Empirical Economics
https://doi.org/10.1007/s00181-024-02634-8

Modelling green knowledge production and environmental policies with semiparametric panel data regression models

Antonio Musolesi[1] · **Davide Golinelli**[1] · **Massimiliano Mazzanti**[1]

Received: 14 January 2023 / Accepted: 12 June 2024

Abstract
This paper introduces a novel semiparametric econometric framework for policy evaluation and estimates a green knowledge production function for a large, 30-year panel dataset of high-income countries. Due to the substantial uncertainty in the data-generating process and the potential presence of nonlinearities and latent common factors, the paper explores semiparametric panel specifications that go beyond interactive fixed effects fully parametric models. The findings suggest that (i) the semiparametric additive specification with individual time trends is the preferred model, (ii) threshold effects and nonlinearities are salient features of the data that parametric specifications fail to capture, and (iii) the impact of environmental policy is noteworthy and exhibits clear heterogeneity when modelled as a nonparametric function of specific knowledge inputs. The evidence reveals a significant nonlinear policy inducement effect stemming from R&D investments.

Keywords Green knowledge generation · Environmental policy · Heterogeneous policy effect · Large panels · Interactive Fixed effects · Spline functions · Model selection

JEL Classification C14 · C23 · O3 · Q5

1 Introduction

A decarbonized, resource- and energy-efficient economy relies heavily on the global dissemination of technological innovations (UNIDO 2018). Key technological drivers over the past three decades include advancements in information and communications technologies, the rise of the Internet of Things, and the development of automation and robotics. The development of the green economy represents a significant shift,

✉ Antonio Musolesi
mslntn@unife.it

1 Department of Economics and Management, University of Ferrara and SEEDS, Ferrara, Italy

expanding the production possibilities frontier for human wants (Griliches 1990). This development encompasses both gradual changes and marked discontinuities, warranting explanations from market and policy perspectives.

Green technological development has also been affected by various environmental policies over the last few decades. Landmark policies include the US Clean Air Act, the 1987 UN Convention on Sustainable Development, the 1992 Rio Convention, the 2015 COP21 in Paris, and the 2019 European Green Deal. Additionally, globalization and global macroeconomic factors, such as economic crises, can significantly influence technological development, with economic crises that may generate market destruction and creation effects.

Figure 1 displays a bivariate plot of green patents and R&D stock, both presented in logarithmic scale, for selected OECD countries over the period 1982–2012. The bivariate plots, along with the corresponding locally estimated scatterplot smoothing (LOESS), unveil intricate and heterogeneous correlation patterns. Most countries showcase nonlinear patterns, with the U.S. exhibiting an exponential-type pattern. Germany presents a concave shape, while Italy and Sweden display highly complex nonlinearities.

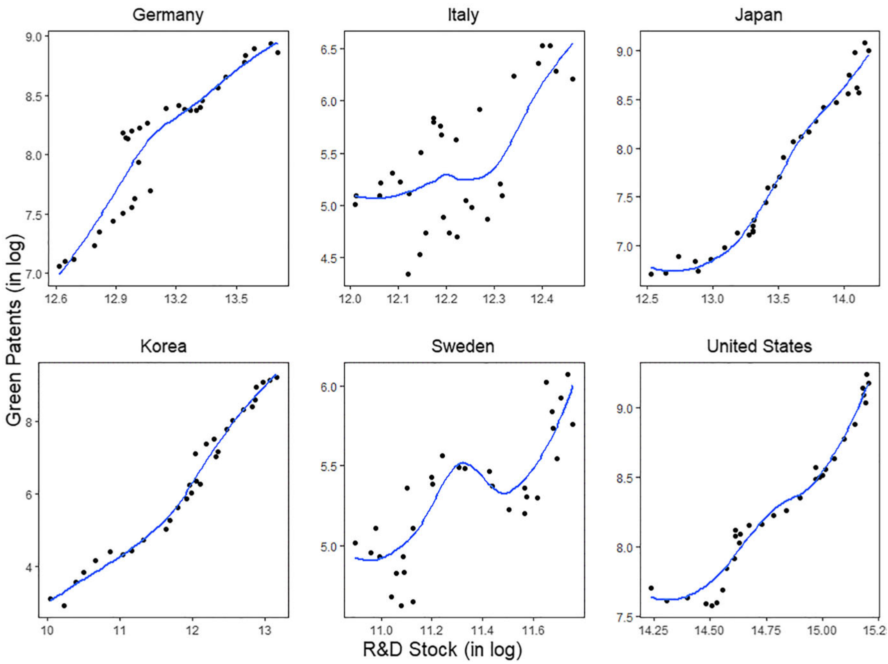

Fig. 1 Bivariate plots. **Legend** Green Patents (y-axis) and R&D Stock (x-axis) for a selected set of OECD countries (1982–2012). These countries comprise Germany, representing the EU technological leader; Sweden and Italy, representing nations from the northern and southern regions of the EU; Japan and Korea as Asian leaders, and the U.S. as the global leader. The relationship is estimated using locally estimated scatterplot smoothing (LOESS)

This paper aims to explain green technological development by capturing the complexity of green knowledge generation, emphasizing nonlinear dynamics, global latent factors, and environmental policies. Specifically, it bridges the knowledge production function (KPF) framework (Griliches 1990) with the study of green inventions and environmental policy. This integration seeks to enhance the analysis of green inventions and policy assessments.

Since Griliches' seminal contribution, the KPF concept has become central in innovation analysis. Micro-level studies have explored firm-level innovation relationships (Mairesse and Monhen, 2010), while regional studies have assessed the impact of regional inputs on local knowledge creation, focusing on spillover effects (Charlot et al. 2015). In the context of green innovations, research has evolved from identifying drivers of eco-innovations to assessing their performance impacts (Kemp 1997; Costantini and Mazzanti 2012). Notably, the literature on policy-induced innovation effects (Popp 2002, 2019; Acemoglu et al. 2016; Aghion et al. 2016) explores the relationship between clean technologies and policies. Studies also examine firm behaviour under policies like the EU emission trading system (Calel 2020).

However, existing research predominantly focuses on firm-level or sectoral data, with notable exceptions like Porter and Stern (2000) adopting a macroeconomic perspective. Additionally, much of the literature relies on (log) linear specifications, with some exceptions using nonparametric approaches (Charlot et al. 2015) to allow for a flexible form and threshold regression models (Nesta et al., 2014) to capture discontinuities in policy effectiveness.

This paper diverges from the existing literature by adopting a macroeconomic and long-term perspective on green technological dynamics, potentially yielding new insights into the historical role of policies in stimulating green technologies. Additionally, it employs flexible semi-nonparametric estimators to model the green knowledge production functions (GKPF), addressing functional form bias and correlated unobservable factors bias, which is associated with standard parametric fixed effects approaches, through the approach proposed by Su and Jin (2012). This method allows for a nonparametric relationship between dependent and explanatory variables while considering latent common factors parametrically through interactive fixed effects.

A specific focus of this paper is to assess the effect of environmental policy on green knowledge. To overcome the common practice to estimate a homogenous effect across units we considered a binary-by-continuous interaction model with a spline regression framework (Ruppert et al. 2003; Cardot and Musolesi 2020). With respect to other approaches like matching estimators, control functions, and machine learning methods, it is a simple and flexible way to account for heterogeneous policy effect, which has been largely employed also within an instrumental variables framework (Antonioli et al. 2020; Cerulli 2014; Das 2005; Wooldridge 2003, 2010).

This modelling choice is motivated by the presence of absorptive capacity and the potential non-neutral impact of environmental policy, suggesting that the transmission of policy to technological performance varies based on country-specific investments in knowledge-driving factors.

In summary, the primary research objectives of this paper are: (i) analysing the functional form of green knowledge generation with a focus on model uncertainty, nonlinearities, and unobserved common factors; (ii) assessing the heterogeneous policy effects underpinning the GKPF.

A key result is that environmental policies have significantly driven green inventions since the early 1980s. Another important result is that the proposed semiparametric framework offers evidence of relevant nonlinearities, threshold effects and complementarities. It indeed suggests the need for a critical mass of knowledge inputs to sustain technological development through a direct effect of these inputs on knowledge but also via an induced policy effect occurring through R&D activity.

The paper is organized as follows. Section2 presents the data. Section3 introduces the econometric framework, focusing specifically on issues related to specification, identification, and estimation. Sections 4, 5, and 6 discuss the main estimation results, focusing on homogeneous policy effects, heterogeneous policy effects, and providing robustness checks, respectively. Section 7 concludes and offers suggestions for further research. Supplementary information is provided in the appendices.

2 Data

2.1 The data

We build a new dataset that is a balanced panel dataset covering the period 1982–2012 for 19 OECD countries. Notably, the discrete policy indicator derived from raw OECD data has a longer time dimension than the OECD Environmental Policy Stringency Index (EPSI; Albrizio et al. 2017), which excludes the 1980s. This motivated our decision to construct a new indicator from OECD sources. Opting for a categorical, specifically binary, variable is also grounded in econometric rationale within a nonparametric framework. It facilitates easier interpretation and mitigates the curse of dimensionality problem associated with using a continuous policy variable when considering interactions, as in the present case.

