Empirical Economics
https://doi.org/10.1007/s00181-021-02029-z

Predicting recessions with a frontier measure of output gap: an application to Italian economy

Camilla Mastromarco[1] · **Léopold Simar**[2] · **Valentin Zelenyuk**[3]

Received: 18 January 2020 / Accepted: 11 January 2021

Abstract
Despite the long and great history, developed institutions, and high level of physical and human capital, the Italian economy has been fairly stagnant during the last three decades. In this paper, we merge two streams of literature: nonparametric methods to estimate frontier efficiency of an economy, which allows us to develop a new measure of output gap, and nonparametric methods to estimate probability of an economic recession. To illustrate the new framework, we use quarterly data for Italy from 1995 to 2019 and find that our model, using either nonparametric or the linear probit model, is able to provide useful insights.

Keywords Forecasting · Output gap · Robust nonparametric frontier · Generalized nonparametric quasi-likelihood method · Italian recession

JEL Classification C5 · C14 · C13 · C32 · D24 · E37 · O4

Valentin Zelenyuk: Research support from his employer and from ARC Grants (DP130101022 and FT170100401) is gratefully acknowledged.

✉ Camilla Mastromarco
camilla.mastromarco@unical.it

Léopold Simar
leopold.simar@uclouvain.be

[1] Dipartimento di Economia, Statistica e Finanza - Giovanni Anania - DESF, Universitá della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy

[2] Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Voie du Roman Pays 20, 1348 Louvain-La-Neuve, Belgium

[3] School of Economics and Centre for Efficiency and Productivity Analysis (CEPA), The University of Queensland, Colin Clark Building (39), St Lucia, Brisbane, QLD 4072, Australia

Published online: 04 March 2021

1 Introduction

How to predict economic recessions of a country? This is a very important and challenging question which is of interest to a fairly wide audience. Many papers in the empirical macroeconomic literature have proposed various methods to predict economic recessions, mainly focusing on the USA. Here we follow one of the paradigms, started by Estrella and Mishkin (1995, 1998) and further elaborated in various papers (e.g., see Duecker 1997; Kauppi and Saikkonen 2008, and references cited therein), and we try to elaborate further by adapting some newly advanced methods in nonparametric statistics and in productivity and efficiency analysis.

In this paper, we focus on Italian economy, one of the oldest in the World, with roots going back to at least the Roman Empire. Notwithstanding the long and great history, developed institutions, and high level of physical and human capital, Italian Economy has been stagnant during the last decades. Semiparametric and nonparametric methods are increasingly popular to analyze data in economics, business, and other fields (e.g., see Horowitz 2009; Henderson and Parmeter 2015). Specifically, we use a nonparametric version of the dynamic probit for time series (Park et al. 2017) to model the dependent variable (recession vs. non-recession). Meanwhile, for the explanatory variables, besides the standard predictor such as the spread, we try to develop a method to incorporate the estimates of the efficiency scores of a country. For this purpose, we use the method of frontier estimation in nonparametric location-scale models (Florens et al. 2014) and robust conditional frontier methods (Cazals et al. 2002; Daraio and Simar 2005; Daouia and Gijbels 2011; Mastromarco and Simar 2018, etc.). We illustrate our approach on the case of the Italian economy.

Our paper is also related to and in the spirit of the work of Wheelock and Wilson (1995), who pioneered the use of efficiency estimates among predictors in the parametric probability models, in their case for predicting bank failures. Besides the focus on macroeconomic recessions rather than banks, the major distinctive features of our paper relative to theirs include (i) the use of recent nonparametric estimation methods for the discrete choice model (rather than a parametric one), (ii) the use of time-series data, with a dynamic component modeled explicitly, and (iii) the use of more advanced methods for efficiency estimation that have become available very recently.

1.1 Predicting recessions

Among the variety of different approaches attempting to model and forecast economic recessions, we will focus on those that employed the parametric binary choice approach and find that a good model for the prediction of the US recessions is a parsimonious model with only one of a few predictors, the most important of which is the interest rate spread and one discrete variable, the lagged dependent variable. The roots of this approach go back to at least the seminal work of Estrella and Mishkin (1995, 1998), who thoroughly investigated various parametric models with many variables and concluded that the best forecasts resulted from a parsimonious probit model involving only one explanatory variable, the lagged spread. Duecker (1997) confirmed this result, yet also found that including the lagged dependent variable among regres-

sors substantially improved the predicting power of the Estrella and Mishkin (1995, 1998) approach, especially for the recessions of the 1970s and 1990s that were missed by various other forecasting methods. Overall, the analyses in Estrella and Mishkin (1995, 1998) and Duecker (1997) suggest that their parsimonious model outperforms many alternative models that included many variables to gain a high in-sample fit, yet happened to be poorly forecasting the future. Also see Kauppi and Saikkonen (2008) for further refinements and more references and discussions.

This paper contributes to the empirical literature on predicting recessions by adding two novelties: (i) we apply a nonparametric dynamic time series discrete response model suggested by Park et al. (2017) and (ii) we use a new measure of output gap as one of the recession predictors. In particular, we employ a robust nonparametric frontier panel data model proposed by Mastromarco and Simar (2015) to estimate the time-dependent conditional efficiency of countries and use this as a measure of output gap.[1] In a macroeconomics context, where countries are producers of output (i.e., GDP), given inputs (e.g., capital, labor), and technology, inefficiency can be identified as the distance of the individual production from the frontier. This frontier can be estimated by the maximum output of the reference country regarded as the empirical counterpart of an optimal boundary of the production set. Hence, at least on intuitive grounds, we might interpret the inefficiency as a measure of output gap with respect to the potential output of the technological frontier.

1.2 Existing measures of output gap

Output gap is traditionally obtained as a deviation from a statistical measure of trend. One of the earliest and currently widely used statistical methods for measuring the output gap is based on measuring the output trend calculated by fitting a polynomial in time to the output, the residual being the estimated cycle. This method imposes a strong prior on the smoothness of the trend. Another popular statistical approach uses a filter, Hodrick and Prescott (1997), to identify the trend and the cycle. The trend measure in this case is smooth but not deterministic. The Baxter and King (1999) filter defines the cycle as having spectral power in pre-specified frequencies. However, Murray (2003) stresses that this filter extracts an estimate of the cycle which includes some trend shock. Other statistical approaches need a model to identify the stochastic trend component. These statistical methods do not require smoothness but impose the restriction of no correlation between the cycle and the trend, which may lack theoretical support. Beveridge and Nelson (1981) suggest a measure of trend as a long run forecast of an ARMA model. The unobserved components model extracts an estimate of the trend and cycle using the Kalman filter (Harvey 1985; Watson 1986; Clark 1987).

Differently from the statistical methods, the economic approaches estimate the output gap in the framework of the production function (for example Galí and Gertler 1999). Recently, various studies (Kuttner 1994; Gerlach and Smets 1999; Apel and Jansson 1999; Roberts 2001; Basistha and Nelson 2007; Basu and Fernald 2009) tried

[1] Also see Cazals et al. (2002), Daraio and Simar (2005), Daouia and Gijbels (2011) for related discussions on robust nonparametric frontier.

to combine the statistical approach with the economic approach by estimating the unobserved components of the multivariate model. These approaches do not impose smoothness or restrictive correlation structure, but estimate the output gap based on the empirical implications of the forward-looking Phillips curve.

1.3 Inefficiency as an alternative measure of output gap

Often, potential output is referred to as the production capacity of the economy. In our framework of the frontier model, potential output refers to the maximum level of output that can be produced for a given level of inputs, using full employment and capital utilization. The gap between the potential and actual outputs is interpreted as a measure of inefficiency which in our paper also captures the varying factor utilization over the cycle. The approach is closely linked to the production theory based approach in measuring the output gap. We cast our empirical model in frontier form, treating the gap as an unobserved variable—efficiency scores—estimated using nonparametric frontier methods. In pursuing an economic based approach, we avoid imposing strong priors on the smoothness of the trend or cycle, and the restrictive correlation structure between the trend and the cycle shocks.

Furthermore, parametric modeling may suffer from misspecification problems when the data generating process is unknown, as is usual in the applied studies. We propose a unified nonparametric framework for accommodating simultaneously the problem of model specification uncertainty and time dependence in the panel data frontier model. Specifically, we estimate the panel data frontier model using a flexible nonparametric two step approach to take into account the time dependence. Following recent developments in nonparametric conditional frontier literature (Florens et al. 2014; Mastromarco and Simar 2015, 2018), we adapt the nonparametric location-scale frontier model, where we link production inputs and output to time. In the first step we clean the dependence of inputs and outputs on time factors. These time factors capture the correlation among units. By eliminating the effect of these factors on the production process, we mitigate the problem of dependence across our time units and we are able to estimate a nonparametric frontier model from the panel data. (In the application we illustrate this approach for the data on 16 OECD countries.) In the second step, we estimate the frontier and the efficiency scores using inputs and outputs whitened from the influence of time.

1.4 The contribution in a nutshell and a roadmap

The main idea of this paper is to merge the interesting streams of literature described above: the novel nonparametric methods to estimate frontier efficiency of an economy as a new measure of output gap and the novel nonparametric method to estimate the probability of an economic recession. We do this by deploying a generalized nonparametric quasi-likelihood method in the context of dynamic discrete choice models for time series data (Park et al. 2017). To illustrate the new framework, we use data from 1995 to 2019 with quarterly frequency and find that our model using either

nonparametric or the linear probit model, applied frequently in this context, is able to offer additional insights into the literature.

The paper is organized as followed. Section 2 presents the methodology. Specifically, Sect. 2.1. explains the nonparametric discrete choice models for time series to predict recessions. Section 2.2. introduces our proposed measure of output gap and explains time-dependent conditional efficiency scores and the nonparametric estimation. This section elucidates the location-scale models to eliminate the influence of common time factors and external variables. Section 3 illustrates an empirical application for the case of Italian economy. Section 4 gives concluding remarks.

2 Methodology

2.1 Forecasting model

In this section, we summarize the elements from Park et al. (2017) (hereafter PSZ) that are needed in our setup to forecast economic recessions. The model should provide the elements for analyzing the behavior of a discrete variable in a time series setup. The approach is nonparametric.

Suppose we observe $(\boldsymbol{X}^t, \boldsymbol{Z}^t, Y^t)$, $t = 1, \ldots, T$, where $\left\{(\boldsymbol{X}^t, \boldsymbol{Z}^t, Y^t)\right\}_{t=-\infty}^{\infty}$ is a stationary random process. We assume as in PSZ that the process satisfies strong mixing conditions that typically allows time dependence which disappears at a geometrical rate when the time lags are too large.[2]

The response variable is binary taking the values 0 and 1; in our set-up, $Y = 1$ for a recession and $Y = 0$ is otherwise. The vector of covariates $\boldsymbol{X}^t$ is of dimension r and of continuous type, whereas $\boldsymbol{Z}^t$ is a discrete vector of dimension k. The components of $\boldsymbol{Z}^t$ may be lagged values of the response Y, e.g., Y^{t-1}, Y^{t-2}. The idea is to estimate the mean function

$$m(\boldsymbol{x}, z) = \mathbb{E}\left(Y | \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = z\right). \tag{2.1}$$

Since Y is binary, we have

$$\mathbb{P}\left(Y = y | \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = z\right) = m(\boldsymbol{x}, z)^y \left[1 - m(\boldsymbol{x}, z)\right]^{1-y}, \text{ for } y \in \{0, 1\}. \tag{2.2}$$

A key ingredient in these discrete choice models is the link function g, which is a strictly increasing function, defining the function f as

$$f(\boldsymbol{x}, z) = g(m(\boldsymbol{x}, z)). \tag{2.3}$$

In parametric models, it is assumed that $f(\boldsymbol{x}, z)$ takes a parametric form, and then, $m(\boldsymbol{x}, z) = g^{-1}(f(\boldsymbol{x}, z))$. Thus, a wrong choice may jeopardize the estimation of m. In nonparametric settings, $f(\boldsymbol{x}, z)$ will be locally approximated by some local polynomial around $(\boldsymbol{x}, z)$, so the choice of g is much less important. Approximating locally the

[2] See PSZ, Sect. 3.1, for mathematical details and additional references.

functions $g_1(m(\boldsymbol{x}, \boldsymbol{z}))$ or $g_2(m(\boldsymbol{x}, \boldsymbol{z}))$ for two different link functions g_1 and g_2 does not make much difference. One may simply take the identity function, though since the range of the target m is [0, 1], we will choose a link that guarantees the correct range (like Probit or Logit). Now, given the link g and the sample $\left\{(\boldsymbol{X}^t, \boldsymbol{Z}^t, Y^t)\right\}_{t=1}^{T}$, we see from (2.2) that the log-likelihood of f is given by $\sum_{t=1}^{T} \ell\left(g^{-1}\left(f(\boldsymbol{X}^t, \boldsymbol{Z}^t)\right), Y^t\right)$ where $\ell(\mu, y) = y \log\left(\frac{\mu}{1-\mu}\right) + \log(1-\mu)$.

