\sum_{t=1}^{T} \boldsymbol{\ell}_{t}^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \widehat{\phi}_{\boldsymbol{\xi} T}) &= \sum_{t=1}^{T} \boldsymbol{\ell}_{t}^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi_{\boldsymbol{\xi}}) + \frac{\partial}{\partial \phi} \sum_{t=1}^{T} \boldsymbol{\ell}_{t}^{c}(\boldsymbol{\xi}, \boldsymbol{\psi}, \phi_{\boldsymbol{\xi}})(\widehat{\phi}_{\boldsymbol{\xi}}(z_t) - \phi_{\boldsymbol{\xi}}(z_t)) + r^{(4)}(\boldsymbol{\psi}),
\end{aligned}$$

such that, as $T \to \infty$

$$\left|\left| \frac{1}{\sqrt{T}} \frac{\partial r^{(2)}}{\partial \boldsymbol{\xi}_i}(\boldsymbol{\xi}_{i0}) \right|\right|_2 = o_P(1), \quad \textit{and} \quad \left|\left| \frac{1}{\sqrt{T}} \frac{\partial r^{(4)}}{\partial \boldsymbol{\psi}}(\boldsymbol{\psi}_0) \right|\right|_2 = o_P(1).$$

Proof. The proof is in Lemma 2 and 3 by Severini and Wong (1992) and we use the Central Limit Theorem 2.11 in Wooldridge and White (1988).

Lemma 4 *Under assumptions C.1 and K, assumption N is satisfied with $\widetilde{\phi}_{\boldsymbol{\xi}} = \phi_{\boldsymbol{\xi}}$, i.e., as $T \to \infty$, $Th_{NT} \to \infty$, $h_{NT} \to 0$,*

$$\begin{aligned}
\sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} ||\widehat{\phi}_{\boldsymbol{\xi} NT} - \phi_{\boldsymbol{\xi}}|| = \sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} \sup_{z_\tau \in [0,1]} |\widehat{\phi}_{\boldsymbol{\xi} NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)| &= o_P(1), \\
\sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} ||\widehat{\phi}'_{\boldsymbol{\xi} NT} - \phi'_{\boldsymbol{\xi}}|| = \sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} \sup_{z_\tau \in [0,1]} |\widehat{\phi}'_{\boldsymbol{\xi} NT}(z_\tau) - \phi'_{\boldsymbol{\xi}}(z_\tau)| &= o_P(1), \\
\sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} ||\widehat{\phi}''_{\boldsymbol{\xi} NT} - \phi''_{\boldsymbol{\xi}}|| = \sup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} \sup_{z_\tau \in [0,1]} |\widehat{\phi}''_{\boldsymbol{\xi} NT}(z_\tau) - \phi''_{\boldsymbol{\xi}}(z_\tau)| &= o_P(1).
\end{aligned}$$

If also assumption C.2 holds then the same results are valid also when $N \to \infty$.

Proof. Consider $\bar{\lambda}_N(\boldsymbol{\xi}, \phi(z_\tau))$ and $\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \phi(z_\tau))$ defined respectively in (A-4) and (A-1). Let

$R(z_t) = \phi_0(z_t)/\phi(z_t)$. We then have

$$
\begin{aligned}
\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \phi(z_\tau)) &= \sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}\left(\frac{z_\tau - z_t}{h_{NT}}\right) \boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_0(z_t)R(z_t)) = \\
&= \sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}(u_t)\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_0(z_t)(R(z_\tau) + R'(z_\tau)h_{NT}u_t)) = \\
&= \sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}(u_t)\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_0(z_t)R(z_\tau)) + \\
&\quad + \sum_{t=1}^{T}\sum_{i=1}^{N} \mathrm{K}(u_t)R'(z_\tau)h_{NT}u_t \frac{\partial}{\partial R}\boldsymbol{\ell}^m_{i\,t}(\boldsymbol{\xi}_i, \phi_0(z_t)R(z_\tau)) = \\
&= A_{NT} + B_{NT}. \qquad \text{(A-47)}
\end{aligned}
$$

Notice that A_{NT} can be written as $A_{NT} = \widetilde{A}_{NT} + O_P(T^{-1})$, where $\widetilde{A}_{NT}$ is a sum of mixing random variables. To see this consider for simplicity the case $\gamma_i = 0$ and notice that

$$
\mu_{i\,t}(\boldsymbol{\delta}_i) = \omega_i + \alpha_i \frac{x_{i\,t-1}}{a_i\phi(z_{t-1})} + \beta_i\mu_{i\,t-1} = \frac{\omega_i}{1-\beta_i} + \sum_{k=1}^{t} \alpha_i\beta_i^{k-1}\frac{x_{i\,t-k}}{a_i\phi(z_{t-k})} \qquad \text{(A-48)}
$$

where $\omega_i = 1 - \alpha_i - \beta_i$. We then have

$$
\begin{aligned}
\mu_{i\,t}(\boldsymbol{\delta}_i) &= \frac{\omega_i}{1-\beta_i} + \sum_{k=1}^{t}\alpha_i\beta_i^{k-1}\frac{x_{i\,t-k}}{a_i\phi_0(z_{t-k})}R(z_{t-k}) = \\
&= \left[\frac{\omega_i}{1-\beta_i} + \sum_{k=1}^{t}\alpha_i\beta_i^{k-1}\frac{x_{i\,t-k}}{a_i\phi_0(z_{t-k})}R(z_\tau)\right] + \\
&\quad + \sum_{k=1}^{t}\alpha_i\beta_i^{k-1}\frac{x_{i\,t-k}}{a_i\phi_0(z_{t-k})}R'(z_{t-\bar{k}})(z_{t-k} - z_\tau) = \chi_{1i\,t} + \chi_{2i\,t}, \qquad \text{(A-49)}
\end{aligned}
$$

where $z_{t-k} < z_{t-\bar{k}} < z_t$, i.e. $0 \le \bar{k} \le k$. Then, since $z_t = t/T$ we have that $\mathrm{E}_0[\chi_{2i\,t}] = O(T^{-1})$ and

$$
\begin{aligned}
\mathrm{E}_0[\chi^2_{2i\,t}] &= \frac{1}{T^2}\mathrm{E}_0\left[\left(\sum_{k=1}^{t}\alpha_i\beta_i^{k-1}\frac{x_{i\,t-k}}{a_i\phi_0(z_{t-k})}R'(z_{t-\bar{k}})\,(t-k-\tau)\right)^2\right] \le \\
&\le \frac{1}{T}\mathrm{E}_0\left[\frac{1}{T}\left\{\sum_{k=1}^{t}\left(\frac{x_{i\,t-k}}{a_i\phi_0(z_{t-k})}\right)^2\right\}\left\{\sum_{k=1}^{t}(\alpha_i\beta_i^{k-1}R'(z_{t-\bar{k}})\,(t-k-\tau))^2\right\}\right].
\end{aligned}
$$

Thus $\mathrm{E}_0[\chi^2_{2i\,t}] = O(T^{-1})$ implying that $\mathrm{Var}_0[\chi_{2i\,t}] = O(T^{-1})$ and, finally, $\chi_{2i\,t} = O_P(T^{-1})$.

