Sovereign risk and corporate cost of borrowing: Evidence from a counterfactual study

Marcin Wolski*

European Investment Bank

April 2018

Abstract

We assess the impact of the sovereign risk spill-overs onto corporate cost of borrowing in selected euro area countries. We utilize a novel nonparametric dependence filtering framework to remove the effects of sovereign risk in the interest rate pass-through context. The main findings confirm the heterogeneity in sovereign risk spill-overs. We also find divergence in sovereign risk transmission between core and peripheral Member States during financial and sovereign debt crises. We discover that the standard linear models may underestimate the underlying pass-through distortions, suggesting the existence of nonlinear sovereign risk effects.

JEL-Classification: C14, E43, E52, G21

Keywords: counterfactual distributions, nonparametric methods, sovereign risk, cost of borrowing, pass-through

*Contact details: 98-100 Boulevard Konrad Adenauer, L-2950 Luxembourg. Email: M.Wolski@eib.org. I would like to thank Cees Diks, Laurent Maurin, Michiel van de Leur, Debora Revoltella, Christoph Rothe and participants of the 23rd International Conference Computing in Economics and Finance in New York and the 26th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics in Tokyo for useful comments. Numerical simulations were run on the Lisa Compute Cluster in Amsterdam. The opinions expressed herein are those of the author and do not necessarily reflect those of the European Investment Bank.

1 Introduction

The purpose of the paper is to evaluate to what extent the sovereign risk distorted the pass-through channel in the euro area (EA). To this end we utilize a novel fully nonparametric framework to construct a counterfactual scenarios with no sovereign risk distortions. Nonparametric methods have been proved to deliver substantial precision gains under possible model misspecification (Rothe, 2010) or irregularities of the underlying data processes (Diks and Wolski, 2016), which we aim to pass on to our empirical findings.

The counterfactual framework has its starting point in hypothesis testing literature (Hiemstra and Jones, 1994; Diks and Panchenko, 2006; Diks and Wolski, 2016), and further capitalizes on data sharpening methods proposed by Choi and Hall (1999) and Hall and Minnotte (2002). The goal of the procedure is to provide a sample from counterfactual distribution which is independent from the effects of a given set of covariates. The method can be viewed as a dependence filtering procedure.

To provide liquidity to the troubled EA banking sector and to alleviate the supply-side credit constraints, the European Central Bank (ECB) started a series of interest rates cuts in October 2008. The effects on the retail funding costs heavily differed across the member states revealing the heterogeneity in the ECB monetary policy transmission mechanism and the segmentation of the Euro Area (EA) banking system (EIB, 2016).

A sizeable part of cross-country transmission asymmetries can be attributed to the differences in sovereign risk (Arnold and van Ewijk, 2014). Elevated risk levels raise the borrowing costs of the sovereign, which often serve as a benchmark market rate for loan pricing. Additionally, weak fiscal positions reduce the ability of a sovereign to support financial sector in times of distress. Disturbances to this implicit sovereign backstop can be perceived as an untight safety net, putting an upward pressure on the required rate of return and increasing consequently the lending rates.

We focus on three EA countries, i.e. France, Italy and Spain, for which we filter out the sovereign risk effect on the retail costs of borrowing. Our main findings confirm the pass-through heterogeneity across three EA Member States. During the sovereign debt crisis, the funding costs were substantially more distorted in Spain and Italy, compared to France, although the latter also observed non-negligible sovereign risk pass-through in that period. The sovereign distortions were also less pronounced in the pricing of the short-term loans, which provides some degree of comfort for smaller corporates. We also confirm that the standard linear models may underestimate the underlying pass-through effects, suggesting an existence of nonlinear sovereign risk spill-overs.

The paper is organized as follows. Section 2 outlines the filtering framework. We provide the estimation techniques for independent counterfactual distributions and we derive its basic in-sample properties. For brevity reasons, numerical verification of the procedure is given in the Appendix. We apply the methodology to filter out the sovereign risk effects on corporate borrowing costs in Section 3. Finally, Section 4 concludes.

2 Methodology

Counterfactual distributions have become an important tool for policy practicioners to evaluate the performance of proposed strategies, decompose specific policy outcomes, select the most efficient implementation mechanisms, and tailor-make policy instruments. In the context of impact assessment, Gertler et al. (2011) point out that counterfactual scenarios constitute an industry benchmark for evidence-driven policy making.

In general terms, one can think of a counterfactual distribution, denoted by Y', as representing the behaviour of distribution Y in the context of information spill-overs from a given set of covariates X. Chernozhukov et al. (2013) point out that distribution Y' can reflect the effects of changes in a covariate distribution X onto the outcome distribution Y, or it can echo a relationship change between X and Y.

Despite the wide array of applications, the vast majority of counterfactual scenarios are user-designed, suffering from a natural over-simplification and potential model misspecification biases. Nevertheless, the recent advances in counterfactual distributions aim at providing possibly assumption-free inference techniques. Chernozhukov et al. (2013) offer a complete toolkit to study counterfactual distributions through a prism of regression methods. Rothe (2010) extends the approach to a fully nonparametric setup and demonstrates that nonparametric estimation has superior Mean Squared Error (MSE) performance in the case of (functional) model misspecification. Rothe (2012) further extends the nonparametric aproach to cover partial distributional effects.

In those techniques, the identification of counterfactual distribution Y' reflects a change in the covariate distribution from X to, say, X', both of which need to be known to the end user. Additionally, for Y' to be well-defined, the support of X' needs to be included in the support of X, $X' \subseteq X$. In this paper we propose a different identification strategy, which follows from the independence and conditional independence principles between random variables.

For a formal introduction of the method, let us consider a random real-valued outcome variable Y, which is potentially influenced by a d_X-dimensional vector of random real-valued covariates $\boldsymbol{X} = \{X_1, ..., X_{d_X}\}$. Let us further assume that both Y and $\boldsymbol{X}$ have smooth and well-defined Probability Density Functions (PDFs) over their domains, denoted by $f_Y(y)$ and $f_{\boldsymbol{X}}(\boldsymbol{x})$, respectively.[1] There also exist joint PDF, denoted by $f_{Y,\boldsymbol{X}}(y, \boldsymbol{x})$. For clarity of exposition and without loss of generality let us imagine that the outcome variable and covariates are not independent, i.e. $f_{Y,\boldsymbol{X}}(y, \boldsymbol{x}) \neq f_Y(y) f_{\boldsymbol{X}}(\boldsymbol{x})$ for some $y \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^{d_X}$.[2] In principle, if the variables of interest are mutually independent, the following reasoning still holds, and the filtering map is an identity transformation.

Assume that we have a sample $\{(Y_i, \boldsymbol{X_i}) : i = 1, ..., n\}$, consisting of independent and identically distributed (i.i.d.) draws from $(Y, \boldsymbol{X})$ with properties described above. By filtering out the effects of $\boldsymbol{X}$ onto Y we mean finding a counterfactual distribution of a random variable Y', characterized

[1] The Cumulative Distribution Functions (CDFs) are given by $F_Y(y)$ and $F_{\boldsymbol{X}}(\boldsymbol{x})$, respectively.

[2] PDF-representation of the condition proved to be computationally more attractive and has been widely used in hypothesis testing (Hiemstra and Jones, 1994; Diks and Panchenko, 2006; Diks and Wolski, 2016). Although our main results can be also derived from the CDF functionals, the PDF approach allows to link the methodology to the testing environment explicitly.

by a sample $\{Y_i' : i = 1..n\}$, which is independent from $\boldsymbol{X}$. By design, counterfactual distribution Y' is defined over the subset of the support of Y, $Y' \subseteq Y$. In terms of PDFs, one can write the independence condition as

$$f_{Y',\boldsymbol{X}}(y, \boldsymbol{x}) = f_Y(y) f_{\boldsymbol{X}}(\boldsymbol{x}) \tag{1}$$

for all y and $\boldsymbol{x}$. Note that in this setup the independence is obtained by adjusting the joint probability space rather than the respective unconditional distributions. We motivate this by the information theory. In case of dependent random variables, variables $\boldsymbol{X}$ provide extra information about variable Y. Consequently, the conditional distribution of Y given $\boldsymbol{X}$ is not equal to the unconditional distribution of Y. Filtering this extra information should be then naturally defined over the conditional rather than unconditional distribution of Y, which is reflected in Eq. (1). Additionally, such a setting builds an easy link to the literature on counterfactual analysis (see, for instance, Chernozhukov et al. (2013)).

The sample from a counterfactual distribution Y' is obtained by solving Eq. (1) for sample realizations $\{Y_i' : i = 1, ..., n\}$. Using the data sharpening framework, one can show that sample $\{Y_i'\}$ is a function of $\{Y_i\}$ for given realizations $\{\boldsymbol{X}_i\}$ for each $i = 1, ..., n$, i.e. $Y_i' = \phi(Y_i | \boldsymbol{X}_i = \boldsymbol{x}_i)$, where ϕ is a sharpening function. The main result is summarized in Theorem 1.

Theorem 1. *Suppose that we have an i.i.d. sample* $\{(Y_i, \boldsymbol{X_i}) : i = 1, ..., n\}$ *from a continuous distribution with well-defined and sufficiently smooth PDFs. Then, the counterfactual distribution* Y'*, satisfying the independence condition given in Eq. (1), follows asymptotically*

$$F_Y(y') = F_{Y|\boldsymbol{X}}(y|\boldsymbol{x}), \tag{2}$$

where $F_{Y|\boldsymbol{X}}$ *is the conditional distribution function of* Y *given* $\boldsymbol{X} = \boldsymbol{x}$*, for any* y *and* $\boldsymbol{x}$ *in the support of* $(Y, \boldsymbol{X})$*.*

Proof of Theorem 1 is built upon kernel and data-sharpening methods (see Wand and Jones (1995); Silverman (1998); Hall and Minnotte (2002)), and can be found in Appendix A.1. Data sharpening changes, or perturbs, the points along given marginals to meet user-defined conditions. Since it closely follows the narrative of the proposed idea, it is a preferred line of arguments in the proof of Theorem 1, where the points along marginal Y are shifted until the independence condition in Eq. (1) is satisfied. Nonetheless, a similar conclusion can be reached by decomposing the copula of the joint distribution of (Y, X).

Theorem 1 complements and extends the growing body of literature on counterfactual distributions (Chernozhukov et al., 2013; Rothe, 2010, 2012). There are two main novelties in the proposed framework. Firstly, the setup does not require a counterfactual covariate distribution $\boldsymbol{X}'$. It makes the procedure an attractive alternative for practical applications, where such distributions are often unknown or unobservable. Secondly, since the RHS of Eq. (2) is identified by the data, and under an additional assumption that F_Y is strictly increasing, Theorem 1 is also identified for any point in the support of (Y, X).

2.1 Estimation

We propose a coherent, and fully nonparametric, framework to estimate the counterfactual distribution Y'. To generalize the notation from the beginning of Section 2, let us consider a multivariate random variable $\boldsymbol{W}$, with a corresponding sample satisfying Assumption 1.

Assumption 1. *Data* $\{\boldsymbol{W}_i : i = 1, ..., n\}$, *where* $\boldsymbol{W}_i = \{W_{1i}, ..., W_{d_W i}\}$, *are i.i.d. as a* d_W*-variate smooth continuous distribution* $F_{\boldsymbol{W}}(\boldsymbol{w})$ *with well-defined PDF* $f_{\boldsymbol{W}}(\boldsymbol{w})$ *and respective derivatives, up to a finite order* r*, which are finite, continuous and uniformly bounded on the support.*

Following Wand and Jones (1995), the kernel PDF estimator, around point $\boldsymbol{w}$, is given by

$$\hat{f}_{\boldsymbol{W}}(\boldsymbol{w}) = n^{-1} \sum_{i=1}^{n} K_{\mathbf{H}} \left(\boldsymbol{w} - \boldsymbol{W_i} \right), \tag{3}$$

where $\mathbf{H}$ is a symmetric positive definite $d_W \times d_W$ bandwidth matrix. To simplify the presentation and without the loss of generality, let us further impose that $\mathbf{H}$ is diagonal with elements $\text{diag}(h_1^2, ..., h_{d_W}^2)$. Expression

$$K_{\mathbf{H}}(\boldsymbol{w}) = (\det \mathbf{H})^{-1/2} K(\mathbf{H}^{-1/2}\boldsymbol{w}), \tag{4}$$

is the scaled kernel with 'det' denoting the determinant and K being a d_W-variate kernel function satisfying Assumption 2.

Assumption 2. *Kernel function* $K : \mathbb{R}^{d_W} \to \mathbb{R}$ *behaves as*

$$\begin{aligned} &\int K(\boldsymbol{w}) \mathrm{d}\boldsymbol{w} = 1, \\ &\int K(\boldsymbol{w}) \boldsymbol{w}^c \mathrm{d}\boldsymbol{w} = 0 \quad \text{for} \quad c = 1, ..., r-1, \\ &\int K(\boldsymbol{w}) \boldsymbol{w}^c \mathrm{d}\boldsymbol{w} = \kappa_r I_{d_W} < \infty \quad \text{for} \quad c = r, \end{aligned} \tag{5}$$

and $K(\boldsymbol{w})$ *is* r*-times differentiable, where* I_{d_W} *is a* $d_W \times d_W$ *identity matrix.*

Assumption 2 provides a set of regularity conditions on the smoothing function, commonly used in nonparametric setups. More specifically, they require that the kernel function integrates to 1 and that it is of order $r > 1$, which can be specified by the end-user. The latter condition determines the point-wise rate of convergence of the kernel estimator to the true density (Wand and Jones, 1995), which turns out to fully determine the asymptotic properties discussed later.

The corresponding CDF kernel estimator is given by integrating the kernel function

$$\hat{F}_{\boldsymbol{W}}(\boldsymbol{w}) = n^{-1} \sum_{i=1}^{n} \bar{K}_{\mathbf{H}_0} \left(\boldsymbol{w} - \boldsymbol{W_i} \right), \tag{6}$$

where $\bar{K}_{\mathbf{H}_0}(\boldsymbol{w}) = \int_{-\infty}^{\boldsymbol{w}} K(\mathbf{H}_0^{-1/2}\boldsymbol{u})\mathrm{d}\boldsymbol{u}$. We introduce here a convention that bandwidth subscript 0 corresponds to CDF estimates.

