A Multi-Level Panel Smooth Transition Autoregression for US Sectoral Production*

Dennis Fok, Dick van Dijk and Philip Hans Franses†

Econometric Institute
Erasmus University Rotterdam

Econometric Institute Report EI 2003-43

November 2003

Abstract

We introduce a multi-level smooth transition model for a panel of time series variables, which can be used to examine the presence of common non-linear features across many such variables. The model is positioned in between a fully pooled model, which imposes such common features, and a fully heterogeneous model, which might render estimation problems for some of the panel members. To keep the model tractable, we introduce a second-stage model, which links the parameters in the transition functions with observable explanatory variables. We discuss representation, estimation by concentrated simulated maximum likelihood and inference. We illustrate our model for data on industrial production of 18 US manufacturing sectors, and document that there are subtle differences across sectors in leads and lags for business cycle recessions and expansions.

Keywords: Panel of time series, business cycle, non-linearity
JEL Classification Codes: C23, C51, E32

*We thank participants of the Tinbergen Conference (Rotterdam, April 2003) and of seminars at the University of Western Australia and Curtin University (Perth, Australia, October 2003) and Richard Paap for their helpful comments. The usual disclaimer applies.

†Address for correspondence: Dennis Fok, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, `email: dfok@few.eur.nl`

1 Introduction

Characterizing business cycle dynamics has since long been a topic of intense research. The stylized empirical fact that recessions last much shorter but are more severe than expansions recently has led to the application of a wide range of non-linear time series (regression) models to macroeconomic variables such as industrial output and (un)employment. Most studies in this area make use of aggregate, often nationwide, variables. Relatively little attention has been given to business cycle asymmetries in disaggregated variables, such as sector-level output or state-level employment, although Cooper (1998), Bidarkota (1999), and Owyang, Piger and Wall (2003) are notable exceptions.

When using disaggregated output or unemployment data, a question of prime interest is the extent to which there is co-movement across different states or sectors, and in particular the possible presence of common cyclical components. This issue is usually analyzed by means of factor models, where for example state-level output growth is decomposed into national, regional and state-specific components, see Quah (1996), Clark (1998), Forni and Reichlin (1998) and Del Negro (2002) for illustrations of this approach. Note that the different components in these models typically are assumed to be independent and, more importantly, linear. Although extensions to factor models with non-linear components may be feasible, to the best of our knowledge this has not yet been pursued in practice. Instead, the presence and relevance of common non-linear components is often analyzed by means of comparing results from univariate non-linear models, as in Owyang *et al.* (2003) for Markov-Switching models applied to state-level output. Such an analysis may render useful insights, but obviously, using univariate models is not efficient as they ignore the information contained in the common cyclical component(s).

So far, the only attempt to develop a formal methodology to investigate the presence of common non-linear components in multivariate time series models was made by Anderson and Vahid (1998), building upon ideas from the literature on common features, see Vahid and Engle (1993). A drawback of this approach might be that it is limited to small-sized systems though, as multivariate time series models typically become unreliable for systems consisting of many variables.

In this paper, we propose a novel approach that can be used to examine common non-linear cyclical components in possibly many disaggregated variables. For this purpose, we develop a panel smooth transition autoregressive (STAR) model that imposes a common regime-switching mechanism while allowing for considerable heterogeneity in the timing of the regime changes across series. This makes the model particularly useful for situations where the non-linear dynamics are driven by a common regime-switching component, but where the response to this component can be different across variables. For example, probably all sectors in an economy are affected by nationwide recessions, but some sectors may enter into (or get out of) recessions earlier than others. In order to arrive at a parsimonious model, we assume a second-level model for the parameters in the regime-switching mechanism of the STAR model, where these are then related to, for example, sector-specific characteristics.

It should be remarked that we completely abstain from the notion of cointegration in our context, as this amounts to yet another additional degree of complexity. In principle, our panel STAR model can be extended to incorporate common trends, but we believe that the relevant statistical theory should first be derived and this is beyond the scope of the present paper. Furthermore, with respect to our empirical application to US sectoral output, we note that Pesaran, Pierse and Lee (1993) and Engle and Issler (1995) find little evidence for the existence of common trends in these series.

Our paper is organized as follows. In Section 2, we present our multi-level panel STAR model and describe its main features. Parameter estimation is feasible but not straightforward and therefore we dedicate a full Section 3 to this topic. We also briefly discuss inference in the panel STAR model in this section, and suggest several ways to extract relevant information concerning business cycles. In Section 4, we present our application to US industrial production in 18 manufacturing sectors, a number that certainly cannot easily be dealt with by vector autoregressive time series models. We document that a partially heterogeneous panel STAR model outperforms a fully pooled model, leading to subtle differences across sectors in leads and lags for business cycle recessions and expansions. In particular, we find that the four sectors machinery, aerospace, primary metal and fabricated metal products enter recessions earlier than other manufacturing sectors. In Section 5 we conclude with several suggestions for further research.

2 A Multi-Level Panel STAR model

The basic smooth transition autoregressive [STAR] model, as discussed extensively in Granger and Teräsvirta (1993), Teräsvirta (1994), Franses and van Dijk (2000), and van Dijk, Teräsvirta and Franses (2002), embraces two regimes, where the prevailing regime at time t is determined by the value of an observable variable s_t. A STAR model of order P for a univariate time series $y_{i,t}$, which may represent for example quarterly output growth in sector i, $i = 1, \dots, N$ and $t = 1, \dots, T$, is given by

$$y_{i,t} = \alpha_{i,0} + \sum_{j=1}^{P} \alpha_{i,j} y_{i,t-j} + G(s_t; \gamma_i, \tau_i)(\beta_{i,0} + \sum_{j=1}^{P} \beta_{i,j} y_{i,t-j}) + \varepsilon_{i,t}, \tag{1}$$

or

$$y_{i,t} = \boldsymbol{\alpha}_i' \mathbf{x}_{i,t} + \boldsymbol{\beta}_i' \mathbf{x}_{i,t} G(s_t; \gamma_i, \tau_i) + \varepsilon_{i,t}, \tag{2}$$

where $\mathbf{x}_t = (1, \tilde{\mathbf{x}}_t')'$ with $\tilde{\mathbf{x}}_t = (y_{t-1}, \dots, y_{t-p})'$, $\boldsymbol{\alpha}_i = (\alpha_{i,0}, \alpha_{i,1}, \dots, \alpha_{i,p})'$, $\boldsymbol{\beta}_i$ is similarly defined, and the properties of $\varepsilon_{i,t}$ are discussed in detail below.

