Sustainability Environmental Economics and Dynamics Studies

Working Paper Series

Modelling the Green Knowledge Production Function with Latent Group Structures for OECD countries

by

Saptorshee Kanto Chakraborty, Massimiliano Mazzanti

07/2019

SEEDS Working Paper 07/2019
August 2019
By Saptorshee Kanto Chakraborty, Massimiliano Mazzanti

Modelling the Green Knowledge Production Function with Latent Group Structures for OECD countries[1]

Saptorshee Kanto Chakraborty

University of Ferrara

Massimilliano Mazzanti

University of Ferrara, SEEDS

Abstract

We explore the green knowledge production function and human capital spillovers in the OECD region using a latent group structure. The number of groups and the group membership are both unknown, we determine these unknowns using a penalized regression technique in the presence of cross-sectional dependence in error terms and nonstationarity. We find substantial heterogenous groups classified under three distinctive groups and their efficient estimates. We try to model the green knowledge production function with Latent-Group Structures using PPC- base method with one unobserved global non-stationary factor, we find heterogeneous behaviour in green technologies using a Cup-Lasso estimate. Human capital and expenditure in Research and Development plays an important part in our findings

Keywords: Green Innovation, Human Capital Spillover, Gross Research and Development, OECD, C-Lasso

[1]Presented at EAERE- 2019 Annual conference at Manchester

1. Introduction

The biggest daunting challenge faced by human civilization in the post-war era is to maintain suitable ecological well-being without harming levels of economic growth patterns, this requires moving from dirtier to clean technologies without disrupting the engine of economic development (Stokey, 1998; Aghion et al., 1998, chapter 5; Acemoglu et al., 2012). Following the footsteps of Griliches, 1979 a lot of work has been done both at macro, mezo and micro level which relates knowledge to analyze the pathways towards creation of innovation Eberhardt et al., 2013, Charlot et al., 2015, including green innovation. To achieve sustainable targets as defined by Kyoto protocol and recently Paris Accords, countries can adopt multiple measures like taxation, tax reliefs and also public spending to incentivize green innovation through research and development.

So our research focuses the mainly on the aspect to understand the evolution of green knowledge and the spillover from acquirement of knowledge production, since knowledge itself is a complex factor to be quantified we try to model green knowledge through somewhat agreed upon units of research and development along with human capital which might be considered as basic innovation inputs.

We intend to quantify the concept of Green Knowledge Production function for selected OECD countries. We assume that cross-section dependence is generated by unobserved common factors which are both stationary and nonstationary in nature. Using a C-Lasso estimator as proposed by Su et al., 2016 and Huang et al., 2018b, we find distinctive groups with their coefficient being responsive and also non-responsive to green innovation. The remainder of this paper is organized as follows, Section 2 outlines the contemporaneous ongoing research regarding spillover studies and empirical methodologies in spillover studies. Section 3 outlines the evolution of Green knowledge production function from traditional knowledge production function. Section 4 describes the model and methodology applied. Section 5 lays down the empirical equations and data used. Section 6 presents the results and section 7 concludes.

2. Literature review and outline

Our paper can be related a burgeoning literature of evaluation of spillover effects without any prior assumption of the framework of synergy among individual units (countries in our case). Lam and Souza, 2014 uses a random effects approach to determine structural interaction by integrating over a class of network formation models. de Paua, Rasul and Souza (2016) identifies spillover effects and interactivity in structures by using reduced form equations for both endogenous and contextual effects. Using an adaptive elastic net approach Bonaldi et al., 2015 measures systemic conditional on estimated network for banks inside the European union via liquidity auctions of European Central Bank. Rose (2018) uses a Self Tuning Instrumental Variable (STIV) approach to identify and estimate spillover effects of R&D in an oliogopolistic model for US firms. Manresa, 2016a uses a pooled Lasso estimator (a panel variant Lasso as proposed by Tibshirani, 1996) by keeping unrestricted relationship among interaction structures and observables/unobservables (time-invariant) to study R&D spillovers of US firms, her objective is a noble one since the methodology estimates the reference group and the magnitude of spillover effects without any prior information.

But all of these methodologies do not consider or considers only weak cross-sectional dependence among error terms and limited time-series dependence, through our method we try to consider strong cross-sectional dependence in error-terms as defined by Bailey et al., 2016 and presence of non-staionarity in the estimation framework itself. Our focus is mainly on OECD countries which is a heterogeneous group compared to the G-7 or Eurozone, so to deal with unobservable heterogeneity especially cross-sectional dependence in error terms a prevalent phenomenon in spillover and innovation studies (Pesaran, 2015b) we apply latent group structure methodology following Huang et al., 2018b which employs a variant form of Lasso to handle unobserved heterogeneity in the form of non-stationarity and cross-sectional dependence. The technique also employs penalized principal component to identify individual group membership to estimate group-specific long-run relations.

Traditional fixed-effects panel data model assumes cross-sectional units are heterogeneous in terms of time-varying intercepts with a homogeneous slope coefficient, but this assumption of homogeneity in slope has been a debatable issue in econometric literature. To deal with this issue, the traditional view is to split the data into similar groups and apply standard fixed-effects model to each of them in this type of models unobserved heterogeneity enters the model additively. Time and again this method has been criticized in studies Hsiao and Tahmiscioglu, 1997, Lee et al., 1997, Phillips and Sul, 2007, Su and Chen, 2013. Over the years different approaches have emerged to deal with unknown group structure with respect to inferencing unobserved slope heterogeneity. The first one being finite mixture models, Sun, 2005 proposes a finite parametric linear mixture model; Kasahara and Shimotsu, 2009, Browning and Carro, 2013 uses nonparametric discrete mixture distributions to identify finite number of groups in a discrete choice panel data. Another concept is cluster analysis by using K-means algorithm. Quite a lot of progress has been made in this regard, Lin and Ng, 2012, Sarafidis and Weber, 2015, Bonhomme and Manresa, 2015, Ando and Bai, 2016 have all worked using a K-means algorithm to deal with slope based heterogenenity. Su et al., 2016 have used a variant form of Lasso, C-Lasso [Classifier-Lasso] to identify latent group pattern when the slope coefficients exhibit group structure. Due to increase in availability of macro-economic data there has been a surge in theoretical econometric papers dealing with unobserved heterogeneity by imposing latent group patterns in panels of large dimensions, see Su et al., 2016, Su and Ju, 2018, Huang et al., 2018a and Huang et al., 2018b, Lu and Su, 2017, Bonhomme et al., 2016, Wang et al., 2019. In this it is notable to mention Huang et al., 2018b [HPS, (2018) hereafter] have extended the technique as proposed by Su et al., 2016 [SSP, (2016) hereafter] to deal with cross-sectional dependence in non-stationary time series, irrespective of I(0) or I(1) order, SSP (2016) introduces Classifier-Lasso (C-Lasso, hereafter) to study unobserved grouped patterns and HPS (2018) included penalized principal component analysis (PPC) to deal with this cross-sectional dependence and obtain three types of estimators Classifier-Lasso, post-Lasso and continuous-updated Lasso (Cup-Lasso). They asymptotically establish efficiency in estimation technique and consistency in presence cross-sectional dependence in error terms, non-stationarity and unknown group patterns by including PPC technique from Bai, 2009, this also can be viewed as an extension of multi-factor error structure approach of Bai and Ng, 2002, Pesaran, 2006 and Moon and Weidner, 2015 along with others.

3. The Green Knowledge Production Function

Starting from (Schumpeter, 1939, p 100) an enormous amount of literature in growth theory has pointed the role of innovation, modern growth theorists like (Romer, 1986, Rebelo, 1991, Grossman and Krueger, 1991, Aghion and Howitt, 1992) all have recognized the importance of innovation and capital (physical and human) in long-run economic growth. The important factor of innovation is that it helps in converting knowledge both assets and processes into suitable economic payoffs (McCann and Ortega-Argiles, 2013). Though knowledge creation and diffusion are distinctive phenomenons (Schumpeter, 1942) and can help in tracing the difference in between knowledge spillover and externality created by knowledge (Dominicis et al., 2013). These spillovers are very much intertwined with human capital (Romer, 1986, Lucas, 1988, Dakhli and De Clercq, 2004, Aghion et al., 1998). So in short it can be easily commented that human capital and innovation are very much interactive of nature. An overwhelming amount of studies focus on the interaction in between human capital and innovation with various types of human capital (Dakhli and De Clercq, 2004, Chellaraj et al., 2008, Bottazzi and Peri, 2007, Ang, 2011, Ali and Alpaslan, 2017) or at different geographic levels (Griliches, 1979, Coe, 2005, Mairesse and Mohnen, 2010, Charlot et al., 2015).

Recently there has been a surge in literature of environmental innovative activities, a good review paper Barbieri et al., 2016 and some notable mentions in this regards are (Fu et al., 2018, Hepburn et al., 2018, Verdolini and Bosetti, 2017, Vona et al., 2018, Costantini et al., 2017, Marin and Zanfei, 2019, Gagliardi et al., 2016, Gilli and Mazzanti, 2019, Kim et al., 2013, Popp, 2017, Parry et al., 2017, Fisher-Vanden et al., 2014, Popp et al., 2013). However, little research has been done on the connectivity of human capital and innovation which can be used to counter global climate change (though there exists various debate on terminology of such innovative activities, for simplicity in this paper and beyond we define this type of innovative activities as **green innovation**). The main focus of this paper is to deal with the complexity of green innovation and human capital .

We start with a variant version of Cobb-Douglas production function like Griliches, 1979,

$$Y = f(L, K, R) \tag{1}$$

where, Y is a value-added output, inputs are labor represented by L, tangible capital represented by K and knowledge capital represented by R and f(.) is assumed to be Cobb-Douglas, Griliches, 1979 assumes knowledge capital R as a complement to standard inputs. Griliches, 1979 also defines knowledge capital as a function of present and past research and development expenditure,

$$R = G[W(B)RD] \tag{2}$$

where W(B) is the lag polynomial and B is the lag operator (refer to Crepon et al., 1998, and Eberhardt et al., 2013, for more details). Griliches, 1979 then re-writes (1) as

$$Y = AL^{\alpha}K^{\beta}R^{\gamma}exp^{\lambda t+e} \tag{3}$$

where A being a constant, t being the time index which captures a common linear trend λ, e is the stochastic error term and α, β, γ and λ are to estimated. Hall and Mairesse, 1995 uses (3) to obtain output using current and past R&D levels. It is always preferred to use R&D stock instead of lagged values to take into account previous years impact. (3) can be re-written as in logarithmic terms (where lower-case variables denotes logarithmic counterpart of variables in (3)) following, (Hall and Mairesse, 1995, Hall et al., 2010, Eberhardt et al., 2013)

$$y_{it} = \alpha l_{it} + \beta k_{it} + \gamma r_{it} + \lambda_y + \psi_i + e_{it} \tag{4}$$

One of the main drawbacks of Griliches, 1979 knowledge production function, were assumption of perfectly competitive factor markets, full capacity utilization, cross-sectional independence of the error term and absence of spillover effects. Eberhardt et al., 2013 tries to update Griliches, 1979 knowledge production function by adopting newer econometric technique to deal with this shortcomings. In this regard it is interesting to mention theory of *non-excludability and non-rivalry* of knowledge as mentioned by Arrow, 1962 and the implications of such knowledge transfer in a macroeconomic framework.