To estimate of the conditional CDF, corresponding to the RHS of Eq. (2), let us introduce variable V for which the joint i.i.d. sample is given by $\{(V_i, \boldsymbol{W}_i) : i = 1, ..., n\}$, satisfying Assumption 1. Following Li and Racine (2013), the conditional CDF can be estimated as

$$\hat{F}_{V|\boldsymbol{W}}(v|\boldsymbol{w}) = \frac{n^{-1} \sum_{i=1}^{n} \bar{K}_{\mathrm{H}_0}(v - V_i) K_{\mathbf{H}} \left(\boldsymbol{w} - \boldsymbol{W_i} \right)}{\hat{f}_{\boldsymbol{W}}(\boldsymbol{w})}, \tag{7}$$

where $\mathrm{H}_0 = h_0^2$ describes the bandwidth behaviour over CDF marginal V. Throughout the paper we take that bandwidth subscript 0 refers to the CDF marginal, whereas subscripts $1, ..., d_{\boldsymbol{W}}$ refer to PDF marginals. Consequently, the bandwidth matrix for conditional CDF estimates can be written as

$$\mathbf{H}_{V|W} = \begin{bmatrix} \mathrm{H}_0 & 0 \\ 0 & \mathbf{H} \end{bmatrix} = \begin{bmatrix} h_0^2 & 0 & \dots & 0 \\ 0 & h_1^2 & 0 & \\ \vdots & 0 & \ddots & 0 \\ 0 & & 0 & h_{d_{\boldsymbol{W}}}^2 \end{bmatrix}. \tag{8}$$

For consistency of the estimates we require that all bandwidths decrease sufficiently slowly with the sample size, as required by Assumption 3.

Assumption 3. *As $n \to \infty$,*

(i) $n^{1/2} h_0 / (\log n)^{1/2} + n^{1/2} h_0^r \to 0,$

(ii) $n^{1/2} \det \mathbf{H}^{1/2} / \log n + n^{1/2} \max \mathbf{H}^{r/2} \to 0.$

To guarantee uniform convergence of estimates, we require moderately more restrictive bandwidth behaviour than for the point-wise convergence. Assumption 3 suggests that along the CDF marginal the sequence of bandwidths of the form $h_0 = c_0 n^{-\gamma}$, with $c_0 > 0$, should converge at the rate $\gamma \in ((2r)^{-1}, 1)$. For the PDF marginals, by symmetry, setting $h_1 = c_1 n^{-\beta}, ..., h_{d_{\boldsymbol{W}}} = c_{d_{\boldsymbol{W}}} n^{-\beta}$, with $c_1, ..., c_{d_{\boldsymbol{W}}} > 0$, we have that $\beta \in ((2r)^{-1}, (2d_{\boldsymbol{W}})^{-1})$. It follows that for conditional estimates of $V|\boldsymbol{W} = \boldsymbol{w}$ we need $r > d_{\boldsymbol{W}}$, which is satisfied for the simplest example with a second-order kernel and $d_{\boldsymbol{W}} = 1$. Clearly, for multivariate applications, the kernel smoothing would require a higher-order kernel, to reduce the bias of the estimates, and create a non-empty set of possible β values.

Assumption 3 guarantees that CDF bandwidths do not converge slower than PDF ones as $n \to \infty$. In fact, the dominant terms in the MSE rates of convergence of the smoothed conditional estimators are of the same order of convergence as in the unsmoothed equivalents, as suggested by Theorem 2.2 of Li and Racine (2008). Since the convergence in MSE implies convergence in probability (Li and Racine, 2007), the corresponding uniform rates of convergence should also be similar in their dominant terms. This argument is built along the lines suggested by the uniform rates of convergence for conditional PDFs, as demonstrated by Ferraty et al. (2010), and it opens a door for smooth, instead of step-wise, CDF estimators in the setup.

In practice, Eq. (7) is often expressed as a weighted unconditional CDF of V, with weights determined by

$$\theta_{\mathbf{H}}(w, W_i) = \frac{K_{\mathbf{H}}(\boldsymbol{w} - \boldsymbol{W_i})}{\hat{f}_{\boldsymbol{W}}(\boldsymbol{w})}. \tag{9}$$

This, in fact, reduces computetional complexity of the estimators (Rothe, 2010).

Estimation of independent counterfactuals It follows that the plug-in representation of Theorem 1 requires two CDF estimates with two separate bandwidth matrices. To standardize the notation, let us denote the corresponding bandwidth matrices by superscripts referring to the specific distribution. Consequently, we estimate the draw from the independent counterfactual distribution by solving

$$\sum_{i=1}^{n} \bar{K}_{\mathrm{H}_0^Y}(y' - Y_i) = \sum_{i=1}^{n} \bar{K}_{\mathrm{H}_0^{Y|X}}(y - Y_i)\theta_{\mathbf{H}^{Y|X}}(\boldsymbol{x}, \boldsymbol{X_i}). \tag{10}$$

We will also use the bandwidth block matrix notation, highlighted by corresponding subscripts, i.e. we have that $\mathrm{H}_Y = \mathrm{H}_0^Y$ and, after Eq. (8), $\mathrm{H}_{Y|X}$ consists of blocks $\mathrm{H}_0^{Y|X}$ and $\mathrm{H}^{Y|X}$.

A counterfactual sample is obtained by solving Eq. (10) for y'. In this respect, following Jones (1992), Eq. (10) is equivalent to the plug-in representation of Eq. (2), where one can use a smooth empirical quantile function $\hat{F}_Y^{-1} = \hat{Q}_Y$.[3]

The method can be flexibly readjusted to match more complex setups, as we demonstrate in Section 2.3. Nevertheless, a close attention should be paid to the increased dimensionality and corresponding convergence properties of the (conditional) estimates.

2.2 Asymptotic properties

The (smooth) filtering framework described above can be directly linked to the literature on empirical processes and their convergence. Let us denote the estimator of the independent counterfactual distribution in terms of the quantile transformation as

$$\hat{Y}' \equiv \hat{Y}'(y, \boldsymbol{x}) = \hat{F}_Y^{-1}(\hat{F}_{Y|\boldsymbol{X}}(y|\boldsymbol{x})). \tag{11}$$

There are two main difficulties in deriving convergence properties of $\hat{Y}'$. Firstly, the quantile estimator is evaluated at a random element $\hat{F}_{Y|\boldsymbol{X}}$, adding an extra noise to the process. To guarantee consistency of the procedure, this extra stochasticity needs to be contained by the uniform convergence of the estimates (see Assumption 3). Secondly, the convergence takes place in a smooth probability space. Although, under extra assumptions, smooth empirical processes have been shown to converge in distribution, they require special care. Consequently, we firstly derive convergence properties for the standard, i.e. stepwise, estimator of the empirical process characterized in Eq. (11), and then, following Giné and Nickl (2008), we extend the results to smooth estimators. For the Reader's convenience, the main findings are presented jointly below.

Convergence of empirical processes requires certain additional regularity conditions which are outlined as Assumptions 4. In particular, we need to impose bounds on the quantile function as otherwise the process would be non-convergent in the tails.

Assumption 4. *We assume that (i) distribution functions F_Y and $F_{Y|\boldsymbol{X}}$ are Hadamard differentiable, (ii) F_Y^{-1} is uniformly Lipschitz and bounded by $[a, b] \in \mathbb{R}$, (iii) Y is supported by a compact interval on $J \in \mathbb{R}$ for which $F_{Y|\boldsymbol{X}}(y|\boldsymbol{x})$ is uniformly bounded by $[p_1, p_2] \in (0, 1)$.*

The assumption on Hadamard differentiability allows to linearize the estimate functionals in converging sequences by the functional delta method (van der Vaart, 2000). As a result, the estimates can be shown to converge to a Gaussian process, as demonstrated by Theorem 2.

[3]Because of better computational performance, in this paper we rely on numerical methods rather than on a quantile function transformation. Jones (1992) shows that both methods are MSE-equivalent.

Theorem 2. *Suppose that Assumptions 1-4 hold. Then*

$$\sqrt{n}\left(\hat{Y}' - Y'\right) \xrightarrow{d} N(0, \sigma^2), \tag{12}$$

conditional on the data, where σ^2 *is given by*

$$\sigma^2 = \frac{F_Y(y)(1 - F_Y(y)) + F_{Y|\boldsymbol{X}}(y|x)(1 - F_{Y|\boldsymbol{X}}(y|x))}{f_Y\left(F_Y^{-1}(F_{Y|\boldsymbol{X}}(y|x))\right)}. \tag{13}$$

The proof of Theorem 2 is given in Appendix A.2. This result supports the use of standard inference techniques on $\hat{Y}'$ estimates, as the distribution of the difference $\hat{Y}' - Y'$ contains all the information necessary to assess the precision of $\hat{Y}'$. Following Theorem 2, under the assumption that σ^2 is estimated consistently, the sequence $(\hat{Y}' - Y')/\hat{\sigma}$ tends to a standard Gaussian curve. It is then straightforward to calculate the confidence bounds on standard normal quantiles $z_\alpha = Q_N(\alpha)$, as $[\hat{Y}' + z_{\alpha/2}\hat{\sigma}, \hat{Y}' + z_{1-\alpha/2}\hat{\sigma}]$, where α is the statistical significance level.

In practical applications it may be difficult to estimate $\hat{\sigma}^2$ consistently, however. Therefore, we follow the approach of Chernozhukov et al. (2013) and Rothe (2010), a propose a naive bootstrap method as an alternative to approximate the true distribution of $\hat{Y}' - Y'$. Let us generate a random sample by re-sampling (i.e. sampling with replacement) from the original distribution $\{(Y_i, \boldsymbol{X}_i) : i = 1, ..., n\}$. For the specific realization i under the bootstrap draw $b = 1, ..., B$, we then estimate the $\{\hat{Y}'_{i,b}\}$ on the sample $\{(Y_{i,b}, \boldsymbol{X}_{i,b}) : i = 1, ..., n\}$ according to Theorem 2. This gives the bootstrap distribution $\{\hat{Y}'_b\} = \{(\hat{Y}'_{i,b}) : i = 1, ..., n\}$. Theorem 3 verifies that the bootstrap procedure is asymptotically valid, i.e. the distribution of the filtered realizations under the bootstrap $\{\hat{Y}'_b\}$ mimics well the distribution of filtered values under the true measure Y' as $n \to \infty$.

Theorem 3. *Suppose that Assumptions 1-4 hold. Then*

$$\sqrt{n}\left(\hat{Y}'_b - \hat{Y}'\right) \xrightarrow{d} N(0, \sigma^2), \tag{14}$$

conditional on the data, where σ^2 *is the same as in Theorem 2 but taken over the bootstrap sample.*

Proof of Theorem 3 is given in Appendix A.3. Lemma 23.3 of van der Vaart (2000) implies that under Theorems 2 and 3, bootstrap confidence intervals, given by $[\hat{Y}' + \hat{\zeta}_{\alpha/2}\hat{\sigma}, \hat{Y}' + \hat{\zeta}_{1-\alpha/2}\hat{\sigma}]$, are asymptotically consistent at the $1-\alpha$ level, where ζ_α is taken as α-quantile from the bootstrap distribution. Formally, confidence bounds for the counterfactual realization around any point y and $\boldsymbol{x}$ can be expressed as

$$\lim_{n\to\infty} \inf_{\hat{\zeta}_{\alpha/2}, \hat{\zeta}_{1-\alpha/2}} P_b\left(\hat{Y}'(y, \boldsymbol{x}) + \hat{\zeta}_{\alpha/2}\hat{\sigma} \le Y'(y, \boldsymbol{x}) \le \hat{Y}'(y, \boldsymbol{x}) + \hat{\zeta}_{1-\alpha/2}\hat{\sigma}\right) \ge 1 - \alpha. \tag{15}$$

In practice, the unknown quantities $\hat{\zeta}$ and $\hat{\sigma}^2$ can be obtained from the bootstrap distribution directly, simplifying the derivations (see, for instance, the percentile confidence interval method proposed by Efron and Tibshirani (1993)).

2.3 Conditionally independent counterfactuals

The setup described above can be extended to account for partial dependence between the variables, corresponding to Rothe (2012). Conditionally independent counterfactual distributions offer even a more flexible framework, removing the dependence against some of the covariates while preserving it for others, being particularly interesting for practical applications.

We again begin with the hypothesis testing framework. Hiemstra and Jones (1994) provide a coherent framework to test for conditional independence.[4] Extending the setup developed in Section 2, let us introduce an extra variable, denoted by $\boldsymbol{Q}$, corresponding to a $d_{\boldsymbol{Q}}$-dimensional set of covariates for which the dependence against Y is to be preserved. After Hiemstra and Jones (1994), any additional information about the conditional distribution $Y|\boldsymbol{Q} = \boldsymbol{q}$ which is contained in variable $\boldsymbol{X}$ can be reflected as a distance between the respective conditional distributions. It follows that the conditional independence condition can be written as

$$f_{Y''|\boldsymbol{Q},\boldsymbol{X}}(y|\boldsymbol{q},\boldsymbol{x}) = f_{Y|\boldsymbol{Q}}(y|\boldsymbol{q}), \tag{16}$$

where by Y'' we distinguish the conditionally, from unconditionally, independent counterfactuals.

Following closely the steps described in Section 2, we can apply the data sharpening technique to derive the closed-form solution for the counterfactual distribution of Y''. Theorem 4 summarizes the findings.

Theorem 4. *Suppose that we have an i.i.d. sample $\{(Y_i, \boldsymbol{Q_i}, \boldsymbol{X_i}) : i = 1, ..., n\}$ from a continuous distribution with well-defined and sufficiently smooth PDFs. Then, the counterfactual distribution Y'', satisfying the conditional independence condition given in Eq. (16), follows asymptotically*

$$F_{Y|\boldsymbol{Q}}(y''|\boldsymbol{q}) = F_{Y|\boldsymbol{Q},\boldsymbol{X}}(y|\boldsymbol{q},\boldsymbol{x}). \tag{17}$$

The proof of Theorem 4 follows directly from the steps outlined in Appendix A.1, by noting that the sharpened estimate of the joint density should be equal to $f_{Y|\boldsymbol{Q}}(y|\boldsymbol{q})f_{\boldsymbol{X}(x)}$. Theorem 4 suggests also that the conditionally independent counterfactual distribution is identified by data at any point, by design. Having pointed this out, the exogeneity condition ensures efficiency of the filtering procedure, in the same way as argued in Section 2.