By using a Taylor approximation of A_{NT} we obtain $A_{NT} = \widetilde{A}_{NT} + O_P(T^{-1})$. Where each term of the sum in $\widetilde{A}_{NT}$ is the log–likelihood with the generic trend $\phi(z_t)$ replaced by $\phi_0(z_t)R(z_\tau)$ and $\mu_{i\,t}$ replaced by $\chi_{1i\,t}$. With this substitution $\widetilde{A}_{NT}$ is just function of the process

$$
\widetilde{x}_{i\,t} = \frac{x_{i\,t}}{a_i\phi_0(z_t)} = \frac{a_{i\,0}\epsilon_{i\,t}\mu_{i\,t}(\boldsymbol{\delta}_{i0})}{a_i},
$$

which is mixing by assumption A.1, and of $R(z_\tau)$ which for a fixed z_τ can be treated as a constant. By using the Weak Law of Large Numbers by McLeish (1975) and Lemma 2.1 of White and Domowitz (1984) for $\widetilde{A}_{NT}$, we have, as $T \to \infty$ and $Th_{NT} \to \infty$,

$$\frac{\widetilde{A}_{NT} + O_p(T^{-1})}{NTh_{NT}} \xrightarrow{P} \mathrm{E}_0\left[\frac{A_{NT}}{NTh_{NT}}\right]. \tag{A-50}$$

Now let us compute

$$\begin{aligned}\mathrm{E}_0\left[\frac{A_{NT}}{NTh_{NT}}\right] &= \frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}(u_t)\mathrm{E}_0\left[\sum_{i=1}^{N}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi_0(z_\tau)R(z_\tau))\right] + \\ &+\frac{1}{NTh_{NT}}\sum_{t=1}^{T}\mathrm{K}(u_t)h_{NT}u_t\phi_0^{'}(z_\tau)\mathrm{E}_0\left[\sum_{i=1}^{N}\frac{\partial}{\partial\phi}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi_0(z_\tau)R(z_\tau))\right].\end{aligned}$$

Notice that $\phi_0(z_\tau)R(z_\tau) = \phi(z_\tau)$. When $T \to \infty$ and $Th_{NT} \to \infty$ and using expressions analogous to (A-8) in the proof of Theorem 2 we have

$$\mathrm{E}_0\left[\frac{A_{NT}}{NTh_{NT}}\right] \to \frac{1}{N}\mathrm{E}_0\left[\sum_{i=1}^{N}\boldsymbol{\ell}_{it}^{m}(\boldsymbol{\xi}_i, \phi(z_\tau))\right] = \bar{\lambda}_N(\boldsymbol{\xi}, \phi(z_\tau)).$$

Then, it is possible to show that $\mathrm{Var}_0[B_{NT}/NTh_{NT}] = O(h_{NT}^2)$. Therefore, as $h_{NT} \to 0$ and by assumption C.1, $B_{NT}/NTh_{NT} \xrightarrow{P} \mathrm{E}_0[B_{NT}/NTh_{NT}]$. But, as $T \to \infty$ and $Th_{NT} \to \infty$ and using expressions analogous to (A-8) in the proof of Theorem 2 we have

$$\mathrm{E}_0\left[\frac{B_{NT}}{NTh_{NT}}\right] \to 0.$$

Therefore, we have, for any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, any $z_\tau \in [0, 1]$, and any $\phi \in \mathcal{P}$,

$$\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \phi(z_\tau))}{NTh_{NT}} \xrightarrow{P} \bar{\lambda}_N(\boldsymbol{\xi}, \phi(z_\tau)), \text{ as } T \to \infty,\ Th_{NT} \to \infty,\ h_{NT} \to 0.$$

Following the same reasoning as in the proof of Lemma 8 by Severini and Wong (1992), we can also prove that, as $T \to \infty$, $Th_{NT} \to \infty$, and $h_{NT} \to 0$,

$$\sup_{\boldsymbol{\xi}_i\in\boldsymbol{\Xi}_i}\ \sup_{\phi\in\Gamma}\ \sup_{z_\tau\in[0,1]}\left|\frac{\partial^k}{\partial\boldsymbol{\xi}_i^k}\frac{\partial^l}{\partial z_t^l}\frac{\partial^j}{\partial\phi^j}\left(\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \phi(z_\tau))}{NTh_{NT}} - \bar{\lambda}_N(\boldsymbol{\xi}, \phi(z_\tau))\right)\right| = o_P(1), \tag{A-51}$$

for $k, l, j = 0, 1, 2$.

For any $\varepsilon > 0$, there exists $\delta > 0$ such that, as $T \to \infty$, $Th_{NT} \to \infty$, and $h_{NT} \to 0$,

$$\begin{aligned}&P\left\{\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}}\ \sup_{z_\tau\in[0,1]}\left|\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)\right| > \varepsilon\right\} \le P\left\{\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}}\ \sup_{z_\tau\in[0,1]}\left|\frac{\partial\bar{\lambda}_N(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial\phi}\right| > \delta\right\} = \\ &= P\left\{\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}}\ \sup_{z_\tau\in[0,1]}\left|\frac{1}{NTh_{NT}}\frac{\partial\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial\phi} - \frac{\partial\bar{\lambda}_N(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial\phi}\right| > \delta\right\} \le \\ &\le P\left\{\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}}\ \sup_{z_\tau\in[0,1]}\ \sup_{\phi\in\Gamma}\left|\frac{1}{NTh_{NT}}\frac{\partial\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \phi(z_\tau))}{\partial\phi} - \frac{\partial\bar{\lambda}_N(\boldsymbol{\xi}, \phi(z_\tau))}{\partial\phi}\right| > \delta\right\} \to 0,\end{aligned}$$

