Parameter Regimes in Partially Functional Linear Regression for Panel Data

Fabian Walders and Dominik Liebl

Abstract We introduce a novel semiparametric partially functional linear regression model for panel data. The parametric model part is completely time varying, whereas the functional non-parametric component is allowed to vary over a set of different (functional) parameter regimes. These parameter regimes are assumed latent and need to be estimated from the data additionally to the unknown model parameters. We develop asymptotic theory for the suggested estimators including rates of convergence as $n, T \to \infty$. Our statistical model is motivated from economic theory on asset pricing. It allows to identify different risk regimes, governing the pricing of idiosyncratic risk in stock markets. For our application we develop necessary theoretical ground and offer a vast empirical study based on high-frequency stock-level data for the S&P 500 Index.

1 Introduction

This work contributes a novel semiparametric regression model for panel data. The suggested approach allows a scalar response to be affected by a random function as well as by real-valued predictors in a time varying manner. The nonparametric functional parameter changes over time by switching between different parameter regimes which have to be estimated from the data. The real-valued parametric parameters are allowed to vary over time independently from the regimes. Estimation relies on estimators from the functional data literature and a recent nonparametric classification strategy that allows to identify the different parameter regimes. We develop asymptotic theory for the estimators as n and T diverge simultaneously. Given a set of standard assumptions, we prove classification consistency and derive rates of convergence.

Fabian Walders, University of Bonn, BGSE & Institute for Financial Economics and Statistics, Bonn, Germany. fwalders@uni-bonn.de · Dominik Liebl, University of Bonn, Institute for Financial Economics and Statistics, Bonn, Germany. dliebl@uni-bonn.de

Our approach relates to the recent literature on limited parameter instabilities in panel data models. For instance, [3] consider regime varying parameters in a linear panel model with unknown regime classes. Further important works are those of [9] and [11] which postulate group structures governing dissimilarities between regression functions in the cross section. Another important reference is the contribution of [8] who introduce the partially functional regression model in a cross section context. Our work can be understood as an extension of that work to a panel data context with additional unknown parameter regimes. A second important reference in the functional data literature is the work of [6], introducing parameter instabilities in classical functional linear regression.

The specific form of our suggested statistical model is motivated by relevant economic theory. It is particularly well suited to identify pricing regimes of idiosyncratic risk in stock markets. In our application we offer a combination of economic and statistical theory. Beyond that we provide a vast empirical study for the US stock market examining the idiosyncratic volatility puzzle (see, e.g., [2]).

The remainder of this paper is structured as follows. In Sections 2 and 3 we introduce the model and present the estimation procedure. In Section 4 we present the asymptotic theory, including rates of convergence. Subsequently, we provide economic theory for the application as well as a glance at the corresponding empirical work in Section 5. A last Section 6 briefly concludes.

2 Model

We suggest a linear regression model which is formally obtained as a partially functional regression in a panel data context. Generically, the task is to model the effect of a square integrable random function $X_{it} \in L^2[0,1]$ on a scalar response y_{it} in the presence of a real-valued random variable z_{it}. Indexing the cross section and time dimensions $1 \leq i \leq N$ and $1 \leq t \leq T$ respectively, the ultimate statistical model reads as

$$y_{it} = \alpha_{0,t} + \int_0^1 \alpha_t(s) X_{it}(s) \mathrm{d}s + \beta_t z_{it} + \varepsilon_{it}, \tag{1}$$

with $\alpha_{0,t}$ being a t-specific intercept parameter and ε_{it} being random disturbance. The unknown functional parameters $\alpha_t \in L^2[0,1]$ differ across K different time regimes $G_1,\ldots,G_K$ which form a latent partition of the index set $\{1,\ldots,T\}$. These time regimes are mutually exclusive, i.e., $G_k \cap G_l = \emptyset$ for all $k,l \in \{1,\ldots,K\}$, form a complete partition, i.e., $\bigcup_{k=1}^K G_k = \{1,\ldots,T\}$, and may consist of non-adjacent time points t. Each regime G_k is associated with a square integrable parameter function $A_k \in L^2[0,1]$ governing the effect of X_{it} on y_{it}, i.e.,

