Heterogeneous Returns to Scale of Wind Farms in Northern Europe

Giacomo Benini, Maria Carvalho,** Ludovic Gaudard,*** Patrick Jochem,* and Kaveh Madani*****

ABSTRACT

The present paper tries to identify the optimal size of a wind farm using North European data. An empirical analysis of 61 sites constructed between 2004 and 2014 suggests that economies-of-scale are highly heterogeneous across on-shore and off-shore projects. A Varying Coefficient Model captures such diversity by making the impact of the farm site on the amount of its potential capacity a non-linear function of the number of installed turbines. The resulting scale elasticities suggest that small on-shore farms have a bigger per-turbines output than off-shore ones, while the opposite is true for big projects.

Keywords: Wind, Returns to Scale, Cross-Sectional Heterogeneity, Semiparametric Methods, Varying Coefficient Models

https://doi.org/10.5547/01956574.40.SI1.gben

1. INTRODUCTION

In order to achieve the Paris Agreement's goals, reducing greenhouse emissions from the power sector has become a priority of the global energy agenda (Stern, 2007). Among different renewable technologies, wind seems to be the most competitive (IEA et al., 2015). Therefore, it might become the primary option to achieve ambitious climate policy targets (Lantz et al., 2012).

The industry learning processes and economies-of-scale are the main factors which allowed wind to become competitive over the years (Junginger et al., 2005a). The link between the gains in efficiency and the investment's effort has been studied in detail (Ibenholt, 2002; Neij et al., 2004; Junginger et al., 2005b; Neij, 2008). Less attention has been given to the relation between the average costs and the number of installations. Among the few notable exceptions, (Berry, 2009a) highlighted the existence of two types of returns. The first ones are the economies-of-scale of a single turbine. Here, the per-unit cost declines till three Mega Watts (MW) are installed, then it increases (Wiser and Bolinger, 2011). The second ones are the economies-of-scale of an entire wind farm.

For nearly a century, it has been a widespread conjecture that large-scale power generation installations can deliver lower cost electricity. This idea is indeed right for fossil fuel and nuclear power stations where: 1) larger volume components add more usable space than the related materials costs; and 2) the fixed costs associated with the generation of power tend to be extremely high.

* Chair of Energy Economics, Karlsruhe Institute of Technology (KIT). Corresponding author: Giacomo Benini. E-mail: giagi.benini@gmail.com.

** Grantham Research Institute on Climate Change & Environment, London School of Economics and Political Science (LSE).

*** Precourt Energy Efficiency Center, Management Science and Engineering, Stanford University.

**** Centre for Environmental Policy, Imperial College London.

However, the coexistence of two contrasting effects makes the analogy inappropriate for wind. On the one hand, taller turbines can support bigger rotors, which generate more electricity. On the other, each turbine generates a wake effect directly proportional to the size of the rotor's arm, preventing the installation of tall turbines close to one another (González-Longatt et al., 2012; Kim et al., 2012). Furthermore, off-shore farms tend to produce more electricity than on-shore ones, as they often use bigger turbines which usually are more productive than the on land ones. However, they need to be connected with the grid increasing the capital expenditures of this type of projects (Wüstemeyer et al., 2015).

Traditionally, studies which try to explore the economies-of-scale of an entire wind farm, either concentrate on the existence of different economies-of-scale for different cost components (Blanco, 2009), or on the returns of one particular type of installation (Morgan et al., 2003; Dismukes and Upton Jr, 2015). The present paper aims to overcome these limitations introducing a production function able to calculate the amount of per-turbine installed capacity both for on-shore and off-shore farms using a unique econometric specification. The key to analyse different projects using a single regression model is to allow for high degrees of cross-farm heterogeneity.

Varying Coefficient Models (VCM) are a generalization of traditional parametric regressions specifically designed to include diversity across economic units (Hastie and Tibshirani, 1993). In the case of the wind energy sector, the number of turbines is the factor which changes the impact of the site on the amount of installed capacity. This means that a VCM can quantify how a delta in the number of installed turbines impacts the competitiveness of off-shore and on-shore platforms shading light on the trade-off between the size of the rotors and the total number of installed turbines.

Different regions have instituted market support policies for wind. While several jurisdictions support on-shore wind, fewer sponsor off-shore projects. According to the Bloomberg New Energy Finance's database, among the few areas which sustain both technologies, the North of Europe is the one with the biggest number of installations. The widespread use of on-shore and off-shore farms combined with a rather homogeneous market design makes this region an interesting case study to analyse the economies-of-scale of the wind industry.

2. ECONOMETRIC ANALYSIS

2.1 Returns to Scale in Wind Farms

A production plan is a bundle of m non-negative inputs $(X_1, X_2, ..., X_m)$ which delivers a single output Y. In the case of a wind farm, the factors of production include labour, capital and turbines' technology. The outcome is the amount of installed capacity. The resulting technological set (TS),

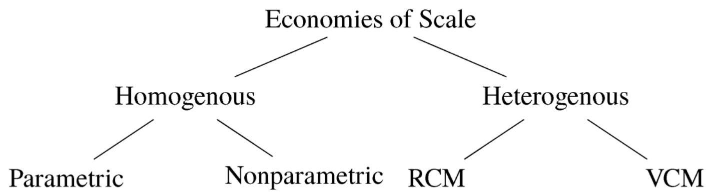
$$\mathrm{TS} = \{(X_1, ..., X_m; Y) \,|\, Y \leq v(X_1, ..., X_m) \quad \text{and} \quad X_1 \geq 0, ..., X_m \geq 0\}, \tag{1}$$

links the quantity of installed MWs to the different inputs via the production function $v(.)$ (Mas-Colell et al., 1995). In order to empirically estimate the shape of $v(.)$, we analyse the Bloomberg New Energy Finance's database. This collection of farm-level data provides information about 1307 investment projects realized during the decade 2004–2014 in 61 countries. For each of them, we know the installed capacity, the number of turbines, the site (on-shore or off-shore), the height of the wind turbines, the time required to complete the installation, the country of destination, the manufacturer, the status of the signed contract and the final buyer. The dataset contains only one region where both on-shore and off-shore projects are widespread, and the set of public incentives

Table 1: Number of Wind Farms per Country

	Belgium	Germany	Netherlands	Denmark	Norway	Sweden	Ireland	UK
Off-Shore	4	19	2	3	0	3	0	16
On-Shore	3	35	5	1	5	35	13	65
Total	7	54	7	4	5	38	13	81

Figure 1: Econometric Taxonomy of the TS Estimation.



is comparable, the North of Europe. The high amounts of wind in the North Sea, the Baltic Sea, the North European Plain and the Scandinavian Peninsula, combined with the willingness of the European Union to promote renewable energy markets, makes this part of the world an interesting area where to compare the performances of off-shore and on-shore turbines. Table 1 reports the selected subset of the data.

Once the area of interest is selected the study proceeds with an increasingly more sophisticated econometric analysis of the data. First, economies-of-scale are assumed to be homogeneous across the two technologies. Under this paradigm two enquiries are made. The first one is a fully parametric regression which connects the estimated coefficients to the economic theory. The second one is a fully nonparametric model which, while failing to deliver a unique elasticity of the inputs' demand, captures the non-linear nature of the relation $Y \leftrightarrow (X_1,\ldots,X_m)$.