Let $(\boldsymbol{x}, \boldsymbol{z})$ be a fixed point of interest at which we want to estimate the value of the mean function m, or equivalently of its transformed function f. In a nonparametric approach, we will apply local smoothing techniques to the observations $(\boldsymbol{X}^t, \boldsymbol{Z}^t)$, which are in the neighborhood of $(\boldsymbol{x}, \boldsymbol{z})$. As explained in PSZ, this leads to weighting the observation $(\boldsymbol{X}^t, \boldsymbol{Z}^t)$ near $(\boldsymbol{x}, \boldsymbol{z})$ by some kernel. For the continuous variables (X), usual continuous kernels (Gaussian, Epanechnikov, etc.) can be used, while for the discrete variables (Z), some appropriate discrete kernels have to be used. Here we use the product kernel $w_c^t(\boldsymbol{x}, \boldsymbol{z}) \times w_d^t(\boldsymbol{z})$ defined as

$$w_c^t(\boldsymbol{x}, \boldsymbol{z}) = \prod_{j=1}^{r} K_{h_j}(x_j, X_j^t, \boldsymbol{z}). \tag{2.4}$$

$$w_d^t(\boldsymbol{z}) = \prod_{l=1}^{k} \gamma_l^{\mathbb{1}(Z_l^t \neq z_l)} \tag{2.5}$$

where $\mathbb{1}(A)$ denotes the indicator function such that $\mathbb{1}(A) = 1$ if A holds and zero otherwise and $\gamma_l \in [0, 1]$ is the bandwidth for the jth discrete variable, while for the continuous kernels, we have

$$K_{h_j}(x_j, X_j^t, \boldsymbol{z}) = \frac{1}{h_j(1)} K\left(\frac{X_j^t - x_j}{h_j(1)}\right) \times \mathbb{1}(Z^t = z(1)) + \frac{1}{h_j(2)} K\left(\frac{X_j^t - x_j}{h_j(2)}\right) \times \mathbb{1}(Z^t = z(2))$$

for a symmetric kernel function K and two bandwidth, $h_j(1) > 0$ and $h_j(2) > 0$, corresponding to the two groups denoted as $z(1)$ and $z(2)$, for each jth continuous variable. The discrete kernel is in the spirit of Aitchison and Aitken (1976), except that it is standardized to be between 0 and 1. The continuous kernel is a generalized kernel proposed by Li et.al. (2016), which allows different bandwidths for the continuous variables across various groups defined by the values of $\boldsymbol{Z}$, thus allowing for more flexibility in terms of the fitted curvatures in the two groups. It is worth noting that when $\gamma_l = 0$, one performs a separate estimation for each group identified by the values of Z_l. When $\gamma_l = 1$, one considers that Z_l is irrelevant and so all the groups are pooled together, although different bandwidths for continuous variables may still imply different curvatures in the two groups.

For approximating $f(\cdot, \cdot)$ locally near the point $(\boldsymbol{x}, \boldsymbol{z})$, we will not make use of the link function, nor of the likelihood function. The local approximation is linear in

the direction of the continuous variable and constant in the direction of the discrete variables. To be specific, we have

$$f(\boldsymbol{u}, \boldsymbol{v}) \approx f(\boldsymbol{x}, \boldsymbol{z}) + \sum_{j=1}^{r} f_j(\boldsymbol{x}, \boldsymbol{z})(u_j - x_j), \tag{2.6}$$

where $f_j(\boldsymbol{x}, \boldsymbol{z}) = \partial f(\boldsymbol{x}, \boldsymbol{z})/\partial x_j$. So the local approximation can be viewed as a first order Taylor's expansion of f in $\boldsymbol{x}$, near $(\boldsymbol{x}, \boldsymbol{z})$.

To estimate $f(\boldsymbol{x}, \boldsymbol{z})$ and its partial derivatives $f_j(\boldsymbol{x}, \boldsymbol{z})$, we thus maximize

$$T^{-1} \sum_{t=1}^{T} w_c^t(\boldsymbol{x}, \boldsymbol{z}) w_d^t(\boldsymbol{z}) \ell \left(g^{-1} \left(\beta_0 + \sum_{j=1}^{r} \beta_j (X_j^t - x_j) \right), Y^t \right) \tag{2.7}$$

with respect to β_0 and β_j, $j = 1, \ldots, r$. The solutions $\hat{\beta}_0 = \widehat{f}(\boldsymbol{x}, \boldsymbol{z})$ and $\hat{\beta}_j = \widehat{f_j}(\boldsymbol{x}, \boldsymbol{z})$ for $j = 1, \ldots, r$. Then, an estimator of the mean function $m(\boldsymbol{x}, \boldsymbol{z})$ is obtained by inverting the link function: $\widehat{m}(\boldsymbol{x}, \boldsymbol{z}) = g^{-1}(\hat{\beta}_0)$.

The theory in PSZ shows that the asymptotic properties of the estimators do not much depend on the choice of the link function, as long it is smooth enough and strictly increasing, because the estimation is performed locally. We will choose below the probit link, i.e., $g(s) = \Phi^{-1}(s)$, where Φ is the cumulative distribution function of the standard normal distribution. So we have to maximize in (β_0, β_j), $j = 1, \ldots, r$

$$T^{-1} \sum_{t=1}^{T} w_c^t(\boldsymbol{x}, \boldsymbol{z}) w_d^t(\boldsymbol{z}) \left[Y^t \log \left(\frac{\Phi\left(\beta_0 + \sum_{j=1}^{r} \beta_j (X_j^t - x_j)\right)}{1 - \Phi\left(\beta_0 + \sum_{j=1}^{r} \beta_j (X_j^t - x_j)\right)} \right) + \log \left(1 - \Phi \left(\beta_0 + \sum_{j=1}^{r} \beta_j (X_j^t - x_j) \right) \right) \right]. \tag{2.8}$$

The properties of the resulting estimators follow from PSZ. In summary, under certain regularity assumptions and with the optimal order of the bandwidths, $h_{c,j} := (h_j(1) + h_j(2))/2 \propto T^{-1/(r+4)}$ and $\gamma_l \propto T^{-2/(k+4)}$, Theorem 3.1 in PSZ establishes

$$\sqrt{T \bar{h}_c} \left(\widehat{f}(\boldsymbol{x}, \boldsymbol{z}) - f(\boldsymbol{x}, \boldsymbol{z}) + \sum_{j=1}^{r} O\left(h_{c,j}^2\right) + \sum_{l=1}^{k} O\left(\gamma_l\right) \right) \xrightarrow{\mathcal{L}} N(0, V(\boldsymbol{x}, \boldsymbol{z})), \tag{2.9}$$

where $\bar{h}_c = \prod_{j=1}^{r} h_{c,j}$ and the variance V has a complicated expression which depends on the properties of the data generation process (DGP) (see PSZ for details). We see from (2.9) that the optimal bandwidths balance, as often the case, is between the square of the bias terms and the variance.

Remark 1 It is worth noting that if the bandwidths for continuous variables increase such that they cover all the observations on those variables, the nonparametric approach yields very similar estimates as the parametric approach that assumes (2.6) holds exactly. In this sense, the parametric approach can be viewed as a special case of the nonparametric approach, in the sense that the latter allows for much more flexibility and can be 'reduced' to the former by removing the flexibility through tuning the bandwidths to be large enough.

Remark 2 The nonparametric approach can also be viewed as a tool for validation of a suitable parametric approach. Indeed, when a parametric approach that assumes a particular (and perhaps very restrictive) functional form yields very similar results or conclusions as the nonparametric approach that allows for much more flexibility, this should give more confidence in the results or conclusions from the parametric approach, despite its restrictive assumptions. We will find this consideration very useful in our empirical application section for the particular data we use there.

2.2 Efficiency and estimation of the output gaps

We propose as an output gap our measure of inefficiency. The output gap is an economic measure of the difference between the actual output of an economy and its potential output. Potential output is the maximum amount of goods and services an economy can turn out when it is most efficient—that is, at full capacity. Often, potential output is referred to as the production capacity of the economy. In the context of this paper, we assume that a country is the producer of an output (i.e., GDP), given inputs (e.g., capital, labor), and available technology. The inefficiency is defined as the distance between the actual production and its maximum or frontier potential, given the inputs and technology.[3]

As explained above, we would like to use the level of inefficiency of the country for a particular year by considering the so-called conditional inefficiency (Cazals et al. 2002; Daraio and Simar 2005; Mastromarco and Simar 2015). Inputs here are Capital (K) and Labor (L), and the output is the GDP (Q), and we have quarterly data $t = 1, \ldots, T$ for 16 OECD countries. Evaluating the marginal efficiency measures by considering the so-called meta-frontier of the 3-dimensional cloud of T points $\{(K_t, L_t, Q_t)\}_{t=1}^{T}$ would not make too much sense since the technology certainly varies over the years. We will rather consider the conditional efficiency measure where we condition on the time period. This enables us to take into account that production factors adjust to fluctuations of aggregate demand and supply with time delays due to market regulations and price stickiness.[4]

As suggested in Mastromarco and Simar (2015), to introduce the time dimension we consider indeed, with some abuse of notation, time as a conditioning variable W

[3] In principle, each country may have their own frontier characterized by their uniqueness, which in turn could imply each country is 100% efficient relative to their own frontier. This is a different paradigm of thinking. The type of inefficiency we are measuring is relative to a common frontier and we acknowledge that other paradigms (e.g., which include country effects or group-specific effects) are also interesting to explore and we leave that for the future research.

[4] We thank anonymous reviewer for comment.

and we define the attainable production set at time t as the support of the conditional probability

$$H_{K,L,Q|W}(\xi, \zeta, \eta|W = t) = \mathbb{P}(K \leq \xi, L \leq \zeta, Q \geq \eta \mid W = t), \tag{2.10}$$

which can be interpreted as the probability of observing, at time t, a production plan dominating a given point (ξ, ζ, η). So, the feasible technology Ψ^t can be defined as

$$\Psi^t = \{(\xi, \zeta, \eta) \in \mathbb{R}^3_+ | H_{K,L,Q|W}(\xi, \zeta, \eta \mid W = t) > 0\}. \tag{2.11}$$

Finally, this leads to consider for the output orientation the conditional efficiency score

$$\lambda(\xi, \zeta, \eta|t) = \sup\{\lambda|(\xi, \zeta, \lambda\eta) \in \Psi^t\} \geq 1, \tag{2.12}$$

which is known as the Farrell–Debreu output oriented efficiency measure (see, e.g., Kumar and Russell 2002, for its use in a related context but using a simpler estimator). Nonparametric estimators of these efficiency scores have been developed and their asymptotic properties are well known (see e.g. Jeong et al. 2010). Here, we will follow the approach suggested by Florens et al. (2014) which has some advantages described below.

In the first step, a flexible nonparametric model is used to whiten the inputs (K, L) and the output Q from the effect of time W. We have the following model

$$\begin{aligned} K_{it} &= \mu_K(t) + \sigma_K(t)\varepsilon_{K,t} \\ L_{it} &= \mu_L(t) + \sigma_L(t)\varepsilon_{L,t} \\ Q_{it} &= \mu_Q(t) + \sigma_Q(t)\varepsilon_{Q,t}, \end{aligned} \tag{2.13}$$

where we assume that $(\varepsilon_K, \varepsilon_L, \varepsilon_Q)$ are 'independent' of time W, with $\mathbb{E}[\varepsilon_\ell] = 0$ and $\mathbb{V}[\varepsilon_\ell] = 1$ for $\ell = K, L, Q$. The estimation of the mean and variance functions is done by local polynomial smoothing as explained in detail in Florens et al. (2014). They suggest also a bootstrap test for testing the assumption of independence, but in our application below we will evaluate various correlations (Spearman, Pearson, and Kendall) to check if this assumption is reasonable.