In general, the so-called transition function $G(s_t; \gamma_i, \tau_i)$ in (1) is a continuous function that is bounded between 0 to 1. Two interpretations of the STAR model then are possible. On the one hand, the STAR model can be thought of as a regime-switching model that allows for two regimes, associated with the extreme values of the transition function, $G(s_t; \gamma_i, \tau_i) = 0$ and $G(s_t; \gamma_i, \tau_i) = 1$, where the transition from one regime to the other is smooth. On the other hand, the STAR model can be said to allow for a "continuum" of regimes, each associated with a different value of $G(s_t; \gamma_i, \tau_i)$ between 0 and 1. In this paper we will use the two-regime interpretation.

A popular choice for $G(s_t; \gamma_i, \tau_i)$, which we also employ in the present paper, is the logistic function

$$G(s_t; \gamma_i, \tau_i) = \frac{1}{1 + \exp(-\gamma_i(s_t - \tau_i))} \qquad \text{with } \gamma_i > 0, \tag{3}$$

where the parameter restriction $\gamma_i > 0$ is an identifying restriction. The parameter τ_i in (3) can be interpreted as the threshold between the two regimes, in the sense that the logistic function changes monotonically from 0 to 1 as the transition variable s_t increases and $G(s_t; \gamma_i, \tau_i) = 0.5$ when $s_t = \tau_i$. The parameter γ_i determines the smoothness of the change in the value of the logistic function and, thus, the smoothness of the transition from one regime to the other. As $\gamma_i \to \infty$, the logistic function $G(s_t; \gamma_i, \tau_i)$ approaches the indicator function $\mathtt{I}[s_t > \tau_i]$, defined as $\mathtt{I}[A] = 1$ if A is true and $\mathtt{I}[A] = 0$ otherwise, and, consequently, the change of $G(s_t; \gamma_i, \tau_i)$ from 0 to 1 becomes instantaneous at $s_t = \tau_i$. Finally, when $\gamma_i \to 0$, the logistic function approaches a constant (equal to 0.5) and when $\gamma_i = 0$, the STAR model reduces to a linear dynamic model with parameters $(\alpha_{i,j} + \beta_{i,j})/2$, $j = 0, 1, \ldots, p$.

As the logistic function (3) is a monotonic transformation of the transition variable s_t, the two regimes in the STAR model (2) are associated with small and large values of s_t (relative to τ_i). This makes the model convenient for modelling business cycle asymmetry where, through a suitable choice for s_t, the regimes of the STAR can be related to expansions and recessions, see Teräsvirta and Anderson (1992) for an empirical example. The value of the switching function then can be interpreted as an indicator function of the business cycle.

In the STAR model (2) with (3), one usually assumes that $\varepsilon_{i,t}$ is a martingale difference with respect to the history of the time series up to time $t-1$, which is denoted as $\Omega_{i,t-1} = \{y_{i,t-1}, y_{i,t-2}, \ldots, y_{i,1-P}\}$, that is, $\mathsf{E}[\varepsilon_{i,t}|\Omega_{i,t-1}] = 0$. For simplicity, we also assume that the conditional variance of ε_t is constant, $\mathsf{E}[\varepsilon_t^2|\Omega_{t-1}] = \sigma_i^2$. An extension of the STAR model which allows for (possibly asymmetric) autoregressive conditional heteroscedasticity [ARCH] is considered in Lundbergh and Teräsvirta (1998). For our particular application, we additionally assume that the $\varepsilon_{i,t}$ are mutually uncorrelated across time and across sectors. It is possible to allow for various unrestricted covariance structures, but this would seriously complicate parameter estimation. Another approach to allow for contemporaneous correlation across the errors of the N equations is to introduce common exogenous variable $z_{1,t}, \ldots, z_{k,t}$ as additional regressors in all equations, such as, for example, the world economy growth rate, see Paap, Franses and van Dijk (2003). In the univariate context, the resultant smooth transition regression (STR) model is discussed in detail in Teräsvirta (1998).

The specification of the STAR model in (2) with (3), with no cross-equation restrictions on the parameters $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$, γ_i and τ_i, $i = 1, \ldots, N$, leads to a fully heterogeneous panel time series model. Furthermore, given the particular assumptions on the shocks $\varepsilon_{i,t}$ discussed above, the model can be estimated sector by sector and the resulting regime-switching dynamics for each sector can subsequently be compared and inspected for similarities and differences, as in Owyang *et al.* (2003).

Obviously, the fully heterogeneous model could amount to a huge amount of parameters (namely N(2(1+P)+3)) to be estimated, if N (and/or P) is reasonably large. Hence, it is of interest to see if the panel model can be restricted to become more parsimonious. There are at least two important reasons for imposing more structure on the panel STAR model. First, estimating sector-specific STAR models might not work for all sectors considered. It can happen that the likelihood function does not have a well-defined maximum, which makes it impossible to obtain reliable parameter estimates. This behavior of the STAR model is well-known to practitioners, and in most of the cases a few outliers are found to be responsible for these difficulties. Instead of states representing recessions and expansions, one of the two regimes of the STAR model will then capture the outliers, see van Dijk, Lucas and Franses (1999) for example.

Second, if the time series in the panel $y_{i,t}$ represent sector-level output or state-level employment, it is obvious that their dynamics may bear close similarities, in particular in terms of timing of the regime switches if these regimes represent business cycle recessions and expansions, for example. To impose a common regime-switching mechanism across panel members, one might pool the parameters in the switching functions across the STAR models, that is, set $\gamma_i = \gamma$ and $\tau_i = \tau$ for all $i = 1, \ldots, N$. However, this pooling approach is rather restrictive. In particular, the assumption that the timing of regime switches is exactly the same across all sectors may be unrealistic. Even though all sectors in an economy probably are affected by nationwide recessions, some sectors may enter into (or get out of) recessions earlier than others.