Then a more elaborate analysis of the data, combined with the statistical difficulty to manage a homogeneous interaction between a factor and a continuous variable in a nonparametric context, suggests introducing heterogeneity between off-shore and on-shore farms. A first attempt to model the difference between cross-sectional units, which are displaced in different sites, is done using a semiparametric Random Coefficient Model (RCM). Once the RCM assumptions are discarded by the specificities of the wind industry, a semiparametric Varying Coefficient Model (VCM) is introduced. This last model can quantify the impact of the number of turbines on the installed capacity, shedding light on how the sites' characteristics determine the competitiveness of adding an additional turbine, see Figure 1.

2.2 Intermediate Homogeneous Returns to Scale

Assuming that all firms are profit-maximizing units, which operate at $Y = v(X_1,\ldots,X_m)$, it is possible to overcome the lack of data about the canonical factors of production postulating that all the unobserved inputs deliver an observed 'intermediate' good: the number of turbines. This 'final' factor of production can be seen as the only capital good of the installed capacity. Hence, the production function $v(\cdot)$ becomes an univariate relation between the nominal power Y and the total number of turbines X,

$$Y_i = v(X_i) \qquad i = 1,\ldots,n, \tag{2}$$

where n is the dimension of the sample (Aigner and Chu, 1968). The first step needed to estimate the shape of $v(.)$ is to replace the theoretical unobserved variables with the collected sample. When $(Y = y, X = x)$, it is possible to redefine both theoretical and observed variables as stochastic objects and make the acceptance-rejection region function of the stochastic nature of the sample. This second step allows us to transform equation (2) into $y_i = v(x_i, \varepsilon_i)$, where ε_i is a random error with finite variance (Haavelmo, 1943, 1944). Equations like $y_i = v(x_i, \varepsilon_i)$ present a lot of statistical shortcomings (Imbens, 2007; Imbens and Newey, 2009). Therefore, it is common to impose $v(x_i, \varepsilon_i) = f(x_i) + \varepsilon_i$ and work with the simplified relation $y_i = f(x_i) + \varepsilon_i$. Once the error term has been transformed into an additive disturbance it is possible to assume an *a priori* functional form for $f(.)$ and describe the relation $Y \leftrightarrow X$ using a finite number of parameters β^T, such that $y_i = f(x_i; \beta^T) + \varepsilon_i$. Given that the aim of the study is to estimate the elasticity of the installed capacity to the number of turbines, and not the ratio between different inputs, a natural choice for $f(x_i; \beta^T)$ is the Cobb-Douglas production function $Y = AX^{\beta_1}$, where A is the technological level. The empirical match of this last specification, $y_i = Ax_i^{\beta_1} e_i$, with $E[e_i \mid x_i] = 1$ and $V[e_i] < \infty$, can be log-transformed into

$$\log y_i = \beta_0 + \beta_1 \log x_i + \varepsilon_i \quad \text{where} \quad \beta_0 = \log A \quad \text{and} \quad \log e_i = \varepsilon_i \,. \tag{3}$$

Equation (3) secures the concept of *output elasticity* into a single unit-free coefficient. There are increasing returns to scale if $\hat{\beta}_1 > 1$, decreasing if $\hat{\beta}_1 < 1$ and constant if $\hat{\beta}_1 = 1$.

Despite being elegant and concise, the previous equation ignores any form of decoupling. The introduction of an interaction term between the number of turbines and the site of the farm,

$$\log y_i = \beta_0 + \beta_1 \log x_i + \beta_2 \log x_i sec_i + \alpha^T D^j_{\alpha,i} + \varepsilon_i, \tag{4}$$

where $sec_i = [\text{Off} - \text{Shore}, \text{On} - \text{Shore}]$, changes drastically the interpretation of all the coefficients. In equation (3), β_1 is the unique effect of $\log x$ on $\log y$, while in equation (4) the effect of the number of turbines on the amount of produced electricity is allowed to be different for on-shore and off-shore projects. This means that the returns to scale, $\dfrac{\partial \log y}{\partial \log x} = \beta_1 + \beta_2 sec_i$, become function of the type of installation. In particular, off-shore returns are equal to β_1, while on-shore ones are $\beta_1 + \beta_2$. Equation (4) further expands on (3) controlling for a set of country-specific effects, $\alpha = [\alpha_1, \cdots, \alpha_8]^T$, which are identified using the dummies $D^j_{\alpha,i} = 0$ for $i \neq j$ and $D^j_{\alpha,i} = 1$ for $i = j$. Assuming the conditional moment restriction $E[\varepsilon_i \mid x_i, z_i] = 0$, where $z_i = [sec_i, \alpha^T]$, it is possible to estimate equation (4) using Ordinary Least Squares (OLS), see Table 2.

Regressions which use $\log x$ as the only explanatory variable convey increasing returns to scale. For example in equation (3), an increase of 1% of the number of turbines leads to a 1.11% increase in the installed capacity. The same is true when the dummies are introduced, an increase of 1% of the input leads to a 1.12% increase in the output. This first result changes with the introduction of the interaction term. While off-shore farms still have an expected return of 1.12, on-shore farms have almost constant returns to scale $1.12 - 0.10 = 1.02$. The magnitude of the estimated country-specific effects is negative for Denmark ($\hat{\alpha}_D = -0.346$), Germany ($\hat{\alpha}_G = -0.057$), Ireland ($\hat{\alpha}_I = -0.045$), Sweden ($\hat{\alpha}_S = -0.144$) and United Kingdom ($\hat{\alpha}_{UK} = -0.129$) and positive for the Netherlands ($\hat{\alpha}_{NE} = 0.021$) and Norway ($\hat{\alpha}_{NO} = 0.047$). None of them is statistically significant, suggesting that there is no relevant regional effect. This preliminary conclusion might be an indirect consequence of the rigid functional form imposed by equation (4). For example, if we standardize the residuals of (4), $\varepsilon^{SD} = [\dfrac{\hat{\varepsilon}_1 - E[\varepsilon]}{s.e.(\hat{\varepsilon}_1)}, \dfrac{\hat{\varepsilon}_2 - E[\varepsilon]}{s.e.(\hat{\varepsilon}_2)}, \ldots, \dfrac{\hat{\varepsilon}_N - E[\varepsilon]}{s.e.(\hat{\varepsilon}_N)}]$, and we order them, we can compare their values with the ones of an ordered vector of random draws pulled from a standardized normal dis-

Table 2: Parametric Homogeneous Returns to Scale

	Dependent variable: log y			
	No Country Effect		Country Effect	
	(1)	(2)	(3)	(4)
β_0	0.567***	0.741***	0.792***	0.838***
	(0.076)	(0.077)	(0.164)	(0.152)
log(x)	1.114***	1.126***	1.107***	1.127***
	(0.025)	(0.023)	(0.025)	(0.024)
log(x) · On-Shore		–0.102***		–0.103***
		(0.017)		(0.018)
Observations	209	209	209	209
Adjusted R^2	0.91	0.92	0.91	0.92

