Green Technologies, Environmental Policy and Regional Growth

Philip Kerner[a,*], Torben Klarl[a,b], and Tobias Wendler[a]

[a]University of Bremen, Faculty of Business Studies and Economics, Max-von-Laue-Straße 1, 28359 Bremen, Germany

[b]SPEA, Indiana University, Bloomington, United States

Monday 31[st] January, 2022

Abstract

Green technologies are at the core of endeavors to combine economic and environmental targets. In this article, we aim to determine the impact of green technology development on total factor productivity of European regions. We advance methodologically on the pertinent literature by accounting for cross-sectional dependence in our empirical approach. Additionally, we provide a theoretical framework to link our results to implications of environmental policies for capital accumulation. Our results for 270 European regions imply that general technology development is associated with positive economic returns, but our data is not supportive of positive economic returns to green technologies.

Keywords: Regional Growth, Green Technologies, Environmental Policy, Cross-Sectional Dependence

JEL Classification: C23, O1, O33

*Corresponding author: kerner@uni-bremen.de. Earlier versions of this paper were presented at the European Association of Environmental and Resource Economists (EAERE) Annual Conference on the 25th of June 2021, the Scottish Economic Society (SES) Annual Conference on the 28th of April 2021, the workshop "Regional Inequality in Europe and the United States – Are there New Empirical Measures and Policy Approaches?" on the 29th of October 2020 and at the HSU Empirical Research in Economics Seminar at the Helmut Schmidt University in Hamburg on the 30th of January 2020. We thank all organizers and participants for valuable comments. The card material presented in this paper is based on the shapefiles from Eurostat, ©EuroGeographics for the administrative boundaries.

1 Introduction

Green technologies are at the very core of endeavors to combine economic and environmental targets to achieve sustainable growth, one of the aims of the *European Green Deal* (European Commission, 2019). First, green technical progress might substantially contribute to increase environmental productivity (e.g. Popp, 2010). At the same time, green technologies might enhance economic productivity (e.g. Xepapadeas and de Zeeuw, 1999). If green technologies are indeed fostering economic productivity, they can serve to stimulate regional growth and perhaps be a tool for regional inclusion. Indeed, technological progress provides the foundation of Europe's regional development strategies (e.g. McCann and Ortega-Argilés, 2015). In this article, we aim to determine the impact of green technology development on total factor productivity of European regions.

Our paper contributes to the literature on technological change and regional growth[1] in various ways. First, our paper is, to the best of our knowledge, the first to assess the specific role of green technologies for regional growth on a broad empirical base. Second, we advance methodologically on the pertinent literature by explicitly accounting for cross-sectional dependence (CSD) in our empirical approach. Third, by providing a simple theoretical framework, we directly link our results to implications of environmental policies for capital accumulation and composition dynamics, contributing to the ongoing debate revolving around the strong version of the Porter hypothesis (Porter and van der Linde, 1995).

We focus on the economic returns that occur within the same region the technology is developed, which we call the private returns. This contrasts the public returns that include potential positive influences on neighboring regions that occur, for example, through knowledge spillovers, a result of the public goods nature of knowledge (Keller, 2004). The analysis of the private returns to green technological knowledge has important policy dimensions. It gives insights whether policies promoting regional green technology development also promote economic development and competitiveness of regions, and hence whether they contribute to both green and perhaps inclusive growth.

To guide our empirical analysis, we develop a simple theoretical growth model with local (regional) knowledge and environmental externalities. Using this simple framework, we emphasize the relevance of the aggregate output elasticities of both green and non-green (polluting)[2] knowledge capital by exemplifying how they influence the effects of an environmental tax in the short-, medium- and long-run. Additionally, the theoretical foundation provides a possible interpretation of the identified productivity parameters. In our empirical approach, we estimate the implied aggregate production function of the model economy for European regions to get consistent estimates of the output elasticities.

Our empirical analysis builds upon a panel of 270 European NUTS-2 regions in 28 countries for the period 1991-2015. By relying on the flexible common correlated effects (CCE) approach (Pesaran, 2006), we are able to effectively control for different forms of CSD and other challenges in the estimation of production functions, such as heterogeneous impacts of green technologies on output between regions. Additionally, we employ various alternative estimation techniques to get a comprehensive view. Our main results comprise the following. First, we highlight the importance to account for CSD between European regions in the variables of the production function. Second, while general technology development is mostly associated with positive economic returns, our data is not supportive of positive economic returns to green technologies. This insight is generally robust for all applied estimation procedures and a battery of econometric extensions.

Empirical studies have mostly been conducted on the firm or sector-country level. Firm level evidence points to lower returns of environmentally friendly innovation compared to other innovation (Marin and Lotti, 2017) or positive effects only for specific types of green technologies (resource-saving) (Ghisetti and Rennings, 2014; Rexhäuser and Rammer, 2014; Van Leeuwen and Mohnen, 2017). Sector-country level evidence suggests positive, albeit rather small returns (Stucki and Woerter, 2019) and a possibly U-shaped relationship between green knowledge and productivity (Soltmann et al., 2015; Stucki and Woerter, 2019). To the best of our knowledge, a broad based empirical analysis on the regional level is yet missing.

Furthermore, our empirical approach directly builds upon the ever more growing econometric literature on cross-sectional dependence (CSD) in panel estimation. Appropriately accounting for CSD is especially important in the empirical setup at hand, as above mentioned knowledge spillovers and unobserved common shocks make it a very likely feature of the data. In this regard, Ertur and Musolesi (2017) highlight the importance to account for CSD, even if potential channels of knowledge spillovers are explicitly controlled for. Furthermore, Mitze et al. (2016) and Eberhardt et al. (2013) detect only limited returns to general knowledge capital at the industry level, when unobserved spillovers and common factors are accounted for.

The remainder of the paper is organized as follows. In Section 2 we sketch a simple growth model with knowledge externalities and environmental externalities. In Section 3 the empirical framework is outlined in detail, with a focus on estimating CSD and capturing potential CSD in the main model. Section 4 contains a detailed description and discussion of the data. Section 5 provides the empirical results of the tests for CSD, the main model, a battery of robustness checks and a discussion. Finally, Section 6 concludes.

2 Theoretical Framework

In this section, we outline a simple dynamic general equilibrium growth model, which shall provide a foundation for our empirical investigation of the private returns to green knowledge capital. In our model, the population of N identical agents grows exponentially at the steady rate n. We assume that each agent has an infinite planning horizon, possesses perfect foresight and maximizes lifetime utility from consumption. Moreover, each agent produces a single output that can be costlessly transformed into green (g) or polluting (p) knowledge capital investment and into a consumption good. At each point in time, and given the stocks of green and polluting knowledge capital, each agent inelastically supplies one unit of labor and optimally chooses the rate of consumption and investment. A specific feature of this model is that the accumulation of polluting knowledge capital negatively affects environmental quality that in turn affects agents' productivity. On the other hand, the accumulation of both sorts of knowledge increases the economy-wide knowledge stock, which in turn positively affects productivity.[3]

2.1 Production

Output of the individual firm, indexed by i, is determined by the Cobb-Douglas production function:

$$Y_i = K_i^{\alpha} K^{\beta} E^{\mu}, \quad \alpha, \beta, \mu > 0 \tag{2.1}$$

where $K_i \equiv K_{g,i}^{\kappa} K_{p,i}^{1-\kappa}$ with $\kappa \in [0,1]$.[4] Both forms of knowledge capital are private goods. Further note that, ex-ante, there are no restrictions imposed on returns to scale.

Production is further subject to two externalities: First, through the existence of an economy-wide, aggregate stock of knowledge, K, that is related to the individual knowledge stocks by $K_i = \frac{K}{N}$. This externality can be interpreted as knowledge spillovers in the spirit of Romer (1986). Second, the economy-wide, aggregate environmental quality, E, which is taken as given by the single producer, affects output as well.

2.2 Consumer Optimization

The representative agent's welfare is given by the intertemporal, isoelastic utility function:

$$\Omega \equiv \left(\frac{1}{1-\theta}\right) \int_0^{\infty} [C_i]^{(1-\theta)} e^{-(\rho-n)t} dt, \ \rho - n > 0, \tag{2.2}$$

with $n \geq 0$ and $\theta > 1$. $C_i = \frac{C}{N}$ denotes the consumption per worker.

In performing the optimization, the agent is constrained by the following flow budget constraint:

$$\dot{K_{p,i}} = [r_p(1-\xi_p) - \delta_p - n]K_{p,i} + r_g(1+\xi_g)K_{g,i} + w - C_i - \tau I_i - T_i, \tag{2.3}$$

where r_p and r_g are the gross return to polluting respectively green knowledge capital, w is the wage rate, ξ_p is the tax on polluting knowledge capital, ξ_g is the subsidy on green knowledge, $T_i \equiv \frac{T}{N}$ is the agent's share of lump-sum taxes (transfers if T is negative).