To derive asymptotic properties of the estimates of Y'', one needs to extend the set of Assumptions 4 to cover the extra set of covariates $\boldsymbol{Q}$. Assumption 5 replaces necessary conditions on Hadamard differentiability of $F_{Y|\boldsymbol{Q}}$ and $F_{Y|\boldsymbol{Q},\boldsymbol{X}}$.

Assumption 5. *We assume that (i) conditional distribution functions $F_{Y|\boldsymbol{Q}}$ and $F_{Y|\boldsymbol{Q},\boldsymbol{X}}$ are Hadamard differentiable, (ii) $F^{-1}_{Y|\boldsymbol{Q}}$ is uniformly Lipschitz and bounded by $[a, b] \in \mathbb{R}$, (iii) Y is supported by a compact interval on $J \in \mathbb{R}$ for which $F_{Y|\boldsymbol{Q},\boldsymbol{X}}(y|\boldsymbol{q},\boldsymbol{x})$ is uniformly bounded by $[p_1, p_2] \in (0, 1)$.*

Let us denote the estimates of Y'' by $\hat{Y}''$. They can be directly retrieved through the conditional quantile functionals as

$$\hat{Y}'' \equiv \hat{Y}''(y, \boldsymbol{q}, \boldsymbol{x}) = \hat{F}^{-1}_{Y|\boldsymbol{Q}}(\hat{F}_{Y|\boldsymbol{Q},\boldsymbol{X}}(y|\boldsymbol{q},\boldsymbol{x})|\boldsymbol{q}), \tag{18}$$

[4] Diks and Panchenko (2006) show that the test statistic developed by Hiemstra and Jones (1994) is misspecified and propose a weighted test statistic as an alternative. Nevertheless, here we are not interested in the asymptotic properties of the test but we use it merely as a starting point to describe the point-wise dependence between variables, which is the same for Hiemstra and Jones (1994) and Diks and Panchenko (2006).

where, following Li and Racine (2007), we define a conditional quantile function as $F^{-1}_{Y|\boldsymbol{Q}}(e|\boldsymbol{q}) = \inf\{u : F_{Y|\boldsymbol{Q}}(u|\boldsymbol{q}) \geq e\}$.

The only major difficulty, compared with the unconditionally independent counterfactual distributions, lies in the need to estimate conditional CDFs for both sides of the baseline condition given by Theorem 4. The consistency of the conditional quantile estimates holds by the same asymptotic principles as suggested by Jones (1992). Consequently, we can extend the arguments of the smooth estimation strategy described in detail in Sections 2.1 and 2.2 to conditionally independent counterfactuals. Theorem 5 summarizes this reasoning.

Theorem 5. *Suppose that Assumptions 1-3 and 5 hold. Then*

$$\sqrt{n}\left(\hat{Y}'' - Y''\right) \xrightarrow{d} N(0, \sigma^2), \tag{19}$$

conditional on the data, where σ^2 *is given by*

$$\sigma^2 = \frac{F_{Y|\boldsymbol{Q}}(y|q)(1 - F_{Y|\boldsymbol{Q}}(y|q)) + F_{Y|\boldsymbol{Q},\boldsymbol{X}}(y|q,x)(1 - F_{Y|\boldsymbol{Q},\boldsymbol{X}}(y|q,x))}{f_{Y|\boldsymbol{Q}}\left(F^{-1}_{Y|\boldsymbol{Q}}(F_{Y|\boldsymbol{Q},\boldsymbol{X}}(y|q,x))\right)}. \tag{20}$$

To prove Theorem 5, one can follow the steps in Appendix A.2, by noting that the order of the uniform convergence of the conditional quantile function is the same as for the conditional distribution functions.

Since, in the practical applications, the extra conditioning variable $\boldsymbol{Q}$ can make the estimates of σ^2 even more cumbersome, we propose a naive bootstrap. We generate a random sample by re-sampling from the original distribution $\{(Y_i, \boldsymbol{Q}_i, \boldsymbol{X}_i) : i = 1, ..., n\}$. For the specific realization i under the bootstrap draw $b = 1, ..., B$, we then estimate $\hat{Y}''_{i,b}$ on the sample $\{(Y_{i,b}, \boldsymbol{Q}_{i,b}, \boldsymbol{X}_{i,b}) : i = 1, ..., n\}$ according to Theorem 4. Theorem 6 verifies that the bootstrap distribution $\hat{Y}''_b$ is asymptotically valid.

Theorem 6. *Suppose that Assumptions 1-3 and 5 hold. Then*

$$\sqrt{n}\left(\hat{Y}''_b - \hat{Y}''\right) \xrightarrow{d} N(0, \sigma^2), \tag{21}$$

conditional on the data, where σ^2 *is the same as in Theorem 5 but taken over the bootstrap sample.*

The proof again is a direct implication of the delta method for bootstrap (van der Vaart, 2000) and the result of Theorem 5.

We provide a numerical verification of the methodology, together with a comparison between smooth and non-smooth kernels, on an example of a stylized ARCH process in Appendix B.

3 Empirical application

To show a practical application of the filtering procedure we choose the interest rate environment in the EA. In normal times, the changes in the bank funding rates are well represented in the retail lending rates, and are passed on to the corporate costs of borrowing. In abnormal times, as a result of heterogeneous risk components, pricing differentials and market segmentation, the changes in bank lending conditions do not have to correspond to the changes in the market funding rates. In the largely banking-dependent EA this can have a detrimental influence on the real economy as the ECB rate cuts have not been fully passed to the bank lending rates in several member states and regions, where accommodative monetary policy would be particularly welcomed.

The recent financial crisis and the subsequent sovereign debt crisis decompressed the levels of risk perceived on the sovereign accounts, and their possible non-linear transmission within the financial system (Darracq Paries et al., 2014). Despite the fact that standard (linear) pass-through models can account for the former (Kok and Werner, 2006), they can underestimate the latter effects, offering an incomplete picture of the heterogeneity in sovereign risk transmission.

There is vast international evidence in favour of nonlinear dependencies in the interest rate transmission mechanisms. Apergis and Cooray (2015) estimate a Nonlinear Auto-Regressive Distributed Lag (NARDL) model and confirm asymmetric interest rate pass-through in the U.S., the U.K. and Australia. In particular, they observe that the lending rates responded more strongly to the hikes in policy rates, rather than to the falls. Using a time-varying parameter state space approach Arnold and van Ewijk (2014) find that heterogeneity in credit risk had a negligible effect on retail lending rates in the EA, in contrast to the sovereign risk. Additionally, they report that the sovereign distortions affect deposit rates to a greater extent than the lending rates. On the example of Japan, Kitamura et al. (2016) find that banks with high shares of relationship lending appear to be more restrained in pass-through of policy rate hikes, compared to the falls. This is somehow in line with a more general finding of Huang et al. (2010) and He and Krishnamurthy (2012) who suggest that banking risk is a nonlinear function of asset exposure.

In this study we focus to what extent the interest rate pass-through was distorted by country-specific sovereign characteristics. We choose three large EA economies, i.e. France, Italy and Spain. The sovereign costs are benchmarked against the German equivalent. The sample is representative of the EA's dynamics during the financial and sovereign debt crises. In particular, it has been observed that two latter countries suffered more during the crisis and consequently the monetary policy transmission channel was there less effective, than in the core EA countries (Darracq Paries et al., 2014).

We consider the Error Correction Model (ECM), which has been widely applied to evaluate the interest rate pass-through dynamics (see Darracq Paries et al. (2014)). We adapt the framework to match the filtering procedure described in Sections 2 and 2.3. In particular, the goal of the exercise is to provide counterfactual realizations of country-specific corporate costs of borrowing as if there was no distortions from sovereign risk. By comparing counterfactuals with actual realizations one can approximate by how much sovereign risk escalated, or alleviated, the pass-through rigidities in a given country.

3.1 Estimation strategy

We begin by adjusting the parametric setup to match the procedure to estimate the conditionally independent counterfactuals. The standard pass-through model, comprising the sovereign risk factors, can be written as

$$\Delta C_t = \sum_{k=0}^{lR} \beta_{kR} \Delta R_{t-k} + \sum_{j=1}^{lC} \beta_{jC} \Delta C_{t-k} + \sum_{m=1}^{lS} \beta_{mS} \Delta S_{t-m} + \alpha \nu_{t-1} + \varepsilon_t, \tag{22}$$

where C_t is a corporate cost of borrowing, R_t is the monetary policy reference rate, S_t is the sovereign risk component, ν_t is the error correction factor and ε_t is the standard error term. In such a specification coefficients β reflect the short-run pass-through effects, whereas α represents the speed of adjustment to the long-run equilibrium path given by

$$C_t = \mu_0 + \mu_R R_t + \mu_S S_t + \nu_t, \tag{23}$$

with μ_R and μ_S reflecting the long-run elasticity of corporate funding costs to the reference rate and sovereign risk factors, respectively. Finally, subscript $t = 1, ..., T$ denotes the time dimension and Δ is the first difference operator.

To match the notation described in Sections 2 and 2.3, let R_t^{lR} describe a $(lR+1)$-dimensional vector of lags given by $R_t^{lR} = \{R_{t-lR}, ..., R_t\}$, and similarly $C_t^{lC} = \{C_{t-lC}, ..., C_t\}$ and $S_t^{lS} = \{S_{t-lS}, ..., S_t\}$. The data sample is given by $\{(C_t^{lC}, R_t^{lR}, S_t^{lS}) : t = 1, ..., T\}$. Let us denote the conditionally independent counterfactual distribution of ΔC by $\Delta C''$. In this case, the counterfactual describes the dynamics of ΔC under no spillovers from the sovereign risk variables. For each realization t, the corresponding conditional independence condition is given by[5]

$$f_{\Delta C''|.}\left(\Delta C_t|\Delta R_t^{lR}, \Delta C_{t-1}^{lC}, \Delta S_{t-1}^{lS}, \nu_{t-1}\right) = f_{\Delta C|.}\left(\Delta C_t|\Delta R_t^{lR}, \Delta C_{t-1}^{lC}, \xi_{t-1}\right), \tag{24}$$

where ξ_t is the the error correction term of the model without a sovereign risk component.

Finally, by Theorem 4, the realizations $\{\Delta C_t''\}$ of a conditionally independent distribution can be estimated by solving for each t

$$\hat{F}_{\Delta C|.}\left(\Delta C_t''|\Delta R_t^{lR}, \Delta C_{t-1}^{lC}, R_{t-1}, C_{t-1}\right) = \hat{F}_{\Delta C|.}\left(\Delta C_t|\Delta R_t^{lR}, \Delta C_{t-1}^{lC}, \Delta S_{t-1}^{lS}, R_{t-1}, C_{t-1}, S_{t-1}\right), \tag{25}$$

where we nest the long-run equilibrium dynamics, as commonly done in practice (Darracq Paries et al., 2014). This, in fact, allows us to account for short- and long-term effects of sovereign risk on the costs of borrowing explicitly. For easier interpretation of the results, we reconstruct the levels of counterfactual dynamics, instead of first differences.

In line with the numerical examples, we utilize leave-one-out kernel estimates. We choose the Gaussian kernels of the order necessary to meet the uniform convergence criteria under the given dimensionality. We also evaluate statistical significance of each counterfactual estimate at the customary 5% level.

3.2 Data description and results

As the variable of interest we focus on the composite indicator of the cost of borrowing for non-financial corporations, as reported by the ECB in the monthly reports on the interest rates developments across the EA. The variable takes into account interest rates on bank loans and overdrafts in the SMEs across the EA. As argued by ECB (2013), the composite indicator is more representative of corporate borrowing situation as merely interest rates on loans, as in some of the countries, like Italy, the proportion of overdrafts may be substantial. The variable is annualised and covers new businesses.

[5]The notation subscripts are shortened for transparency.

Under the baseline specification we look at the average cost of borrowing across maturities.[6] As the reference policy rate we take 3-month EURIBOR rate as being the most liquid benchmark. Both time series come from the ECB. As a proxy for the sovereign risk we take average 10-year sovereign yield spread against the German sovereign in the same period. The time series come from Bloomberg.

The data sample is at a monthly frequency and covers the period from January 2003 until May 2017. The basic summary statistics, together with the stationarity results, are depicted in Table 1.

Table 1: Summary statistics.

	Obs.	Mean	St. dev.	Min	Max	ADF	ADF (Δ)
France							
Corporate borrowing cost	173	3.051	1.118	1.450	5.820	0.678	0.000
Sovereign risk	173	0.321	0.298	-0.007	1.450	0.530	0.000
EURIBOR (3 month)	173	1.570	1.542	-0.330	5.113	0.403	0.000
Italy							
Corporate borrowing cost	173	3.945	1.085	1.850	6.390	0.728	0.000
Sovereign risk	173	1.256	1.163	0.098	4.833	0.645	0.000
EURIBOR (3 month)	173	1.570	1.542	-0.330	5.113	0.403	0.000
Spain							
Corporate borrowing cost	173	3.528	0.952	1.900	5.960	0.818	0.000
Sovereign risk	173	1.195	1.307	-0.021	5.512	0.899	0.000
EURIBOR (3 month)	173	1.570	1.542	-0.330	5.113	0.403	0.000

Notes: Time span covers January 2003 - May 2017. Corporate borrowing cost is taken as the composite indicator of the cost of borrowing for non-financial corporations across maturities. Sovereign risk is taken as 10-year sovereign yield spread against German equivalent. ADF and ADF (Δ) denote the p-values from the Augmented Dickey-Fuller test on levels and first differences, respectively. Sources: ECB and Bloomberg.

