CREST
CENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

WORKING PAPER SERIES

Continuous-time Impulse Response Functions with functional approaches and mixed-frequency data

Catherine DOZ, Laurent FERRARA, Anna SIMONI

N°4/May 2026

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

Continuous-time Impulse Response Functions with functional approaches and mixed-frequency data*

Catherine Doz† Laurent Ferrara‡ Anna Simoni§

Abstract

The impulse response function (IRF) characterizes how a given structural shock propagates over time through the economy. To measure the impact of a high-frequency shock on low-frequency macroeconomic aggregates, the usual approach relies on aggregating all data at the lowest frequency, potentially leading to a loss of information. This paper proposes a novel concept to measure this macroeconomic impact directly at high frequency, without requiring aggregation: the continuous-time IRF (CT-IRF). We express the response of the low-frequency target variable y_t to the input signal as a convolution integral between the impulse-response and the input signal. Our approach is similar in spirit to local projections, with the key difference being that we can construct the entire IRF in a single step and handle mixed frequencies. The estimation problem is an ill-posed inverse problem, which we address using a penalized least squares estimator with a Sobolev-type penalty. We derive the convergence rate of our estimator as the sample size T grows and demonstrate its excellent finite-sample performance via a Monte Carlo study. Finally, we apply our method to estimate high-frequency IRFs of quarterly U.S. business investment to uncertainty shocks and of U.S. GDP growth rate to financial shocks.

*May 2026 - First version December 2024. An earlier version of this paper circulated under the title "Impulse Response Function with functional approaches and mixed-frequency data".

†Paris School of Economics (PSE) and University Paris 1 Panthéone-Sorbonne, Paris (France). e-mail: Catherine.Doz@univ-paris1.fr

‡SKEMA Business School - University Côte d'Azur, Nice (France). e-mail: laurent.ferrara@skema.edu

§CREST, CNRS, École Polytechnique, ENSAE, 5 Avenue Henry Le Chatelier, 91120 Palaiseau (France). Phone: +33(0)170266837. e-mail: anna.simoni@polytechnique.edu

Keywords: Impulse response function, high-frequency data, alternative data, covariance operators, penalized least squares.

1 Introduction

Economic time series used in macroeconomic models for prediction and impulse-propagation studies are often sampled at different frequencies; for instance, monthly, quarterly, and annual frequencies. With the rise of data from non-traditional sources, such as social networks, narratives, satellite imagery, or the internet (see *e.g.* Ferrara and Simoni 2023), frequency mismatches among series have become even more pronounced because these datasets are typically sampled at higher frequencies, such as weekly or daily. In this context, economists can either aggregate high-frequency series to the lowest frequency or incorporate mixed frequencies into the model using an appropriate weighting scheme. The first approach risks losing potentially valuable high-frequency information.

Suppose one is interested in measuring the impact of a shock that occurs in a high-frequency series, such as a daily financial shock, on a low-frequency series, like the quarterly GDP growth rate. For example, if the shock occurs on the first day of a quarter, its impact must be measured 60 days later; if it occurs on the third day, its impact must be measured 57 days later, and so on. To obtain such a measure, we need to construct an impulse response function (IRF) and estimate it at every high-frequency period. Given the brevity of the high-frequency period, this IRF can be thought of as a discretization of a continuous-time IRF.

We first introduce the concept of continuous-time IRF (CT-IRF) in macroeconomics by borrowing from the theory of linear time-invariant systems in engineering and signal processing (see for example Margaris (2024)), where the impulse response describes how a system reacts to a brief input signal, referred to as an *impulse*, in a continuous time stochastic process $\{x(u)\}$. In continuous time, this impulse is mathematically represented by the Dirac delta function, which models an infinitely short impulse with unit area. In macroeconomics, the continuous-time stochastic process $\{x(u)\}$ models the process that underlies the high-frequency series, and we see the latter as a discretization, on a fine grid, of $\{x(u)\}$. The key idea to recover the CT-IRF is to consider the high-frequency observations of the series as the discretized realization (the discretized trajectory) of a stochastic process in continuous time. This is particularly true when the number of times the high-frequency period enters the

low-frequency period is large; that is, when the frequency mismatch is wide, as, for instance, when dealing with daily and quarterly frequencies.

Then, we consider the linear projection of the low-frequency target series y_t onto the path segment of $\{x(u)\}$ on $[t - p_x, t]$, and onto a set of controls that include lagged values of all variables in the system up to the period $t - p_x$ at which the impulse occurs. To account for the continuous nature of $\{x(u)\}$, we express the response of the low-frequency target variable y_t to the input signal $\{x(u)\}$ through a convolution integral between the impulse response, denoted by β, and the input signal $\{x(u)\}$. In the context of causal impact analysis, our model measures the dynamic influence of $\{x(u)\}$ on the target variable y_t for a given v which is the gap between the forecast horizon and the moment at which the shock occurs. We recover the impact path as a function of $v \mapsto \beta(v)$. Our proposed methodology is similar in spirit to the local projections approach (*e.g.* Jordà (2005), Plagborg-Møller and Wolf (2021),Chudik and Georgiadis (2022)) in that we do not require the specification of the unknown true multivariate dynamic system. On the other hand, we extend the local projection literature by incorporating mixed-frequency into the setup. This has two major benefits. First, we provide a method that enables inference without requiring the aggregation of observations, which may entail a potential loss of information. Second, we do not have to run a separate regression for each forecasting horizon, as is done in the local projections literature. Instead, we propose a one-step estimation for all high-frequency forecasting horizons.

In order to make inference on the CT-IRF, which is a deconvolution problem for which we observe only one discretized trajectory, we write the problem as a linear functional regression. While the resulting regression is related to the functional linear regression model (*e.g.* Cai and Hall (2006); Li and Hsing (2007); Crambes et al. (2009) and references therein), it differs from it in many respects. First, the high-frequency series and its functional coefficient are not expressed here as a function of time but as a function of the time-gap between the forecasting horizon and a particular time lag in the past. This specific way of writing the functional linear regression is due to the fact that our initial model is a convolution. This is new in the linear functional regression literature, to the best of our knowledge. Second, we consider a multivariate setting that, in addition to the continuous time variable, includes lagged values of the dependent variable as well as other variables in the system as control variables. Most of the previous literature on functional regression models considers only one covariate. The few papers that consider the multivariate extension, such as Fan et al. (2015), do not consider either lagged values or mixed frequency. The consideration of mixed-frequency changes the technique of identification and the expression of the estimator

because we have to project onto the space orthogonal to the low-frequency variables. Finally, in contrast to the functional linear regression literature, we do not observe many realizations of the functional signal but only one discretized trajectory of $\{x(u)\}$. We then construct many functional observations from it by cutting partially overlapping sample paths.

The estimation of the IRF belongs to the class of ill-posed inverse problems. Ill-posedness refers to the fact that a small change in the data may result in a large change in the causal estimate, a problem also known as overfitting (*e.g.* Florens (2003), Florens and Simoni (2012, 2016)). To solve this challenge, our estimator involves regularization, which consists of penalizing the least squares criterion with a Sobolev-norm penalty after representing the continuous-time stochastic process $\{x(u)\}$ and the β function in terms of a basis function expansion. The estimator is called the Penalized Least Squares (PLS) estimator. While this estimator is well known in the literature, to the best of our knowledge, its performance has not been analyzed against the background of local projections with mixed-frequency.

We study the asymptotic behavior of the PLS estimator of β under the assumptions of joint stationarity and β-mixing of $\{x(u)\}$. We establish the rate of convergence with respect to the L^2-norm. This rate depends on the regularity of the true CT-IRF and on the decay rate of the eigenvalues of the covariance operator of $\{x(u)\}$. The latter operator plays a central role in our analysis and the construction of the estimator.

Finite sample properties of our procedure are studied through Monte Carlo simulations. We show in simulations that our model largely outperforms existing methods designed to deal with mixed frequency datasets and to recover the true IRF, even in the more challenging case of non-smoothness. This result is robust across different specifications of the Data Generating Process (DGP). The basis expansions that we consider are Fourier basis functions, B-splines, and the eigenfunctions of the empirical covariance operator of the high-frequency series.

We then study the performance of our approach in two real data applications. The first one focuses on the empirical question of measuring the impact of uncertainty on quarterly business investment in equipment for the U.S. economy. We focus on uncertainty shocks estimated at a daily frequency and use the Economic Policy Uncertainty (EPU) daily index computed by Baker et al. (2016). We show that the growth of equipment investment negatively responds to a daily economic policy uncertainty shock on impact, with an amplitude of about 0.75%. This result is very much in line with standard responses of investment to an uncertainty shock.

The second real data application examines the extent to which a high-frequency financial

condition index, as measured by the Composite Indicator of Systemic Stress (CISS) developed by the European Central Bank (Hollo et al. (2012)), affects the U.S. business cycle. We find that the impact becomes statistically different from zero after 15 business days have passed since the impact. The trough of the impact is reached around 60 business days after the initial impact. Afterward, the effect gradually dissipates and returns to zero after 7 months.

In addition to the literature on local projections, our paper relates to the vast literature on impulse response functions, *e.g.* Hamilton (1994), Koop et al. (1996), Inoue and Rossi (2021), and Chudik and Kilian (2026). Finally, due to the treatment of mixed frequency, our paper relates to the literature on mixed frequency sampling, which has been developed mainly for the purpose of nowcasting rather than for estimating the impulse response function. This literature includes parametric weighting schemes (MIDAS, Ghysels et al. (2007) and UMIDAS, Foroni et al. (2015)), as well as their nonparametric extensions (Mogliani and Simoni (2021); Babii et al. (2022); Mogliani and Simoni (2024)). We extend this literature by proposing a method that works particularly well when the gap in sampling frequencies is wide. Note also that there is a potentially large field of applications for our model, as it is able to assess the dynamic effects of multiple shocks within a given time period, in contrast to standard approaches that can only account for one shock at a time. In practice, the occurrence of a sequence of shocks is frequent, especially during turbulent periods, such as wars or recessions.

The rest of the paper is organized as follows. We start by providing the definition of continuous-time IRF in Section 2. In Section 3, we present the model and establish the conditions under which identification holds. The estimation procedure is described in Section 4. We also describe the effects of discretization. Section 5 discusses the theoretical performance of our model. Results of the real data applications are illustrated in Sections 6 and 7. Section 8 presents the results of a Monte Carlo experiment. Extensions to the multivariate case are presented in Section 9. All proofs are provided in the Supplementary Appendix.

2 Continuous Time Impulse Response Functions

In this section, we provide the definition of the impulse response function in continuous time, referred to as CT-IRF. To start, let's recall the definition of the IRF in discrete time as presented in Hamilton (1994) and Koop et al. (1996). An IRF captures the dynamic variation in the conditional expectation of the target variable y_t in response to an isolated

variation in one of the variables affecting the system while keeping all the other variables fixed. For a fixed $\Delta t > 0$, define $t_n := n\Delta t$, for $n \geqslant 1$. Suppose that in one scenario, the system is hit between t_1 and t_{1+h}, for $h \in \mathbb{N}_+$ only by a shock in the variable x_{t_1} of size ϱ_Δ at period t_1 and in the second scenario the system is not hit by any shock between t_1 and t_{1+h}. Then, the discrete time IRF is defined as the difference between two conditional forecasts:

$$\begin{aligned}&\mathcal{R}_{x\to y}(t_1, h, \varrho_\Delta; \mathbf{y}_{t_1-1}) := \\ &\mathbf{E}[y_{t_{1+h}}|x_{t_1} = \varrho_\Delta, x_{t_2} = 0, \ldots, x_{t_{1+h}} = 0; \mathbf{y}_{t_1-1}] - \mathbf{E}[y_{t_{1+h}}|x_{t_1} = 0, \ldots, x_{t_{1+h}} = 0; \mathbf{y}_{t_1-1}], \quad (1)\end{aligned}$$

where $\mathbf{E}[\cdot|\cdot]$ denotes the best mean squared error predictor, and the other variables affecting the system are, for simplicity, reduced to $\mathbf{y}_{t_1-1}$. The shock ϱ_Δ can be modeled as a measure on $\mathbb{R}$: for a fixed $\Delta t > 0$ and $\varrho > 0$, $\varrho_\Delta(A) := \varrho\Delta t 1_A(t_1)$, where $A \subset \mathbb{R}$ is a measurable set, and $1_A(t_1)$ is the indicator function equal to 1 if $t_1 \in A$ and 0 otherwise. This characterization is useful in order to view $\varrho_\Delta(\cdot)$ as a finite approximation of the Dirac delta in the continuous-time limit.

In the continuous-time framework, the series on which the shock occurs is a continuous-time stochastic process $x(\cdot) : \mathbb{R} \to \mathbb{R}$, and we observe a discretized version of $x(\cdot)$ at high-frequency. The variable $\{y_t\}_{t\geqslant 1}$ is discrete and observed at low-frequency. To model an instantaneous shock at time t_1, we represent it as an impulse $\rho\delta_{t_1}(t)$, where $\delta_{t_1}(t)$ is the Dirac delta function, which is a generalized function equal to zero everywhere except at $t = t_1$, where it is infinite, and has an integral over $\mathbb{R}$ equal to 1, see *e.g.* Kolmogorov and Fomin (1975). The Dirac delta function can be thought of as representing the limit of a pulse ϱ_Δ as $\Delta t \to 0$, while keeping its integral constant.

If there is a shock (or impulse) at $t = t_1$ in the stochastic process $x(\cdot)$, then we can represent it as $x(t) + \varrho\delta_{t_1}(t)$. If we take the linear best mean squared error predictor, that is, $\mathbf{E}[y_t|\{x(u); t - p_x \leqslant u \leqslant t\}] = \int_{t-p_x}^t x(u)\beta(t-u)du$ for some $p_x > 0$, then the function $u \mapsto \beta(t-u)$ measures this dynamic response to a unit shock and is the *continuous time IRF* (CT-IRF). More precisely, we can write $\mathbf{E}[y_t|\{x(u); t - p_x \leqslant u \leqslant t\}] = (x * \beta)(t)$, where '*' denotes the convolution product.

Thus, we define the *continuous time IRF* (CT-IRF) as the image of the Dirac's delta function under the transformation $\int_{t-p_x}^t \cdot\beta(t-u)du$, which is the dynamic response of y_t at time t to a unit impulse in $x(\cdot)$ at time t_1. The response to an arbitrary input signal $x(\cdot)$ is given by the convolution $(x * \beta)(t)$. That is, if we consider the impulse $\varrho\delta_{t_1}(t)$, then $\int_{t-p_x}^t \varrho\delta_{t_1}(u)\beta(t-u)du = \varrho\beta(t-t_1) = \varrho(\delta_{t_1} * \beta)(t-t_1)$, for every $t_1 \in [t-p_x, t]$.