where we used assumption I.2, and (A-51). Hence,

$$\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \sup_{z_\tau\in[0,1]} |\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)| = o_P(1),$$

which implies that, for any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ and any $z_\tau \in [0,1]$, $\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) \xrightarrow{P} \phi_{\boldsymbol{\xi}}(z_\tau)$ as $T \to \infty$, $Th_{NT} \to \infty$ and $h_{NT} \to 0$. Then, we prove (A-47). From a Taylor approximation of the first order conditions, we have

$$\begin{aligned}
0 &= \frac{1}{NTh_{NT}} \frac{\partial \widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial \phi} - \frac{\partial \bar{\lambda}_N(\boldsymbol{\xi}, \phi_{\boldsymbol{\xi}}(z_\tau))}{\partial \phi} = \\
&= \frac{1}{NTh_{NT}} \frac{\partial \widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial \phi} - \frac{\partial \bar{\lambda}_N(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial \phi} + \frac{\partial \bar{\lambda}_N(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial \phi} - \frac{\partial \bar{\lambda}_N(\boldsymbol{\xi}, \phi_{\boldsymbol{\xi}}(z_\tau))}{\partial \phi} = \\
&= R_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau)) + \Delta_N(\boldsymbol{\xi}, \bar{\phi}(z_\tau)) \left(\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)\right), \qquad \text{(A-52)}
\end{aligned}$$

where $\bar{\phi}(z_\tau)$ lies between $\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau)$ and $\phi_{\boldsymbol{\xi}}(z_\tau)$, and

$$R_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau)) = \frac{1}{NTh_{NT}} \frac{\partial \widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial \phi} - \frac{\partial \bar{\lambda}_N(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau))}{\partial \phi},$$

and

$$\Delta_N(\boldsymbol{\xi}, \bar{\phi}(z_\tau)) = \frac{\partial^2 \bar{\lambda}_N(\boldsymbol{\xi}, \bar{\phi}(z_\tau))}{\partial \phi^2}.$$

By differentiating (A-52) with respect to $\boldsymbol{\xi}$, we have

$$\begin{aligned}
\mathbf{0} &= \frac{\partial}{\partial \boldsymbol{\xi}} R_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau)) + \frac{\partial}{\partial \boldsymbol{\xi}} \Delta_N(\boldsymbol{\xi}, \bar{\phi}(z_\tau)) \left(\widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi_{\boldsymbol{\xi}}(z_\tau)\right) + \\
&\quad + \Delta_N(\boldsymbol{\xi}, \bar{\phi}(z_\tau)) \left(\widehat{\phi}'_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi'_{\boldsymbol{\xi}}(z_\tau)\right). \qquad \text{(A-53)}
\end{aligned}$$

From assumption C.1 and from (A-51) we have, respectively, that

$$\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \sup_{z_\tau\in[0,1]} \left| \frac{\partial}{\partial \boldsymbol{\xi}} \Delta_N(\boldsymbol{\xi}, \bar{\phi}(z_\tau)) \right| = O_P(1), \quad \sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \sup_{z_\tau\in[0,1]} \left| \frac{\partial}{\partial \boldsymbol{\xi}} R_{NT}(\boldsymbol{\xi}, \widehat{\phi}_{\boldsymbol{\xi}\,NT}(z_\tau)) \right| = o_P(1).$$

Therefore, from (A-53), we have

$$\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \sup_{z_\tau\in[0,1]} |\widehat{\phi}'_{\boldsymbol{\xi}\,NT}(z_\tau) - \phi'_{\boldsymbol{\xi}}(z_\tau)| = o_P(1),$$

which implies that, for any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ and any $z_\tau \in [0,1]$, $\widehat{\phi}'_{\boldsymbol{\xi}\,NT}(z_\tau) \xrightarrow{P} \phi'_{\boldsymbol{\xi}}(z_\tau)$, as $T \to \infty$, $Th_{NT} \to \infty$ and $h_{NT} \to 0$.

Finally, as $N, T \to \infty$, $Th_{NT} \to \infty$, $h_{NT} \to 0$, (A-51) becomes

$$\sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}_i} \sup_{\phi\in\Gamma} \sup_{z_t\in[0,1]} \left| \frac{\partial^k}{\partial \boldsymbol{\xi}_i^k} \frac{\partial^l}{\partial z_t^l} \frac{\partial^j}{\partial \phi^j} \left(\frac{\widetilde{\mathcal{L}}_{NT}(\boldsymbol{\xi}, \phi(z_\tau))}{NTh_{NT}} - \lambda(\boldsymbol{\xi}, \phi(z_\tau)) \right) \right| = o_P(1),$$

where $\lambda(\boldsymbol{\xi}, \phi(z_\tau)) = \lim_{N\to\infty} \bar{\lambda}_N(\boldsymbol{\xi}, \phi(z_\tau))$ which exists by assumption C.2. Using this assumption the lemma can be proved as well for $N \to \infty$. $\square$

Appendix B - Data Description and Detailed Estimation Results

Table 10: S&P100 constituents

Ticker	Name	Sector
AA	Alcoa Inc	Materials
AAPL	Apple Inc.	Information Technology
ABT	Abbott Labs	Health Care
AEP	American Electric Power	Utilities
ALL	Allstate Corp.	Financials
AMGN	Amgen	Health Care
AMZN	Amazon Corp.	Consumer Discretionary
AVP	Avon Products	Consumer Staples
AXP	American Express	Financials
BA	Boeing Company	Industrials
BAC	Bank of America Corp.	Financials
BAX	Baxter International Inc.	Health Care
BHI	Baker Hughes	Energy
BK	Bank of New York Mellon Corp.	Financials
BMY	Bristol-Myers Squibb	Health Care
BNI	Burlington Northern Santa Fe C	Industrials
CAT	Caterpillar Inc.	Industrials
C	Citigroup Inc.	Financials
CL	Colgate-Palmolive	Consumer Staples
CMCSA	Comcast Corp.	Consumer Discretionary
COF	Capital One Financial	Financials
COST	Costco Co.	Consumer Staples
CPB	Campbell Soup	Consumer Staples
CSCO	Cisco Systems	Information Technology
CVS	CVS Caremark Corp.	Consumer Staples
CVX	Chevron Corp.	Energy
DD	Du Pont (E.I.)	Materials
DELL	Dell Inc.	Information Technology
DIS	Walt Disney Co.	Consumer Discretionary
DOW	Dow Chemical	Materials
DVN	Devon Energy Corp.	Energy
EMC	EMC Corp.	Information Technology
ETR	Entergy Corp.	Utilities
EXC	Exelon Corp.	Utilities
FDX	FedEx Corporation	Industrials
F	Ford Motor	Consumer Discretionary
GD	General Dynamics	Industrials
GE	General Electric	Industrials
GILD	Gilead Sciences	Health Care
GS	Goldman Sachs Group	Financials
HAL	Halliburton Co.	Energy
HD	Home Depot	Consumer Discretionary
HNZ	Heinz (H.J.)	Consumer Staples
HON	Honeywell Int'l Inc.	Industrials
HPQ	Hewlett-Packard	Information Technology
(cont.)		