Note: *p<0.1; **p<0.05; ***p<0.01

tribution, $\epsilon = [\epsilon_1, \epsilon_2, ..., \epsilon_M]$, with $\epsilon \overset{iid}{\sim} \mathcal{N}(0,1)$, and compare their respective rankings, $[r_1, r_2, ..., r_N]$ and $[s_1, s_2, ..., s_M]$, calculating the Anderson's T,

$$T = \frac{N\sum_{i=1}^{N}(r_i - 1)^2 + M\sum_{j=1}^{M}(s_j - 1)^2}{NM(N+M)} - \frac{4MN-1}{6(M+N)}. \tag{5}$$

The values returned by (5) reject the hypothesis that y and $\hat{y}$ originate from the same distribution (Anderson, 1962). In the same way, a Kolmogorov-Smirnov statistics based on the obtained standardized residuals ε_i^{SD} and the simulated ones ϵ,

$$D_{N,M} = \sup_{\varepsilon} \left| \frac{1}{N}\sum_{i=1}^{N} I_{[-\infty,\varepsilon^{SD}]}\varepsilon_i^{SD} - \frac{1}{M}\sum_{j=1}^{M} I_{[-\infty,\epsilon]}\epsilon_i \right| \tag{6}$$

where

$$I_{[-\infty,\varepsilon^{SD}]} = \begin{cases} 1 & \text{if } \varepsilon_i^{SD} < \varepsilon^{SD} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad I_{[-\infty,\epsilon]} = \begin{cases} 1 & \text{if } \epsilon_i < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

rejects the hypothesis that y and $\hat{y}$ originate from the same distribution. Given that both statistics reject the hypothesis that the dependent variable and the response of regression (4) originate from the same distribution, the estimates $(\hat{\beta}_1, \hat{\beta}_2)$ could be far away in probability to the true unobserved population parameters (β_1, β_2). Hence, it would be risky to rely on Table 2 in order to understand the economies-of-scale of the wind industry. A possible way out of this shortcoming is to avoid any specific functional form and estimate a nonparametric regression,

$$\log y_i = \beta_0 + f(x_i) + \varepsilon_i \qquad E[\varepsilon_i \mid x_i] = 0, \tag{7}$$

able to bypass some of the rigidities imposed by equation (4). The price to pay for this increased flexibility is the necessity to use relatively complex estimation techniques to identify the unknown shape of the function $f(.)$ (Hastie and Tibshirani, 1990). Among the different available options, thin regression splines tend to outperform traditional alternatives like local polynomial approximations, marginal integration and kernels (Wood, 2004). Since the rest of the paper will use this methodol-

Table 3: Nonparametric Homogeneous Returns to Scale

NAM	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
$\hat{f}(x)$	–2.069	–0.760	0.115	0.000	0.528	2.199

ogy, it is useful to see how it works in practice. Like for the parametric specification, the value of the splines can be obtained minimizing the sum of the squared errors

$$\min\sum_{i=1}^{n}\varepsilon_i^2 = \min\sum_{i=1}^{n}(\log y_i - \beta_0 - f(x_i))^2 + \lambda J_m(f), \tag{8}$$

where λ is a smoothing parameter which weights the penalty term $J_m(f) = \int_{\Re}\sum_{\nu_1=m}\frac{m!}{\nu_1!}(\frac{\partial^m f}{\partial x_1^{\nu_1}})^2 dx_1$, added to measure the 'wiggliness' of the estimated function $f(.)$ and $n+1$ is a sequence of non-decreasing real numbers which define the domain of each basis function (Duchon, 1977). The smoothing parameter $\lambda \in [0,\infty)$ controls for the trade-off between model fit and smoothness. If $\lambda \to \infty$, $\hat{f}(.)$ is a straight line which converges to the LS estimate of equation (3). To the contrary, if $\lambda \to 0$, $\hat{f}(.)$ is an un-penalized regression spline which does not correct any form of over-fitting. It can be shown that the function minimizing equation (8), given a fixed λ, is

$$\hat{f}_\lambda(x) = \sum_{j=1}^{M}\theta_j\phi_j(x) + \sum_{i=1}^{n}\delta_i\frac{\Gamma(1/2-m)}{2^{2m}\pi^{1/2}(m-1)!}r^{2m-1}(\| x - x_i \|), \tag{9}$$

where δ_i and θ_j are coefficients to be estimated, $\phi_j(x)$ is the vector of $M = \binom{m-1}{1}$ un-penalized polynomials of x with degrees up to $m-1$ and $\Gamma(1/2-m)/2^{2m}\pi^{1/2}(m-1)!$ is a function which models additional non-linearity using the Gamma probability density $\Gamma(.)$ and the Euclidean distance r between any x and the observed x_i.[1]

The resulting estimates for equation (7) are summarized in Table 3. The estimated $\hat{f}(.)$ is growing in x. Its minimum value is –2.0687 for projects which have only one turbine, this means that, given the value of the intercept $\hat{\beta}_0 = 3.83$, the shape of $\hat{f}(.)$ returns an estimated amount of installed capacity of four MWs. As the number of turbines increases, the values associated to the splines of $f(.)$ increase achieving a maximum value of 3.32, see Figure 2.

The estimated function is very similar to a text book production function which has $f'(.) > 0$ and $f''(.) < 0$. However, unlike for the parametric regression, $\hat{f}(.)$ does not have a single parameter which directly measures the instantaneous rate of change of production. Furthermore, there is no single spline from which to obtain a derivative, making an immediate interpretation of $\hat{f}(.)$ impossible. One way out of this shortcoming is to approximate $\hat{f}'(.)$ using the method of finite differences. The first step is to select a set of equidistant points $[x_1, x_2, ..., x_z]$ and then to obtain a linear projection (LP) of $\hat{f}(.)$,

$$[\hat{f}^{LP}(x_1), \hat{f}^{LP}(x_2), ..., \hat{f}^{LP}(x_z)], \tag{10}$$

for each of them. Moving marginally to the right, it is possible to obtain a new set of values from which the LPs of $\hat{f}(.)$,

1. After a series of rearrangements, it can be shown that a penalized likelihood can be used to estimate (δ_i, θ_j). For all the mathematical details in a frequentist and in a Bayesian framework see (Milivinti and Benini, 2018).

Figure 2: Non-linear fit of $\log \hat{y}$ (Left plot). Non-linear fit of $\hat{f}(.)$ (Right plot).