In this paper we therefore propose a model specification in between the pooled model and the fully heterogeneous model, which has these two cases occurring at the boundaries of the model specification. The basic idea is to introduce a second level regression model for the parameters γ_i and τ_i, which makes these a function of observed explanatory variables, the value of which differ across sectors, and an unobserved error term. The resulting model allows for common features across sectors as well as for (unexplained) differences between sectors. The common features are captured by observable sector characteristics, while the differences are captured by random effects. To be more precise, in this paper we describe the switching parameters by the regression model

$$\begin{pmatrix} \log(\gamma_i) \\ \tau_i \end{pmatrix} = \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i, \qquad \boldsymbol{\eta}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \tag{4}$$

where $\mathbf{w}_i$ denotes a $(Q \times 1)$ vector consisting of a constant and $Q-1$ observable characteristics of sector i, and where $\boldsymbol{\delta}$ is a $(Q \times 2)$ matrix of unknown coefficients. Note that we model $\log(\gamma_i)$ instead of γ_i directly to ensure that $\gamma_i > 0$. Furthermore, note that the specification where the switching parameters are restricted to be equal across sectors is nested in (4). That is, in case $\mathbf{w}_i$ only contains a constant and at the same time $\Sigma_\eta = 0$, we obtain the pooled specification. In sum, our model allows for meaningful flexibility in a panel time series model for non-linear data, which at the same time allows us to examine the potential presence of common non-linear features.

3 Parameter estimation

Estimation of the parameters in the multi-level panel STAR model (2)-(4) is not straightforward, as the model contains a latent second level regression for the parameters γ_i and τ_i that appear in non-linear functions in the first-stage model. For clarity, let us first restate the complete model, which reads

$$y_{i,t} = \boldsymbol{\alpha}_i'\mathbf{x}_{i,t} + \boldsymbol{\beta}_i'\mathbf{x}_{i,t}G(s_t;\gamma_i,\tau_i) + \varepsilon_{i,t}, \qquad i = 1,\ldots,N, \tag{5}$$

$$\begin{pmatrix} \log(\gamma_i) \\ \tau_i \end{pmatrix} = \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i, \tag{6}$$

$$\boldsymbol{\eta}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \tag{7}$$

$$\varepsilon_{i,t} \sim N(0, \sigma_i^2), \qquad i = 1,\ldots,N. \tag{8}$$

The likelihood function for this model equals $\mathcal{L} = \prod_i \mathcal{L}_i$, where $\mathcal{L}_i$ is the contribution of sector i, which is given by

$$\mathcal{L}_i = \int_{\boldsymbol{\eta}_i} \prod_t \phi(e_{i,t}(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i); 0, \sigma_i^2)\phi(\boldsymbol{\eta}_i; \mathbf{0}, \boldsymbol{\Sigma}_\eta)d\boldsymbol{\eta}_i, \tag{9}$$

where $\phi(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes the (bivariate) normal density function with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ evaluated at $\mathbf{x}$ and where $e_{i,t}(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\boldsymbol{\theta}_i)$ gives the error for sector i and period t, given the parameters associated with the two regimes, that is $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$, and the switching parameters $\boldsymbol{\theta}_i = (\log(\gamma_i),\tau_i)'$. To be precise, the error is given by

$$e_{i,t}(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i,\boldsymbol{\theta}_i) = y_{i,t} - \boldsymbol{\alpha}_i'\mathbf{x}_{i,t} - \boldsymbol{\beta}_i'\mathbf{x}_{i,t}G(s_t;\gamma_i,\tau_i). \tag{10}$$

Parameter estimates may be obtained by maximizing the log-likelihood function $\sum_{i=1}^{N}\log\mathcal{L}_i$. This maximization problem poses two difficulties however. First of all, the likelihood function contains many parameters. As we allow the AR-parameters to be different across sectors, we have $2(1+P)+1$ parameters per sector, comprising two intercepts, P AR-parameters for each regime, and the error variance σ_i^2. Furthermore, we have $2Q$ parameters relating the switching parameters to the sector characteristics. Finally, we have three parameters contained in $\boldsymbol{\Sigma}_\eta$. Directly maximizing the likelihood over all parameters therefore is very difficult. To circumvent this problem, we concentrate the likelihood function with respect to the sector-specific variables, see for example Davidson and MacKinnon (1993) for a general discussion on likelihood concentration. Such a concentration approach is often used in estimating STAR models for single series, see van Dijk *et al.* (2002). In our case, concentrating the likelihood is more complicated as we also have to deal with the stochastic nature of the switching parameters.

The second problem in maximizing the likelihood is the fact that calculating the likelihood function requires solving N two-dimensional integrals, one for each sector. Numerical integration may seem to be a good solution as for each sector we have to integrate over only two dimensions. However, when using numerical integration it is not possible to concentrate the likelihood for the sector-specific parameters. We would then have to maximize the complex likelihood function over a very large number of parameters. Moreover, the calculation of the likelihood function will probably be very time-consuming due to the numerical integration.

For the above-mentioned reasons, we will instead use simulation to approximate the likelihood, see Gourieroux and Monfort (1993), Lee (1995) and Hajivassiliou and Ruud (1994) for discussions of Simulated Maximum Likelihood [SML] estimation. The resulting simulated likelihood is in turn concentrated with respect to the sector-specific parameters. Finally, we numerically optimize the concentrated simulated likelihood over $\boldsymbol{\delta}$ and $\boldsymbol{\Sigma}_\eta$. We end up with an estimation routine that can be labeled as "Concentrated Simulated Maximum Likelihood".