3.1. Green Knowledge and measuring R&D and Human capital spillovers

The economic intuition behind knowledge spillovers (both green and non-green, we only consider green) of R&D at any level (micro, mezo or macro, we only consider macro) is that benefits from knowledge created can be appropriated irrespective of boundary. This knowledge absorption can take many forms, like educated personnels' (like scientists, engineers) might meet and exchange ideas, or move around units (firms, universities, industries); researchers might read publications or patents of other scholars or through the novel process of reverse engineering. These flows are not very easy to be taken into account especially in numeric terms so that it can be beneficiary for econometricians. In order to estimate and measure R&D spillovers across countries or within countries, traditional research can be categorized into two ways, firstly by estimating a knowledge production function within each country which considers R&D activity and knowledge spillover and then estimate the effect of that knowledge or real outcomes. So, following (1), (2), (3) we use a reduced form equation to quantify our relationship

$$G_{n,t} = f(RD_{n,t}, HC_{n,t}, U_{n,t}) \tag{5}$$

where G can be expressed in terms of Green patents count, f is real function (which can be assumed to be a Cobb-Douglas production function for generality), t = 1,....,T represents time, n = 1,....,N represents cross-sectional units (in our case OECD countries), RD represents Research and Development spending both at Business level and gross country levels. HC stands for Human Capital levels. In this regard it is convenient to comment that HC plays an important role because green innovative activities requires specific skills to support and enhance innovation, which also plays an important role in documenting absorptive capacity of knowledge. Unobservable characteristics as represented by $U_{n,t}$ is an important aspect in this regard. Some characteristics might affect both knowledge and primary innovation inputs, for example, highly skilled workers might non-randomly choose regions or countries for better opportunities by some prior knowledge which are cannot be captured by econometric techniques.

4. Model and Methodology

Spillovers which can also be considered as a sub-class of externality is a common phenomenon especially in network economics, and addressing the issue related have always intrigued policy-makers and economists. Manski, 1993 defines spillover broadly into two categories: *endogenous*, in which outcomes of interest are simultaneous of nature and *contextual* the ones whose interaction takes place through covariate of others. Spillover studies are relevant nowadays in every economic studies like education, crime, consumption, technologyy adoption and productivity (De Giorgi and Pellizzari, 2014, Liu et al., 2012, De Giorgi et al., 2016, Conley and Udry, 2010; Griliches, 1998). Most of these studies assume some pre-defined notion through which interaction among units takes place, Manresa, 2016a defines such notion as *structure of interaction* and also comments that it is somehow misleading to use such pre-defined structures and proposes a new Pooled-Lasso technique in which she treats such structure of interaction to be unobserved and using a sparse interaction structure. This is very much consistent with latent variable framework.

4.1. A. Latent Variables, Measurement errors and Unobservables

A very popular econometric methodology commonly used by applied economists is General Linear Model (GLM), but GLMs' have some shortcomings when there is a presence of unmodelled dependence among units, like temporal, spatial or network. Econometricians try to consider these dependence in the category of unobserved heterogeneity and have been trying long enough to deal with such (Stewart, 2014). In the literature of panel data latent factors or variables play an important role to provide consistent estimators form of time variant cross-sectional data (Bai, 2009, Pesaran et al., 1999, Pesaran, 2006, Moon and Weidner, 2015).

The term "latent variables" is not new in economic/econometric literature, Koopmans, 1949 used the term as distinct from "observed variables" in reference to stochastic disturbances in a standard simultaneous model of supply and demand. Kmenta, 1991 solidified the definition of latent variables and defines latent variables as unobserved variables except stochastic disturbances, and classifies them under following categories:

- Variables for which exact measurements are unavailable and are represented by error contaminated substitutes. Example- National Income.
- Unobservables which can be represented only through closely related substitutes termed as "proxies". Example: Capital stock in Production function.
- Variables that are intrinsically not measurable. Example: Intelligence.

Over the years the use of latent variables in almost all three categories as mentioned above has been utilized in various economic scenarios (Griliches, 1974, Aigner and Goldberger, 1977, Doran and Kmenta, 1986, Kmenta, 1997, Greene, 1990). Innovation of any kind (Green or Non-green) is a difficult measure since it is unobservable of nature, so economists quantify innovation through Research and Development statistics, number of scientific personnel and also via patent counts, but this brings the question of measurement error in innovation studies (the existing literature lacks to deal with such). For example, R&D be it gross, government or business contains presence of measurement error since this statistics is based on finite samples and imperfect sources. This is also true for measuring human capital, since human capital is itself not measurable but taken into consideration through measures by years of schooling or level of educational attainment, one can easily comment the problems with such measures and presence of error in these measures. Where as patents which are mostly considered as an outcome of an innovative activity are an imperfect measure (Johnstone et al., 2010 page 138). So, given the fact our decided variables theoretically has presence of measurement error, we choose our model carefully to deal with such, in fact in this subsection we illustrate how to deal with them.

Tibshirani (1996) proposed a l^1 penalization term for least-square regression,

$$y_i = \sum_{j=1}^{J} \beta_j x_{i,j} + \epsilon_i$$

where the lasso estimate can be defined as:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{J} \beta_j x_{ij} \right)^2 + \lambda_N \sum_{j=1}^{J} |\beta_j|$$

where $\lambda_N \geq 0$ is a penalty parameter, and β = $(\beta_1, ..., , \beta_J)$. Because of the features of efficiency, convexity and sparsity, the Lasso technique have gained a huge amount of popularity. One must mention, the condition of sparsity is important since fact the values of many elements of $\hat{\beta}_j$ can be exactly zero (Society et al., 2010- Chapter 9). The Lasso technique is also useful in reducing measurement error for a variable selection process both in parametric and non-parametric terms (Stefanski et al., 2014, Society et al., 2010- Chapter 8). Lasso technique is also very much efficient in spillover studies and growth studies, because of its variable selection efficiency, (Society et al., 2010- Chapter 7, Manresa, 2016a). The lasso estimator and its various forms also posses the oracle property, i.e., it performs well enough even if its true underlying model was defined beforehand (Zou, 2006).

4.2. B. Model

Consider a following simple fixed-effects model:

$$y_{it} = \alpha_i + \beta_i x_{it} + \sum_{j \neq i} \gamma_{ij} x_{jt} + w_{it}'\theta + e_{it} \tag{6}$$

assuming the outcomes of i =1,....,N, y_{it} is only affected by its own x_{it} and also is influenced by other x's like $x_{1t}, ..., x_{Nt}$. Furthermore, α_i is the intercept used to capture unobserved individual-specific heterogeneity, β_i's are the slope capturing the effect of own heterogenity. γ_{ij} are used to capture the pair-specific of individual j on individual i. We assume, spillover exists, i.e., $\gamma_{ij} \neq 0$ interaction between i & j exists. w_{it} is introduced to capture all aggregate level variables to measure correlated shocks through θ, but w_{it} do not measure any kind of spillover effect and e_{it} captures the idiosynctratic shocks. This model can be used to measure various types of spillovers both overlapping and non-overlapping groups using a linear mean model, (Ballester et al., 2006). The linear models has a shortcoming since, it assumes homogeneity and symmetricity of spillover withing groups.

To overcome this problem, following Manresa, 2016a assumes sparsity, which can deal with heterogeneity, so that the curse of dimensionality can be dealt with ease. Another way of dealing with such heterogeneity is by heterogeneous group structures which may be called a mid-way in between heterogeneity and sparsity (Hahn and Moon, 2010, Bonhomme and Manresa, 2015, Bonhomme et al., 2016, Su et al., 2016). We choose the later one, in group heterogeneity theory the assumption is links within each group are far more predominant than across groups. That is spillover within groups are homogenous of nature whereas across groups are heterogeneous.

4.2.1. i. Group Heterogeneity

Under a group heterogeneity pattern spillover effects are not pair-individual specific, but rather group specific, we explain the partitioning of the group using an Information criterion in a separate section later. Assuming K$_{MAX}$ is the maximum number of groups attainable, N being the total number of individuals irrespective of partitioning and N_{gK} being the maximum number of individuals in each group, where g and K are not fixed, let Υ:1,....,N$\longrightarrow$1,....,N$_{gK}$ be the mapping for individuals i and the group which to it belongs, Υ(i) = K$_{MAX}$. So for and i,j,

$$\gamma_{ij} = \gamma_{\Upsilon(i)\Upsilon(j)} \tag{7}$$

Also, for each g, being the number of groups 1,....,K$_{MAX}$

$$\sum_{g'=1}^{K_{MAX}} \mathbb{S}\{\gamma_{g_g}' \neq 0\} = s_g << T \tag{8}$$

(8) is derived from the sparsity condition, as in Manresa, 2016a Eq (2) and Eq (S3) of Manresa, 2016b, assuming spillover exists i.e, the sum of spillover effect s_i is positive, ($>$ 0), so now s_i becomes relatively smaller than the time-series dimension.

Intuitively, the identities of each individual collapses and remains confined within a group, and spillover within group is assumed to be similar of magnitude. If chosen K$_{MAX}$ becomes N, or the number of individuals becomes one group each, the heterogenity can be explained by (6). So, choosing the number of groups is quiet important to understand the spillover effects especially dealing with heterogeneity and not-so large time dimension.

The group structure methodology also has an attractive feature to deal with omitted variable bias, due to the fact it deals with the problem of sampling. Since, spillover effects are being grouped the representative group elements have enough consistency.