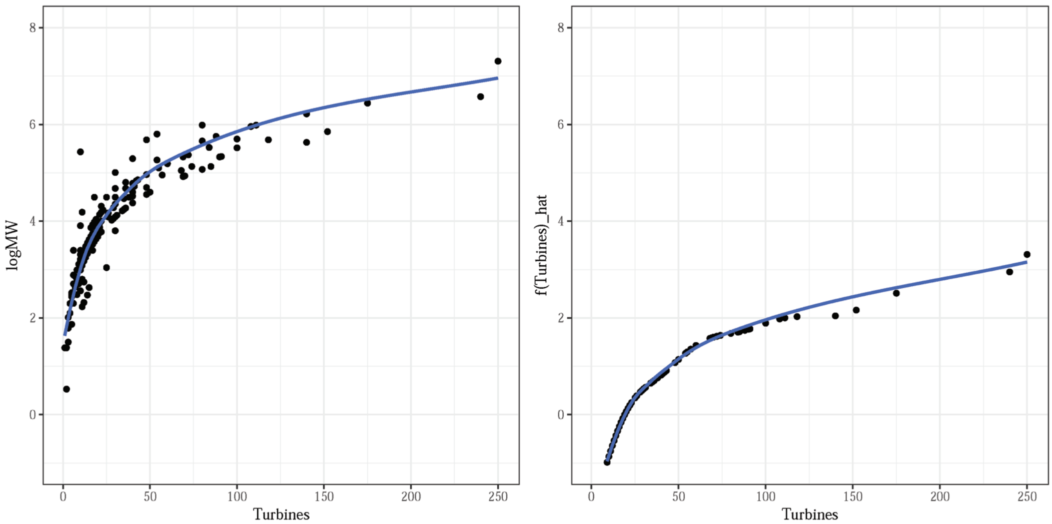
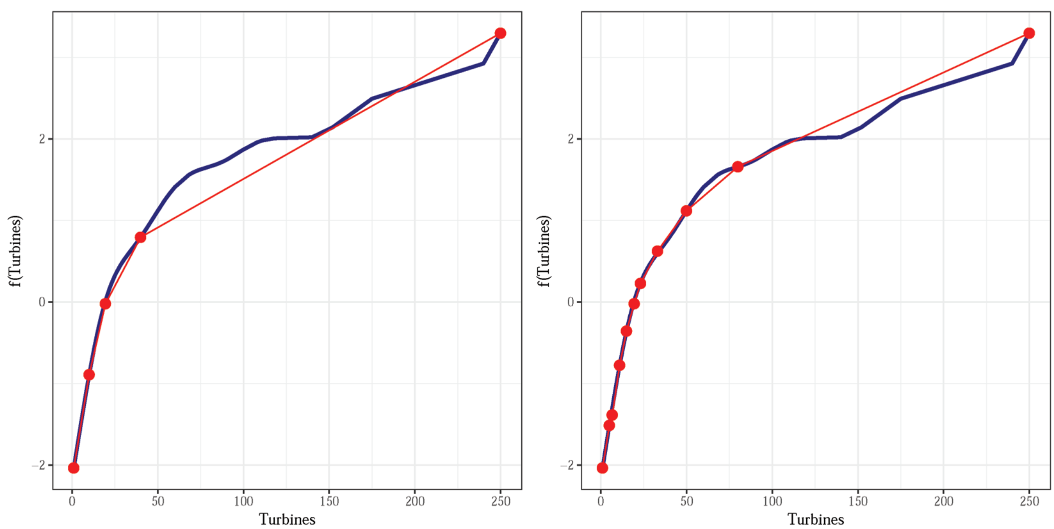


Figure 3: Five points evaluation of $\hat{f}'(.)$ (Left plot). Eleven points evaluation of $\hat{f}'(.)$.



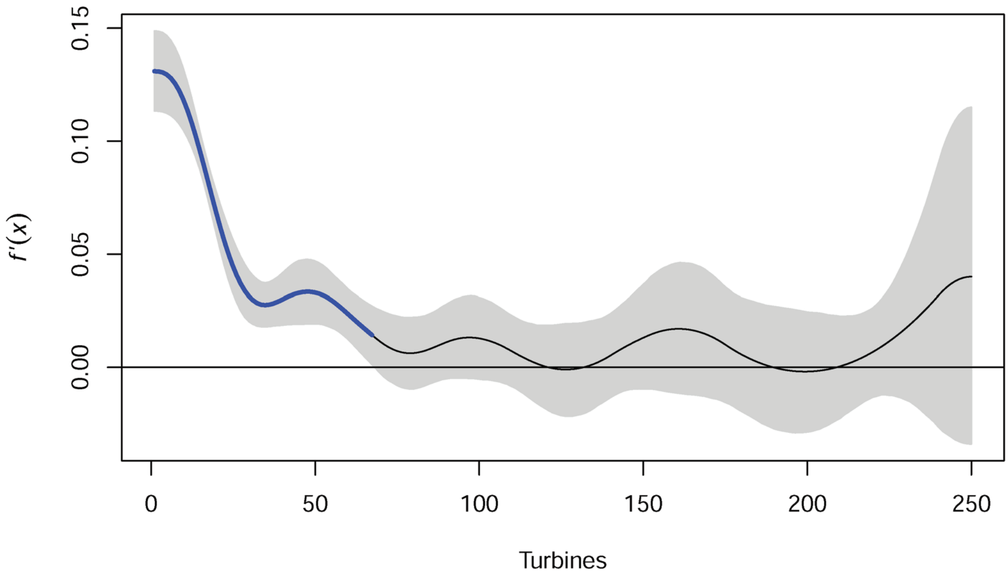
$$[\hat{f}^{LP}(x_1+d), \hat{f}^{LP}(x_2+d), \ldots, \hat{f}^{LP}(x_z+d)] \quad \text{with} \quad d = 1\times 10^{-7},$$

can be recomputed. For every couple of points, the approximated derivative becomes

$$\frac{\hat{f}^{LP}(x_r+d) - \hat{f}^{LP}(x_r)}{d}. \tag{11}$$

The precision of the approximation is function of the total amount of points. For example, Figure 3 shows how an increase in the number of points makes the approximated derivative (the red lines) closer and closer to the true non-observable derivatives of $\hat{f}(.)$.

Figure 4: LP approximation of $\hat{f}'(.)$ obtained using 10,000 simulations evaluated at 200 equidistant points.



Once this first step has been completed, it is possible to extract from the estimated splines the values of the parameters $\hat{\theta}_j$ which multiply the basis $\phi_j(x)$. These parameters can be used to generate 'new' simulated values derived from a multivariate normal distribution, which has the mean of the vector of estimated parameters $\hat{\theta}_j$ and as variance-covariance matrix the one obtained again from the values of $(\hat{\theta}_j, \hat{\theta}_{j+1})$. For example, if we have three points to determine the function, such that $M=3$, the simulation would be

$$\begin{bmatrix} \theta_0^S \\ \theta_1^S \\ \theta_2^S \end{bmatrix} \overset{iid}{\sim} N\left[\begin{pmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix}, \begin{pmatrix} Var(\hat{\theta}_0) & Cov(\hat{\theta}_0,\hat{\theta}_1) & Cov(\hat{\theta}_0,\hat{\theta}_2) \\ Cov(\hat{\theta}_1,\hat{\theta}_0) & Var(\hat{\theta}_1) & Cov(\hat{\theta}_1,\hat{\theta}_2) \\ Cov(\hat{\theta}_2,\hat{\theta}_0) & Cov(\hat{\theta}_2,\hat{\theta}_1) & Var(\hat{\theta}_2) \end{pmatrix}\right]. \tag{12}$$

The simulated θ_j^S can be multiplied by equation (11) in order to obtain the value of the derivative. A t-test can check whether the resulting posterior simulation is statistically significant.