3.1 Concentrated Simulated Maximum Likelihood

We approximate the likelihood contribution of sector i by simulation

$$\tilde{\mathcal{L}}_i = \frac{1}{L}\sum_{l=1}^{L}\prod_t \phi(e_{i,t}(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\Sigma}_\eta^{1/2}\tilde{\boldsymbol{\eta}}_{i,l}); 0, \sigma_i^2) \tag{11}$$

where $\tilde{\boldsymbol{\eta}}_{i,l}$, $l = 1, \ldots, L$ denotes a draw from $N(0, \mathbf{I})$ and where $\boldsymbol{\Sigma}_\eta^{1/2}$ denotes the Choleski decomposition of $\boldsymbol{\Sigma}_\eta$ such that $\boldsymbol{\Sigma}_\eta^{1/2}\tilde{\boldsymbol{\eta}}_{i,l} \sim N(0, \boldsymbol{\Sigma}_\eta)$. Next, we concentrate (11) with respect to $\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i$ and σ_i^2. To simplify the notation, we denote the error associated with the l-th draw as $e_{i,t,l} \equiv e_{i,t}(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\Sigma}_\eta^{1/2}\tilde{\boldsymbol{\eta}}_{i,l})$. Given $\boldsymbol{\delta}$ and $\boldsymbol{\Sigma}_\eta$, the objective is to find the optimum of the (simulated) likelihood contribution of sector i with respect to $\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i$ and σ_i^2. That is, we need to solve

$$\max_{\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \sigma_i^2} \frac{1}{L}\sum_{l=1}^{L}\prod_t \phi(e_{i,t,l}; 0, \sigma_i^2) \tag{12}$$

which is equivalent to

$$\max_{\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \sigma_i^2} \frac{1}{L}\sum_l \exp\left(\sum_t \log\phi(e_{i,t,l}; 0, \sigma_i^2)\right) \tag{13}$$

The first order conditions for (13) can be compactly written as

$$\frac{1}{L}\sum_l\sum_t w_{i,l}\frac{e_{i,t,l}}{\sigma_i^2}\mathbf{x}'_{i,t,l} = \mathbf{0}_{2P+1}, \tag{14}$$

$$\frac{1}{L}\sum_l\sum_t \frac{w_{i,l}}{2\sigma_i^2}(\frac{e_{i,t,l}^2}{\sigma_i^2} - 1) = 0, \tag{15}$$

where $\mathbf{0}_n$ denotes a $n \times 1$ vector of zeros and

$$\mathbf{x}_{i,t,l} = ([1, G(s_t; \gamma_i, \tau_i)] \otimes [1, y_{i,t-1}, \ldots, y_{i,t-p}]),$$
$$w_{i,l} = \prod_t \phi(e_{i,t,l}; 0, \sigma_i^2).$$

For numerical stability, we compute $w_{i,l}$ as $w_{i,l} = \exp\left(\sum_t \log \phi(e_{i,t,l}; 0, \sigma_i^2)\right)$ and, as only the relative size of $w_{i,l}$ is important, we can use $w_{i,l} = \exp\left(\sum_t \log \phi(e_{i,t,l}; 0, \sigma_i^2) + \kappa\right)$, where κ is a fixed arbitrary constant, which can be chosen such that we avoid taking the exponent of a very small number.

Loosely speaking, $w_{i,l}$ gives the relative importance of the l-th simulation draw. As $w_{i,l}$ depends on $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$ and σ_i^2 a closed-form solution of the problem in (14)-(15) is difficult to find. However, it turns out that it can quite easily be solved iteratively. Given starting values for $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$ and σ_i^2, we first calculate $w_{i,l}$. If we ignore the dependence of $w_{i,l}$ on the parameters, the first order conditions (14)-(15) become a special case of a weighted least squares [WLS] problem. Denoting $\mathbf{x}_{i,l} = (\mathbf{x}'_{i,1,l}, \ldots, \mathbf{x}'_{i,T,l})'$ and $\mathbf{y}_i = (y_{i,1}, \ldots, y_{i,T})'$, we obtain

$$\begin{pmatrix} \hat{\boldsymbol{\alpha}}_i \\ \hat{\boldsymbol{\beta}}_i \end{pmatrix} = \left(\frac{1}{L} \sum_{l=1}^{L} w_{i,l} \mathbf{x}'_{i,l} \mathbf{x}_{i,l} \right)^{-1} \left(\frac{1}{L} \sum_{l=1}^{L} w_{i,l} \mathbf{x}'_{i,l} \mathbf{y}_i \right) \tag{16}$$
$$\hat{\sigma}_i^2 = \frac{\frac{1}{L} \sum_l \sum_t w_{i,l} e_{i,t,l}}{\frac{T}{L} \sum_l w_{i,l}}. \tag{17}$$

Given these estimates we can update the weights $w_{i,l}, l = 1, \ldots, L$. This WLS procedure is then iterated until convergence of the sector-specific parameters. At this point, we have found the optimal parameters $\hat{\boldsymbol{\alpha}}_i$, $\hat{\boldsymbol{\beta}}_i$ and $\hat{\sigma}_i^2$ given values for $\boldsymbol{\delta}$ and $\boldsymbol{\Sigma}_\eta$.

Finally, to obtain estimates of $\boldsymbol{\delta}$ and $\boldsymbol{\Sigma}_\eta$, we maximize the concentrated likelihood over these parameters. For each evaluation of the concentrated likelihood, we repeat the WLS procedure described above to obtain conditional estimates of $\hat{\boldsymbol{\alpha}}_i$, $\hat{\boldsymbol{\beta}}_i$ and $\hat{\sigma}_i^2$. However, in the numerical optimization routine the parameter values in consecutive iterations tend to be close together so that only few iterations will be necessary to achieve convergence.

Under the usual regularity conditions, the Simulated Maximum Likelihood estimator is consistent for $N \to \infty$ and $L \to \infty$, see Hajivassiliou and Ruud (1994). Furthermore, the SML estimator is asymptotically efficient and normally distributed, where the asymptotic covariance matrix is equal to the inverse of the information matrix. Note that, to estimate the variance of the parameters of the switching function, we only have to calculate the Hessian of the concentrated likelihood, see Davidson and MacKinnon (1993). The estimated covariance matrix equals

$$\widehat{\text{Var}(\vartheta)} = \left(\frac{\partial^2 \mathcal{L}^c}{\partial\vartheta'\partial\vartheta'}\right)^{-1} \tag{18}$$

where ϑ contains the parameters in δ and Σ_η and where $\mathcal{L}^c$ denotes the concentrated likelihood function.

3.2 Conditional inference

Often one is interested in the value of the switching function $G(s_t; \gamma_i, \tau_i)$ for a particular sector at a point in time given the sector characteristics $\mathbf{w}_i$. For this purpose, in our multi-level panel data model we can use the unconditional expectation of $G(s_t; \gamma_i, \tau_i)$ assuming the model parameters are known and fixed. This expectation equals $\mathsf{E}_{\boldsymbol{\eta}_i}[G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i)]$. On the other hand, we can also consider the conditional expectation $\mathsf{E}_{\boldsymbol{\eta}_i}[G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i)|\mathbf{y}_i]$. In this conditional expectation, the expected level of the switching function is calculated conditional on the observed time series. Hence, the conditional expectation of $G(s_t; \gamma_i, \tau_i)$ can differ across sectors even when they have equal observed characteristics $\mathbf{w}_i$.