(cont.)

IBM	International Bus. Machines	Information Technology
INTC	Intel Corp.	Information Technology
JNJ	Johnson & Johnson	Health Care
JPM	JPMorgan Chase & Co.	Financials
KO	Coca Cola Co.	Consumer Staples
LMT	Lockheed Martin Corp.	Industrials
LOW	Lowe's Cos.	Consumer Discretionary
MCD	McDonald's Corp.	Consumer Discretionary
MDT	Medtronic Inc.	Health Care
MMM	3M Company	Industrials
MO	Altria Group, Inc.	Consumer Staples
MRK	Merck & Co.	Health Care
MSFT	Microsoft Corp.	Information Technology
MS	Morgan Stanley	Financials
NKE	NIKE Inc.	Consumer Discretionary
NSC	Norfolk Southern Corp.	Industrials
ORCL	Oracle Corp.	Information Technology
OXY	Occidental Petroleum	Energy
PEP	PepsiCo Inc.	Consumer Staples
PFE	Pfizer, Inc.	Health Care
PG	Procter & Gamble	Consumer Staples
QCOM	QUALCOMM Inc.	Information Technology
RF	Regions Financial Corp.	Financials
SGP	Schering-Plough	Health Care
SLB	Schlumberger Ltd.	Energy
SLE	Sara Lee Corp.	Consumer Staples
SO	Southern Co.	Utilities
S	Sprint Nextel Corp.	Telecommunications Services
T	AT&T Inc.	Telecommunications Services
TGT	Target Corp.	Consumer Discretionary
TWX	Time Warner Inc.	Consumer Discretionary
TXN	Texas Instruments	Information Technology
TYC	Tyco International	Industrials
UNH	UnitedHealth Group Inc.	Health Care
UPS	United Parcel Service	Industrials
USB	U.S. Bancorp	Financials
UTX	United Technologies	Industrials
VZ	Verizon Communications	Telecommunications Services
WAG	Walgreen Co.	Consumer Staples
WFC	Wells Fargo	Financials
WMB	Williams Cos.	Energy
WMT	Wal-Mart Stores	Consumer Staples
WY	Weyerhaeuser Corp.	Materials
XOM	Exxon Mobil Corp.	Energy
XRX	Xerox Corp.	Information Technology