4.2.2. ii. Green Knowledge Spillover

Now, let us bring in the concept of green knowledge in our analysis, which is somehow similar to knowledge as conceptually. Knowledge is a non-rival good and is a process which is assumed to generate externalities. Using the same concept from (1) and (2) we try to structure unobservable spillover effects in a green macro-economic sense, via a Cobb-Douglas production function.

$$OUTPUT_{it} = \alpha_i + \beta RD_{i,t-1} + \sum_{j \neq i} \gamma_{ij} RD_{j,t-1} + \theta controls_{i,t} + e_{i,t} \tag{9}$$

since, we are also dealing with human capital and assuming free movement of skilled labour, (9) can be re-written as

$$OUTPUT_{it} = \alpha_i + \beta_1 RD_{i,t-1} + \sum_{j \neq i} \gamma_{ij} RD_{j,t-1} + \beta_2 HC_{i,t} + e_{i,t} \tag{10}$$

introducing the concept of R&D stock instead of flow and introducing logged values, (10) can be re-written as

$$output_{it} = \alpha_i + \beta_1 rd_{i,t} + \sum_{j \neq i} \gamma_{ij} rd_{j,t} + \beta_2 hc_{i,t} + e_{i,t} \tag{11}$$

but we also use a debt indicator which is not idiosyncratic of nature in (11) as explained in later section. The spillover effects, γ_{ij} which can be simply written as

$$\gamma_{ij} = \gamma.w_{ij}$$

this can be implemented via a spatial proximity or a multi-factor error structure, we chose the later one since of its uniqueness to capture strong error-cross sectional dependence along with non-stationarity.

$$output_{it} = \beta_1 rd_{it} + \beta_2 hc_{it} + \lambda_i' f_t + u_{it} \tag{12}$$

4.3. C. Methodology

One of the easiest ways to deal with unit-specific heterogeneity is time-invariant fixed-effects, but the basic assumption behind fixed-effects is that the unobserved heterogeneity is constant over time, which is strict assumption in regard to spillover studies. Presence of unobserved heterogeneity and cross-sectional dependence can cause inferential problems in nonstationary panels. We borrow from a novel estimation technique as proposed by HPS (2018) to deal with unobserved parameter heterogeneity together with cross-sectional dependence in a nonstationary panel.

The two central goals in statistical based economic modelling are ensuring high prediction accuracy and detecting relevant predictors. Moreover variable selection gains importance if the true underlying model is sparse. Identifying relevant predictors increases the prediction performance of the fitted model. Ordinary Least Square (OLS) gives nonzero estimates to all coefficients, traditional subset selection is based upon manual selection to select significant variables, but this selection procedure bears two limitations. Firstly, if the number of predictors are large, computationally it is improbable to perform subset selection and secondly subset selection bears deep-rooted distinctiveness (Breiman, 1995, Fan and Li, 2001). Subset selection also becomes difficult in presence of stochastic errors or presence of uncertainty in the variable (Fan and Li, 2001, Shen and Ye, 2002, Zou, 2006).

The Lasso model proposed by Tibshirani, 1996 [as briefly explained in section 4.1 A] is a regularization technique for simultaneous estimation and variable selection. It can be defined in the following form. A l^1 penalization term for least-square regression, i.e., absolute value of magnitude of coefficient as penalty term to the loss function

$$y_i = \sum_{j=1}^{J} \beta_j x_{i,j} + \epsilon_i$$

where the lasso estimate can be defined as:

$$\hat{\beta} = \underset{\beta}{\text{argmin}} \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{J} \beta_j x_{ij} \right)^2 + \lambda_N \sum_{j=1}^{J} |\beta_j|$$

where $\lambda_N \geq 0$ is a penalty parameter, and β = $(\beta_1, ..., , \beta_J)$. If λ is zero then we will get back OLS, whereas very large value of λ will make coefficients zero hence it will under-fit. Another type of common penalization technique is l^2 norm, where squared magnitude of coefficient is added as penalty term to the loss function. This type of regression technique is called as Ridge Regression, which follows an euclidean norm penalization. Here, if λ is zero then one can imagine to get back OLS. However, if λ is very large then it will add too much weight and it will lead to under-fitting. So choosing λ is important . This technique works very well to avoid over-fitting issue. Lasso technique have gained a huge amount of popularity due to the features of efficiency, convexity and sparsity, sparsity is important since fact the values of many elements of $\hat{\beta}_j$ can be exactly zero (Society et al., 2010- Chapter 9). Lasso has been heavily used in various fields of applied statistics like signal processing, facial recognition, text mining, genetics, genomics, biomedical imaging, social media analysis and high-frequency finance. Recently it has gained usage in applied economics like Belloni et al., 2011, Belloni et al., 2013, Belloni et al., 2014, and Chernozhukov et al., 2015. The Lasso is a regularization estimation procedure, in which a regression is estimated via an objective function whose purpose is to balance the in-sample goodness of fit using a penalty term, the value of the penalty term depends on the sum of the magnitude of the coefficients used in the regression (Athey, 2018). Due the penalty term, many covariates effectively becomes zero and hence gets dropped from the regression. The magnitude of this penalty term is selected by using cross-validation.

The l^1 is crucial for lasso and is used both in variable selection and regularization, the continuous shrinkage increases the predictability of Lasso as a technique due to bias-variance trade-off. The lasso technique with help of orthogonal predictors provides near-minimax optimality of soft-thresholding (Donoho et al., 1996). The lasso technique can also locate 'right' sparse representation of the model under some pre-defined conditions (Donoho and Elad, 2003), it is also consistent in variable selection provided the model satisfies some condition (Meinshausen and Bhlmann, 2006). So it is safe to assume the lasso technique satisfies oracle properties (*an oracle estimator must be consistent in parameter estimation and variable selection*).

4.3.1. i. Latent group structure with presence of nonstationarity and cross-sectional dependence

We adopt the estimation technique of HPS (2018) for our empirical purpose and a brief explanation of the technique is given below.

Let a dependent variable y_{it} is measured for individuals i = 1...N over time t = 1....T, and x_{it} is p$\times$ 1 vector of non-stationary regressors (irrespective of I(1) or I(0)) and e_{it} is the error term which might be composed of unobserved common factors ϵ_{it} with zero mean and finite variance. β_i's are homogenous within a group but heterogenous across groups represents long-run cointegrating relations, for p$\times$ 1 vectors of dependent variables are represented by β_i^0

$$\begin{cases} y_{it} = {\beta^0}'_i x_{it} + e_{it} \\ x_{it} = x_{it-1} + \epsilon_{it} \end{cases} \tag{13}$$

Now we assume, true values of β_i as β_i^0 which follows a latent group structure,

$$\beta_i^0 = \begin{cases} \alpha_1^0, \text{if } i \in G_1^0 \\ . \\ . \\ \alpha_K^0, \text{if } i \in G_K^0 \end{cases} \tag{14}$$

where, $\alpha_j^0 \neq \alpha_k^0$ for any j$\neq$k, $\bigcup_{k=1}^{K} G_k^0 = 1, 2, ...N$ and $G_k^0 \bigcup G_j^0 = \emptyset$ for any j$\neq$k, we assume the number of groups to be known and the members at this instance, and we calculate the number using an Information criterion as following HPS (2018).

To account for unobserved common patterns Bai and Ng(2004) proposes a multi-factor error structure on e_{it} which is assumed to be cross-sectionally dependent, so

$$e_{it} = {\lambda^0}'_i f_t^0 + u_{it} = {\lambda^0}'_{1i} f_{1t}^0 + {\lambda^0}'_{2i} f_{2t}^0 + u_{it} \tag{15}$$

where f_t^0 is a vector of r$\times$1 of unobserved common factors, which can be broken down into f_{1t}^0 of non-stationary I(1) and f_{2t}^0 stationary I(0) processes, λ_i is a vector of factor loadings and u_{it} is the cross-sectionally independent idiosyncractic component with zero-mean and finite variance. The assumption in this regard is the cross-sectional dependence arises due to common factors.

So, (14) can be re-written as

$$y_{it} = {\beta^0}'_i x_{it} + {\lambda^0}'_i f_t^0 + u_{it} \tag{16}$$

A penalized principal component method is used to to estimate (17), so that the true values of α, β, Λ and f are represented by $\alpha^0, \beta^0, \Lambda^0$ and f^0 , where

$$\alpha \equiv (\alpha_1, ...\alpha_{K_0}),\ \beta \equiv (\beta_1, ...\beta_N),\ \Lambda \equiv (\Lambda_1, ...\Lambda_N)',\ and\ f \equiv (f_1, ...f_T)'$$

The main purpose is to obtain consistent estimators for group-specific long-run relations α_k and unobserved common factors f_t. For that $\alpha_k^0, \beta_i^0, \lambda_i^0$ and f_t^0 are denoted by $\alpha_k, \beta_i, \lambda_i$ and f_t

4.3.2. ii. Penalized Principal Component

(17) can be re-written as

$$y_{it} = x_{it}\beta_i^0 + f_t^0\lambda_i^0 + u_{it} = x_{it}\beta_i^0 + {f_1^0}_t\lambda_{1i}^0 + {f_2^0}_t\lambda_{2i}^0 + u_{it} \tag{17}$$

where $f^0 = (f_1^0, f_2^0)$, $\lambda_i^0 = (\lambda^{0\prime}{}_1, \lambda^{0\prime}{}_2)$, $y_i = (y_{i1}, ..., y_{iT})'$ the definitions of x_i, f_1^0, f_2^0 and u_{it} follows from before.

Following Bai (2009), the least square objective function is defined by

$$SSR(\beta_i, f_1, \Lambda_1) = \sum_{i=1}^{N}(y_i - x_i\beta_i - f_i\lambda_{1i})'(y_i - x_i\beta_i - f_i\lambda_{1i}) \tag{18}$$

the constraint of this equation follows, $\frac{f_1' f_1}{T^2} = I_{r_1}$ also $\Lambda_1'\Lambda_1$ is diagonal. The projection matrix can be defined as $Mf_1 = I_T - Pf_1 = I_T - \frac{f_1' f_1}{T^2}$.

So, now the least square estimates of β_i for each given f_i becomes

$$\hat{\beta}_i = (x_i' M_{f_1} x_i)^{-1} x_i' M_{f_1} y_i$$

Therefore, for given β_i, $e_i = y_i - x_i\beta_i = f\lambda_i + u_i$ tends to posses a pure factor structure. If we defines, $e = (e_1, e_2, ..., e_N)$ as T×N matrix and $\Lambda_1 = (\lambda_{11}, ..., \lambda_{1N})'$ a N $\times$ r_1 matrix, then f_1 can be obtained through least square using

$$tr[(e - f_1\Lambda_1')(e - f_1\Lambda_1')']$$

Using principal component analysis of pure factor models following Conor and Korajzcyk (1986) and Stock and Watson (2002), Bai (2009) and HPS (2018) defines Λ_1 can be concentrated out by its least square estimator $\Lambda_1 = e' f_1 (f_1' f_1)^{-1} = e' f_1 / T^2$, so now (10) can be re-written as

$$tr(e' M_{f_1} e) = tr(e'e) - tr(f_1' e e' f_1 / T^2) \tag{19}$$

So given f we can estimate β and given β we can estimate f, the final least squares estimator $(\hat{\beta}, \hat{f}_1)$, which is the solution for a system of non-linear equations,

$$\hat{\beta}_i = \left(x_i' M_{\hat{f}_1} x_i\right)^{-1} (x_i' M_1 x_i) \tag{20}$$

$$\hat{f}_1 V_{1,N\,T} = \left[\frac{1}{NT^2}\sum_{i=1}^{N}(y_i - x_i\hat{\beta}_i)(y_i - x_i\hat{\beta}_i)'\right]\hat{f}_1 \tag{21}$$

where $V_{1,N\,T}$ is a diagonal matrix consisting of the r_1 largest eigen value of the matrix inside the brackets, arranges in a decreasing order and $M_{\hat{f}_1} = I_T - \frac{1}{T^2}\hat{f}_1 \hat{f}_1{}'$, $\frac{1}{T^2}\hat{f}_1{}'\hat{f}_1 = I_{r_1}$.