$\tau \in [0,1]$ is the fraction of investment that goes into green knowledge, i.e. $I_g = \tau I$, while $(1-\tau)I = I_p$ represents polluting knowledge investments. The rate of accumulation of green knowledge is given by:

$$\dot{K_{g,i}} = \tau I_i - (\delta_g + n)K_{g,i}. \tag{2.4}$$

2.3 Government

The government imposes a Pigouvian tax subsidy scheme on the rental income aiming to correct for the externalities. In other words, maintaining the assumption of a balanced budget at each point in time, the tax revenue exactly covers green subsidies and lump-sum transfers T_t to the private sector. As the tax and subsidy rates are given, T_t adjust so as to balance the governmental budget. Thus, the government budget constraint is

$$\xi_p r_p K_p + T = \xi_g r_g K_g. \tag{2.5}$$

2.4 Environmental Quality

We assume that polluting knowledge capital negatively impacts environmental quality, while green knowledge capital contributes in a positive way to aggregate environmental quality, E. Thus, we assume the simple relationship (see Klarl, 2016)

$$E = K_g^{\phi_g} K_p^{-\phi_p},\ \phi_g, \phi_p > 0. \tag{2.6}$$

By combining the individual production functions with the equations for the externalities, aggregate production reads as

$$Y = K_g^{\sigma_g} K_p^{\sigma_p} N^{\sigma_n}, \tag{2.7}$$

with $\sigma_g = \kappa(\alpha+\beta) + \phi_g\mu$, $\sigma_p = (1-\kappa)(\alpha+\beta) - \phi_p\mu$ and $\sigma_n = 1-\alpha$.

2.5 Decentralized Equilibrium

The agent chooses the rate of consumption to maximize (2.2) subject to (2.3), (2.4). The first order conditions for an optimum are

$$C_i : (C_i)^{-\theta} = \lambda_{p,i} e^{-(\rho-n)t}, \tag{2.8}$$

$$\tau : \lambda_{p,i} = \lambda_{k,i}, \tag{2.9}$$

$$K_{p,i} : -\lambda_{p,i} \left[r_p(1-\xi_p) - \delta - n \right] = \dot{\lambda}_{p,i}, \tag{2.10}$$

$$K_{g,i} : -\lambda_{g,i} \left[r_g(1+\xi_g) - \delta - n \right] = \dot{\lambda}_{g,i}, \tag{2.11}$$

where $\lambda_{p,i}$ and $\lambda_{g,i}$ denote the private shadow values to agent i associated with the polluting and the green capital stock, respectively. Moreover, the transversality conditions read as

$$\lim_{t\to\infty} \lambda_{p,i} K_{p,i} = \lim_{t\to\infty} \lambda_{g,i} K_{g,i} = 0. \tag{2.12}$$

The interpretation of the optimality conditions is standard. (2.8) equates the marginal utility of consumption to the marginal utility of present value polluting wealth, $\lambda_{p,i}$; (2.10) equates the marginal return on consumption to the marginal product of polluting capital; (2.9) says that the marginal product of polluting capital must be equal to the marginal product of green capital. This is an indifference condition: the two types of capital are equally productive in the steady-state. In other words: the marginal utility of consumption must also be equal to the marginal utility of green wealth, $\lambda_{g,i}$.

Using (2.8)-(2.11) and aggregating over N identical representative agents leads to the macroeconomic equilibrium of the decentralized economy. Specifically, assuming firms maximize profits under perfect competition, and inputs get paid their marginal product, using (2.5), the aggregate versions of (2.3) and (2.4) can be written as:

$$\dot{K}_p = Y - C - \tau I - \delta_p K_p, \tag{2.13}$$

and

$$\dot{K}_g = \tau I - \delta_g K_g. \tag{2.14}$$

Under plausible conditions (that now involve tax rates), we define a balanced-growth equilibrium as a growth path along which all variables grow at a constant rate. With polluting and green capital being accumulated from final output, along such a path, the polluting capital-output ratio remains constant. From the aggregate production function $Y = NY_i$, the long-run equilibrium balanced growth path of output and polluting capital is

$$\gamma_{K_p} = \gamma_Y = \frac{\sigma_n}{1-\tilde{\sigma}} n, \tag{2.15}$$

with

$$\tilde{\sigma} \equiv \sigma_p + \sigma_g = (\alpha + \beta) + \mu(\phi_g - \phi_p). \tag{2.16}$$

In the following, we shall make the plausible assumption that polluting externalities do not dominate, i.e. $\tilde{\sigma} > 0$. Moreover, as the ratio of polluting and green capital is constant as well (for constant tax rates), i.e.

$$\frac{\tilde{K}_p}{\tilde{K}_g} = \frac{(1-\kappa)(1-\xi_p)}{\kappa(1+\xi_g)}, \tag{2.17}$$

the long-run equilibrium balanced growth path of green capital is identical to the long-run equilibrium balance growth path of output and polluting capital. The equilibrium growth rate is only determined by technological factors, summarized by the term $\frac{\sigma_n}{1-\tilde{\sigma}}$. Due to the non-scale nature of the production technique, the growth rate is independent of all demand characteristics (see Jones, 1995; Turnovsky and Monteiro, 2007). The scale adjusted per capita growth rates are given as:

$$\gamma_C - n = \gamma_{K_g} - n = \gamma_{K_p} - n = \gamma_Y - n = \left(\frac{\sigma_n + \tilde{\sigma} - 1}{1-\tilde{\sigma}}\right) n. \tag{2.18}$$

As directly immediate from expression (2.18), for the case of constant returns to scale, i.e. $\sigma_n + \tilde{\sigma} = 1$, there is no long-run per capita growth.

Expressing the dynamics in terms of scale-adjusted stationary variables

$$k_j \equiv \frac{K_j}{N^{\left(\frac{\sigma_n}{1-\tilde{\sigma}}\right)}}, \quad c \equiv \frac{C}{N^{\left(\frac{\sigma_n}{1-\tilde{\sigma}}\right)}}, \tag{2.19}$$

for $j = \{p, g\}$ allows us to express the core equilibrium dynamics of the decentralized economy in terms of the redefined stationary variables

$$\dot{c} = \frac{c}{\theta}\left[(1-\kappa)(1-\xi_p)k_p^{\sigma_p-1}k_g^{\sigma_g} - \delta - \rho - n\left(1-\theta+\frac{\sigma_n\theta}{1-\tilde{\sigma}}\right)\right], \tag{2.20}$$

$$\dot{k_p} = \left[k_p^{\sigma_p}k_g^{\sigma_g}(1-\tau) - c(1-\tau) - \delta k_p - \frac{\sigma_n}{1-\tilde{\sigma}}nk_p\right], \tag{2.21}$$

$$\dot{k_g} = \left[k_p^{\sigma_p}k_g^{\sigma_g}\tau - c\tau - \delta k_g - \frac{\sigma_n}{1-\tilde{\sigma}}nk_g\right]. \tag{2.22}$$

Imposing the steady state condition, $\dot{c} = \dot{k_p} = \dot{k_g} = 0$, we can recursively solve for the steady-states of k_p, k_g and c:

$$\tilde{c} = \frac{1}{\alpha(1-\kappa)(1-\xi_p)}\left[\rho + \delta(1-\tilde{\sigma}) + n\left(1-\theta+\frac{\sigma_n(\theta-\tilde{\sigma})}{1-\tilde{\sigma}}\right)\right]\tilde{k}_p, \tag{2.23}$$

$$\tilde{k}_p = \left\{\frac{1}{\alpha(1-\kappa)(1-\xi_p)}\left[\delta+\rho+n\left(1-\theta+\frac{\sigma_n\theta}{1-\tilde{\sigma}}\right)\right]\left(\frac{(1-\kappa)(1-\xi_p)}{\kappa(1+\xi_g)}\right)^{\sigma_g}\right\}^{\frac{1}{\tilde{\sigma}-1}} \tag{2.24}$$

$$\tilde{k}_g = \left(\frac{\kappa(1+\xi_g)}{(1-\kappa)(1-\xi_p)}\right)\tilde{k}_p, \tag{2.25}$$

where we have used the fact that $\tau = \frac{k_g}{k_g+k_p}$.

Proposition 1. *The steady-state equilibrium is a unique hyperbolic equilibrium point which is a (degenerate) saddle with a two dimensional stable manifold if and only if* $\tilde{\sigma} < 1$.

Proof. *See Appendix.*

Corollary. *For* $\tilde{\sigma} < 1$ *we find that* $\tilde{k}_p > 0$ *and* $\tilde{k}_g > 0$.

2.6 Tighter Environmental Policy: Long-run Response

In this section, we discuss the long-run responses of a positive environmental tax shock. The key insight from this analysis is that the effect of a positive environmental tax on the absolute evolution as well as on the relative composition of the economy's knowledge capital stock is ambiguously determined by the relative size of non-internalized environmental externalities in σ_p and σ_g, as shown with the next proposition. In turn, we take up this ambiguous result to motivate our empirical application, namely to estimate σ_p and σ_g in order to check whether or not a more restrictive environmental policy not only significantly increases the share of green capital in total capital but also leads to an increase of the *absolute* stock of green capital in the post-shock long-run.