It can be readily observed that the average corporate borrowing costs were the highest in Italy, followed by Spain and France. Interestingly, the volatility of the borrowing costs has been the highest in France, followed by Italy and Spain. After the interest rate cuts in 2008 and further non-conventional monetary measures introduced by the ECB, the borrowing costs in the France adjusted much faster than in the other two countries. In fact, only recently the corporate funding costs in Italy and Spain started converging to the levels observed in France. This, in fact, builds an avenue which we aim at exploring in the empirical setup, i.e. to what extent the sluggishness in these adjustments was driven by elevated sovereign risk.

[6] The results for the short-term costs are presented in Appendix C.

The sovereign risk in France was at negligible levels. To the contrary, the risk in Italy and Spain was nearly four times higher, on average, suggesting substantial tensions in respective sovereign sectors.

Verifying the nonparametric methodological assumptions, we test the stationarity of the time series with the Augmented Dickey-Fuller test (Fuller, 1995). The results suggest that all time series are $I(1)$. The first differences are, however, all stationary at conventional significance levels. We also find evidence for co-integration of all the variables in France and Italy. Although the error correction framework is a customary estimation strategy in the pass-through models, there is little evidence in our data that the Spanish time series are co-integrated. In fact, Johansen co-integration tests (both trace and maximum eigenvalue) doesn't deliver statistically significant results on co-integration between corporate borrowing costs, sovereign yield spread and reference rates in that country. We therefore re-estimate the model without the error correction component in Spain, i.e. we run the standard Auto-Regressive Distributed Lag (ARDL) model. Since the estimation results are parallel to the ones discussed below our focus is on the commonly applied error correction specification.

In the estimation of conditionally independent counterfactual distributions we use the multivariate Gaussian kernel of order 8, which allows us to jointly include up to 7 variables. We choose a simple specification with the lag order equal to 1 for all variables to maximize the explanatory power of our relatively short time frame. We apply the smooth estimators with normal-reference bandwidths with rates of convergence being in the middle of the allowed interval. We estimate the counterfactual differences in the corporate costs of borrowing and then remap the filtered estimates into levels. The effects of the sovereign risk factors on the corporate cost of borrowing is taken as a difference between the observed and counterfactual values, which we plot in the bottom panels. To test for the statistical significance of the sovereign risk pass-through distortions, we calculate standard confidence intervals as implied by Theorem 2.[7] We additionally compare the sovereign risk contributions with their linear equivalents, calculated from Eq. (22), as being the literature benchmark on the pass-through transmission effects.[8] The results are presented in Figs 1-3.

Several interesting features can be readily observed from the results. Firstly, the sovereign risk substantially disturbed the interest rate pass-through in Spain, compared to France and Italy. On average, the corporate borrowing costs would have been lower by 38 bp in France, by 27bp in Italy, and by 65bp in Spain, if we remove the sovereign risk transmission effects. Looking at period after 2011 only, the difference would amount to 43bp in France, 55bp in Italy and 157bp in Spain. One may observe that France record relatively less significant sovereign pass-through transmission after 2009. To the contrary, both Italy and Spain observe statistically significant sovereign distortions in the months during the financial and sovereign debt crises. This is in fact a confirmation of asymmetric interest rate transmission in various EA jurisdictions (Arnold and van Ewijk, 2014; Darracq Paries et al., 2014).

Secondly, we observe a change in sovereign risk contributions around the period of the global

[7]Because of relatively small sample size, the bootstrap procedure gave relatively weak approximation of the underlying distribution with very wide confidence bounds.

[8]The linear model is estimated on the same set of variables as its nonparametric equivalent, without demeaning the variables. For brevity, the estimation results are reported in Appendix C.

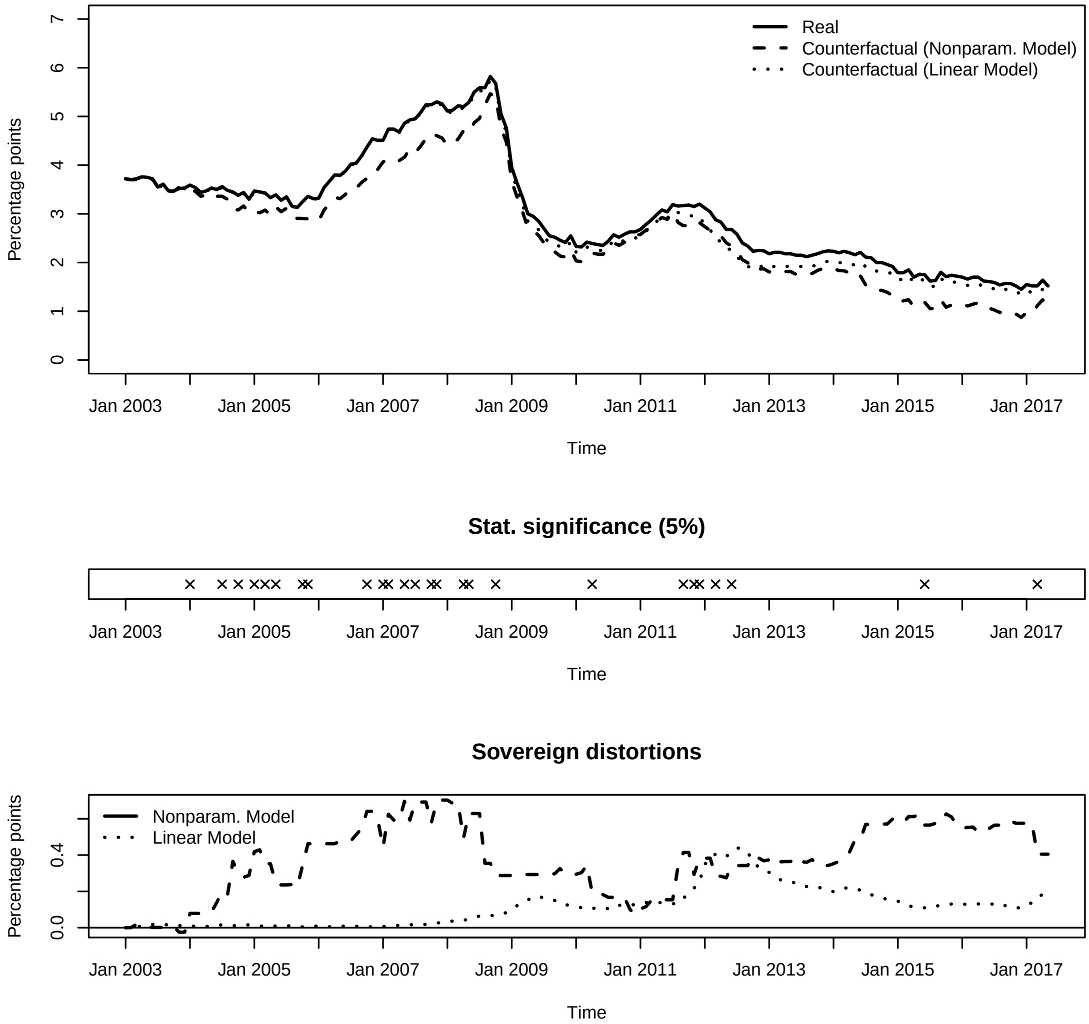

Figure 1: Sovereign risk transmission onto corporate cost of borrowing in France.

Notes: Top figure plots the realized and the counterfactual costs of borrowing, estimated by nonparametric and linear frameworks. Middle figure shows periods when the difference between the observed and nonparametric counterfactuals was statistically significant at 5% level according to the standard confidence intervals. Bottom figure shows the sovereign risk contributions derived as a difference between the observed and counterfactual rates.

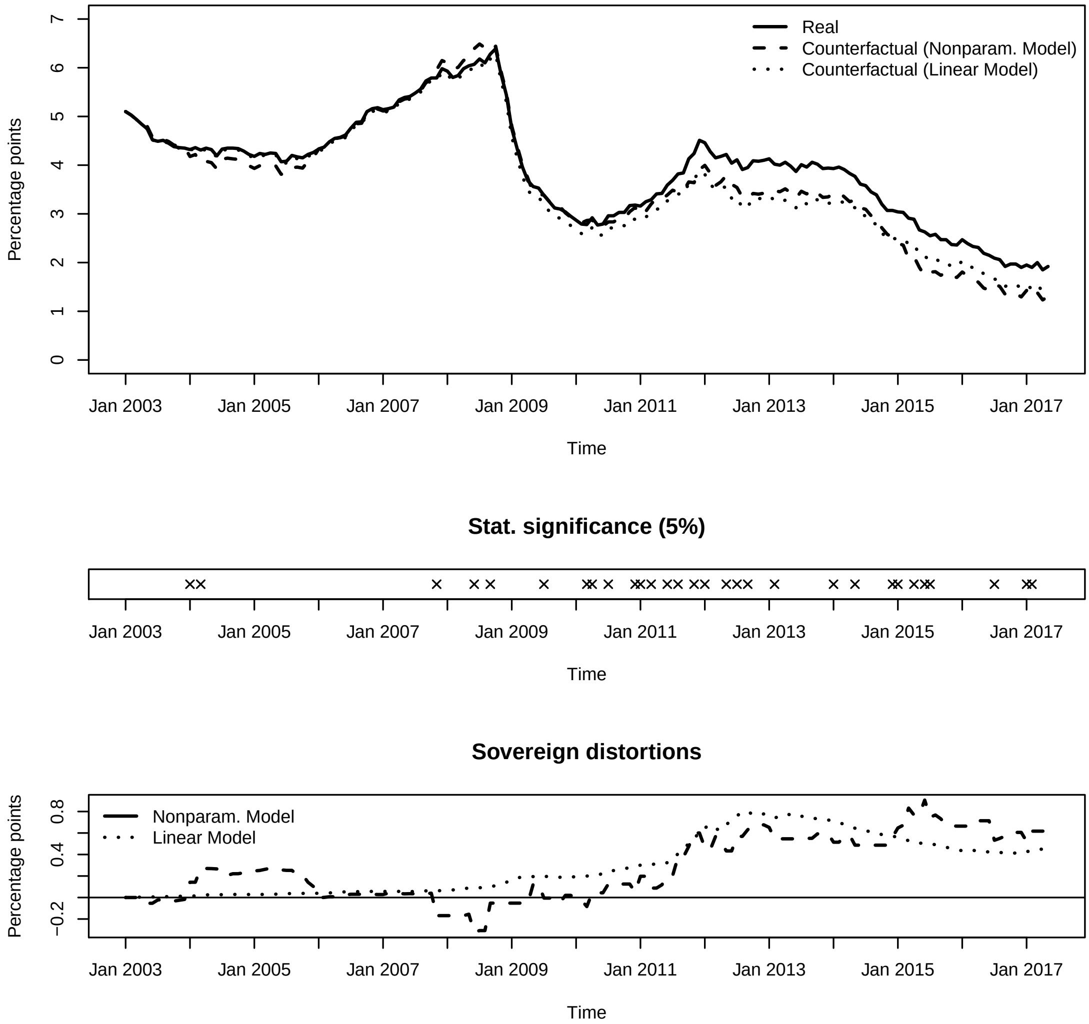

Figure 2: Sovereign risk transmission onto corporate cost of borrowing in Italy.

Notes: Top figure plots the realized and the counterfactual costs of borrowing, estimated by nonparametric and linear frameworks. Middle figure shows periods when the difference between the observed and nonparametric counterfactuals was statistically significant at 5% level according to the standard confidence intervals. Bottom figure shows the sovereign risk contributions derived as a difference between the observed and counterfactual rates.

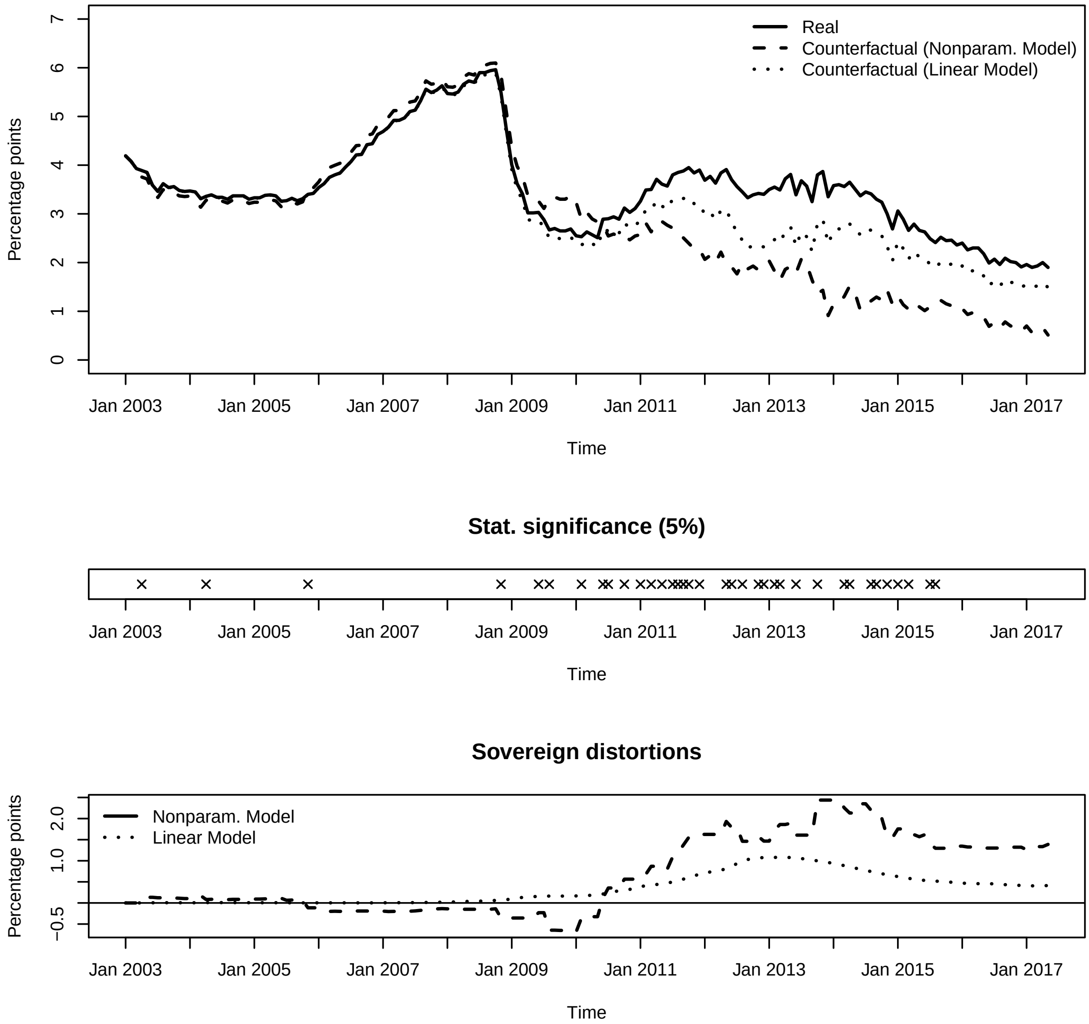

Figure 3: Sovereign risk transmission onto corporate cost of borrowing in Spain.