In our application, we first use the local-linear methods to estimate the mean functions $\mu_\ell(t)$, $\ell = K, L, Q$. From the squared residuals, we estimate the variance functions $\sigma^2_\ell(t)$ by local constant methods (to avoid negative variances). Finally, Florens et al. (2014) define the estimated 'pure' inputs and the estimated 'pure' outputs as

$$\begin{aligned} \widehat{\varepsilon}_{K,it} &= \frac{K_{it} - \widehat{\mu}_K(t)}{\widehat{\sigma}_K(t)}, \\ \widehat{\varepsilon}_{L,it} &= \frac{L_{it} - \widehat{\mu}_L(t)}{\widehat{\sigma}_L(t)}, \\ \widehat{\varepsilon}_{Q,it} &= \frac{Q_{it} - \widehat{\mu}_Q(t)}{\widehat{\sigma}_Q(t)}, \end{aligned} \tag{2.14}$$

which are 'pure' in the sense of being filtered from time dependence. In this 'pure units space,' we can compute the output directional distance to the efficient frontier.[5] Since the output here is univariate, the efficient frontier in pure units is the function

$$\varphi(e_K, e_L) = \sup\{e_Q | \mathbb{P}(\varepsilon_K \leq e_K, \varepsilon_L \leq e_L, \varepsilon_Q \geq e_Q) > 0\}, \tag{2.15}$$

so that the directional distance of a point (e_K, e_L, e_Q) to the frontier is simply given by

$$\delta(e_K, e_L, e_Q) = \varphi(e_K, e_L) - e_Q \geq 0, \tag{2.16}$$

where the value zero indicates the point (e_K, e_L, e_Q) is on the efficient frontier. Under the location-scale assumptions, it can be proved that the conditional frontier in original units can be recovered as (see Florens et al. 2014, for details)

$$\tau(\xi, \zeta | t) = \mu_Q(t) + \sigma_Q(t)\varphi\left(\frac{\xi - \mu_K(t)}{\sigma_K(t)}, \frac{\zeta - \mu_L(t)}{\sigma_L(t)}\right), \tag{2.17}$$

so that the gap in the output to reach the frontier level is given by

$$G_Q(\xi, \zeta, \eta | t) = \sigma_Q(t)\delta(e_K, e_L, e_Q). \tag{2.18}$$

The nonparametric estimators of these various elements are obtained by plugging the estimators of the mean and variance functions derived above. One of the main advantages of this location-scale approach is that for estimating the functions $(\mu_\ell(t), \sigma_\ell(t))$ we require only smoothing in the center of the data in a standard regression setup. As pointed out in Bădin et al. (2019), a direct estimation of $\lambda(\xi, \zeta, \eta | t)$ requires delicate problems of optimal bandwidths selection for estimating the support of the conditional $H_{K,L,Q|W}(\xi, \zeta, \eta \mid W = t)$.

So at the end of this step of efficiency estimations, we end up in practice with estimated efficiency scores in the pure units $\delta(e_{K_t}, e_{L_t}, e_{Q_t})$ and, if wanted, the measures of the gaps in original units of the DGP, i.e., $G_Q(K_t, L_t, Q_t | t)$ at each observation $t = 1, \ldots, T$. These values (eventually lagged) will be used to improve the prediction of a recession in our application below.

Real data samples contain in general some anomalous data, and the estimated frontier obtained by these nonparametric techniques can be fully determined by these outliers or extreme data points, jeopardizing the measurement of inefficiencies, potentially leading to unrealistic results. Cazals et al. (2002), Daouia and Simar (2007), in the frontier literature, propose an approach which aims to keep all the observations in the sample but replace the frontier of the empirical distribution by (conditional) quantiles or by the expectation of the minimum (or maximum) of a sub-sample of the data. This latter method defines the order-m frontier that we will use here.

In brief, the partial output frontier of order-m is defined for any integer m and for input values e_{K_t}, e_{L_t}, as the expected value of the maximum of the output of m units

[5] We need to use directional distance here since the 'pure' inputs and the 'pure' outputs may take negative values.

drawn at random from the populations of units such that $\varepsilon_K \leq e_K, \varepsilon_L \leq e_L$. Formally,

$$\tau_m(\xi, \zeta | t) = \mathbb{E}\left[\max\left(\varepsilon_{Q,1t}, \ldots, \varepsilon_{Q,mt}\right)\right], \tag{2.19}$$

where the $\varepsilon_{Q,it}$ are drawn from the empirical conditional survival function $\widehat{S}_{\varepsilon_Q|\varepsilon_x}(e_Q|\varepsilon_{x,it} \leq e_x)$. This can be computed by Monte Carlo approximation or by solving a univariate numerical integral (for practical details see Simar and Vanhems 2012).

If m increases and converges to ∞ and $n \to \infty$, it has been shown (see Cazals et al. 2002) that the order-m frontier and its estimator converge to the full frontier, but for a finite m, the frontier will not envelop all the data points and so is much more robust than the Free Disposal Hull (FDH) to outliers and extreme data points (see, e.g., Daouia and Gijbels 2011, for the analysis of these estimators from a theory of robustness perspective). Another advantage of these estimators is that besides the fact that their limiting distribution is normal, they achieve the parametric rate of convergence ($\sqrt{n}$).

3 Empirical illustration: the case of modern Italy

3.1 Data in brief

There are different ways to measure the spread to be used in the models that we consider here. For the US economy, it is often (albeit not always) measured as the difference between the 10-year US Treasury bond rate and the 3-month US Treasury bill rate, though there are other variants (e.g., see Park et al. (2020) and references there in). For other countries, including those in the EU, there appears to be no 'one-fit-all' rule on how to best measure the spread, as it may depend largely on the country of interest or even the time period considered. Here we choose to measure it as the difference between the 10-year Italy Treasury bond rate and the 10-year Germany Treasury bond rate, in per cent per annum. The logic behind using this measure of spread is grounded in the belief that the 10-year yield on German bonds is typically considered as the benchmark for the Euro area since they are viewed by investors as a risk-free market asset, at least in relative terms. [6] The data for this variable were sourced from OECD.stat Monthly Monetary and Financial Statistics (MEI).[7] However, again, we acknowledge that other measures of the spread can be tried, and some of them potentially may work better for some countries, yet not others, or differ across different periods for the same country. In fact, finding such a measure of spread that would serve as the best predictor for a given country may be a research question in itself and we leave it for future research endeavors.

The variables on recessions are constructed as following. We use the Composite Leading Indicators from OECD Reference Turning Points and Component Series data, which is analogous to the information from the Business Cycle Dating Committee of

[6] E.g., for related discussion for a broad audience, see *The Economist* https://www.economist.com/blogs/buttonwood/2014/03/investing.

[7] The observation period is selected by the data availability.

NBER typically used for timing the recessions in the USA.[8] In particular, note that the OECD identifies months of the so-called *turning points* (peaks and troughs) of the business cycle. The periods between a peak and a trough that follows it are then deemed as the recessionary periods ($Y_t = 1$), while the periods between a trough and a peak that follows it are deemed as the expansionary periods ($Y_t = 0$). To be more precise, since the turning points are announced for a particular month while we use quarterly data, to construct this time series we use the following rule: the recession begins on the quarter of the month of the peak and ends on the quarter of the month on the trough.[9]

To construct our measure of output gap, we need to go beyond the data on Italy and consider a few other countries that may be deemed as relevant peers for Italy, to estimate a relevant technological frontier. For this illustrative exercise, we choose the following OECD countries: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Israel, Italy, South Korea, Netherlands, New Zealand, Norway, Spain, Sweden and UK.[10]

The data for these countries were sourced from OECD.stat (OECD Quarterly National Accounts) and include 99 quarterly observations from $(1995 : Q1)$ till $(2019 : Q2)$, on capital, labor and output.[11] To be precise, the output Q is proxied by gross domestic product (GDP), and is measured in millions of US dollars, at 2015 constant price level. For the labor input L, we use the number of employed persons (in thousands) seasonally adjusted. Meanwhile, the capital K is also measured in millions of US dollars at 2015 constant price and constructed applying the perpetual inventory method (PIM) by using the real investment series (gross fixed capital formation).[12] As often suggested with macroeconomic data, all these variables are transformed in logarithms before the frontier estimation.

[8] This information can be found at: www.oecd.org/sdd/leading-indicators/oecdcompositeleadingindicatorsreferenceturningpointsandcomponentseries.htm.

[9] It is worth noting that (as with many other time series macroeconomic data) the historical information about the turning points is sometimes updated, which potentially can change the months or even quarters for when (according to the OECD) the recession started with respect to different editions (or 'vintages') of the data. In this version of the paper, we use the latest edition available (February 2020), which is slightly different from its previous edition used in the earlier version of this paper. Also note that the last turning point reported by OECD in this edition was in November 2017, which was a peak, and for simplicity we treat all the remaining quarters in the data (till 2019:Q2) as recessionary periods. This simplification was also confirmed by forecasting the value of Y_t for periods beyond 2019:Q2.

[10] We acknowledge that there could be many reasons for why more or less countries can be selected here and there appears to be no theoretical or empirical rule to decide on this, besides the data availability which was an obvious consideration here. Trying other sets of countries may be a fruitful avenue for future research on this topic.

[11] See "Appendix A1" for data description.

[12] PIM is necessitated by the lack of capital stock data across all the countries. The capital stock is constructed as $K_t = K_{t-1}(1-\theta) + I_t$, where I_t is investment and θ the rate of depreciation assumed to be 6% (*e.g.*, Hall and Jones 1999; Iyer et al. 2008). Repair and maintenance are assumed to keep the physical production capabilities of an asset constant during its lifetime. Initial capital stocks are constructed, assuming that capital and output grow at the same rate. Specifically, for countries with investment data beginning in 1995, we set the initial stock, $K_{1995} = I_{1995}/(g+\theta)$, where g is output growth rate from 1995 to 2019. Estimated capital stock includes both residential and non-residential capital.

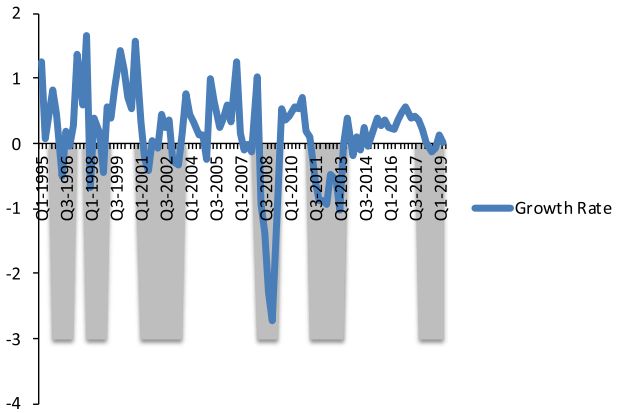

Fig. 1 Growth rate of real GDP in Italy during 1995–2019

3.2 Brief economic background on Italy

The Italian economy is one of the oldest in the World, with roots going thousands years back to at least the Roman Empire. Through its long evolution on its way to the modern days, it has witnessed a myriad of 'ups and downs' of its economy—what is now usually referred to as Business Cycles. In a broad sense, even a book cannot give a full picture of this interesting country and its economy, yet a brief snapshot on the recent years might be useful here.[13]

Despite the long and great history, fairly developed institutions, and relatively high level of physical and human capital, the Italian economy has been fairly stagnant during the last three decades, the period we focus on in this study. For example, in Fig. 1 we depict the growth rate of Italian GDP during 1995–2019.[14] Note that for the late 1900s, the figure exhibits negative growth in Q2 (2nd quarter) of 1996 and Q1 (1st quarter) of 1998. In the Q1 (1st quarter) of 2009, as the figure reveals, GDP growth registers the largest negative value, and by the Q3 (3rd quarter) of 2009 the economy began to re-grow slightly. In the Q3 (3rd quarter) of the year 2011, Italy's growth was negative till Q1 (1st quarter) of 2013; then, Italy's economy recovered with positive economic growth rates, but in Q1 of 2019 it starts to contract again. Here it is worth noting that, similarly as with the NBER data on recessions in the US, the OECD data on recessions in Italy (highlighted with gray shadow in Fig. 1) are not the same as the casual definition of a recession being two consecutive quarters of negative growth, but are based on the identification of the turning points of the business cycle, as described above.

Various reasons have been advocated in the literature as explanations for such poor economic performance of Italy. One of them is the lagging productivity growth relative to its peer countries. In particular, it was argued that insufficient productivity

[13] E.g., see Locke (1995), Malanima (2005) and Pellegrino and Zingales (2017) for more detailed exposition and discussions.

[14] Here we use GPSA measure (growth rate compared to previous quarter, seasonally adjusted), from OECD.stat.