Proposition 2. *Assume that Proposition 1 holds. A tighter environmental tax policy leads to a reduction of the stock of green and polluting capital that is more pronounced in the short- than in the long-run, provided that environmental spillovers are sufficiently low so that $\sigma_g < 1$ and $\sigma_p > 0$. In turn, if environmental spillovers are sufficiently pronounced so that $\sigma_g > 1$ and $\sigma_p < 0$, a tighter environmental tax reduces the stock of green and polluting capital in the short-run but leads to an increase of both types of capital in the long-run.*

Corollary. *Proposition 2 holds even if knowledge spillover are absent, i.e. $\beta = 0$.*

Proof. *Assume that Proposition 1 holds. Differentiating (2.24) and (2.25) for ξ_p, we find that*

$$\frac{d\tilde{k}_p}{d\xi_p} \begin{cases} < 0, & 0 < \sigma_g < 1, \quad -\sigma_g < \sigma_p < 1 - \sigma_g \\ = 0, & \sigma_g = 1, \quad -1 < \sigma_p < 0, \\ > 0, & \sigma_g > 1, \quad -\sigma_g < \sigma_p < 1 - \sigma_g \end{cases} \tag{2.26}$$

and

$$\frac{d\tilde{k}_g}{d\xi_p} \begin{cases} < 0, & 1 > \sigma_p > 0, \quad 0 < \sigma_g < 1 - \sigma_p \\ = 0, & \sigma_p = 0, \quad 0 < \sigma_g < 1. \\ > 0, & \sigma_p < 0, \quad -\sigma_p < \sigma_g < 1 - \sigma_p \end{cases} \tag{2.27}$$

The intuition behind Proposition 2 is as follows. Assume that environmental spillovers are sufficiently moderate ($\sigma_g < 1$ and $\sigma_p > 0$) which might be a realistic benchmark scenario. A stricter environmental policy reduces the polluting knowledge capital and thereby the growth rate of output. The growth rate of green capital begins to fall as well, so that green as well as polluting capital follow declining paths. In the short-run, the private agents substitute away from capital investment into consumption, leading to an instantaneous increase of consumption, as capital cannot be adjusted immediately. As green capital increases more in relative scarcity than polluting capital, its productivity rises, inducing investment in green capital and thereby restoring its growth rate in the medium-run. The rising stock of green capital raises the productivity of polluting capital, and thus, the private agent starts accumulating polluting capital as well. This is paralleled by a fall of consumption in the medium-run, while the growth rate of consumption increases in the longer-run due to an increase of the growth rate of output initiated by increased capital accumulation in the longer-run. Although both stocks of capital are lower in the new-steady state, we observe a change in the composition of both kinds of capital, where the economy's capital stock is gradually changed from polluting to green capital. However, for specific parameter constellations, our economy also ends up with higher stocks of both types of capital or with a higher stock of green capital relative to the pre-shock scenario. Figure A1 in the Appendix shows the transitional dynamics after perturbating the economy with an environmental tax shock for a calibrated example. Right after the tax shock, we observe that polluting capital gains overall importance in the overall capital stock, as green capital decreases more rapidly than its polluting counterpart. In the medium- and long-run, however, as green capital is accumulated more rapidly than polluting capital, the importance of green capital increases from 32 % (10 periods after the shock) to 45% close to the new steady-state. Hence, this example also mirrors the fact that short-run implications of policy shocks might dramatically differ from their medium or long-run counterpart. Note that our model implies, by contrast, that a positive shock to green knowledge subsidies has always a positive effect on both knowledge classes if $\sigma_g > 0$ and $\sigma_p < 1$.

Reflecting our model, we see that the effect of a more restrictive environmental policy on the evolution of both capital stocks and output crucially depends on the parameters σ_p and σ_g (see Proposition 2). Hence, in our empirical approach, we attempt to obtain consistent estimates of these parameters from the aggregate production function (2.7) by employing a unique panel data set consisting of European regions. In the theoretical model we have implicitly assumed, for the ease of exposition, that the environmental and knowledge externalities are limited to a single country (or region in our case). As a consequence, the externalities in our model are of rather local nature, as only the agents within a given region are affected by them. Many (environmental) endogenous growth models impose a specific structure on spillovers and their respective channels, normally based on ad-hoc assumptions. Of course, in this regard, our theoretical model makes no exception with regard to within-region spillovers: in particular, spillovers affect productivity in a linear fashion[5] and they are additively separable from the own (positive) productivity effect of green (and polluting) capital. Naturally, an important aspect are as well spillovers between regions, which we have abstracted from for the sake of clarity. Consequently, an important econometric challenge is to adequately control for these unobserved sources of spillovers between regions. To do so, we choose a flexible approach that is able to account for unobserved spillovers of unknown form between regions. The approach, which is outlined in the following section, has the advantage that we do not have to make any ex-ante assumptions on the nature of these spillovers. Moreover, our employed empirical strategy should also allow for heterogeneous production technologies across regions and should, finally, treat dynamics and properties of time-series appropriately. In the following section, we outline our empirical approach in due detail.

3 Empirical Framework

3.1 Aggregate Production Function

To empirically model the productivity effects of green technologies, we estimate the aggregate production function (2.7). In logarithmic form and adding an error term, economic output (y_{rt}) is produced according to

$$y_{rt} = \sigma_g k_{g,rt} + \sigma_p k_{p,rt} + \sigma_n n_{rt} + \sigma_k k_{k,rt} + v_{rt}, \tag{3.1}$$

with

$$v_{rt} = \psi_r + e_{rt}, \tag{3.2}$$

where t and r now index years and European regions, respectively, ψ_r is an individual-specific effect constant over time, n_{rt} is labor input, and $k_{g,rt}$ and $k_{p,rt}$ are measures of green and polluting knowledge capital, respectively. A detailed exposition of how the two stocks are classified can be found in Section 4. Compared to our model, we additionally control for physical investment ($k_{k,rt}$). Note that this kind of Cobb-Douglas functions with included knowledge capital is of Griliches (1979)-form, which is a standard approach in the literature (e.g. Eberhardt et al., 2013; Mitze et al., 2016; Stucki and Woerter, 2019).

Building their empirical models on endogenous growth theory, Coe and Helpman (1995) and Ertur and Musolesi (2017) arrive at similar specifications for general knowledge capital and its impact on total factor productivity. Latter studies attribute an important role to the effects of external knowledge capital, which can in principal stem from different sources: first, they can be the result of targeted knowledge transfer. Secondly, they might occur because of the public goods nature of knowledge, which means that knowledge might spill from one actor to another in an unintended way (e.g. Keller, 2004; Eberhardt et al., 2013). These knowledge spillovers might be modeled by including stocks of external knowledge to equation (3.1), where the external knowledge capital of a region is a weighted average of the internal knowledge capital of all other regions (e.g. Coe and Helpman, 1995; Ertur and Musolesi, 2017). Weights might be, for example, based on trade relations (e.g.

Coe and Helpman, 1995) or on geographic distance (e.g. Ertur and Musolesi, 2017). However, estimation ordinarily rests on a priori assumptions on the nature of the dependence, i.e. the weighting matrix (Chudik and Pesaran, 2015a) and on the assumption that CSD is effectively controlled for when the chosen channels are included (Eberhardt et al., 2013). Additionally, in the case of green technologies, as highlighted in our model, spillovers might as well occur through positive effects of green knowledge capital on environmental quality of other regions. However, we are explicitly modeling within-region externalities only and controlling for these between-region spillovers is consequently of primary importance to consistently identify the parameters of the production function. Hence, we adopt an approach that accounts for unobserved spillovers of unknown form without explicitly modeling them (e.g. Eberhardt et al., 2013).[6] The drawback of this procedure is that we can not quantify the contribution of different sources of spillovers directly (Mitze et al., 2016) and we thus focus on the private returns of knowledge, as done, e.g., by Eberhardt et al. (2013) and Stucki and Woerter (2019). Hence, we focus on consistently identifying the parameters of the aggregate production function (2.7) with implicit within-region spillovers. The next section briefly recaptures the notion of CSD and ways to estimate the degree of CSD in the data, followed by a detailed exposition of our estimators of choice.

3.2 Cross-Sectional Dependence

For illustrative purpose, we adopt the depiction by Ertur and Musolesi (2017), which highlights two potential sources of CSD in the error term of equation (3.2)

$$e_{rt} = \boldsymbol{\varrho}_r' \boldsymbol{f}_t + \xi \sum_{s \neq r} \omega_{rs} e_{st} + \varepsilon_{rt}, \tag{3.3}$$

where $\boldsymbol{f}_t = (f_{1t}, f_{2t}, ..., f_{mt})'$ is a $m \times 1$ vector of unobserved factors, $\boldsymbol{\varrho}_r = (\varrho_{r1}, \varrho_{r2}, ..., \varrho_{rm})'$ is a $m \times 1$ vector of factor loadings, ω_{rs} is a spatial weight matrix satisfying specific conditions[7], ξ is a spatial autoregressive parameter and ε_{rt} is an idiosyncratic error. The first term on the RHS is related to factor models and typically to so-called strong CSD, whereas the second term on the RHS is a spatial error process satisfying so-called weak

CSD (Ertur and Musolesi, 2017; Sarafidis and Wansbeek, 2012; Chudik et al., 2011).[8] Simply put, weak CSD might be thought of as rather local, spatial dependence, whereas common effects that are due to unobserved, global factors are a form of strong CSD (e.g. Bailey et al., 2016a). This implies that while former dependence is restricted to units that are somehow connected to each other, latter is not (Mitze et al., 2016). The implications for estimation are related to the degree of CSD. For example, the spatial error process in equation (3.3) does itself not affect consistency and unbiasedness of conventional panel estimators, whereas strong CSD, represented by a factor model, does if factors and/or loadings are correlated to the regressors (Sarafidis and Wansbeek, 2012). As Ertur and Musolesi (2017) argue, there is neither a theoretical nor an empirical reason in the context of international technology spillovers to assume the mere prevalence of weak or strong CSD. We argue that this reasoning applies to regional technological and environmental spillovers as well. European regions could likely be driven by European-wide factors with region-specific responses to them or by rather local spillover effects. For example, one might think of a global technology trend from which regions profit depending on their individual characteristics, and/or very local clusters consisting of few proximate regions.