2.1.1 Dependent variable

As far as the dependent variable is concerned, data on green patents (GK) are collected from the OECD Stat databases. We consider patents that fall under the 'selected environment-related technologies' category as defined by the OECD (IPC: ENV_TECH) and that were granted by the USPTO (United States Patent & Trademark Office), and we calculate the number of country-wise patents according to the inventor or inventors' country or countries of residence. Patents of an agent belonging to country i but submitted to country j are accounted for as i-related patents. For each patent, we have information on the patent family, year of filing and geographical location of the inventors.[1]

[1] Fractional counts are used to avoid double counting of the same inventions across different geographical areas. This means that if a patent family is developed by more than one inventor, we weight that patent family according to the geographical areas of the inventors. The patent family also captures the 'quality' of patents at the macroeconomic level.

2.1.2 Knowledge inputs

As far as the explanatory variables are concerned, we use the knowledge input set following Charlot et al. (2015), among others. Research and development (RD) and human capital (HK) are both included as input factors. Specifically, we use gross domestic expenditure on research and development (GERD) flow values, collected from the OECD-Stats database, using total data as a source of funds. The data are in 2010 dollars – constant prices and PP. Missing values are filled in using a similar method to that of Coe et al. (2009), and then we calculate GERD stock values using the perpetual inventory method as in Coe and Helpman (1995), assuming the depreciation rate to be 0.05. The HK is also computed as a stock and is collected from the Penn World Table version 9.0 (Feenstra et al., 2015).[2]

To account for knowledge spillovers, consistent with Griliches (1992) and with a relevant corpus of literature on international R&D spillovers, we consider foreign RD and foreign HK (WRD and WHK, respectively). We use geographic proximity (between capitals) as a channel of technology diffusion because of its consistency with theory (Keller 2002) and for exogeneity reasons, as it may be considered an exogenous proxy for some endogenous measures of socioeconomic, institutional, cultural, or linguistic similarities that might enhance the diffusion of technology. We use an exponential decay function that is very common in the spatial econometric literature concerning international technology diffusion and adopt the same expression as Ertur and Musolesi (2017), where the foreign capital and human capital stocks, $\mathrm{WRD}_{\mathrm{it}}$ and $\mathrm{WHK}_{\mathrm{it}}$, are defined as the weighted arithmetic mean of $\mathrm{RD}_{\mathrm{it}}$ and $\mathrm{HK}_{\mathrm{it}}$, respectively, for $j \neq i$, using an exponential distance decay function, $w_{ij} = \exp\left(-d_{ij}\right)$ such that:

$$\mathrm{WRD}_{it} = \sum_{j \neq i} \exp\left(-d_{ij}\right)\mathrm{RD}_{jt}$$
$$\mathrm{WHK}_{it} = \sum_{j \neq i} \exp\left(-d_{ij}\right)\mathrm{HK}_{jt}$$

where d_{ij} is the spherical distance between two capitals.

2.2 Binary policy variable

A binary environmental policy variable, denoted as EP_{it}, is considered. A binary index is chosen for several reasons: (i) the difficulty, as underscored by Johnstone et al. (2010), in establishing continuous variables across diverse policy types and countries; (ii) it covers a longer time span compared to commonly adopted continuous measures such as *EPSI*, which does not cover the eighties; (iii) it is less susceptible to the curse of dimensionality problem when estimating a model with interactions, compared to multiple-category or continuous variables.

Specifically, starting from OECD raw data sources, policy indices are obtained (see also Nesta et al. 2014). The information refers to three main domains: air pollution,

[2] The stock measure of human capital is grounded on earlier works on the return to schooling and then it is assumed a log-(piecewise) linear relationship between human capital stock and the average number of years of schooling.

climate change, and energy efficiency. We focus on the domain of air pollution regulations (Berman and Bui 2001) given the relatively longer history of air pollution policies which also justify the rationale for opting for a more extensive time span. Air pollution regulation strategy has always represented a pillar of environmental policy since the US clean air act in the 70's. It has represented a complement to climate policy as well (Markandya et al. 2018).

Within the air pollution domain, we first consider a multiple-category policy indicator. These categories are: (i) deposit refund schemes, (ii) fees, (iii) tax rates of environmentally related taxes, (iv) tradable permits, (v) voluntary approaches, and (vi) environmentally motivated subsidies. These categories cover a broad spectrum that accommodate the multidimensionality of policy efforts over market and non-market instruments (Baumol and Oates 1988; OECD 2016).

Based on the country implementation of regulations over those six categories, the multiple-category policy indicator, which is labelled as EP_{base}, specifically assumes a value ranging from 0, potentially up to six, $EP_{\text{base}} = j$, with $j = 0,1,..,6$, where j indicates that at least one policy in the air pollution domain is introduced in j categories (further details in Appendix C).

From that indicator, we derive the binary environmental policy variable, EP_{it}, that takes the value 1 if one or more policies are introduced at least in two categories, and 0 otherwise,

$$EP_{it} = \begin{cases} 1 \text{ if } \; 2 \leq EP_{\text{base}} \leq 6 \\ 0 \text{ otherwise} \end{cases}$$

The adoption of the above presented binary policy variable is motivated from both an economic and a statistical perspective (see also Lanoie et al. 2011, p. 824). From an economic viewpoint, it can be argued that when two or more policy categories are implemented, the intensity of the policy action increases, and the complementarity of different actions is potentially introduced. The estimated coefficient may capture the joint effect (Costantini et al. 2017; Braathen 2007), which may include trade-offs, distortions, and complementarities (Requate and Unold 2003).

From a statistical perspective, such a choice avoids the estimation problems (the curse of dimensionality) arising when categories have few observations, as the binary *EP* variable presents two homogeneous categories in terms of size, while the basis six-category variable has categories with a very small number of observations (supplementary information supporting the choice of the binary indicator are available upon request). Moreover, the estimation problems arising with cells having few observations are, à priori, particularly severe when allowing heterogeneous effects within a similar framework to that of Cardot and Musolesi (2020).

We now turn our attention to providing some descriptive statistics for the *EP* variable. In total, 45% of the observations are treated units. Importantly, the prevalence of treated units has changed over time. Initially, very few observations were treated (0% in 1982, 5% in 1983, etc.), but this proportion increased significantly, reaching almost 80% by 2012. In the analysis of cross-country variation, Nordic nations consistently emerge as proponents of the most stringent environmental policies (Botta and Kozluk

2014). However, heterogeneity prevails. Notably, Germany exhibits a lagged trajectory, instituting policies from the mid-2000s. In contrast, the United States, an early adopter since the early 1980s, consistently surpasses the average. Meanwhile, Italy, Korea, and Sweden demonstrate similar patterns since the early 1990s. Among the selected countries, Japan is the only nation that did not introduce policies in alternative categories during the considered period, remaining the only untreated statistical unit. Detailed results can be found in Appendix C.

3 Semiparametric modelling of green knowledge production and environmental policies

3.1 Econometric modelling

To model the GKPF, we exploit recent advances in nonparametric panel data regression with interactive fixed effects (Su and Jin 2012; Gioldasis et al. 2023) and consider the following model:

$$
\begin{aligned}
GK_{it} &= \beta EP_{it} + g(\mathrm{RD}_{it}, \mathrm{HK}_{it}, \mathrm{WRD}_{it}, \mathrm{WHK}_{it}) + \nu_{it} \\
\nu_{it} &= \gamma_i' f_t^* + \varepsilon_{it}
\end{aligned}
\tag{1}
$$

where GK_{it} measures green patents, *g(.)* is a real unknown function, and RD_{it} and HK_{it} refer to the two main factors behind inventions, namely, R&D and human capital stocks.[3] WRD_{it} and WHK_{it} are introduced to consider spillover effects that may arise from foreign countries. Finally, EP_{it} is the binary indicator of policy intensity described in the previous section.

The response and the continuous explanatory variables are expressed in logarithmic values. This facilitates the economic interpretation in terms of elasticity. This also makes the Gaussian assumption more plausible and allows for a straightforward comparison with parametric models, which are often expressed in log–log form.

The errors ν_{it} in (1) have a multifactor structure, where $f_t^{*'} = [1, f_t\prime]$ with f_t being a vector of unobservable common factors with heterogeneous factor loadings γ_i', and ε_{it} is the idiosyncratic error term. Interestingly, in such an interactive fixed effects framework, the explanatory variables are allowed to be correlated with both the individual component and the common factors (Su and Jin 2012; Gioldasis et al. 2023). This dependence allows addressing endogeneity problems due to endogenous selection and, under fairly weak assumptions, β identifies the average treatment effect. Moreover, the idiosyncratic error terms are allowed to be serially correlated as well as weakly cross-sectionally dependent (see Gioldasis et al. 2023 for additional details).

Model (1) contains many others as special cases and allows for a direct comparison with standard parametric models. With the aim of exploring in more detail the effect

[3] It is worth citing Griliches (1990, p.1674) here, who stresses that ‘patents tend to be taken out relatively early in the life of a research project’. The empirical literature has noted that in the green realm, patenting activities arose quite early in the first environmental policy phases of the 1980s and 1990s (Jaffe et al. 1995; Jaffe and Palmer, 1997).

of EP_{it}, we also consider in this paper a variant of Model (1) that allows for a heterogeneous treatment effect via a possible interaction between the policy and the knowledge inputs, that is:

$$\begin{aligned} GK_{it} &= \beta \mathrm{EP}_{it} + g_{EP_{it}}(RD_{it}, HK_{it}, WRD_{it}, \mathrm{WHK}_{it}) + \nu_{it} \\ \nu_{it} &= \gamma_i' f_t^* + \varepsilon_{it} \end{aligned} \tag{2}$$

where there are two distinct nonparametric functions (one for each level of EP_{it}).