$$\alpha_t = A_k \text{ if } t \in G_k.$$

Model (1) compromises two extreme specifications. On the one hand it might be the case that $K = 1$ and hence $G_1 = \{1,\dots,T\}$. In this situation the effect of the random functions X_{it} on the response y_{it} is time invariant. On the other hand the model nests complete heterogeneity if $K = T$ and all regimes G_k are singletons which turns (1) into a collection of T different cross section models. Further purely functional or purely parametric specifications are possible in case $\beta_t = 0$ or $\alpha_t = 0$ for all $1 \le t \le T$ respectively.

3 Estimation

The objective is to estimate the parameters A_k, β_t and $\alpha_{0,t}$ as well as the regimes $G_1,\dots,G_K$ from realizations of the random variables $\{(y_{it}, X_{it}, z_{it}) : 1 \le i \le n, 1 \le t \le T\}$. Without loss of generality all variables are assumed to be centered (see Assumption 1). We suggest a four step estimation procedure which comprises an initial auxiliary estimation (Step 1) that serves as a basis for the subsequent classification (Step 2) of the regression functions. The two final estimation steps (Step 3 and 4) re-estimate the model parameters borrowing strength from the identified regime structure. In the following we give a detailed description:

Step 1 Estimate the parameters α_t, β_t and $\alpha_{0,t}$ separately for each $1 \le t \le T$ using the (modified) estimators in [8] as well as within-t-averages. For the intercepts simply employ the estimators $\hat{\alpha}_{0,t} = n^{-1}\sum_{i=1}^n y_{it}$ and use these in turn to form centered responses $\check{y}_{it} = y_{it} - \hat{\alpha}_{0,t}$. The estimators for α_t and β_t are based on cross section estimates of the relevant covariance objects. Consider for a fixed t the cross section $\{(y_{it}, X_{it}, z_{it}) : 1 \le i \le n\}$. For each t, the empirical covariance operator of the $\{X_{it} : 1 \le i \le n\}$ is the integral operator $\hat{\Gamma}_t : L^2[0,1] \to L^2[0,1]$ defined via its kernel, namely, the empirical covariance function $\hat{K}_{X,t}(u,v) := n^{-1}\sum_{i=1}^n X_{it}(u)X_{it}(v)$. Let $(\hat{\lambda}_{1,t}, \hat{\phi}_{1,t}),\dots,(\hat{\lambda}_{n,t}, \hat{\phi}_{n,t})$ denote the eigenvalue-eigenfunction pairs of $\hat{\Gamma}_t$ ordered according to $\hat{\lambda}_{1,t} \ge \dots \ge \hat{\lambda}_{n,t}$ and $\langle \cdot,\cdot \rangle$ the inner product in $L^2[0,1]$. In a similar t-wise fashion we define the following estimators:

$$\hat{K}_{zx,t}(s) := n^{-1}\sum_{i=1}^n z_{it}X_{it}(s), \qquad \hat{K}_{yx,t}(s) := n^{-1}\sum_{i=1}^n \check{y}_{it}X_{it}(s),$$
$$\hat{K}_{z,t}(s) := n^{-1}\sum_{i=1}^n z_{it}^2, \qquad \hat{\Phi}_t(g) := \sum_{j=1}^m \frac{\langle \hat{K}_{zx,t}, \hat{\phi}_{j,t}\rangle\langle \hat{\phi}_{j,t}, g\rangle}{\hat{\lambda}_{j,t}} \text{ for } g \in L^2[0,1].$$