We follow Ertur and Musolesi (2017) and Ciccarelli and Elhorst (2018) and employ diagnostics to gauge the magnitude and the nature of CSD in the data. These include the CD test (Pesaran, 2004, 2015a) and the estimation of the exponent of CSD (Bailey et al., 2016b). Both measures are applied as well to the residuals to validate and compare the estimation approaches. As Pesaran (2015a) shows, the CD test has the implicit null hypothesis of weak CSD. Specifically, if the panel dimension T is almost fixed as $N \to \infty$, as it is roughly the case in our sample, the implicit null is given by $0 \leq \alpha < 1/2$ (Ertur and Musolesi, 2017), where α refers to the exponent of CSD (Bailey et al., 2016b). The exponent is a measure of the convergence rate of the variance of the cross-sectional average of a specific variable (Bailey et al., 2016b). α can be in the range $[0, 1]$ and can only be identified if $\alpha > 1/2$. Any process with $\alpha < 1$ fulfills the definition of weak CSD, whereas $\alpha = 1$ corresponds to strong CSD. However, values of $\alpha < 1$ indeed indicate different magnitudes of CSD (Chudik and Pesaran, 2015a). A special threshold is furthermore marked by $\alpha = 0.75$, because for values of $\alpha \in [0.75, 1)$ convergence of the average correlation coefficient is still slower than $N^{-1/2}$, which indicates common factors as well (Ciccarelli and Elhorst, 2018). To get a better understanding of the nature and the magnitude of the CSD, we estimate α as described in Bailey et al. (2016b). In our application, we estimate two different versions of the bias-adjusted estimator given by equation (13) of Bailey et al. (2016b): The first one (denoted by $\hat{\alpha}$) is the standard version assuming no temporal structure in the factors and no weak CSD in the error term. The second one is the version robust against both issues (denoted by $\tilde{\alpha}$). An estimation approach that is able to consistently estimate a model with multifactor error structure and spatial error correlations as in equation (3.3) is the Pesaran (2006) common correlated effects (CCE) approach (Pesaran and Tosetti, 2011), which we introduce in due brevity in the following section.

3.3 Estimation Strategy

Our estimation strategy follows broadly Eberhardt et al. (2013) and Eberhardt and Teal (2013) in that we contrast several estimators that make different assumptions regarding the data generating process. We do this to get a comprehensive view and to ensure that the results are not driven by specific a priori assumptions. As our main approach, we choose a flexible framework that accounts for several important aspects connected to the estimation of production functions (see Eberhardt and Teal, 2011, for an overview). The CCE approach explicitly models an unobserved common factor structure in the residuals (Pesaran, 2006), and is a very convenient way to capture unobserved spillovers that are potentially complex and non-symmetric (Eberhardt et al., 2013). Drawing on Pesaran (2006) and Pesaran (2015b) we specify the logarithmic aggregate production function (3.1) as follows:

$$y_{rt} = \boldsymbol{a}_r' \boldsymbol{d}_t + \boldsymbol{\beta}_r' \boldsymbol{x}_{rt} + e_{rt} \tag{3.4}$$

$$e_{rt} = \boldsymbol{\varrho}_r' \boldsymbol{f}_t + \epsilon_{rt} \tag{3.5}$$

$$\boldsymbol{x}_{rt} = \boldsymbol{A}_r'\boldsymbol{d}_t + \boldsymbol{\Gamma}_r'\boldsymbol{f}_t + \boldsymbol{v}_{rt}. \tag{3.6}$$

Where $\boldsymbol{x_{rt}} = [k_{g,rt}, k_{p,rt}, n_{rt}, k_{k,rt}]'$, $\boldsymbol{d}_t = 1$, $\boldsymbol{a}_r' = \psi_r$, and $\boldsymbol{\beta_r} = [\sigma_{g,r}, \sigma_{p,r}, \sigma_{n,r}, \sigma_{k,r}]'$ collects the coefficients. The errors have a multifactor structure, where $\boldsymbol{f}_t$ is a vector of unobserved common effects, $\boldsymbol{\varrho}_r$ is a vector of factor loadings and ϵ_{rt} are idiosyncratic errors. The explanatory variables are driven by a deterministic component, the factors and an idiosyncratic component, where $\boldsymbol{A}_r$ and $\boldsymbol{\Gamma}_r$ are factor loading matrices, and $\boldsymbol{v}_{rt}$ is the idiosyncratic component. Note that the error structure nests time-specific, individual-invariant effects, by defining $\boldsymbol{\varrho_r} = 1$ and $\boldsymbol{f_t} = \lambda_t$ (Sarafidis and Wansbeek, 2012). Pesaran (2006) shows that such a model can be estimated consistently by including cross-sectional averages of the dependent and independent variables to the regression. Two estimators are possible: first, the mean group version (CCEMG) in which the coefficients are assumed to be heterogeneous and are hence estimated separately for each region and then averaged. Second, the pooled version (CCEP), in which the average coefficient is identified directly under the assumption of slope homogeneity. Both estimators are consistent for the average coefficient irrespective of whether the parameters are heterogeneous or homogeneous, although the relative efficiency might be different (Pesaran, 2006). Notably, the idiosyncratic terms $\boldsymbol{v}_{rt}$ and ϵ_{rt} are allowed to contain additional weak CSD (Pesaran and Tosetti, 2011; Chudik and Pesaran, 2015a), an important feature in our empirical setting, as discussed in the previous section. The CCE approach is appealing since it is robust against several additional potential properties of our data. The first issue is possible nonstationarity of the variables of the production function. Kapetanios et al. (2011) show that the CCE approach remains valid if the factors contain unit roots and are possibly cointegrated. More recently, the examinations by Westerlund (2018) suggest that the requirements on the factors are very flexible, including factors with unknown but finite order of integration and structural break dummies. Furthermore, the approach allows by definition for endogeneity of the input variables, since both $\boldsymbol{x_{rt}}$ and y_{rt} are driven by the unobserved factors. Hence, the approach offers a way to control for endogeneity brought in by unobservables (Ertur and Musolesi, 2017), as long as the endogeneity can be captured by the unobserved factors.

Additionally, we employ several alternative estimators for the static benchmark model. First, we employ two pooled estimation techniques, which restrict the coefficients to be identical for each region. These comprise the standard two-way fixed effects (2FE) estimator and the first-difference estimator (FD) with time dummies. As noted above, time-specific individual-invariant fixed effects are a special case of the general factor structure, in which the effect on regions is homogeneous. In fact, evidence from Monte Carlo simulations by Eberhardt and Bond (2009) suggests that including time dummies can remarkably decrease the bias induced by unobserved common factors. As an alternative technique that assumes heterogeneous parameters, we implement the mean group (MG) estimator (Pesaran and Smith, 1995). For the latter we subtract the cross-sectional mean from each variable each year. This procedure removes the impact of common factors entirely, if their effect is region-invariant. If the effect of the factors is heterogeneous across regions, their impact might still be reduced (e.g. Pesaran et al., 1999; Bond et al., 2010).