3.2 Identification issues

For identification purposes, the following condition is considered:

$$E\big(g\big(X_{it}^*\big)\big) = 0,$$

where $X_{it}^* = [\mathrm{RD}_{it}, \mathrm{HK}_{it}, \mathrm{WRD}_{it}, \mathrm{WHK}_{it}]'$. Such a constraint is deemed necessary in this framework due to the inclusion of the intercept term in the model (refer to Su and Jin 2012, for detailed explanations). Moreover, when contemplating an additive structure for $g(.)$, as further explored in this paper, it becomes imperative to impose an identifiability constraint to the smooth terms even in standard cross-sectional specifications, as outlined by Wood (2017). Specifically, a centering constraint, akin to the one described above, is deemed optimal for Gaussian responses (Stringer 2023).

To identify β in model (1), let first define $X_i^* = \left[X_{i1}^*, \ldots, X_{iT}^*\right]$ and note that:

$$\begin{aligned} E\big(GK_{it}|EP_{it} = 1, X_i^* = x\big) &= \beta + g(x) + E\big(v_{it}|EP_{it} = 1, X_i^* = x\big), \\ E\big(GK_{it}|EP_{it} = 0, X_i^* = x\big) &= g(x) + E\big(v_{it}|EP_{it} = 0, X_i^* = x\big) \end{aligned}$$

Taking the difference between the two conditional expectations yields

$$\begin{aligned} &E\big(GK_{it}|EP_{it} = 1, X_i^* = x\big) - E\big(GK_{it}|EP_{it} = 0, X_i^* = x\big) \\ &\quad = \beta + E\big(v_{it}|\mathrm{EP}_{it} = 1, X_i^* = x\big) - E\big(v_{it}|EP_{it} = 0, X_i^* = x\big) \end{aligned}$$

Hence, we obtain that

$$\begin{aligned} \beta = &E\left(GK_{it}|EP_{it} = 1, X_i^* = x\right) - E\left(GK_{it}|EP_{it} = 0, X_i^* = x\right) \\ &- E\left(v_{it}|EP_{it} = 1, X_i^* = x\right) + E\left(v_{it}|EP_{it} = 0, X_i^* = x\right) \end{aligned}$$

This implies that a condition for identification of β in model (1) is that:

$$E\big(v_{it}|EP_{it}, X_i^* = x\big) = E\big(v_{it}|X_i^* = x\big)$$

The same condition applies to model (2).

3.3 Estimation method and spline modelling

3.3.1 Estimation procedure

Econometric theory underlying the estimation has been developed by Su and Jin (2012). The estimation procedure consists of the following steps.

Following Pesaran (2006), the unobservable common factors f_t in (1), are proxied by the cross-sectional averages $\overline{Z}_t = N^{-1}\sum_{i=1}^{N} Z_{it}$, where $Z_{it} = \left[GK_{it}, X_{it}^{*'}\right]'$.

The nonparametric component of the model, $g(.)$, is approximated using sieves and specifically splines, as they typically provide better approximations (see, for example, Hansen, 2014). Sieve approximation proceeds as follows. First, an infinite sequence of known basis functions that can approximate any square-integrable function of x very well should be chosen. Let K denote the order of approximation (Refer to Su and Jin 2012 and Gioldasis et al. 2023 for details), under fairly weak conditions, the unknown function $g(.)$ in (1) can be approximated very well by a linear combination of the first K elements of the chosen basis, i.e., $g\left(X_{it}^{*}\right) \approx \theta' \pi^K\left(X_{it}^{*}\right)$, with $\pi^K\left(X_{it}^{*}\right) = \left[\pi_1\left(X_{it}^{*}\right), \pi_2\left(X_{it}^{*}\right), ..., \pi_K\left(X_{it}^{*}\right),\right]'$. With the above approximations for f_t and $g(.)$, to estimate (1), the following auxiliary regression equation is considered,

$$GK_{it} = \beta EP_{it} + \theta' \pi^K(\mathrm{RD}_{it}, \mathrm{HK}_{it}, \mathrm{WRD}_{it}, \mathrm{WHK}_{it}) + \gamma_i' \overline{Z}_t^{*} + u_{it},$$

where $\overline{Z}_t^{*'} = \left[1, \overline{Z}_t'\right]$, the term u_{it} includes ε_{it} and two approximation errors, one for the unobservable factors f_t and the other for the unknown function $g(.)$. The former approximation error is given by $f_t - \overline{Z}_t$ while the latter is the difference between the unknown function and the sieve approximation, i.e. $g\left(X_{it}^{*}\right) - \theta' \pi^K\left(X_{it}^{*}\right)$. In that auxiliary regression framework, the unknown smooth function is estimated as $\hat{g}\left(X_{it}^{*}\right) = \hat{\theta}' p^K\left(X_{it}^{*}\right)$.

3.3.2 Computation

We follow Gioldasis et al. (2023) and employ penalized regression splines (PRS), as they combine the features of both regression splines, which use fewer knots than data points but do not penalize roughness, and smoothing splines, which control the smoothness of the fit through a penalty term but use all data points as knots. PRS have proven to be useful empirically in many respects (see, for example, Ruppert et al. 2003), and Gioldasis et al. (2023) who showed that they perform better than regression (unpenalized) splines. We specifically employ thin plate regression splines (TPRS), which were introduced by Wood (2003) and are optimal low rank eigen-approximation to thin plate splines. Thin plate splines are somehow ideal smoothers but are not computationally attractive because their computation requires the estimation of as many parameters as the number of data points. TPRS avoids the problem of knot placement that usually complicates modelling with splines and more generally has some optimality properties, while it is also computationally efficient. TPRS can be used for both univariate and multivariate functions. In the latter case, however, one main feature of TPRS is the isotropy of the wiggliness penalty, i.e., wiggliness in all directions is treated equally, with the fitted spline entirely invariant to rotation of the coordinate system. Isotropy is considered desirable when the explanatory variables have the same units. When this is not the case, isotropy can be avoided by considering a tensor product basis (Wood 2006), which is constructed by assigning TPRS as the basis for the marginal smooth function of each covariate and then creating their Kronecker product. The tensor product smooths are invariant to the linear rescaling of covariates, and for this reason, they are appropriate when the arguments of a smooth function have different units. Finally, the smoothing parameter is selected by the restricted maximum likelihood (REML) estimation, which, relative to other approaches, is less likely to develop multiple minima or to undersmooth at finite sample sizes.[4] A detailed technical discussion can be found in Gioldasis et al. (2023).

3.4 Methodological contribution

The adopted econometric model specifically addresses five main issues: (*i*) *functional form and nonlinearities; (ii) latent common factors and cross-sectional dependence; (iii) model uncertainty; (iv) endogeneity; and (v) heterogeneous treatment effects.* We briefly discuss below why these issues are relevant.

Functional form and nonlinearities: Although a parametric log–log specification is customary in the literature on the KPF, the precise functional form cannot be straightforwardly defined on a theoretical basis, and alternative functional forms could better approximate the unknown DGP. This relevant issue was recognized, even at the firm level, in an early work by Griliches (1990, p. 303), who pointed out the *"noisiness in this relation".* Moreover, it can be expected that a critical mass of R&D or human capital is necessary to make such inputs truly effective. These considerations suggest that estimating a nonparametric relation between knowledge and its main inputs could be important to avoid a *functional form bias*.

Latent common factors and cross-sectional dependence: The literature on the KPF has generally adopted a two-way (individual and common time) fixed effects approach to handle unobserved heterogeneity. Charlot et al. (2015) use a *random trend* specification (see Wooldridge 2005). Both approaches are special cases of the factor model considered in (1), which introduces cross-sectional dependence as a result of a finite number of unobservable common factors that may have different effects on knowledge creation across countries.

Model uncertainty: We recognize the existence of high uncertainty surrounding the true DGP. In general, there is a bias-efficiency trade-off when comparing parsimonious to complex models. Considering flexible models is appealing but may come at the price of unfeasible or extremely inefficient estimates. For these reasons, we perform model selection by comparing some alternative models. This is done by exploiting recent smooth model selection (Wood et al. 2016; Wood 2020) and placebo tests (Heckman

[4] Computations are performed within the R environment, and in particular, the semiparametric specifications are estimated by exploiting the mgcv package.

and Hotz 1989; Cardot and Musolesi 2020) to assess whether selection bias is properly accounted for.

Endogeneity of the environmental policy variable and of the knowledge inputs: A relevant issue to address is the endogeneity of both the knowledge inputs and the policy variable. As far as the endogeneity of the knowledge inputs is concerned, the literature motivated such a problem due to the existence of omitted variables, which are likely to be correlated with R&D and/or human capital, while the existence of reverse causality was discarded, as implied by the knowledge production function framework proposed by Griliches (1990, p. 1671) and suggesting a unidirectional link between patents and R&D.[5]

Similarly, as far as the *EP* dummy variable is concerned, the existence of selection on both observables and unobservable (Heckman and Hotz 1989), suggests that correlation between the probability to be 'treated', and that both observable and unobservable variables is likely to be present. Since we perform model selection, we more formally discuss this issue in the next section for the selected specification.

Heterogeneous treatment effects: As far as specification (2) is concerned, we relax the possibly too restrictive assumption of homogeneous treatment effects.