Given a truncation parameter m, with $1 \le m < n$, one obtains least squares estimators for (a^*_{jt}) $1 \le j \le m$ and β_t in the approximate empirical model $\check{y}_{it} \approx \sum_{j=1}^m \langle X_{it}, \hat{\phi}_{j,t}\rangle a^*_{jt} + z_{it}\beta_t + \varepsilon_{it}$. These estimators read as

$$\hat{\beta}_t = \left[\hat{K}_{z,t} - \hat{\Phi}_t\left(\hat{K}_{zx,t}\right)\right]^{-1}\left[\hat{K}_{zy,t} - \hat{\Phi}_t\left(\hat{K}_{xy,t}\right)\right]$$

$$\hat{a}_{j,t} = \hat{\lambda}_{j,t}^{-1}\frac{1}{n}\sum_{i=1}^{n}\langle X_{it}, \hat{\phi}_{j,t}\rangle(\check{y}_{it} - \hat{\beta}_t z_{it}), \quad 1 \le j \le m,$$

where the final estimate for α_t is obtained as $\hat{\alpha}_t = \sum_{j=1}^{m}\hat{a}_{j,t}\hat{\phi}_{j,t}$. For the next step we use the scaled estimator $\hat{\alpha}_t^{(\Delta)} := \sum_{j=1}^{m}\frac{\hat{\lambda}_{j,t}^{1/2}}{\hat{\sigma}_{t,\varepsilon}}\hat{a}_{j,t}\hat{\phi}_{j,t}$, where $\hat{\sigma}_{t,\varepsilon}^2$ is the empirical residual variance from the above regression for time t.

Step 2 Classify the t-specific regression functions into K regimes using a thresholding procedure along the lines of [9]. To do so, select a subset $S \subset \{1,\dots,T\}$ and some $s \in S$. Suppose s is in some regime G_k. The squared L^2 distances of the modified estimators $\hat{\alpha}_t^{(\Delta)}$ from other periods $t \in S$ to $\hat{\alpha}_s^{(\Delta)}$ are denoted as $\hat{\Delta}_{ts} := ||\hat{\alpha}_t^{(\Delta)} - \hat{\alpha}_s^{(\Delta)}||_2^2$. The corresponding order statistics $\hat{\Delta}_{t(1)} \le \hat{\Delta}_{t(2)} \le \cdots \le \hat{\Delta}_{t(|S|)}$ now form the basis for classification.

Given a pre-selected threshold τ_{nT}, see discussion below, define the regime estimate $\hat{G}_k$ as $\{(1),\dots,(\hat{p})\}$, where $(\hat{p})$ is defined according to $\hat{\Delta}_{t(\hat{p})} \le \tau_{nT} < \hat{\Delta}_{t(\hat{p}+1)}$. Iteratively proceed with this procedure for the remaining points in time, i.e., for $t \in \{1,\dots,\mathrm{T}\} \setminus \hat{G}_k$. This naturally provides an estimate $\hat{K}$.

Step 3 Re-estimate the t-specific parameters β_t by borrowing strength from the identified regime structure. For each $1 \le k \le \hat{K}$, the regime specific empirical covariance operator $\tilde{\Gamma}_k$ of the $\{X_{it} : 1 \le i \le n,\ t \in \hat{G}_k\}$ is defined via its integral kernel $\tilde{K}_{X,k}(u,v) := \left(n|\hat{G}_k|\right)^{-1}\sum_{i=1}^{n}\sum_{t\in\hat{G}_k}X_{it}(u)X_{it}(v)$. The eigenfunction-eigenvalue-pairs of this operator are denoted by $(\tilde{\lambda}_{j,k}, \tilde{\phi}_{j,k})$, $1 \le j \le n|\hat{G}_k|$ ordered according to $\tilde{\lambda}_{1,k} \ge \cdots \ge \tilde{\lambda}_{n|\hat{G}_k|,k} \ge 0$. Further, necessary regime specific quantities are obtained according to

$$\tilde{K}_{zx,k}(s) := \left(n|\hat{G}_k|\right)^{-1}\sum_{i=1}^{n}\sum_{t\in\hat{G}_k}z_{it}X_{it}(s), \qquad \tilde{K}_{z,k}(s) := \left(n|\hat{G}_k|\right)^{-1}\sum_{i=1}^{n}\sum_{t\in\hat{G}_k}z_{it}^2,$$