Given the dynamic nature of the series that we consider, we can include pre-shock information and enrich the model with lagged predetermined values of the dependent variable y_t, so that we obtain:

$$\mathbf{E}[y_t|\{x(u)\}_{t-p_x\leqslant u\leqslant t}, y_{t-p_x}, \ldots, y_{t-p_y}] = \gamma(L)y_{t_1} + \int_{t_1}^{t} x(s)\beta(t-s)ds,$$

where $\gamma(L) := 1 - \gamma_1 L - \gamma_2 L^2 - \ldots - \gamma_{p_y} L^{p_y}$ is a polynomial in the lag operator L for some $p_y > 0$. For every $u \in [t - p_x, t]$, $\beta(t-u)$ gives the marginal effect of a unit shock at time u on y_t.

The integral $\int_{t-p_x}^{t} x(s)\beta(t-s)ds$ represents the continuous-time aggregation of shocks in $x(\cdot)$ over the interval $[t-p_x, t]$, weighted by the impulse response kernel $\beta(t-s)$. This formulation enables us to assess the cumulative dynamic effects on the low-frequency variable y_t of more than one shock within the interval $[t-p_x, t]$. Accounting for multiple shocks in the IRF is particularly valuable in practice, as economists often face a sequence of shocks during turbulent times, such as financial crises or periods of high uncertainty, as illustrated by the war in Iran that began in early 2026. However, standard approaches to estimating IRFs, such as Structural Vector Autoregressive (SVAR) models or Local Projections, only permit one shock at a time. Our approach addresses this limitation, offering multiple potential applications in macroeconomics.

To conclude this section, it is noteworthy that, in practice, the shock observed will not be in continuous time but in discrete time, even if it is measured in high frequency. Therefore, one must consider the discrete time shock as the discretization of the continuous time shock; that is, the integral over the high-frequency period of the continuous time shock.

3 The model

As made clear in Jordà (2005), equation (1) implies that to recover the IRF, one must first obtain the best mean-squared multi-step predictions. We consider the linear prediction model, which is an approximation of the true conditional expectation given past values of y_t and the past trajectory of $x(\cdot)$. We denote by $\mathbf{EL}[\cdot|\cdot]$ the linear conditional expectation and by $\mathbf{E}[\cdot]$ the expectation with respect to the true probability distribution.

Consider a low-frequency variable y_t indexed by the low-frequency time unit $t = 1, \ldots, T$, and a continuous-time stochastic process $x(\cdot) : \mathbb{R} \to \mathbb{R}$, whose trajectory is discretely observed at high frequency. Note that $\{x(u)\}$ does not directly model the high-frequency

variable, but rather the underlying process generating it. For a given $p_x \leqslant p_y \in \mathbb{N}_+$, consider projecting y_t linearly onto the linear space generated by $\{y_{t-p_x}, y_{t-p_x-1}, \ldots, y_{t-p_y}\}$. For every $t \in \{p_y, \ldots, T\} \subset \mathbb{N}_+$

$$\mathbf{EL}\left[y_t | y_{t-p_x}, \ldots, y_{t-p_y}, \{x(\tau)\}_{\tau \in (t-p_x, t]}\right] = \sum_{i=p_x}^{p_y} \gamma_{i+1-p_x} y_{t-i} + \int_{t-p_x}^{t} \beta(t-\tau) x(\tau) d\tau. \quad (2)$$

We implicitly assume that:

$$y_t \underline{\perp\!\!\!\perp} \{y_s\}_{s<t-p_y}, \{x(s)\}_{s \leqslant t-p_x} \Big| y_{t-p_x}, \ldots, y_{t-p_y}, \{x(\tau)\}_{\tau \in (t-p_x, t]}.$$

We exclude the regressors $y_{t-1}, \ldots, y_{t-p_x+1}$ in order to isolate the direct and indirect effects of the signal $x(\cdot)$ at the forecasting horizon t, where by indirect effect we mean the effect of the shock that passes to y_t through $y_{t-1}, \ldots, y_{t-p_x+1}$. By denoting ε_t as the linear projection residual, we can equivalently write: for every $t = \{p_y, \ldots, T\} \subset \mathbb{N}_+$,

$$y_t = \sum_{i=p_x}^{p_y} \gamma_{i+1-p_x}\, y_{t-i} + \int_{t-p_x}^{t} \beta(t-\tau) x(\tau) d\tau + \varepsilon_t, \qquad \mathbf{E}\left[\varepsilon_t(\mathbf{y}_{t-p_x}, x(s))\right] = 0, \quad \forall s \in \mathbb{R}, \quad (3)$$

where $\mathbf{y}_{t-p_x} := (y_{t-p_x}, \ldots, y_{t-p_y})'$ is the $(p_y - p_x + 1)$-vector of lagged values of the dependent variable. The functional parameter $v \mapsto \beta(v)$ is unknown and must be estimated. Here, the argument v denotes the forecast horizon gap, which is the time elapsed between the measurement point y_t and the moment the shock arises in the signal $\{x(u)\}$, and it characterizes the CT-IRF.

By making the change of variable $v = t - \tau$, (3) writes: for every $t = \{p_y, \ldots, T\} \subset \mathbb{N}_+$,

$$y_t = \sum_{i=p_x}^{p_y} \gamma_{i+1-p_x}\, y_{t-i} + \int_{0}^{p_x} \beta(v) X_t(v) dv + \varepsilon_t, \qquad \mathbf{E}\left[\varepsilon_t(\mathbf{y}_{t-p_x}, X_t(s))\right] = 0, \quad \forall s \in \mathbb{R}, \quad (4)$$

where $X_t(\cdot) : [0, p_x] \to \mathbb{R}$ is defined as $X_t(v) := x(t-v)$ for every $v \leqslant t$, and as $X_t(v) := 0$ for every $v > t$. In addition to $\beta(\cdot)$, the $(p_y - p_x + 1)$-vector $\gamma := (\gamma_1, \ldots, \gamma_{p_y - p_x + 1})'$ is also an unknown parameter that must be estimated. In the following, we work with model (4), show its identification, and then explain how to estimate it. Model (4) is a linear regression model, and its distinctive feature compared to classical linear models is the presence of a functional covariate. This modeling is helpful when the frequency gap is very large, such that the high-frequency series is better approximated by a stochastic process in continuous time rather than in discrete time.

Let L^2 denote the space of real-valued square integrable functions on $[0, p_x]$ with respect to the Lebesgue measure. We denote by $\langle \cdot, \cdot \rangle$ the inner product in L^2 and by $\| \cdot \|$ the norm induced by it. By assuming $\beta \in L^2$ and $X_t \in L^2$ for every $t \in \{p_y, \ldots, T\}$ almost surely, the model (4) can be written as

$$y_t = \mathbf{y}'_{t-p_x}\gamma + \langle \beta, X_t \rangle + \varepsilon_t, \qquad \mathbf{E}\left[\varepsilon_t(\mathbf{y}_{t-p_x}, X_t(s))\right] = 0, \quad \forall s \in \mathbb{R}, \tag{5}$$

which is a functional linear regression model. As pointed out in the Introduction, this is not a classical linear regression model because the signal $X_t(\cdot)$ and its coefficient $\beta(\cdot)$ are not expressed as functions of time but as a function of the time-gap, and because we do not observe many realizations of the process $\{x(u)\}$. Instead, we observe only one trajectory of $\{x(\cdot)\}$, which we divide into (partially) overlapping path segments. Each of these partially overlapping path segments is an observation of $X_t(\cdot)$.

We introduce the following assumption on the stochastic processes $\{y_t\}_{t\in\mathbb{Z}}$ and $\{x(t)\}_{\in\mathbb{R}}$.

Assumption 3.1. [*a*]. *The discrete-time stochastic process $\{y_t\}_t$ is a zero-mean weakly stationary process. The AR polynomial $1 - \sum_{i=p_x}^{p_y} \gamma_{i+1-p_x} z^i$ has all roots outside the unit circle.*
[*b*]. *The continuous-time stochastic process $\{x(t), t \in \mathbb{R}\}$ is a zero-mean L^2-valued weakly stationary process, that is, (i) $x \in L^2$ with probability 1, (ii) $\mathbf{E}[x(t)] = 0$, for all $t \in \mathbb{R}$, and (iii) for all $t_1, t_2 \in \mathbb{R}$, $\mathbf{E}[x(t_1)x(t_2)] = \gamma_x(t_1 - t_2)$ and $\gamma_x(t_1 - t_2) = \gamma_x(t_2 - t_1)$. In addition, $\mathbf{E}\|x\|^4 < \infty$.*
[*c*.] *The cross-covariance function between y_t and $x(t-s)$ is time invariant, that is, for all $t \in \mathbb{Z}$ and all $s \in \mathbb{R}$, $Cov(y_t, x(t-s)) = \gamma_{yx}(s)$, where $\gamma_{yx}(s)$ depends on the lag s and not on t.*
[*d*.] *Both $\{y_t\}$ and $\{x(t)\}$ are purely nondeterministic processes, and their joint dynamics do not contain any deterministic components.*

3.1 Identification of the Linear Functional Regression Model

In this subsection, we establish the identification of $(\gamma, \beta(\cdot))$. Let us consider the model (4). The orthogonality condition on ε_t implies the following set of functional moment restrictions that characterize the true values of (γ, β) denoted by (γ_0, β_0):

$$\mathbf{E}\left[y_t \begin{pmatrix} \mathbf{y}_{t-p_x} \\ X_t(\cdot) \end{pmatrix} \right] = \begin{pmatrix} \mathbf{E}\left[\mathbf{y}_{t-p_x}\mathbf{y}'_{t-p_x}\right] \\ \mathbf{E}\left[X_t(\cdot)\mathbf{y}'_{t-p_x}\right] \end{pmatrix} \gamma_0 + \begin{pmatrix} \mathbf{E}\left[\mathbf{y}_{t-p_x}\langle X_t, \beta_0 \rangle\right] \\ \mathbf{E}\left[X_t(\cdot)\langle X_t, \beta_0 \rangle\right] \end{pmatrix}, \tag{6}$$

where we recall that $\mathbf{E}[\cdot]$ denotes the expectation with respect to the true data distribution. We use the notation $p_{yx} := (p_y - p_x + 1)$ in the following discussion.

Denote $\Omega_{yy} := \mathbf{E}\left[\mathbf{y}_{t-p_x}\mathbf{y}'_{t-p_x}\right]$, which is a matrix of dimension $(p_{yx} \times p_{yx})$. The cross-covariance operator Ω_{yx} from L^2 to $\mathbb{R}^{p_{yx}}$ is defined by the covariance between the high-frequency variable $X_t(\cdot)$ and $\mathbf{y}_{t-p_x}$ as $\Omega_{yx}\beta := \mathbf{E}\left[\mathbf{y}_{t-p_x}\langle X_t, \beta\rangle\right]$, for every $\beta \in L^2$. The adjoint cross-covariance operator of Ω_{yx} is denoted by Ω^*_{yx} and defined as the operator from $\mathbb{R}^{p_{yx}}$ to L^2 that satisfies $(\Omega_{yx}\beta)'\gamma = \langle \beta, \Omega^*_{yx}\gamma\rangle$ for every $\gamma \in \mathbb{R}^{p_{yx}}$ and every $\beta \in L^2$, that is, $(\Omega^*_{yx}\gamma)(\cdot) := \mathbf{E}[X_t(\cdot)\mathbf{y}'_{t-p_x}]\gamma$. The covariance operator of $x(\cdot)$, denoted by $\Omega_{xx} : L^2 \to L^2$, is defined as $(\Omega_{xx}\phi)(\cdot) := \mathbf{E}\left[X_t(\cdot)\langle X_t, \phi\rangle\right], \forall \phi \in L^2$.

With these notations, equation (6) can be expressed as:

$$\mathbf{E}\left(y_t\mathbf{y}_{t-p_x}\right) = \Omega_{yy}\gamma_0 + \Omega_{yx}\beta_0 \tag{7}$$

$$\mathbf{E}\left(y_t X_t\right) = \Omega^*_{yx}\gamma_0 + \Omega_{xx}\beta_0. \tag{8}$$

In order to address identifiability issues for the model, we need to recall some properties of the covariance operator Ω_{xx} (see, for instance, Bosq (2000) for a general presentation). Under Assumption (3.1) and by using the Fubini's theorem we can write the operator Ω_{xx} as: for every $\phi \in L^2$,

$$(\Omega_{xx}\phi)(u) = \mathbf{E}\left[X_t(u)\langle X_t, \phi\rangle\right] = \int_0^{p_x} \mathbf{E}\left[x(t-u)x(t-v)\right]\phi(v)dv = \int_0^{p_x} \phi(v)\gamma_x(v-u)dv.$$

This operator is self-adjoint, non-negative, nuclear and therefore Hilbert-Schmidt and compact, see *e.g.* Cardot et al. (2007). In particular, Ω_{xx} admits the following decomposition: for every $\phi \in L^2$,

$$(\Omega_{xx}\phi)(u) = \sum_{j=1}^{+\infty} \lambda_j \langle \phi, \omega_j\rangle \omega_j(u),$$

where $(\omega_j)_{j\geqslant 1}$ is an orthonormal basis of L^2 and $(\lambda_j)_{j\geqslant 1}$ is a positive sequence such that $\sum_{j=1}^{+\infty}\lambda_j = \mathbf{E}\|x\|^2 < \infty$. Thus, $(\lambda_j, \omega_j)_{j\geqslant 1}$ is the eigensystem of Ω_{xx}; see Bosq (2000, page 6). We now make the following assumption, which ensures the identifiability of β_0 and then of γ_0.

Assumption 3.2. *Let Ω_{xx} be the covariance operator defined above. The following holds:*
[a.] Ω_{xx} is injective.
[b.] $\forall \gamma \in \mathbb{R}^{p_{yx}}$, there is no $\phi \in L^2$, such that $\gamma'\mathbf{y}_{t-p_x} = \langle \phi, X_t\rangle$.

Assumption 3.2 [a.] means that Ω_{xx} has no zero eigenvalue. If this assumption was not

satisfied, then equations (7) and (8) would have an infinite number of solutions. We shall discuss this point in more detail at the end of this section.

Assumption 3.2 [b.] means that no linear combination of the components of $\mathbf{y}_{t-p_x}$ can be written as a linear functional of X_t, that is, as $\langle \phi, X_t \rangle$, for a $\phi \in L^2$. It generalizes the assumption made in the discrete time case, namely that the explanatory variables are not collinear. If it were not satisfied, equations (7) and (8) would also have an infinite number of solutions.

The next proposition characterizes the true values of γ and β that satisfy model (4).