In summary, from a methodological perspective, this paper aims to provide an incremental contribution with respect to previous works and in particular to the work by Charlot et al. (2015), as (i) we here consider a model with a multifactor error structure, which is more general than the random trend specification, by exploiting more recent works on nonparametric panel data models with multifactor error (Su and Jin 2012; Gioldasis et al. 2023); (ii) we perform a sound model selection by exploiting both recent advances on smooth model selection (Wood et al. 2016; Wood 2020) and placebo tests (Heckman and Hotz 1989; Cardot and Musolesi 2020) to assess whether selection bias is properly accounted for; and (iii) we focus here on the effect of a binary policy, and in particular, we allow for heterogeneous effects, via a nonparametric interaction between the discrete policy variable and the knowledge inputs, which appears, to the best of our knowledge, as totally new in the literature.

4 Homogeneous policy effect

4.1 Model selection

As far as Model (1) is concerned, we perform model selection. When performing model selection, it is typically assumed that there exists a finite-dimensional "true model". Within this view, many model selection methods are defined in terms of an appropriate information criterion that aims to find the best model for the unknown true DGP among the set of models under consideration. Information criteria are based on the idea of balancing fit with complexity in a logic of bias-variance trade-off (see Claeskens and Hjort, 2008; Konishi and Kitagawa, 2008).

[5] Another related issue is the possible existence of a link between current patents and past R&D spending. Employing an R&D variable that is built as a stock allows for such a link, while drastically reducing the curse of the dimensionality problem; this is a common practice within the KPF literature.

While both AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) can be used interchangeably for selecting among nested or non-nested models (Claeskens and Hjort, 2008; Konishi and Kitagawa, 2008), they exhibit different properties. The AIC, derived as an asymptotic approximate estimate of the Kullback–Leibler information discrepancy, is not strongly consistent, although it is efficient. In contrast, the BIC is strongly consistent, making it a preferred choice. Studies (e.g., Koheler and Murphee, 1988; Granger and Jeon, 2004) have suggested the superiority of BIC over AIC in handling macroeconomic data and applications. For these reasons, we adopt the BIC, as recently proposed by Wood et al. (2016) and Wood (2020), which overcomes some limitations of previously proposed smooth model selection criteria. Specifically, we compare models that incorporate alternative specifications for g(.) and for the unobserved time effects.

Moreover, to detect misspecified models, we also consider 'preprogram' tests along the lines depicted by Heckman and Hotz (1989) within a program evaluation framework. As underlined by Heckman et al. (1999), the fact that various methods produce different inferences would suggest that selection bias is important and that some of the adopted estimators are likely to be misspecified. Preprogram tests are based on the idea that a valid estimator would correctly adjust for differences in pre-program outcomes between future participants and nonparticipants; otherwise, the estimator is rejected.

The results, which are detailed in Appendix A, indicate that the preferred model presents additive smooth terms for *g(.)* and individual time trends (random trend) to represent the latent common factors. We thus adopt the following *semiparametric random trend model*:

$$GK_{it} = c_i + \beta EP_{it} + g_1(\mathrm{RD}_{it}) + g_2(\mathrm{HK}_{it}) + g_3(\mathrm{WRD}_{it}) + g_4(\mathrm{WHK}_{it}) + \gamma_i t + \varepsilon_{it} \tag{3}$$

Three main remarks are in order. Firstly, it is essential to note that the preferred model permits an intermediate-high level of flexibility. This is achieved by allowing for smooth additive effects of the regressors along with individual linear trends. The excessively simple models, such as parametric or two-way fixed effects models, are likely to suffer from severe estimation bias and are rejected. On the other hand, overly complex specifications like multifactor error models are the most inefficient and are also rejected (see also Baltagi et al. 2002, 2003).

Second, it is worth noting that allowing for smooth additive terms provides a more credible identification of the policy effect than that provided by parametric models. This is because the underlying identification condition (see, for a detailed and more formal discussion, and Lechner, 2010a, 2015) is a nonparametric one and imposing linearity or a specific, possibly wrong, functional form may invalidate the estimate.

4.2 Endogeneity issues

The random trend model identifies the parameters and functions of interest under a *conditional* strict exogeneity assumption. Indeed, given (3) the strict exogeneity assumption conditional to the unobserved effects c_i and γ_i can be formulated as

$$E\left(\varepsilon_{it}|EP_{it}, X_{i1}^*, ..., X_{iT}^*, c_i, \gamma_i\right) = 0, \tag{4}$$

for $t = 1,...,T$. Clearly, such a conditional strict exogeneity assumption provides a more credible identification than the one arising from the individual (one-way) fixed effects specification, i.e.,

$$E\left(\varepsilon_{it}|EP_{it}, X_{i1}^*, ..., X_{iT}^*, c_i\right) = 0, \tag{5}$$

and this is because while condition (5) does not restrict the correlation between c_i and $\left[\mathrm{EP}_{it}, X_{i1}^*, ..., X_{iT}^*\right]$, condition (4) does not restrict the correlation between $\left[c_i, \gamma_i\right]$ and $\left[EP_{it}, X_{i1}^*, ..., X_{iT}^*\right]$, thus providing a more credible identification. Allowing such an arbitrary correlation handles s*election on unobservables* (see Heckman and Hotz 1989; Wooldridge 2005; and Charlot et al. 2015, for detailed discussions). With a specific focus on treatment effects, it is important to note that the random trend model, unlike the two-way model (which is equivalent to the difference-in-differences estimator), does not assume that treated and untreated units would have followed a common trend in the absence of the policy. This assumption, even when conditioned on some explanatory variables, is often too restrictive (Abadie 2005).

Moreover, as far as endogenous selection is concerned, *EP* implementation can also be related to some observable variables that are also related to GK generation, such as RD and HK, whose inclusion in the model solves the *selection on observables* problem (see Cardot and Musolesi 2020, among many others, for a formal discussion within the potential outcome framework).

For all these reasons, *interactive fixed effect models*, such as the random trend model or the factor model, are typically considered a sound and robust framework for treatment effect estimation under selection on both observables and unobservables and have been employed for program evaluation in many empirical frameworks that are typically characterized by a severe endogenous selection process (Fujiki and Hsiao 2015; Heckman and Hotz 1989; Papke 1994; Gobillon and Magnac 2016; Brown et al. 2006).

Interestingly, placebo tests (Appendix A.2) indicate that the selected specification provides a good alignment, suggesting that the inclusion of country-specific trends and allowing for additive smooth effects of the knowledge inputs is crucial for properly addressing the problem of selection bias.

4.3 Semiparametric estimation of the GKPF: nonlinearities, threshold effects and average policy effect

The results concerning the additive nonparametric components of Model (3) are presented in Fig. 2. Here, we mainly focus our attention on the functional relation between green patents and their main inputs, eventually highlighting possible nonlinearities and threshold effects. The three plots depict the estimated univariate smooth functions for the inputs that are found to be statistically significant. Note that the smooths are subject to a sum-to-zero identifiability constraint, implying that the estimated smooths must be constrained to have a zero mean; confidence intervals are obtained following Wood

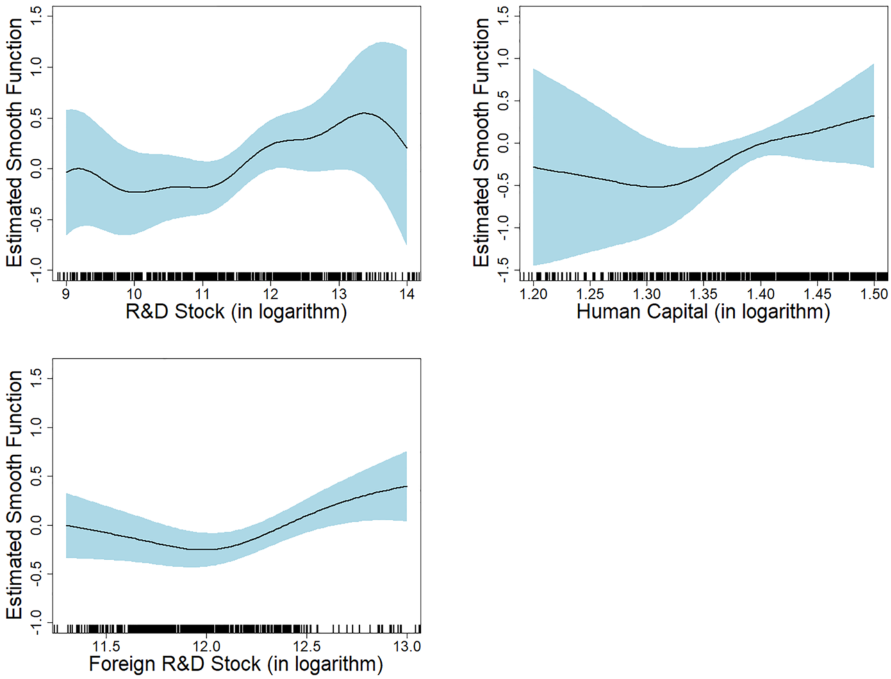

Fig. 2 R&D, human capital, and RD spillover effects on green inventions: estimated univariate smooth functions. **Legend** The continuous line shows the estimate of the smooth; the blue area represents 95% confidence bands, which are obtained following Wood (2006). The black vertical line at bottom of the plots represents the frequency of the observations over the range of the explanatory variable. The smooths are subject to sum-to-zero identifiability constraint implying that the estimated smooths must be constrained to have zero mean

(2006). Moreover, statistical significance is assessed by computing the p-values for the smooth terms using Wald test statistics as suggested by Wood (2012).[6]

The estimated smooths (except WHK) appear to be highly significant, showing extremely low p-values on the Wald tests. Moreover, using an approximate ANOVA test procedure (Wood 2017), linearity is always rejected for all explanatory variables except WHK. WHK instead presents a positive linear effect, which is not statistically significant (p value $= 0.27$). The nonsignificance of the WHK is consistent with the literature focusing on international technology diffusion, which stresses the role of spillovers arising from R&D (Ertur and Musolesi 2017).