$$\tilde{\Phi}_k(g) := \sum_{j=1}^{\tilde{m}}\frac{\langle\tilde{K}_{zx,k}, \tilde{\phi}_{j,k}\rangle\langle\tilde{\phi}_{j,k}, g\rangle}{\tilde{\lambda}_{j,k}} \quad \text{for } g \in L^2[0,1],$$

where $\tilde{m}$ denotes a regime specific truncation parameter with $1 \le \tilde{m} < n|\hat{G}|$. From these objects the final slope estimate is calculated according to

$$\tilde{\beta}_t := \left[\tilde{K}_{z,k} - \tilde{\Phi}_k\left(\tilde{K}_{zx,k}\right)\right]^{-1}\left[\hat{K}_{zy,t} - \tilde{\Phi}_k\left(\hat{K}_{xy,t}\right)\right] \quad \text{for all } t \in \hat{G}_k.$$

Step 4 Estimate the regime specific parameter functions A_k. To do so, the response variables need to be transformed for each $t \in \hat{G}_k$ according to $\tilde{y}_{it} = \check{y}_{it} - \tilde{\beta}_t z_{it}$. For any $1 \le k \le \hat{K}$ the pairs $\{(\tilde{y}_{it}, X_{it}) : 1 \le i \le n,\ t \in \hat{G}_k\}$ are then pooled in order to estimate A_k by the following regression:$\tilde{y}_{it} = \langle X_{it}, A_k\rangle + \varepsilon_{it}^*$ using the procedure in [10]. Here, the error term is composed according to $\varepsilon_{it}^* := \varepsilon + (\beta_t - \tilde{\beta}_t)z_{it}$. The resulting estimator $\tilde{A}_k$ obtains as

$$\tilde{A}_k(s) = \sum_{j=1}^{\tilde{m}} \tilde{a}_{j,k} \tilde{\phi}_{j,k}(s), \quad \text{where} \quad \tilde{a}_{j,k} = \tilde{\lambda}_{j,k}^{-1} \left(n|\hat{G}_k|\right)^{-1} \sum_{i=1}^{n} \sum_{t \in \hat{G}_k} \langle X_{it}, \tilde{\phi}_{j,k} \rangle \tilde{y}_{it}.$$

In summary, the final estimates are $\hat{G}_1, \ldots, \hat{G}_k$ for the regimes as well as $\tilde{\beta}_t$, $\tilde{A}_k$ and $\hat{\alpha}_{0,t}$ for the regression parameters.

Practical choice of τ_{nt}. For large n and given m, it can be shown that the law of $2^{-1} \cdot n \cdot \Delta_{ts}$ can be reasonably well approximated by a χ_m^2-distribution. For the actual implementation we thus suggest to set the threshold to $\hat{\tau}_{nT} := 2 \cdot n^{-1} \cdot q_{0.99}(\chi_m^2)$, with $q_{0.99}(\chi_m^2)$ being the 99% quantile of a χ_m^2-distribution.

4 Asymptotic Theory

Two types of problems add to the well understood estimation in functional linear regression. The first one is the additional classification error contaminating the estimation of A_k. The second one is that estimation of A_k suffers from the presence of nuissance parameters $\alpha_{0,t}$ and β_t. For our asymptotic analysis we rely on a set of standard assumptions present in the literature[1]. In the following we present a detailed list of our assumptions.

Assumption 1. Suppose that

1. $\{(\varepsilon_{it}, X_{it}, z_{it}) : 1 \leq i \leq n, 1 \leq t \leq T\}$ are centered and iid over i and t,
2. $E\left[||X||_2^4\right] < \infty$, $E\left[z_{it}^4\right] < \infty$, $E\left[\varepsilon_{it}^4\right] < \infty$
3. ε_{it} is independent from X_{js} and z_{js} for any $1 \leq i, j \leq n$ and $1 \leq t, s \leq T$.