Figure 4 identifies only one segment where the first order derivative is increasing in a statistically significant way, $\hat{f}'(.)$ grows together with x almost without interruption till 70 turbines, then it is statistically insignificant.

Incorporating the decoupling in equation (7), while controlling for country-specific effects,

$$\log y_i = \beta_0 + f(x_i, sec_i) + \alpha^T D_{\alpha,i}^j + \varepsilon_i \quad E[\varepsilon_i \mid x_i, z_i] = 0\,, \tag{13}$$

requires to model the interaction between x and sec. If both variables were continuous, it would be possible to reuse the methodology just described and evaluate $\hat{f}'(.)$ applying the finite difference method (Wood, 2017). However, given that sec is a categorical variable, it would be 'meaningless' to use a thin plate regression spline to fit $f(x_i, sec_i)$ because it is nonsensical to smooth over a factor. There are different ways to model the interaction. The next section shows one which originates from the data.

Table 4: Heterogeneous Returns to Scale

	Dependent variable: log y							
	On-Shore				Off-Shore			
Quantile	1st	2nd	3rd	4th	1st	2nd	3rd	4th
β_0	0.350**	1.691*	0.221	1.211**	1.592***	0.194	1.099***	0.291
	(0.140)	(0.988)	(0.937)	(0.492)	(0.241)	(1.487)	(0.000)	(0.900)
log(x)	1.272***	0.603	1.217***	0.935***	0.917***	1.272***	1.000***	1.218***
	(0.092)	(0.406)	(0.322)	(0.114)	(0.096)	(0.404)	(0.000)	(0.195)
Economies-of-Scale	Yes	No	Yes	No	No	Yes	Yes	Yes
Observations	40	45	30	20	8	18	4	17
Adjusted R^2	0.831	0.027	0.315	0.778	0.928	0.345	1.000	0.704

Note: *p<0.1; **p<0.05; ***p<0.01

2.3 Intermediate Heterogeneous Returns-to-Scale

All the previous specifications imply a homogeneous impact of the number of turbines on the installed capacity. In the first case, it is the sum of coefficients $\hat{\beta}_1 + \hat{\beta}_2 sec$, in the second one the value taken by $\hat{f}'(x)$ at $X=x$. A quantile analysis suggests that this simplification is not justified. Dividing x into quantiles and making a simple regression of log y on log x for the different values of *sec* returns heterogeneous intercepts and slopes, see Table 4.

The previous analysis gives an intuition on how to model the interaction between X and SEC. Given that the regressions' returns are non-constant, there must be a significant amount of cross-sectional heterogeneity. Varying Coefficient Models (VCM) are a semiparametric extension of linear regressions specifically designed to acknowledge the possibility that different units respond differently to the same explanatory variables (Hastie and Tibshirani, 1993). They can be used when the impact of one regressor on the dependent variable changes smoothly over different groups characterized by a certain feature, the so called effect modifier(s). In energy economics, they are mostly implemented to fit and forecast electricity demand, often assuming that the effect modifier is time (Harvey and Koopman, 1993; Chang et al., 2003; Mount et al., 2006; Karakatsani and Bunn, 2008). Their use on the supply side is similar. In the case of the wind industry, they are either used to understand the optimal placement of the turbines (Sanchirico and Wilen, 2005; Pookpunt and Ongsakul, 2013) or to forecast the wind power (Nielsen et al., 2002; Sanchez, 2006), but always using time as the only effect modifier. However, given that VCMs allow to quantify how a change in a continuous variable can affect the impact of an other on the response, it might be useful to propose other smoothers to study cross-fields trade-offs.

Table 4 suggests to re-adapt an intuition which originated in agricultural economics. Agrarian, as well as farming, firms are able to extract different amounts of output given different extensions of their operating facilities. This means that changing the extensive margin of the field changes its intensive margin. For example, enlarging the size of a grain field by 1 square kilometre changes the crop yield per square kilometre. In other words, the amount of crops obtainable per acre is a function of the total number of ploughed kilometres. In a similar way, the amount of milk obtainable per-cow is a function of the total amount of livestock managed by a dairy industry (Løyland and Ringstad, 2001). One of the reasons which might be behind the heterogeneous returns-to-scale is the intervention of non-market forces in the decisions taken by farmers (Severance-Lossin and Sperlich, 1999).

Table 5: Varying Coefficient Heterogeneous Returns to Scale

VCM		Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
$\hat{\beta}_1 + \hat{g}(x)$	Off-Shore	–1.535	0.045	0.579	0.457	1.101	2.754
$\hat{g}(x)$	On-Shore	–2.695	–1.374	–0.546	–0.655	–0.160	1.572

The case of the wind power industry is more complex. On one side, the amount of installed capacity depends upon the number of turbines, and on the other, depends upon the height and the size of the rotors. Given the presence of significant wake effects, a project manager faces a critical trade-off: either few tall turbines with installed large rotors, or many small ones with small engines.

One way to portray this complex situation is to use the continuous variable x_i a thin plate regression spline with a modified penalty, which shrinks to zero when there is a high enough smoothing parameter, and a random effect for sec_i. In this case, we would produce a full tensor product where the first base is a thin plate regression spline, with extra shrinkage, while the second one is a random effect transforming (13) into

$$\log y_i = \beta_0 + g(x_i)\gamma_i sec_i + \alpha^T D^j_{\alpha,i} + \varepsilon_i \quad \gamma_i \sim N(0,1), \tag{14}$$

a Random Coefficient Model (RCM), where $f(x_i, sec_i) = g(x_i)\gamma_i sec_i$. An implicit assumption of equation (14) is that the impact of *sec* on log *y* is independent from *x*. Therefore, any difference between on-shore and off-shore is uncorrelated with the total number of turbines. However, this assumption does not hold in the wind power industry because the number of installed turbines is a function of the site of deployment. Furthermore, equation (14) does not explain why different cross-sectional units behave heterogeneously. A less restrictive technique is to have a separate smoother for each level of the factor *sec*, and associated with it, a single smoother $\hat{g}(.)$ (Rose et al., 2012). The resulting specification is a VCM,

$$\log y_i = \beta_0 + \beta_1 sec_i + g(x_i) sec_i + \alpha^T D^j_{\alpha,i} + \varepsilon_i \qquad E[\varepsilon_i \mid x_i, z_i] = 0, \tag{15}$$

where $f(x_i, sec_i) = \beta_1 sec_i + g(x_i) sec_i$. Even though equation (15) might look similar to (14) or even to (13), it is not. The VCM transforms the individual returns of every wind farm into a non-random function of the number of turbines discovering eventual hidden paths of the dataset while maintaining the linearity of the *sec*. Consequently, the impact of the factor on the response changes jointly with the specific qualities of the cross-sectional observations. In particular, since each coefficient is no longer an average but a function, it is possible to detect latent patterns while keeping an interpretation of the estimate very similar to that of a linear specification. Furthermore, this new specification relaxes one of the assumptions about the information set of the decision makers. In particular, assuming that equation (13) is the result of an optimization process, decision-makers need to know the joint impact of (*X*,*SEC*) on *Y*. To the contrary, in (15) they need to know only the impact of the turbines on the amount of generated electricity, while being unaware of the effect of the site (Benini et al., 2016). The resulting scale elasticities, obtainable using the estimation technique presented in Section 2.2, are summarized in Table 5 and portrait in Figure 5.