Figure 2 shows that as expected, the estimated smooth functions are highly nonlinear, with relevant threshold effects. Indeed, for all three significant variables (RD, HK, and WRD), a critical mass is necessary to ensure an effective impact on green patenting. This result is consistent with some related literature focusing on nongreen knowledge

[6] These are p-values associated with the Wald test of the hypothesis that the whole function equals zero. Low p-values indicate a low likelihood that the splines of the function are jointly zero.

creation and/or different levels of aggregation (Charlot et al. 2015).[7] Evidence for RD and HK attests that policy targets on these inputs are justified to support innovation and growth, as the number of countries making substantial knowledge investments (e.g., 3% of GDP or higher) enhances the spillover effect. Countries that have structurally invested in R&D over the last decades, such as Japan, the USA, Germany, South Korea, and China, as well as Denmark and Sweden on a relatively smaller scale, present large shares of top inventions in climate-related technologies. The ascending phase of the climate change patent trend was around the mid-1990s, influenced by new policy expectations (Dechezleprêtre et al. 2009). Nevertheless, investing in R&D matters if substantial thresholds are structurally surpassed. It is worth noting that R&D/human capital and policy factors could be entangled. The fact that the mid-1990s witnessed a turning point in R&D-intense countries that proactively responded to new policies is a signal of an interconnection: to induce inventions, policies need a pre-existent, dense R&D environment. Section5.1 below focuses on the analysis of interactive policy effects to shed further light. Spillovers are the other key component of technological inducement (Verdolini and Galeotti, 2011).

As for the effect of the policy, which is assumed here to be constant both across countries and over time, the parameter β, which also identifies the average treatment effect, is estimated to be 0.09 and is almost significant at the 5% level (p value $= 0.055$). This result thus indicates a positive effect, albeit rather small in magnitude.[8] Provided that the dependent variable is logarithmic, the estimated policy effect represents the percentage change in the predicted GK when $EP = 1$ versus when $EP = 0$, holding the other factors constant.[9]

What we do substantially find through this analysis is that over the time span that embraces the second US Clean Air Act, the Rio Convention, the Kyoto Protocol and the EU 2020 climate and energy package, all key examples of policy steps of international relevance, a high level of policy intensity brought about specific effects on green inventions. This outcome enriches the macroeconomic evidence on policy effectiveness that, among others, Johnstone et al. (2010, 2012) provide, first over a similarly long-time span (1978–2003), but focusing on renewable energy policies, and then over a more restricted time span and using opinion survey-based policy indicators. McKitrick (2007), who takes a long-run perspective and specifically focuses on air pollution trends, finds that more than oil price effects, the 1970 Clean Air Act, a milestone policy to abate air pollution, was a relevant structural factor that induced accelerated abatement technologies.

[7] It is worth noting that the domain of the variables has been appropriately reduced to the regions where the effects are significant. Indeed, in the regions of the domain of the variables where data are sparse, large confidence interval bands are present, since it is not possible to precisely estimate the functions of interest. These regions in which the plots cannot be easily interpreted correspond to low levels of HK and to very low levels of RD and WRD.

[8] Jaffe and Palmer (1997) find little evidence of effects on innovative outputs using regulation compliance costs as a proxy for policies in a panel of US manufacturing industries. Brunnermier and Cohen (2003) find mixed evidence on the effects of environmental regulations on environmental innovations proxied by patents for a panel of US industries.

[9] More formally, this percentage difference is given by $100\left[\exp\left(\hat{\beta}\right) - 1\right]$

In summary, the estimation of a *semiparametric random trend model* indicates that the long-run evolution of green inventions was affected by environmental policy and mostly by the continuous variables R&D, human capital, and foreign R&D, which show significant nonlinear monotonic patterns and relevant threshold effects. Only foreign human capital has, at least in the present time span, no influence on green inventions.

5 Heterogeneous policy effect

5.1 A binary-by-continuous interaction framework

Common practices in program evaluation consist of focusing attention on the mean effect of a policy or imposing a homogeneous effect across units and overtime, as we did in the previous section. Adopting Model (3) was extremely useful because it allowed us to (i) conduct a direct comparison with parametric models (see Appendix B) and (ii) focus attention on the functional relation between patents and knowledge inputs, which is one of the goals of this paper.

In this section, we extend the previous analysis by searching for possible heterogeneous policy effects. Specifically, we consider a binary-by-continuous interaction model within a spline regression framework (Ruppert et al. 2003; Cardot and Musolesi 2020), as in Eq. (2). This allows us to obtain two distinct nonparametric functions (one for each level of the policy) for each continuous explanatory variable, representing a simple and flexible way to account for heterogeneous policy effects.

Given the identification condition that was previously discussed in Sect. 3.2, i.e.

$$E\left(v_{it}|EP_{it}, X_i^* = x\right) = E\left(v_{it}|X_i^* = x\right),$$

the policy effect can be defined as follows:

$$\beta_{it} = E\left[GK_{it}|X_{it}^*, EP_{it} = 1\right] - E\left[GK_{it}|X_{it}^*, \mathrm{EP}_{it} = 0\right]$$

In the binary-by-continuous interaction framework of Eq. (2), the policy effect β_{it} is a function of the continuous explanatory variables. If additivity is imposed, it can be expressed, given the vector of continuous covariates X_{it}^*, with the following specification:

$$\beta_{it} = \beta + \sum_{j=1}^{4} m_j\left(X_{jit}^*\right) \tag{6}$$

where $m_j\left(X_{jit}^*\right) = g_{j,EP_{it}=1}\left(X_{jit}^*\right) - g_{j,EP_{it}=0}\left(X_{jit}^*\right)$ are unknown smooth functions satisfying the centering identifiability constraint that was previously discussed, i.e., $E\left[m_j\left(X_{jit}^*\right)\right] = 0$, $j = 1, ..., 4$.

Notably, in such a framework, the policy effect β_{it} is decomposed as the sum of a constant β plus a nonparametric function of the explanatory variables. This nonparametric function is represented by the difference between two smooth functions, each one corresponding to a different level of the binary policy variable.

As noted in Ruppert et al. (2003, see p. 225 for a thorough discussion), the fitted curves do not depend on which constraint is adopted, but the interpretation of the parameters does. Importantly, under the centering constraint $E\left[m_j\left(X^*_{jit}\right)\right] = 0,\ j = 1, \ldots, 4,\ \beta$ represents the average treatment effect, and the function $m_j(.)$ indicates how the mean effect of the policy varies with the knowledge input j.

The rationale behind such a modelization lies in the presence of absorptive capacity and nonneutral or even localized (to some specific input domain) inducement effects of environmental policy. Indeed, we can expect that environmental policy is transmitted to technological performance (green inventions, in this case) with a significance and strength that depend on the country-specific investments in knowledge-driving factors, which become types of 'innovation endowments' (see also Jaffe et al. 1995, 2002).

The economic system's absorptive capacity is the ability to recognize the value of new external 'information', a policy in this case, assimilate it, and apply it to the ends of invention. Absorptive capacity can be regarded as an important factor for general competitive advantages, as stressed in Dosi (1982).

Ex ante, it can be expected that the larger knowledge investments are, the stronger the possible role of policy in inducing new inventions. The higher the combination of any R&D/human capital sources, the stronger is the sociotechnical system's capacity to absorb the effect of the policy, translating this into inventions. Moreover, this may happen with possibly complex nonlinear shapes. To the best of our knowledge, the possible existence of (possibly complex) nonlinear induced effects has never been addressed in previous studies.

5.2 Selecting nonparametric interaction terms

To precisely define the model to be estimated, we adopted the following procedure:

i) We first adopt a backward selection procedure to select the variables to be considered to estimate the model. The initial model contained all four knowledge inputs interacted with the binary policy variable, i.e.

$$GK_{it} = c_i + \beta EP_{it} + g_{1EP_{it}}(RD_{it}) + g_{2EP_{it}}(\mathrm{HK}_{it}) + g_{3EP_{it}}(WRD_{it}) + g_{4EP_{it}}(WHK_{it}) + \gamma_i t + \varepsilon_{it}.$$

The nonsignificant binary-by-continuous smooth terms $g_{jEP_{it}}(.)$ were replaced with univariate terms, $g_j(.)$, and then they were finally removed if still nonsignificant. This procedure led us to retain only two significant binary-by-continuous smooth terms, i.e., domestic and foreign R&D, while domestic and foreign human capital entered the model without interacting with the binary policy:

$$GK_{it} = c_i + \beta EP_{it} + g_{1EP_{it}}(RD_{it}) + g_2(HK_{it}) + g_{3EP_{it}}(WRD_{it}) + g_4(WHK_{it}) + \gamma_i t + \varepsilon_{it},$$

so that

$$\beta_{it} = \beta + m_1(\mathrm{RD}_{it}) + m_2(W\mathrm{RD}_{it}).$$

ii) The model above presents only two binary-by-continuous interaction terms, $g_{1EP_{it}}(\mathrm{RD}_{it})$ and $g_{3EP_{it}}(\mathrm{WRD}_{it})$, and such a small dimension allows considering a bivariate smooth function instead of additive functions. For this reason, we employed an approximate ANOVA test procedure (see Wood 2017) to compare the additive specification

$$GK_{it} = c_i + \beta EP_{it} + g_{1EP_{it}}(RD_{it}) + g_2(HK_{it}) + g_{3EP_{it}}(WRD_{it}) + g_4(WHK_{it}) + \gamma_i t + \varepsilon_{it}$$

with the following one, which introduces a bivariate smooth function $GK_{it} = c_i + \beta EP_{it} + g_{1EP_{it}}(RD_{it}, WRD_{it}) + g_2(HK_{it}) + g_4(WHK_{it}) + \gamma_i t + \varepsilon_{it}$.