The unconditional expectation of $G(s_t; \gamma_i, \tau_i)$ can easily be calculated using simulation as follows,

$$\begin{aligned}\mathsf{E}_{\boldsymbol{\eta}_i}[G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i)] &= \int_{\boldsymbol{\eta}_i} G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i)\phi(\boldsymbol{\eta}_i; 0, \boldsymbol{\Sigma}_\eta)d\boldsymbol{\eta}_i \\ &= \frac{1}{L}\sum_l G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_{i,l}),\end{aligned} \tag{19}$$

where $\boldsymbol{\eta}_{i,l}$ is a draw from $N(0, \boldsymbol{\Sigma}_\eta)$.

The conditional expectation of $G(s_t; \gamma_i, \tau_i)$ is more difficult to compute, as one needs to consider

$$\begin{aligned}
\mathsf{E}_{\boldsymbol{\eta}_i}[G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i)|\mathbf{y}_i] &= \int_{\boldsymbol{\eta}_i} G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i) g(\boldsymbol{\eta}_i|\mathbf{y}_i) d\boldsymbol{\eta}_i \\
&= \frac{\int_{\boldsymbol{\eta}_i} G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_i) g(\mathbf{y}_i|\boldsymbol{\eta}_i) \phi(\boldsymbol{\eta}_i; \mathbf{0}, \boldsymbol{\Sigma}_\eta) d\boldsymbol{\eta}_i}{\int_{\boldsymbol{\eta}_i} g(\mathbf{y}_i|\boldsymbol{\eta}_i) \phi(\boldsymbol{\eta}_i; \mathbf{0}, \boldsymbol{\Sigma}_\eta) d\boldsymbol{\eta}_i} \\
&= \frac{\frac{1}{L}\sum_l G(s_t; \boldsymbol{\delta}'\mathbf{w}_i + \boldsymbol{\eta}_{i,l}) w_{i,l}}{\frac{1}{L}\sum_l w_{i,l}}, \qquad (20)
\end{aligned}$$

with $w_{i,l}$ the weights as defined earlier and where, in general, $g(\mathbf{x}|\mathbf{z})$ denotes the density function of $\mathbf{x}$ given $\mathbf{z}$.

4 US sectoral production

We consider quarterly growth rates in industrial production for 18 main manufacturing sectors at the three-digit level[1] in the new North American Industry Classification System (NAICS) over the period 1972Q1-2002Q4.[2] The sectors are listed in Table 1. In addition, we obtain industry characteristics using the NBER-CES Manufacturing Industry Database, containing data on output, employment, payroll and other input costs, investment, capital stocks, TFP, and various industry-specific price indexes.[3] Finally, we decide to use the term spread, computed as the difference between the 10-year Treasury bond rate and the 3-month T-bill rate, lagged two quarters as transition variable s_t in all STAR models reported below. This choice is motivated by the general finding that the term spread is among the most powerful US business cycle indicators, see Estrella and Mishkin (1998), among many others.

[1] We exclude the sectors "Wood products", "Computers and electronic products" from the analysis.

[2] See `http://www.census.gov/epcd/www/naics.html` for more information about NAICS and its relation to the US Standard Industrial Classification (SIC).

[3] See `http://www.nber.org/nberces/nbprod96.htm` and Bartelsman and Gray (1996) for detailed information.

We start with estimating univariate STAR models for each of the 18 sectors, which can be interpreted as components in a fully heterogeneous panel STAR model (2)-(3) with no cross-equation restrictions on the parameters γ_i and τ_i. For simplicity, we impose an identical autoregressive order P in the STAR models for all sectors. We set $P = 4$ as this is the value preferred by conventional information criteria such as the Akaike and Schwarz criteria for most sectors.

It turns out that it is extremely difficult to obtain reliable parameter estimates for some sector-specific STAR models. Figure 1 shows the surfaces of the (concentrated) log-likelihood function for the sectors "Paper", "Food, beverage and tobacco", "Electrical equipment" and "Chemical" as functions of $\log(\gamma_i)$ and τ_i. These graphs reveal the reason for these difficulties in estimation: the log-likelihood is essentially flat in the direction of $\log(\gamma_i)$, and especially for the sectors "Paper" and "Food, beverage and tobacco" a maximum is not well-defined. This illustrates the point made in Section 2, that univariate STAR models may be difficult to estimate, and (partial) pooling may be necessary. Detailed estimation results of the univariate STAR models are not shown here to save space, but these are available upon request from the corresponding author.

Our model selection strategy now proceeds as follows. First we test whether the restriction $\tau_i = \tau$ or $\gamma_i = \gamma$ can be imposed. To this end we compare the likelihood values of the sector-specific model to the likelihood of a model in which one of the two restrictions is imposed. A difficulty in this testing is that, as mentioned earlier, sector-specific estimates can not be obtained for some sectors. Proper Likelihood Ratio testing is therefore not possible. Instead, we apply these tests in a more informal manner to see whether there is evidence of differences in switching parameters or not. For the sectors for which proper estimates can not be found we use the maximum of likelihood function over a grid of reasonable parameter values. The p-values of the tests should therefore not be taken literally. We find that the restriction $\tau_i = \tau$

cannot be maintained (p-value = 0.0000). There is however no strong evidence that the γ parameter differs across sectors (p-value = 0.120). Next we estimate our two-level model where we only allow the thresholds (τ_i) to differ across sectors. Among the different sector characteristics we explore as possible explanatory variables in the level-2 model (6), total 5-factor productivity (TFP5) was found to be the most discriminatory. This model can be formally tested against a model in which both switching parameters are pooled. This restriction corresponds to setting and σ_η^2 and the parameter corresponding to the factor productivity to zero. This restriction is rejected at the 5%-level (p-value 0.0432).[4] Hence, our preferred model is a "partial" multi-level panel STAR model, where the slope parameter in the transition functions is pooled, while the threshold variable can differ across sectors according to the value of TFP5.