(21) and (22) can be shown that $\hat{\Lambda}_1'\hat{\Lambda}_1$ is a diagonal matrix which becomes,

$$\frac{1}{N}\hat{\Lambda}_1'\hat{\Lambda}_1 = T^{-2}\hat{f}_1'\left(\frac{1}{NT^2}\sum_{i=1}^{N}(y_i - x_i\hat{\beta}_i)(y_i - x_i\hat{\beta}_i)'\hat{f}_1\right) = \left(\frac{1}{T^2}\hat{f}_1'\hat{f}_1\right)V_{1,N\,T} = V_{1,N\,T}$$

Since, β_i and f_1 can be estimated from (21) and (22), we follow a penalized principal component method to estimate $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, where $\boldsymbol{\beta}$ tends to exhibit latent group properties. So,

$$Q_{NT}^{\lambda,K}(\boldsymbol{\beta},\boldsymbol{\alpha},f_1) = Q_{NT}(\boldsymbol{\beta},f_1) + \frac{\lambda}{N}\sum_{i=1}^{N}\prod_{k=1}^{K}||\beta_i - \alpha_k|| \tag{22}$$

in this scenario, $Q_{NT}(\boldsymbol{\beta},f_1) = \frac{1}{NT^2}\sum_{i=1}^{N}(y_i - x_i\beta_i)' M_{f_1}(y_i - x_i\beta_i)$, $\lambda = \lambda(N,T)$ is the tuning parameter. Using PPC criterion to minimize (23) we get the Classifier-Lasso estimators of β_i and α_k

So now we update the estimates of non-stationary common factors by minimizing f_1 as in (24) and for stationary common factors by minimizing f_2 as in (25); for (24) the restriction for identification is $\frac{1}{T^2}\hat{f}_1'\hat{f}_1 = I_{r_1}$ and similarly from above $\hat{\Lambda}_1'\hat{\Lambda}_1$ is a diagonal matrix

$$\hat{f}_1 V_{1,N\,T} = \left[\frac{1}{NT^2}\sum_{k=1}^{K}\sum_{i\in\hat{G}_k}(y_i - x_i\hat{\alpha_k})(y_i - x_i\hat{\alpha_k})'\right]\hat{f}_1 \tag{23}$$

and for (16) the identification restrictions are, $\frac{1}{T^2}\hat{f}_2'\hat{f}_2 = I_{r_2}$ and a diagonal matrix $V_{2,N\,T}$

$$\hat{f}_2 V_{2,N\,T} = \left[\frac{1}{NT}\sum_{k=1}^{K}\sum_{i\in\hat{G}_k}(y_i - x_i\hat{\alpha_k} - \hat{f}_1\hat{\lambda}_{1i})(y_i - x_i\hat{\alpha_k} - \hat{f}_1\hat{\lambda}_{1i})'\right]\hat{f}_2 \tag{24}$$

So, now we apply bias-correction in post-lasso estimators of β and α. This tend to take into consideration of unobserved stationary common factors, endogeneity and serial correlation arising from weakly dependent error terms.

SSP (2016) and HPS (2018) have also established the oracle properties of the C-Lasso and its variant estimators, so that the sparsity of this type of estimators are well enough for applied purposes.

4.4. D. Estimating number of unobserved factors and groups

The theory is to assume the r_1^0 and r^0 as true values of generic number of nonstationary factors r_1 and r as the generic value of total number of nonstationary and stationary factors. This two values are estimated by an Information criterion

$$IC_1(r) = logV_1(r,\hat{G}^r) + rg_1(N,T) \tag{25}$$

and

$$IC_2(r) = logV_2(r_1,\hat{f}_1^{r_1}) + r_1 g_2(N,T) \tag{26}$$

where $g_1(N,T)$ and $g_2(N,T)$ are two penalty functions, and following Bai, 2009, Bai and Ng, 2002, HPS (2017) and SSP (2016) one can determine the values of $g_1(N,T)$ and $g_2(N,T)$ as $g_1(N,T) = (\frac{N+T}{(NT)})log(min(N,T)$ and $g_2(N,T) = g_1(N,T) \times \frac{T}{4log(log(T)}$.

Whereas to determine the number of groups K a BIC type criterion is followed, it is assumed being the true number of groups K_0 is bounded by a finite-integer from above K_{MAX} , so for this a new Information criterion is proposed,

$$IC_3(K, \lambda) = logV_3(K) + pKg_3(N, T) \tag{27}$$

where g_3(N,T) is a penalty function. The minimizer IC_3 (K, λ) with respect to K is assumed to be K_0 for values of λ. HPS(2018) proves $\lambda = c_\lambda \times T^{-3/4}$.

5. Taking the model to the Data

5.1. A. Data

Dealing with the empiric of Green Knowledge spillover on a macro-level is of extreme challenge, due to problems of identification of measurement error in output, endogeneity among inputs and perplexity of spillover. We use green patent data of IPC-CPC (**Y02**) category (Development of environment-related technologies) in ratio terms with respect to total patents from OECD Patents in environment-related technologies database (OECD, 2018), (for better understanding please refer to section A in appendix at the end of the chapter). We also include Human capital from PWT 9.0 (Feenstra et al., 2016) (for better understanding please refer to section B in appendix at the end of the chapter). One can say this solves the biasness in measurement error of outputs, due to the fact we are specifying specific types of patents (Griliches, 1998 page-319). Following Manresa, 2016a we proceed with given Knowledge i.e., past R&D investments, so it can be said to be uncorrelated from any type of future or ongoing shocks so the estimator is unbiassed. R&D data is collected from OECD-stats database following the methodology proposed by Coe et al., 2009 as explained in Section C of appendix. We compute both Business Enterprise Reseach and Development (BERD) and Gross domestic Expenditure on Research and Development (GERD). Human capital data is from Penn World Table and Debt to GDP ratio is fro IMF statistics.

We choose a sample of 25 OECD countries as mentioned below in the table (1) for a time-period of 1971-2014.

Table 1: Country sample

Australia	Finland	Ireland	Mexico	Spain
Austria	France	Israel	Netherlands	Sweden
Belgium	Germany	Italy	New Zealand	Switzerland
Canada	Greece	Japan	Norway	UK
Denmark	Iceland	Korea	Portugal	USA

5.2. B. Estimation Strategy

We divide our estimation into two categories, first we check the heterogeneous effects of research and development (business enterprise research and development, BERD and gross expenditure on research and development, GERD) and human capital on green innovative activities, then we introduce a debt to gdp ratio along-side to understand the dynamics of fiscal deficit effect on green innovation at a macro-level.

So from (17) in short our analysis can be written as following:

$$gp_{it} = \beta_i^g gerd_{it} + \beta_i^h hc_{it} + \lambda_i' f_t + u_{it} \tag{28}$$

$$gp_{it} = \beta_i^b berd_{it} + \beta_i^h hc_{it} + \lambda_i' f_t + u_{it} \tag{29}$$

$$gp_{it} = \beta_i^g gerd_{it} + \beta_i^h hc_{it} + \beta_i^d dgdp_{it} + \lambda_i' f_t + u_{it} \tag{30}$$

$$gp_{it} = \beta_i^b berd_{it} + \beta_i^h hc_{it} + \beta_i^d dgdp_{it} + \lambda_i' f_t + u_{it} \tag{31}$$

where, i is the country index, t is the time index, gp is the ratio of green patents to total patents, gerd and berd are reseach and development stock for gross and business enterprise respectively, hc represents human capital for innovation outside the R&D sector and other aspects of human capital not captured by formal R&D. dgdp represents debto-to-gdp ratio. In the following section we explain the construction of the variables. The fixed effects are captured by the factor structure and the unobserved common patterns are modelled by the multi-factor error structure $\lambda_i' f_t + u_{it}$, assuming errors are cross-sectionally dependent with unobserved common patterns. We also consider $(\beta_i^g, \beta_i^b, \beta_i^h, \beta_i^d)$ as long-run cointegrating relations with latent group structures.

6. Results

6.1. A. Cross-sectional dependence test

We use Pesaran, 2015a, Bailey et al., 2016 and Ertur and Musolesi, 2017 to calculate the degree of the Cross-sectional Dependence statistic along with estimated confidence bands of α, the exponent of cross-sectional dependence defined over the range [0,1] for our required variables as depicted in Table 3, the null of the CD test depending upon the increase of T and N. When T is fixed and N$\longrightarrow \infty$, the null for CD test is given by $0 \leqslant \alpha \leqslant 0.5$ and when T and N$\longrightarrow \infty$ at the same rate, the null for CD test is given by $0 \leqslant \alpha \leqslant 0.25$ (which is our case). So, the value of α in the range of [0.5,1] depicts different degree of strong cross-sectional dependence and in between [0, 0.5] depicts different degree of weak cross-sectional dependence.

Table 2: CD Results

Variables	CD statistic	$\widehat{\alpha}_{0.5}$	$\widehat{\alpha}$	$\widehat{\alpha}_{0.95}$
gp	34.42	0.8202	0.8917	0.9633
gerd	85.01	0.9229	0.9801	1.0373
hc	107.98	0.9507	1.0033	1.0559
berd	84.34	0.9201	0.9759	1.0318
dgdp	39.78	0.84846	0.9111	0.9736

In our case for all the variables, the CD statistic strongly rejects the null hypothesis suggesting the fact that the exponent of cross-sectional dependence lies in the range [0.25, 1]. To figure out the degree of cross-sectional dependence, one has to look at the bias-corrected estimates of α and the 90% confidence bands around it. In our case the exponent of cross-sectional dependence is estimated at approximately one for all variables at levels and more than 0.90 for all variables in first differences. In addition, the 90% confidence bands are highly above 0.5 and include unity. This confirms our preliminary finding and suggests presence of strong cross-sectional dependence in both dependent and explanatory variables for our analysis.