These conditions are even standard for analyzing classical regression models.

The next assumption contains the regularity conditions on the covariance structure of the functional regressor, on the complexity of the corresponding parameter functions, and on the covariance between functional and scalar regressors.

Assumption 2. Suppose there exist constants $0 < C_\lambda, C'_\lambda, C_\theta, C_a, C_{zX}, C_\beta < \infty$, such that

1. $C_\lambda^{-1} j^{-\mu} \leq \lambda_j \leq C_\lambda j^{-\mu}$ and $\lambda_j - \lambda_{j+1} \geq C'_\lambda j^{-(\mu+1)}$, $j \geq 1$ for the eigenvalues $\lambda_1 > \lambda_2 > \ldots$ of the covariance operator Γ of X_{it} and a $\mu > 1$,
2. $E\left[\langle X_{it}, \phi_j \rangle^4\right] < C_\theta \lambda_j^2$ for the eigenfunction ϕ_j of Γ corresponding to the j-th eigenvalue,
3. $|\langle A_k, \phi_j \rangle| \leq C_a j^{-\nu}$ for all $1 \leq k \leq K$,
4. $|\langle K_{zX}, \phi_j \rangle| \leq C_{zX} j^{-(\mu+\nu)}$, where $K_{zX} := E[X_{it} z_{it}]$, and
5. $\sup_{1 \leq t \leq T} \beta_t^2 < C_\beta$.

[1] See, for instance, [8], [9], [6] and [10]

Statements 1 and 3 are traditional in the literature (see, e.g., [10] or [7] among others) while [8] introduces (variant of) statements 2 and 3.

The following assumption specifies the panel asymptotics we are considering.

Assumption 3. Suppose that

1. $|G_k| \propto T$ and
2. $T \propto n^{\delta}$ for some $\delta > 0$.

By writing $n, T \to \infty$ it is meant that n and T diverge simultaneously on the path specified by the second point in Assumption 3. The only restriction on δ is indirectly formulated in terms of ν and μ.

Assumption 4. Suppose that $\nu > \max\{r_1, r_2, r_3\}$, where $r_1 := 3(1+\mu)$, $r_2 := (\delta)^{-1}(1+\delta+\mu/2)$ and $r_3 := ((1+\delta)+(1+2\delta)\mu)/2$.

Assumption 4 can be understood as setting an upper bound on the complexity of the parameter functions A_k in terms of the complexity of the regressor X and the rate at which nuisance parameters β_t are added. Another issue which is generic to many nonparametric problems is the specification of truncation, i.e., smoothing parameters, addressed in the following Assumption.

Assumption 5. Suppose for the truncation parameters $m = m(n)$ and $\tilde{m} = \tilde{m}(n,T)$ that $m \propto n^{\frac{1}{\mu+2\nu}}$ and $\tilde{m} \propto (n|G_k|)^{\frac{1}{\mu+2\nu}}$. While m is the standard truncation parameter as in the related literature on cross sections, $\tilde{m}$ is a logical extension to within-regime observations.

The following assumption is borrowed from [8] and serves as the central regularity condition in the estimation of $\hat{\beta}_t$ in step 1.

Assumption 6. The random variables $q_{it} := z_{it} - \int_0^1 X(u)\left(\sum_{j=1}^{\infty} \frac{\langle K_{zX}, \phi_j\rangle}{\lambda_j}\phi_j(u)\right)\mathrm{d}u$ are iid and $E[q_{it}|X_{1t},\dots,X_{nt}] = 0$ as well as $E[q_{it}^2|X_{1t},\dots,X_{nt}] > 0$.

A last assumption is designed to ensure identification of different the regimes given the thresholding procedure.

Assumption 7.