Table 11: S&P100 Parameter Estimates

	SPvMEM						Univariate MEM					
	a_i	α_i	γ_i	β_i	ν_i	π_i	a_i	α_i	γ_i	β_i	ν_i	π_i
AA	0.35 (0.041)	0.28 (0.020)	0.08 (0.013)	0.63 (0.019)	0.27 (0.000)	0.94 (0.029)	0.16 (0.036)	0.28 (0.034)	0.08 (0.021)	0.65 (0.029)	0.28 (0.027)	0.97 (0.045)
AAPL	0.55 (0.037)	0.34 (0.018)	0.12 (0.012)	0.53 (0.018)	0.48 (0.000)	0.93 (0.027)	0.34 (0.080)	0.34 (0.054)	0.12 (0.035)	0.55 (0.051)	0.51 (0.037)	0.95 (0.076)
ABT	0.10 (0.013)	0.28 (0.016)	0.09 (0.011)	0.65 (0.013)	0.34 (0.000)	0.98 (0.021)	0.07 (0.020)	0.29 (0.037)	0.09 (0.027)	0.65 (0.030)	0.34 (0.030)	0.98 (0.049)
AEP	0.10 (0.012)	0.26 (0.013)	0.09 (0.010)	0.67 (0.012)	0.38 (0.000)	0.98 (0.018)	0.07 (0.019)	0.26 (0.029)	0.10 (0.023)	0.67 (0.027)	0.38 (0.026)	0.99 (0.041)
ALL	0.09 (0.011)	0.29 (0.015)	0.10 (0.012)	0.65 (0.014)	0.34 (0.000)	0.98 (0.021)	0.05 (0.016)	0.29 (0.034)	0.10 (0.028)	0.65 (0.030)	0.34 (0.028)	0.99 (0.048)
AMGN	0.17 (0.025)	0.38 (0.024)	0.07 (0.018)	0.54 (0.025)	0.28 (0.000)	0.95 (0.036)	0.13 (0.025)	0.38 (0.034)	0.08 (0.025)	0.53 (0.032)	0.29 (0.023)	0.96 (0.049)
AMZN	0.43 (0.043)	0.31 (0.016)	0.10 (0.014)	0.59 (0.016)	0.32 (0.000)	0.95 (0.023)	0.17 (0.053)	0.31 (0.036)	0.10 (0.032)	0.62 (0.030)	0.35 (0.034)	0.98 (0.049)
AVP	0.13 (0.014)	0.41 (0.013)	0.06 (0.009)	0.55 (0.013)	0.67 (0.000)	0.98 (0.019)	0.30 (0.080)	0.40 (0.052)	0.08 (0.043)	0.48 (0.061)	0.64 (0.041)	0.92 (0.083)
AXP	0.10 (0.014)	0.35 (0.016)	0.12 (0.013)	0.57 (0.016)	0.30 (0.000)	0.98 (0.024)	0.04 (0.013)	0.36 (0.033)	0.11 (0.025)	0.58 (0.030)	0.30 (0.027)	1.00 (0.047)
BA	0.11 (0.018)	0.24 (0.020)	0.11 (0.014)	0.68 (0.018)	0.24 (0.000)	0.98 (0.028)	0.06 (0.017)	0.24 (0.030)	0.11 (0.021)	0.69 (0.026)	0.25 (0.024)	0.99 (0.041)
BAC	0.12 (0.009)	0.37 (0.020)	0.14 (0.013)	0.52 (0.017)	0.42 (0.000)	0.97 (0.027)	0.05 (0.014)	0.37 (0.058)	0.14 (0.038)	0.54 (0.048)	0.44 (0.043)	0.99 (0.077)
BAX	0.07 (0.008)	0.31 (0.011)	0.09 (0.009)	0.64 (0.009)	0.65 (0.000)	1.00 (0.015)	0.08 (0.035)	0.31 (0.063)	0.09 (0.047)	0.63 (0.049)	0.64 (0.048)	0.99 (0.083)
BHI	0.20 (0.045)	0.27 (0.023)	0.06 (0.014)	0.67 (0.022)	0.20 (0.000)	0.98 (0.033)	0.14 (0.036)	0.26 (0.025)	0.07 (0.015)	0.68 (0.024)	0.21 (0.022)	0.98 (0.036)
BK	0.12 (0.014)	0.32 (0.009)	0.07 (0.012)	0.63 (0.006)	0.59 (0.000)	0.99 (0.013)	0.08 (0.049)	0.32 (0.051)	0.08 (0.066)	0.63 (0.033)	0.58 (0.057)	1.00 (0.069)
BMY	0.16 (0.018)	0.26 (0.011)	0.02 (0.009)	0.70 (0.011)	0.46 (0.000)	0.97 (0.016)	0.08 (0.027)	0.26 (0.031)	0.02 (0.026)	0.71 (0.028)	0.46 (0.028)	0.98 (0.043)
BNI	0.22 (0.028)	0.34 (0.019)	0.08 (0.015)	0.58 (0.019)	0.28 (0.000)	0.96 (0.028)	0.19 (0.041)	0.33 (0.036)	0.09 (0.029)	0.58 (0.036)	0.27 (0.026)	0.96 (0.053)
C	0.10 (0.015)	0.33 (0.021)	0.16 (0.015)	0.56 (0.017)	0.25 (0.000)	0.98 (0.028)	0.03 (0.012)	0.34 (0.033)	0.16 (0.024)	0.58 (0.025)	0.25 (0.025)	1.00 (0.043)
CAT	0.19 (0.029)	0.35 (0.023)	0.07 (0.016)	0.57 (0.023)	0.21 (0.000)	0.96 (0.033)	0.12 (0.028)	0.35 (0.034)	0.08 (0.024)	0.58 (0.032)	0.21 (0.028)	0.97 (0.048)
CL	0.07 (0.009)	0.35 (0.016)	0.05 (0.010)	0.61 (0.014)	0.39 (0.000)	0.98 (0.022)	0.07 (0.026)	0.36 (0.054)	0.06 (0.033)	0.59 (0.048)	0.39 (0.037)	0.98 (0.074)
CMCSA	0.13 (0.021)	0.33 (0.020)	0.08 (0.013)	0.60 (0.019)	0.27 (0.000)	0.97 (0.028)	0.09 (0.031)	0.33 (0.042)	0.08 (0.027)	0.60 (0.041)	0.28 (0.034)	0.98 (0.060)
COF	0.20 (0.030)	0.37 (0.019)	0.13 (0.014)	0.55 (0.018)	0.27 (0.000)	0.98 (0.027)	0.08 (0.025)	0.37 (0.033)	0.13 (0.025)	0.56 (0.030)	0.27 (0.025)	1.00 (0.047)
COST	0.23 (0.031)	0.37 (0.022)	0.05 (0.015)	0.55 (0.023)	0.26 (0.000)	0.94 (0.032)	0.13 (0.047)	0.38 (0.052)	0.04 (0.038)	0.57 (0.051)	0.26 (0.039)	0.96 (0.075)
CPB	0.09 (0.006)	0.27 (0.013)	0.04 (0.010)	0.68 (0.011)	0.58 (0.000)	0.98 (0.018)	0.09 (0.031)	0.29 (0.077)	0.05 (0.063)	0.66 (0.069)	0.56 (0.065)	0.97 (0.109)
CSCO	0.31 (0.040)	0.49 (0.022)	0.10 (0.018)	0.40 (0.021)	0.25 (0.000)	0.94 (0.032)	0.15 (0.040)	0.50 (0.043)	0.10 (0.036)	0.42 (0.036)	0.26 (0.034)	0.97 (0.059)
CVS	0.14 (0.015)	0.24 (0.010)	0.00 (0.008)	0.74 (0.011)	0.55 (0.000)	0.98 (0.015)	0.10 (0.031)	0.24 (0.033)	0.01 (0.024)	0.74 (0.032)	0.56 (0.032)	0.98 (0.047)
CVX	0.08 (0.015)	0.35 (0.022)	0.08 (0.015)	0.59 (0.020)	0.22 (0.000)	0.98 (0.031)	0.12 (0.027)	0.34 (0.035)	0.09 (0.025)	0.56 (0.034)	0.22 (0.027)	0.95 (0.050)
DD	0.17 (0.022)	0.30 (0.018)	0.10 (0.015)	0.60 (0.018)	0.25 (0.000)	0.96 (0.026)	0.09 (0.027)	0.30 (0.037)	0.10 (0.031)	0.62 (0.034)	0.26 (0.033)	0.97 (0.052)
DELL	0.21 (0.036)	0.39 (0.027)	0.09 (0.018)	0.52 (0.026)	0.19 (0.000)	0.95 (0.039)	0.09 (0.028)	0.40 (0.037)	0.08 (0.027)	0.54 (0.035)	0.19 (0.029)	0.98 (0.053)
DIS	0.11 (0.010)	0.27 (0.013)	0.09 (0.011)	0.66 (0.013)	0.38 (0.000)	0.98 (0.019)	0.06 (0.019)	0.27 (0.044)	0.09 (0.037)	0.66 (0.041)	0.39 (0.043)	0.99 (0.063)
DOW	0.17 (0.017)	0.25 (0.012)	0.11 (0.011)	0.67 (0.010)	0.37 (0.000)	0.97 (0.016)	0.08 (0.032)	0.26 (0.039)	0.10 (0.036)	0.68 (0.029)	0.39 (0.040)	0.99 (0.052)