Proposition 3.1. *Suppose Assumptions 3.1 and 3.2 hold, and that the operator $\Omega_{xx} : L^2 \to L^2$ has an eigensystem $(\lambda_j, \omega_j)_{j \geqslant 1}$ with $\lambda_j > 0$ for every $j \geqslant 1$. Assume that*

$$\sum_{j \geqslant 1} \frac{\langle \mathbf{E}[\mathbf{y}_{t-i} X_t], \omega_j \rangle^2}{\lambda_j^2} < \infty \qquad \forall i \in \{0, p_x, p_x + 1, \ldots, p_y\}. \tag{9}$$

Then model (4) is identified in the sense that there exists a unique value of (γ, β) in $\mathbb{R}^{p_{yx}} \times L^2$ that satisfies the moment restriction (6). This value is denoted by (γ_0, β_0) and is given by

$$\gamma_0 = \left(\Omega_{yy} - \Omega_{yx}\Omega_{xx}^{-1}\Omega_{yx}^*\right)^{-1} \left(\mathbf{E}[y_t \mathbf{y}_{t-p_x}] - \Omega_{yx}\Omega_{xx}^{-1}\mathbf{E}[y_t X_t]\right) \tag{10}$$

$$\beta_0 = \left(\Omega_{xx} - \Omega_{yx}^*\Omega_{yy}^{-1}\Omega_{yx}\right)^{-1} \left(\mathbf{E}[y_t X_t] - \Omega_{yx}^*\Omega_{yy}^{-1}\mathbf{E}[y_t \mathbf{y}_{t-p_x}]\right). \tag{11}$$

The proof of this result is deferred to the Supplementary Appendix A.1. Condition (9) in Proposition 3.1 is known in the Inverse Problems literature as the Picard condition (see *e.g.* Engl et al. (2000)). It means that $\mathbf{E}[y_{t-i} X_t] \in \mathcal{R}(\Omega_{xx})$ for all $i \in \{0, p_x, \ldots, p_y\}$, where $\mathcal{R}(\cdot)$ denotes the range of an operator, and it guaranties that $\Omega_{xx}^{-1}\mathbf{E}[y_t X_t]$ and $\Omega_{xx}^{-1}\mathbf{E}[y_{t-i} X_t]$ are well-defined in L^2 for $i = p_x, \ldots, p_y$. Therefore, together with Assumptions 3.1 and 3.2, it guarantees the existence and uniqueness of a solution (γ_0, β_0).

Denote by $\breve{X}_t$ the residual of the projection of X_t on the subspace of L^2 spanned by the vector of lags $\mathbf{y}_{t-p_x}$, that is $\breve{X}_t(\cdot) := X_t(\cdot) - \mathbf{EL}[X_t(\cdot)|\mathbf{y}_{t-p_x}]$ [1] and by $\Omega_{xx|y}$ the covariance operator associated with $\breve{X}_t$. Then, it can be shown that $\breve{X}_t = X_t - \Omega_{yx}^*\Omega_{yy}^{-1}\mathbf{y}_{t-p_x}$ and that $\Omega_{xx|y} = \Omega_{xx} - \Omega_{yx}^*\Omega_{yy}^{-1}\Omega_{yx}$. Thus, we can write $\beta_0 = \Omega_{xx|y}^{-1}\mathbf{E}(y_t \breve{X}_t)$ so that both (10) and (11) can be seen as extensions of the Frish-Waugh theorem to our framework. In the framework of local projections, the residual $\breve{X}_t$ plays the role of a "shock" that is given by the residual from the projection of the signal on the control variables (*e.g.* Montiel Olea et al. 2025).

[1] The projection is defined pointwise for any $v \in [0, p_x]$ as $\mathbf{EL}[X_t(v)|\mathbf{y}_{t-p_x}]$ by considering $y_{t-p_x}, \ldots, y_{t-p_y}$ as constant random variables in L^2.

We conclude this section with a discussion about the identification in case Assumption 3.2 [a.] is violated. Since Ω_{xx} is self-adjoint, then L^2 may be decomposed as $L^2 = \mathcal{N}(\Omega_{xx}) + \overline{\mathcal{R}(\Omega_{xx})}$, where $\mathcal{N}(\Omega_{xx})$ is the null space of Ω_{xx}, $\mathcal{R}(\Omega_{xx}) := \{f \in L^2; f = \Omega_{xx}x, x \in L^2\}$ is the range of Ω_{xx}, and $\overline{\mathcal{R}(\Omega_{xx})}$ is the closure in L^2 of $\mathcal{R}(\Omega_{xx})$. If $\mathcal{N}(\Omega_{xx}) \neq \{0\}$, then (6) has an infinite number of solutions in β. More precisely, the β_0 characterized in Proposition 3.1 belongs to $\overline{\mathcal{R}(\Omega_{xx})}$ and it is the best-approximate solution, that is, the least-squares solution of minimal norm. Any function $\beta = \beta_{0,1} + \beta_0$ with $\beta_{0,1} \in \mathcal{N}(\Omega_{xx})$ is a solution of (6). In the following, we focus on recovering (γ_0, β_0). When Ω_{xx} is not injective, we recover a particular solution for (6).

4 Estimation

We first present the estimation procedure, assuming we observe the continuous sample path $\{x(t)\}_{t\in[p_y-p_x,T]}$ without discretization. In practice, $\{x(t)\}$ is only observed at discrete times; however, the sampling frequency is high enough to closely approximate the continuous path. We will present, in Section 4.2, the estimation procedure that takes this discretization into account.

Let us denote $p := p_y$, by $\overline{T} := T - p + 1$ the number of in-sample observations, and by $\widehat{\mathbf{E}}[\cdot] := \sum_{t=p}^{T}[\cdot]/\overline{T}$ the sample average operator based on $\overline{T}$ observations. A naïve estimator for (γ_0, β_0) consists of replacing the population moments in (10) - (11) with empirical moments constructed as follows:

$$\begin{aligned}
\widehat{\Omega}_{yy} &:= \frac{1}{\overline{T}}\sum_{t=p}^{T}\mathbf{y}_{t-p_x}\mathbf{y}'_{t-p_x}, \qquad \widehat{\Omega}_{yx} := \frac{1}{\overline{T}}\sum_{t=p}^{T}\mathbf{y}_{t-p_x}\langle X_t, \cdot\rangle \\
\widehat{\Omega}_{xx} &:= \frac{1}{\overline{T}}\sum_{t=p}^{T}X_t(\cdot)\langle X_t, \cdot\rangle, \qquad \widehat{\Omega}^*_{yx} := \frac{1}{\overline{T}}\sum_{t=p}^{T}X_t\mathbf{y}'_{t-p_x},
\end{aligned}$$

$\widehat{\mathbf{E}}[y_tX_t] := \frac{1}{\overline{T}}\sum_{t=p}^{T} y_tX_t$, and $\widehat{\mathbf{E}}[y_t\mathbf{y}_{t-p_x}] := \frac{1}{\overline{T}}\sum_{t=p}^{T} y_t\mathbf{y}_{t-p_x}$. Unfortunately, this estimator is not consistent because, even if it can be shown that $\widehat{\Omega}_{xx|y} \xrightarrow{p} \Omega_{xx|y}$ under some conditions (*e.g.* Dauxois et al. (1982, Proposition 1) for the i.i.d. case and Bosq (2000) and Hörmann and Kokoszka (2012) for time series), the operator $\Omega_{xx|y}$ is not continuously invertible on L^2 because zero is a limit point of the eigenvalues of $\Omega_{xx|y}$ (see *e.g.* Florens (2003), Crambes et al. (2009), Florens and Simoni (2012, 2016, 2021)). This follows from the fact that Ω_{xx} is compact, and thus, its spectrum accumulates at zero, and $\Omega^*_{yx}\Omega^{-1}_{yy}\Omega_{yx}$ is a finite-rank operator on L^2. Moreover, the sum of a compact operator and a finite-rank operator is

compact. Therefore, estimation of 4 belongs to the class of ill-posed inverse problems. Ill-posedness refers to the fact that a small change in the data may result in a large change in the resulting estimator. Therefore, any consistent estimator of γ_0 and β_0 must involve some form of regularization.

4.1 The penalized least squares (PLS) estimator

We construct our regularized estimator using penalized least squares (PLS) applied to the projections of the signal $X_t(\cdot)$ and β onto the span of basis functions, as in *e.g.* in Ramsay and Silverman (2007) and Li and Hsing (2007). Let $(\phi_j)_{j\geqslant 1}$ be a set of orthonormal basis (o.n.b.) functions of L^2 (for instance B-splines basis functions) and let us express $\beta(\cdot)$ and $X_t(\cdot)$ in terms of their generalized Fourier series as

$$\beta(\cdot) = \sum_{j\geqslant 1} \beta_j \phi_j(\cdot) \qquad \text{and} \qquad X_t(\cdot) = \sum_{j\geqslant 1} c_{t,j}\phi_j(\cdot),$$

where $\beta_j := \langle \beta, \phi_j \rangle$ and $c_{t,j} := \langle X_t, \phi_j \rangle$. In practice, we project β and X_t onto the finite dimensional subspaces of L^2: $\text{span}\{\phi_1, \ldots, \phi_{J_\beta}\}$ and $\text{span}\{\phi_1, \ldots, \phi_{J_x}\}$, respectively, for $J_\beta, J_x \in \mathbb{N}_+$, yielding the projections: $\widetilde{\beta}(\cdot) := \sum_{j=1}^{J_\beta} \beta_j \phi_j(\cdot)$ and $\widetilde{X}_t(\cdot) := \sum_{j=1}^{J_x} c_{t,j}\phi_j(\cdot)$. There is a trade-off between parsimony, which demands small values for J_β and J_x, and approximation, which demands large values for J_β and J_x. The optimal balance between these two goals defines an optimal value for these tuning parameters.

Let $\mathcal{L} : \mathcal{D}(\mathcal{L}) \subseteq L^2 \to L^2$ be a densely defined self-adjoint, strictly positive-definite, unbounded operator (for instance, a differential operator with boundary constraints) suitably chosen for the problem under consideration, where $\mathcal{D}(\cdot)$ denotes the domain of the operator. In particular, $(\phi)_{j\geqslant 1}$ must be such that $\|\mathcal{L}\phi_j\| < \infty$ for every $j \geqslant 1$. For given $J_\beta, J_x \in \mathbb{N}_+$, consider the following penalized sum of squares of residuals from (4) with β and X_t replaced by their projections $\widetilde{\beta}$ and $\widetilde{X}_t$: for every $\beta \in \mathcal{D}(\mathcal{L}) \subset L^2$,

$$\begin{aligned} &C_{PLS}(\gamma, \beta_1, \ldots, \beta_{J_{\min}}; J_\beta, J_x, \alpha_T) \\ &\qquad := \frac{1}{\overline{T}} \sum_{t=p}^{T} \left(y_t - \mathbf{y}'_{t-p_x}\gamma - \sum_{j=1}^{J_\beta}\sum_{j'=1}^{J_x} \beta_j c_{t,j'} \langle \phi_j, \phi_{j'} \rangle \right)^2 + \alpha_T \int_0^{p_x} \left[(\mathcal{L}\widetilde{\beta})(v) \right]^2 dv, \quad (12) \end{aligned}$$

where $\alpha_T > 0$ is a tuning parameter and $J_{\min} := \min\{J_\beta, J_x\}$. The subindex of α_T emphasizes that α_T is a function of T and, in general, it must converge to zero when $T \to \infty$. For

simplicity, take $J_\beta = J_x =: J$. Because the basis functions $(\phi_j)_{j\geqslant 1}$ are orthonormal, the criterion simplifies as follows:

$$C_{PLS}(\gamma, \beta_1, \ldots, \beta_J; J, \alpha_T) = \frac{1}{\overline{T}} \sum_{t=p}^{T} \left(y_t - \mathbf{y}'_{t-p_x}\gamma - \sum_{j=1}^{J} \beta_j c_{t,j} \right)^2 + \alpha_T \int_0^{p_x} \left[(\mathcal{L}\widetilde{\beta})(v) \right]^2 dv. \quad (13)$$

Additionally, the penalty can be expressed as:

$$\int_0^{p_x} \left[(\mathcal{L}\widetilde{\beta})(v) \right]^2 dv = \sum_{j=1}^{J} \sum_{j'=1}^{J} \beta_j \beta_{j'} \int_0^{p_x} (\mathcal{L}\phi_j)(v) \, (\mathcal{L}\phi_{j'})(v) dv.$$

If $\mathcal{L}$ is the negative of the differential operator of order 2 (with boundary conditions, see Example 2 below), then $(\mathcal{L}\phi_j)(v) = -\phi_j^{(2)}(v)$, which is well defined if the basis functions $(\phi_j)_{j\geqslant 1}$ are 2-times differentiable. Since we use the same basis functions for expanding β and x, we are implicitly assuming that both β and x are b-times differentiable. In practice, the choice of the basis functions depends on the nature of the problem and the data, and we do not need to use the same basis for expanding β and x. However, in that case, the simplified expression in (13) no longer holds. In this paper, we will consider basis functions such as Fourier, B-splines, and eigenfunctions of the empirical covariance operator $\widehat{\Omega}_{xx}$. In the latter case, let $\widehat{\omega}_j \in L^2$, and $\widehat{\lambda}_j \geqslant 0$ be such that $\widehat{\Omega}_{xx}\widehat{\omega}_j = \widehat{\lambda}_j \widehat{\omega}_j$, for every $j \geqslant 1$. Thus, $(\widehat{\omega}_j, \widehat{\lambda}_j)_{j\geqslant 1}$ are the eigenfunctions and eigenvalues of $\widehat{\Omega}_{xx}$, respectively, ordered such that $\widehat{\lambda}_1 \geqslant \widehat{\lambda}_2 \geqslant \ldots \geqslant \widehat{\lambda}_j = 0$ for every $j \geqslant T - p + 1$, and $\widehat{\Omega}_{xx} = \sum_{j=1}^{T-p} \widehat{\lambda}_j \widehat{\omega}_j \langle \widehat{\omega}_j, \cdot \rangle$.

The PLS estimator for γ_0 and β_0 is obtained by first solving the following minimization problem, where $p_{yx} := p_y - p_x + 1$:

$$\widehat{\gamma}^P, \widehat{\beta}_1^P, \ldots, \widehat{\beta}_J^P := \text{argmin}_{\gamma \in \mathbb{R}^{p_{yx}}, \beta_1 \in \mathbb{R}, \ldots, \beta_J \in \mathbb{R}} C_{PLS}(\gamma, \beta_1, \ldots, \beta_J; J, \alpha_T), \quad (14)$$

Then, we construct the estimator of the entire function β_0 as follows:

$$\widehat{\beta}^P(\cdot) := \sum_{j=1}^{J} \widehat{\beta}_j^P \phi_j(\cdot) := \Phi_J(\cdot)' \underline{\widehat{\beta}}^P,$$

where $\Phi_J := (\phi_1, \ldots, \phi_J)'$ is a J-vector of L^2-o.n.b. functions and $\underline{\widehat{\beta}}^P := (\widehat{\beta}_1^P, \ldots, \widehat{\beta}_J^P)'$.