1. The threshold parameter $\tau_{nT} \to 0$ satisfies $\mathbb{P}\left(\max_{t,s\in G_k}\hat{\Delta}_{ts} \leq \tau_{nT}\right) \to 1$ for all $1 \leq k \leq K$.
2. There exists some $C_{\Delta} > 0$ such that for any $1 \leq k \leq K$ and any $t \in G_k$

$$||\alpha_t^{(\Delta)} - \alpha_s^{(\Delta)}||_2^2 =: \Delta_{ts} \begin{cases} \geq C_{\Delta} & \text{if } s \notin G_k \\ = 0 & \text{if } s \in G_k, \end{cases}$$

where $\alpha_t^{(\Delta)} := \sigma_{\varepsilon}^{-1}\sum_{j=1}^{\infty}\lambda_j^{1/2}\langle\alpha_t,\phi_j\rangle\phi_j$ and $\sigma_{\varepsilon}^2 := E[\varepsilon_{it}^2]$.

Assumption 7 consist of the same ingredients as the corresponding requirements in [9]. It ensures that different regimes have different parameter functions in the relevant L^2-metric and that the threshold tends to zero sufficiently slowly.

Based on the above assumptions we conclude with the following results starting with a Lemma extending Theorems 3.1 and 3.2 in [8].

Lemma 1. *Given Assumptions 1,2,4,5 and 6 it holds for all* $1 \leq t \leq T$ *as* $n \to \infty$ *that*

$$(\hat{\alpha}_{0,t} - \alpha_{0,t})^2 = O_p\left(n^{-1}\right), \tag{2}$$

$$\left(\hat{\beta}_t - \beta_t\right)^2 = O_p\left(n^{-1}\right), \tag{3}$$

$$||\hat{\alpha}_t - \alpha_t||_2^2 = O_p\left(n^{\frac{1-2\nu}{\mu+2\nu}}\right). \tag{4}$$

Based on the consistency of the $\hat{\alpha}_t$ for a fixed t our estimation procedure achieves classification consistency in the sense of the following theorem.

Theorem 1. *Given Assumptions 1–7 it holds that*

$$\mathbb{P}\left(\{\hat{G}_1, \ldots, \hat{G}_{\hat{K}}\} \neq \{G_1, \ldots, G_K\}\right) = o(1) \quad \text{as } n, T \to \infty. \tag{5}$$

The following result establishes rates of convergence for the suggested estimators.

Theorem 2. *Given Assumptions 1–7 it holds for all* $1 \leq k \leq \hat{K}$ *that*

$$\left|\hat{G}_k\right|^{-1} \sum_{t \in \hat{G}_k} (\hat{\alpha}_{0,t} - \alpha_{0,t})^2 = O_p\left(n^{-1}\right), \tag{6}$$

$$\left|\hat{G}_k\right|^{-1} \sum_{t \in \hat{G}_k} \left(\tilde{\beta}_t - \beta_t\right)^2 = O_p\left(n^{-1}\right), \tag{7}$$

$$\left|\left|\tilde{A}_k - A_k\right|\right|_2^2 = O_p\left(n^{-1}\right) \quad \text{as } n, T \to \infty. \tag{8}$$

Here, root-n consistency of $\tilde{A}_k$ is mainly a consequence of Assumption 4.

5 Regime Dependent Pricing of Idiosyncratic Risk

The specific form of our suggested statistical model is strongly motivated by economic theory. According to a standard asset pricing approach, underdiversified investors ask for a premium compensating for the idiosyncratic risk of an asset. Proxying idiosyncratic risk with idiosyncratic volatility, our statistical model offers a tailor-made tool to uncover dynamics in such premiums from price data[2]. In Section 5.1, we motivate our statistical model in (1) from a theoretical viewpoint and argue how to construct suitable functional regressors from discrete data. We provide a brief outline of the ongoing empirical study in Section 5.2.

[2] See, e.g., [1] for the notion of time varying risk premiums.