It is worth emphasizing that the assumptions of our main approach are strict exogeneity of the regressors and that no relevant dynamics are missed in the static approach. This includes the CCE approach, which does not allow for lagged feedback from y_{rt} on $\boldsymbol{x}_{\boldsymbol{rt}}$ and for lagged dependent variables among the regressors (Chudik and Pesaran, 2015b). In case of heterogeneous slopes and weakly exogenous regressors, CCEMG might be biased for small T, and CCEP even becomes inconsistent (Pesaran, 2015b). Omitting relevant dynamics might furthermore lead to a situation where the results do not correspond to long-run responses (Eberhardt et al., 2013). Unfortunately, we can not apply the dynamic version of the CCE estimator (CS-ARDL), which allows for lagged endogeneous variables and weakly exogenous regressors (Chudik and Pesaran, 2015b). This is because especially estimation of the lagged dependent variable suffers considerable bias for small T (Pesaran, 2015b), and with the moderate time dimension in our setup, mean group estimation quickly becomes infeasible as degrees of freedom per group approach zero. Additionally, Everaert and De Groote (2016) show for the homogeneous dynamic panel case that CCEP displays considerable small sample bias for values of T up to 50. As De Vos and Everaert (2021) summarize, in dynamic panels the time series dimension mainly matters for reliability of the CCE estimators. However, to mitigate concerns over misspecification with regard to weak exogeneity of the regressors and potentially missed dynamics, we provide different robustness checks to our main model. First, we employ the CS-DL approach to estimate long-run coefficients, which is especially well suited in moderate T samples and seems to outperform the CS-ARDL approach in these circumstances (Chudik et al., 2016). In a nutshell, the CS-DL approach is a reformulation of a dynamic panel approach with common factors, in which the long-run coefficients can be estimated directly by CCE, without estimating the coefficient of the lagged dependent variable explicitly. Purging the cross-sectional dependence requires the addition of cross-sectional averages of regressors and the dependent variable, as well as lags of the cross-sectional averages of the regressors (Chudik et al., 2016). It is a quite flexible approach as it allows for heterogeneous slopes and offers a mean group and a pooled variant, just as the CCE approach. However, the CS-DL estimation procedure maintains the strict exogeneity assumption. Additionally, to relax latter condition, we estimate a dynamic version of the main model with the well-known dynamic panel approach by Blundell and Bond (1998), henceforth BB. The BB approach can be seen as an extension of the estimator by Arellano and Bond (1991) with the attempt to better handle very persistent data series by adding additional moment conditions based on a mean-stationarity assumption on the initial observations (e.g. Roodman, 2009a; Sarafidis and Robertson, 2009). Both approaches can handle weakly exogenous and even endogenous regressors, given appropriate moment conditions formed by lagged observations (e.g. Kiviet, 2020). However, the standard approach allows to control for CSD only via time dummies, and assumes homogeneous parameters. The bias from unaccounted error CSD in these approaches can be reduced by including time-fixed effects (Sarafidis and Robertson, 2009), but violation of the homogeneous slope condition can lead to inconsistent results (Pesaran et al., 1999).

4 Data and Descriptive Statistics

The data set covers a time period of $T = 25$ years between 1991-2015 for $R = 270$ European NUTS-2 regions in 28 countries, resulting in a balanced panel of 6750 observations.[9]

The main sources of data are, first, the Cambridge Econometrics database, from which we obtain industrial gross value added (GVA), industrial gross fixed capital formation (GFCF) and total hours worked in the industry sector. Investment and value added series are deflated to constant 2005 prices and given in Millions of Euro, total hours worked are given in Millions of hours worked. We include the flow measure gross fixed capital formation as physical capital input directly in our model instead of computing physical capital stocks, for example with the perpetual inventory method (e.g. Caselli, 2005). By including the flow series, we circumvent controversial decisions on starting values.[10]

Secondly, we use the EPO (European Patent Office) Worldwide Patent Statistical database (PATSTAT) to gather information on patent applications and to construct regional knowledge stocks by accumulating patent counts into patent stocks. Patents are one of the most commonly used measures of innovation (Barbieri et al., 2016) as they represent an advantageous indicator (Griliches, 1990), not the least due to their wide and detailed data provision (Haščič and Migotto, 2015). Only few economically significant inventions have not been patented (Dernis and Khan, 2004). Nevertheless, patent data faces some relevant drawbacks that can hardly be circumvented, such as the accounting of strategic patents or the restriction to technological innovation (Barbieri et al., 2016) as well as very limited information on diffusion (Kemp, 2010). Further concerns, such as differing patent quality (Johnstone et al., 2010) or mistakes when searching environmental patents (Lanjouw and Mody, 1996) can be substantially mitigated by the choices made in the search of patents.

We rely on multinational patent applications filed at the EPO to create robust measures with respect to patent value and comparability, as only innovations of sufficient expected commercial profitability justify the relatively high application costs (Johnstone et al., 2010). In order to avoid counting technologies multiple times we restrict our search to the first EPO patent application of a patent family. We follow Costantini et al. (2017) by using patent applications with their earliest filing year in order to timely capture the innovative effort. Further, we decide to assign patents based on the residence of the inventor, thus capturing inventive activity (Kruse and Wetzel, 2016; Wurlod and Noailly, 2018). In case of multiple inventors from different regions or countries, the patent is allocated using fractional counts (Kruse and Wetzel, 2016; Wurlod and Noailly, 2018). The accumulation into knowledge stocks follows the method proposed by Popp et al. (2011), such that

$$K_{j,rt} = \sum_{s=0}^{\infty} e^{-\beta_1(s)}(1 - e^{-\beta_2(s+1)})PAT_{j,r,t-s}, \tag{4.1}$$

where $PAT_{j,r,t-s}$ is the patent count in period $t-s$ for region r for the patent group $j = \{g, p\}$. The rate of knowledge depreciation is set to 0.1 (β_1) and the rate of diffusion to 0.25 (β_2), as proposed by Popp et al. (2011). Thus, the relevance of a patent application peaks after 4 years (Popp et al., 2011), which seems to be a reasonable dynamic for diffusion patterns of knowledge capital to affect productivity. To mitigate the influence of the initial observation on the knowledge stocks, we calculate all stocks with pre-sample patent data from 1985 onwards.[11]

Since our analysis is performed for the industry sector, we utilize the concordance table provided by Schmoch et al. (2003) that links technology classes to economic sectors to match the patents to it. This concordance table is frequently used in the empirical literature (Costantini et al., 2017). It should be noted that the industry data as classified by Cambridge Econometrics is more comprehensive than the scope of the classification by Schmoch et al. (2003), which encompasses the manufacturing sector only. The industry sector as defined in the Cambridge Econometrics data further encompasses mining, recycling, energy provision, and water provision. However, the technology classes defined by Schmoch et al. (2003) lead to similar patent counts as those obtained without any restriction on the technology classes.

To define the patent groups green (g) and non-green (p), we differentiate green technologies based on the technology classes a patent belongs to. Two established options are the Green Inventory (GI) and the OECD EnvTech (EnvTech), which both define technology classes that are considered to correspond to environmentally sound technologies. As noted by Lanjouw and Mody (1996), there are two potential errors when searching environmental patents. First, to include too many patents that are not actually qualifying as environmental patents. Second, to include too few patents, as some environmental patents are not encompassed. However, with respect to the soundness of results, we perceive the latter to be less problematic (Lanjouw and Mody, 1996; Wurlod and Noailly, 2018). Hence, we choose the GI, which is considered as being more narrow (Ghisetti and Quatraro, 2017). Consequently, as soon as a patent belongs to a technology class encompassed by the GI, it is considered as a green patent. Non-green patent counts are constructed by subtracting environmental patents from the overall patent count. Finally, since we are using logarithmic variables in estimation, we use $k_{j,rt} = log(1 + K_{j,rt})$ for $j = \{g, p\}$ (e.g. Stucki and Woerter, 2019), since knowledge stocks might be zero for some regions and years.

Table 1 displays some summary statistics for the main variables we employ in the empirical analysis. In the Appendix (Table B1 and Table B2), we present the unconditional correlation matrix of our main variables, including the cross-sectional means that approximate the common factors. As expected, the green and non-green knowledge stocks are highly correlated. This holds true for both the two specific stocks as well as the cross-sectional averages. Furthermore, in the Appendix (Table B5) we report the results of panel unit root tests of the second generation (Pesaran, 2007) that allow for one unobserved factor. The results are somehow mixed but suggest that the presence of unit roots can not be rejected for all variables when a higher lag order is added. As noted in Section 3, the CCE estimators are robust against nonstationarity in the factors and different scenarios of cointegration. Since our approach does not rely on cointegration but is robust to various scenarios, we do not test for cointegration, as similarly argued by Eberhardt and Teal (2011). However, to validate the estimation approaches in the empirical part, we test whether the residuals are integrated of order one. Finally, to get an overview of the majority of the regions comprising the sample and the spatial patterns of the dependent variable and the explanatory variable of main interest, Figure B1 and Figure B2 in the Appendix present maps with regions colored according to their relative position of labor productivity and labor deflated green knowledge stocks for the first and the second half of the time series.

5 Results

5.1 Estimation of Cross-Sectional Dependence

In this section, we discuss the results of the CD test (Pesaran, 2004, 2015a) and the exponent of cross-sectional dependence (α) (Bailey et al., 2016b) applied to the variables of our model. Table 2 contains the CD test statistics, the point estimates of the bias-adjusted version of α and 90% confidence intervals.[12]

As Ertur and Musolesi (2017) note, the exponent of cross-sectional dependence is originally developed for stationary variables. Hence, we adopt their proposed robustness test and estimate both the CD statistic and α for first-differenced variables as well. First, it is evident that the implicit null of the CD test, $\alpha \in [0, 0.5)$, is strongly rejected for all variables, based on conventional standard normal critical values. This holds true for the variables in log-levels as well as in first log-differences. Secondly, the point estimate of α (denoted $\hat{\alpha}$) is above 0.9 and close to 1 for all considered variables in log-levels. In first differences, the point estimates are lower, but still considerably above the turning point of 0.75. Furthermore, the lower bound of the confidence interval is well above the threshold value of 0.75 in all cases.[13] These observations imply that it is very likely that a factor structure is driving the data, and that we have to take this into account when estimating the aggregate production function. Hence, the results indicate that the approach outlined in Section 3.3 should be well suited in the empirical context at hand.