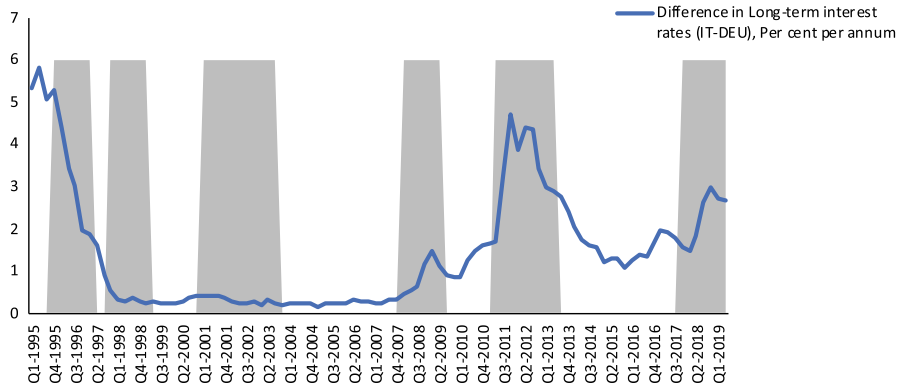

Fig. 2 Spread variable during 1995–2019, calculated as the difference between the 10-year Italy Treasury bond rate and 10-year Germany Treasury bond rate in per cent

growth may be pivotal to Italy's competitiveness problem, witnessed by the continual erosion of world export market shares and the limited ability to attract foreign direct investment (Faini et al. 2004). These problems appear to be particularly relevant in Italian manufacturing industries where productivity has been low and international competitiveness has worsened over the recent decades (Bassanetti et al. 2004; Aiello et al. 2011; Pellegrino and Zingales 2017). For example, Pellegrino and Zingales (2017) credit the inability of Italian firms to take full advantage of the information and communication technology revolution as one of the key reasons for the poor productivity or what they dubbed as 'Italy's productivity disease.' In turn, and as for many other failures or successes of a country, the existence and persistence of this 'disease' appear to be due to specific institutional aspects; or, as Pellegrino and Zingales (2017) put it:

> "While many institutional features can account for this failure, a prominent one is the lack of meritocracy in the selection and rewarding of managers. ...the prevalence of loyalty-based management in Italy is not simply the result of a failure to adjust, but an optimal response to the Italian institutional environment. Italy's case suggests that familism and cronyism can be serious impediments to economic development even for a highly industrialized nation."

Clearly, disentangling the true reasons for the recessions in Italy is well beyond the scope of this paper, if at all possible. What seems more feasible, however, is to compare or benchmark Italy to some of its peers—as we do via the proposed output gap measure explained above—in the hope that it may potentially help in providing some useful information for predicting upcoming recessions via the dynamic choice models.

Turning the attention to the spread dynamics, one can also note that Fig. 2 reveals the spread between the 10-year Italy Treasury bond rate and 10-year Germany Treasury bond rate increases during the period of low economic growth. This indicates a lack of confidence of investors in the Italian economy due to the deterioration of potential determinants of the spread, namely the current or expected macroeconomic fundamentals, such as fiscal policy, international risk, liquidity conditions, sovereign

Table 1 Correlation between $W = time$ and pure inputs ε_{X1}, ε_{X2} and pure output ε_Y

	ε_{X1}	ε_{X2}	ε_Y
Pearson correlations			
W	0.000184	0.000102	0.000192
Spearman rank correlations			
W	− 0.003584	− 0.018066	− 0.009535
Kendall correlations			
W	0.002918	− 0.011346	− 0.004158

credit ratings, to mention a few. Again, note that while in some periods the dynamics of the spread to some extent matches the upcoming changes in the recession indicators (highlighted with gray shadow), the relationship appears to be not very strong, e.g., relative to what we found in the literature for the recessions in the USA (see Park et al. 2020 and references therein).

3.3 Filtering the inputs/output and efficiency estimates

Here, we first have to run three location-scale models for K, L, Q, respectively, to clean the effect of time W.[15] This provides the 'pure' inputs and 'pure' output, $\{(\widehat{\varepsilon}_{K_t}, \widehat{\varepsilon}_{L_t}, \widehat{\varepsilon}_{Q_t})\}_{t=1}^{T}$ as explained above. The correlations of these 'pure' inputs/output with time are given in Table 1 (where $X_1 = K, X_2 = L, Y = Q$ and $Z = W$). Clearly these correlations are very small so we can infer that the assumption of independence between $(\varepsilon_K, \varepsilon_L, \varepsilon_Q)$ and W, which is part of our location-scale model, seems reasonable.

Robust measures of efficiency scores, providing the gaps in 'pure' units were computed with $m = 1500$. This choice was done for letting less than 25% of points above the order-m frontier, as shown in Fig. 3. Note that from the values of $m = 1500$ and $m \to \infty$ (the full FDH frontier), all the results are quite similar.

The resulting efficiency scores $\widehat{\delta}_{m,t}$ are shown in Fig. 4, which illustrates that most of the time, the time effect has indeed been cleaned from the production process. We see also that most of the values of $\widehat{\delta}_{m,t}$ are positive and some take very small (near zero) negative values. Figure 5 exhibits the time path of output gaps in original units (in logs and re-scaled by their mean). Figure 6 reports the values of the gap in original units for each country in our sample at the first period of observation (1995: Q1) and last period (2019: Q2).

We give in "Appendix C" the full table of results for all the time periods. The table also indicates the gaps G_t in original units of the DGP, as defined above (in log scale and re-scaled by their mean). Figure 9 in "Appendix C" reports the values of output gap in original units for all countries in our analysis for the first and last year of the observation period.

[15] For numerical convenience, all variables are scaled by their means, including the conditioning variable time, denoted here by W.

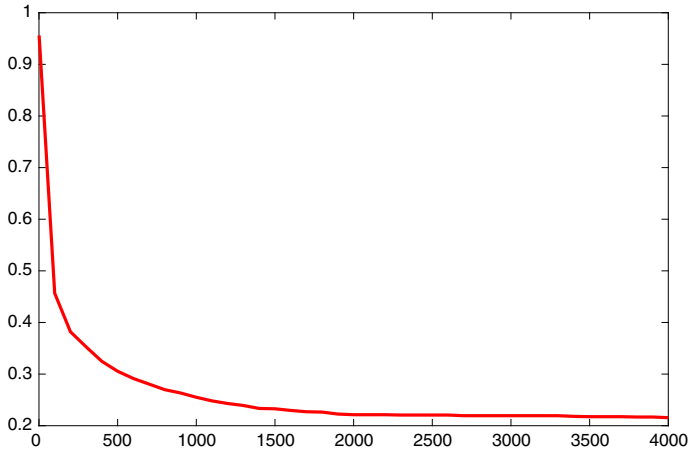

Fig. 3 Percentage of points left above the order-m frontier, as a function of m. We selected $m = 1500$ letting less than 25% of data points above the frontier

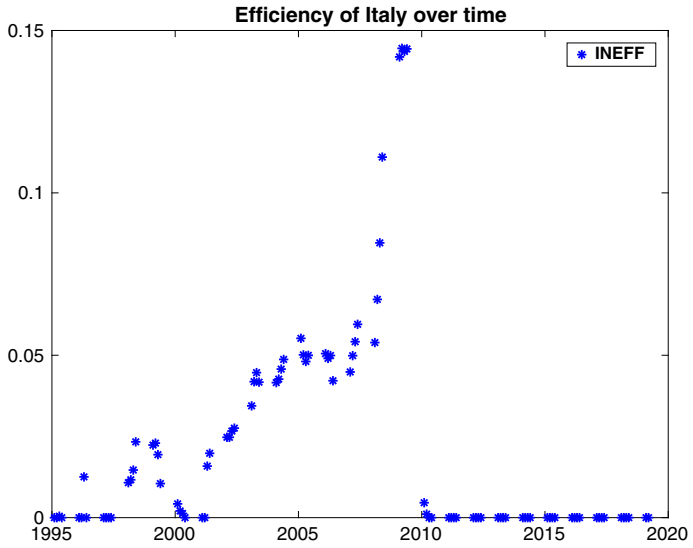

Fig. 4 Evolution of estimated inefficiency scores of Italy ($\widehat{\delta}_{m,t}$) over time

3.4 In-sample fit of the model

Our next step is to fit the prediction model described above by estimating the parametric linear probit model and the nonparametric model of PSZ to the data described above. In particular, we fit the following model:

$$\mathbb{E}\,(Y|\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}) = m(\boldsymbol{x}, \boldsymbol{z}) = m(X_1, X_2, Z), \tag{3.1}$$

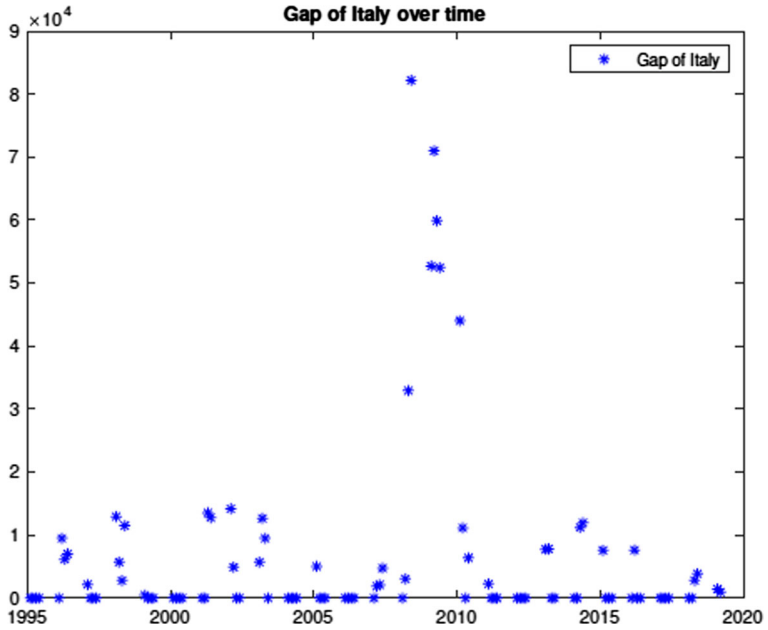

Fig. 5 Evolution of GAP of Italy over time

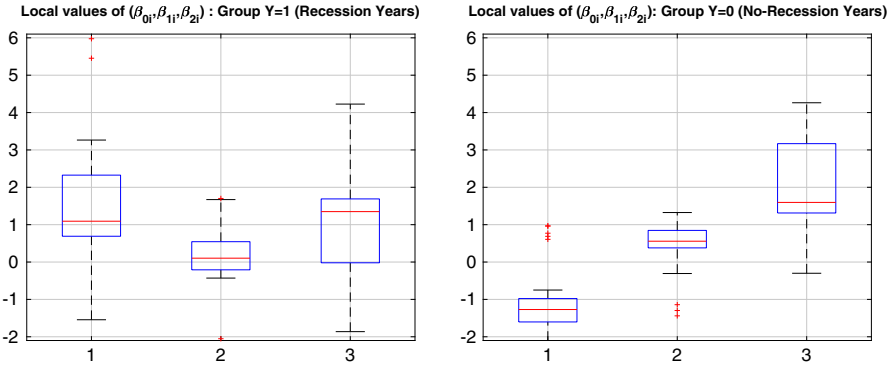

Fig. 6 Boxplots of the estimated local β's, with the full sample of $n = 98$ data points

where $X_{1,t} = Sp_{t-r_1}$ is the spread lagged by r_1 periods, $X_{2,t} = \Delta_{G,t-r_2}$ is the first difference of the estimates of output gaps (production efficiency) lagged by r_2 periods. Finally, $Z_t = Y_{t-r_3}$, where we recall that Y_t is the dichotomous dependent variable, defined as $Y_t = 1$ if "Italian economy is in recession" in the quarter t and 0 otherwise and r_3 is its chosen lag. Finally, for smoothing Z_t in the nonparametric approach we use the complete smoothing technique suggested by Li et.al. (2016), allowing different bandwidths for the continuous variables in the two groups determined by the values of Z, as described in Sect. 2.1.

Even though there are only three potential predictors in the general specification (3.1), many variations in it are possible that are based on different subsets of predictors and different choices of lags for each predictor. In the following sub-sections, using data on Italy we briefly show and discuss how a model selection can be done in such situations.[16]

3.4.1 Selection of lags

As is typical in empirical time-series studies, there is no theory on what lags should be chosen—it is largely an empirical issue. Here we will focus our discussion on the case when $r_1 = r_2 = 2$ and $r_3 = 1$. Thus, intuitively, our model assumes that the first difference of our measure of output gap affects the probability of an economy to be in recession with some delays, e.g., due to market imperfections and frictions. In particular, in this model it is expected to act as an indicator of recession two periods in advance, similarly as the other indicator, the spread, often used to forecast the recession, which in our case is expected to indicate two periods before a recession.