6.2. B. Second-generation panel unit root tests

The literature related to Panel Unit root tests have evolved over time, Quah, 1994 and Breitung and Meyer, 1994 introduced unit root testing in panel framework which was based on similar analysis from time-series literature. The so-called first generation of panel unit root tests do not consider correlation in between cross-sectional error components which were developed by Levin et al., 2002and Im et al., 2003. The second-generation panel unit-root tests do consider of errors being cross-sectionally correlated. Three main approaches in this regard are, Maddala and Wu, 1999 and developed by various other authors thereafter, which applies bootstrapping to panel unit root test but this approach is mainly feasible for large T and relatively small N. Bai and Ng, 2004, Bai and Ng, 2010proposed to decompose the observed series into two unobserved components, common factors and idiosyncratic errors and test for unit roots in both of these components, the test is also known as PANIC (Panel Analysis of Nonstationarity in Idiosyncratic and Common Components) it provides indirect test for unit roots in observed series. The third approach was put forward by Pesaran, 2006, Pesaran, 2007 and extended by Pesaran et al., 2013, in Pesaran, 2006, Pesaran, 2007 a new test is proposed underlying the idea of cross-section average (CA) augmentation approach which augments individual Dickey-Fuller (DF) regressions with cross-section averages to take into account of error of cross-section dependence, then these cross-sectionally augmented DF regressions can be further augmented with lagged changes, to deal with possible serial correlation in the residuals. These doubly augmented DF regressions are referred to as CADF regressions. The panel unit root test statistic is then computed as the average of the CADF statistics. The average statistic is free of nuisance parameters but, due to non-zero cross correlation of the individual, CADF; statistics, the average statistic has a non-normal limit distribution as N and T tends towards infinity. Pesaran et al., 2013 extends this approach to the case of multi-factor error structure using Sargan-Bhargava type statistics. But the problem with CA approach is it is complicated to implement, since this test is implemented when testing for unit roots in test statistics with nonstandard asymptotic distributions. Recently, Reese and Westerlund, 2016 has put forward a new approach combining PANIC and CA, since PANIC approach uses Principal Component (PC) analysis, so it might present distorted results when N is small. PANICCA on the other hand leads to much improved small sample performance, when N is small or medium.

Weak cross-sectional dependence can be addressed with simple-correction of the tests but strong cross-sectional dependence causes the test statistic to be more divergent (Westerlund and Breitung, 2013). Pesaran et al., 2013 states, that the effect of cross-sectional dependence can be reduced by demeaning the data in first-generation unit root tests if the pair-wise errors covariances do not strongly digress across individuals.

Table 4 reports the outcomes of three first-generation unit root tests with cross-sectionally demeaned data Im et al., 2003test and Choi, 2001 the alternative P, Z, L* and Pm tests. The result shows the dependent variable gp is of stationary in nature and the independent variables gerd, berd, hc and dgdp are of non-stationary in nature i.e., being generated by unit root stochastic processes.

Table 3: First-generation unit root test**

Variables	IPS	P	Z	L*	P_M
gp	0.0000	0.0000	0.0000	0.0000	0.0000
gerd	0.0234	0.0007	0.0020	0.0010	0.0001
berd	0.0111	0.0002	0.0050	0.0008	0.0000
hc	0.9958	0.6462	0.9958	0.9954	0.6662
dgdp	0.4334	0.7327	0.9164	0.9070	0.7445

(**) p-values. Variables gp is in ratio terms and gerd and hc are in logarithmic terms

Due to the presence of strong cross-sectional dependence in our data we also employ second-generation unit root tests, these tests use multi-factor error structure using heterogeneous factor loadings to model various forms of cross-sectional dependence. We employ Pesaran, 2007 (CADF, CIPS)[Table: 5]; Bai and Ng, 2004 (PANIC) and Reese and Westerlund, 2016(PANICCA) [Table: 6] to investigate more in-depth sources of unit roots among the variables. PANIC decomposes each variable into deterministic, common and idiosyncratic components, so that the origin of the cause of non-stationarity can be traced i.e., whether it arises from common component or the idiosyncratic component or both.

Table 4: Second-generation unit root test- CADF, CIPS**

Variables	$CADF^+$	$CIPS^+$
gp	-2.108	-2.819
gerd	-1.907	-1.799
berd	-1.700	-1.981
hc	-1.614	-2.419
dgdp	-1.985	-1.786

[(+: statistics). Variables gp, in ratio terms and gerd and hc are in logarithmic terms.
Critical values, CADF: -2.080 (cv10), -2.160 (cv5) -2.300 (cv1) and CIPS: -2.04 (10%),-2.11(5%) -2.23 (1%)]

Bai and Ng, 2004 requires the number of common factors needed to represent the cross-sectional dependence, we assume only one common factor following Westerlund and Urbain, 2015 which indicates small number of unobserved common factors are sufficient enough to deal in macroeconomic examples. The test PANICCA is mix of both Bai and Ng, 2004 and Pesaran, 2007, in which they use Cross-sectional Averages instead of Principal component estimates as used by Bai and Ng, 2004 to proxy for factors by pooling individual ADF t statistics on defactored residuals to test for nonstationarity of the idiosyncratic components

Table 5: Second-generation unit root test- PANIC, PANICCA**

Variables			PANIC**	
	ADF	P_a	P_b	PMSB
gp	0.0001	0.1214	0.1415	0.3799
gerd	0.0001	0.997	1	1
berd	1	0.9728	1	1
hc	0.0001	0	0	0.0207
dgdp	0.0001	0.4945	0.4943	0.6574
			PANICCA**	
	ADF	P_a	P_b	PMSB
gp	0.0849	0	0.003	0.517
gerd	1	0.1805	0.1747	0.5852
berd	0.001	0.8257	0.8844	0.9569
hc	0.001	0	0	0.0231
dgdp	0.001	0.2752	0.2828	0.4458

(**) p-values. Variables gp is in ratio terms and gerd and hc are in logarithmic terms

6.3. C. Estimation results

6.3.1. i. Information criteria

To estimate the number of unobserved factors we employ the BIC type penalty function following (Bai and Ng, 2004 , SSP 2016, HPS 2018) and set $g_1(N,T) = (\frac{N+T}{(NT})log(min(N,T)$ to determine the total number of unobserved common factors and $g_2(N,T) = g_1(N,T) \times \frac{T}{4log(log(T)}$ to determine the number of unobserved non-stationary factors, where N = 25 and T = 44, we find the level and differenced indicates one unobserved common factor, which also verifies Westerlund and Urbain, 2015 which indicates small number of unobserved common factors are sufficient enough to deal in macroeconomic example.

6.3.2. ii. Determining the number of groups

Group selection is one of the most important criteria in this kind of estimation technique, we select the number of groups following previous literature, SSP 2016, HPS 2018. The exact number of groups are typically unknown but a finite integer K_{max} is assumed which is considered to be an upper bound to the true number of groups K_0. The tuning parameter is chosen as $\lambda = c\lambda \times T^{-3/4}$ where c takes five candidates 0.01, 0.02, 0.05, 0.10 and 0.20. We fix K_{max} arbitrarily at 6. For each combination of the number of groups and the tuning parameter c, we compute the information criterion value accordingly. The results are reported in table 7, 8, 9, 10.

Table 6: Information Criterion values: eq 29

	0.01	0.02	0.05	0.1	0.2
1	-5.9114	-5.9114	-5.9114	-5.8571	-5.9114
2	-5.9638	-5.9638	-5.8314	-5.7426	-5.7421
3	-5.8153	**-5.9981**	-5.6651	-5.5738	-5.5774
4	-5.7004	-5.6902	-5.5016	-5.4107	-5.4033
5	-5.5287	-5.5185	-5.4092	-5.2746	-5.2571
6	-5.3559	-5.3468	-5.2789	-5.1771	-5.1315

We choose 3 groups and set $c_\lambda = 0.02$, and apply it to C-Lasso technique for equation 29, since of the minimum value of I.C., **-5.9981** being accredited.

Table 7: Information Criterion values: eq 30

	0.01	0.02	0.05	0.1	0.2
1	-5.9725	-5.9726	-5.9726	-5.9726	-5.9726
2	-5.9572	-5.9748	-5.9745	-5.8346	-5.8147
3	-5.7877	**-5.9819**	-5.8055	-5.716	-5.7989
4	-5.6672	-5.6307	-5.5376	-5.6302	-5.626
5	-5.4709	-5.5265	-5.3779	-5.4815	-5.5305
6	-5.2175	-5.3425	-5.2062	-5.2822	-5.3926

We choose 3 groups and set $c_\lambda = 0.02$, and apply it to C-Lasso technique for equation 30, since of the minimum value of I.C., **-5.9819** being accredited. We choose 3 groups and set $c_\lambda = 0.05$, and apply it to C-Lasso technique for equation 31, since of the minimum value of I.C., **-5.6936**.

Table 8: Information Criterion values eq 31

	0.01	0.02	0.05	0.1	0.2
1	-5.4423	-5.4424	-5.4425	-5.4424	-5.4425
2	-5.6402	-5.6068	-5.5869	-5.5935	-5.5833
3	-5.6926	-5.6711	**-5.6936**	-5.4433	-5.3511
4	-5.1703	-5.1817	-5.1449	-5.1809	-5.2907
5	-4.9359	-5.1438	-5.1441	-4.8042	-4.9629
6	-4.6977	-4.7995	-4.8464	-4.5467	-4.6406

We choose 3 groups and set $c_\lambda = 0.05$, and apply it to C-Lasso technique for equation 32, since of the minimum value of I.C., **-5.8818**.

6.3.3. iii. Post Classifier Lasso results

Table 11, 13, 15, 17 report the Cup-Lasso estimates with one unobserved non-stationary common factors for equations 29, 30, 31, 32, we also report the group classification for each of the

Table 9: Information Criterion values eq 32

	0.01	0.02	0.05	0.1	0.2
1	-5.8007	-5.8007	-5.8008	-5.8008	-5.8005
2	-5.7747	-5.7749	-5.7778	-5.7218	-5.8466
3	-5.7405	-5.7335	**-5.8818**	-5.5135	-5.62
4	-5.1636	-5.1512	-5.2083	-5.3021	-5.325
5	-4.9347	-4.9935	-4.9531	-5.0539	-5.1125
6	-4.6685	-4.7024	-4.7174	-4.8427	-4.8334

equations in Table 12, 14, 16, 18. We will explain the results in a concise form in the next subsection.

Table 10: POST- Classifier-LASSO results : eq 29

	Group 1	Group 2	Group 3
gerd	0.005323312*** (0.007874287)	-0.041036067 (0.012613428)	-0.016168494 (0.004013536)
hc	-0.345069953 (0.057083789)	-0.449300837 (0.105114664)	0.032991704*** (0.034164233)

[Values inside parenthesis indicates values for standard errors. Symbols ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.]

For equation 29, when we apply C-Lasso to understand the effect of human capital (hc) and gross- expenditure of r&d (gerd) on green innovation measure through green patents, we find gerd significant for Group 1 with 10% level and positive value and hc being significant for Group 3 with 10% level and positive value, for other two groups in their category results are not significant with negative sign.

Table 11: GROUP MEMBERSHIP : eq 29

Group 1 membership = **12**	Australia, Austria, Denmark, Germany, Ireland, Israel, Korea, Mexico, Norway, Portugal, Switzerland, United States
Group 2 membership = **10**	Belgium, Canada, Finland, France, Italy, Japan, Netherlands, Spain, Sweden, United Kingdom,
Group 3 membership = **3**	Greece, Iceland, New Zealand

As mentioned earlier we found three groups with membership in each group being 12, 10 and 3 respectively. The members countries are given in table 12.