Tables 1 and 2 give the parameter estimates of the preferred model. Of particular interest, of course, are the parameter estimates of the level-2 model for τ_i, which show that the threshold between the two regimes varies positively with total factor productivity. Given that recessions are associated with small values of the term spread and, hence, values of the logistic function close to 0, this shows that sectors with higher total factor productivity are likely to enter recessions faster. The second column of Table 1 shows that this is the case for the sectors "Machinery", "Aerospace", "Primary metal" and "Fabricated metal products". These sectors appear to be particularly sensitive to common cyclical movements and might be used to signal the onset of nationwide recessions.

This point is further illustrated by Figure 2, showing the conditional expectation of the transition functions $G(s_t; \gamma_i; \tau_i)$ for all sectors in the multi-level panel STAR model. In fact, these plots reveal several interesting findings. First, note that the

[4]As the restriction on the variance parameter has an one-sided alternative ($\sigma_\eta^2 = 0$ versus $\sigma_\eta^2 > 0$) the critical value of the joint test is not standard. The distribution of the test statistic is $\frac{1}{2}\chi^2(1) + \frac{1}{2}\chi^2(2)$ instead of $\chi^2(2)$, see Wolak (1989, pp. 19-20)

regimes $G(s_t;\gamma_i;\tau_i) = 0$ and $G(s_t;\gamma_i;\tau_i) = 1$ correspond quite closely with business cycle recessions and expansions, respectively, as dated by the NBER. Second, these graphs show that the four aforementioned sectors generally enter recessions slightly before the other sectors. Third, only the sectors Primary metal, Aerospace, and Apparel and leather appear to experience a regime switch around 1990. Apparently, this shallow final recession of the previous century did not affect all sectors in the economy equally. This last point also illustrates the potential use of the conditional expectation of $G(s_t;\gamma_i,\tau_i)$ given the time series $\mathbf{y}_i$ for inference. Even though the expected value of the threshold τ_i for "Apparel and leather" is close to the sector-average (see the entries in Table 1), the conditional expectation of the transition function as shown in Figure $G(s_t;\gamma_i;\tau_i) = 0$ shows that this sector did experience a regime switch around 1990 not shared by other sectors with comparable values of total factor productivity such as "Motor vehicles and parts" and "Paper" (for which the expected value of τ_i is close to -1.10 as well).

5 Conclusion

In this paper we have developed a multi-level panel smooth transition autoregressive model, which is capable of describing joint regime-switching behavior in time series variables while allowing for heterogeneity in the exact timing of the regime changes. This makes the model particularly useful for situations where the non-linear dynamics are driven by a common regime-switching component, but where the response to this component can be different across variables. For example, probably all sectors in an economy are affected by nationwide recessions, but some sectors may enter into (and get out off) recessions earlier than others. By using a second-level model for the parameters that appear in the regime-switching mechanism, the model is kept relatively parsimonious.

Our application of the model to US sector-level industrial production demonstrates its potential. We see that fully pooling is not effective in understanding possible common non-linear features across the series, and that a fully heterogeneous model introduces estimation problems for 4 out of the 18 series. Hence, our restrictions due to the second-level regression model for the parameters in the individual regime-switching functions not only allows for interpretability but it also facilitates empirical analysis.

Several topics for further research are worth considering. Further applications of the model to, for example, state-level output or employment are necessary to establish its usefulness. The model may also be used to address the question whether forecasts of aggregate industrial production constructed from the panel model are more accurate than those obtained from a univariate (non-linear) model for the aggregate, as documented for example by Lee (1997) in the case of linear models. Finally, the model could be extended to incorporate idiosyncratic cyclical components.

References

Anderson, H.M. and F. Vahid (1998), Testing multiple equation systems for common nonlinear components, *Journal of Econometrics* **84**, 1–36.

Bartelsman, E.J. and W. Gray (1996), The NBER manufacturing productivity database, NBER Technical Working Paper No. 205.

Bidarkota, P.V. (1999), Sectoral investigation of asymmetries in the conditional mean dynamics of the real US GDP, *Studies in Nonlinear Dynamics and Econometrics* **3**, 191–200.

Clark, T.E. (1998), Employment fluctuations in US regions and industries: the roles of national, region-specific and industry-specific shocks, *Journal of Labor Economics* **16**, 202–229.

Cooper, S.J. (1998), Multiple regimes in US output fluctuations, *Journal of Business and Economic Statistics* **16**, 92–100.

Davidson, R. and J.G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.

Del Negro, M. (2002), Asymmetric shocks among U.S. states, *Journal of International Economics* **56**, 273–297.

Engle, R.F. and J.V. Issler (1995), Estimating common sectoral cycles, *Journal of Mone-*

tary Economics **35**, 83–113.

Estrella, A. and F.S. Mishkin (1998), Predicting U.S. recessions: financial variables as leading indicators, *Review of Economics and Statistics* **80**, 45–61.

Forni, M. and L. Reichlin (1998), Let's get real: A factor analytical approach to disaggregated business cycle dynamics, *Review of Economic Studies* **65**, 453–473.

Franses, P.H. and D. van Dijk (2000), *Nonlinear Time Series Models in Empirical Finance*, Cambridge: Cambridge University Press.

Gourieroux, C. and A. Monfort (1993), Simulation-based Inference: A survey with special reference to panel data models, *Journal of Econometrics* **59**, 5–33.

Granger, C.W.J. and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships*, Oxford: Oxford University Press.

Hajivassiliou, V.A and P.A. Ruud (1994), Classical estimation methods for LDV models using simulation, in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics IV*, Amsterdam: North-Holland, pp. 2383–2411.

Lee, L. (1995), Asymptotic bias in simulated maximum likelihood estimation of discrete choice models, *Econometric Theory* **11**, 437–483.

Lee, K.C. (1997), Modelling economic growth in the UK: An econometric case for disaggregated sectoral analysis, *Economic Modelling* **14**, 369–394.

Lundbergh, S. and T. Teräsvirta (1998), Modelling economic high-frequency time series with STAR-GARCH models, Working Paper Series in Economics and Finance No. 291, Stockholm School of Economics.

Owyang, M.T., J. Piger, and H.J. Wall (2003), Business cycle phases in U.S. states, Federal Reserve Bank of St. Louis Working Paper No. 2003-011C.