We also considered other combinations of lags and none of them have dominated the one we focus on here in the main text of the paper (see "Appendix A" for the related results). In particular, as suggested in the literature for measuring the quality of the model fit, we used the values of the achieved Maximum Likelihood and the Estrella Pseudo-R^2 to compare the models, although alternative measures of goodness-of-fit can also be used.[17]

3.4.2 Selection of predictors

For each combination of lags, we tried several specifications to check the sensitivity of results with respect to dropping/adding of predictors of interest. The estimation results are shown in Table 2 for the case when $r_1 = r_2 = 2$ and $r_3 = 1$, and analogous results for other choices of the lags are presented in "Appendix A". Specifically, the first column indicates which coefficients of the index function were estimated: β_0 is the constant, β_1 is the coefficient for the spread, β_2 is the coefficient for the output gap and β_3 is the coefficient for lagged dependent variable. The second column reports the parametric estimates and the third column presents their standard errors, while the fourth and the fifth columns present the corresponding t statistics and p values

[16] Each model required re-estimation of all the bandwidths. To simplify the computations, we used the rule-of-thumb bandwidths from PSZ for all of them, which have correct theoretical rates and showed good performance in simulations in PSZ. Specifically, for a continuous predictor X_j, we used $h_j(0) = 1.06T^{-1/(4+r)} \times \hat{\sigma}_{j,0}$ and $h_j(1) = 1.06T^{-1/(4+r)} \times \hat{\sigma}_{j,1}$, where $\hat{\sigma}_{j,0}$ and $\hat{\sigma}_{j,1}$ are the estimated standard deviations from data on X_j corresponding to $Y = 0$ and $Y = 1$, respectively. Meanwhile, for the discrete variable we use $\gamma = 0.1n^{-2/(d+4)}$. We also tried the maximum likelihood cross-validation approach (adapted from PSZ), yet it exhibited some instability and sensitivity to starting values, running into the problem of 'spurious optima.' This caveat is known for these methods to often occur particularly for small samples like ours, which was also noticed in the simulations and is discussed in Sect. 4 of PSZ, where the rule-of-thumb bandwidths often outperformed the cross-validation bandwidths. In any case, improving bandwidth selection would therefore be another natural direction for future research.

[17] Specifically, we also used the Efron Pseudo-R^2, which gave similar results. For details of these measures and related discussion, see PSZ and references therein.

Table 2 Parametric and nonparametric estimates of the dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 2$, lag of $X_2 = 2$, lag of $Z = 1$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.430	0.186	− 2.309	0.021	− 0.033
β_1	0.364	0.140	2.598	0.009	0.184
Max.log-likelihood	− 0.654				− 0.615
Estrella pseudo-R^2	0.074				0.148
Specification 2					
β_0	− 0.165	0.136	− 1.209	0.227	0.091
β_2	1.444	0.534	2.704	0.007	1.531
Max.log-likelihood	− 0.629				− 0.610
Estrella Pseudo-R^2	0.123				0.158
Specification 3					
β_0	− 0.657	0.211	− 3.116	0.002	0.147
β_1	0.470	0.152	3.082	0.002	0.230
β_2	1.809	0.600	3.018	0.003	2.002
Max.log-likelihood	− 0.573				− 0.522
Estrella Pseudo-R^2	0.229				0.321
Specification 4					
β_0	− 1.198	0.228	− 5.263	0.000	NaN
β_3	2.406	0.337	7.136	0.000	NaN
Max.log-likelihood	− 0.356				NaN
Estrella pseudo-R^2	0.600				NaN
Specification 5					
β_0	− 1.443	0.287	− 5.025	0.000	− 0.040
β_1	0.267	0.180	1.487	0.137	0.671
β_3	2.366	0.343	6.893	0.000	NaN
Max.log-likelihood	− 0.344				− 0.303
Estrella pseudo-R^2	0.619				0.679
Specification 6					
β_0	− 1.190	0.230	− 5.166	0.000	0.222
β_2	0.807	0.680	1.187	0.235	0.854
β_3	2.296	0.346	6.635	0.000	NaN
Max.log-likelihood	− 0.346				− 0.332
Estrella pseudo-R^2	0.615				0.638
Specification 7					
β_0	− 1.500	0.300	− 5.000	0.000	0.001

Table 2 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
β_1	0.326	0.184	1.771	0.076	0.355
β_2	1.096	0.753	1.456	0.145	4.801
β_3	2.211	0.355	6.225	0.000	NaN
Max.log-likelihood	− 0.329				− 0.286
Estrella pseudo-R^2	0.641				0.705

for the two-sided tests (relying on the asymptotic normality), respectively. Since the nonparametric estimates of the coefficients vary across the observations, in this table we only present their *averages* (where available)—reported in the last column of the table and discussed in more detail further on.[18]

In principle, it is possible to automatize the model selection process, e.g., using currently popular statistical approaches in Machine Learning (e.g., forward step-wise selection, best subset selection, various LASSO approaches, etc.) to our modeling, to arrive at a final parsimonious model suggested by "the machines" based on some pre-specified statistical criteria. However, in the case of a small number of predictors like ours, it might be more valuable, at least for illustration/pedagogical purposes, to discuss how a model can be selected by practitioners, also in the spirit of forward step-wise and best subset selection methods, yet thoughtfully rather than automatically.

We start with the Specification 1 in Table 2, which considers the spread as the only predictor, i.e., it is the specification that Estrella and Mishkin (1997) used for the US economy and the originator of the paradigm we tried to adapt and extend here. One can see that the parametric estimate of β_1 is 0.364 and is statistically significant at 1%. Note, however, that the average of the nonparametric estimates is substantially smaller, around 0.184. Moreover, the Pseudo-R^2 of the parametric model is quite low, about 0.074, while for the nonparametric it is 0.15, i.e., about 2 times higher than for the parametric model, yet still relatively low from a perspective of predictive power, thus encouraging to try other or more predictors, as we do below.

The Specification 2 considers the output gap as the only predictor. One can see that the parametric estimate of β_2 is 1.444 and is also statistically significant at 1%. Notably, it is substantially larger in magnitude than the estimate of β_1 in the Specification 1 (note that the data on both variables were standardized). Interestingly, the average of the nonparametric estimates is 1.531, i.e., has the same sign of the relationship and is also similar in magnitude to the parametric estimate. Meanwhile, the Pseudo-R^2 of the parametric and nonparametric models are still fairly low, about 0.123 and 0.158, respectively, which is somewhat better than for Specification 1, especially for the parametric model.

[18] The nonparametric approach does not provide an estimate of β_3 and so it is not presented. In principle, what is possible is to estimate $\Pr\left(Y_t = 1|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 1\right) - \Pr\left(Y_t = 1|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 0\right)$ and $\Pr\left(Y_t = 0|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 1\right) - \Pr\left(Y_t = 0|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 0\right)$ as well as $\Pr\left(Y_t = 1|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 1\right) - \Pr\left(Y_t = 0|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 0\right)$ and $\Pr\left(Y_t = 0|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 1\right) - \Pr\left(Y_t = 1|\boldsymbol{X} = \boldsymbol{x}, Y_{t-1} = 0\right)$. As with other nonparametric estimates, these may vary across different values of $\boldsymbol{x}$ and so we do not present them for the sake of brevity.

The Specification 3 considers both the spread and the output gap as the two predictors. The parametric estimate of the coefficient of the spread (β_1) is now 0.470, which is a bit larger relative to what it was in Specification 1, and continues to be statistically significant at 1%. Note that the average of the nonparametric estimates for this coefficient is about two times smaller. Meanwhile, the parametric estimate of the output gap (β_2) is now 1.81, which is slightly larger than what it was in Specification 2 (and continues to be a statistically significant predictor at 1%), while the average of the nonparametric estimates of β_2 is higher, about 2.0, which is still fairly similar in magnitude to the parametric estimate. The Pseudo-R^2 of the parametric and nonparametric models is now 0.23 and 0.32, i.e., both improved substantially relative to Specifications 1 and 2, suggesting that both variables have something 'valuable to tell us' in terms of predictions of the recession for these data.

Specification 4 is analogous to Specifications 1 and 2, except that it takes the lagged value of the dependent variable as the only predictor. Since there is no continuous variable in this specification we only use a parametric approach here, which gives 2.406 as the estimate of β_3, with a high statistical significance (well under 1%). Moreover, the Pseudo-R^2 here is 0.6, which is the highest so far.

Specification 5 has the spread variable and the lagged recession indicator, i.e., this is the model analogous to Duecker (1997), Kauppi and Saikkonen (2008), Park et al. (2020) and many others. Specification 6 has the output gap variable and the lagged recession indicator, while Specification 7 has all three variables in the model. In all three cases, the estimate of β_3 remained similar (albeit slightly lower) relative to Specification 4, while the Pseudo-R^2 increased to some extent, with the highest one for Specification 7 (about 0.64 for the parametric and 0.71 for the nonparametric approaches). Meanwhile, relative to those from Specifications 1, 2 and 3, the magnitudes of the estimates of β_1 and β_2 decreased further (especially relative to Specification 3, which had both of them), while their standard errors increased further. In turn, this led to a substantial increase in p-values to around 0.137 and 0.235 for β_1 and β_2 in Specifications 5 and 6, respectively, and 0.076 and 0.145 in Specification 7, for the two-sided tests or half of those values for the one-sided tests.

Note that while we presented the two-sided test results, the one-sided tests might be indeed more relevant here: *a priori* we would expect that the increasing output gap of a country (i.e., its further lagging behind relative to peer countries) could serve as an early signal of the country entering a recession. Similarly with our definition of the spread: the increase in the difference between the Italian bonds and the German bonds is a cumulative signal of what investors sense about the Italian economy, which may reflect the true dynamics or contribute to the 'self-fulfilling prophecies' by forcing local businesses to pay higher local interest rates or reduce the local investments. Even more evident is the expected sign for the relationship between the recession indicator and its lagged value: the majority of the quarters are where $Y_t = 0$ (i.e., no recession) and most of them are also followed by $Y_t = 0$, until the switch to $Y_t = 1$ (recession) that stays for a few quarters as $Y_t = 0$ until it switches back to $Y_t = 1$ and so on, i.e., implying a positive relationship between Y_t and Y_{t-1}.

The phenomenon where a powerful predictor of the dependent variable is the lagged dependent variable, and possibly dominating all other predictors is, of course, very common in time series. However, while it appears as the most powerful predictor of

the three variables that we considered here, it is important to note that the precise information on our lagged value of the recession indicator is often not available for the most recent periods in real time, which are also the periods that are the most important for the prediction of future periods. This is because the OECD decisions on the turning points of a business cycle (peak or trough) from which the variable is constructed usually come with some delay (similarly as for NBER data about USA), which may be as long as a few months to a few quarters. That is, while there is a lot of useful historical information in this variable, most of it is 'too old' for the actual prediction of the future. And, this is where the other two predictors might be useful, although their overall predictive power is partially taken over by the lagged dependent variable *once it becomes available*, making β_1 and β_2 significant only at 5% and 10%, respectively, in the one-sided tests. Overall, considering this phenomenon and the relatively small sample (96 observations), we deem these two continuous variables as useful predictors for the case of Italy and in what follows we will focus on Specification 7, which we will refer to as the 'final specification.'[19]

3.4.3 Insights from the final specification

While focusing on the final specification, one can see that the nonparametric complete smoothing approach offers similar (and slightly better) results as the parametric probit on both the achieved maximum likelihood value and of the Pseudo-R^2. Indeed, the Pseudo-R^2 is around 64% for the parametric approach, while it is 71% for the nonparametric approach. This suggests that the linearity assumption in the parametric approach may be a reasonable approximation for both X_1 (the Spread, $S_{p,t-2}$) and X_2 (the output gap, $\Delta_{G,t-2}$) for these data. Although this simplification led to a slightly lower attained goodness of fit, its simplicity of estimation, especially due to the readily available inference procedures, may warrant it a status of the preferred approach for these data and specification. Meanwhile, the nonparametric approach can serve here as a robustness check tool and so a few words on this are in order.

Figure 6 exhibits the boxplots of the resulting local estimates of β_0, β_1 and β_2 for the two states of the economy $Y_t = 1$ (recessions) and $Y_t = 0$ (expansions).[20] It is interesting to see some similarity as well as substantial difference in the local estimates of β_0, β_1 and β_2 across the two groups of observations. In particular, note that only the medians of β_2 are somewhat similar, suggesting about some stability of the relationship between the predictors and the response variable regardless of the state of the economy. (The estimates are very different for β_0, which is expected since the estimate of β_0 determines the estimate of the probability of recession, via the link function.) Also note that the median of the nonparametric estimates for β_1 is nearly zero for $Y_t = 1$ and positive (around 0.5) when $Y_t = 0$. This suggests that in these data the spread variable ($X_{1,t} = S_{p,t-2}$) appears to be a more powerful predictor during expansionary periods relative to recessionary periods, which is somewhat intuitive and resembles the so-called liquidity-trap phenomenon in macroeconomics. One can

[19] The final specification, of course, may be different for other countries or even for Italy with a much larger sample—verifying this could be an interesting avenue for future research, and the goal of a somewhat detailed description above is to show a possible (albeit not only) algorithm for such research.

[20] The axes have been trimmed to visualize all boxplots on one scale.