Example 1 (Construction of $\mathcal{L}$). A simple way to construct the operator $\mathcal{L}$ defining the penalty is as follows. Let $\mathcal{D}(\mathcal{L})$ be the subspace of L^2 such that for all $\varphi \in \mathcal{D}(\mathcal{L})$, the series $\sum_{j\geqslant 1} \ell_j \langle \varphi, \phi_j \rangle^2$ converges. Consider an increasing sequence $(\ell_j)_{j\geqslant 1}$ such that $\ell_j \to \infty$ as

$j \to \infty$. Then, one possible construction of $\mathcal{L}$ is: $\forall \varphi \in \mathcal{D}(\mathcal{L})$, $\mathcal{L}\varphi = \sum_{j \geqslant 1} \ell_j^{1/2} \langle \varphi, \phi_j \rangle \phi_j$. The range of this operator is contained in L^2 because, $\forall \varphi \in \mathcal{D}(\mathcal{L}) \subset L^2$, $\|\mathcal{L}\varphi\|^2 = \sum_{j \geqslant 1} \ell_j \langle \varphi, \phi_j \rangle^2$ which is bounded under the assumption above.

Example 2 (Differential operator as $\mathcal{L}^2$). Let $\mathcal{D}(\mathcal{L}^2)$ be the subspace of L^2 defined as $\mathcal{D}(L^2) := \{f \in L^2; f' \in L^2, f'' \in L^2, f(0) = f(p_x) = 0\}$, where f' and f'' denote the first and second derivatives of f, respectively. Define the operator $\mathcal{L}^2 : \mathcal{D}(\mathcal{L}^2) \to L^2$ by $\mathcal{L}^2 f := -d^2 f(x)/dx^2$ for every $f \in \mathcal{D}(\mathcal{L}^2)$. By integration by parts, it is straightforward to verify that $\mathcal{L}^2$ is self-adjoint. Its eigenfunctions are $\phi_j(x) = \sqrt{2/p_x} \sin(j\pi x/p_x)$, for $j = 1, 2, \ldots$, with corresponding eigenvalues $\ell_j = (j\pi/p_x)^2$, for $j = 1, 2, \ldots$. Thus, $\mathcal{L}^2$ is positive definite.

For a given o.n.b. $(\phi_j)_{j \geqslant 1}$, we denote by $P_J \cdot := \Phi_J' \langle \Phi_J, \cdot \rangle$ the projection operator from L^2 to V_J, where $V_J := \{P_J f, f \in L^2\} \subset L^2$ is the linear span of Φ_J. For an operator A, we denote by $\langle \Phi_J, A\Phi_J' \rangle$ the $(J \times J)$-matrix with generic element $\langle \phi_k, A\phi_j \rangle$. The next proposition characterizes the first order conditions of the minimization problem (14).

Proposition 4.1. *Let $J_\beta = J_x =: J$. Denote by $\widehat{\widetilde{\Omega}}_{xx}$, $\widehat{\widetilde{\Omega}}_{yx}$ and $\widehat{\widetilde{\Omega}}^*_{yx}$ the estimators $\widehat{\Omega}_{xx}$, $\widehat{\Omega}_{yx}$ and $\widehat{\Omega}^*_{yx}$, respectively, with X_t replaced by $\widetilde{X}_t$. Moreover, denote by Λ the operator that maps $\widehat{\beta}^P$ to $P_J \mathcal{L}^2 \widehat{\beta}^P$. If $\widehat{\Omega}_{yy}$ is invertible, then $(\widehat{\gamma}^P, \widehat{\beta}^P)$ satisfies the following first order conditions:*

$$\begin{aligned}
&\left(\widehat{\mathbf{E}}[y_t \mathbf{y}_{t-p_x}] - \widehat{\mathbf{E}}[\langle \widehat{\beta}^P, \widetilde{X}_t \rangle \mathbf{y}_{t-p_x}]\right) = \widehat{\Omega}_{yy} \widehat{\gamma}^P, \\
&\widehat{\mathbf{E}}\left[y_t \widetilde{X}_t\right] - \widehat{\mathbf{E}}[y_t \mathbf{y}'_{t-p_x}] \widehat{\Omega}_{yy}^{-1} \widehat{\mathbf{E}}\left[\widetilde{X}_t \mathbf{y}_{t-p_x}\right] = \left(\widehat{\widetilde{\Omega}}_{xx} - \widehat{\widetilde{\Omega}}^*_{yx} \widehat{\Omega}_{yy}^{-1} \widehat{\widetilde{\Omega}}_{yx}\right) \widehat{\beta}^P + \alpha_T \Lambda \widehat{\beta}^P. \qquad (15)
\end{aligned}$$

Equivalently,

$$\begin{aligned}
&\left(\widehat{\mathbf{E}}[y_t \mathbf{y}_{t-p_x}] - \widehat{\mathbf{E}}[\langle \widehat{\beta}^P, X_t \rangle \mathbf{y}_{t-p_x}]\right) = \widehat{\Omega}_{yy} \widehat{\gamma}^P, \\
&P_J \left(\widehat{\mathbf{E}}\left[y_t X_t\right] - \widehat{\mathbf{E}}\left[X_t \mathbf{y}'_{t-p_x}\right] \widehat{\Omega}_{yy}^{-1} \widehat{\mathbf{E}}[y_t \mathbf{y}_{t-p_x}]\right) = P_J \left(\widehat{\Omega}_{xx} - \widehat{\Omega}^*_{yx} \widehat{\Omega}_{yy}^{-1} \widehat{\Omega}_{yx} + \alpha_T \mathcal{L}^2\right) \widehat{\beta}^P. \qquad (16)
\end{aligned}$$

A proof of this result is provided in the Supplementary Appendix A.2. In the following, we denote $\widehat{\Omega}_{xx|y} := \widehat{\Omega}_{xx} - \widehat{\Omega}^*_{yx} \widehat{\Omega}_{yy}^{-1} \widehat{\Omega}_{yx}$ as the empirical counterpart of $\Omega_{xx|y}$. Let $\mathbf{X} := (X_p, \ldots, X_T)'$ be $\overline{T}$-dimensional, $\mathbf{y} := (y_p, \ldots, y_T)' \in \mathbb{R}^{\overline{T}}$, and $\mathbf{y}_{-p_x} := (\mathbf{y}_{p-p_x}, \ldots, \mathbf{y}_{T-p_x})'$ be a $(\overline{T} \times p_{yx})$-matrix of lagged dependent variables. Moreover, denote by $P_y : \mathbb{R}^{\overline{T}} \to \mathbb{R}^{\overline{T}}$ the projection operator $\mathbf{y}_{-p_x} \widehat{\Omega}_{yy}^{-1} \mathbf{y}'_{-p_x} / \overline{T}$. We can then write the left hand side of the second

equation in (16) as:

$$P_J \frac{\mathbf{X}'(I-P_y)\mathbf{y}}{\overline{T}} = \Phi'_J \left\langle \Phi_J, \frac{\mathbf{X}'(I-P_y)\mathbf{y}}{\overline{T}} \right\rangle = \Phi'_J \frac{\underline{c}'(I-P_y)\mathbf{y}}{\overline{T}},$$

where $\underline{c} := \langle \mathbf{X}, \Phi'_J \rangle$ is the $(\overline{T} \times J)$-matrix with (t,k) element $c_{t,k}$ and $P_J \mathbf{X}' = \Phi'_J \underline{c}'$. The right hand side of the second equation in (16) can be written as:

$$P_J \left(\widehat{\Omega}_{xx} - \widehat{\Omega}_{yx}^* \widehat{\Omega}_{yy}^{-1} \widehat{\Omega}_{yx} + \alpha_T \mathcal{L}^2 \right) \widehat{\beta}^P =: \Phi'_J \langle \Phi_J, \left(\widehat{\Omega}_{xx|y} + \alpha_T \mathcal{L}^2 \right) \Phi'_J \underline{\widehat{\beta}}^P \rangle.$$

By replacing the left and right hand sides of the second equation in (16) with these last expressions, by taking the inner product of the left and right hand sides with Φ_J, and then solving for $\underline{\widehat{\beta}}^P$, we get:

$$\underline{\widehat{\beta}}^P = \left(\langle \Phi_J, (\widehat{\Omega}_{xx|y} + \alpha_T \mathcal{L}^2) \Phi'_J \rangle \right)^{-1} \frac{\underline{c}'(I-P_y)\mathbf{y}}{\overline{T}}.$$

4.2 Discretization

In practice, the stochastic process $x(\cdot)$ is observed only at a finite number of dates. Let K be the number of observation times for $x(\cdot)$ in the interval $(t-p_x, t]$, and denote these observation times as $t_1 < t_2 < \ldots < t_{K-1} < t_K \equiv t$. Let m denote the number of times the higher sampling frequency appears in the low-frequency time unit t, that is, in the interval $(t-1, t]$. Hence, in the interval $(t-1, t]$ the stochastic process $x(\cdot)$ is observed at the high-frequency time points: $t-1+1/m, t-1+2/m, \ldots, t-1+(m-1)/m, t$. Extending this to the entire interval $(t-p_x, t]$, the high-frequency observation times are $t_1 = t-p_x+1/m$, $t_2 = t-p_x+2/m$, $t_3 = t-p_x+3/m$, $\ldots$, $t_m = t-p_x+1$, $\ldots$, $t_{K-1} = t-1+(m-1)/m$, $t_K = t$. Since there are m observations per unit interval and p_x unit intervals, the total number of observations is $K = m(p_x-1)+m = mp_x$. For instance, if $p_x = 2$ and $m = 3$ then we have $K = 6$, and the observation times in $(t-2, t]$ are: $t-2+1/3, t-2+2/3, t-1, t-1+1/3, t-1+2/3, t$.

Let us denote by $\mathfrak{X}_t$ the following K-dimensional vector of observations (with $t_K \equiv t$):

$$\mathfrak{X}_t := \begin{pmatrix} x(t_K) \\ x(t_{K-1}) \\ \vdots \\ x(t_{m+1}) \\ \vdots \\ x(t_1) \end{pmatrix} = \begin{pmatrix} X_t(0) \\ X_t(1/m) \\ \vdots \\ X_t(p_x - 1) \\ \vdots \\ X_t(p_x - 1/m) \end{pmatrix}.$$

The K-vector arising from applying this discretization to β is denoted by β_K, that is, $\beta_K := (\beta(0), \ldots, \beta(p_x - 1/m))'$ and the corresponding discretized true value is denoted by $\beta_{0,K}$.

Remark. *With this discretization, one could use the alternative approach that consists of treating $X_t(0), \ldots, X_t(p_x - 1/m)$ as K real-valued covariates and then applying least squares to estimate the corresponding coefficient vector β_K. However, this approach has two major drawbacks. First, the K covariates exhibit a high degree of collinearity. Second, as m increases, the number of covariates $K = mp_x$ grows rapidly. Together, these issues render the problem ill-conditioned, causing the least squares solution to perform poorly. While a regularization procedure might mitigate these issues, it fails to account for the functional nature of the data, specifically the natural ordering of the time points $t_1, \ldots, t_K$. See Marx and Eilers (1999) and Cardot et al. (2003) for further discussions on this point.*

The PLS estimator introduced above has to be modified in order to take into account the discretization. For this, we first replace the inner product in L^2 with its discretized version, denoted by $\langle \cdot, \cdot \rangle_K$. Let $\{t_k\}_{k=1}^K$ be the discretization points in the interval $(t - p_x, t]$, defined as: $t_j := t - p_x + j/m$, $j = 1, \ldots, K$, where $K = mp_x$. The discretized inner product is then defined for every $f, g \in L^2$ as

$$\langle f, g \rangle_K = \frac{p_x}{K} \sum_{k=1}^{K} f(t_k) g(v_k) = \frac{1}{m} \sum_{k=1}^{K} f(t_k) g(t_k).$$

This approximation is sufficiently accurate only if m, and then K, is large enough. Therefore, we assume to be in this case from now on, which means that we consider cases where the difference between the high and the low sampling frequencies is sufficiently large. The

discretized version of model (4) is:

$$y_t = \mathbf{y}'_{t-p_x}\gamma + \frac{1}{m}\beta'_K\mathfrak{X}_t + \widetilde{\varepsilon}_t, \qquad \widetilde{\varepsilon}_t := \varepsilon_t + \Delta_t \tag{17}$$

and $\Delta_t := \langle \beta, X_t\rangle - \frac{1}{m}\beta'_K\mathfrak{X}_t$ is the discretization error, see Crambes et al. (2009). The discretized empirical cross moments are denoted by adding a subscript 'K'.

$$\begin{aligned}
\widehat{\Omega}_{yx,K} &:= \frac{1}{\overline{T}}\sum_{t=p}^{\overline{T}} \mathbf{y}_{t-p_x}\frac{\mathfrak{X}'_t}{m} : \mathbb{R}^K \to \mathbb{R}^{p_{yx}},\\
\widehat{\Omega}_{xx,K} &:= \frac{1}{\overline{T}}\sum_{t=p}^{\overline{T}} \mathfrak{X}_{t-h}\frac{\mathfrak{X}'_t}{m} : \mathbb{R}^K \to \mathbb{R}^K, \qquad \widehat{\Omega}^*_{yx,K} := \frac{1}{\overline{T}}\sum_{t=p}^{\overline{T}} \mathfrak{X}_{t-h}\mathbf{y}'_{t-p_x} : \mathbb{R}^{p_{yx}} \to \mathbb{R}^K,
\end{aligned}$$

$\widehat{\mathbf{E}}[y_t\mathfrak{X}_t] := \frac{1}{\overline{T}}\sum_{t=p}^{\overline{T}} y_t\mathfrak{X}_t$. The discretized PLS estimator for γ_0 and β_0, denoted by $\widehat{\gamma}^P_K \in \mathbb{R}^{p_{yx}}$ and $\widehat{\beta}^P_K \in \mathbb{R}^K$, is obtained by minimizing the discretized criterion $C_{PLS,K}(\gamma, \beta_1, \ldots, \beta_J; J, \alpha_T)$ defined as (for $(\mathcal{L}\phi_j)(v) = \phi_j^{(b)}(v)$):

$$\begin{aligned}
&C_{PLS,K}(\gamma, \beta_1, \ldots, \beta_J; J, \alpha_T)\\
&\qquad := \frac{1}{\overline{T}}\sum_{t=p}^{\overline{T}}\left(y_t - \mathbf{y}'_{t-p_x}\gamma - \sum_{j=1}^{J}\beta_j c_{t,j}\right)^2 + \frac{\alpha_T}{m}\sum_{j=1}^{J}\sum_{j'=1}^{J}\beta_j\beta_{j'}\boldsymbol{\phi}_j^{(\ell)'}\boldsymbol{\phi}_{j'}^{(\ell)},
\end{aligned}$$

where $\boldsymbol{\phi}_j^{(\ell)} := (\phi_j^{(\ell)}(0), \ldots, \phi_j^{(\ell)}(p_x - 1/m))'$ is a K-dimensional vector.