5.1 Economic Modeling and Volatility-Curve Estimation

Our approach is motivated by the stochastic volatility framework of [4]. Let us denote the underlying probability space as $(\Omega, \mathscr{A}, \mathbb{P})$, where randomness is emphasized by making the dependence on an $\omega \in \Omega$ explicit. For a point in time $s \geq 0$ we suggest considering the following differential equation

$$\mathrm{d}\log\left(P(s,\omega)\right) = \mu(s,\omega)\mathrm{d}s + \alpha(s)\sigma^2(s,\omega)\mathrm{d}s + \sigma(s,\omega)\mathrm{d}W_s \tag{9}$$

to model the log returns corresponding to the asset price P, where μ and W_s denote a stochastic drift and a Wiener process. The object $\sigma(s,\omega)$ is the stochastic instantaneous volatility and α a time varying risk premium. Model (9) generalizes the log-price process in [4].

For our purpose it is of interest to deal with the returns between the beginning and the end of a certain period such as a trading day, a week or a month. Choosing without loss of generality a unit interval, the start and end points of the tth period are given by $t-1$ and t. It follows then immediately from (9) that

$$y_t := \log\left(\frac{P(t,\omega)}{P(t-1,\omega)}\right) = \gamma_t + \int_0^1 \alpha_t(v)\sigma_t^2(v)\mathrm{d}v + \varepsilon_t. \tag{10}$$

with $\gamma_t := \int_{t-1}^t \mu(v,\omega)\mathrm{d}v$ and $\varepsilon_t := \int_{t-1}^t \sigma(v,\omega)\mathrm{d}W_v$. Further $\alpha_t(v)$ and $\sigma_t^2(v)$ are defined as $\alpha_t(v) := \alpha(t-1+v)$ and $\sigma_t^2(v) := \sigma^2(t-1+v)$ to emphasize the t-dependence of the two objects.

Beyond that we postulate that the drift term is at each time $s > 0$ proportional to a stochastic process $Z(s)$, describing pricing relevant market frictions. The proportionality factor is a deterministic step function $b(s)$ taking values β_t in the interval $[t-1,t]$. That is, we assume that $\mu(s) \propto b(s)Z(s)$ which implies that $\gamma_t = \beta_t z_t$ with $z_t := \int_{t-1}^t Z(v)\mathrm{d}v$. As a consequence (10) reads as

$$y_t = \beta_t z_t + \int_0^1 \alpha_t(v)\sigma_t^2(v)\mathrm{d}v + \varepsilon_t. \tag{11}$$

Above, we implicitly operated on two time scales: the discrete time periods indexed by t and the intra-period times over the generic interval $[0,1]$. The notation in (11) indicates that the parameters (α_t, β_t) might change on the first time scale. We postulate that the parameter of interest, α_t, is *regime*-specific with different possible parameter values $A_1, \dots, A_K$, where each parameter A_k can be interpreted as the regime specific idiosyncratic risk premium. Such regimes are then collections of periods with similar investor's perception of the asset-specific risk relevance.

As the Log-Volatility (LV)-trajectories σ_t^2 are latent, they need to be recovered from observed discrete log-returns. Along the lines of [5] this proceeds as follows. Without loss of generality, intraday trading time is indexed in the unit interval $[0,1]$. Denote the incremental log-return over an interval of length $0 < \Delta \ll 1$ for some $s \geq 0$ as

$$Y_\Delta(s) := \Delta^{-\frac{1}{2}} \log\left(\frac{P(s+\Delta)}{P(s)}\right) = \beta_t \Delta^{\frac{1}{2}} Z(s) + \Delta^{\frac{1}{2}} \alpha(s)\sigma^2(s) + \sigma(s) W_\Delta(s) + \sum_{j=1}^{3} R_{j,\Delta},$$

where $W_\Delta(s) := \Delta^{-\frac{1}{2}} (W(s) - W(s-\Delta))$ and $R_{1,\Delta}, R_{2,\Delta}, R_{3,\Delta}$ are discretization errors as described in the appendix . The latter are negligible in the following sense.

Theorem 3. *Given Assumption 8 in the appendix,* $\sum_{j=1}^{3} R_{j,\Delta} = O_p(\Delta^{1/2})$ *as* $\Delta \to 0$.