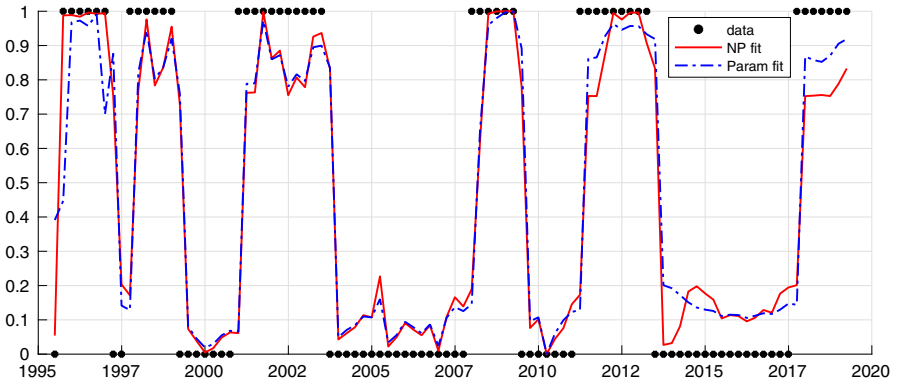

Fig. 7 In-sample fit for recessions in Italy, 1995–2019, with $n = 98$ data points

also observe a greater range and the interquartile range, as well as more outliers in the recessionary periods for this variable, suggesting about greater possible estimation noise. It is also coherent with the fact that while all recessions are coded here as $Y_t = 1$, many (if not all) of them have many unique features as well as those caused by possible different compositions of factors and triggers, which in turn make predicting recessions a very challenging task.

We see an even greater difference for the first difference of our measure of output gap ($X_{2,t} = \Delta_{G,t-2}$). In particular, note that while the median of estimates for β_2 are similar in the two states of the economy, the range is very different and is larger for the recessionary periods. In both cases, the median is around 1.5, suggesting that the positive growth in inefficiency, our measure of output gap, is associated with an increase in the probability to be in a recession. This positive association seems to be more pronounced in expansionary periods of the economy, as we see mostly a positive range there.[21]

We now look at the in-sample fit for modeling the probability of recessions in Fig. 7. We can indeed observe that both the nonparametric and parametric approaches fit the data well (as seen with various measures described above). In particular, note that most of the recession periods, as established by the turning points of OECD.stat, are successfully captured by our model both using the parametric and the nonparametric approaches.

[21] Here it is also worth contrasting the difference between the mean and the median of the estimates: the former is substantially larger due to two extreme estimates of β_2, which were around 155 (corresponding to the first two periods in the data and not seen on the boxplots due to trimming of the axes), distorting the mean substantially away from the median. Observing such anomalous estimates in nonparametric approaches for some observations in real data is not uncommon, especially for relatively small samples like ours, and so the median would be a more reliable indicator of central tendency here. For the sake of illustrating the point, we reported both.

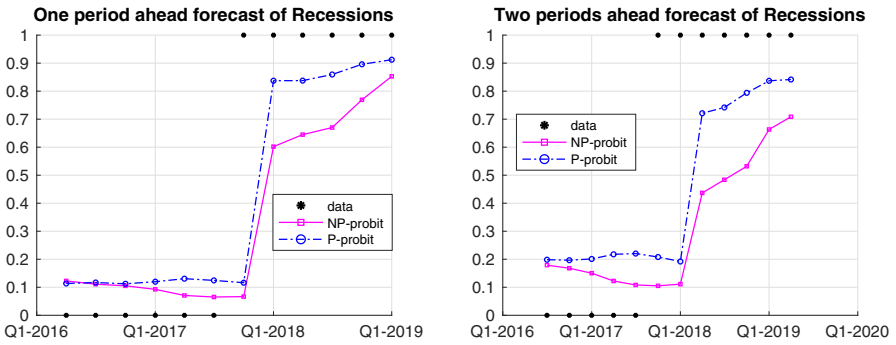

Fig. 8 Forecast of the Recessions, starting after $T = Q1 - 2016$ periods

3.5 Out-of-sample forecasts

We now proceed with the out-of-sample forecasts, to see if we can have a reasonably good prediction of the recession periods (one-period and two-periods ahead), using the data from the beginning till 2016:Q1 from either the parametric or nonparametric approaches.[22]

The forecasts of the recessions are displayed in Fig. 8. In most cases (and on average), we can observe a slightly better forecast value for the parametric approach, both for the case of the one-period ahead and for the two-periods ahead forecasts. In particular, note that, with the one-period ahead forecasts, both approaches correctly and somewhat similarly warn about the recession in Q1-2018–Q2-2019, with the parametric approach slightly outperforming. Both approaches with the one-period ahead forecasts correctly alert us to the non-recession (or expansions) in Q1-2016 through Q3-2017, though both miss on warning about the start of the recession in Q4-2017, while they manage to warn correctly about the subsequent quarters being in the recession.

Finally, it is worth recalling here that the parametric approach can be viewed as a special case of the nonparametric approach, in the sense that the latter allows for much more flexibility and can be restricted further to obtain the former through reducing this flexibility. Interestingly, for this data set we see that despite assuming a naive (linear) and quite restrictive (e.g., constancy of the first derivative) functional form for the index function, the parametric approach still produced very similar conclusions and very similar or even slightly better forecasts than the nonparametric approach which allows for much more flexibility. This suggests that, for this sample and the specifications we considered, we can have more confidence in the results and conclusions from the parametric approach, even though it imposes fairly restrictive assumptions. Of course for other data (e.g., for other countries or even the same country but for different time periods or with different variables), this similarity of parametric and nonparametric approaches may or may not hold *a priori* and so needs to be verified and validated on a case-by-case basis. Indeed, it is very easy to construct an example when parametric

[22] Available sample size (past) at this time is 81.

and nonparametric approaches deliver very different results and conclusions, (e.g., see Monte Carlo examples in PSZ).

4 Concluding remarks

In this paper, we have attempted to merge two so far largely unrelated streams of literature. The first stream is about the non-parametric methods to estimate frontier efficiency of an economy, which we tailor to estimate the output gap of a country. The other stream is the literature on predicting economic recessions. We considered various methods among the myriad of approaches, selecting and tailoring one that currently appears to be the most suitable for a new measure of output gap to be used, *inter alia*, for estimating the probability of economic recessions. For the latter goal, we have chosen the paradigm started by Estrella and Mishkin (1995, 1998), further refined by Duecker (1997) and Kauppi and Saikkonen (2008) as well as their nonparametric version recently developed by Park et al. (2017, 2020). Naturally, endeavoring to merge the economic efficiency literature with other methods from the many paradigms for forecasting of economic recessions would be a natural direction for future research.

To illustrate our proposed framework that resulted from the merger of two different literatures, we apply it to the context and data on the Italian economy. In particular, we utilize the data from 1995 to 2019 and find that the proposed approach (using both the linear probit model and its non-parametric version) is capable of giving useful insights, although of course it is not a 'crystal ball' and more work is needed to refine and further improve this method and, possibly, synthesize it with other methods as well as try it on other data sets. In particular, it appears that our measure of output gap, based on efficiency measures in general and via the estimation approach we considered here, is sound conceptually and can be useful as a predictor (or a proxy) in the models for forecasting recessions and perhaps other macroeconomic models. We acknowledge that there, of course, could be many other good predictors or proxies for similar or different reasons and they could be fruitful avenues for future research. Also, development of the asymptotic theory for the statistical inference in the nonparametric approach (e.g., via bootstrap) would be an important direction for future theoretical research.

Acknowledgements We thank the Editors, anonymous referees and all others who gave feedback to various versions of this paper, including Adrian Pagan, Artem Prokhorov, Peter Schmidt as well as Bao Hoang Nguyen, Evelyn Smart, Zhichao Wang, and other colleagues and participants of conferences (NAPW2018 at University of Miami, Time-Series and Forecasting Symposium at the University of Sydney, ANZESG2018 at The University of Queensland, ANZESG2020 and Monash University, Eighth Italian Congress of Econometrics and Empirical Economics (ICEEE) at the University of Salento ICEEE2019, The Econometric Society and Bocconi University Virtual World Congress 2020, 2nd Italian Workshop of Econometrics and Empirical Economics, University Ca' Foscari, Venice, etc.) and seminars at the Reserve Bank of Australia, Reserve Banks of New Zealand. These individuals and organizations are not responsible for the views expressed here.

Funding Open access funding provided by Universitá della Calabria within the CRUI-CARE Agreement.

A Sensitivity analysis and model validation

Table 3 Parametric and nonparametric estimates of dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 1$, lag of $X_2 = 2$, lag of $Z = 1$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.432	0.188	− 2.298	0.022	0.020
β_1	0.358	0.140	2.562	0.010	− 0.277
Max.log-likelihood	− 0.654				− 0.580
Estrella pseudo-R^2	0.073				0.214
Specification 2					
β_0	− 0.165	0.136	− 1.209	0.227	0.091
β_2	1.444	0.534	2.704	0.007	1.531
Max.log-likelihood	− 0.629				− 0.610
Estrella pseudo-R^2	0.123				0.158
Specification 3					
β_0	− 0.624	0.209	− 2.985	0.003	0.154
β_1	0.434	0.148	2.939	0.003	− 0.033
β_2	1.730	0.594	2.913	0.004	1.466
Max.log-likelihood	− 0.578				− 0.479
Estrella pseudo-R^2	0.219				0.397
Specification 4					
β_0	− 1.198	0.228	− 5.263	0.000	NaN
β_3	2.406	0.337	7.136	0.000	NaN
Max.log-likelihood	− 0.356				NaN
Estrella pseudo-R^2	0.600				NaN
Specification 5					
β_0	− 1.424	0.285	− 4.998	0.000	0.007
β_1	0.249	0.177	1.403	0.161	0.177
β_3	2.360	0.343	6.888	0.000	NaN
Max.log-likelihood	− 0.346				− 0.309
Estrella pseudo-R^2	0.617				0.671

Table 3 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 6					
β_0	− 1.190	0.230	− 5.166	0.000	0.222
β_2	0.807	0.680	1.187	0.235	0.854
β_3	2.296	0.346	6.635	0.000	NaN
Max.log-likelihood	− 0.346				− 0.332
Estrella pseudo-R^2	0.615				0.638
Specification 7					
β_0	− 1.465	0.294	− 4.978	0.000	0.066
β_1	0.293	0.180	1.634	0.102	− 0.021
β_2	1.024	0.730	1.402	0.161	1.446
β_3	2.215	0.355	6.248	0.000	NaN
Max.log-likelihood	− 0.332				− 0.287
Estrella pseudo-R^2	0.637				0.703

Table 4 Parametric and nonparametric estimates of dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 3$, lag of $X_2 = 2$, lag of $Z = 1$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.441	0.187	− 2.359	0.018	0.360
β_1	0.397	0.144	2.760	0.006	0.653
Max.log-likelihood	− 0.650				− 0.604
Estrella pseudo-R^2	0.083				0.171
Specification 2					
β_0	− 0.152	0.137	− 1.109	0.267	0.101
β_2	1.444	0.537	2.690	0.007	1.525
Max.log-likelihood	− 0.629				− 0.611
Estrella pseudo-R^2	0.123				0.157
Specification 3					
β_0	− 0.699	0.217	− 3.226	0.001	0.505
β_1	0.535	0.164	3.265	0.001	0.757
β_2	1.885	0.616	3.059	0.002	1.628
Max.log-likelihood	− 0.564				− 0.505
Estrella pseudo-R^2	0.245				0.353
Specification 4					
β_0	− 1.187	0.229	− 5.189	0.000	NaN
β_3	2.394	0.338	7.088	0.000	NaN
Max.log-likelihood	− 0.358				NaN
Estrella pseudo-R^2	0.597				NaN
Specification 5					
β_0	− 1.385	0.282	− 4.908	0.000	0.140
β_1	0.233	0.191	1.220	0.222	0.614

Table 4 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
β_3	2.330	0.343	6.803	0.000	NaN
Max.log-likelihood	− 0.350				− 0.294
Estrella pseudo-R^2	0.611				0.694
Specification 6					
β_0	− 1.178	0.231	− 5.091	0.000	0.241
β_2	0.810	0.682	1.188	0.235	0.845
β_3	2.284	0.347	6.587	0.000	NaN
Max.log-likelihood	− 0.349				− 0.334
Estrella pseudo-R^2	0.613				0.635
Specification 7					
β_0	− 1.453	0.298	− 4.871	0.000	0.128
β_1	0.311	0.201	1.548	0.122	0.600
β_2	1.113	0.761	1.463	0.144	1.578
β_3	2.163	0.356	6.077	0.000	NaN
Max.log-likelihood	− 0.334				− 0.282
Estrella pseudo-R^2	0.634				0.711

Table 5 Parametric and nonparametric estimates of dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 3$, lag of $X_2 = 3$, lag of $Z = 1$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.441	0.187	− 2.359	0.018	0.360
β_1	0.397	0.144	2.760	0.006	0.653
Max.log-likelihood	− 0.650				− 0.604
Estrella pseudo-R^2	0.083				0.171
Specification 2					
β_0	− 0.117	0.134	− 0.869	0.385	− 0.128
β_2	0.665	0.385	1.726	0.084	1.073
Max.log-likelihood	− 0.666				− 0.630
Estrella pseudo-R^2	0.051				0.122
Specification 3					
β_0	− 0.559	0.200	− 2.792	0.005	0.261
β_1	0.446	0.149	2.993	0.003	0.491
β_2	0.822	0.405	2.028	0.043	1.764
Max.log-likelihood	− 0.616				− 0.541
Estrella pseudo-R^2	0.149				0.288