The ANOVA test indicates that the additive structure is strongly rejected in favour of the model above that exploits a bivariate regression function for RD and WRD, so that the policy effect can be expressed as

$$\beta_{it} = \beta + m(RD_{it}, WRD_{it}).$$

5.3 Estimation: unveiling nonlinear policy inducement effects

According to the model selection procedure discussed above, we adopt the following specification:

$$GK_{it} = c_i + \beta EP_{it} + g_{1EP_{it}}(RD_{it}, WRD_{it}) + g_2(HK_{it}) + g_4(WHK_{it}) + \gamma_i t + \varepsilon_{it} \tag{7}$$

The results indicate that the estimated average effect of the policy is positive and sizeable, with $\hat{\beta} = 0.15$, and is significant in statistical terms (p value $= 0.011$). Thus, compared to the results in the previous section, allowing for a heterogeneous policy effect produces a higher estimated average effect and higher significance level.

For the estimated function $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it})$, Fig. 3 draws the contour plot of the estimated bivariate function $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it})$. Note that the contours indicate the *varying* component of the policy effect, i.e., the effect of the policy that varies non-parametrically with RD and WRD, while the total estimated effect of the policy is given by:

$$\hat{\beta}_{it} = 0.15 + \hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it}).$$

Overall, over the whole domain of RD and WRD (Fig. 3, top), there is evidence that the effect of the policy nonlinearly and monotonically increases with both domestic and foreign R&D, pointing again to the joint relevance of internal and external knowledge

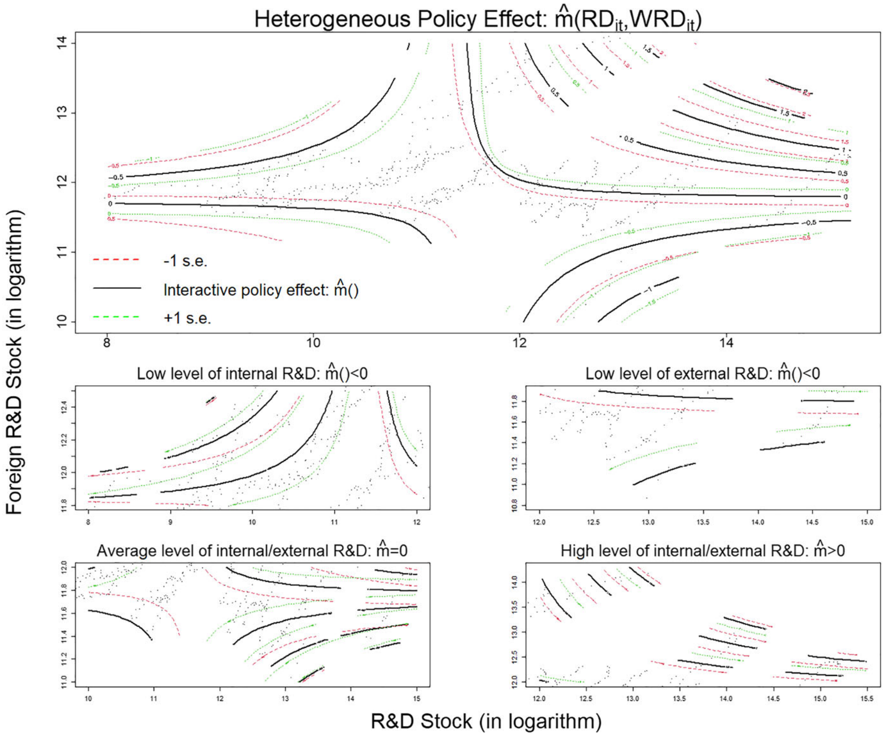

Fig. 3 Heterogeneous policy effect: contour plot of $\hat{m}(RD_{it}, WRD_{it})$.**Legend**. Heterogeneous Policy Effect (Interactive Policy Effect). Contour plot for the entire domain (on the top), specific domain's ranges (bottom panels)

for the support of technological development through their joint effect with the policy. We can identify two macro regions, which are described below.

The first macro region corresponds to very low levels of internal/foreign R&D, where $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it}) < 0$ (Fig. 3, middle), and at the very extreme $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it}) \simeq -0.5$, so that the total effect $\hat{\beta}_{it}$ is negative over part of the analysed green knowledge generation domain. From an economic viewpoint, this points to a kind of coordination failure or insufficient investment in the innovation drivers. At very low levels of these knowledge inputs, the costs outweigh the benefits. The absorptive capacity of the system, represented by the investments in knowledge, is insufficient to provide an effective framework where invention can arise through the inducement effect of the policy. From a statistical point of view, it can also be observed that for low levels of internal/foreign R&D, the data are sparse, and the estimates lack precision.

In the second, and most relevant, macro region, we observe $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it}) \geq 0$ (Fig. 3, bottom right). This region corresponds to most of the domain of RD and WRD. Specifically, it can be expected that environmental policies' effects on innovation require a critical amount of core investments in knowledge to exert the dynamic efficiency effects that theory predicts (Milliman and Prince 1989; Requate and Unhold, 2003), and indeed, for average levels of internal/foreign R&D, we find that $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it}) \simeq 0$ so that $\hat{\beta}_{it} \simeq 0.15$ (Fig. 3, bottom left).

Then, by increasing the level of RD/WRD, we find that threshold effects again matter and are likely connected to some complementarity with the underlying innovation function (Charlot et al. 2015). In fact, for high levels of internal/foreign R&D, the estimated function $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it})$ turns positive and monotonically increases with both variables, up to a maximum at which $\hat{m}(\mathrm{RD}_{it}, \mathrm{WRD}_{it}) \simeq 1.5$, so that in that region of the domain, $0.15 \leq \hat{\beta}_{it} \leq 1.65$.

Two main highlights of econometric and economic relevance arise. First, the results are clear-cut in showing how the effect of the policy nonlinearly and monotonically increases with both domestic and foreign R&D, suggesting, more specifically, that a critical mass of these inputs is necessary to make the policy effective. The proposed semiparametric model reveals heterogeneous policy effects, which operate through R&D layers, signalling that the effects of environmental policy on knowledge are significantly mediated by country investments in R&D. Domestic and foreign R&D act as knowledge absorptive capacity factors and enhance the effectiveness of the environmental policy: the denser the market and institutional environment is in R&D, the more effective the air pollution policy is in driving green patents. Second, this evidence suggests that green patenting dynamics conceptually connect to two main relevant dimensions of a GKPF: (i) the complementarity of various invention drivers, here domestic and foreign R&D, and (ii) the crucial role of R&D spillovers, mediated by geographic distance, in directly contributing to green knowledge generation and allowing the policy to become effective. Further studies may consider alternative transmission channels such as trade, technological proximity, language, or genetic distance.

The significance of R&D spillovers compared to HK effects is possibly due to the greater historical emphasis of technological innovation flows in the international setting (e.g., environmental international agreements), especially after the Kyoto Protocol on climate change was signed in 1998, a key milestone, and its ratification across most countries in the years that followed. It is also worth noting that over the 1998–2012 Kyoto protocol period (with 2008–2012 as the first commitment period), schemes such as the Clean Development Mechanisms and Joint implementation boosted various project-based technological exchanges aimed at abating emissions (Costantini and Sforna 2014). The focus of the international community was on the (more mobile) role of technologies to mitigate emissions along the development convergence path. Although it is true that technological flows can also incorporate human capital as 'embodied' in the exchanges of technologies (Marin and Mazzanti 2021), one could argue that overall, human capital investments are generally more localized in geographical terms with respect to both the governance level of education and training expenses, partially more specific (training is general and specific in kind) and their effects largely on state/national labour markets (Popp et al. 2020). On the 'features' of R&D and training as knowledge-based capital, the OECD (2013) highlights that 'While R&D exhibits properties of partial excludability and nonrivalry, other forms of KBC may have a smaller impact on growth (and have also been less studied). For instance, firm-specific human capital, and much of brand equity, are highly excludable and rivalrous" (p.22).

6 Sensitivity analysis: assessing the effect of OECD EPS Index

In this section, we check the sensitivity of our results with respect to the consideration of a continuous environmental policy measure, i.e. the *EPSI*, which has been often employed in the literature. In doing so, the sample must be restricted to the period 1990–2012.