Notes: Top figure plots the realized and the counterfactual costs of borrowing, estimated by nonparametric and linear frameworks. Middle figure shows periods when the difference between the observed and nonparametric counterfactuals was statistically significant at 5% level according to the standard confidence intervals. Bottom figure shows the sovereign risk contributions derived as a difference between the observed and counterfactual rates.

financial crisis and sovereign debt crisis between the countries. The majority of sovereign risk distortions in France was present before 2009, around the period of the global financial crisis. To the contrary, in Spain and Italy we observe highly significant effects in the period after 2009, largely corresponding to the sovereign debt crisis. These findings are in line with EIB (2016). Interestingly, the counterfactual exercise reveals that the levels of sovereign risk distortions are still elevated in more recent years in all the countries. It suggests that despite the reduction in sovereign yields, there are still non-negligible transmissions to the corporate costs of borrowing.

Thirdly, one can clearly observe a difference between the dependence filtering approach and linear specification. The latter appears to substantially underestimate the sovereign risk transmission mechanisms for France and Spain, with Italy being more aligned, on average. It suggests that, since the counterfactual approach does not assume any underlying model parametrization, the sovereign risk transmission may be highly nonlinear in nature.

4 Conclusions

This study proposes a novel dependence filtering framework. Under mild regularity conditions, and without assuming any specific parametric structure, we derive counterfactual distributions which are either independent or conditionally independent from the effects of given covariates. We show that the methodology has desired asymptotic properties and can be easily applied in a multivariate setting. We provide bootstrap confidence intervals and we additionally confirm the accuracy of the method numerically, on an example of a stylized nonlinear process with volatility spillovers between variables.

In an empirical application, we demonstrate that the proposed framework can be easily adjusted to match standard parametric regressions. In particular, we look at the interest rate pass-through model, widely used by central bankers (Darracq Paries et al., 2014), and we demonstrate that the (conditionally) independent counterfactual distributions can be used to estimate the degree of sovereign risk distortions onto corporate borrowing costs.

The main findings confirm the heterogeneity in the ECB interest rate pass-through between France, Italy and Spain. Interestingly, we find that the standard linear models can underestimate the sovereign risk distortions, especially for the more recent periods. One can speculate that the elevated risk levels of France can coincide with the implicit fiscal backstop role it plays in the EU, together with Germany. We additionally find that the sovereign risk contagion was less of a problem on the short end.

The results build an interesting avenue for policy makers and practitioners. With non-responsive retail rates the standard monetary policy actions are largely ineffective. The fact that the ECB operates near to the zero-lower bound environment makes the bank lending only more difficult to stimulate as the conventional tools run out of scope. Strategies which could improve the effectiveness of the transmission channels should gain more attention. These could include policies aiming at restoring market confidence, cross-border risk-sharing and fostering financial integration in the EA.

To get a better understanding of the influence of risk factors on bank-specific characteristics and lending, further research is therefore desired. In particular, one could extrapolate the filtering methodology to a panel framework or fit the pass-through model to a different set of risk proxies, controlling for, for instance, macroeconomic and financial risks.

References

N. Apergis and A. Cooray. Asymmetric interest rate pass-through in the U.S., the U.K. and Australia: New evidence from selected individual banks. *Journal of Macroeconomics*, 45:155 – 172, 2015. doi: 10.1016/j.jmacro.2015.04.010.

I. J. Arnold and S. E. van Ewijk. A state space approach to measuring the impact of sovereign and credit risk on interest rate convergence in the euro area. *Journal of International Money and Finance*, 49, Part B:340 – 357, 2014. doi: 10.1016/j.jimonfin.2014.04.005.

P. Asdrubali and S. Signore. The economic impact of EU guarantees on credit to SMEs. Technical Report 2015/29, European Investment Fund, 2015.

V. Chernozhukov, I. Fernndez-Val, and B. Melly. Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268, 2013. doi: 10.3982/ECTA10582.

E. Choi and P. Hall. Data sharpening as a prelude to density estimation. *Biometrika*, 86(4): 941–947, 1999. doi: 10.1093/biomet/86.4.941.

M. Darracq Paries, D. Moccero, E. Krylova, and C. Marchini. The retail bank interest rate pass-through: The case of the euro area during the financial and sovereign debt crisis. Technical Report no 154, European Central Bank, 2014.

C. Diks and V. Panchenko. A new statistic and practical guidelines for nonparametric Granger causality testing. *Journal of Economic Dynamics and Control*, 30(9–10):1647–1669, 2006. doi: 10.1016/j.jedc.2005.08.008.

C. Diks and M. Wolski. Nonlinear Granger causality: Guidelines for multivariate analysis. *Journal of Applied Econometrics*, 31(7):1333–1351, 2016. doi: 10.1002/jae.2495.

ECB. Assessing the retail bank interest rate pass-through in the euro area at times of financial fragmentation. Technical Report August, European Central Bank, 2013.

B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Monographs on Statistics & Applied Probability, Chapman and Hall/CRC: London – New York, 1993.

EIB. *Investment and Investment Finance in Europe*. European Investment Bank, Luxembourg, 2016.

F. Ferraty, A. Laksaci, A. Tadj, and P. Vieu. Rate of uniform consistency for nonparametric estimates with functional variables. *Journal of Statistical Planning and Inference*, 140(2):335 – 352, 2010. doi: 10.1016/j.jspi.2009.07.019.

W. A. Fuller. *Introduction to Statistical Time Series (Wiley Series in Probability and Mathematical Statistics)*. Wiley-interscience: New York, 1995. doi: 10.1002/9780470316917.

P. J. Gertler, S. Martinez, P. Premand, L. B. Rawlings, and C. M. J. Vermeersch. *Impact Evaluation in Practice*. World Bank, Washington, D.C., 2011.

E. Giné and R. Nickl. Uniform central limit theorems for kernel density estimators. *Probability Theory and Related Fields*, 141(3):333–387, 2008. doi: 10.1007/s00440-007-0087-9.

P. Hall and M. C. Minnotte. High order data sharpening for density estimation. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64(1):141–157, 2002. doi: 10.1111/1467-9868.00329.

Z. He and A. Krishnamurthy. A macroeconomic framework for quantifying systemic risk. Working Paper 233, National Bank of Belgium, 2012.

C. Hiemstra and J. D. Jones. Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance*, 49(5):1639–1664, 1994. doi: 10.2307/2329266.

X. Huang, H. Zhou, and H. Zhu. Systemic risk contributions. Working Paper 60, Bank for International Settlements, 2010.

M. C. Jones. Estimating densities, quantiles, quantile densities and density quantiles. *Ann Inst Stat Math*, 44(4):721–727, 1992. doi: 10.1007/BF00053400.

T. Kitamura, I. Muto, and I. Takei. Loan interest rate pass-through and changes after the financial crisis: Japans evidence. *Journal of the Japanese and International Economies*, 42:10 – 30, 2016. doi: 10.1016/j.jjie.2016.10.002.

C. Kok and T. Werner. Bank interest rate pass-through in the euro area: a cross country comparison. Working Paper Series 580, European Central Bank, 2006.

Q. Li and J. S. Racine. *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Princeton and Oxford, 2007.

Q. Li and J. S. Racine. Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business & Economic Statistics*, 26(4): 423–434, 2008. doi: 10.1198/073500107000000250.

Q. Li and J. S. Racine. Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *Journal of Business & Economic Statistics*, 31(1):57–65, 2013. doi: 10.1080/07350015.2012.738955.

C. Rothe. Nonparametric estimation of distributional policy effects. *Journal of Econometrics*, 155 (1):56 – 70, 2010. doi: 10.1016/j.jeconom.2009.09.001.

C. Rothe. Partial distributional policy effects. *Econometrica*, 80(5):2269–2301, 2012.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis.* Monographs on Statistics & Applied Probability, Chapman and Hall/CRC: London – New York, 1998.

A. W. van der Vaart. Weak convergence of smoothed empirical processes. *Scandinavian Journal of Statistics*, 21(4):501–504, 1994.

A. W. van der Vaart. *Asymptotic Statistics.* Cambridge University Press, Cambridge, 2000.

A. W. van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes.* Springer, New York, 1996.

A. W. van der Vaart and J. A. Wellner. Empirical processes indexed by estimated functions. *Lecture Notes-Monograph Series*, 55:234–252, 2007.

M. P. Wand and M. C. Jones. *Kernel Smoothing.* Monographs on Statistics & Applied Probability, Chapman and Hall: New York, 1995.

A Proofs

A.1 Theorem 1

Proof. Let us write that the filtered sample $\{Y_i'\}$ is defined as a function of $\{Y_i\}$ conditional on realization of $\{\boldsymbol{X}_i\}$, for each $i = 1, ..., n$, such that $Y_i' = \phi(Y_i|\boldsymbol{X}_i = \boldsymbol{x_i}) \equiv \phi(Y_i)$. Let us further assume that ϕ is invertible around point y. Transformation $\phi : \mathbb{R} \to \mathbb{R}$ might be viewed as a perturbation, or following Choi and Hall (1999) and Hall and Minnotte (2002), a sharpening function. Consequently, the plug-in estimator of the Left-Hand Side (LHS) of Eq. (1), which is defined over Y' marginal, can be rewritten as

$$\hat{f}_{Y',\boldsymbol{X}}(y, \boldsymbol{x}) = n^{-1} \sum_{i=1}^{n} K_{\mathbf{H}} \left(y - \phi(Y_i), \boldsymbol{x} - \boldsymbol{X_i}\right), \tag{A.1.1}$$

where $\mathbf{H}$ is a $(1+d_X) \times (1+d_X)$ bandwidth matrix and $K_{\mathbf{H}}$ is a scaled multivariate kernel function satisfying the standard regularity conditions (see Assumption 2 or Wand and Jones (1995)). For presentation, in notation we explicitly distinguish between variables Y and $\boldsymbol{X}$ as sharpening is defined over the Y marginal only. Consequently, let us assume that bandwidth matrix consists of two blocks: one for marginal Y and one for $\boldsymbol{X}$.

By taking the expected value of the estimator from Eq. (A.1.1) with respect to the continuous joint distribution of $(Y, \boldsymbol{X})$, and substituting into Eq. (1) we get

$$\int K_{\mathbf{H}} \left(y - \phi(Y), \boldsymbol{x} - \boldsymbol{X}\right) \mathrm{d}F(Y, \boldsymbol{X}) = f_Y(y) f_{\boldsymbol{X}}(\boldsymbol{x}), \tag{A.1.2}$$

as the sample size $n \to \infty$. By substituting $V = \phi(Y)$

$$\int K_{\mathbf{H}} \left(y - V, \boldsymbol{x} - \boldsymbol{X}\right) \mathrm{d}F(\phi^{-1}(V), \boldsymbol{X}) = f_Y(y) f_{\boldsymbol{X}}(\boldsymbol{x}). \tag{A.1.3}$$

Changing variables we observe that

$$\int K\left(u, \boldsymbol{z}\right) g\left(y - \mathbf{H}_Y^{1/2} u \middle| \boldsymbol{x} - \mathbf{H}_{\boldsymbol{X}}^{1/2} \boldsymbol{z}\right) \mathrm{d}\boldsymbol{z}\mathrm{d}u = f_Y(y) f_{\boldsymbol{X}}(\boldsymbol{x}). \tag{A.1.4}$$

where $\mathbf{H}_Y$ and $\mathbf{H}_X$ are the block matrices associated with marginals Y and $\boldsymbol{X}$, respectively. We also defined $g = \partial G/\partial a$ where

$$G(a|\boldsymbol{b}) = \int_{-\infty}^{a} f_{Y,\boldsymbol{X}}(\phi^{-1}(u), \boldsymbol{b}) \mathrm{d}u. \tag{A.1.5}$$

Assuming g is sufficiently smooth, the LHS of Eq. (A.1.4) is equal $g(y|\boldsymbol{x}) + o(\mathrm{tr}(\mathbf{H}))$. By substituting it back into Eq. (A.1.4) and integrating, we conclude that $G + o(\mathrm{tr}(\mathbf{H})) = D$ where

$$D(y|\boldsymbol{x}) = F_Y(y) f_{\boldsymbol{X}}(\boldsymbol{x}). \tag{A.1.6}$$

Combining it with Eq. (A.1.5), up to the order of $o(\mathrm{tr}(\mathbf{H}))$, we get

$$\int_{-\infty}^{y} f_{Y,\boldsymbol{X}}(\phi^{-1}(u), \boldsymbol{x}) \mathrm{d}u = D(y|\boldsymbol{x}). \tag{A.1.7}$$

Denote LHS of Eq. (A.1.7) as $J(\phi^{-1}(y))$ for any point $\boldsymbol{X} = \boldsymbol{x}$. Consequently, one finds that $\phi(y) = D^{-1}(J(y)|\boldsymbol{x})$ or

$$\begin{aligned}\phi(y) &= D^{-1}\left(\int_{-\infty}^{y} f_{Y,\boldsymbol{X}}(u,\boldsymbol{x})\mathrm{d}u\right)\\ &= F_Y^{-1}\left(\frac{\int_{-\infty}^{y} f_{Y,\boldsymbol{X}}(u,\boldsymbol{x})\mathrm{d}u)}{f_{\boldsymbol{X}}(\boldsymbol{x})}\right),\end{aligned} \tag{A.1.8}$$

or alternatively, by substituting $y' \equiv \phi(y)$,

$$F_Y(y') = F_{Y|\boldsymbol{X}}(y|\boldsymbol{x}). \tag{A.1.9}$$

□

A.2 Theorem 2

Proof. To derive the asymptotic properties of the process indexed by Eq. (11), let us introduce the basic concepts from the (smooth) empirical process literature. Therefore, let $Z_1, ..., Z_n$ be an i.i.d. random elements in a measurable space $(\mathcal{Z}, \mathcal{A})$ with law (distribution) P, and for a measurable function $g : \mathcal{Z} \to \mathbb{R}$ let the expectation, the empirical measure and empirical process at g be denoted by

$$Pg = \int g\mathrm{d}P, \qquad \mathbb{P}_n g = \frac{1}{n}\sum_{i=1}^{n} g(Z_i), \qquad \mathbb{G}_n g = \sqrt{n}(\mathbb{P}_n - P)g. \tag{A.2.1}$$