Table 12: POST- Classifier-LASSO results : eq 30

	Group 1	Group 2	Group 3
berd	0.146690311***	-0.004902611	-0.003437177
	(0.006696195)	(0.004394472)	(0.007063436)
hc	-3.724429116	0.052223711***	-0.296077921
	(0.136115107)	(0.037190415)	(0.059088037)

[Values inside parenthesis indicates values for standard errors. Symbols ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.]

For equation 30, when we apply C-Lasso to understand the effect of human capital (hc) and business- expenditure of r&d (berd) on green innovation measure through green patents, we find berd significant for Group 1 with 10% level and positive value and hc being significant for Group 2 with 10% level and positive value, for other two groups in their category results are not significant with negative sign.

Table 13: GROUP MEMBERSHIP : eq 30

Group 1 membership = **1**	Iceland
Group 2 membership = **17**	Australia, Austria, Belgium, Canada Denmark, Finland, France, Germany Italy, Japan, Netherlands, Norway Spain, Sweden, Switzerland United Kingdom, United States
Group 3 membership = **7**	Greece, Ireland, Israel Korea, Mexico New Zealand, Portugal

As mentioned earlier we found three groups with membership in each group being 1, 17 and 7 respectively. The members countries are given in table 14.

For equation 31, when we apply C-Lasso to understand the effect of human capital (hc) and gross- expenditure of r&d (gerd) and debt-gdp ratio (dgdp) on green innovation measure through green patents, we find gerd significant for Group 2 with 10% level and positive value and dgdp being significant for Group 1 and Group 3 with 10% level and positive value, for other groups in their category results are not significant with negative sign. For hc, in all three groups the coefficients are negative without any level of significance.

We found three groups with membership in each group being 10, 10 and 5 respectively. The members countries are given in table 16.

For equation 32, when we apply C-Lasso to understand the effect of human capital (hc)

Table 14: POST- CUP-LASSO results : eq 31

	Group 1	Group 2	Group 3
gerd	-0.018574315 (0.026929496)	0.007917289*** (0.008299823)	-0.01114053 (0.003216006)
hc	-0.438794061 (0.25208975)	-0.254873843 (0.05476204)	-0.012894626 (0.03031962)
dgdp	0.017674734*** (0.012685904)	-0.003638012 (0.011581591)	0.009035279*** (0.003569003)

[Values inside parenthesis indicates values for standard errors. Symbols ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.]

Table 15: GROUP MEMBERSHIP : eq 31

Group 1 membership = **10**	Austria, Denmark, Finland, Germany Greece, Iceland, Ireland Korea, New Zealand, Portugal
Group 2 membership = **10**	Australia, Belgium, Canada, Italy Japan, Netherlands, Norway Spain, United Kingdom, United States
Group 3 membership = **5**	France, Israel Mexico, Sweden, Switzerland

Table 16: POST- CUP-LASSO results : eq 32

	Group 1	Group 2	Group 3
berd	0.003003989*** (0.009507318)	0.001030705*** (0.004433412)	-0.027267814 (0.014607955)
hc	-0.186506212 (0.112361387)	0.01748895*** (0.0409458)	0.412704465*** (0.150333382)
dgdp	0.035235995*** (0.015161108)	0.005649531*** (0.004917282)	-0.064634375 (0.028137968)

[Values inside parenthesis indicates values for standard errors. Symbols ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.]

and business- expenditure of r&d (berd) and debt-gdp ratio (dgdp) on green innovation measure through green patents, we find berd significant for Group 1 and Group 2 with 10% level and positive values and dgdp being significant for Group 1 and Group 2 with 10% level and positive value, hc has positive coefficient with 10% level significance for Group 2 and Group 3. For other groups in their category results are not significant with negative signed coefficients.

Table 17: GROUP MEMBERSHIP : eq 31

Group 1 membership = **14**	Belgium, Canada, Denmark, Finland Greece, Ireland, Israel, Italy Japan, Netherlands, Norway Portugal, Sweden, Switzerland
Group 2 membership = **6**	Austria, Germany, Korea Mexico, United Kingdom United States
Group 3 membership = **5**	Australia, France Iceland, New Zealand, Spain

We found three groups with membership in each group being 14, 6 and 5 respectively. The members countries are given in table 18.

6.4. iv. Classification Results

Considering the fact **BERD** (Business Expenditure on Research and Development) is the biggest share on **GERD** (Gross Expenditure on Research and Development) OECD, 2015, results in Table 11 and 13 and their respective Group classifications in Table 10 and 12 are similar. When using the PPC based estimation methods, we find similar heterogeneous behaviour for model (29) and (30). Considering Iceland becomes the sole member of Group 1 in table 14 with significant berd is not surprising to the fact, due to the of its economic composition which is more based on private research and development. Infact, Iceland is one of the few countries which meets EU Barcelona target of 3% and the private sector accounts for more than 40% of R&D expenditure. In other cases in Table 12 and 14, results are nearly correlatable. In Table 12 and 14 Group 3 consists of countries which accounts for less than 1% of global R&D stock, but for Group 2 in Table 14 and Group 1 and 2 in Table 12, the membership is mostly of countries which are leaders in Innovation both Green and Non-green and also account for more than 60% of global R&D stock. When we introduce debt-gdp ratio in our analysis, the group membership does not vary a lot according to results in Table 15 and 17 and their respective Group classifications in Table 16 and 18 are similar. In this regard, it is very necessary to mention a strange clustering can be found in G7 countries in most of the cases, which is very accordance to Keller, 2004 which states major technical change leading to productivity growth in OECD countries are mostly originating from from abroad through channel of international technology diffusion.

In summary, while we try to model the green knowledge production function with Latent-Group Structures using PPC- base method with one unobserved global non-stationary factor, we find heterogeneous behaviour in green technologies using a Cup-Lasso estimate. Human capital and expenditure in Research and Development plays an important part in our findings.

7. Conclusion

The main contribution of this paper was to quantify the concept of Green Knowledge Production function for selected OECD countries. We assumed that cross-section dependence is generated by unobserved common factors which are both stationary and nonstationary in nature. Using a C-Lasso estimator as proposed by SSP (2016) and HPS (2018), we find distinctive groups with their coefficient being responsive and also non-responsive to green innovation.

Appendix

(A) We use OECD (OECD, 2018), we only use Development of environment-related technologies (**Y02** IPC-CPC class) as a ratio to All technologies. Which can be describes as to inventions or more jurisdictions (with family size 1 or greater) or in two or more jurisdictions (family size 2 or greater). The data is collected based on the definition of Paris Convention applicable to as the set of all patent applications protecting the same 'priority'.

(B) We use Penn World Table version 9.0 (Feenstra et al., 2016) to consider Human Capital data in our analysis, in short the data can be defined as educational attainment data across countries. For a better understanding refer to Feenstra et al., 2016

(C) The BERD and GERD data were collected from the OECD-STAT database, which is based on the OECD, 2015, we tried to extend Coe and Helpman, 1995 and Coe et al., 2009. The time-series for the countries were from 1970-2015 in terms of National Currency, some countries had some missing values. The missing time series values for each country are given below in Table 21.

In the sample of countries, Canada, France, Italy, Japan, Netherlands and USA had no missing observations, Germany and Spain had one missing observations in both BERD and GERD those were linearly interpolated. For other countries we followed the methodology from Coe et al., 2009.

OECD also supplies BERD data under ANBERD (Analytical Business Enterprise Research and Development) database which is under the STAN family database. The ANBERD data is of three versions ISIC Rev. 4 (1987 onwards), ISIC Rev.3 (1987-2010/11) and ISIC Rev. 2 (1973-1997/98) with the country-list expanding over time. For fifteen countries the business research development data is more than the BERD data. For these countries the correlation between ANBERD and BERD lies between 0.99 and 1. We interpolated the missing values of BERD with those available from ANBERD. This reduced the missing values for these countries indicated by the values inside the brackets in Table 21. For Austria, the GERD data is more than BERD data, and after the substitution from ANBERD for some of the BERD, the ratio of BERD/GERD was seem to be constant at 0.55 from 1970-1993 and increased from 0.63 to 0.71 in between 1998-2015, by using the ratio we linearly interpolated the missing values for Austria and 12 more values of BERD were filled bring down the missing years to 8.

We then converted the BERD and GERD values from national currency using the following formula,

$$BERD_{cppp} = [(BERD_{nc})/GDPP]/PPP_{2010}$$
$$GERD_{cppp} = [(GERD_{nc})/GDPP]/PPP_{2010}$$

Table .18: GERD and BERD: 1970-2014

Country	BERD	corr. BERD and ANBERD	GERD
Australia	12	0.99 (3)	24
Austria	24	1 (20) [8]	9
Belgium	6	0.99 (6)	9
Canada	0	-	0
Denmark	4	0.99 (4)	4
Finland	7	1(2)	7
France	0	-	0
Germany	1	-	1
Greece	20	1 (16)	22
Iceland	13	0.99 (10)	12
Ireland	6	0.99 (2)	6
Israel	21	1 (21)	21
Italy	0	-	0
Japan	0	-	0
Korea	25	1 (25)	21
Mexico	20	1 (20)	23
Netherlands	0	-	0
New Zealand	22	1 (13)	22
Norway	11	0.99 (2)	11
Portugal	7	1 (7)	7
Spain	1	-	1
Sweden	17	0.99 (8)	17
Switzerland	25	0.99 (11)	25
UK	10	0.99 (2)	10
USA	0	-	0

where (BERD_{cppp} and GERD_{cppp}) are data transformed to PPP exchange rates, PPP_{2010} is the purchasing power parity exchange rate in local currency per US dollar in 2010 and GDPP is the GDP price deflator, 2010=100.

Remaining missing observations were estimated using OLS prediction, of berd on gdpbv and ibv. Where berd is natural logarithm value of BERD_{cppp} and gdpbv is natural logarithm value of real value added in business sector, ibv is logarithm value of real non-residential private investment, these data of gdpbv and ibv were collected from World Bank data archive. The predicted values were collected from these regressions, which had R^2 were in between 0.95-0.99. This way the missing values of BERD_{cppp} were filled.

For GERD_{cppp}, we again used predicted values from OLS to fill up missing values, using gerd on rb, gdpv and itv. Where gerd is natural logarithm value of GERD_{cppp}, berd is natural logarithm value of BERD_{cppp}, gdpv is real value-added and itv is real private investment. The R^2 in these case were mostly above 0.98, and so the predicted values were used to fill missing GERD_{cppp}.

To calculate the stock of R&D for both GERD_{cppp} and BERD_{cppp} we used perpetual inventory rate as proposed by Coe and Helpman, 1995 with a depreciation rate of δ to be 0.05 or 5% level.

References

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous (2012). "The Environment and Directed Technical Change". In: *The American Economic Review* 102.1, pp. 131–166.

Aghion, Philippe and Peter Howitt (1992). "A Model of Growth through Creative Destruction". In: *Econometrica* 60.2, pp. 323–51.

Aghion, Philippe, Peter Howitt, Maxine Brant-Collett, and Cecilia Garca-Pealosa (1998). *Endogenous growth theory*. Cambridge, Mass: MIT Press.