Such variable is obtained from the OECD database; it spans from 0, indicating the least stringent, to 6, representing the highest degree of stringency, and displays a bimodal distribution, with most observations concentrated between the two peaks around 1 and 3. This phenomenon likely mirrors the temporal progression of environmental policies, with one mode indicative of early adopters and the other mode representing countries that embraced such policies at a subsequent stage in the timeline.

Our attention in the estimation is directed towards the preferred semiparametric random trend model. Given the continuous nature of EPSI, all explanatory variables (all in logs) are incorporated into the model through a nonparametric univariate smooth term, i.e.:

$$\begin{aligned} GK_{it} = c_i + g_1\left(EPSI_{it}\right) + g_2\left(RD_{it}\right) + g_3\left(\mathrm{HK}_{it}\right) \\ + g_4\left(WRD_{it}\right) + g_5\left(WHK_{it}\right) + \gamma_i t + \varepsilon_{it} \end{aligned}$$

The results suggest that **EPSI** is statistically significant at a 10% level (p value $= 0.08$). The estimated smooth function for the policy variable **EPSI** (depicted in Fig. 4a) exhibits a moderately non-linear pattern. Specifically, up to a certain threshold, corresponding approximately to the fourth quantile of **EPSI**, the estimated smooth is relatively flat, and afterward, it shows a slight positive pattern. This finding aligns with the results obtained when considering a binary policy variable, where a small positive and marginally significant effect was observed (p value $= 0.055$).

Moreover, the first derivative of the estimated smooth function (Fig. 4b) gives information about the elasticity that may be directly compared to the estimated parameter of *EP* in (3), i.e. $\hat{\beta}$. It is worth noting that the estimated mean elasticity of **EPSI** is 0.1 (Fig. 4c). Notably, this value is very close to the estimated effect that was obtained when employing the binary policy variable *EP* ($\hat{\beta} = 0.09$; see Sect. 4.3).

Finally, akin to the approach in Sect. 5.3, we endeavored to estimate models involving the interaction of the continuous policy variable **EPSI** with knowledge inputs. Regrettably, the results were notably influenced by sparse data. The estimated smooth function exhibits substantial gaps, encompassing at least one-third of the domain (additional details available upon request), without yielding any supplementary insights compared to what we gleaned from the binary variable.

In summary, using the binary policy variable *EP* not only enables consideration of a larger time span but also facilitates easier interpretation and mitigates the curse of dimensionality problem associated with using a continuous policy variable when considering interactions, which is one of the main focuses of the paper.

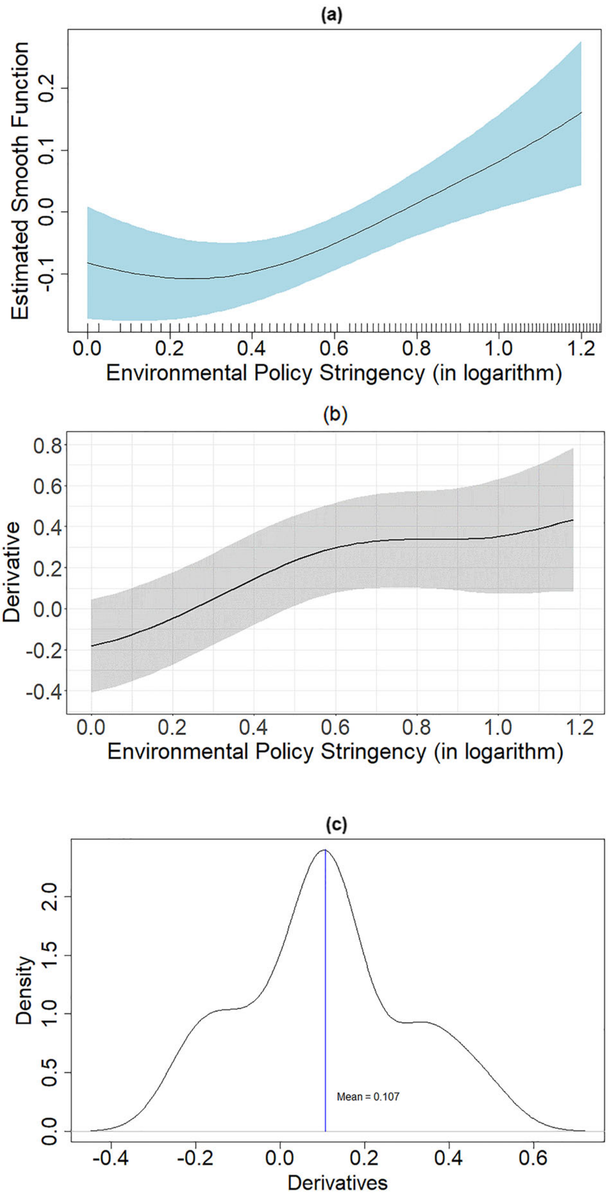

Fig. 4 Estimated smooth function of **EPSI** (**a**); Estimated first derivative (**b**) and density plot (**c**). **Legend** (**a**) The continuous line shows the estimate of the smooth; the blue area represent 95% confidence bands. The black vertical lines at bottom of the upper plot represents the frequency of the observations over the range of the explanatory variable. The smooth is subject to sum-to-zero identifiability constraint implying that the estimated smooths must be constrained to have zero mean. (**b**) The estimated first derivative of the estimated smooth function of **EPSI** via finite differences and 95% confidence bands. (**c**) The kernel density plot of the estimated first derivative, with bandwidth selected using biased cross-validation

7 Conclusion

The paper takes a macroeconometric long-run perspective to examine green knowledge production functions. The main methodological issues the investigation addresses are the possible *functional form bias* and *correlated unobservable factor* bias that may arise when adopting standard parametric fixed effects approaches. Consequently, the work has considered semiparametric panel data specifications with interactive fixed effects. The modelling framework that the paper developed also aims at enhancing the understanding of long-run innovation phenomena and the setting of flexible and sound policy assessment tools. In this regard, we consider a flexible specification that allows us to relax the hypothesis of homogeneous policy effects, considering effects that nonparametrically interact with some knowledge inputs, such as research activity.

A first relevant result is that the effects of R&D, human capital and foreign R&D are characterized by relevant nonlinearities and thresholds. The specifications that emerge from the model selection reinforce the idea that nonlinearities and thresholds are relevant components of knowledge creation in the specific case of green innovation activities.

Another important result is that environmental policies have significantly driven green inventions since the early 1980s. The effect is significant from both economic and statistical points of view. It is specifically found that to fully unveil the significance and strength of environmental policy, allowing for heterogeneous/interactive policy effects is necessary. In fact, the consideration of a model in which the effect of policy is expanded as a nonparametric function of some knowledge inputs indicates that the estimated induced policy effect, which is mediated by domestic and foreign R&D, is on average positive and highly significant. It is also found that such an effect nonlinearly increases with both domestic and foreign R&D and that a critical mass of both variables is necessary to make the policy effective. Threshold effects are again relevant, as at extremely low levels of R&D activity, environmental policy may even produce a negative effect on green patenting. The emergence of complementarity between innovation policies and R&D activity connects a methodological issue (the heterogeneous policy effect) with a real-world policy issue (the necessary R&D investments for green technological development). The economic meaning is that those countries with excessively low levels of domestic and foreign R&D are not providing a favourable setting for substantial policy inducement effects to appear.

Further analyses could consider extending the coverage to other policy domains, introducing nonbinary policy index structures, comparing the KPF for green and non-green inventions, and considering other spillover transmission channels.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00181-024-02634-8.

Funding Project funded under the National Recovery and Resilience Plan (NRRP), Mission 04 Component 2 Investment 1.5 – NextGenerationEU, Call for tender n. 3277 dated 30/12/2021. Award Number: 0001052 dated 23/06/2022.

We thank scholars who provided comments on previous versions of the paper, that has been presented in various seminars and conferences since 2019 (Econometric Models of Climate Change, International Panel data conference IPDC, UEA East Anglia, Wien University, UCLY Lyon Business school, European Economic association conference, EAERE and IAERE conferences, AISRE conference, SIE Italian association conference).