5 Theoretical properties

In this section we study asymptotic properties of $\widehat{\beta}^P$ as the number of in-sample observations $\overline{T}$ increases. We denote by $\|\cdot\|_{op}$ the operator norm and by $\|\cdot\|_2$ the Euclidean norm. The symbols $\asymp$, $\lesssim$ and $\gtrsim$ denote equality and inequalities up to a constant.

Recall the notation $\Omega_{xx|y} := \Omega_{xx} - \Omega^*_{yx}\Omega^{-1}_{yy}\Omega_{yx}$ which is the covariance operator of the residualized X_t (after projecting it on $\mathbf{y}_{t-p_x}$):

$$\Omega_{xx|y}\cdot := \mathbf{E}[X_t\langle X_t, \cdot\rangle] - \mathbf{E}[X_t\mathbf{y}_{t-p_x}]\Omega^{-1}_{yy}\mathbf{E}[\mathbf{y}_{t-p_x}\langle X_t, \cdot\rangle] : L^2 \to L^2$$

which is bounded.

To investigate the asymptotic behaviour of the penalized least squares estimator $\widehat{\beta}^P$, we recall that the penalty is induced by a densely defined self-adjoint, strictly positive-

definite, and unbounded operator $\mathcal{L}^2 : \mathcal{D}(\mathcal{L}^2) \subseteq L^2 \to L^2$ (such as a differential operator with boundary constraints). We assume that the o.n.b. $(\phi_j)_{j\geqslant 1}$ of L^2 used to construct $\widehat{\beta}^P$ consists of the eigenfunctions of $\mathcal{L}^2$. We denote by $(\ell_j)_{j\geqslant 1}$ the corresponding eigenvalues and they are such that $\ell_j \to \infty$. We also recall the definition given in Proposition 4.1 of $P_J \cdot := \sum_{j=1}^{J} \phi_j \langle \phi_j, \cdot \rangle$ which is the orthogonal projection onto V_J. Finally, we use the notation $\Omega_{JJ,\alpha_T} := \langle \Phi_J, (\Omega_{xx|y} + \alpha_T \mathcal{L}^2)\Phi'_J \rangle$.

We start by introducing the following assumption which imposes regularity conditions on the process $\{X_t\}$ and the errors ε_t.

Assumption 5.1. *The process $\{X_t\}_{t\in\mathbb{Z}}$ is stationary and β-mixing with mixing coefficients $\beta(k) \to 0$ as $k \to \infty$ at a polynomial or exponential rate, i.e., $\sum_{k=1}^{\infty} k^{\delta}\beta(k) < \infty$ for some $\delta > 0$. Moreover, $\mathbf{E}[\varepsilon_t^2] < \infty$, and $\mathbf{E}[\varepsilon_t \varepsilon_\tau] < \infty$, for every $t \neq \tau$.*

This assumption ensures asymptotic independence and finite second moments, which are crucial for our asymptotic analysis. The next assumption quantifies the smoothness of the true parameter β_0 in terms of the regularity of the (integral) operator $\mathcal{L}^{-1}$.

Assumption 5.2. *For some $\eta \geqslant 0$, there exists a $\omega_0 \in L^2$ such that $\beta_0 = \mathcal{L}^{-\eta}\omega_0$.*

The parameter η characterizes the regularity of the true function β_0 and is generally unknown. Assumption 5.2 is satisfied by regular functions β_0. For instance, if $\mathcal{L}$ is the b-differential operator, then Assumption 5.2 is assuming that β_0 has at least $b\eta$ square integrable derivatives.

Example 1 (continued). Suppose that the operator $\mathcal{L}$ is constructed as in Example 1. Then, Assumption 5.2 means that for some $\eta \geqslant 0$, there exists a $\omega_0 \in L^2$ such that $\beta_0 = \sum_{j\geqslant 1} \ell_j^{-\eta/2} \langle \phi_j, \omega_0 \rangle \phi_j$. Assumption 5.2 can also be expressed equivalently by requiring that $\sum_{j\geqslant 1} \ell_j^{\eta} \langle \beta_0, \phi_j \rangle^2 < \infty$. This is a stronger condition than simply requiring $\sum_{j\geqslant 1} \langle \beta_0, \phi_j \rangle^2 < \infty$, as it imposes that the sequence of coefficients $(\langle \beta_0, \phi_j \rangle)_{j\geqslant 1}$ decays fast enough to compensate for the growth of $(\ell_j)_{j\geqslant 1}$.

Assumption 5.2 is classical in Inverse Problems literature, *e.g.* Chen and Reiss (2011) and Florens and Simoni (2016), and it is similar to the *source condition* which expresses the regularity of β_0 according to the spectral representation of $\Omega_{xx|y}$ defining the Inverse Problem (11) (see *e.g.*Engl et al. 2000). Here, we express the regularity of β_0 according with the spectral representation of $\mathcal{L}$ and the latter is linked to the smoothness of the operator $\Omega_{xx|y}$ through Assumption 5.3 below which makes this link explicit.

Assumption 5.3. *Let* $B := \Omega_{xx|y}^{1/2}\mathcal{L}^{-1}$ *and* $[diag(\ell)]_J := diag(\ell_1, \ldots, \ell_J)$ *for* $J \in \mathbb{N}_+$. *There are constants* $0 < d < D < \infty$ *such that: for every* $|\nu| \leqslant 1$,

$$d \sup_{J\in\mathbb{N}_+} \left\|[diag(\ell^{-\nu(a+1)})]_J\, h\right\|_2 \leqslant \left\|(\langle \Phi_J, B^*B\Phi_J'\rangle)^{\nu}\, h\right\|_2 \leqslant D \sup_{J\in\mathbb{N}_+} \left\|[diag(\ell^{-\nu(a+1)})]_J\, h\right\|_2 \quad (18)$$

holds for every vector $h \in \mathbb{R}^J$ *such that* $\left\|[diag(\ell)^{-\nu(a+1)}]_J\, h\right\|_2 < \infty$.

This assumption is similar to Johannes and Schwarz (2010, Assumption 2.7) and to Chen and Reiss (2011, Assumptions 2 and 5). It establishes a norm equivalence between a weighted ℓ^2-norm and a norm induced by the positive-definite matrix $(\langle \Phi_J, B^*B\Phi_J'\rangle)^{\nu}$ (which in turn is induced by the operator B^*B). We refer to the parameter a in Assumption 5.3 as the *degree of ill-posedness* of the estimation problem under study and a is determined by the rate of decreasing of the spectrum of $\Omega_{xx|y}^{1/2}$ with respect to the rate of decreasing of the spectrum of $\mathcal{L}^{-1}$. This assumption is used to determine the rate of convergence of the PLS estimator.

Theorem 5.1. *Let the basis functions* $(\phi_j)_{j\geqslant 1}$ *be the eigenfunctions of the operator* $\mathcal{L} : \mathcal{D}(\mathcal{L}) \to L^2$. *Suppose that Assumptions 3.1 - 5.3 hold. Then,*

$$\|\widehat{\beta}^P - \beta_0\| \asymp \mathcal{O}_p\left(\alpha_T^{-2a/(a+1)}\sqrt{\frac{J}{\overline{T}}}\right) + \ell_J^{-\eta/2}\alpha_T^{-a/(2(a+1))}.$$

In the next theorem we strengthen our assumptions to improve the rate of convergence and assume that $\mathcal{L}^2$ and $\Omega_{xx|y}$ are simultaneously diagonalizable. This assumption means that they share the same eigenfunctions, that is, there exists a positive sequence $(\mu_j)_{j\geqslant 1}$ such that $\sum_{j=1}^{+\infty}\mu_j < \infty$ and for every $j \geqslant 1$,

$$\Omega_{xx|y}\phi_j = \mu_j\phi_j,$$

where $(\phi_j)_{j\geqslant 1}$ are such that $\mathcal{L}^2\phi_j = \ell_j\phi_j$ for every $j \geqslant 1$. This is equivalent to $\Omega_{xx|y}$ commuting with $\mathcal{L}^2$: $\Omega_{xx|y}\mathcal{L}^2 = \mathcal{L}^2\Omega_{xx|y}$. For instance, one can choose the operator $\mathcal{L}$ for the penalty such that $\mathcal{L} = \Omega_{xx|y}^s$, for some $s > 0$. This guarantees simultaneous diagonalizability and it is justified by the fact that the penalization is adapted to the covariance structure of the covariate $\widetilde{X}_t$. In this case, the penalty $\alpha_T\|\mathcal{L}\beta\|^2$ penalizes directions of low variance of X_t more heavily and Assumption 5.3 means that $\mu_j \asymp \ell_j^{-a}$. We then have the following theorem.

Theorem 5.2. *Let the operators $\Omega_{xx|y} : L^2 \to L^2$ and $\mathcal{L}^2 : \mathcal{D}(\mathcal{L}) \to L^2$ have the same eigenfunctions $(\phi_j)_{j\geqslant 1}$. Suppose that Assumptions 3.1 - 5.3 hold with $\eta > 1$. Then,*

$$\mathbf{E}\|\widehat{\beta}^P - \beta_0\|^2 \asymp \frac{J}{\overline{T}}\alpha_T^{-a/(a+1)} + \left(\ell_J^{-\eta} + \alpha_T^{\widetilde{\eta}/(a+1)}\right)\left(1 + \alpha_T^{-(2a+1)/(a+1)}\frac{1}{\overline{T}}\right) \tag{19}$$

where $\widetilde{\eta} := \min\{\eta, 2(a+1)\}$.

The rate is made of three term. The first term accounts for the variance term, the second term in the rate $\left(\ell_J^{-\eta} + \alpha_T^{\widetilde{\eta}/(a+1)}\right)$ is due to the two biases introduced by the penalty and the truncation, while the last rate $\left(\ell_J^{-\eta} + \alpha_T^{\widetilde{\eta}/(a+1)}\right)\alpha_T^{-(2a+1)/(a+1)}\frac{1}{\overline{T}}$ is the rate of the estimation error. Notice that the last rate is negligible if $\alpha_T \asymp T^{-(a+1)/(2a+1)}$.

The optimal rate is obtained by first choosing α_T in such a way to balance the two terms in the second rate, that is, $\alpha_T \asymp \alpha_T^* := \ell_J^{-\eta(a+1)/\widetilde{\eta}}$. In the mildly ill-posed case: $\ell_j \asymp j^\zeta$ for some $\zeta > 0$ and for every $j \geqslant 1$, so that the optimal α_T is $\alpha_T^* := J^{-\zeta\eta(a+1)/\widetilde{\eta}}$. By equating with the first term, we get the optimal J:

$$J^* \asymp \begin{cases} \overline{T}^{1/(1+\zeta(a+2))} & \text{if } \eta < 2(a+1) \\ \overline{T}^{2(a+1)/(1+\zeta\eta(a+2))} & \text{if } \eta \geqslant 2(a+1). \end{cases}$$

The corresponding rate is:

$$\mathbf{E}\|\widehat{\beta}^P - \beta_0\|^2 \begin{cases} \overline{T}^{-\eta/((a+2+1/(\eta\zeta)))} & \text{if } (a+1) - \frac{1}{\zeta} \leqslant \eta < 2(a+1) \\ \overline{T}^{-2(a+1)/((a+2)+1/(\zeta\eta))} & \text{if } \eta \geqslant 2(a+1). \end{cases}$$

6 Application: Inference on CT-IRFs to high-frequency uncertainty shocks

There has recently been a strong focus in the news on a sudden and sharp increase in economic and financial uncertainty around the globe, related to unexpected changes in the global trade system initiated by the U.S. administration and to increasing geopolitical tensions . The concept of uncertainty is not new and dates back to the seminal work by Knight (1921), who suggests that "uncertainty must be taken in a sense radically distinct from the familiar notion of risk". Uncertainty exists when people cannot assign probabilities to imaginable outcomes. This differs from the concept of risk, for which we assume we can assign a given probability distribution, even when there are fat tails and asymmetries.

The literature on the macroeconomic effects of uncertainty shocks is extensive and dy-

namic, generally converging on the notion that uncertainty shocks lead to short- to medium-term adverse effects on economic activity, followed by a rebound (see the review by Bloom 2014). For instance, "uncertainty about the economic outlook can lead firms to defer spending projects until prospects for economic activity will become clearer" (Bloom 2009). A standard framework for assessing how fluctuations in uncertainty affect the economy is provided by irreversible investments, as discussed in the seminal contribution of Bernanke (1983). The basic idea is that when investment projects are irreversible—that is, they cannot be "canceled" or "modified" without incurring very high costs—there exists a trade-off for investors between additional returns from the immediate launch of an investment project and the benefits of waiting to gather more information in the future. The value of waiting is described in the literature as *real-option value*. Against this background, a rise in uncertainty would clearly tilt the balance in favor of a wait-and-see approach. From an econometric perspective, assessing the macroeconomic impact of uncertainty shocks has generally been conducted via SVAR models, as exemplified in the influential paper by Bloom (2009). As is standard in SVAR models, identifying the uncertainty shock can sometimes be challenging (see Caldara et al. (2016), who propose a penalty function to disentangle uncertainty from financial shocks). Strikingly, econometric approaches are typically conducted at the same frequency, either quarterly or monthly. One notable exception is the paper by Ferrara and Guérin (2018), which attempts to assess the impact of high-frequency uncertainty shocks measured by financial volatility (the VIX measure) and economic policy uncertainty (the EPU index developed by Baker et al. 2016), utilizing SVAR models.

In this paper, we pursue this line of research by proposing a new method to estimate impulse response functions of low-frequency macroeconomic variables in response to high-frequency uncertainty shocks. In practice, we focus on uncertainty shocks estimated at a daily frequency, and we aim to assess their impact on business investment. Empirical results from the literature tend to indicate that the most irreversible investment categories, such as infrastructure or equipment investments, respond most negatively to uncertainty shocks compared to, for example, investments in intellectual property products (see Ferrara and Guérin 2018). In this regard, we focus in this empirical exercise on the response of quarterly business investment in equipment for the U.S. economy to high-frequency uncertainty shocks, as proxied by the Economic Policy Uncertainty (EPU) daily index computed by Baker et al. (2016). This index is constructed by reflecting the frequency of articles in ten leading US newspapers that simultaneously contain at least one word referring to the economy (*e.g.*, *economy* or *economics*), one word related to policy (*e.g.*, *deficit*, *central bank*, or *taxes*), and

one word indicating uncertainty (e.g., *uncertain* or *uncertainty*). After some normalization steps, an index is computed, allowing for comparison over time.