Aigner, Dennis J. and Arthur Stanley Goldberger, eds. (1977). *Latent variables in socio-economic models*. Contributions to economic analysis 103. Amsterdam : New York: North-Holland Pub. Co. ; sole distributors for the U.S.A. and Canada, American Elsevier Pub. Co. 383 pp.

Ali, Abdilahi and Baris Alpaslan (2017). "Is There an Investment Motive Behind Remittances?: Evidence From Panel Cointegration". In: *The Journal of Developing Areas* 51.1, pp. 63–82.

Ando, Tomohiro and Jushan Bai (2016). "Panel Data Models with Grouped Factor Structure Under Unknown Group Membership:" in: *Journal of Applied Econometrics* 31.1, pp. 163–191.

Ang, James B. (2011). "Financial development, liberalization and technological deepening". In: *European Economic Review* 55.5, pp. 688–701.

Arrow, Kenneth J. (1962). "The Economic Implications of Learning by Doing". In: *The Review of Economic Studies* 29.3, p. 155.

Athey, Susan (2018). "The Impact of Machine Learning on Economics". In: *The Economics of Artificial Intelligence: An Agenda*. National Bureau of Economic Research, Inc.

Bai, Jushan (2009). "Panel Data Models With Interactive Fixed Effects". In: *Econometrica* 77.4, pp. 1229–1279.

Bai, Jushan and Serena Ng (2002). "Determining the Number of Factors in Approximate Factor Models". In: *Econometrica* 70.1, pp. 191–221.

— (2004). "A PANIC Attack on Unit Roots and Cointegration". In: *Econometrica* 72.4, pp. 1127–1177.

— (2010). "Panel Unit Root Tests with Cross-Section Dependence: A Further Investigation". In: *Econometric Theory* 26.4, pp. 1088–1114.

Bailey, Natalia, George Kapetanios, and M. Hashem Pesaran (2016). "Exponent of CrossSectional Dependence: Estimation and Inference". In: *Journal of Applied Econometrics* 31.6, pp. 929–960.

Ballester, Coralio, Antoni Calv-Armengol, and Yves Zenou (2006). "Who's Who in Networks. Wanted: The Key Player". In: *Econometrica* 74.5, pp. 1403–1417.

Barbieri, Nicol, Claudia Ghisetti, Marianna Gilli, Giovanni Marin, and Francesco Nicolli (2016). "A SURVEY OF THE LITERATURE ON ENVIRONMENTAL INNOVATION BASED ON MAIN PATH ANALYSIS: A SURVEY OF THE LITERATURE ON ENVIRONMENTAL INNOVATION". In: *Journal of Economic Surveys* 30.3, pp. 596–623.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). "Inference on Treatment Effects after Selection among High-Dimensional Controls". In: *The Review of Economic Studies* 81.2, pp. 608–650.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2011). "Inference for High-Dimensional Sparse Econometric Models". In: *arXiv:1201.0220 [econ, stat]*.

Belloni, Alexandre, Victor Chernozhukov, and Christian B. Hansen (2013). "Inference for High-Dimensional Sparse Econometric Models". In: *Advances in Economics and Econometrics*. Ed. by Daron Acemoglu, Manuel Arellano, and Eddie Dekel. Cambridge: Cambridge University Press, pp. 245–295.

Bonaldi, Pietro, Ali Hortasu, and Jakub Kastl (2015). *An Empirical Analysis of Funding Costs Spillovers in the EURO-zone with Application to Systemic Risk*. w21462. Cambridge, MA: National Bureau of Economic Research.

Bonhomme, Stphane and Elena Manresa (2015). "Grouped Patterns of Heterogeneity in Panel Data: Grouped Patterns of Heterogeneity". In: *Econometrica* 83.3, pp. 1147–1184.

Bonhomme, Stphane, Koen Jochmans, and Jean-Marc Robin (2016). "Estimating multivariate latent-structure models". In: *The Annals of Statistics* 44.2, pp. 540–563.

Bottazzi, Laura and Giovanni Peri (2007). "The International Dynamics of R&D and Innovation in the Long Run and in the Short Run". In: *The Economic Journal* 117.518, pp. 486–511.

Breiman, Leo (1995). "Better Subset Regression Using the Nonnegative Garrote". In: *Technometrics* 37.4, p. 373.

Breitung, Jrg and Wolfgang Meyer (1994). “Testing for unit roots in panel data: are wages on different bargaining levels cointegrated?” In: *Applied Economics* 26.4, pp. 353–361.
Browning, Martin and Jesus M. Carro (2013). “The Identification of a Mixture of First-Order Binary Markov Chains*:” in: *Oxford Bulletin of Economics and Statistics* 75.3, pp. 455–459.
Charlot, Sylvie, Riccardo Crescenzi, and Antonio Musolesi (2015). “Econometric modelling of the regional knowledge production function in Europe”. In: *Journal of Economic Geography* 15.6, pp. 1227–1259.
Chellaraj, Gnanaraj, Keith E. Maskus, and Aaditya Mattoo (2008). “The Contribution of International Graduate Students to US Innovation”. In: *Review of International Economics* 16.3, pp. 444–462.
Chernozhukov, Victor, Christian Hansen, and Martin Spindler (2015). “Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach”. In: *Annual Review of Economics* 7.1, pp. 649–688.
Choi, In (2001). “Unit root tests for panel data”. In: *Journal of International Money and Finance* 20.2, pp. 249–272.
Coe, David T. and Elhanan Helpman (1995). “International R&D spillovers”. In: *European Economic Review* 39.5, pp. 859–887.
Coe, David T., Elhanan Helpman, and Alexander W. Hoffmaister (2009). “International R&D spillovers and institutions”. In: *European Economic Review* 53.7, pp. 723–741.
Coe, Neil M. (2005). “Putting knowledge in its placea review essay”. In: *Journal of Economic Geography* 5.3, pp. 381–384.
Conley, Timothy G and Christopher R Udry (2010). “Learning about a New Technology: Pineapple in Ghana”. In: *American Economic Review* 100.1, pp. 35–69.
Costantini, Valeria, Francesco Crespi, Giovanni Marin, and Elena Paglialunga (2017). “Eco-innovation, sustainable supply chains and environmental performance in European industries 1 1We gratefully acknowledge the support by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 649186 ISIGrowth. The comments and suggestions by three anonymous referees are also acknowledged. The usual disclaimers apply.” In: *Journal of Cleaner Production* 155, pp. 141–154.
Crepon, Bruno, Emmanuel Duguet, and Jacques Mairessec (1998). “Research, Innovation And Productivi[Ty: An Econometric Analysis At The Firm Level”. In: *Economics of Innovation and New Technology* 7.2, pp. 115–158.
Dakhli, Mourad and Dirk De Clercq (2004). “Human capital, social capital, and innovation: a multi-country study”. In: *Entrepreneurship & Regional Development* 16.2, pp. 107–128.
De Giorgi, Giacomo and Michele Pellizzari (2014). “Understanding Social Interactions: Evidence from the Classroom”. In: *The Economic Journal* 124.579, pp. 917–953.
De Giorgi, Giacomo, Anders Frederiksen, and Luigi Pistaferri (2016). *Consumption Network Effects*. w22357. Cambridge, MA: National Bureau of Economic Research.
Dominicis, Laura de, Raymond J.G.M. Florax, and Henri L.F. de Groot (2013). “Regional clusters of innovative activity in Europe: are social capital and geographical proximity key determinants?” In: *Applied Economics* 45.17, pp. 2325–2335.
Donoho, D. L. and M. Elad (2003). “Optimally sparse representation in general (nonorthogonal) dictionaries via 1 minimization”. In: *Proceedings of the National Academy of Sciences* 100.5, pp. 2197–2202.
Donoho, David L., Iain M. Johnstone, Grard Kerkyacharian, and Dominique Picard (1996). “Density estimation by wavelet thresholding”. In: *The Annals of Statistics* 24.2, pp. 508–539.
Doran, Howard E. and Jan Kmenta (1986). “A Lack-of-Fit Test for Econometric Applications to Cross-Section Data”. In: *The Review of Economics and Statistics* 68.2, p. 346.
Eberhardt, Markus, Christian Helmers, and Hubert Strauss (2013). “Do Spillovers Matter When Estimating Private Returns to R&D?” In: *Review of Economics and Statistics* 95.2, pp. 436–448.
Ertur, Cem and Antonio Musolesi (2017). “Weak and Strong CrossSectional Dependence: A Panel Data Analysis of International Technology Diffusion”. In: *Journal of Applied Econometrics*. Journal of Applied Econometrics 32.3, pp. 477–503.
Fan, Jianqing and Runze Li (2001). “Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties”. In: *Journal of the American Statistical Association* 96.456, pp. 1348–1360.
Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (2016). *Penn World Table 9.0*. Groningen Growth and Development Centre.