5.2 Main Estimation Results

In this section we present the main estimation results for the aggregate production functions. Table 3 contains the estimation results for the static baseline model estimated with two-way fixed effects (2FE), first-difference OLS (FD), CCEP, CCEMG and the mean-group estimator (MG). The first three approaches pool the data under the assumption of common slope coefficients, whereas the mean group estimators run separate regression for each unit under the assumption of fully heterogeneous coefficients. It is worth emphasizing that both assumptions are likely to be violated, as they present the two diametrical notions of either all regions having the exact same coefficient or all regions having different coefficients. In reality, a mixed case seems to be more plausible (Maddala et al., 1997).

We also report various diagnostics for the residuals. First, we apply the CIPS test (Pesaran, 2007) to the residuals to gauge whether the residuals are stationary. Secondly, we report the CD statistic and both estimates of α, introduced in Section 3.2, in order to get an impression of the degree of CSD that is left in the errors. As noted by Sarafidis and Wansbeek (2012), however, the CD statistic might lose power if time-dummies are included in estimation or, equivalently, the data is expressed as deviations from a time-specific mean, since the positive and negative correlations in the residuals cancel. As Millo (2019) notes, the same effect applies in the CCEP case because of the augmentation with cross-sectional averages. This might lead to a situation where the average (pairwise) correlation coefficient is near zero, and so will be the CD statistic. To detect such a situation, we also report, as suggested by Millo (2019), the average (pairwise) cross-correlation coefficient ($\bar{\rho}$) as well as the average absolute (pairwise) cross-correlation coefficient ($|\bar{\rho}|$).

As evident from Table 3, the estimated elasticity of value added with respect to physical capital input is significantly positive in all employed approaches. The magnitude ranges from $0.099 - 0.138$, being very similar to comparable studies based on the country-industry level (Mitze et al., 2016; Stucki and Woerter, 2019). The same observation can be made for labor input, the parameter estimates range from $0.239 - 0.605$ for all approaches. With respect to the estimated returns to the two differentiated knowledge classes, the following pattern emerges. The coefficient for non-green knowledge capital is estimated to be significantly positive in all cases. The range of magnitudes spans $0.078 - 0.221$. On the other hand, the parameter estimates for the green knowledge stock are negative or not significantly different from zero in all approaches. In all pooled models, the negative coefficient is significant. Overall, non-green knowledge has a positive association with value added in European regions, whereas the parameter estimates for the green knowledge stock are insignificant or even significantly negative.

Turning to the residual diagnostics, the CIPS test rejects the null hypothesis of a unit root for all residuals except the ones obtained from the 2FE model. Interestingly, the

CD statistic is below the common critical values of the standard normal distribution only in the 2FE case, indicating that α should be above 0.5 in all other cases. However, the low value of the CD statistic seems to be a result of the situation previously elaborated on. The average pairwise correlation coefficient is close to zero, whereas the average pairwise absolute correlation coefficient is relatively high, indeed higher then in all other models. This fact is as well represented in the estimation of the exponent of cross-sectional dependence. Both versions of the bias-adjusted point estimate of α are indeed larger then 0.5 for all models. The 2FE and the FD model perform quite well in this regard, displaying low estimated degrees of CSD in the residuals. Surprisingly, the estimation of the exponent of CSD suggests that the CCEP estimator is not able to account for the factor structure effectively, since the point estimate is quite high and above the threshold value of 0.75. Taken together, the residual diagnostics suggest that for most approaches strong CSD in the form of common factors is controlled for, and the errors remain weak to semi-strong cross-sectionally correlated. Based on the diagnostics, FD in the pooled case and CCEMG in the mean group case appear to be the preferable estimators.

To summarize, while our results point to a robust positive effect of non-green knowledge of considerable magnitude, they imply significantly negative returns to green technologies in the pooled models. In both mean group models, the coefficients for green knowledge are insignificantly different from zero. While these results line up well with firm level evidence (Marin and Lotti, 2017), industry-level evidence by Stucki and Woerter (2019) points to rather similar effects between green and traditional knowledge. Our results are more pessimistic with regard to the economic returns to green technologies on the regional level. With regard to our simple growth model, these results give rise to the following interpretation. First, as the estimated coefficient of green knowledge capital comprises both spillover effects, there is either no significant impact of environmental quality on output as sketched in our model ($\mu \approx 0$), or there is no significant positive effect of green capital on local environmental quality ($\phi_g \approx 0$). Second, the direct effect on output is not significantly different from zero ($\kappa \approx 0$). Furthermore, with regard to a more stringent environmental policy, it is more likely that the first scenario applies. This implies that a tax would likely change the relative composition of green and non-green capital, but it might result in a scenario in which, in the long-run, both capital stocks decline. We discuss this implication in detail in Section 5.4.

5.3 Robustness and Extensions

In the following section we discuss a battery of robustness checks and extensions to the benchmark approach. First, we consider dynamic production functions and methods that allow for weakly exogenous and endogenous production inputs. Second, we analyze whether the main approach is robust against different categorizations of green knowledge and utilize different, more narrow subgroups of green technologies. Third, we further investigate possible heterogeneity across regions by splitting the sample to EU15 and non-EU15 regions. Additionally, we perform numerous further robustness exercises the results of which can be found in the Appendix. Specifically, to validate the results against the choice of the method to compute the knowledge stocks, we also consider knowledge stocks that are computed with the perpetual inventory method[14] and knowledge stocks in which we assign the patents to the address of the applicant instead of the inventor. The results of the main approach are generally qualitatively robust against this alternative technique to compute the stocks and to assign patents. Furthermore, the results are qualitatively robust against the use of lagged production inputs, the inclusion of a physical capital stock, computed with the perpetual inventory method, instead of investment flows, and the exclusion of physical investment. Finally, we do not find convincing empirical support for a nonlinear relationship between green knowledge and regional growth.

Dynamic Estimation and Weak Exogeneity

To mitigate concerns that the results of the main approach are biased because of the presence of weakly exogenous / endogenous regressors or because the coefficients do not correspond to long-run responses, we apply the dynamic panel approaches discussed in Section 3. Specifically, we specify our empirical models as follows. For the BB approaches, we perform one step estimation in all three displayed specifications. With regard to the lag structure, we start with the most parsimonious specification in which we only include one lag of the dependent variable (BB(1)). Furthermore, we treat each variable as potentially endogenous and use all available lags from period $t-2$ onward as instruments. To reduce problems of potential instrument proliferation, we follow Eberhardt et al. (2013) and collapse the instrument matrix as suggested by Roodman (2009b). Next, we consider a specification with a richer lag structure by including all first and second lags of the dependent and all independent variables, while retaining the same instruments set (BB(2)). Finally, in the third specification (BB(3)), we retain the lag structure of the second one but use a smaller instruments set, including lags $t-4$ to $t-8$. Cross-sectional dependence is captured with a full set of time dummies in all three specifications. With regard to the CS-DL approach, we use contemporaneous first-differences of the input variables, the contemporaneous cross-sectional average of the dependent variable and contemporaneous and one-period lagged cross-sectional averages of the production inputs in levels.[15]

Table 4 reports the implied long-run coefficients for the three different BB specifications, the CS-DL mean group (CS-DLMG) and CS-DL pooled (CS-DLP) approaches. The results remain qualitatively similar for the knowledge stocks when applying the dynamic specifications compared to our benchmark approach, and are even of more pronounced magnitude. Notably, while second order autocorrelation of the residuals can not be rejected at the 10% level for the most parsimonious specification, it can be for both higher order models. Furthermore, reducing the instruments count clearly improves instrument validity, such that BB(3) is our preferred specification. Finally, the results of the CS-DL approach are both very similar to the static CCE results in the main approach. Again, although remaining cross-sectional dependence seems to be reduced compared to the main approach, the pooled version appears to leave higher degrees of CSD behind.

Different Subgroups of Green Technologies

Despite our findings for green technologies in general, it appears to be possible that certain subclasses of green technologies have a positive effect on productivity, while others have not. For example, while end-of-pipe innovation (Frondel et al., 2007) only provides economic returns if emissions are priced, reductions of resource use provide economic returns, since (properly) priced natural resources qualify as a cost factor (O'Mahony and Timmer, 2009). If the new technology saves an input that is not (or not sufficiently) priced, the willingness to pay for these technologies will likely be low. This intuition corresponds to findings in the literature on the firm level, which suggest that only specific technologies (e.g. resource saving) provide distinct returns (Ghisetti and Rennings, 2014; Rexhäuser and Rammer, 2014). Hence, we define four subgroups of green technologies, namely: alternative energy production, energy efficiency, transportation, and recycling and reuse. We utilize the operationalization by Wendler (2019) and include those IPC classes which are both defined for the green technology field and relate to the industry sector according to Schmoch et al. (2003). These subgroups capture relevant distinct fields and could provide further insights on the above mentioned considerations. While alternative energy production and transportation could relate to long-term large scale processes (Wendler, 2019) and potentially suffer from the existence of externalities, both energy efficiency and resource saving innovations should directly relate to cost reductions. Hence, we estimate the baseline model for each subclass separately. For the sake of brevity, we report the results only for the preferred pooled and mean group model from the main approach, FD and CCEMG, and refrain from reporting residual diagnostics.[16]

As shown in Table 5, the parameter estimates for the green technology subgroups are generally insignificantly different from zero for all estimation methods and technology subgroups, while the estimated coefficients for the non-green knowledge stock are in most cases similar to the ones in the benchmark model. Hence, the results obtained from the subgroups generally mirror the ones obtained for the overall green class.