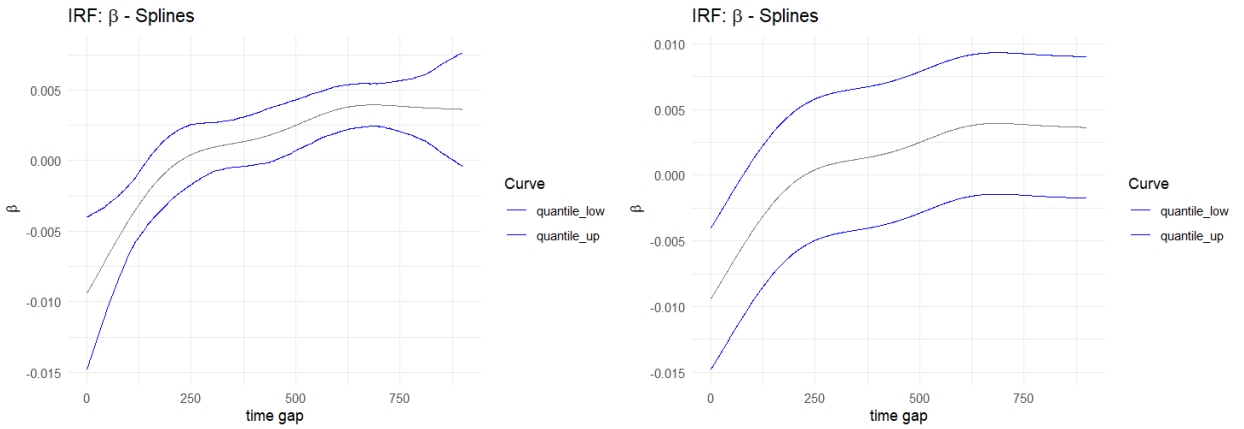
Empirical results arising from the local projection of the quarterly growth rate of investment in equipment on the daily EPU index are presented in Figure 1. Both graphs are obtained using the PLS estimator based on B-splines with $p_y = 11$ and $p_x = 10$. In order to statistically interpret the CT-IRFs, we also compute 68% confidence bands around the response based on bootstrap resampling of residuals from the local projection (see details in the Supplementary Appendix). Two strategies are implemented to compute the confidence bands: (i) pointwise (left panel of Figure 1) and (ii) uniform (right panel of Figure 1). All variables are standardized.

From Figure 1, we observe that the growth of equipment investment responds negatively to a daily economic policy uncertainty shock upon impact, with an amplitude of approximately 0.09%. Both methods for computing confidence bands point to a significant drop upon impact. Following this initial impact, the shape of the IRF is very similar to the standard response to an uncertainty shock on investment, as highlighted by Bloom (2009). Indeed, the initial drop upon impact is followed by a bounce-back after about 250 days, when the growth rate turns positive. However, this rebound does not appear to be statistically significant.

These results are very much in line with standard responses of investment to an uncertainty shock, keeping in mind that we observe here the response to a *daily* uncertainty shock, while usual results from the literature relate to *monthly* or *quarterly* shocks, which are an average of daily shocks over the month or quarter. Thus, it makes sense to observe a much lower amplitude in the responses to a one standard deviation shock. From a policy perspective, we highlight that an uncertainty shock, even if created in a single day, is likely to produce negative effects on equipment investment growth. Thus, policy-makers should be extremely careful in their decisions and communication to avoid creating large uncertainties, even in the short run.

7 Application: Inference on macroeconomic CT-IRFs to high-frequency financial shocks

In this section, we focus on the effects of high-frequency financial shocks on economic activity. In the wake of the global financial crisis that led to a world economic recession in $2008 - 09$, many theoretical and empirical papers have highlighted the significant impact of



(a) Pointwise (b) Uniform

Figure 1: *CT-IRF*. Estimation of the $\beta(\cdot)$ function from the local projection of the quarterly growth rate of investment in equipment on the daily EPU index. PLS estimator with B-splines. Pointwise and uniform 68%-confidence bands obtained with 1000 Bootstrap samples.

financial shocks on macroeconomic cycles. In this respect, an influential paper is the one put forward by Adrian et al. (2009), in which the authors propose forecasting the distribution of GDP growth over a horizon by using both past GDP values and a measure of financial conditions. They illustrate that a tightening of financial conditions tends to amplify the effects of negative shocks to the real economy, as notably emphasized in theory by Bernanke and Gertler (1989).

One limitation of standard macro-financial models aiming to assess such effects is that data with a higher frequency are generally aggregated at the lowest frequency. For example, if quarterly GDP growth is the target of a forecasting model, then all variables involved in the model are sampled at the same quarterly frequency. This approach implicitly discards valuable within-period information, especially for financial indicators that are generally available on a high-frequency basis (intra-day or daily). Such data aggregation is likely to lead to biased estimates if the underlying data generating process does not feature a flat aggregation scheme from high to low frequencies. In this respect, mixed-frequency models have recently been proposed to optimally account for financial dynamics. Various approaches show that higher-frequency financial series are helpful for more accurately assessing economic conditions compared to various benchmark models and survey forecasts, as evidenced by Clements and Galvão (2009), Andreou et al. (2013), Ghysels (2016), Carriero et al. (2020), or Ferrara et al. (2022).

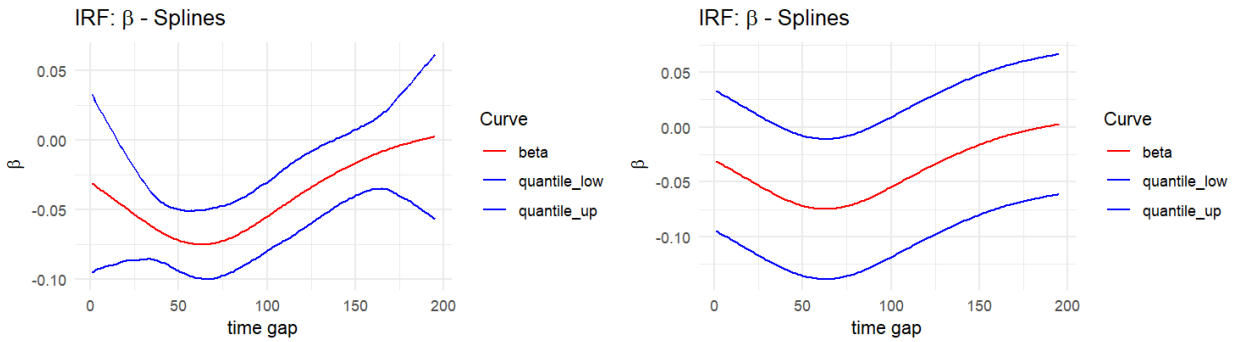
Here, we assess to what extent a shock in a high-frequency financial conditions index can affect U.S. business cycles. In this respect, we estimate our model to compute the CT-

IRF of the quarterly GDP growth rate in response to a daily financial shock, as measured by the Composite Indicator of Systemic Stress (CISS) developed by the European Central Bank (Hollo et al. (2012)). This index is the aggregation of five market-specific sub-indexes, namely the foreign exchange market, the equity market, the money market, the bond market, and financial intermediaries. The aggregation procedure accounts for time-varying cross-correlations between the five sub-indexes, thereby allowing the CISS index to place greater weight on situations in which stress prevails in several market segments simultaneously. The CISS index can be seen as a common factor of tensions in several markets, providing a measure of systemic financial stress that is likely to be damaging to economic activity (see Ferrara et al. (2022) for an application to the euro area).

In our application, we use CISS data from January 1st, 1980, to September 30, 2025 for the United States, available at a daily frequency. The U.S. quarterly GDP growth rate ranges from $1980q1$ to $2025q3$. In our modeling approach, we project current GDP onto the lagged GDP, as well as onto the past quarters for the CISS. The estimate of the CT-IRFs of GDP growth to the CISS index is presented in Figure 2 together with bootstrapped pointwise and uniform 68% confidence bands around the response. The estimation is obtained using the PLS estimator based on B-Splines basis functions with $p_y = 5$ and $p_x = 2$. All variables are standardized.

When looking at the CT-IRFs presented in the left panel of Figure 2 (Pointwise confidence bounds), we note that the impact is not statistically different from zero until approximately 15 business days after the shock occurs, highlighting the rapid economic response to financial stress. This fast response is likely to occur through expectations and confidence that rapidly decline in the wake of a financial crisis. The trough of the CT-IRF is reached approximately 60 business days (3 months) after the initial impact, at -0.075%. Afterward, the effect gradually dissipates and returns to near baseline after 140 days; that is about 7 months. When computing confidence bounds using the Uniform strategy (right panel of Figure 2), it turns out that the effect becomes significant after 38 days (instead of 15 days using the Pointwise strategy) and dissipates after 90 days (instead of 140). Obviously, the dynamic response of the conditional mean remains unchanged.

These dynamics mirror recession patterns following financial stress episodes, suggesting that a financial shock occurring at the beginning of the quarter will have a fully negative macroeconomic impact within that specific quarter. However, if the shock occurs in the middle of a quarter, the effect will be divided between the current and subsequent quarters.



(a) Pointwise (b) Uniform

Figure 2: *CT-IRF*. Estimation of the $\beta(\cdot)$ function from the local projection of the quarterly U.S. GDP growth rate on the CISS index. PLS estimator with B-Splines basis functions. Uniform 68%-confidence bands obtained with 1000 Bootstrap samples.

8 Simulations

In this section, we study the finite-sample performance of our approach to make inference on the CT-IRF and compare different basis functions to construct the PLS estimator. We also compare the performance of our estimators with the Functional Principal Component (FPC) estimator described in Doz et al. (2025), as well as with the MIDAS approach, which accommodates both Almon and Beta polynomial functions.

The first data generating process (DGP) that we consider is as follows. For every $t = 1, \dots, T$,

$$y_t = \int_0^{p_x} \beta_0(v) X_t(v) dv + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{20}$$

where $\sigma_\varepsilon^2 > 0$ and $\beta_0(\cdot)$ denotes the true beta function. We consider two specifications for $\beta_0(\cdot)$: smooth and piecewise. In the *smooth case*, $\beta_0(\cdot)$ is proportional to the density function of a Generalized beta distribution of the first kind: for every $a, b > 0$,

$$\beta_0(v; a, b) = \left(\frac{v}{p}\right)^{a-1} \left(1 - \frac{v}{p_x}\right)^{b-1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mathbb{1}\{v \in [0, p_x]\}$$

which has support $(0, p_x)$. In the *piecewise case*, $\beta_0(\cdot)$ is generated as $\beta_0(v) := \left[\left(\frac{v}{p_x} - 0.5\right)^2 - 0.025\right] * 100$ for $v \in [0, 0.342 p_x)$, $\beta_0(v) := \left[-\left(\frac{v}{p_x} - 0.5\right)^2 + 0.025\right] * 100$ for $v \in (0.658 p_x, p_x]$, and zero otherwise.

The covariate $x(v)$ follows an Ornstein-Uhlenbeck (OU) process, that is:

$$dx(v) = -\theta x(v)dv + \sigma_x dW(v), \qquad \theta > 0, \quad \sigma_x > 0, \tag{21}$$

and $W(v)$ denotes a Wiener process. We then discretize the integral in (21) by using the solution of the OU process, which is $x(v) = x(s)e^{-\theta(v-s)} + \sigma_x \int_s^v e^{-\theta(v-u)}dW(u)$, $\forall 0 \leqslant s \leqslant v$. By the properties of the Itô integral with respect to Brownian motion (see *e.g.* Klebaner 2012, Theorem 4.11), we know that $\int_s^v e^{-\theta(v-u)}dW(u)$ is a Gaussian process with zero mean and a covariance function given by

$$Cov\left(\int_s^v e^{-\theta(v-u)}dW(u), \int_s^{v+\ell} e^{-\theta(v-u)}dW(u)\right) = \int_s^v e^{-2\theta(v-u)}du = \frac{1}{2\theta}(1 - e^{-2\theta(v-s)}).$$

The process $x(\cdot)$ is observed at K points $t_1 < t_2 < \ldots < t_K$ in the interval $(t - p_x, t]$, as described in Section 4.2. To generate the data, we take a fine grid with $g * K$ equispaced points in the interval $(t - p_x, t]$: $t_{0,1} < t_{0,2} < \ldots < t_{0,g-1} < t_{0,g} \equiv t_{1,0} \equiv t_1 < t_{1,1} < \ldots < t_K$. For instance, with $g = 3$, the points in the fine grid are: $t_{0,1} < t_{0,2} < t_{1,0} \equiv t_1 < t_{1,1} < t_{1,2} < t_{2,0} \equiv t_2 < \ldots < t_{K-1,2} < t_K$. If we denote by Δ the difference between two of these points, then for every $i = 0, 1, 2$ and each $j = 0, 1, \ldots, K$ and each $t = 1, 2, \ldots$

$$x(t_{j,i}) = e^{-\theta\Delta}x(t_{j,i-1}) + u_{j,i}, \qquad u_{j,i} \sim \mathcal{N}\left(0, \frac{1}{2\theta}\sigma_x^2(1 - e^{-2\theta\Delta})\right)$$

which is a stationary $AR(1)$ whose stationary distribution has a zero mean and a variance equal to $\frac{1}{2\theta}\sigma_x^2$. Therefore, $X_0 \sim \mathcal{N}(0, \sigma_x^2/(2\theta))$. Notice that as θ increases, the autocovariance of $x(t)$ decreases and approaches that of white noise. Equivalently, we can write

$$X_t(t - t_{j,i}) = e^{-\theta\Delta}X_t(t - t_{j,i-1}) + u_{j,i}, \qquad u_{j,i} \sim \mathcal{N}\left(0, \frac{1}{2\theta}\sigma_x^2(1 - e^{-2\theta\Delta})\right).$$

Data are generated according to the following equation for each $t = 1, \ldots, T$:

$$y_t = \frac{1}{gK}\Big(\sum_{k=0}^{K-1}\sum_{i=0}^{g-1}\beta_0(t - t_{k,i})X_t(t - t_{k,i}) + \beta_0(t - t_K)X_t(t - t_K) - \beta_0(t - t_{0,0})X_t(t - t_{0,0})\Big) + \widetilde{\varepsilon}_t, \qquad \widetilde{\varepsilon}_t \sim \mathcal{N}(0, 1).$$

As in Section 4.2, m denotes the number of times the high-frequency period enters the low-frequency period. Therefore, $t_1 = t - p_x + 1/m$, $t_2 = t - p_x + 2/m$, $t_3 = t - p_x + 3/m$, $\ldots$,

$t_m = t - p_x + 1, \ldots, t_{K-1} = t - 1 + (m-1)/m$, $t_K = t$ and the grid to generate the data is given by the dates $t_{j,i} = t_{j+1} - (g-i)/(g*m)$ for $j = 0, 1, \ldots, K$ and $i = 0, 1, \ldots, g-1$. We fix the parameters at: $\theta = 1$, $\sigma_x = 1$, $p_x = 10$. For the *smooth case* we have tried the following different DGPs: $(a, b) = (2, 10)$ (bell-shaped, DGP 1.1), $(a, b) = (10, 1)$ (fast-decaying, DGP 1.2), $(a, b) = (4, 1)$ (slow decaying, DGP 1.3), and $(a, b) = (1, 1)$ (flat, DGP 1.4). We fix different values for T: $T = 300$, $T = 500$, $T = 1000$ and for g, m: $g = 3$, $m = 90$ (to reproduce daily observations) and $g = 30$, $m = 12$ (to reproduce monthly observations). Figure 3 illustrates different shapes of β_0 functions for the smooth case.