Paap, R., P.H. Franses and D. van Dijk (2003), Does Africa grow slower than Asia and Latin America?, Econometric Institute Research Report 2003-7/A.

Pesaran, M.H., R.G. Pierse and K.C. Lee (1993), Persistence, cointegration, and aggregation - A disaggregated analysis of output fluctuations in the United States' economy, *Journal of Econometrics* **56**, 57–88.

Quah, D.T. (1996), Aggregate and regional disaggregate economic fluctuations, *Empirical Economics* **21**, 137–159.

Teräsvirta, T. (1994), Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**, 208–218.

Teräsvirta, T. (1998), Modelling economic relationships with smooth transition regressions, in A. Ullah and D.E.A. Giles (eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, pp. 507–552.

Vahid, F. and R.F. Engle (1993), Common trends and common cycles, *Journal of Applied Econometrics* **8**, 341–360.

Vahid, F. and R.F. Engle (1997), Codependent cycles, *Journal of Econometrics* **80**, 199–221.

van Dijk, D., P.H. Franses and A. Lucas (1999), Testing for smooth transition nonlinearity in the presence of additive outliers, *Journal of Business and Economic Statistics* **17**, 217–235.

van Dijk, D., T. Teräsvirta and P.H. Franses (2002), Smooth transition autoregressive models - a survey of recent developments, *Econometric Reviews* **21**, 1–47.

Wolak, F.A. (1989), Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models, *Econometric Theory*, **5**, 1–35.

Table 1: Estimates of parameters in panel STAR model for quarterly growth rates of US sectoral production

Sector	$\mathsf{E}[\log \gamma_i]$	$\mathsf{E}[\tau_i]$	$\alpha_{i,0}$	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,3}$	$\alpha_{i,4}$	$\beta_{i,0}$	$\beta_{i,1}$	$\beta_{i,2}$	$\beta_{i,3}$	$\beta_{i,4}$	$\log(\sigma_i^2)$
Nonmetallic mineral products	4.61	-1.13	-2.44	0.06	0.52	0.88	-0.18	3.06	0.13	-0.55	-0.79	0.07	1.15
Primary metal	4.61	-1.06	-1.48	0.89	-0.03	0.10	-0.26	1.98	-0.81	0.05	-0.14	0.16	2.85
Fabricated metal products	4.61	-1.07	-1.41	0.80	0.35	0.20	-0.05	1.84	-0.27	-0.52	-0.03	-0.14	0.58
Machinery	4.61	-1.01	-2.46	0.39	0.96	0.30	-0.18	2.72	0.36	-1.04	-0.27	-0.01	1.12
Electrical equipments	4.61	-1.13	-2.75	0.71	0.26	1.55	-0.48	3.29	-0.27	-0.34	-1.41	0.26	1.16
Motor vehicles and parts	4.61	-1.08	-7.29	-0.03	-0.15	0.00	-0.08	9.30	0.00	-0.08	0.04	0.08	3.36
Aerospace and other misc. transportation	4.61	-1.05	-0.99	0.20	0.00	0.31	0.11	1.10	0.23	0.12	-0.13	-0.20	1.42
Furniture and related products	4.61	-1.13	-1.41	0.76	-0.06	-0.06	0.06	2.09	-0.39	-0.01	0.11	-0.14	1.22
Miscellaneous	4.61	-1.10	-0.61	0.94	-0.93	0.86	-0.64	1.49	-0.65	0.85	-0.91	0.52	0.52
Food, beverage, and tobacco	4.61	-1.08	0.00	0.56	-0.33	0.00	-0.47	0.47	-0.55	0.35	0.03	0.20	-0.16
Textile and product mills	4.61	-1.17	-2.43	0.35	0.17	0.20	-0.62	2.83	0.18	-0.57	0.01	0.36	1.48
Apparel and leather	4.61	-1.10	-0.87	0.35	-0.63	1.04	-0.86	0.69	0.25	0.52	-1.33	0.99	0.80
Paper	4.61	-1.10	-0.37	0.59	-0.06	-0.29	-0.71	0.96	-0.42	-0.02	0.09	0.65	1.17
Printing and support	4.61	-1.13	-0.54	0.58	0.05	0.49	-0.07	1.13	-0.41	0.15	-0.40	-0.09	0.59
Petroleum and coal products	4.61	-1.07	-2.05	0.17	0.57	0.27	-0.10	2.66	-0.32	-0.46	-0.37	-0.10	1.45
Chemical	4.61	-1.17	-0.90	0.86	-0.15	-0.05	-0.33	1.64	-0.52	0.04	0.05	0.25	0.61
Plastics and rubber products	4.61	-1.20	-2.46	0.34	0.61	0.09	0.12	4.02	-0.18	-0.90	-0.06	-0.26	1.88
Other manufacturing (non-NAICS)	4.61	-1.13	-0.18	0.47	0.44	-0.03	0.00	0.38	-0.11	-0.48	0.17	0.07	0.26

Notes: The table reports concentrated SML estimates of the multi-level panel STAR model (5)-(8), for quarterly growth rates in 18 manufacturing sectors for the period 1972Q1-2002Q4, where the term spread lagged two quarters is used as transition variable, total (5-)factor productivity is used as regressor in the level-2 model for the threshold parameter τ_i, and the slope parameter γ_i is pooled across sectors.

Table 2: Estimates of parameters in level-2 model in panel STAR model for quarterly growth rates of US sectoral production

	Intercept	TFP5	$\sqrt{\text{diag}\Sigma_\eta}$
$\log \gamma_i$	4.607	0	0
	(2.198)	–	–
τ_i	-2.041	0.973	0.1097
	(0.690)	(0.714)	(0.0375)

Notes: The table reports concentrated SML estimates of the level-2 model (6) in the panel STAR model for quarterly growth rates in 18 manufacturing sectors for the period 1972Q1-2002Q4, where the term spread lagged two quarters is used as transition variable. Total (5-)factor productivity (TFP5) is used as regressor in the level-2 model for the threshold parameter τ_i, and the slope parameter γ_i is pooled across sectors.