A.2.1 Empirical measures

Let us firstly consider empirical measures, i.e. not smoothed, of F_Y and $F_{Y|\boldsymbol{X}}$. Consequently, in the notation in this section we treat the estimators as defined over non-convoluted probability space. Let us write

$$\sqrt{n}(\hat{Y}' - Y') = \sqrt{n}(\hat{F}_Y^{-1} \circ \hat{F}_{Y|\boldsymbol{X}} - F_Y^{-1} \circ F_{Y|\boldsymbol{X}}). \tag{A.2.2}$$

We observe that $\hat{F}_{Y|\boldsymbol{X}}$ can be viewed as a pseudo-observation, i.e. an estimate rather than a point realization. After van der Vaart and Wellner (2007), we note that $F_Y^{-1} \circ F_{Y|\boldsymbol{X}}$ is Donsker under Assumption 3, and we can decompose Eq. (A.2.2) as

$$\begin{aligned}\sqrt{n}(\hat{F}_Y^{-1} \circ \hat{F}_{Y|\boldsymbol{X}} - F_Y^{-1} \circ F_{Y|\boldsymbol{X}}) &= \mathbb{G}_n(F_Y^{-1} \circ \hat{F}_{Y|\boldsymbol{X}} - F_Y^{-1} \circ F_{Y|\boldsymbol{X}}) + \mathbb{G}_n F_Y^{-1} \circ F_{Y|\boldsymbol{X}}\\ &+ \sqrt{n}P(F_Y^{-1} \circ \hat{F}_{Y|\boldsymbol{X}} - F_Y^{-1} \circ F_{Y|\boldsymbol{X}})\\ &= A1 + A2 + A3.\end{aligned} \tag{A.2.3}$$

Before handling Eq. (A.2.3) it is useful to remind that under Assumptions 1-4, both empirical quantile function and empirical conditional distribution functions are uniformly consistent, i.e.

$$\sup_{y\in\mathbb{R}} \sup_{\boldsymbol{x}\in\mathbb{R}^d} \left|\hat{F}_{Y|\boldsymbol{X}}(y|\boldsymbol{x}) - F_{Y|\boldsymbol{X}}(y|\boldsymbol{x})\right| = O_P\left(\left(\frac{\log n}{nh^{d_{\boldsymbol{X}}}}\right)^{1/2} + h^r\right), \tag{A.2.4}$$

$$\sup_{w\in J} \left|\hat{F}_Y^{-1}(w) - F_Y^{-1}(w)\right| = O_P\left(\left(\frac{\log n}{n}\right)^{1/2}\right), \tag{A.2.5}$$

where J is a compact subset of $\mathbb{R}$ (Li and Racine, 2007).

To deal with term $A1$ we observe that, by Assumption 4 (ii), if F_Y^{-1} is uniformly bounded by $[a,b] = [F_Y^{-1}(p_1) - \epsilon, F_Y^{-1}(p_2) + \epsilon]$ for some $\epsilon > 0$ where $0 < p_1 < p_2 < 1$, and that F_Y^{-1} is uniformly Lipschitz, i.e.

$$\left|F_Y^{-1}(r_1) - F_Y^{-1}(r_2)\right| \leq C\left|r_1 - r_2\right|, \tag{A.2.6}$$

for every pair of points $r_1, r_2 \in [p_1, p_2]$ and a constant $C > 0$. Following van der Vaart and Wellner (1996) (page 157), functional F_Y^{-1} has finite uniform entropy integral and by Theorem 3.1 of van der Vaart and Wellner (2007)

$$\sup_{y \in J}\left|\mathbb{G}_n\left(F_Y^{-1}\left(\hat{F}_{Y|\boldsymbol{X}}\right) - F_Y^{-1}\left(F_{Y|\boldsymbol{X}}\right)\right)\right| \xrightarrow{p} 0, \tag{A.2.7}$$

in $\ell^\infty(J, \mathbb{R})$ where $J \subset \mathbb{R}$ for which $F_{Y|\boldsymbol{X}}$ is uniformly bounded by $[p_1, p_2]$.

To handle terms $A2$ and $A3$ we note that the quantile function is defined over distribution F_Y, i.e. for $\Gamma_Y \equiv F_Y^{-1}$, the empirical quantile function is given by

$$\Gamma_Y(\hat{F}_Y(\tau)) = \inf\{y : \hat{F}_Y(y) \geq \tau\}, \tag{A.2.8}$$

with $\Gamma_Y : D[a,b] \to \ell^\infty[p_1, p_2]$ where $D[a,b]$ is the set of cadlag functions on $[a,b]$. Term $A2$ is the usual empirical process for i.i.d. one-dimensional transformation of a random variable, given by $F_{Y|\boldsymbol{X}}(Y_1|\boldsymbol{x}), \ldots, F_{Y|\boldsymbol{X}}(Y_n|\boldsymbol{x})$, and it converges weakly by Eq. (A.2.5). Since, by Assumption 4, Γ_Y is Hadamard differentiable at F_Y tangentially to $S[a,b]$, where $S[a,b]$ is the set of continuous functions on $[a,b]$, by the distributive property of considered functions, functional delta method and continuous mapping theorem, respectively, we have that

$$\begin{aligned}
\mathbb{G}_n F_Y^{-1} \circ F_{Y|\boldsymbol{X}} &= \sqrt{n}(\hat{F}_Y^{-1} - F_Y^{-1}) \circ F_{Y|\boldsymbol{X}} \\
&= \sqrt{n}(\Gamma_Y(\hat{F}_Y) - \Gamma_Y(F_Y)) \circ F_{Y|\boldsymbol{X}} \\
&= \Gamma_Y'(\sqrt{n}(\hat{F}_Y - F_Y)) \circ F_{Y|\boldsymbol{X}} + o_P(1) \\
&= \Gamma_Y'(\mathbb{G}_{n,F_Y}) \circ F_{Y|\boldsymbol{X}} + o_P(1) \\
&\xrightarrow{d} \Gamma_Y'(\mathbb{G}_{F_Y}) \circ F_{Y|\boldsymbol{X}},
\end{aligned} \tag{A.2.9}$$

where $\mathbb{G}_{F_Y}$ is the F_Y-Brownian bridge, defined as $\mathbb{G}_{F_Y} = \mathbb{G}_\lambda \circ F_Y$, where $\mathbb{G}_\lambda$ is the standard Brownian bridge process, and convergence is in $\ell^\infty(J)$. As a result we have that

$$A2 \xrightarrow{d} N\left(0, \frac{F_Y(1 - F_Y)}{f_Y^2(F_Y^{-1}(F_{Y|\boldsymbol{X}}))}\right), \tag{A.2.10}$$

and convergence is in $\ell^\infty(J, \mathbb{R})$.

Following similar reasoning for the term $A3$, we have

$$\begin{aligned}
\sqrt{n}P(F_Y^{-1}(\hat{F}_{Y|\boldsymbol{X}}) - F_Y^{-1}(F_{Y|\boldsymbol{X}})) &= P\left[\sqrt{n}\left(\Gamma_Y\left(F_Y(\hat{F}_{Y|\boldsymbol{X}})\right) - \Gamma_Y\left(F_Y(F_{Y|\boldsymbol{X}})\right)\right)\right] \\
&= P\left[-\Gamma_Y'\left(\sqrt{n}(\hat{F}_{Y|\boldsymbol{X}} - F_{Y|\boldsymbol{X}})\right)\right] + o_P(1) \\
&= P\left[-\Gamma_Y'\left(\mathbb{G}_{n,F_{Y|\boldsymbol{X}}}\right)\right] + o_P(1)
\end{aligned} \tag{A.2.11}$$

where we exploited the fact that $\mathrm{d}/\mathrm{d}F_{Y|\boldsymbol{X}}\Gamma_Y(F_Y) = -\Gamma'_Y$. We further build an argument along the lines of Corollaries 5.1 and 5.3 of van der Vaart and Wellner (2007). By Hadamard differentiability assumption we remind that

$$P\left[-\Gamma'_Y\left(\mathbb{G}_{F_{Y|\boldsymbol{X}}}\right)\right] = L(\sqrt{n}(\eta_n - \eta)), \tag{A.2.12}$$

where L is continuous and linear map and

$$\eta_n - \eta \equiv \eta_n(y, \boldsymbol{x}) - \eta(y, \boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{I}\{Y_i \leq y\}\hat{\theta}_{\mathbf{H}}(x, X_i) - F_{Y|\boldsymbol{X}}(y|x)\right). \tag{A.2.13}$$

In line with Lemma 4.2 from van der Vaart and Wellner (2007), we have that

$$\begin{aligned} L(\sqrt{n}(\eta_n - \eta)) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} L\left(\mathbb{I}\{Y_i \leq y\}\hat{\theta}_{\mathbf{H}}(x, X_i) - F_{Y|\boldsymbol{X}}(y|x)\right) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} L\left(\mathbb{I}\{Y_i \leq y\}\hat{\theta}_{\mathbf{H}}(x, X_i)\right) - \sqrt{n}L(\eta) \\ &= \mathbb{G}_\eta L\left(\mathbb{I}\{Y_i \leq y\}\hat{\theta}_{\mathbf{H}}(x, X_i)\right) \end{aligned} \tag{A.2.14}$$

Substituting back to the original notation, and expanding function Γ'_Y, we receive

$$\begin{aligned} P\left[-\Gamma'_Y\left(\mathbb{G}_{F_{Y|\boldsymbol{X}}}\right)\right] &= \frac{\mathbb{G}_{F_{Y|\boldsymbol{X}}}}{f_Y\left(F_Y^{-1}\left(E\left[\mathbb{I}\{Y \leq y\}\hat{\theta}_{\mathbf{H}}(x, X)\right]\right)\right)} \\ &= \frac{\mathbb{G}_{F_{Y|\boldsymbol{X}}}}{f_Y\left(F_Y^{-1}\left(F_{Y|\boldsymbol{X}}\right)\right)} + o(\mathrm{tr}(\mathbf{H})), \end{aligned} \tag{A.2.15}$$

and convergence is in $\ell^\infty(J, \mathbb{R})$. By the properties of a Brownian bridge it follows that

$$A3 \xrightarrow{d} N\left(0, \frac{F_{Y|\boldsymbol{X}}(1 - F_{Y|\boldsymbol{X}})}{f_Y^2(F_Y^{-1}(F_{Y|\boldsymbol{X}}))}\right). \tag{A.2.16}$$

Putting the results together we find that $\sqrt{n}(\hat{Y}' - Y)$ is asymptotically tight for the empirical (non-smoothed) estimators.

A.2.2 Smooth empirical measures

To extend the results to the convoluted probability spaces, let us firstly remind that, in the current setup, the probability law $\mathbb{P}$ has a density with respect to the Lebesgue measure. The smooth empirical measure and smooth empirical process indexed at g can be written as

$$\tilde{\mathbb{P}}_n g = \mathbb{P}_n * \mu_n(g), \qquad \tilde{\mathbb{G}}_{\mu_n} g = \sqrt{n}(\mathbb{P}_n - P) * \mu_n(g), \tag{A.2.17}$$

where μ_n is a sequence of probability measures converging weakly to zero. Following van der Vaart (1994), in our setup μ_n is random and satisfies $\mathrm{d}\mu_n(\boldsymbol{w}) = K_{\mathbf{H}}(\boldsymbol{w})\mathrm{d}\boldsymbol{w}$. Since convoluted estimates converge in MSE, conditions in Eqs. (A.2.4) and (A.2.5) are also satisfied for smooth estimates. (In this section all the estimates are taken as smooth.)

It is useful to restate the uniform convergence rates for the smooth empirical measures, as

$$\sup_{y \in \mathbb{R}} \sup_{\boldsymbol{x} \in \mathbb{R}^d} \left|\hat{F}_{Y|\boldsymbol{X}}(y|\boldsymbol{x}) - F_{Y|\boldsymbol{X}}(y|\boldsymbol{x})\right| = O_P\left(\left(\frac{\log n}{nh^{d_{\boldsymbol{X}}}}\right)^{1/2} + h^r + h_0^r\right), \tag{A.2.18}$$

$$\sup_{w \in J} \left| \hat{F}_Y^{-1}(w) - F_Y^{-1}(w) \right| = O_P \left(\left(\frac{\log n}{n} \right)^{1/2} + h_0^r \right). \tag{A.2.19}$$

Following the argument under Theorem 1 proposed by Rothe (2010), we observe that the classes of functions $\mathcal{F}_1$ and $\mathcal{F}_2$, given by $\mathcal{F}_1 = \{F_Y(y) : y \in \mathbb{R}\}$ and $\mathcal{F}_2 = \{F_{Y|\boldsymbol{X}}(y, .) : y \in \mathbb{R}\}$ are r-times differentiable by Assumption 1. Hence, $\mathcal{F}_1 \in C^r(J)$ and $\mathcal{F}_2 \in C^r(J, \mathbb{R})$ with bounded derivatives up to order r. Class $\mathcal{F}$ is Donsker by definition since $d_Y = 1$ in our setting, and class $\mathcal{F}_2$ is Donsker for $r > (d_{\boldsymbol{X}})/2$, which is satisfied as Assumption 3 implies that $r > d_{\boldsymbol{X}}$, as $\mathbf{H}_0$ vanishes faster than $\mathbf{H}$. In fact, this condition allows for application of the usual second order Kernels in the bi-variate settings with $d_Y = d_{\boldsymbol{X}} = 1$.