Like in the simpler nonparametric case, it is not possible to directly measure the instantaneous rate of change of the production function. At the same time, due to the presence of the factor sec_i, it is not possible to reuse the method of finite differences presented in Section 2.2. An alternative solution is to evaluate the obtained smooths$\phi_j(x)$ for the two set of values [On-Shore, Off-Shore] (Rose et al., 2012). In order to do that, the first step is to re-evaluate the model using a set of LPs,

Figure 5: VCM fit of $\log \hat{y}$ (Left plot). VCM fit of $\hat{g}(.)$ (Right plot).

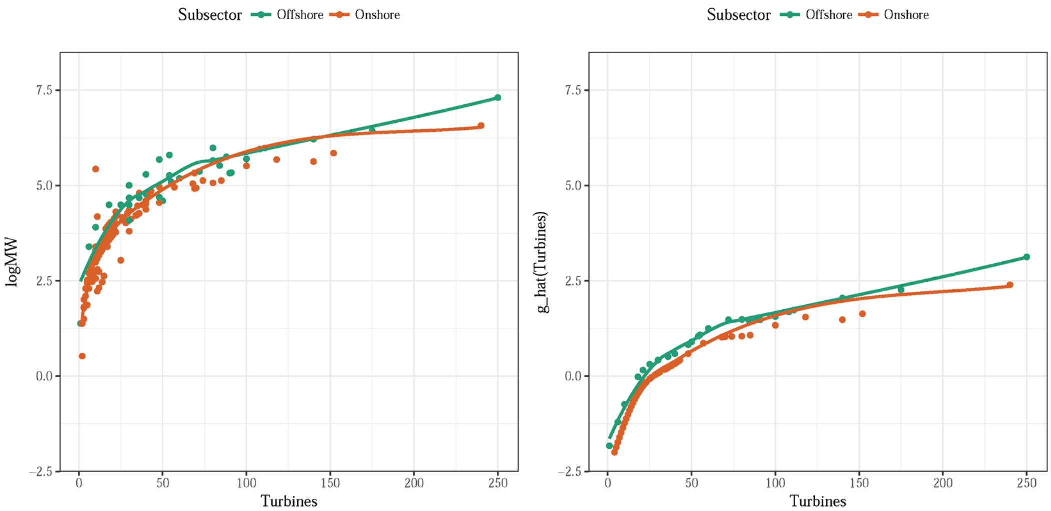
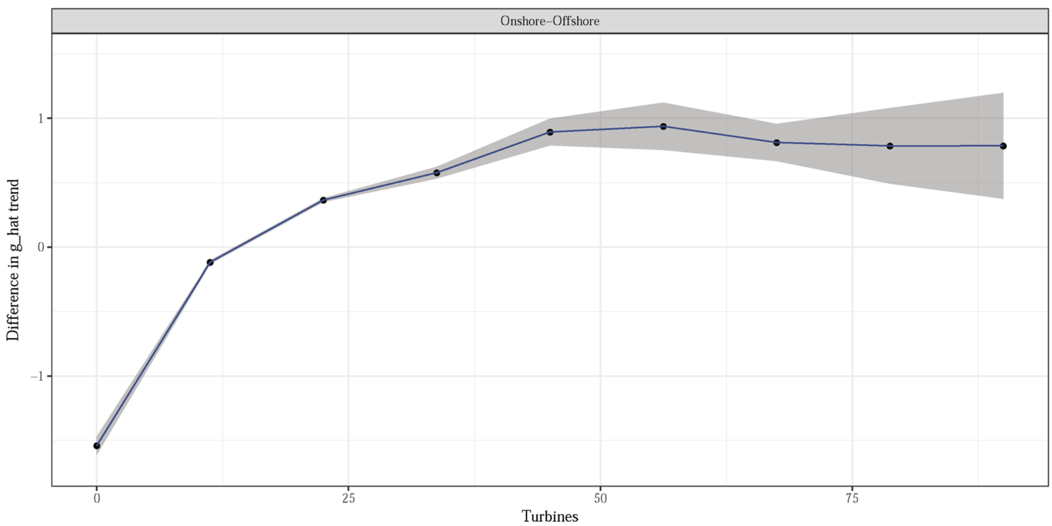


Figure 6: LP approximation of the differences between the estimated $g'(.)$ for the British Isles.



like in (10), for each factor. These new values can be multiplied by the basis coefficients obtained from (15). Once the LPs have been multiplied by $(\hat{\delta}_i, \hat{\theta}_j)$, it is possible to row-sum them and to use the resulting matrix to compute differences between the $\hat{\phi}_j$ before the actual splines are calculated[2].

Using this estimation technique, the VCM captures regional differences which go behind the introduction of a simple dummy variable. In particular, equation (15) computes a separate smoother for every level of the factor *sec*, so we can compare the ratio of the two smoothers for every sub-region of the dataset. Figures 6–8 report the outcome for the British Isles (Ireland and

2. Note that this process should be repeated for every pair of factors. In this case, the factor *sec* only takes two values, so only one comparison is needed.

Figure 7: LP approximation of the differences between the estimated $g'(.)$ for Scandinavia.