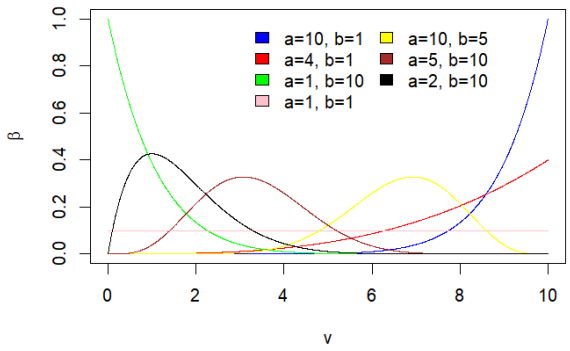


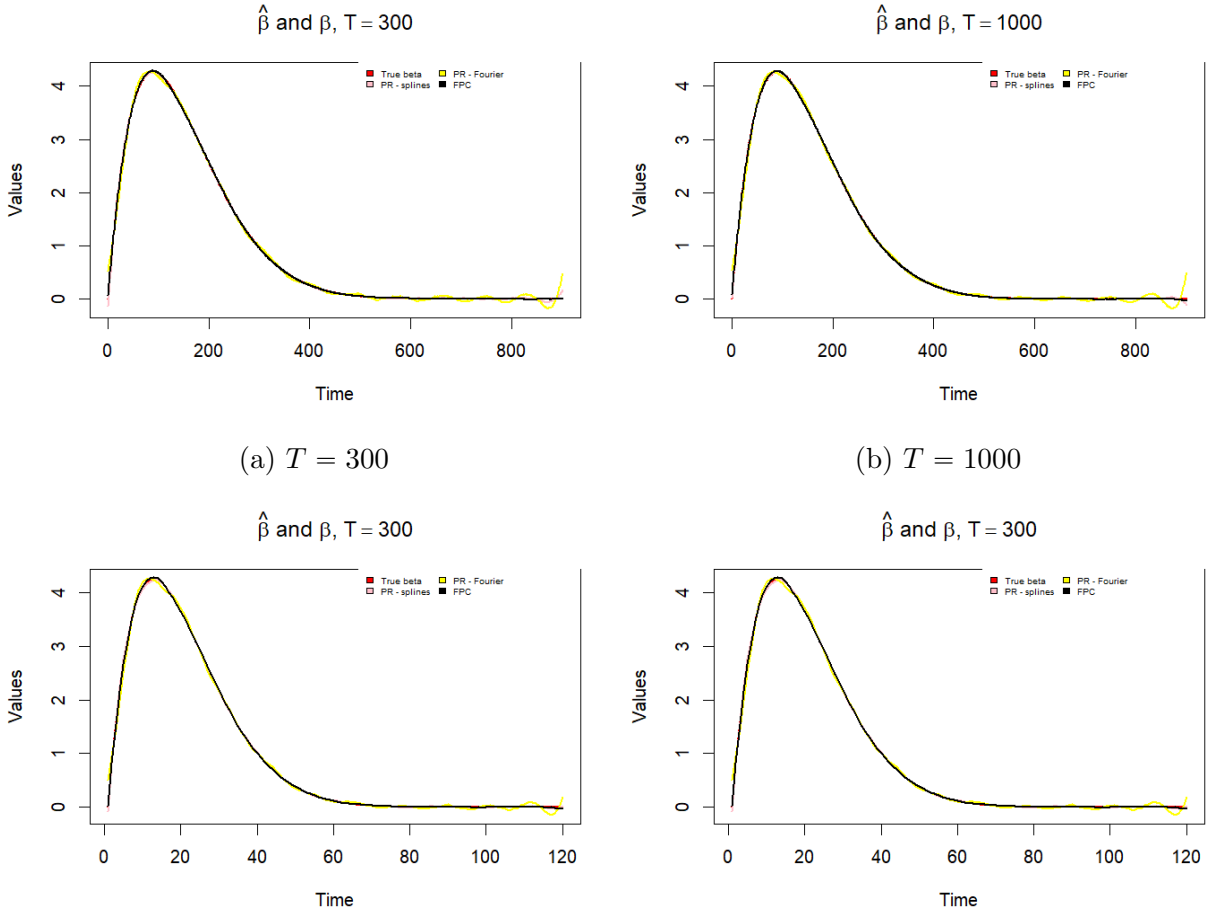
Figure 3: Different parametrization for the function $\beta(\cdot)$ in the *smooth case*.

Results for the *smooth case* are presented in Figure 4. Results for the *piecewise case* are presented in Figure 5. We see that the PLS estimator with B-splines always performs better than the competitors. The MIDAS estimator, whether using the Almon polynomial or the Beta function, performs poorly in terms of recovering the function β. In some figures, we do not report the estimation with MIDAS because it is too far from the true function.

9 Additional controls

In this section, we aim to add control variables to our model as given by equation (4). As emphasized in Montiel Olea et al. (2025), the aim of a local projection is to estimate impulse responses with respect to a shock, defined as the residual of the signal $x(\cdot)$ after projecting it onto the control variables. So far, we have only considered the lagged values $y_{t-p_x}, \ldots, y_{t-p_y}$ of the target variable y_t as control variables. However, one may choose to include additional controls depending on the macroeconomic context.

To address this issue, we explain how our model (4) can be extended to account for



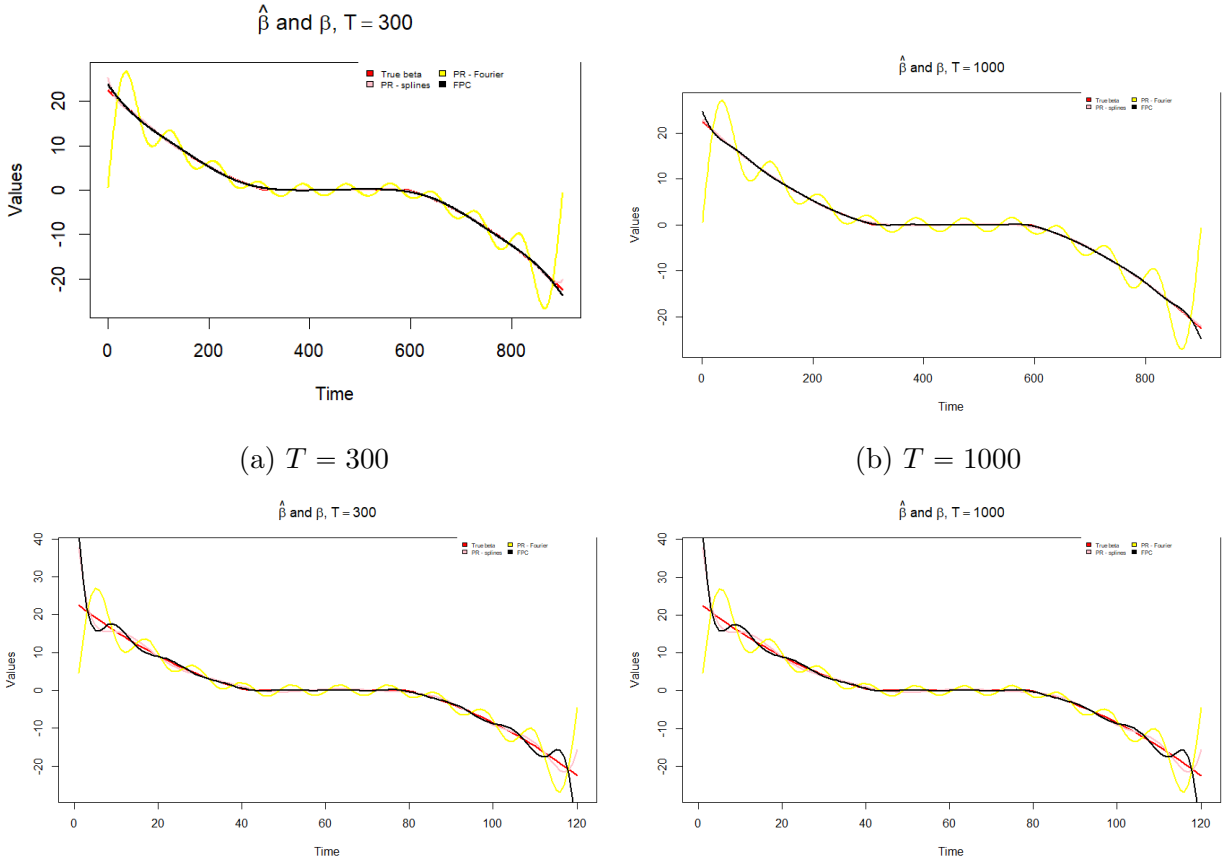
(a) $T = 300$ (b) $T = 1000$

(c) $T = 300$ (d) $T = 1000$

Figure 4: *Smooth case - Daily and Monthly.* Estimation of the $\beta(\cdot)$ function. Monte Carlo with 1000 iterations. Panels (a)and (b): $g = 3$, $m = 90$ (daily). Panels (c), (d) and (f): $g = 30$, $m = 12$ (monthly). FPC is Functional Principal Components, PR - F is penalized regression with Fourier basis functions, PR - BS is penalized regression with B-Splines.

both contemporaneous and lagged controls. For simplicity, and since aggregating the control variables is less critical for our purposes than aggregating the signal, we focus on the case where all controls are low-frequency. While we could address different low-frequencies for the controls, we omit this case to avoid overly cumbersome notation.

Let us consider enriching model (4) with a d_1-vector $\mathbf{z}_t := (z_{1,t}, \ldots, z_{d_1,t})'$ of time series that are included as lagged controls, and with a d_2-vector $\mathbf{r}_t := (r_{1,t}, \ldots, r_{d_2,t})'$ of time series that are included as contemporaneous controls. All the controls are observed at low-frequency, and they are supposed to have a mean of zero.Let us consider the following model:



(a) $T = 300$ (b) $T = 1000$

(c) $T = 300$ (d) $T = 1000$

Figure 5: *Piecewise case - Daily and Monthly.* Estimation of the $\beta(\cdot)$ function. Monte Carlo with 1000 iterations. Panels (a) and (b): $g = 3$, $m = 90$ (daily). Panels (c) and (d): $g = 30$, $m = 12$ (monthly). FPC is Functional Principal Components, PR - F is penalized regression with Fourier basis functions, PR - BS is penalized regression with B-Splines.

for each $t \in \{p, \ldots, T\} \subset \mathbb{N}_+$,

$$y_t = \sum_{i=p_x}^{p_y} \gamma_{i+1-p_x} y_{t-i} + \sum_{i=p_x}^{p_y} \beta'_{z,i+1-p_x} z_{t-i} + \beta'_r \mathbf{r}_t + \int_{t-p_x}^{t} \beta(t-\tau)x(\tau)d\tau + \eta_t,$$

$$\mathbf{E}\left[\eta_t(\mathbf{y}_{t-p_x}, \mathbf{z}_{t-p_x}, \ldots, \mathbf{z}_{t-p_y}, \mathbf{r}_t, x(s))\right] = 0, \quad \forall s \in \mathbb{R}, \quad (22)$$

where η_t is a projection residual, $\beta_{z,i} \in \mathbb{R}^{d_1}$ for every $i = 1, \ldots, p_y - p_x + 1$, $\beta_z := (\beta'_{z,1}, \ldots, \beta'_{z,p_{yx}})' \in \mathbb{R}^{d_1 p_{yx}}$, $\beta_r \in \mathbb{R}^{d_2}$, $\gamma := (\gamma_1, \ldots, \gamma_{p_y - p_x + 1})$, $\beta \in L^2$, and $\mathbf{z}_{t-p_x} : (z'_{t-p_x}, \ldots, z'_{t-p_y})'$ is a $d_1 p_{yx}$-vector. Similarly, as in the univariate case, the implicit assumption is:

$$y_t \perp\!\!\!\perp \{y_s\}_{s<t-p_y}, \{\mathbf{r}_s\}_{s<t}, \{\mathbf{z}_s\}_{s<t-p_y}, \{x(s)\}_{s\leqslant t-p_x} | y_{t-p_x}, \ldots, y_{t-p_y}, \mathbf{z}_{t-p_x}, \ldots, \mathbf{z}_{t-p_y}, \mathbf{r}_t, \{x(\tau)\}_{\tau\in(t-p_x,t]}.$$

Let $\mathbf{w}_t := (\mathbf{y}'_{t-p_x}, \mathbf{z}'_{t-p_x}, \mathbf{r}_t)'$ denote the $(p_{yx}(1+d_1)+d_2)$-vector containing all the low-frequency covariates. Define $\Omega_{ww} := \mathbf{E}[\mathbf{w}_t\mathbf{w}'_t]$, the covariance matrix in $\mathbb{R}^{(p_{yx}(1+d_1)+d_2)\times(p_{yx}(1+d_1)+d_2)}$, $\Omega_{wx} := \mathbf{E}[\mathbf{w}_t\langle X_t, \cdot\rangle]$ the cross-covariance operator from L^2 to $\mathbb{R}^{(p_{yx}(1+d_1)+d_2)}$, defined by the covariance between the high-frequency variable $X_t(\cdot)$ and $\mathbf{w}_t$, and $\Omega^*_{wx} := \mathbf{E}[X_t(\cdot)\mathbf{w}'_t] : \mathbb{R}^{(p_{yx}(1+d_1)+d_2)} \to L^2$ the adjoint cross-covariance operator of Ω_{wx}. We introduce the following assumption which modifies Assumption 3.1.