References

Abadie A (2005) Semiparametric difference-in-differences estimators. Rev Econ Stud 72(1):1–19

Acemoglu D, Akcigit U, Hanley D, Kerr W (2016) Transition to clean technology. J Polit Econ 124(1):52–10

Aghion P, Dechezleprêtre A, Hemous D, Martin R, Van Reenen J (2016) Carbon Taxes, path dependency, and directed technical change: evidence from the auto industry. J Political Econ 124:1

Albrizio S, Kozluk T, Zipperer V (2017) Environmental policies and productivity growth: evidence across industries and firms. J Environ Econ Manag 81:209–226

Antonioli D, Gioldasis G, Musolesi A (2020) Estimating a non-neutral production function. Oxf Econ Pap 73:856

Baltagi BH, Bresson G, Pirotte A (2002) Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: some empirical evidence from US electricity and natural-gas consumption. Econ Lett 76(3):375–382

Baltagi BH, Bresson G, Griffin JM, Pirotte A (2003) Homogeneous, heterogeneous or shrinkage estimators? Some Empirical Evidence from French Regional Gasoline Consumption. Empir Econ 28(4):795–811

Baumol W, Oates W (1988) The Theory of Environmental Policy, 2nd edn. Cambridge University Press, Cambridge

Berman E, Bui L (2001) Environmental regulation and productivity: evidence from oil refineries. Rev Econ Stat 83(3):498–510

Botta E, Kozluk T, (2014) Measuring environmental policy stringency in OECD countries: a composite index approach, No 1177, OECD Economics Department Working Papers, OECD Publishing

Braathen NA (2007) Instrument mixes for environmental policy: how many stones should be used to kill a bird? Int Rev Environ Resour Econ 1(2):185–235

Brown JD, Earle JS, Telegdy A (2006) The productivity effects of privatization: longitudinal estimates from Hungary Romania, Russia, and Ukraine. J Political Econ 114(1):61–99

Brunnermeier SB, Cohen MA (2003) Determinants of environmental innovation in US manufacturing industries. J Environ Econ Manag 45(2):278–293

Calel R (2020) Adopt or Innovate: understanding technological responses to cap-and-trade. Am Econ J Econ Pol 12:3

Cardot H, Musolesi A (2020) Modeling temporal treatment effects with zero inflated semi-parametric regression models: the case of local development policies in France. Econ Rev 39(2):135–157

Cerulli G (2014) ivtreatreg: A command for fitting binary treatment models with heterogeneous response to treatment and unobservable selection. The Stata J 14(3):453–480

Charlot S, Crescenzi R, Musolesi A (2015) (2015). Econometric modelling of the regional knowledge production function in Europe. J Econ Geogr 15(6):1227–1259

Coe DH, Helpman E (1995) International R & D spillovers. Eur Econ Rev 39(5):859–887

Coe DT, Helpman E, Hoffmaister AW (2009) International R&D spillovers and institutions. Eur Econ Rev 53(7):723–741

Costantini V, Mazzanti M (2012) On the green and innovative side of trade competitiveness? The impact of environmental policies and innovation on EU exports. Res Policy 41:1

Costantini V, Sforna G (2014) Do bilateral trade relationships influence the distribution of CDM projects? Clim Policy 14:5

Costantini V, Crespi F, Palma A (2017) Characterizing the policy mix and its impact on eco-innovation: a patent analysis of energy-efficient technologies. Res Policy 46(4):799–819

Das M (2005) Instrumental variables estimators of nonparametric models with discrete endogenous regressors. J Economet 124(2):335–361

Dechezleprêtre A, Glachant M, Hašcic I, Johnstone N., Ménière Y. (2009). Invention and transfer of climate change-mitigation technologies on a global scale: a study drawing on patent data, FEEM working paper 82, FEEM, Milan

Dosi G (1982) Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. Res Policy 11(3):147

Ertur C, Musolesi A (2017) Weak and strong cross-sectional dependence: a panel data analysis of international technology diffusion. J Appl Econ 32(3):477–503

Feenstra RC, Inklaar R, Timmer MP (2015) The next generation of the Penn World Table. Am Econom Rev 105(10): 3150–3182

Fujiki H, Hsiao C (2015) Disentangling the effects of multiple treatments—Measuring the net economic impact of the 1995 great Hanshin-Awaji earthquake. J Econ 186(1):66–73

Gioldasis G, Musolesi A, Simioni M (2023) Interactive R&D spillovers: an estimation strategy based on forecasting-driven model selection. Int J Forecast 39(1):144–169

Gobillon L, Magnac T (2016) Regional policy evaluation: Interactive fixed effects and synthetic controls. Rev Econ Stat 98(3):535–551

Granger C, Jeon Y (2004)] Forecasting performance of information criteria with many macro series. J Appl Statis 31(10): 1227–1240

Greaker M, Popp D, (2022) Environmental economics, regulation, and innovation, NBER Working paper 30415, NBER

Griliches Z (1990) Patent statistics as economic indicators: a survey. J Econ Lit Am Econ Assoc 28(4):1661–1707

Griliches Z (1992) The search for R&D spillovers. Scand J Econ 94:29–47

Hansen BE (2014) Nonparametric sieve regression: Least squares, averaging least squares, and cross-validation. In JS Racine, L Su, & A Ullah (Eds.) Handbook of applied nonparametric and semiparametric econometrics and statistics (pp. 215–248). Oxford University Press

Heckman JJ, Hotz VJ (1989) Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training. J Am Stat Assoc 84(408):862–874

Heckman JJ, Smith JA (1999) The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies. The Econ J 109(457):313–348

Jaffe AB, Palmer K (1997) Environmental regulation and innovation: a panel data study. Rev Econ Stat 79(4):610–619

Jaffe AB, Peterson SR, Portney PR, Stavins RN (1995) Environmental regulation and the competitiveness of US manufacturing: what does the evidence tell us? J Econ Lit 33(1):132–163

Jaffe AB, Newell RG, Stavins RN (2002) Environmental policy and technological change. Environ Resour Econ 22(1):41–70

Johnstone N, Haščič I, Popp D (2010) Renewable energy policies and technological innovation: evidence based on patent counts. Environ Resour Econ 45:133

Johnstone N, Haščič I, Poirier J, Hemar M, Michel C (2012) Environmental policy stringency and technological innovation: evidence from survey data and patent counts. Appl Econ 44(17):2157–2170

Keller W (2002) Geographic localization of international technology diffusion. Am Econ Rev 92(1):120–142

Kemp, R., (1997). *Environmental Policy and Technical Change,* Books.

Koehler AB, Murphree ES (1986) A comparison of the AIC and BIC on empirical data. In *6th international symposium on forecasting*, Paris.

Konishi S, Kitagawa G (2008) *Information criteria and statistical modeling*. Springer Science & Business Media

Lanoie P, Laurent-Lucchetti J, Johnstone N, Ambec S (2011) Environmental policy innovation and performance: new insights on the porter hypothesis. J Econ Manag Strategy Wiley Blackwell 20(3):803–842

Lechner M (2010) The relation of different concepts of causality used in time series and microeconometrics. Economet Rev 30(1):109–127

Lechner M (2015) Treatment effects and panel data. Oxford University Press, Oxford, pp 257–284
Mairesse J, Mohnen P (2010) Using innovation surveys for econometric analysis. In Handbook of the Economics of Innovation (Vol. 2, pp. 1129–1155). North-Holland
Marin G, Mazzanti M (2021) Structural change and the environment: unbundling the contributions of structural change to sustainable consumption and production, *new perspectives on structural change: causes and consequences of structural change in the global economy,* chapter 26. Oxford University Press, Oxford
Markandya A, Sampedro J, Smith SJ, Van Dingenen R, Pizarro-Irizar C, Arto I, González-Eguino M (2018) Health co-benefits from air pollution and mitigation costs of the Paris Agreement: a modelling study. The Lancet Planetary Health 2(3):e126–e133
McKitrick R (2007) Why did US air pollution decline after 1970? Empirical Economics 33(3):491–513
Milliman SR, Prince R (1989) Firm incentives to promote technological change in pollution control. J Environ Econ Manag 17(3):247–265
Nesta L, Vona F, Nicolli F (2014) Environmental policies, competition and innovation in renewable energy. J Environ Econ Manag 67:396
OECD, (2013) New sources of growth: knowledge-based capital. key analyses and policy conclusions, OECD, Paris
OECD, (2016) How stringent are environmental policies?, OECD report
Papke LE (1994) Tax policy and urban development: evidence from the Indiana enterprise zone program. J Public Econ 54(1):37–49
Pesaran MH (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74(4):967–1012
Popp D (2002) Induced innovation and energy prices. Am Econ Rev 92(1):160
Popp D, Vona F, Marin G,Chen Z (2020) The employment impact of green fiscal push: evidence from the american recovery act, NBER Working paper 27231, NBER.
Popp D, (2019) Environmental policy and innovation. A decade of research, NBER working paper 25631
Requate T, Unold W (2003) Environmental policy incentives to adopt advanced abatement technology: Will the true ranking please stand up? Eur Econ Rev 47(1):125–146
Ruppert D, Wand MP, Carroll RJ (2003) Semiparametric regression. Cambridge University Press, Cambridge
Stringer A (2023) Identifiability constraints in generalized additive models. Can J Stat 52:461
Su L, Jin S (2012) Sieve estimation of panel data models with cross section dependence. J Econ 169(1):34–47
UNIDO, (2018). Industrial development report. Demand for manufacturing: driving inclusive and sustainable industrial development, UNIDO, Wien
Verdolini E, Galeotti M (2011) At home and abroad: an empirical analysis of innovation and diffusion in energy technologies. J Environ Econ Manag 61(2):119–134
Wooldridge JM (2003) Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model. Econom Lett 79:185–191
Wooldridge J (2010) Econometric analysis of cross section and panel data. Economet Anal Cross Sect Panel Data. MIT Press 16, 17
Wood SN (2003) Thin plate regression splines. J R Stat Society: Series B (statistical Methodology) 65(1):95–114
Wood SN (2006) On confidence intervals for generalized additive models based on penalized regression splines. Aust N Z J Stat 48:445–464
Wood SN (2013) On p-values for smooth components of an extended generalized additive model. Biometrika 100(1):221–228
Wood SN (2017) Generalized additive models: an introduction with R. CRC Press, Boca Raton
Wood SN, Pya N, Säfken B (2016) Smoothing parameter and model selection for general smooth models. J Am Stat Assoc 111(516):1548–1563
Wood SN, (2020) Inference and computation with generalized additive models and their extensions, TEST, pp. 1–33
Wooldridge JM (2005) Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. Rev Econ Stat 87(2):385–390