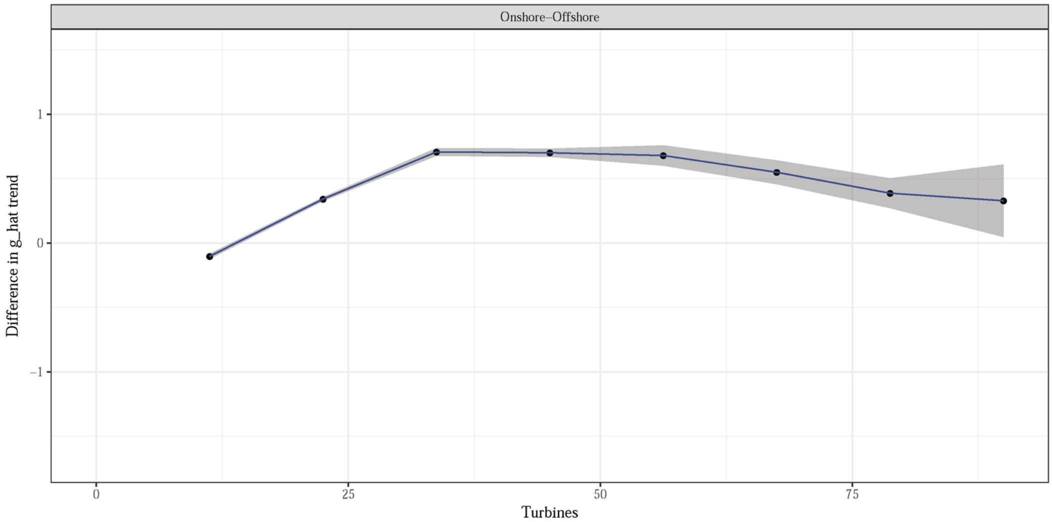
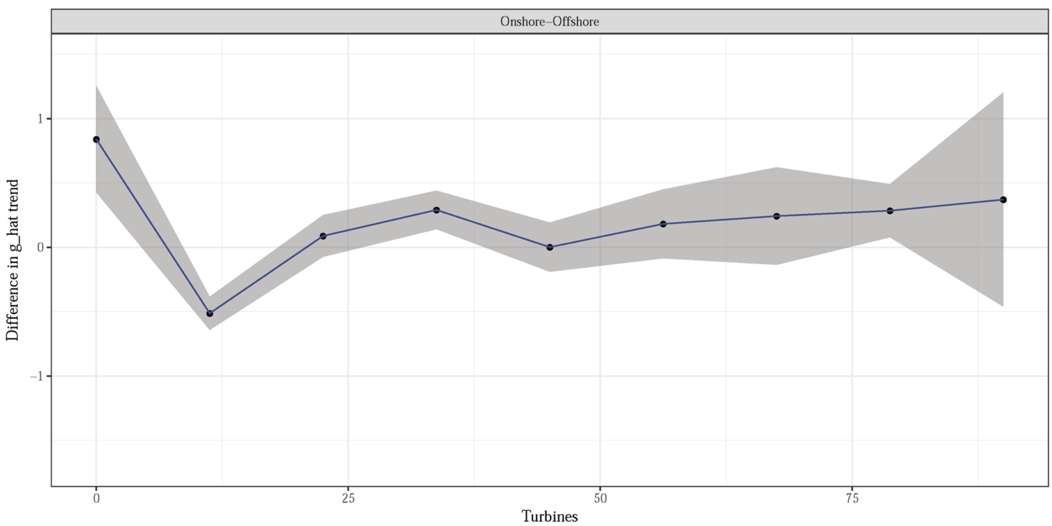


Figure 8: LP approximation of the differences between the estimated $g'(.)$ for Continental Europe.



United Kingdom), Scandinavia (Denmark, Norway, Sweden) and continental Europe (Belgium, Germany, Netherlands).

If the approximated difference is above zero, off-shore installations are more productive than on-shore ones and vice-versa. Like in (Voormolen et al., 2016), we find significant regional differences in the productivity of wind farms. The first two figures show that for small farms the amount of power installed per turbine is favourable to on-shore sites, while for farms with more than 20 turbines, the off-shore sites become more competitive. The difference in extractable power per turbine is much bigger in the case of the British Isles, which has a difference that is almost double to the Scandinavian one. In contrast, continental Europe seems to have almost insignificant difference between the two sites.

3. CONCLUSION AND POLICY RECOMMENDATIONS

Returns-to-scale in the wind sector are more complicated than in most other energy technologies. Using data from the North of Europe, we noticed how standard parametric regressions might overlook their complexity returning estimates which are functional specific and blind to regional differences. More flexible semiparametric alternatives, like the VCM, capture, otherwise unobservable, heterogeneity in the capacity of different installations to amortize per-unit costs. In Continental Europe the highest per-turbine return are achieved by small on-shore installations. To the contrary, in the British Isles big off-shore projects are more productive. For example, off-shore farms with 60 turbines can extract almost one MW more per turbine than their on-shore competitors. The same is true in Scandinavia, where off-shore farms with 30 turbines produce 0.65 MW per turbine more than the on-shore ones.

Climate conditions alone can hardly explain the disparity. National policies probably account for most of the difference between the off-shore and the on-shore performance suggesting that, despite a common EU framework, national stakeholders can influence the trade-off between the two types of installations.

Any further research should incorporate the presence of indirect costs in the the analysis of the farm's productivity. In particular, wind energy is characterized by three major externalities. First, it varies according to the wind's speed, and therefore it is not dispatchable according to the demand (Chandler, 2008). Second, wind turbines have a negative environmental impact, especially in terms of land use (Wang and Wang, 2015; Hadian and Madani, 2015). Third, even if the public attitude toward renewable energies is often positive, the actual construction of farms frequently faces low approval rates among local communities (Wolsink, 2007). Therefore, the empirical analysis should identify, among the projects which have the best economies-of-scale, the ones with the lowest indirect cost focusing on the impact that wind projects have on the grid, the environment, and the communities.

Furthermore this method should be applied to other regions. This research could not do so because of data limitations and market decoupling issues. Nevertheless, if enriched with all the indirect costs, the methodology presented should identify the projects able to extract the highest amount of MW per turbine, across different areas of the globe.

ACKNOWLEDGMENTS

Ludovic Gaudard would like to thank the Swiss National Science Foundation for funding his research (SNSF grant number: P300P1_174397).

REFERENCES

Aigner, D.J. and S.-F. Chu (1968). "On estimating the industry production function." *The American Economic Review* 58(4): 826–839.

Anderson, T.W. (1962). "On the distribution of the two-sample cramer-von mises criterion." *The Annals of Mathematical Statistics* 1148–1159. https://doi.org/10.1214/aoms/1177704477.

Benini, G., S. Sperlich, and R. Theler (2016). *Varying coefficient models revisited: An econometric view*. In *Nonparametric Statistics*, pp. 59–73. Springer. https://doi.org/10.1007/978-3-319-41582-6_5.

Berry, D. (2009a). “Innovation and the price of wind energy in the us.” *Energy Policy* 37: 4493–4499. https://doi.org/10.1016/j.enpol.2009.05.071.

Berry, D. (2009b). “Innovation and the price of wind energy in the us.” *Energy Policy* 37(11): 4493–4499. https://doi.org/10.1016/j.enpol.2009.05.071.

Blanco, M. I. (2009). “The economics of wind energy.” *Renewable and Sustainable Energy Reviews* 13(6): 1372 – 1382. https://doi.org/10.1016/j.rser.2008.09.004.

Chandler, H. (2008). “Empowering variable renewables options for flexible electricity systems.” *International Energy Agency*, Paris, France.

Chang, Y., E. Martinez-Chombo, et al. (2003). *Electricity demand analysis using cointegration and error-correction models with time varying parameters: The mexican case*. Rice University, WP2003-10.

Dismukes, D.E. and G.B. Upton Jr (2015). “Economies of scale, learning effects and offshore wind development costs.” *Renewable Energy* 83, 61–66. https://doi.org/10.1016/j.renene.2015.04.002.

Duchon, J. (1977). “Splines minimizing rotation-invariant semi-norms in sobolev spaces.” *Constructive theory of functions of several variables, Lecture Notes in Mathematics* 571: 85–100. https://doi.org/10.1007/BFb0086566.

González-Longatt, F., P. Wall, and V. Terzija (2012). “Wake effect in wind farm performance: Steady-state and dynamic behavior.” *Renewable Energy* 39(1): 329–338. https://doi.org/10.1016/j.renene.2011.08.053.

Haavelmo, T. (1943). “The statistical implications of a system of simultaneous equations. *Econometrica.*” *Journal of the Econometric Society* 11(1): 1–12. https://doi.org/10.2307/1905714.