	SPvMEM						Univariate MEM					
	a_i	α_i	γ_i	β_i	ν_i	π_i	a_i	α_i	γ_i	β_i	ν_i	π_i
DVN	0.11 (0.020)	0.26 (0.016)	0.05 (0.011)	0.70 (0.014)	0.40 (0.000)	0.98 (0.022)	0.17 (0.058)	0.24 (0.043)	0.07 (0.027)	0.69 (0.040)	0.41 (0.036)	0.97 (0.060)
EMC	0.22 (0.032)	0.25 (0.015)	0.09 (0.011)	0.68 (0.013)	0.35 (0.000)	0.98 (0.021)	0.11 (0.031)	0.26 (0.028)	0.09 (0.022)	0.69 (0.024)	0.36 (0.025)	0.99 (0.038)
ETR	0.09 (0.011)	0.30 (0.017)	0.10 (0.014)	0.63 (0.015)	0.27 (0.000)	0.98 (0.024)	0.07 (0.016)	0.29 (0.031)	0.10 (0.025)	0.63 (0.028)	0.28 (0.025)	0.98 (0.044)
EXC	0.13 (0.016)	0.33 (0.012)	0.08 (0.013)	0.60 (0.012)	0.34 (0.000)	0.97 (0.019)	0.10 (0.035)	0.32 (0.041)	0.09 (0.039)	0.61 (0.037)	0.33 (0.041)	0.97 (0.059)
F	0.35 (0.032)	0.34 (0.010)	0.07 (0.009)	0.60 (0.009)	0.67 (0.000)	0.97 (0.014)	0.20 (0.087)	0.34 (0.054)	0.07 (0.047)	0.61 (0.042)	0.64 (0.046)	0.98 (0.072)
FDX	0.15 (0.026)	0.31 (0.015)	0.05 (0.014)	0.63 (0.017)	0.25 (0.000)	0.96 (0.024)	0.06 (0.028)	0.30 (0.029)	0.05 (0.028)	0.65 (0.028)	0.25 (0.033)	0.98 (0.043)
GD	0.14 (0.021)	0.33 (0.018)	0.07 (0.014)	0.59 (0.018)	0.29 (0.000)	0.96 (0.026)	0.08 (0.024)	0.33 (0.034)	0.07 (0.028)	0.61 (0.031)	0.29 (0.030)	0.98 (0.048)
GE	0.10 (0.014)	0.34 (0.020)	0.12 (0.016)	0.56 (0.018)	0.26 (0.000)	0.96 (0.028)	0.03 (0.014)	0.35 (0.043)	0.11 (0.035)	0.59 (0.038)	0.27 (0.035)	0.99 (0.060)
GILD	0.24 (0.050)	0.41 (0.023)	0.06 (0.018)	0.53 (0.024)	0.20 (0.000)	0.97 (0.034)	0.17 (0.058)	0.41 (0.035)	0.07 (0.028)	0.53 (0.034)	0.20 (0.031)	0.97 (0.051)
GS	0.16 (0.027)	0.43 (0.023)	0.13 (0.018)	0.47 (0.014)	0.22 (0.000)	0.97 (0.028)	0.07 (0.033)	0.44 (0.045)	0.13 (0.036)	0.48 (0.027)	0.23 (0.038)	0.99 (0.055)
HAL	0.23 (0.034)	0.25 (0.014)	0.10 (0.012)	0.68 (0.010)	0.40 (0.000)	0.98 (0.018)	0.13 (0.089)	0.25 (0.059)	0.10 (0.052)	0.69 (0.042)	0.41 (0.058)	0.99 (0.077)
HD	0.17 (0.026)	0.30 (0.018)	0.09 (0.013)	0.62 (0.018)	0.26 (0.000)	0.96 (0.026)	0.07 (0.024)	0.30 (0.033)	0.10 (0.024)	0.64 (0.029)	0.27 (0.028)	0.98 (0.045)
HNZ	0.04 (0.006)	0.19 (0.008)	0.06 (0.006)	0.77 (0.007)	0.56 (0.000)	0.99 (0.011)	0.04 (0.018)	0.19 (0.031)	0.07 (0.025)	0.76 (0.028)	0.56 (0.034)	0.99 (0.043)
HON	0.14 (0.021)	0.33 (0.015)	0.09 (0.011)	0.60 (0.013)	0.33 (0.000)	0.98 (0.021)	0.08 (0.032)	0.33 (0.038)	0.09 (0.028)	0.61 (0.031)	0.33 (0.034)	0.99 (0.051)
HPQ	0.11 (0.015)	0.25 (0.014)	0.12 (0.011)	0.68 (0.012)	0.37 (0.000)	0.99 (0.019)	0.06 (0.020)	0.26 (0.030)	0.11 (0.023)	0.68 (0.026)	0.38 (0.026)	0.99 (0.042)
IBM	0.10 (0.015)	0.30 (0.021)	0.14 (0.015)	0.59 (0.021)	0.22 (0.000)	0.96 (0.031)	0.04 (0.013)	0.31 (0.032)	0.14 (0.024)	0.60 (0.028)	0.24 (0.029)	0.98 (0.044)
INTC	0.28 (0.043)	0.43 (0.022)	0.13 (0.019)	0.45 (0.025)	0.20 (0.000)	0.94 (0.035)	0.14 (0.044)	0.43 (0.046)	0.13 (0.039)	0.47 (0.046)	0.21 (0.040)	0.97 (0.068)
JNJ	0.04 (0.006)	0.26 (0.016)	0.13 (0.011)	0.66 (0.013)	0.37 (0.000)	0.99 (0.021)	0.02 (0.009)	0.27 (0.040)	0.12 (0.027)	0.66 (0.031)	0.38 (0.033)	1.00 (0.052)
JPM	0.14 (0.016)	0.37 (0.020)	0.14 (0.014)	0.54 (0.017)	0.30 (0.000)	0.98 (0.027)	0.05 (0.016)	0.38 (0.039)	0.13 (0.027)	0.55 (0.033)	0.31 (0.028)	1.00 (0.052)
KO	0.09 (0.009)	0.33 (0.012)	0.07 (0.011)	0.61 (0.014)	0.36 (0.000)	0.97 (0.019)	0.04 (0.016)	0.33 (0.041)	0.06 (0.038)	0.62 (0.043)	0.38 (0.048)	0.99 (0.063)
LMT	0.14 (0.016)	0.31 (0.018)	0.09 (0.012)	0.61 (0.016)	0.35 (0.000)	0.96 (0.025)	0.06 (0.020)	0.32 (0.038)	0.09 (0.028)	0.63 (0.032)	0.37 (0.029)	0.99 (0.051)
LOW	0.27 (0.038)	0.31 (0.020)	0.10 (0.014)	0.59 (0.021)	0.25 (0.000)	0.95 (0.030)	0.11 (0.034)	0.32 (0.032)	0.09 (0.024)	0.61 (0.031)	0.25 (0.029)	0.98 (0.047)
MCD	0.08 (0.009)	0.23 (0.011)	0.05 (0.009)	0.73 (0.010)	0.43 (0.000)	0.98 (0.016)	0.04 (0.017)	0.23 (0.037)	0.05 (0.029)	0.74 (0.030)	0.43 (0.037)	0.99 (0.050)
MDT	0.08 (0.012)	0.25 (0.012)	0.09 (0.007)	0.69 (0.011)	0.60 (0.000)	0.98 (0.017)	0.06 (0.026)	0.25 (0.037)	0.09 (0.023)	0.69 (0.034)	0.61 (0.034)	0.98 (0.051)
MMM	0.11 (0.017)	0.37 (0.021)	0.10 (0.016)	0.54 (0.022)	0.23 (0.000)	0.96 (0.031)	0.09 (0.020)	0.37 (0.037)	0.11 (0.026)	0.54 (0.036)	0.23 (0.030)	0.96 (0.053)
MO	0.13 (0.010)	0.32 (0.010)	0.07 (0.009)	0.61 (0.009)	0.70 (0.000)	0.97 (0.014)	0.10 (0.049)	0.33 (0.067)	0.08 (0.059)	0.61 (0.059)	0.67 (0.059)	0.97 (0.094)
MRK	0.17 (0.013)	0.34 (0.014)	0.07 (0.009)	0.59 (0.011)	1.00 (0.000)	0.97 (0.018)	0.22 (0.084)	0.34 (0.089)	0.10 (0.059)	0.55 (0.081)	0.89 (0.057)	0.94 (0.124)
MS	0.30 (0.043)	0.39 (0.021)	0.16 (0.016)	0.49 (0.019)	0.24 (0.000)	0.96 (0.029)	0.12 (0.056)	0.40 (0.052)	0.16 (0.040)	0.51 (0.044)	0.24 (0.042)	0.99 (0.071)
MSFT	0.16 (0.019)	0.37 (0.019)	0.09 (0.016)	0.52 (0.021)	0.23 (0.000)	0.94 (0.029)	0.05 (0.015)	0.38 (0.029)	0.08 (0.026)	0.56 (0.029)	0.24 (0.030)	0.98 (0.043)
NKE	0.24 (0.025)	0.28 (0.015)	0.10 (0.013)	0.61 (0.017)	0.35 (0.000)	0.94 (0.024)	0.08 (0.029)	0.26 (0.049)	0.10 (0.042)	0.67 (0.045)	0.36 (0.044)	0.98 (0.069)