Table 5 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 4					
β_0	− 1.187	0.229	− 5.189	0.000	NaN
β_3	2.394	0.338	7.088	0.000	NaN
Max.log-likelihood	− 0.358				NaN
Estrella pseudo-R^2	0.597				NaN
Specification 5					
β_0	− 1.385	0.282	− 4.908	0.000	0.140
β_1	0.233	0.191	1.220	0.222	0.614
β_3	2.330	0.343	6.803	0.000	NaN
Max.log-likelihood	−0.350				− 0.294
Estrella pseudo-R^2	0.611				0.694
Specification 6					
β_0	− 1.184	0.231	− 5.132	0.000	0.070
β_2	0.017	0.184	0.093	0.926	0.954
β_3	2.387	0.345	6.920	0.000	NaN
Max.log-likelihood	− 0.358				− 0.270
Estrella pseudo-R^2	0.597				0.728
Specification 7					
β_0	− 1.383	0.283	− 4.885	0.000	− 0.098
β_1	0.237	0.191	1.241	0.215	0.762
β_2	0.045	0.213	0.211	0.833	0.988
β_3	2.312	0.352	6.570	0.000	NaN
Max.log-likelihood	− 0.349				− 0.250
Estrella pseudo-R^2	0.611				0.756

Table 6 Parametric and nonparametric estimates of dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 2$, lag of $X_2 = 3$, lag of $Z = 1$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.494	0.191	− 2.587	0.010	0.666
β_1	0.449	0.148	3.025	0.002	0.864
Max.log-likelihood	− 0.640				− 0.594
Estrella Pseudo-R^2	0.102				0.190
Specification 2					
β_0	− 0.117	0.134	− 0.869	0.385	− 0.128
β_2	0.665	0.385	1.726	0.084	1.073
Max.log-likelihood	− 0.666				− 0.630
Estrella pseudo-R^2	0.051				0.122

Table 6 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 3					
β_0	− 0.593	0.202	− 2.940	0.003	0.469
β_1	0.480	0.151	3.175	0.001	0.661
β_2	0.788	0.409	1.927	0.054	1.768
Max.log-likelihood	− 0.608				− 0.537
Estrella pseudo-R^2	0.163				0.295
Specification 4					
β_0	− 1.187	0.229	− 5.189	0.000	NaN
β_3	2.394	0.338	7.088	0.000	NaN
Max.log-likelihood	− 0.358				NaN
Estrella pseudo-R^2	0.597				NaN
Specification 5					
β_0	− 1.467	0.291	− 5.034	0.000	0.184
β_1	0.316	0.193	1.637	0.102	1.132
β_3	2.332	0.345	6.752	0.000	NaN
Max.log-likelihood	− 0.343				− 0.297
Estrella pseudo-R^2	0.621				0.690
Specification 6					
β_0	− 1.184	0.231	− 5.132	0.000	0.070
β_2	0.017	0.184	0.093	0.926	0.954
β_3	2.387	0.345	6.920	0.000	NaN
Max.log-likelihood	− 0.358				− 0.270
Estrella pseudo-R^2	0.597				0.728
Specification 7					
β_0	− 1.463	0.292	− 5.005	0.000	− 0.024
β_1	0.318	0.193	1.648	0.099	1.054
β_2	0.038	0.209	0.181	0.857	1.623
β_3	2.317	0.355	6.534	0.000	NaN
Max.log-likelihood	− 0.343				− 0.265
Estrella pseudo-R^2	0.622				0.735

Table 7 Parametric and nonparametric estimates of dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 1$, lag of $X_2 = 3$, lag of $Z = 1$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.522	0.194	− 2.685	0.007	0.538
β_1	0.464	0.148	3.130	0.002	0.808
Max.Log-Likelihood	− 0.636				− 0.583
Estrella pseudo-R^2	0.109				0.210
Specification 2					
β_0	− 0.117	0.134	− 0.869	0.385	− 0.128
β_2	0.665	0.385	1.726	0.084	1.073
Max.log-likelihood	− 0.666				− 0.630
Estrella pseudo-R^2	0.051				0.122
Specification 3					
β_0	− 0.631	0.206	− 3.071	0.002	0.508
β_1	0.504	0.151	3.333	0.001	0.755
β_2	0.816	0.413	1.976	0.048	2.031
Max.log-likelihood	− 0.602				− 0.518
Estrella pseudo-R^2	0.174				0.330
Specification 4					
β_0	− 1.187	0.229	− 5.189	0.000	NaN
β_3	2.394	0.338	7.088	0.000	NaN
Max.log-likelihood	− 0.358				NaN
Estrella pseudo-R^2	0.597				NaN
Specification 5					
β_0	−1.462	0.292	− 5.001	0.000	0.207
β_1	0.313	0.196	1.596	0.110	1.336
β_3	2.320	0.345	6.718	0.000	NaN
Max.log-likelihood	− 0.344				− 0.304
Estrella pseudo-R^2	0.620				0.680
Specification 6					
β_0	− 1.184	0.231	− 5.132	0.000	0.070
β_2	0.017	0.184	0.093	0.926	0.954
β_3	2.387	0.345	6.920	0.000	NaN
Max.log-likelihood	− 0.358				− 0.270
Estrella pseudo-R^2	0.597				0.728

Table 7 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 7					
β_0	− 1.459	0.293	− 4.983	0.000	− 0.027
β_1	0.318	0.196	1.619	0.105	1.286
β_2	0.052	0.210	0.245	0.806	1.937
β_3	2.299	0.355	6.480	0.000	NaN
Max.log-likelihood	− 0.343				− 0.264
Estrella pseudo-R^2	0.621				0.736

Table 8 Parametric and nonparametric estimates of dynamic probit model of recessions in Italy, 1995–2019 (when lag of $X_1 = 2$, lag of $X_2 = 2$, lag of $Z = 2$)

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 1					
β_0	− 0.430	0.186	− 2.309	0.021	− 0.033
β_1	0.364	0.140	2.598	0.009	0.184
Max.log-likelihood	− 0.654				− 0.615
Estrella pseudo-R^2	0.074				0.148
Specification 2					
β_0	− 0.165	0.136	− 1.209	0.227	0.091
β_2	1.444	0.534	2.704	0.007	1.531
Max.log-likelihood	− 0.629				− 0.610
Estrella pseudo-R^2	0.123				0.158
Specification 3					
β_0	− 0.657	0.211	− 3.116	0.002	0.147
β_1	0.470	0.152	3.082	0.002	0.230
β_2	1.809	0.600	3.018	0.003	2.002
Max.log-likelihood	− 0.573				− 0.522
Estrella pseudo-R^2	0.229				0.321
Specification 4					
β_0	− 0.751	0.191	− 3.930	0.000	NaN
β_3	1.481	0.285	5.206	0.000	NaN
Max.log-likelihood	− 0.538				NaN
Estrella pseudo-R^2	0.292				NaN

Table 8 continued

	ParamEst	SE	*t* stat	*p* val	NPEst
Specification 5					
β_0	− 1.018	0.238	− 4.284	0.000	0.021
β_1	0.296	0.150	1.973	0.049	0.755
β_3	1.430	0.290	4.930	0.000	NaN
Max.log-likelihood	− 0.518				− 0.464
Estrella pseudo-R^2	0.330				0.424
Specification 6					
β_0	− 0.727	0.193	− 3.764	0.000	0.210
β_2	0.962	0.575	1.673	0.094	1.416
β_3	1.325	0.297	4.460	0.000	NaN
Max.log-likelihood	− 0.518				− 0.486
Estrella pseudo-R^2	0.329				0.385
Specification 7					
β_0	− 1.061	0.247	− 4.287	0.000	0.233
β_1	0.366	0.157	2.330	0.020	0.627
β_2	1.263	0.636	1.986	0.047	0.649
β_3	1.207	0.306	3.941	0.000	NaN
Max.log-likelihood	− 0.488				− 0.417
Estrella pseudo-R^2	0.382				0.502

B Data description

Country		*Y*	*K*	*L*
Austria	Mean	3.35E+05	1.25E+06	3963.9
	SD	40359	1.13E+05	235.81
Belgium	Mean	4.11E+05	1.43E+06	4403.5
	SD	48008	2.16E+05	285.89
Denmark	Mean	2.34E+05	9.16E+05	2777.3
	SD	21562	2.10E+05	78.092
Finland	Mean	1.94E+05	7.41E+05	2405.5
	SD	25712	41161	156.8
France	Mean	2.24E+06	6.87E+06	26681
	SD	2.34E+05	2.15E+06	1356.2
Germany	Mean	3.14E+06	1.02E+07	40683
	SD	3.02E+05	6.55E+05	2052.3
Ireland	Mean	1.95E+05	2.48E+06	1854.3
	SD	63151	2.48E+06	250.37
Israel	Mean	2.01E+05	7.37E+05	3193.9
	SD	51294	1.37E+05	596.6
Italy	Mean	2.03E+06	5.78E+06	23710
	SD	92116	1.53E+06	1146
Korea	Mean	1.38E+06	6.84E+06	23429
	SD	3.92E+05	1.08E+06	2170.3
Netherlands	Mean	7.07E+05	2.31E+06	8412.5
	SD	84283	3.30E+05	485.19
New Zealand	Mean	1.27E+05	3.98E+05	2096.1
	SD	24578	1.06E+05	276.78
Norway	Mean	2.71E+05	8.50E+05	2498.4
	SD	34325	2.09E+05	221.85
Spain	Mean	1.36E+06	5.07E+06	18192
	SD	1.89E+05	9.93E+05	2143
Sweden	Mean	3.67E+05	1.24E+06	4480.6
	SD	59666	2.82E+05	336.44
UK	Mean	2.19E+06	5.69E+06	29036
	SD	2.93E+05	7.03E+05	1878.7

C Measures of efficiencies

Table 9 Different measures of efficiency: efficiency δ, Order-m Efficiency δ_m, Time Conditional Efficiency λ, Order-m time conditional efficiency λ_m, pure time conditional efficiency GAP, Order-m pure time conditional Efficiency GAP_m

Date	δ	δ_m	λ	λ_m	GAP	GAP_m
1995.1	0	− 0.21851	1	0.98365	0	− 0.22212
1995.2	0	− 0.26505	1	0.98018	0	− 0.26932
1995.3	0.000446	− 0.033551	1	0.99749	0.000453	− 0.034077
1995.4	0	− 0.035272	1	0.99737	0	− 0.03581
1996.1	0	− 0.21622	1	0.98387	0	− 0.21942
1996.2	0	− 0.215	1	0.98397	0	− 0.21809
1996.3	0.012553	0.012265	1.0009	1.0009	0.012727	0.012435
1996.4	0	− 0.21445	1	0.98402	0	− 0.21732
1997.1	0	− 0.21437	1	0.98404	0	− 0.21714
1997.2	0	− 0.21643	1	0.98391	0	− 0.21913
1997.3	0	− 0.21701	1	0.98388	0	− 0.21961
1997.4	0	− 0.11868	1	0.9912	0	− 0.12004
1998.1	0.010762	− 0.015923	1.0008	0.99882	0.01088	− 0.016098
1998.2	0.011665	− 0.01464	1.0009	0.99892	0.011787	− 0.014793
1998.3	0.014662	0.011138	1.0011	1.0008	0.014808	0.011249
1998.4	0.023339	− 0.003768	1.0017	0.99972	0.023559	− 0.0038035
1999.1	0.022348	− 0.004278	1.0017	0.99968	0.022546	− 0.004316
1999.2	0.022951	0.020456	1.0017	1.0015	0.023142	0.020627
1999.3	0.019401	0.016882	1.0014	1.0012	0.019552	0.017014
1999.4	0.010497	0.00794	1.0008	1.0006	0.010573	0.0079975
2000.1	0.004246	0.002897	1.0003	1.0002	0.0042744	0.0029164
2000.2	0.00207	0.001371	1.0002	1.0001	0.0020827	0.0013794
2000.3	0.001275	0.000977	1.0001	1.0001	0.0012821	0.00098242
2000.4	0	− 0.002169	1	0.99984	0	− 0.0021798
2001.1	0	− 0.00206	1	0.99985	0	− 0.002069
2001.2	0	− 0.000543	1	0.99996	0	− 0.00054506
2001.3	0.015874	0.015159	1.0012	1.0011	0.015925	0.015208
2001.4	0.019816	0.01909	1.0015	1.0014	0.019868	0.01914
2002.1	0.024743	0.02401	1.0018	1.0018	0.024792	0.024058
2002.2	0.024744	0.023996	1.0018	1.0018	0.024778	0.024029
2002.3	0.026664	0.025895	1.0019	1.0019	0.026685	0.025915
2002.4	0.027563	0.026776	1.002	1.002	0.027567	0.02678