Haavelmo, T. (1944). “The probability approach in econometrics.” *Econometrica: Journal of the Econometric Society* 12: iii–115. https://doi.org/10.2307/1906935.

Hadian, S. and K. Madani (2015). “A system of systems approach to energy sustainability assessment: Are all renewables really green?” *Ecological Indicators* 52(Supplement C): 194–206. https://doi.org/10.1016/j.ecolind.2014.11.029.

Harvey, A. and S.J. Koopman (1993). “Forecasting hourly electricity demand using time-varying splines.” *Journal of the American Statistical Association* 88(424): 1228–1236. https://doi.org/10.1080/01621459.1993.10476402.

Hastie, T. and R. Tibshirani (1993). “Varying-coefficient models.” *Journal of the Royal Statistical Society. Series B (Methodological)* 55(4): 757–796.

Hastie, T.J. and R.J. Tibshirani (1990). *Generalized additive models*. CRC Press.

Ibenholt, K. (2002). “Explaining learning curves for wind power.” *Energy Policy* 30(13): 1181–1189. https://doi.org/10.1016/S0301-4215(02)00014-9.

IEA, NEA, and OECD (2015). *Projected costs of generating electricity*. International Energy Agency, Nuclear Energy Agency and Organisation for Economic Co-operation and Development.

Imbens, G.W. (2007). “Nonadditive models with endogenous regressors.” *Econometric Society Monographs* 43: 17.

Imbens, G.W. and W.K. Newey (2009). “Identification and estimation of triangular simultaneous equations models without additivity.” *Econometrica* 77(5): 1481–1512. https://doi.org/10.3982/ECTA7108.

Junginger, H., A. Faaij, and W. Turkenburg (2005a). “Global experience curves for wind farms.” *Energy Policy* 33: 133–150. https://doi.org/10.1016/S0301-4215(03)00205-2.

Junginger, M., A. Faaij, and W.C. Turkenburg (2005b). “Global experience curves for wind farms.” *Energy policy* 33(2): 133–150. https://doi.org/10.1016/S0301-4215(03)00205-2.

Karakatsani, N.V. and D.W. Bunn (2008). “Forecasting electricity prices: The impact of fundamentals and time-varying coefficients.” *International Journal of Forecasting* 24(4): 764–785. https://doi.org/10.1016/j.ijforecast.2008.09.008.

Kim, H., C. Singh, and A. Sprintson (2012). “Simulation and estimation of reliability in a wind farm considering the wake effect.” *IEEE Transactions on Sustainable Energy* 3(2): 274–282. https://doi.org/10.1109/TSTE.2011.2174260.

Lantz, E., M. Hand, and R. Wiser (2012). *Past and future cost of wind energy: Preprint*. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO. https://doi.org/10.2172/1183175.

Løyland, K. and V. Ringstad (2001). “Gains and structural effects of exploiting scale-economies in Norwegian dairy production.” *Agricultural Economics* 24(2): 149–166. https://doi.org/10.1111/j.1574-0862.2001.tb00020.x.

Mas-Colell, A., M.D. Whinston, J.R. Green, et al. (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

Milivinti, A. and G. Benini (2018). “A bayesian semiparametric approach for trend-seasonal interactions: an application to migration forecasts.” Working Paper.

Morgan, C.A., H.M. Snodin, N.C. Scott, and P. Raftery (2003). *Offshore Wind: Economies of scale, engineering resource and load factors*, Volume 18. Citeseer.

Mount, T.D., Y. Ning, and X. Cai (2006). “Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters.” *Energy Economics* 28(1): 62–80. https://doi.org/10.1016/j.eneco.2005.09.008.

Neij, L. (2008). "Cost development of future technologies for power generation—a study based on experience curves and complementary bottom-up assessments." *Energy policy* 36(6): 2200–2211. https://doi.org/10.1016/j.enpol.2008.02.029.

Neij, L., P.D. Andersen, and M. Durstewitz (2004). "Experience curves for wind power." *International journal of energy technology and policy* 2(1–2): 15–32. https://doi.org/10.1504/IJETP.2004.004585.

Nielsen, T.S., H A. Nielsen, and H. Madsen (2002). "Prediction of wind power using time-varying coefficient functions." In *Proceedings of the XV IFAC World Congress*.

Pookpunt, S. and W. Ongsakul (2013). "Optimal placement of wind turbines within wind farm using binary particle swarm optimization with time-varying acceleration coefficients." *Renewable Energy* 55: 266–276. https://doi.org/10.1016/j.renene.2012.12.005.

Rose, N.L., H. Yang, S.D. Turner, and G.L. Simpson (2012). "An assessment of the mechanisms for the transfer of lead and mercury from atmospherically contaminated organic soils to lake sediments with particular reference to Scotland, UK." *Geochimica et Cosmochimica Acta* 82: 113–135. https://doi.org/10.1016/j.gca.2010.12.026.

Sanchez, I. (2006). "Short-term prediction of wind energy production." *International Journal of Forecasting* 22(1): 43–56. https://doi.org/10.1016/j.ijforecast.2005.05.003.

Sanchirico, J.N. and J.E. Wilen (2005). "Optimal spatial management of renewable resources: matching policy scope to ecosystem scale." *Journal of Environmental Economics and Management* 50(1): 23–46. https://doi.org/10.1016/j.jeem.2004.11.001.

Severance-Lossin, E. and S. Sperlich (1999). "Estimation of derivatives for additive separable models." *Statistics: A Journal of Theoretical and Applied Statistics* 33(3): 241–265. https://doi.org/10.1080/02331889908802693.

Stern, N.H. (2007). *The economics of climate change: the Stern review*. Cambridge University Press. https://doi.org/10.1017/CBO9780511817434.

Voormolen, J., H. Junginger, and W. van Sark (2016). "Unravelling historical cost developments of off-shore wind energy in Europe." *Energy Policy* 88: 435–444. https://doi.org/10.1016/j.enpol.2015.10.047.

Wang, S. and S. Wang (2015). "Impacts of wind energy on environment: A review." *Renewable and Sustainable Energy Reviews* 49: 437–443. https://doi.org/10.1016/j.rser.2015.04.137.

Wiser, R. and M. Bolinger (2011). "2010 wind technologies market report." Technical report, National Renewable Energy Laboratory (NREL), Golden, CO.

Wolsink, M. (2007). "Wind power implementation: The nature of public attitudes: Equity and fairness instead of backyard motives." *Renewable and Sustainable Energy Reviews* 11(6): 1188–1207. https://doi.org/10.1016/j.rser.2005.10.005.

Wood, S.N. (2004). "Stable and efficient multiple smoothing parameter estimation for generalized additive models." *Journal of the American Statistical Association* 99(467): 673–686. https://doi.org/10.1198/016214504000000980.

Wood, S.N. (2017). *Generalized additive models: an introduction with R*. CRC press.

Wüstemeyer, C., R. Madlener, and D.W. Bunn (2015). "A stakeholder analysis of divergent supply-chain trends for the European onshore and offshore wind installations." *Energy Policy* 80: 36–44. https://doi.org/10.1016/j.enpol.2015.01.017.