Since the function $F_Y^{-1} \circ F_{Y|\boldsymbol{X}}$ is Donsker, and together with Assumptions 1-3, by Theorem 5 of Giné and Nickl (2008) we have that $\tilde{\mathbb{G}}_{\mu_n} F_Y$ converges in probability in $\ell^\infty(J)$ to the F_Y-Brownian bridge, and $\tilde{\mathbb{G}}_{\mu_n} F_{Y|\boldsymbol{X}}$ converges in probability in $\ell^\infty(J, \mathbb{R})$ to the $F_{Y|\boldsymbol{X}}$-Brownian bridge. As a consequence, all the implications from Eqs. (A.2.7), (A.2.9) and (A.2.15) hold for convoluted probability spaces. In particular we have for the term $A3$ we have

$$\begin{aligned} P\left[-\Gamma_Y'\left(\tilde{\mathbb{G}}_{F_{Y|\boldsymbol{X}}}\right)\right] &= \frac{\mathbb{G}_{F_{Y|\boldsymbol{X}}}}{f_Y\left(F_Y^{-1}\left(E\left[\bar{K}_{H_0}(y-Y)\hat{\theta}_{\mathbf{H}}(x, X)\right]\right)\right)} \\ &= \frac{\mathbb{G}_{F_{Y|\boldsymbol{X}}}}{f_Y\left(F_Y^{-1}\left(F_{Y|\boldsymbol{X}}\right)\right)} + o(H_0 + \operatorname{tr}(\mathbf{H})), \end{aligned} \tag{A.2.20}$$

As a result, we confirm that also for smooth estimates $\sqrt{n}(\hat{Y}' - Y)$ converges in distribution to a zero-mean Gaussian variable with variance as stated in Theorem 2. □

A.3 Theorem 3

Proof. Proof follows directly from the delta method for bootstrap. We observe that both F_Y^{-1} and $F_{Y|\boldsymbol{X}}$ are Donsker and that $\hat{F}_Y^{-1}$ and $\hat{F}_{Y|\boldsymbol{X}}$ converge uniformly to the true functionals. Then in line with Theorem 23.5 of van der Vaart (2000), the proof follows the stages outlined for Theorem 2. □

B Numerical results

B.1 Independent counterfactuals

The goal of the numerical exercise is to verify the performance of the procedure and to assess the advantages of smoothing techniques in constructing independent counterfactual distributions. Below we investigate the non-smooth against smooth rule-of-thumb bandwidth selectors, however, an extensive numerical investigation covering process and data-driven bandwidths are available upon request. We also verify whether the independence between Y' and X holds in the hypothesis testing framework.

Let us consider a time-series setting with a bivariate ARCH process given by

$$\begin{aligned} X_i &\sim \mathrm{N}\left(0, 1\right), \\ Y_i &\sim \mathrm{N}\left(0, d + aX_i^2\right). \end{aligned} \tag{B.1.1}$$

The process is set up such that $\{X_i\}$ is affecting the second moment of $\{Y_i\}$ and it is similar to the one considered by Diks and Wolski (2016). To satisfy the criteria set by Assumption 1, we require that $d > 0$ and $0 < a < 1$. The setting builds an interesting testing environment as it allows for nonlinear effects, which standard linear models may find difficult to capture. For the simulations we choose $d = 1$ and $a = 0.4$.

In line with Li and Racine (2013), we test the accuracy of $\hat{Y}'$ against true values obtained using numerical integration directly from Eq. (B.1.1). The sample MSE is therefore given by

$$\mathrm{MSE}(\hat{Y}') = n^{-1} \sum_{i=1}^{n} \left(\hat{F}_Y^{-1}(\hat{F}_{Y|\boldsymbol{X}}^{-i}(y|\boldsymbol{x})) - F_Y^{-1}(F_{Y|\boldsymbol{X}}(y|\boldsymbol{x})) \right)^2 . \tag{B.1.2}$$

In the filtering procedure, as well as to calculate the MSE estimates, we use leave-one-out kernel estimators given by $\hat{F}_{\boldsymbol{W}}^{-i}(\boldsymbol{w_i}) = (n-1)^{-1} \sum_{i \neq j} \bar{K}_{\mathbf{H}}(w_i - w_j)$ for PDFs and by $\hat{F}_{V|\boldsymbol{W}}^{-i}(v_i|\boldsymbol{w_i}) = (n-1)^{-1} \sum_{i \neq j} \bar{K}_{\mathrm{H}_0}(v_i - v_j) K_{\mathbf{H}}(w_i - w_j) / \hat{f}_{\boldsymbol{W}}^{-i}(\boldsymbol{w}_i)$ for conditional CDFs. We apply the standard Gaussian kernel of order $r = 2$. In line with Assumption 4, we restrict the support of Y' to fall within $[-3.7\hat{\sigma}_y, 3.7\hat{\sigma}_y]$. When a filtered realization $\hat{y}'$ cannot be found numerically for some $i = 1, ..., n$, we set $\hat{y}_i' \equiv y_i$.

We compare two reference estimators. Firstly, we use the empirical CDF and conditional CDF functions, with smoothing only along X marginal. As demonstrated in the proof of Theorem 1, it is a first-line testing benchmark, which relates to the estimators used by Rothe (2010).

Secondly, we test a normal-reference bandwidth selector. After Silverman (1998), a simple rule of thumb, which assumes normality and mean-dependence of the underlying data generating processes and normal kernel, suggests that along marginals Y one should use $c_0^{Y|X} = c_0^Y = 1.59\hat{\sigma}_y$, for the unconditional and conditional distributions, respectively. At the same time, along X marginal we have $c_1 = 1.06\hat{\sigma}_x$, where $\hat{\sigma}_y$ and $\hat{\sigma}_x$ correspond to empirical dispersion measures of Y and X, respectively.

We run 1000 Monte Carlo replications of process in Eq. (B.1.1) for different sample sizes. True filtered realizations y' are calculated using numerical methods. The medians and standard deviation intervals, aggregated over all the runs, are presented in Tables B1 and B2.

Table B1: Median MSE estimates of independent counterfactual distributions.

Bandwidth selector	n=50	n=100	n=200	n=500	n=1000
no smoothing	0.584	0.406	0.274	0.169	0.107
smoothing	0.292	0.232	0.178	0.116	0.080

Notes: Medians taken over 1000 Monte Carlo results of Eq. (B.1.2) for the ARCH process given in Eq. (B.1.1). Bandwidth selectors are chosen as: 'no smoothing' takes $h_1^{Y|X} = 1.06\hat{\sigma}_x n^{-1/3}$, *'smoothing' takes* $h_0^Y = 1.59\hat{\sigma}_y n^{-1/3}$, $h_0^{Y|X} = 1.59\hat{\sigma}_y n^{-1/3}$ *and* $h_1^{Y|X} = 1.06\hat{\sigma}_x n^{-1/3}$.

Table B2: MSE dispersion of estimates of independent counterfactual distributions.

Bandwidth selector	n=50	n=100	n=200	n=500	n=1000
no smoothing	0.236	0.134	0.076	0.039	0.023
smoothing	0.142	0.092	0.052	0.026	0.015

Notes: Standard deviations taken over 1000 Monte Carlo results of Eq. (B.1.2) for the ARCH process given in Eq. (B.1.1). Bandwidth selectors are chosen as: 'no smoothing' takes $h_1^{Y|X} = 1.06\hat{\sigma}_x n^{-1/3}$, *'smoothing' takes* $h_0^Y = 1.59\hat{\sigma}_y n^{-1/3}$, $h_0^{Y|X} = 1.59\hat{\sigma}_y n^{-1/3}$ *and* $h_1^{Y|X} = 1.06\hat{\sigma}_x n^{-1/3}$.

One observes that smooth estimates deliver MSE improvement, both in terms of accuracy and dispersion. The gains are particularly visible for smaller samples, and they narrow as the sample size increases. This finding is, in fact, common for nonparametric estimation (Li and Racine, 2007). The efficiency gains come from the fact that empirical CDF provides only a step-wise representation of a true distribution, which can be possibly smooth. Importantly, both estimators converge as the sample size increases.

B.2 Conditionally independent counterfactuals

Last but not least, we complement the methodology by a multivariate simulation study, capitalizing on the results presented in Appendix B.1. To verify the MSE performance of conditionally independent counterfactual distributions, we need a 3-variate process

$$\begin{aligned} X_i &\sim \mathrm{N}(0,1), \\ Q_i &\sim \mathrm{N}(0,1), \\ Y_i &\sim \mathrm{N}\left(0, d + aX_i^2 + aQ_i^2\right). \end{aligned} \tag{B.2.1}$$

The process is set up such that $\{X_i\}$ and $\{Q_i\}$ are influencing the second moment of $\{Y_i\}$. On purpose, we make the dependence structure symmetric, to simplify the derivations for the MSE-optimal process-driven bandwidths. We set $d = 1$ and $a = 0.4$.

Conditionally independent counterfactuals should remove dependence on $\{X_i\}$ while preserving it for $\{Q_i\}$. The baseline estimation identity for each index $j = 1, ..., n$, from which we get the counterfactual draws, is given by

$$\sum_{i\neq j}^{n} \bar{K}_{\mathrm{H}_0^{Y|Q}}\left(y'-Y_i\right)\theta_{\mathbf{H}^{Y|Q}}(\boldsymbol{Q_j},\boldsymbol{Q_i}) = \sum_{i\neq j}^{n} \bar{K}_{\mathrm{H}_0^{Y|Q,X}}(Y_j-Y_i)\theta_{\mathbf{H}^{Y|Q,X}}\left((\boldsymbol{Q_j},\boldsymbol{X_j}),(\boldsymbol{Q_i},\boldsymbol{X_i})\right). \quad \text{(B.2.2)}$$

We exploit two separate block bandwidth matrices, i.e. $\mathrm{H}_{Y|Q}$ and $\mathrm{H}_{Y|Q,X}$. However, we need to account for an increased dimensionality in the estimation. It can be verified that the estimates $\hat{F}_{Y|Q}$ satisfy the uniform convergence rate for $h_0^{Y|Q}, h_1^{Y|Q} \sim n^{-1/3}$ and second-order kernel, in the same spirit as $h_0^{Y|X}$ and $h_1^{Y|X}$ discussed in Section B. For consistency of $\hat{F}_{Y|Q,X}$ we need slower convergence of bandwidth, which can only be obtained with higher-order kernels. Therefore, we employ the fourth-order Gaussian kernels, allowing the bandwidths to converge at rates $h_0^{Y|Q,X}, h_1^{Y|Q,X}, h_2^{Y|Q,X} \sim n^{-1/5}$.

We study the same bandwidth constants selectors as described in Section 2.1. The constants for stepwise and smooth (rule of thumb) estimators correspond in magnitude to the ones used for independent counterfactuals. The MSE results for $\hat{Y}''$ are depicted in Tables B3 and B4.

Table B3: Median MSE estimates of conditionally independent counterfactual distributions.

Bandwidth selector	n=50	n=100	n=200	n=500	n=1000
no smoothing	0.794	0.657	0.545	0.423	0.346
smoothing	0.58	0.53	0.433	0.253	0.142

Notes: Medians taken over 1000 Monte Carlo results of Eq. (B.1.2) for the ARCH process given in Eq. (B.2.1). Bandwidth selectors are chosen as: ‘no smoothing’ takes $h_1^{Y|Q} = 1.06\hat{\sigma}_q n^{-1/3}$, $h_1^{Y|Q,X} = 1.06\hat{\sigma}_q n^{-1/5}$ *and* $h_2^{Y|Q,X} = 1.06\hat{\sigma}_x n^{-1/5}$, *‘smoothing’ takes* $h_0^{Y|Q} = 1.59\hat{\sigma}_y n^{-1/3}$, $h_1^{Y|Q} = 1.06\hat{\sigma}_q n^{-1/5}$, $h_0^{Y|Q,X} = 1.59\hat{\sigma}_y n^{-1/5}$, $h_1^{Y|Q,X} = 1.06\hat{\sigma}_q n^{-1/5}$ *and* $h_2^{Y|Q,X} = 1.06\hat{\sigma}_x n^{-1/5}$.

Table B4: MSE dispersion of estimates of conditionally independent counterfactual distributions.

Bandwidth selector	n=50	n=100	n=200	n=500	n=1000
no smoothing	0.322	0.2	0.125	0.068	0.045
smoothing	0.273	0.183	0.12	0.075	0.048

Notes: Standard deviations taken over 1000 Monte Carlo results of Eq. (B.1.2) for the ARCH process given in Eq. (B.2.1). Bandwidth selectors are chosen as: ‘no smoothing’ takes $h_1^{Y|Q} = 1.06\hat{\sigma}_q n^{-1/3}$, $h_1^{Y|Q,X} = 1.06\hat{\sigma}_q n^{-1/5}$ *and* $h_2^{Y|Q,X} = 1.06\hat{\sigma}_x n^{-1/5}$, *‘smoothing’ takes* $h_0^{Y|Q} = 1.59\hat{\sigma}_y n^{-1/3}$, $h_1^{Y|Q} = 1.06\hat{\sigma}_q n^{-1/5}$, $h_0^{Y|Q,X} = 1.59\hat{\sigma}_y n^{-1/5}$, $h_1^{Y|Q,X} = 1.06\hat{\sigma}_q n^{-1/5}$ *and* $h_2^{Y|Q,X} = 1.06\hat{\sigma}_x n^{-1/5}$.

There are two main results from the exercise, which complement the ones derived in Section B. Firstly, in both setups, increased dimensionality of conditionally independent counterfactuals comes at a cost of lower efficiency of the estimates, both in terms of accuracy and dispersion.

The efficiency gains from smoothing, in terms of median MSE of the estimates, are still substantial. The average MSE difference between ‘no smoothing’ and ‘smoothing’ bandwidth selectors amount to nearly 0.17. It is consistent with the bi-variate numerical results depicted in Section B.

Secondly, it seems that the smooth estimates' dispersion converges at a slower rate than in the step-wise case. The differences are, however, marginal and vanish with the sample size.

C Additional empirical results

C.1 Linear pass-through estimates

Table C5: Standard (linear) pass-through results.