	SPVMEM						MEM					
	a_i	α_i	γ_i	β_i	ν_i	π_i	a_i	α_i	γ_i	β_i	ν_i	π_i
NSC	0.22 (0.032)	0.28 (0.014)	0.06 (0.013)	0.66 (0.015)	0.28 (0.000)	0.97 (0.022)	0.12 (0.045)	0.28 (0.030)	0.06 (0.027)	0.68 (0.029)	0.29 (0.029)	0.98 (0.044)
ORCL	0.27 (0.037)	0.45 (0.023)	0.08 (0.018)	0.47 (0.022)	0.24 (0.000)	0.96 (0.034)	0.15 (0.042)	0.45 (0.046)	0.09 (0.036)	0.48 (0.041)	0.25 (0.032)	0.98 (0.064)
OXY	0.13 (0.021)	0.25 (0.020)	0.08 (0.014)	0.68 (0.019)	0.25 (0.000)	0.98 (0.028)	0.14 (0.030)	0.25 (0.028)	0.09 (0.020)	0.68 (0.027)	0.26 (0.022)	0.97 (0.040)
PEP	0.06 (0.007)	0.23 (0.014)	0.09 (0.009)	0.70 (0.012)	0.42 (0.000)	0.98 (0.019)	0.03 (0.013)	0.23 (0.042)	0.09 (0.030)	0.71 (0.034)	0.43 (0.040)	0.99 (0.056)
PFE	0.11 (0.016)	0.42 (0.007)	−0.03 (0.009)	0.58 (0.010)	0.70 (0.000)	0.98 (0.013)	0.08 (0.069)	0.41 (0.055)	−0.02 (0.062)	0.58 (0.062)	0.66 (0.075)	0.98 (0.089)
PG	0.07 (0.008)	0.32 (0.018)	0.08 (0.013)	0.61 (0.016)	0.33 (0.000)	0.97 (0.025)	0.06 (0.017)	0.33 (0.050)	0.08 (0.037)	0.60 (0.045)	0.33 (0.042)	0.97 (0.070)
QCOM	0.21 (0.036)	0.29 (0.018)	0.13 (0.015)	0.61 (0.016)	0.27 (0.000)	0.97 (0.025)	0.15 (0.052)	0.30 (0.040)	0.14 (0.034)	0.60 (0.031)	0.28 (0.037)	0.97 (0.054)
RF	0.05 (0.006)	0.26 (0.011)	0.08 (0.009)	0.70 (0.010)	0.55 (0.000)	1.00 (0.016)	0.06 (0.016)	0.26 (0.029)	0.08 (0.025)	0.69 (0.026)	0.53 (0.024)	0.99 (0.041)
S	0.29 (0.028)	0.29 (0.011)	0.12 (0.009)	0.63 (0.008)	0.75 (0.000)	0.98 (0.014)	0.10 (0.043)	0.30 (0.034)	0.11 (0.029)	0.65 (0.025)	0.77 (0.029)	1.00 (0.044)
SGP	0.24 (0.017)	0.27 (0.010)	0.00 (0.009)	0.70 (0.008)	0.90 (0.000)	0.97 (0.013)	0.08 (0.040)	0.27 (0.047)	0.00 (0.041)	0.72 (0.033)	0.96 (0.040)	0.99 (0.061)
SLB	0.24 (0.043)	0.29 (0.023)	0.07 (0.016)	0.64 (0.023)	0.19 (0.000)	0.96 (0.034)	0.15 (0.031)	0.28 (0.024)	0.08 (0.016)	0.65 (0.023)	0.19 (0.021)	0.97 (0.034)
SLE	0.12 (0.010)	0.28 (0.008)	0.02 (0.007)	0.69 (0.008)	0.75 (0.000)	0.98 (0.012)	0.10 (0.047)	0.30 (0.054)	0.02 (0.046)	0.67 (0.051)	0.73 (0.047)	0.98 (0.078)
SO	0.04 (0.008)	0.27 (0.013)	0.07 (0.011)	0.68 (0.009)	0.36 (0.000)	0.99 (0.017)	0.03 (0.015)	0.28 (0.034)	0.07 (0.029)	0.68 (0.025)	0.37 (0.034)	0.99 (0.045)
T	0.15 (0.016)	0.29 (0.012)	0.08 (0.011)	0.65 (0.010)	0.46 (0.000)	0.98 (0.017)	0.06 (0.022)	0.30 (0.035)	0.08 (0.035)	0.66 (0.030)	0.47 (0.032)	1.00 (0.049)
TGT	0.17 (0.024)	0.26 (0.016)	0.10 (0.013)	0.67 (0.015)	0.29 (0.000)	0.97 (0.023)	0.07 (0.028)	0.27 (0.031)	0.09 (0.027)	0.68 (0.029)	0.30 (0.030)	0.99 (0.045)