Assumption 9.1. [*a*]. *The discrete-time stochastic processes* $\{y_t\}_t$, $\{\mathbf{z}_t\}_t$, *and* $\{\mathbf{r}_t\}_t$ *are jointly zero-mean weakly stationary. Additionally, the AR polynomial* $1 - \sum_{i=p_x}^{p_y} \gamma_{i+1-p_x} z^i$ *has all roots outside the unit circle.*

[*b*]. *The continuous-time stochastic process* $\{x(t), t \in \mathbb{R}\}$ *is a zero-mean* L^2*-valued weakly stationary process, that is, (i)* $x \in L^2$ *with probability* 1, *(ii)* $\mathbf{E}[x(t)] = 0$, *for all* $t \in \mathbb{R}$, *and (iii) for all* $t_1, t_2 \in \mathbb{R}$, $\mathbf{E}[x(t_1)x(t_2)] = \gamma_x(t_1 - t_2)$ *and* $\gamma_x(t_1 - t_2) = \gamma_x(t_2 - t_1)$. *In addition,* $\mathbf{E}\|x\|^4 < \infty$ *and* $\mathbf{E}\|\mathbf{r}_t\|4 < \infty$.

[*c*.] *The cross-covariance functions between all pairs of processes* $(y_t, \mathbf{z}_t, \mathbf{r}_t, x(t))$ *are time-invariant, that is, for all* $t \in \mathbb{Z}$ *and all* $s \in \mathbb{R}$, $h \in \mathbb{Z}$: $Cov(y_t, x(t-s)) = \gamma_{yx}(s)$, $Cov(\mathbf{z}_t, x(t-s)) = \gamma_{zx}(s)$, $Cov(\mathbf{r}_t, x(t-s)) = \gamma_{rx}(s)$, $Cov(y_t, \mathbf{z}_{t-h}) = \gamma_{yz}(h)$, $Cov(y_t, \mathbf{r}_{t-h}) = \gamma_{yr}(h)$, $Cov(\mathbf{z}_t, \mathbf{r}_{t-h}) = \gamma_{zr}(h)$.

[*d*.] *The joint process* $(y_t, \mathbf{z}'_t, \mathbf{r}'_t, x(t))$ *is purely nondeterministic and contains no deterministic components.*

Assumption 9.2. *Let* Ω_{xx} *be the covariance operator defined above. The following holds:*

[*a*.] Ω_{xx} *is injective.*

[*b*.] $\forall \gamma \in \mathbb{R}^{p_{yx}}$, $\beta_z \in \mathbb{R}^{d_1 p_{yx}}$, $\beta_r \in \mathbb{R}^{d_2}$ *there is no* $\phi \in L^2$, *such that* $\gamma'\mathbf{y}_{t-p_x} + \beta'_z\mathbf{z}_{t-p_x} + \beta'_r\mathbf{r}_t = \langle \phi, X_t\rangle$.

Proposition 3.1 then trivially modifies as follows.

Proposition 9.1. *Suppose Assumptions 9.1 - 9.2 hold and consider the operator* $\Omega_{xx} : L^2 \to L^2$ *with eigensystem* $(\lambda_j, \omega_j)_{j\geqslant 1}$ *and* $\lambda_j > 0$ *for every* $j \geqslant 1$. *Assume that*

$$\sum_{j\geqslant 1} \frac{\langle \mathbf{E}[\mathbf{w}_{t,j}X_t], \omega_j\rangle^2}{\lambda_j^2} < \infty \qquad \forall j = p_x, p_x+1, \ldots, (p_y+d). \tag{23}$$

Then, model (22) *is identified in the sense that there exists a unique value of* $(\gamma', \beta'_z, \beta'_r, \beta)'$ *in* $\mathbb{R}^p_{yx} \times \mathbb{R}^{d_1 p_{yx}} \times \mathbb{R}^{d_2} \times L^2$ *that satisfies the moment restrictions in* (22). *This value is denoted*

by $(\gamma_0', \beta_{z,0}', \beta_{r,0}', \beta_0)'$ *and is given by*

$$(\gamma_0', \beta_{z,0}', \beta_{r,0}')' = \left(\Omega_{ww} - \Omega_{wx}\Omega_{xx}^{-1}\Omega_{wx}^*\right)^{-1}\left(\mathbf{E}[y_t\mathbf{w}_t] - \Omega_{wx}\Omega_{xx}^{-1}\mathbf{E}[y_t X_t]\right) \quad (24)$$

$$\beta_0 = \left(\Omega_{xx} - \Omega_{wx}^*\Omega_{ww}^{-1}\Omega_{wx}\right)^{-1}\left(\mathbf{E}[y_t X_t] - \Omega_{wx}^*\Omega_{ww}^{-1}\mathbf{E}[y_t\mathbf{w}_t]\right). \quad (25)$$

The estimation of the parameters of this model proceeds as in the basic case treated in Section 4 where in all the formulas we have to replace the vector $\mathbf{y}_{t-1}$ with the vector $\mathbf{w}_t$, γ with $(\gamma', \beta_z', \beta_r')'$, $\widehat{\Omega}_{yy}$ with $\widehat{\Omega}_{ww} := \frac{1}{T}\sum_{t=p}^{T}\mathbf{w}_t\mathbf{w}_t'$, and $\widehat{\Omega}_{yx}$ with $\widehat{\Omega}_{wx} := \frac{1}{T}\sum_{t=p}^{T}\mathbf{w}_t\langle X_t, \cdot\rangle$.

10 Conclusion

This paper proposes an approach to model and make inferences about the impact of a high-frequency shock on low-frequency macroeconomic aggregates. We introduce the concept of the continuous-time impulse response function (CT-IRF) in macroeconomics. The response of the low-frequency variable y_t to a shock in the continuous-time input signal $x(\cdot)$ is characterized as the convolution of the CT-IRF and the input signal.

Similar to the Local Projections (LP) approach of Jordà (2005), we do not assume a true underlying model. Instead, we consider the linear projection of the low-frequency target series y_t onto: (i) the path segment of $\{x(u)\}$ over the interval $[t - p_x, t]$, and (ii) a set of control variables that include lagged values of all variables in the system up to $t - p_x$, the period at which the impulse occurs. However, unlike local projections, our method estimates the entire IRF in a single step, without needing to estimate a separate regression model for each forecast horizon. Furthermore, our model is designed to handle mixed-frequency data, which is not feasible with the LP approach.

Our model has two important features. First, it is specially designed for situations with large gaps in sampling frequencies, such as daily and quarterly data. Second, the high-frequency signal is modeled as a continuous-time stochastic process, of which we observe a discretized version at high frequency. Importantly, our approach allows us to compute IRFs at high frequency without aggregating the information in the high-frequency variable, thereby avoiding the loss of information that aggregation always generates.

Estimation of the CT-IRF is a deconvolution problem that we express as a linear functional regression. Unlike standard functional linear regression, our method requires only a single observed discretized trajectory of $\{x(t)\}$. We construct functional objects by cutting partially overlapping path segments from this trajectory. To address the ill-posed nature of

the deconvolution problem, we propose a penalized least squares estimator as a regularization approach.

We illustrate the usefulness of our approach by first estimating the dynamic impact of daily uncertainty shocks on U.S. equipment investment growth by providing a daily IRF. In our second application, we estimate the effects of high-frequency financial shocks on economic activity. We find that a negative impact on U.S. GDP growth from daily financial shocks rapidly emerges, about 15 business days after the impact, while the full effect is visible around 60 business days after the impact.

References

E. Andreou, E. Ghysels, and A. Kourtellos. Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2):240–251, 2013.

A. Babii, E. Ghysels, and J. Striaukas. Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics*, 40(3):1094–1106, 2022.

S. R. Baker, N. Bloom, and S. J. Davis. Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636, 2016.

B. Bernanke and M. Gertler. Agency costs, net worth, and business fluctuations. *The American Economic Review*, 79(1):14–31, 1989.

B. S. Bernanke. Irreversibility, Uncertainty, and Cyclical Investment. *The Quarterly Journal of Economics*, 98(1):85–106, 1983.

N. Bloom. The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685, 2009.

N. Bloom. Fluctuations in Uncertainty. *Journal of Economic Perspectives*, 28(2):153–176, 2014.

D. Bosq. *Linear Processes in Function Spaces*. Lecture Notes in Statistics, Springer New York, NY, 2000.

T. T. Cai and P. Hall. Prediction in functional linear regression. *The Annals of Statistics*, 34(5):2159 – 2179, 2006.

D. Caldara, C. Fuentes-Albero, S. Gilchrist, and E. Zakrajsek. The macroeconomic impact of financial and uncertainty shocks. *European Economic Review*, 88:185–207, 2016.

H. Cardot, F. Ferraty, and P. Sarda. Spline estimators for the functional linear model. *Statistica Sinica*, 13(3):571–591, 2003.

H. Cardot, A. Mas, and P. Sarda. CLT in functional linear regression models. *Probab. Theory Relat. Fields*, 138:325 – 361, 2007.

A. Carriero, T. E. Clark, and M. Marcellino. Nowcasting tail risks to economic activity with many indicators. Working Paper 20-13R, Federal Reserve Bank of Cleveland, 2020.

X. Chen and M. Reiss. On rate optimality for ill-posed inverse problems in econometrics. *Econometric Theory*, 27(3):497–521, 2011.

A. Chudik and G. Georgiadis. Estimation of impulse response functions when shocks are observed at a higher frequency than outcome variables. *Journal of Business & Economic Statistics*, 40(3):965–979, 2022.

A. Chudik and L. Kilian. Mean group and pooled mixed-frequency estimators of responses of low-frequency variables to high-frequency shocks. Working Paper 2603, Federal Reserve Bank of Dallas, 2026. URL `https://www.dallasfed.org/research/papers/2026/wp2603`.

M. P. Clements and A. B. Galvão. Forecasting us output growth using leading indicators: an appraisal using midas models. *Journal of Applied Econometrics*, 24(7):1187–1206, 2009.

C. Crambes, A. Kneip, and P. Sarda. Smoothing splines estimators for functional linear regression. *The Annals of Statistics*, 37(1):35 – 72, 2009.

J. Dauxois, A. Pousse, and Y. Romain. Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *Journal of Multivariate Analysis*, 12(1):136–154, 1982.

C. Doz, L. Ferrara, and A. Simoni. Nowcasting with functional approaches and mixed-frequency data. mimeo, 2025.

H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications, Springer Dordrecht, 2000.

Y. Fan, G. M. James, and P. Radchenko. Functional additive regression. *The Annals of Statistics*, 43(5):2296 – 2325, 2015.

L. Ferrara and P. Guérin. What are the macroeconomic effects of high-frequency uncertainty shocks? *Journal of Applied Econometrics*, 33(5):662–679, August 2018.

L. Ferrara and A. Simoni. When are google data useful to nowcast GDP? an approach via preselection and shrinkage. *Journal of Business & Economic Statistics*, 41(4):1188–1202, 2023.

L. Ferrara, M. Mogliani, and J.-G. Sahuc. Monitoring high-frequency Growth-at-Risk. *International Journal of Forecasting*, 38:582–595, 2022.

J.-P. Florens. *Inverse Problems and Structural Econometrics: The Example of Instrumental Variables*, page 284–311. Econometric Society Monographs. Cambridge University Press, 2003.

J.-P. Florens and A. Simoni. Nonparametric estimation of an instrumental regression: A quasi-bayesian approach based on regularized posterior. *Journal of Econometrics*, 170(2): 458–475, 2012.

J.-P. Florens and A. Simoni. Regularizing priors for linear inverse problems. *Econometric Theory*, 32(1):71–121, 2016.

J.-P. Florens and A. Simoni. Gaussian processes and bayesian moment estimation. *Journal of Business & Economic Statistics*, 39(2):482–492, 2021.

C. Foroni, M. Marcellino, and C. Schumacher. Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1):57–82, 2015.

E. Ghysels. Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics*, 193(2):294–314, 2016.

E. Ghysels, A. Sinko, and R. Valkanov. MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90, 2007.

J. D. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

D. Hollo, M. Kremer, and M. L. Duca. CISS - a composite indicator of systemic stress in the financial system. Working Paper 1426, European Central Bank, 2012. URL `https://ssrn.com/abstract=2018792`.

S. Hörmann and P. Kokoszka. *Functional Time Series*, volume 30 of *Time Series Analysis: Methods and Applications*. Elsevier, 2012.

A. Inoue and B. Rossi. A new approach to measuring economic policy shocks, with an application to conventional and unconventional monetary policy. *Quantitative Economics*, 12(4):1085–1138, 2021.

J. Johannes and M. Schwarz. Adaptive nonparametric instrumental regression by model selection, 2010. URL `http://hdl.handle.net/2078.1/106172`.

O. Jordà. Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182, 2005.

F. C. Klebaner. *Introduction to Stochastic Calculus with Applications*. IMPERIAL COLLEGE PRESS, 3rd edition, 2012. doi: 10.1142/p821.

F. H. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin Company, Boston & New York,, 1921.

A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis*. Dover, New York, dover edition edition, 1975.

G. Koop, M. Pesaran, and S. M. Potter. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147, 1996.

Y. Li and T. Hsing. On rates of convergence in functional linear regression. *Journal of Multivariate Analysis*, 98(9):1782–1804, 2007.

A. I. Margaris. *Continuous and Discrete-Time Signals and Systems*. Springer, 2024.

B. D. Marx and P. H. Eilers. Generalized linear regression on sampled signals and curves: A p-spline approach. *Technometrics*, 41(1):1–13, 1999.

M. Mogliani and A. Simoni. Bayesian MIDAS penalized regressions: Estimation, selection, and prediction. *Journal of Econometrics*, 222(1, Part C):833–860, 2021.

M. Mogliani and A. Simoni. Bayesian bi-level sparse group regressions for macroeconomic density forecasting, 2024. URL https://arxiv.org/abs/2404.02671.

J. L. Montiel Olea, M. Plagborg-Møller, E. Qian, and C. K. Wolf. Local projections or VARs? a primer for macroeconomists. arXiv working paper, 2025. URL https://arxiv.org/abs/2503.17144.

M. Plagborg-Møller and C. K. Wolf. Local projections and vars estimate the same impulse responses. *Econometrica*, 89(2):955–980, 2021.

J. O. Ramsay and B. W. Silverman. *Applied Functional Data Analysis Methods and Case Studies*. Springer New York, NY, 2007.

CREST
CENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

The Center for Research in Economics and Statistics (CREST) is a leading French scientific institution for advanced research on quantitative methods applied to the social sciences.

CREST is a joint interdisciplinary unit of research and faculty members of CNRS, ENSAE Paris, ENSAI and the Economics Department of Ecole Polytechnique. Its activities are located physically in the ENSAE Paris building on the Palaiseau campus of Institut Polytechnique de Paris and secondarily on the Ker-Lann campus of ENSAI Rennes.

cnrs

GENES Groupe des écoles nationales d'économie et statistique