	France	Italy	Spain
EURIBOR (Δ)	0.624***	0.486***	0.615***
	(0.049)	(0.058)	(0.084)
EURIBOR (Δ, lag)	-	0.187**	0.342***
	-	(0.076)	(0.097)
Corporate borrowing cost (Δ, lag)	-0.112*	0.002	-0.28***
	(0.058)	(0.073)	(0.072)
Sovereign risk (Δ, lag)	-0.143*	0.051*	0.01
	(0.075)	(0.031)	(0.046)
ECM (lag)	-0.337***	-0.048***	-0.098***
	(0.05)	(0.017)	(0.031)
R2	0.66	0.62	0.53
Adj. R2	0.65	0.61	0.51
DF	167	166	166

*Notes: Estimation results of the model in Eq. (22) and (23). First-difference and lags operators are marked by Δ and 'lag', respectively. Lags are selected according to the Bayesian Information Criterion. DF describes the residual degrees of freedom. Standard errors in parentheses. Significance codes: *** for 0.01, ** for 0.05 and * for 0.1 levels.*

C.2 Costs of short-term corporate funding

We extend the setup to a set of short-term indicators on the cost of borrowing. Usually short-term funding is concentrated in smaller enterprises which, due to lower transparency and higher credit risk, cannot afford to pay the term premium. Nevertheless, SMEs are a backbone of European economy (EIB, 2016), representing 99.8% of EU companies, nearly 60% of EU-wide GDP, and almost 70% of employment (Asdrubali and Signore, 2015). We therefore re-estimate the pass-through model proposed above on a set of corporate borrowing costs on funding up to 12 months.

The results are given in Figs C1-C3.

Interestingly, we observe that the sovereign distortions affect short-term loans to a lower extent than the corporate loans overall. Overall, the result doesn't seem surprising taking into account the maturity mismatch between the variables. In France, the sovereign stability alleviated some of the short-term corporate borrowing constraints, and actually decreased the borrowing costs by 32pp, on average. In Italy, they elevated the borrowing costs by around 27bp, and in Spain by 43bp, on average. The statistical significance of the sovereign spillovers is largely preserved.

Analysis of the short-term borrowing costs demonstrates an interesting difference between two south-European countries. Whereas the contagion of sovereign risk has been contained in Spain from mid-2015, it has still been elevated in Italy throughout the period.

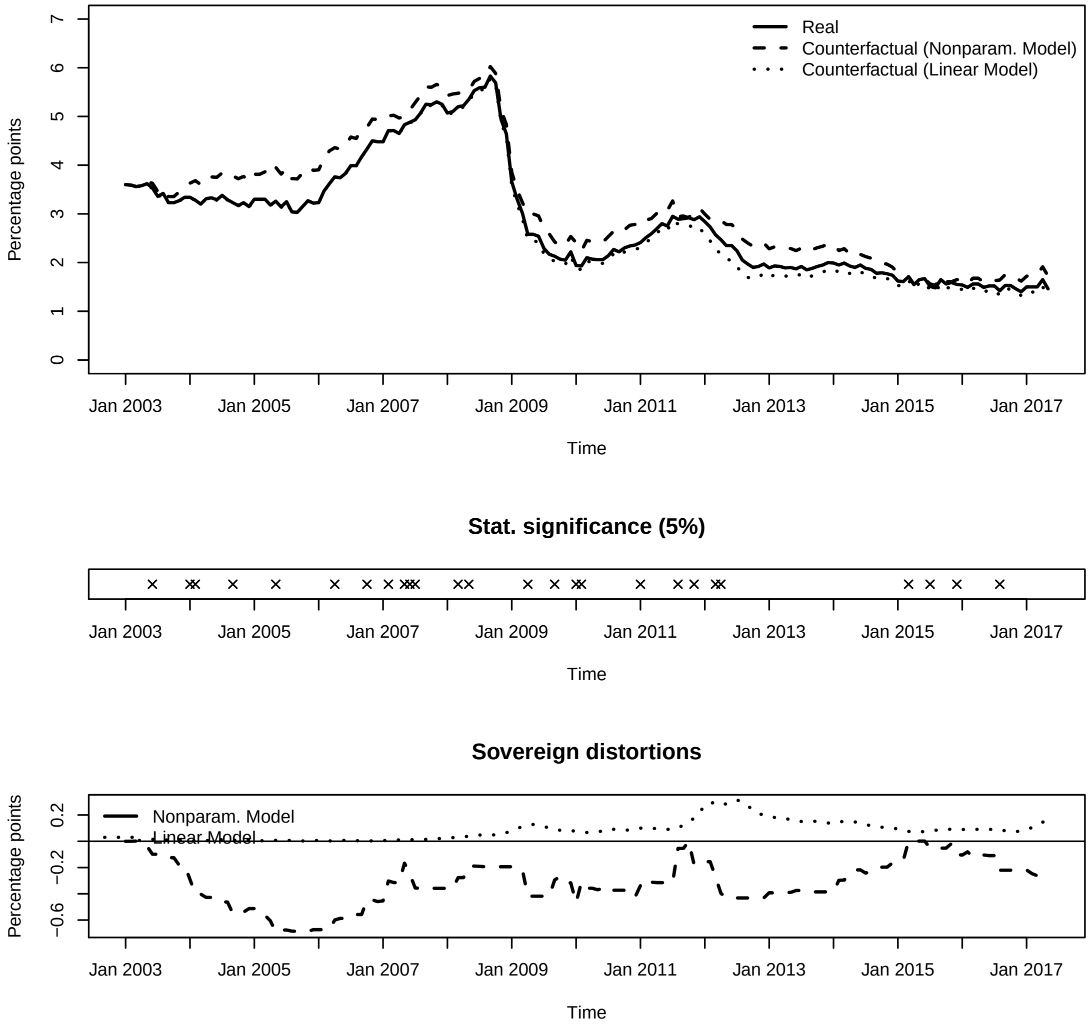

Figure C1: Sovereign risk transmission onto short-term corporate cost of borrowing in France.

Notes: Top figure plots the realized and the counterfactual costs of borrowing, estimated by nonparametric and linear frameworks. Middle figure shows periods when the difference between the observed and nonparametric counterfactuals was statistically significant at 5% level according to the standard confidence intervals. Bottom figure shows the sovereign risk contributions derived as a difference between the observed and counterfactual rates.

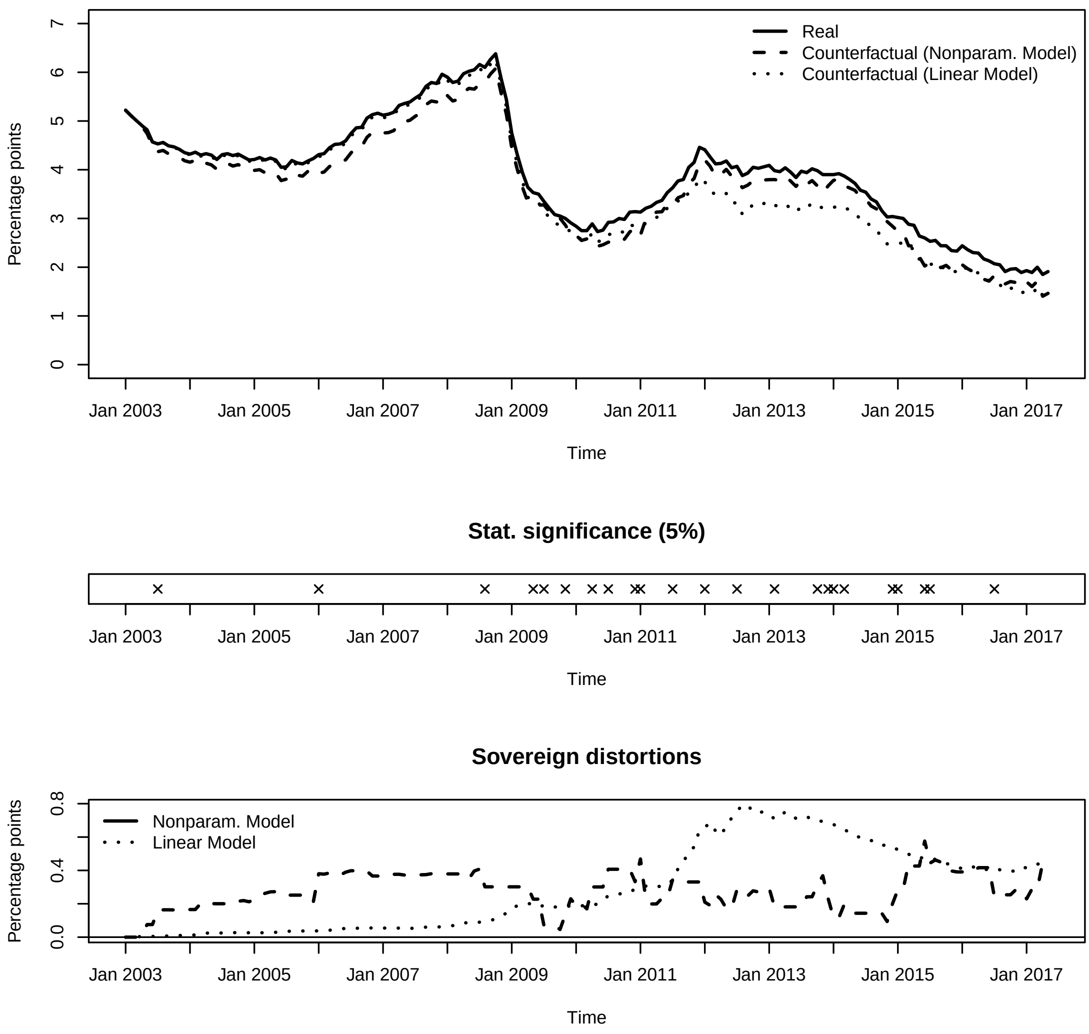

Figure C2: Sovereign risk transmission onto short-term corporate cost of borrowing in Italy.

Notes: Top figure plots the realized and the counterfactual costs of borrowing, estimated by nonparametric and linear frameworks. Middle figure shows periods when the difference between the observed and nonparametric counterfactuals was statistically significant at 5% level according to the standard confidence intervals. Bottom figure shows the sovereign risk contributions derived as a difference between the observed and counterfactual rates.

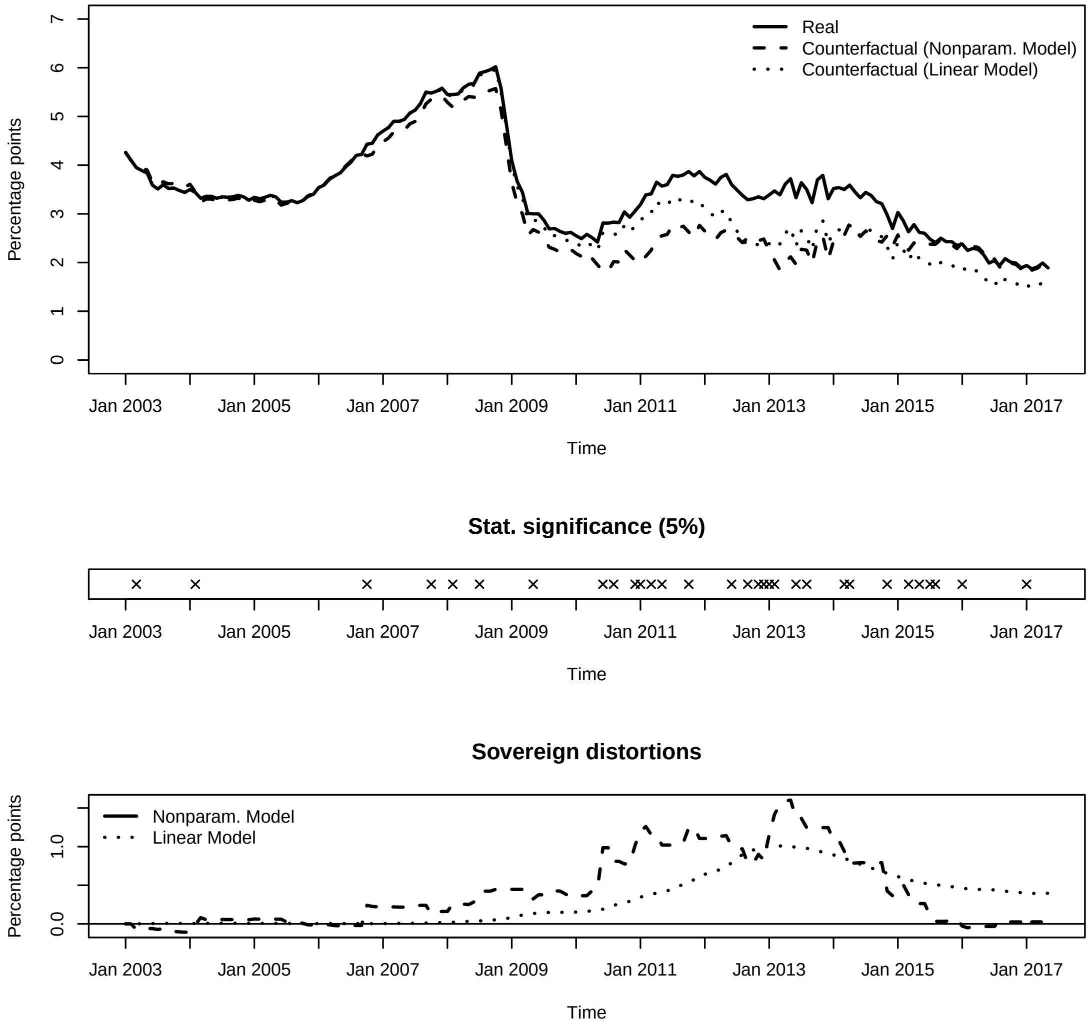

Figure C3: Sovereign risk transmission onto short-term corporate cost of borrowing in Spain.

Notes: Top figure plots the realized and the counterfactual costs of borrowing, estimated by nonparametric and linear frameworks. Middle figure shows periods when the difference between the observed and nonparametric counterfactuals was statistically significant at 5% level according to the standard confidence intervals. Bottom figure shows the sovereign risk contributions derived as a difference between the observed and counterfactual rates.