Subsample Results

In the main approach, although considering heterogeneous parameter models, we are interested in the average coefficient over the full sample. Given the potential heterogeneity of specific groups of European regions, it appears to be interesting to consider averages over specific regional groups in order to gauge whether green knowledge has a different average output elasticity for different, more homogeneous country groups. Hence, we estimate the main model for a EU15 and non-EU15 subsample in the following. In these subsamples, we include the cross-sectional averages over the subsample only. Tables 6 and 7 report the results based on these split sample estimations. Although the results remain qualitatively broadly similar in both split samples, some remarks are in order. First, when concentrating on the CCEMG estimator as preferred approach, the potential negative elasticity of green knowledge with respect to output is somewhat more pronounced for the non-EU15 subsample. Second, the degree of remaining CSD in the errors is lower for the non-EU15 subsample. This might point to greater homogeneity in the non-EU15 group, perhaps because it is considerably smaller, such that the cross-sectional averages / time dummies capture a larger share of regional co-movement.

5.4 Discussion

Based on our theoretical framework we can derive important implications of our empirical findings for the effects of environmental policy. In particular, we propose that environmental policy effects depend on the output elasticities of polluting and green knowledge capital due to their implications for knowledge accumulation. In our empirical investigation, we find that the scenario for which transitional dynamics are displayed in Figure A1 is supported by our empirical findings.

In particular, our empirical findings support that environmental policy induces changes in capital accumulation and composition, similar to the mechanisms isolated by (Xepapadeas and de Zeeuw, 1999). They find that an environmental tax will on the one hand shift the capital composition towards a larger share of modern machines (modernization effect), which corresponds to our finding that the share of green capital increases. On the other hand, they find that the total capital stock will decrease (downsizing effect), again corresponding to our findings that both polluting and green capital will be lower in the new steady-state. These effects on the capital composition have implications for the productivity of the economy. While Xepapadeas and de Zeeuw (1999) assume higher productivity of younger machines, in our application the productivity effects of green capital are estimated to be below those of polluting capital. Hence, in our case it follows that not only output drops, but productivity also decreases due to the shift in capital composition.

Therefore, our findings can be related to the discussion revolving around the strong version of the Porter hypothesis (Porter and van der Linde, 1995). The strong version of the Porter hypothesis (PH) postulates that stricter environmental regulation will positively affect productivity, due to the inducement of innovation by the regulation. Our results are contrary to this in so far as, even if additional innovation is stimulated, productivity will decrease since green capital is less productive than polluting capital and cannot compensate declines in polluting capital. These implications are in line with the bulk of previous literature assessing the validity of the strong PH. Most studies examining the interrelation of environmental regulation and productivity find negative effects of regulation on productivity (Gollop and Roberts, 1983; Smith and Sims, 1985; Gray, 1987; Dufour et al., 1998; Gray and Shadbegian, 2003), though some studies find support for the strong PH for specific samples (Berman and Bui, 2001; Alpay et al., 2002; Murty and Kumar, 2003; Lanoie et al., 2008). With our focus on an encompassing sample of European regions our findings support that, so far, the strong PH does not seem to be a general mechanism.

Evidently, these implications of our results crucially depend on the output elasticity of green capital, with its magnitude surpassing the defined thresholds changing the implications for environmental policy effects. In this vein, it is worthwhile to recall that the output elasticity depends additively on the direct production effect and its effect via the local environmental externality. Hence, our results imply that there is neither a strong direct effect on production nor a productive effect via a local environmental externality. Both of these components are crucial to the future evolution of the productivity effects of green capital. Hence, we shortly discuss the implications and triggers for change in these parameter components.

The direct productivity effect could turn significantly positive if environmental benefits, such as lower emissions, turn into a properly priced input to production. It seems reasonable that, during the timespan under investigation, such environmental input factors have not been relevant cost factors. This is rationalized by the findings on the effectiveness of the EU Emissions Trading System (ETS), which has been considered rather ineffective during our observation period (Ellerman et al., 2016). Nonetheless, with the Paris Agreement in 2015, our sample ends in a year that could mark a shift in these dynamics, such that the direct productivity effect might well increase in the future.

The second component of the output elasticity concerns the reduction of local environmental externalities that reduce production. Hence, the absence of this effect implies that either there is no impact of green capital on local environmental quality or local environmental quality does not pose any restriction on economic activity (yet). It is worth noting that within our theoretical framework, we abstract from global environmental issues and potential externalities from green capital on those. However, under the assumption that these global problems are not yet a restriction on local economic activity, externalities on global environmental problems are negligible for the research question under scrutiny. This assumption seems to be rather reasonable for the chosen set of countries within the timespan of our sample. Beyond these global issues, the relevance of local environmental externalities to production and the contribution of green technologies to resolve environmental limitations might increase in coming decades.

6 Summary and Concluding Remarks

In this paper, we estimate the impact of green technology development on economic productivity for 270 European NUTS-2 regions. To inform our empirical approach, we build a simple growth model with knowledge and environmental externalities within the economy. Abstracting from spillovers between regions, our model highlights that the output elasticities of green and non-green knowledge in the aggregate production function are composed of the within-region environmental and knowledge externalities and the direct productivity parameters. Importantly, in our model, the long-run effects of a tax on non-green knowledge depend crucially on the relative size of both output elasticities. To estimate the aggregate production function, we employ a flexible empirical framework and control for several important econometric challenges that arise in the estimation of macro panels in general and the estimation of production functions in particular.

First, we estimate the degree of cross-sectional dependence between European regions in the variables of the production function and our results give strong indication of the presence of an unobserved common factor structure. Hence, we put special emphasis on appropriately controlling for this feature of the data. Furthermore, we control for heterogeneity in the coefficients and possible cointegration in our main approach and contrast several estimators to get a comprehensive view. The results of our main estimation suggest that the productivity effects of green technologies are insignificantly different from zero in the mean group approaches and even significantly negative for the pooled models. In contrast, our data is robustly indicating significant positive returns to non-green technologies. The results of the main approach are generally robust to a battery of robustness checks. First, we consider dynamic specifications and estimation approaches robust against weak exogeneity and endogeneity of the production inputs. Second, we consider the possibility that only specific subcategories of green technologies have productivity enhancing effects, while others have not. Third, we perform a subsample exercise in which we run separate regressions for regions in the EU15 and in the non-EU15 countries.

A possible methodological advancement might be to control for both common factors and weakly exogenous regressors and/or dynamics, which is beyond the scope of this paper. Furthermore, in this paper, we are interested in a rather general empirical scenario. Nevertheless, it might be an interesting avenue for future research to consider more specific scenarios, for example to analyze if regions that specialize on specific green technologies, according to the smart specialization strategy, can increase economic productivity.

Although our results do not rule out that green technology development could have had positive effects for specific regions with specific capabilities, they question whether green technology development can generally promote economic development of European regions and serve as a "silver bullet" to combine economic and environmental targets. Our simple growth model highlights another mechanism of policy relevance. If productivity effects of between-region spillovers are negligible, the estimated output elasticities imply that a tax on polluting knowledge capital, although shifting the relative compositions towards green knowledge, decreases both knowledge capital stocks in the long-run in absolute terms. To conclude, our empirical results imply that the combination of economic and environmental targets might not be achievable by green technologies alone.

Notes

1. In the following, we use (private) economic returns, growth effects, and productivity effects or variations of these terms interchangeably to refer to the ceteris paribus impacts of green knowledge capital on value added, keeping the other production inputs constant, in line with our empirical approach.

2. We use the terms polluting and non-green interchangeably. Furthermore, we use knowledge capital, knowledge stocks, and technologies interchangeably. For the sake of brevity, we often refer to knowledge capital simply as capital if the meaning is unambiguous.

3. This double externality is neglected by the individual agent. Thus, there is a role for the government in this economy aiming to increase efficiency. The derivation of the optimal policy to implement the first best solution is available upon request.

4. It would equally be possible to regard both capital stocks as an amalgam of physical and knowledge capital, in the spirit of Rebelo (1991). For the sake of clarity with regard to our empirical approach, we consider pure knowledge stocks here.

5. This can be easily seen by linearizing (2.1) and using (2.6).

6. Note that intended transfer is captured either by the modeled spillovers within regions or by the estimation approach in the same manner as unmodeled spillovers are.

7. For details on these "granularity" conditions see, e.g., Chudik et al. (2011).

8. Note that factor models can also generate different forms of weak CSD if there is no strong factor (Chudik et al., 2011). For detailed overviews on weak and strong CSD, the connection to spatial or factor models and to weak and strong factors see, e.g., Chudik et al. (2011); Sarafidis and Wansbeek (2012); Chudik and Pesaran (2015a); Ertur and Musolesi (2017).