TWX	0.11 (0.017)	0.29 (0.015)	0.07 (0.009)	0.66 (0.015)	0.37 (0.000)	0.98 (0.021)	0.06 (0.020)	0.29 (0.038)	0.07 (0.025)	0.66 (0.036)	0.37 (0.033)	0.99 (0.054)
TXN	0.10 (0.027)	0.24 (0.016)	0.11 (0.013)	0.70 (0.015)	0.21 (0.000)	0.99 (0.023)	0.07 (0.029)	0.24 (0.029)	0.11 (0.024)	0.69 (0.026)	0.22 (0.031)	0.99 (0.040)
TYC	0.12 (0.016)	0.34 (0.014)	0.09 (0.012)	0.61 (0.012)	0.39 (0.000)	0.99 (0.020)	0.08 (0.028)	0.34 (0.035)	0.09 (0.030)	0.61 (0.029)	0.39 (0.027)	0.99 (0.048)
UNH	0.14 (0.021)	0.33 (0.013)	0.06 (0.010)	0.61 (0.014)	0.41 (0.000)	0.98 (0.020)	0.13 (0.043)	0.34 (0.035)	0.06 (0.026)	0.60 (0.036)	0.41 (0.032)	0.97 (0.052)
UPS	0.08 (0.008)	0.28 (0.013)	0.12 (0.010)	0.63 (0.013)	0.40 (0.000)	0.97 (0.019)	0.09 (0.030)	0.28 (0.047)	0.12 (0.036)	0.62 (0.046)	0.41 (0.041)	0.96 (0.068)
USB	0.10 (0.005)	0.28 (0.012)	0.11 (0.011)	0.65 (0.011)	0.44 (0.000)	0.99 (0.018)	0.05 (0.009)	0.28 (0.040)	0.11 (0.036)	0.66 (0.035)	0.46 (0.035)	1.00 (0.056)
UTX	0.08 (0.015)	0.28 (0.018)	0.13 (0.014)	0.64 (0.016)	0.28 (0.000)	0.98 (0.025)	0.07 (0.019)	0.28 (0.031)	0.14 (0.025)	0.64 (0.029)	0.28 (0.026)	0.98 (0.044)
VZ	0.12 (0.015)	0.29 (0.018)	0.09 (0.013)	0.63 (0.015)	0.29 (0.000)	0.97 (0.024)	0.05 (0.019)	0.31 (0.041)	0.08 (0.031)	0.64 (0.034)	0.29 (0.035)	0.99 (0.056)
WAG	0.13 (0.018)	0.24 (0.016)	0.06 (0.011)	0.70 (0.014)	0.33 (0.000)	0.97 (0.022)	0.07 (0.025)	0.23 (0.037)	0.07 (0.027)	0.71 (0.031)	0.34 (0.031)	0.98 (0.050)
WFC	0.11 (0.011)	0.32 (0.017)	0.12 (0.014)	0.59 (0.017)	0.29 (0.000)	0.97 (0.025)	0.05 (0.012)	0.32 (0.032)	0.11 (0.027)	0.61 (0.029)	0.31 (0.026)	0.99 (0.046)
WMB	0.12 (0.019)	0.25 (0.008)	0.06 (0.008)	0.71 (0.006)	0.66 (0.000)	0.99 (0.011)	0.09 (0.058)	0.25 (0.031)	0.08 (0.030)	0.71 (0.023)	0.67 (0.034)	1.00 (0.041)
WMT	0.10 (0.016)	0.30 (0.019)	0.09 (0.013)	0.62 (0.019)	0.26 (0.000)	0.97 (0.028)	0.05 (0.014)	0.31 (0.031)	0.09 (0.023)	0.63 (0.030)	0.27 (0.029)	0.98 (0.044)
WY	0.20 (0.029)	0.33 (0.019)	0.10 (0.016)	0.58 (0.020)	0.23 (0.000)	0.96 (0.029)	0.17 (0.039)	0.34 (0.034)	0.10 (0.027)	0.57 (0.035)	0.23 (0.026)	0.96 (0.051)
XOM	0.10 (0.014)	0.30 (0.018)	0.10 (0.016)	0.62 (0.017)	0.25 (0.000)	0.97 (0.026)	0.07 (0.019)	0.29 (0.031)	0.10 (0.029)	0.63 (0.029)	0.26 (0.029)	0.97 (0.045)
XRX	0.12 (0.014)	0.23 (0.009)	0.05 (0.007)	0.73 (0.007)	0.67 (0.000)	0.99 (0.012)	0.07 (0.029)	0.23 (0.027)	0.08 (0.023)	0.73 (0.023)	0.69 (0.025)	1.00 (0.037)

Estimated parameters and standard errors (in parenthesis) for the SPvMEM (left) and the univariate MEM (right).