This justifies the central small-Δ approximation $Y_\Delta(s) \approx \sigma(s) W_\Delta(s)$. From this it can be concluded that $\log\left(Y_\Delta(s)^2\right) + c_0 \approx X(s) + e_s$, where $c_0 \approx 1.27$ and $e_s := \log\left(W_\Delta(s)^2\right) - c_0$ denotes an error term (for details see [5] and the references therein). Most importantly the LV process is defined according to $X(s) := \log(\sigma(s)^2)$. Given returns are observed on an intraday grid $\mathbf{D} := \{0, \Delta, 2\Delta, \dots, 1-\Delta, 1\}$, it is reasonable to understand $\log\left(Y_\Delta(s)^2\right)$ as discrete noisy observations of the LV-process, provided one is willing to assume a suitable dependence structure for $e_s,\ s \in \mathbf{D}$. As a consequence the method from [5] to estimate $X(s)$ from the discrete prices also applies to our setup without further adjustment.

5.2 *Empirical Study: Risk Regimes in the US Stock Market*

Using the presented framework we examine risk pricing in the US stock market. Data for the S&P 500 constituents is recorded over 100 trading days in 2016. Beyond asset prices, which are sampled every 10 minutes, market frictions, e.g. illiquidity, are proxied by a daily bid-ask spread. Using the estimation procedure indicated in the previous section we are able to construct daily LV-trajectories for each (i,t)-combination. These LV-trajectories are employed as functional regressors X_{it}, while y_{it} is the end-of-day price and z_{it} the maximum bid-ask spread reflecting frictions relevant for asset i at day t.

First estimation results indicate that the regimes and risk premiums distinguish times of financial turmoil from tranquil days.

6 Conclusion

In this paper we present a novel regression framework, allowing to examine regime specific effects of a random function on a scalar response in the presence of parametric nuissance terms. Our estimation procedure is designed for a panel data context. We prove consistency and derive rates of convergence. The relevance of our semiparametric model is underlined by an application to idiosyncratic risk pricing in the presence of frictions.

References

[1] **Anderson, R.M.** (2011). "Time-varying risk premia?" *Journal of Mathematical Economics*, 47 (3): 253–259.

[2] **Ang, A. and Hodrick, A.R. and Xing, Y. and Zhang, X.** (2016). "Have we solved the idiosyncratic volatility puzzle?" *Journal of Financial Economics*, 121 (1): 167–194.

[3] **Bada O. and Gualtieri J. and Kneip A. and Sickles R.C.** (2015). "Panel data models with multiple jump discontinuities in the parameters." *Working Paper*.

[4] **Barndorff-Nielsen, O.E. and Shephard, N.** (2002). "Econometric analysis of realized volatility and its use in estimating stochastic volatility models." *Journal of the Royal Statistical Society. Series B*, 64 (2): 253–280.

[5] **Müller, H.G. and Sen, R. and Stadtmüller, U.** (2011). "Functional data analysis for volatility." *Journal of Econometrics*, 165 (2): 233–245.

[6] **Horváth, L. and Reeder, R.** (2012). "Detecting changes in functional linear models." *Journal of Multivariate Analysis*, 111: 310–334.

[7] **Kneip, A. and Poss, D. and Sarda, P.** (2016). "Functional linear regression with points of impact." *The Annals of Statistics*, 44 (1): 1–30.

[8] **Shin, H.** (2009). "Partial functional linear regression." *Journal of Statistical Planning and Inference*, 139 (10): 3405–3418.

[9] **Vogt, M. and Linton, O.** (2016). "Classification of nonparametric regression functions in longitudinal data models." *Forthcomming in: Journal of the Royal Statistical Society: Series B*.

[10] **Hall, P. and Horowitz, J.L.** (2007). "Methodology and convergence rates for functional linear regression." *The Annals of Statistics*, 35 (1): 70–91.

[11] **Su, L. and Shi, Z. and Phillips, P.C.B.** (2016). "Identifying latent structures in panel data." *Econometrica*, 84 (6): 2215–2264.