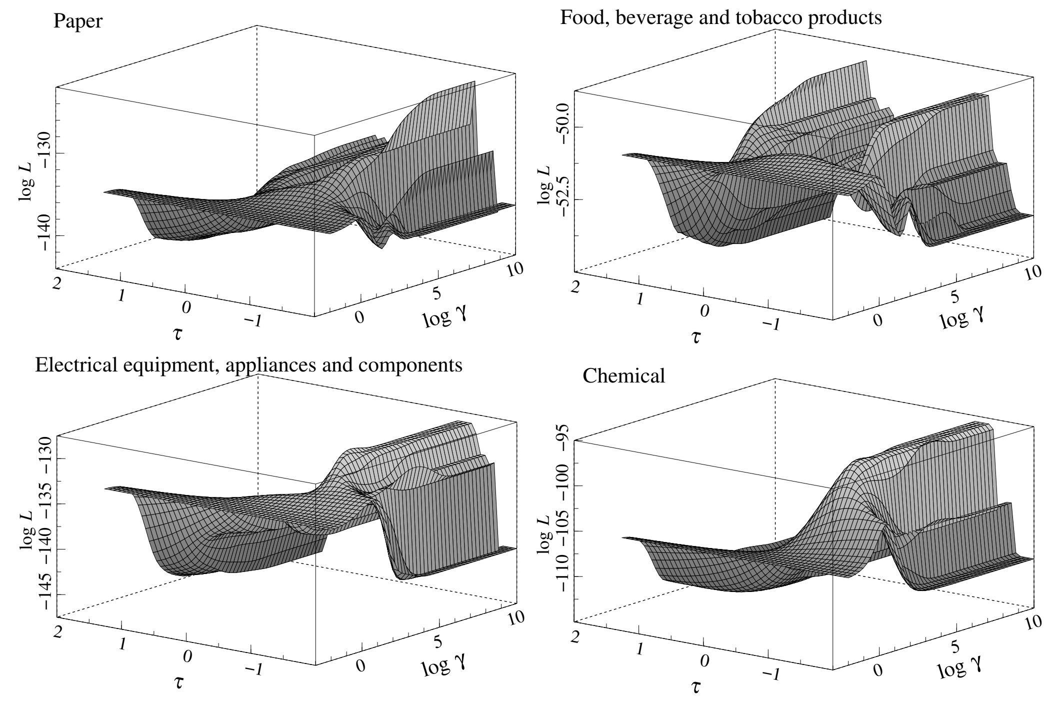

Figure 1: Concentrated log likelihood surfaces for sector-specific STAR models.

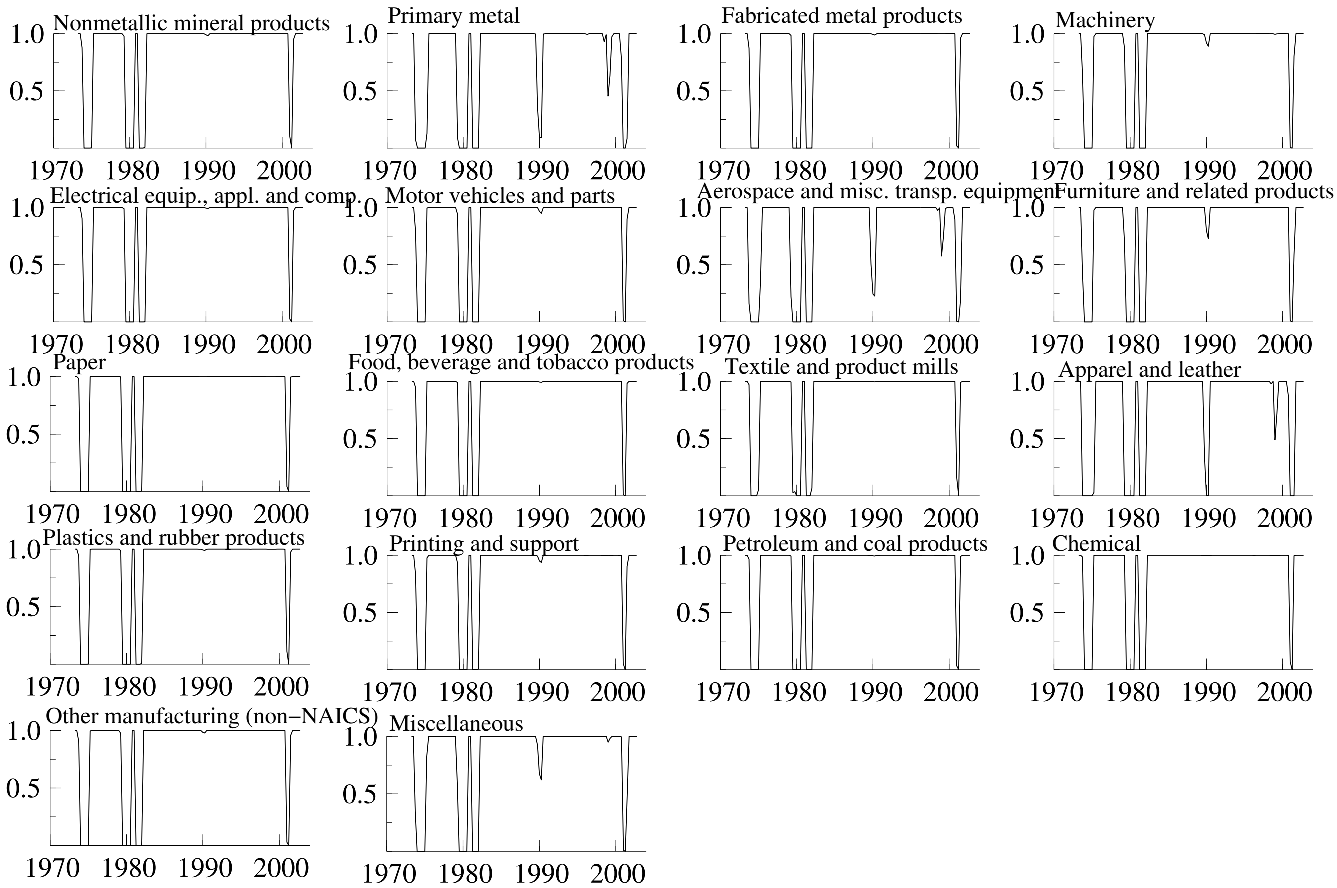

Figure 2: Conditional expectation of switching function.