Fisher-Vanden, Karen, David Popp, and Ian Sue Wing (2014). “Introduction to the Special Issue on Climate Adaptation: Improving the connection between empirical research and integrated assessment models”. In: *Energy Economics* 46, pp. 495–499.
Fu, Wancong, Chong Li, Jan Ondrich, and David Popp (2018). *Technological Spillover Effects of State Renewable Energy Policy: Evidence from Patent Counts*. w25390. Cambridge, MA: National Bureau of Economic Research.
Gagliardi, Luisa, Giovanni Marin, and Caterina Miriello (2016). “The greener the better? Job creation effects of environmentally-friendly technological change”. In: *Industrial and Corporate Change* 25.5, pp. 779–807.
Gilli, Marianna and Massimiliano Mazzanti (2019). *Contextualising Sustainability: Socio-Economic Dynamics, Technology and Policies*. SEEDS, Sustainability Environmental Economics and Dynamics Studies.
Greene, William H. (1990). *Econometric analysis*. New York : London: Macmillan ; Collier Macmillan. 783 pp.
Griliches, Zvi (1974). “Errors in Variables and Other Unobservables”. In: *Econometrica* 42.6, pp. 971–998.
— (1979). “Issues in Assessing the Contribution of Research and Development to Productivity Growth”. In: *The Bell Journal of Economics* 10.1, pp. 92–116.
— (1998). *R&D and productivity: the econometric evidence*. A National Bureau of Economic Research monograph. Chicago: University of Chicago Press.
Grossman, Gene and Alan Krueger (1991). *Environmental Impacts of a North American Free Trade Agreement*. w3914. Cambridge, MA: National Bureau of Economic Research.
Hahn, Jinyong and Hyungsik Roger Moon (2010). “Panel Data Models with Finite Number Of Multiple Equilibria”. In: *Econometric Theory* 26.3, pp. 863–881.
Hall, Bronwyn H. and Jacques Mairesse (1995). “Exploring the relationship between R&D and productivity in French manufacturing firms”. In: *Journal of Econometrics* 65.1, pp. 263–293.
Hall, Bronwyn H., Jacques Mairesse, and Pierre Mohnen (2010). “Measuring the Returns to R&D”. In: *Handbook of the Economics of Innovation*. Vol. 2. Elsevier, pp. 1033–1082.
Hepburn, Cameron, Jacquelyn Pless, and David Popp (2018). “Policy BriefEncouraging Innovation that Protects Environmental Systems: Five Policy Proposals”. In: *Review of Environmental Economics and Policy* 12.1, pp. 154–169.
Hsiao, Cheng and A. Kamil Tahmiscioglu (1997). “A Panel Analysis of Liquidity Constraints and Firm Investment”. In: *Journal of the American Statistical Association* 92.438, pp. 455–465.
Huang, Wenxin, Sainan Jin, and Liangjun Su (2018a). “Identifying Latent Grouped Patterns in Cointegrated Panels”. In: *Singapore Management University- Working Paper*.
Huang, Wenxin, Peter C. B. Phillips, and Liangjun Su (2018b). “Nonstationary Panel Models with Latent Group Structures and Cross-Section Dependence”. In: *Singapore Management University Working Paper*.
Im, Kyung So, M.Hashem Pesaran, and Yongcheol Shin (2003). “Testing for unit roots in heterogeneous panels”. In: *Journal of Econometrics* 115.1, pp. 53–74.
Johnstone, Nick, Ivan Hai, and David Popp (2010). “Renewable Energy Policies and Technological Innovation: Evidence Based on Patent Counts”. In: *Environmental and Resource Economics* 45.1, pp. 133–155.
Kasahara, Hiroyuki and Katsumi Shimotsu (2009). “Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices”. In: *Econometrica* 77.1, pp. 135–175.
Keller, Wolfgang (2004). “International Technology Diffusion”. In: *Journal of Economic Literature* 42.3, pp. 752–782.
Kim, Jung Eun, David Popp, and Andrew Prag (2013). “The Clean Development Mechanism and neglected environmental technologies”. In: *Energy Policy* 55, pp. 165–179.
Kmenta, J. (1991). “Latent variables in econometrics”. In: *Statistica Neerlandica* 45.2, pp. 73–84.
Kmenta, Jan (1997). *Elements of econometrics*. 2nd ed. Ann Arbor: University of Michigan Press.
Koopmans, Tjalling C. (1949). “Identification Problems in Economic Model Construction”. In: *Econometrica* 17.2, p. 125.
Lam, Clifford and Pedro Souza (2014). *Regularization for Spatial Panel Time Series Using the Adaptive LASSO*. Suntory, Toyota International Centres for Economics, and Related Disciplines, LSE.
Lee, Kevin, M. Hashem Pesaran, and Ron Smith (1997). “Growth and convergence in a multi-country empirical stochastic Solow model”. In: *Journal of Applied Econometrics* 12.4, pp. 357–392.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu (2002). “Unit root tests in panel data: asymptotic and finite-sample properties”. In: *Journal of Econometrics* 108.1, pp. 1–24.
Lin, Chang-Ching and Serena Ng (2012). “Estimation of Panel Data Models with Parameter Heterogeneity when Group Membership is Unknown”. In: *Journal of Econometric Methods* 1.1.
Liu, Xiaodong, Eleonora Patacchini, Yves Zenou, and Lung-Fei Lee (2012). *Criminal Networks: Who is the Key Player?* Fondazione Eni Enrico Mattei.
Lu, Xun and Liangjun Su (2017). “Determining the number of groups in latent panel structures with an application to income and democracy: Number of groups in latent panel structures”. In: *Quantitative Economics* 8.3, pp. 729–760.
Lucas, Robert E. (1988). “On the mechanics of economic development”. In: *Journal of Monetary Economics* 22.1, pp. 3–42.
Maddala, G. S. and Shaowen Wu (1999). “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test”. In: *Oxford Bulletin of Economics and Statistics* 61 (s1), pp. 631–652.
Mairesse, Jacques and Pierre Mohnen (2010). *Using Innovations Surveys for Econometric Analysis*. w15857. Cambridge, MA: National Bureau of Economic Research.
Manresa, Elena (2016a). “Estimating the Structure of Social Interactions Using Panel Data”. In: *MIT Sloan Working Paper*, p. 63.
— (2016b). “Supplementary Appendix to Estimating the Structure of Social Interactions Using Panel Data”. In: *MIT Sloan Working Paper*, p. 42.
Manski, Charles F. (1993). “Identification of Endogenous Social Effects: The Reflection Problem”. In: *The Review of Economic Studies* 60.3, p. 531.
Marin, Giovanni and Antonello Zanfei (2019). “Does host market regulation induce cross-border environmental innovation?” In: *The World Economy*.
McCann, P. and R. Ortega-Argiles (2013). “Modern regional innovation policy”. In: *Cambridge Journal of Regions, Economy and Society* 6.2, pp. 187–216.
Meinshausen, Nicolai and Peter Bhlmann (2006). “High-dimensional graphs and variable selection with the Lasso”. In: *The Annals of Statistics* 34.3, pp. 1436–1462.
Moon, Hyungsik Roger and Martin Weidner (2015). “Linear Regression for Panel With Unknown Number of Factors as Interactive Fixed Effects”. In: *Econometrica* 83.4, pp. 1543–1579.
OECD (2015). *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*. The Measurement of Scientific, Technological and Innovation Activities. OECD.
— (2018). *Technology indicators*. OECD.
Parry, Ian W. H., Karen Pittel, and Herman R. J. Vollebergh, eds. (2017). *Energy tax and regulatory policy in Europe: reform priorities*. CESifo seminar series. Cambridge, MA ; London, England: MIT Press. 359 pp.
Pesaran, M. Hashem (2006). “Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure”. In: *Econometrica* 74.4, pp. 967–1012.
— (2007). “A simple panel unit root test in the presence of cross-section dependence”. In: *Journal of Applied Econometrics* 22.2, pp. 265–312.
— (2015a). “Testing Weak Cross-Sectional Dependence in Large Panels”. In: *Econometric Reviews* 34.6, pp. 1089–1117.
— (2015b). *Time series and panel data econometrics*. First edition. Oxford, United Kingdom: Oxford University Press.
Pesaran, M. Hashem, Yongcheol Shin, and Ron P. Smith (1999). “Pooled Mean Group Estimation of Dynamic Heterogeneous Panels”. In: *Journal of the American Statistical Association* 94.446, pp. 621–634.
Pesaran, M. Hashem, L. Vanessa Smith, and Takashi Yamagata (2013). “Panel unit root tests in the presence of a multifactor error structure”. In: *Journal of Econometrics* 175.2, pp. 94–115.
Phillips, Peter C.B. and Donggyu Sul (2007). “Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence”. In: *Journal of Econometrics* 137.1, pp. 162–188.
Popp, David (2017). “From science to technology: The value of knowledge from different energy research institutions”. In: *Research Policy* 46.9, pp. 1580–1594.

Popp, David, Nidhi Santen, Karen Fisher-Vanden, and Mort Webster (2013). "Technology variation vs. R&D uncertainty: What matters most for energy patent success?" In: *Resource and Energy Economics* 35.4, pp. 505–533.
Quah, Danny (1994). "Exploiting cross-section variation for unit root inference in dynamic data". In: *Economics Letters* 44.1, pp. 9–19.
Rebelo, Sergio (1991). "Long-Run Policy Analysis and Long-Run Growth". In: *Journal of Political Economy* 99.3, pp. 500–521.
Reese, Simon and Joakim Westerlund (2016). "Panicca: Panic on Cross-Section Averages: PANICCA: PANIC ON CROSS-SECTION AVERAGES". In: *Journal of Applied Econometrics* 31.6, pp. 961–981.
Romer, Paul M. (1986). "Increasing Returns and Long-Run Growth". In: *Journal of Political Economy* 94.5, pp. 1002–1037.
Sarafidis, Vasilis and Neville Weber (2015). "A Partially Heterogeneous Framework for Analyzing Panel Data". In: *Oxford Bulletin of Economics and Statistics* 77.2, pp. 274–296.
Schumpeter, Joseph A. (1939). *Business cycles: a theoretical, historical, and statistical analysis of the capitalist process*. 1. ed., [reprint]. Chevy Chase, Md.: Bartleby's Books [u.a.]
Schumpeter, Joseph Alois (1942). *Capitalism, socialism, and democracy*.
Shen, Xiaotong and Jianming Ye (2002). "Adaptive Model Selection". In: *Journal of the American Statistical Association* 97.457, pp. 210–221.
Society, Econometric, Daron Acemoglu, Manuel Arellano, and Eddie Dekel, eds. (2010). *Advances in economics and econometrics*. Econometric Society monographs 49-51. Cambridge ; New York: Cambridge University Press. 3 pp.
Stefanski, L. A., Yichao Wu, and Kyle White (2014). "Variable Selection in Nonparametric Classification Via Measurement Error Model Selection Likelihoods". In: *Journal of the American Statistical Association* 109.506, pp. 574–589.
Stokey, Nancy (1998). "Are There Limits to Growth?" In: *International Economic Review* 39.1, pp. 1–31.
Su, Liangjun and Qihui Chen (2013). "Testing Homogeneity in PanelData Models with Interactive Fixed Effects". In: *Econometric Theory* 29.6, pp. 1079–1135.
Su, Liangjun and Gaosheng Ju (2018). "Identifying latent grouped patterns in panel data models with interactive fixed effects". In: *Journal of Econometrics* 206.2, pp. 554–573.
Su, Liangjun, Zhentao Shi, and Peter C. B. Phillips (2016). "Identifying Latent Structures in Panel Data". In: *Econometrica* 84.6, pp. 2215–2264.
Sun, Yixiao X (2005). *Estimation and Inference in Panel Structure Models*. Department of Economics, UC San Diego.
Tibshirani, Robert (1996). "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1, pp. 267–288.
Verdolini, Elena and Valentina Bosetti (2017). "Environmental Policy and the International Diffusion of Cleaner Energy Technologies". In: *Environmental and Resource Economics* 66.3, pp. 497–536.
Vona, Francesco, Giovanni Marin, Davide Consoli, and David Popp (2018). "Environmental Regulation and Green Skills: An Empirical Exploration". In: *Journal of the Association of Environmental and Resource Economists* 5.4, pp. 713–753.
Wang, Wuyi, Peter C.B. Phillips, and Liangjun Su (2019). "The heterogeneous effects of the minimum wage on employment across states". In: *Economics Letters* 174, pp. 179–185.
Westerlund, Joakim and Jrg Breitung (2013). "Lessons from a Decade of IPS and LLC". In: *Econometric Reviews* 32.5, pp. 547–591.
Westerlund, Joakim and Jean-Pierre Urbain (2015). "Cross-sectional averages versus principal components". In: *Journal of Econometrics* 185.2, pp. 372–377.
Zou, Hui (2006). "The Adaptive Lasso and Its Oracle Properties". In: *Journal of the American Statistical Association* 101.476, pp. 1418–1429.