Table 9 continued

Date	δ	δ_m	λ	λ_m	GAP	GAP_m
2003.1	0.034427	0.033625	1.0025	1.0025	0.034411	0.033609
2003.2	0.041866	0.041056	1.0031	1.003	0.04182	0.041011
2003.3	0.044656	0.043837	1.0033	1.0032	0.044579	0.043761
2003.4	0.041716	0.040892	1.003	1.003	0.041617	0.040795
2004.1	0.041579	0.040754	1.003	1.003	0.041454	0.040632
2004.2	0.042656	0.041826	1.0031	1.003	0.042501	0.041674
2004.3	0.045739	0.044903	1.0033	1.0033	0.045543	0.044711
2004.4	0.04872	0.047879	1.0035	1.0035	0.04848	0.047643
2005.1	0.055213	0.054366	1.004	1.0039	0.054905	0.054063
2005.2	0.050126	0.049274	1.0036	1.0036	0.049814	0.048967
2005.3	0.048081	0.047228	1.0035	1.0034	0.04775	0.046903
2005.4	0.049974	0.049115	1.0036	1.0035	0.049598	0.048745
2006.1	0.050511	0.04964	1.0036	1.0036	0.050098	0.049234
2006.2	0.049037	0.04815	1.0035	1.0035	0.048604	0.047725
02006.3	0.049926	0.049049	1.0036	1.0035	0.049452	0.048584
2006.4	0.042162	0.041261	1.003	1.003	0.041735	0.040843
2007.1	0.044862	0.043961	1.0032	1.0032	0.044378	0.043487
2007.2	0.049876	0.04896	1.0036	1.0035	0.049305	0.0484
2007.3	0.054173	0.053248	1.0039	1.0038	0.053518	0.052604
2007.4	0.059529	0.058596	1.0043	1.0042	0.058771	0.05785
2008.1	0.053911	0.052988	1.0039	1.0038	0.05319	0.052279
2008.2	0.067195	0.066277	1.0048	1.0047	0.066253	0.065348
2008.3	0.084608	0.083695	1.0061	1.006	0.083367	0.082468
2008.4	0.11102	0.11012	1.008	1.0079	0.10932	0.10843
2009.1	0.14186	0.14097	1.0102	1.0101	0.1396	0.13872
2009.2	0.1445	0.14278	1.0104	1.0103	0.14211	0.14042
2009.3	0.14355	0.14056	1.0103	1.0101	0.14108	0.13815
2009.4	0.14433	0.10546	1.0103	1.0076	0.14175	0.10358
2010.1	0.004596	0.003892	1.0003	1.0003	0.0045113	0.0038203
2010.2	0.001087	0.00084	1.0001	1.0001	0.0010663	0.000824
2010.3	0	− 0.001903	1	0.99986	0	− 0.0018656
2010.4	0	− 0.001183	1	0.99992	0	− 0.001159
2011.1	0	− 0.00195	1	0.99986	0	− 0.0019093
2011.2	0	− 0.003906	1	0.99972	0	− 0.0038222

Table 9 continued

Date	δ	δ_m	λ	λ_m	GAP	GAP_m
2011.3	0	− 0.004637	1	0.99967	0	− 0.0045348
2011.4	0	− 0.004595	1	0.99967	0	− 0.004491
2012.1	0	− 0.006826	1	0.99951	0	− 0.0066676
2012.2	0	− 0.003307	1	0.99976	0	− 0.0032284
2012.3	0	− 0.003546	1	0.99975	0	− 0.0034597
2012.4	0	− 0.003994	1	0.99972	0	− 0.0038945
2013.1	0	− 0.001276	1	0.99991	0	− 0.0012435
2013.2	0	− 0.000628	1	0.99996	0	− 0.00061167
2013.3	0	− 0.004447	1	0.99968	0	− 0.004329
2013.4	0	− 0.036703	1	0.99739	0	− 0.035709
2014.1	0	− 0.00407	1	0.99971	0	− 0.0039576
2014.2	0	− 0.002331	1	0.99983	0	− 0.0022654
2014.3	0	− 0.001315	1	0.99991	0	− 0.0012773
2014.4	0	− 0.000639	1	0.99995	0	− 0.00062037
2015.1	0	− 0.00888	1	0.99937	0	− 0.0086167
2015.2	0	− 0.037797	1	0.99732	0	− 0.036658
2015.3	0	− 0.038495	1	0.99727	0	− 0.037316
2015.4	0	− 0.006379	1	0.99955	0	− 0.0061805
2016.1	0	− 0.001171	1	0.99992	0	− 0.001134
2016.2	0	− 0.000165	1	0.99999	0	− 0.00015971
2016.3	0	− 0.005117	1	0.99964	0	− 0.0049506
2016.4	0	− 0.000651	1	0.99995	0	− 0.00062954
2017.1	0	− 0.000332	1	0.99998	0	− 0.00032091
2017.2	0	− 0.000683	1	0.99995	0	− 0.00065988
2017.3	0	− 0.000485	1	0.99997	0	− 0.00046837
2017.4	0	− 0.001568	1	0.99989	0	− 0.0015136
2018.1	0	− 0.00429	1	0.9997	0	− 0.0041393
2018.2	0	− 0.001815	1	0.99987	0	− 0.0017505
2018.3	0	− 0.00188	1	0.99987	0	− 0.0018124
2018.4	0	− 0.009783	1	0.99931	0	− 0.0094274
2019.1	0	− 0.009008	1	0.99937	0	− 0.0086771
2019.2	0	− 0.034192	1	0.9976	0	− 0.032923

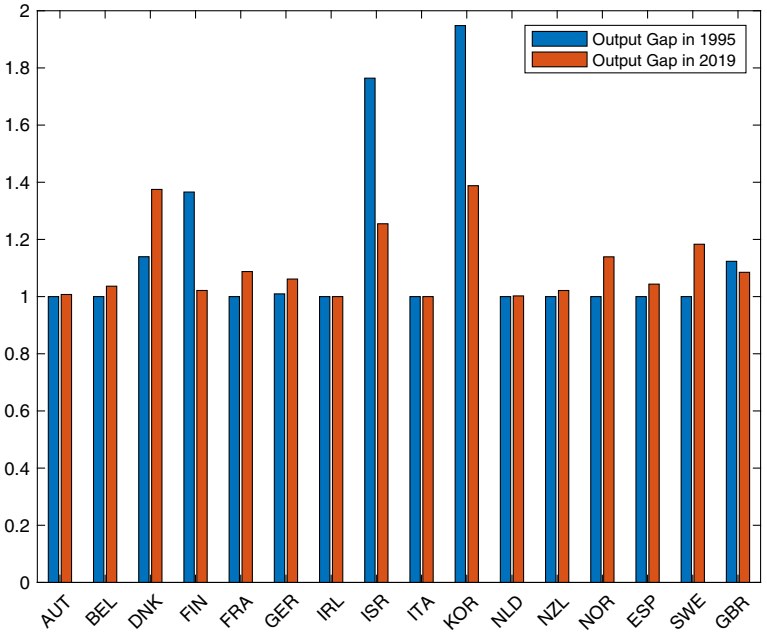

Fig. 9 Our estimates of output gap for a sample of 16 countries from for 1995 and 2019

References

Aiello F, Mastromarco C, Zago A (2011) Be productive or face decline. On the sources and determinants of output growth in Italian manufacturing firms. Empir Econ 41:787–815

Aitchison J, Aitken CGG (1976) Multivariate binary discrimination by the kernel method. Biometrika 63:413–420

Apel M, Jansson P (1999) A theory-consistent system approach for estimating potential output and the NAIRU. Econ Lett 64:271–275

Basistha A, Nelson CR (2007) New measures of the output gap based on the forward-looking new Keynesian Phillips curve. J Monet Econ 54:498–511

Bassanetti A, Iommi M, Jona-Lasinio C, Zollino F (2004) La crescita delleconomia italiana negli anni novanta tra ritardo tecnologico e rallentamento della produttività. Temi di Discussione Banca dItalia

Basu S, Fernald JG (2009) What do we know (and not know) about potential output? Fed Reserve Bank of St Louis Rev 91:187–213

Baxter M, King R (1999) Measuring business cycles: approximate band-pass filters for economic time series. Rev Econ Stat 81:575–593

Beveridge S, Nelson C (1981) A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to the measurement of the business cycle. J Monet Econ 7:151–174

Bădin L, Daraio C, Simar L (2019) A bootstrap based approach for bandwidth selection in estimating conditional efficiency measures. Eur J Oper Res 277:784–797

Cazals C, Florens J, Simar L (2002) Nonparametric frontier estimation: a robust approach. J Econ 106:1–25

Clark P (1987) The cyclical component of US economic activity. Q J Econ 102:797–814

Daouia A, Gijbels I (2011) Robustness and inference in nonparametric partial-frontier modeling. J Econ 161:147–165

Daouia A, Simar L (2007) Nonparametric efficiency analysis: a multivariate conditional quantile approach. J Econ 140:375–400

Daraio C, Simar L (2005) Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J Prod Anal 24:93–121
Duecker M (1997) Strengthening the case for the yield curve as a predictor of U.S. recessions. Fed Reserve Bank St Louis Rev 79:41–51
Estrella A, Mishkin F (1997) The predictive power of the term structure of interest rates in Europe and the United States: implications for the European Central Bank. Eur Econ Rev 41(7):1375–1401
Estrella A, Mishkin FS (1995) Predicting U.S. recessions: financial variables as leading indicators. NBER working paper no. 5379
Estrella A, Mishkin FS (1998) Predicting U.S. recessions: financial variables as leading indicators. Rev Econ Stat 80:45–61
Faini R, Barba NG, Scarpa C, Wey C (2004) Contrasting Europe's decline: do product market reforms help? Fondazione Rodolfo De Benedetti
Florens J, Simar L, van Keilegom I (2014) Frontier estimation in nonparametric location-scale models. J Econ 178:456–470
Galí J, Gertler M (1999) In ation dynamics: a structural econometric analysis. J Monet Econ 44:195–222
Gerlach S, Smets F (1999) Output gaps and monetary policy in the EMU area. Eur Econ Rev 43:801–812
Harvey A (1985) Trends and cycles in macroeconomic time series. J Bus Econ Stat 3:216–227
Henderson DJ, Parmeter CF (2015) Applied nonparametric econometrics. Cambridge University Press, Cambridge
Hodrick R, Prescott EC (1997) Postwar U.S. business cycles: an empirical investigation. J Money Credit Bank 29:1–16
Horowitz JL (2009) Semiparametric and nonparametric methods in econometrics, vol 12. Springer, Berlin
Jeong S, Park B, Simar L (2010) Nonparametric conditional efficiency measures: asymptotic properties. Ann Oper Res 173:105–122
Kauppi H, Saikkonen P (2008) Predicting U.S. recessions with dynamic binary response models. Rev Econ Stat 90:777–791
Kumar S, Russell R (2002) Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. Am Econ Rev 92:527–548
Kuttner K (1994) Estimating potential output as a latent variable. J Bus Econ Stat 12:361–368
Li D, Simar L, Zelenyuk V (2016) Generalized nonparametric smoothing with mixed discrete and continuous data. Comput Statist Data Anal 100:424–444
Locke RM (1995) Remaking the Italian Economy. Cornell University Press, Ithaca
Malanima P (2005) Urbanisation and the Italian economy during the last millennium. Eur Rev Econ Hist 9(1):97–122
Mastromarco C, Simar L (2015) Effect of FDI and time on catching-up: new insights from a conditional nonparametric frontier analysis. J Appl Econ 30:826–847
Mastromarco C, Simar L (2018) Globalization and productivity: a robust nonparametric world frontier analysis. Econ Model 69:134–149
Murray C (2003) Cyclical properties of Baxter–King filtered time series. Rev Econ Stat 85:472–476
Park B, Simar L, Zelenyuk V (2017) Nonparametric estimation of dynamic discrete choice models for time series data. Comput Stat Data Anal 108:97–120
Park BU, Simar L, Zelenyuk V (2020) Forecasting of recessions via dynamic probit for time series: replication and extension of Kauppi and Saikkonen (2008). Empir Econ 58(1):379–392
Pellegrino B, Zingales L (2017) Diagnosing the Italian Disease. NBER working papers 23964, National Bureau of Economic Research, Inc
Roberts J (2001) Estimates of the productivity trend using time-varying parameter techniques. BE J Macroecon 1(3):1–32
Simar L, Vanhems A (2012) Probabilistic characterization of directional distances and their robust versions. J Econ 166:342–354
Watson M (1986) Univariate detrending methods with stochastic trends. J Monet Econ 18:49–75
Wheelock DC, Wilson P (1995) Explaining bank failures: deposit insurance, regulation, and efficiency. Rev Econ Stat 77:689–700