9. In total, 13 regions are excluded, starting from a sample of 283 regions. First, both NUTS-2 regions constituting Croatia are excluded, since data before 1995 is naturally

missing for them. For the same reason five regions in the United Kingdom are excluded, which have a substantial amount of missing data, to end up with a balanced sample. Secondly, we exclude six regions because they have a green patent stock of zero for every single observed year. Hence, we concentrate on regions which had at least any green technology development. The regions distribute among the following 28 countries: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Greece, Spain, Finland, France, Hungary, Republic of Ireland, Italy, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, United Kingdom.

10. We discuss checks whether the main approach is robust against the inclusion of a capital stock computed with the perpetual inventory method in Section 5. To compute the capital stocks, we assume a depreciation rate of 6 % (Caselli, 2005) and initial values are based on the capital/output ratio, which we set to 2.6 (Inklaar and Timmer, 2013).

11. Alternatively, knowledge stocks could be constructed with the perpetual inventory method (Kruse and Wetzel, 2016; Wurlod and Noailly, 2018). Results based on this method are discussed as a robustness check in Section 5. The depreciation rate is set to 10% (Verdolini and Galeotti, 2011). We follow Kruse and Wetzel (2016) by dividing the patent count in the first year observed by 0.25; assuming a previous 15% growth rate of the knowledge stock and the 10% depreciation rate. Table B3 and B4 in the Appendix show that the correlation between the knowledge stocks computed with both approaches is quite high.

12. We implement all estimation steps either in STATA or MATLAB. The estimation procedure for the exponent of CSD (α) is implemented in MATLAB. Codes are based on the GAUSS files obtained from the supplementary material of Bailey et al. (2016b). Some formulations from the panel packages of Álvarez et al. (2017) are adopted as well. CD statistics are also implemented in MATLAB. Any errors in the codes are of course our own. The STATA routines we use include *multipurt* (Eberhardt, 2011), *xtdcce2* (Ditzen, 2019) and *xtabond2* (Roodman, 2009a).

13. Estimates of the version of α that is robust against weak CSD in the error term and autocorrelation in the factors provide very similar point estimates to the ones shown here. Results are available upon request.

14. For details on the construction with the perpetual inventory method, see Section 4.

15. Alternative specifications are available upon request. Note that the CS-DL specifications become quickly very demanding. For example, the addition of a second lag of cross-sectional averages of the regressors turns mean group estimation infeasible as the degrees of freedom per group approach zero. While the pooled point estimates remain very similar, standard errors, the computation of which involve the mean group estimates (see Chudik et al., 2016), become very large.

16. Full results are available upon request.

Tables and Figures

Table 1. Summary statistics

Variable (unit)	RT	Mean	S.D.	Min.	Max.
Value added (Millions of Euro in 2005 prices)	6750	7,450.80	8,256.99	32.707	73,113.81
Physical capital input (Millions of Euro in 2005 prices)	6750	1,834.21	2,060.77	1.37	21,363.70
Labor input (Millions of hours worked)	6750	266.57	250.45	1.55	3446.39
Green knowledge stock (accumulated patent count)	6750	62.63	133.90	0	1,639.10
Non-green knowledge stock (accumulated patent count)	6750	496.90	1,097.00	0	13,018.81

Note: RT: total number of observations; S.D.: standard deviation. The yearly data spans the period 1991 – 2015 ($T = 25$) and comprises 270 European NUTS-2 regions. All variables are, as explained in the text, based on the industry sector.

Table 2. The degree of cross-sectional dependence

	CD statistic	$\hat{\alpha}^*_{0.05}$	$\hat{\alpha}$	$\hat{\alpha}^*_{0.95}$
Log-levels				
Value added	418.88	0.925	0.966	1.007
Physical capital input	157.88	0.870	0.916	0.961
Labor input	355.42	0.928	0.967	1.007
Green knowledge stock	854.18	0.964	1.003	1.042
Non-green knowledge stock	901.26	0.963	1.003	1.043
First log-differences				
Value added	250.23	0.863	0.940	1.017
Physical capital input	127.45	0.800	0.877	0.954
Labor input	186.46	0.836	0.920	1.004
Green knowledge stock	241.33	0.862	0.918	0.973
Non-green knowledge stock	444.31	0.912	0.958	1.004

Note: Estimation of the bias-corrected version of α (Bailey et al., 2016b) and the CD statistic (Pesaran, 2004, 2015a). $\hat{\alpha}$ refers to the point estimate of the exponent of cross-sectional dependence according to equation (13) of Bailey et al. (2016b). $*$ 90% level confidence bands. We follow Bailey et al. (2016b) and Ertur and Musolesi (2017) in preferring Holm's procedure over Bonferroni's. The CD-statistic tends to $\mathcal{N}(0,1)$ under the null of weak CSD as N and $T \to \infty$ (Pesaran, 2015a).

Table 3. Main estimation: static production functions

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.133***	0.0990***	0.107***	0.138***	0.107***
	(0.0215)	(0.0129)	(0.0181)	(0.0135)	(0.0122)
Labor input	0.605***	0.239***	0.423***	0.470***	0.449***
	(0.0592)	(0.0299)	(0.0575)	(0.0377)	(0.0378)
Green knowledge stock	-0.104***	-0.0472**	-0.0623*	0.0112	-0.0308
	(0.0293)	(0.0198)	(0.0333)	(0.0369)	(0.0450)
Non-green knowledge stock	0.214***	0.0784***	0.0724**	0.173***	0.221***
	(0.0238)	(0.0207)	(0.0345)	(0.0437)	(0.0581)
Year dummies	Yes	Yes	No	Demeaned	No
Order of integration	I(1)	I(0)	I(0)	I(0)	I(0)
CD Test	0.79	2.81	6.00	8.13	3.55
$\bar{\hat{\rho}}$	0.001	0.003	0.006	0.009	0.004
$\lvert\bar{\hat{\rho}}\rvert$	0.449	0.202	0.298	0.208	0.210
$\hat{\alpha}$	0.559	0.630	0.858	0.646	0.615
$\tilde{\alpha}$	0.610	0.650	0.871	0.648	0.624
Observations	6750	6480	6750	6750	6750
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. Order of integration refers to the Pesaran (2007) test for unit roots. I(0) refers to the case where the null of a unit root is rejected at 10% level for all lag augmentations until two lags. I(0)/I(1) indicates mixed results, i.e. if null is rejected in some, but not all cases. I(1) refers to the case where the null is never rejected at 10% level. $\hat{\alpha}$ is the bias-corrected version of α given by equation (13) of Bailey et al. (2016b). $\tilde{\alpha}$ refers to the version robust against weak CSD in the errors and autocorrelation of the factors. We use four principal components to construct the estimate robust against weak CSD in the error term. $\bar{\hat{\rho}}$ is the average pairwise correlation coefficient, and $\lvert\bar{\hat{\rho}}\rvert$ the average pairwise absolute correlation coefficient.

Table 4. Dynamic production functions

	BB(1)	BB(2)	BB(3)	CS-DLP	CS-DLMG
Long-run coefficients					
Physical capital input	0.171	-0.193*	-0.222	0.125***	0.171***
	(0.114)	(0.105)	(0.181)	(0.242)	(0.026)
Labor input	0.319***	0.670***	0.326**	0.421***	0.280***
	(0.121)	(0.108)	(0.149)	(0.065)	(0.080)
Green knowledge stock	-0.761***	-0.381**	-0.195	-0.028	-0.086
	(0.151)	(0.159)	(0.215)	(0.051)	(0.104)
Non-green knowledge stock	0.864***	0.766***	0.708***	0.083*	0.265**
	(0.129)	(0.137)	(0.176)	(0.045)	(0.132)
Year dummies	Yes	Yes	Yes	No	No
Order of integration	I(0)	I(0)	I(0)	I(0)	I(0)
CD Test	2.95	3.88	-4.44	3.47	-0.78
$\hat{\alpha}$	0.641	0.585	0.592	0.747	0.524
$\tilde{\alpha}$	0.659	0.608	0.606	0.773	0.582
Observations	6480	6210	6210	6210	6210
Regions	270	270	270	270	270
No. of instruments	144	143	53	-	-
AR2 test	0.063	0.102	0.825	-	-
Sargan test	0.000	0.000	0.094	-	-
Hansen test	0.000	0.000	0.001	-	-

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are calculated with the delta method from the standard errors of the short-run coefficients robust against heteroskedasticity and autocorrelation for BB. CS-DL standard errors are based on the nonparametric variance estimators given in Chudik et al. (2016). Order of integration refers to the Pesaran (2007) test for unit roots. I(0) refers to the case where the null of a unit root is rejected at 10% level for all lag augmentations until two lags. I(0)/I(1) indicates mixed results, i.e. if null is rejected in some, but not all cases. I(1) refers to the case where the null is never rejected at 10% level. $\hat{\alpha}$ is the bias-corrected version of α given by equation (13) of Bailey et al. (2016b). $\tilde{\alpha}$ refers to the version robust against weak CSD in the errors and autocorrelation of the factors. We use four principal components to construct the estimate robust against weak CSD in the error term. Hansen / Sargan test and AR2 test refer to p-values.

Table 5. Static production functions: different subgroups of green technologies

	Alternative energy production		Recycling		Transportation		Energy efficiency	
	FD	CCEMG	FD	CCEMG	FD	CCEMG	FD	CCEMG
Physical capital input	0.100***	0.113***	0.112***	0.113***	0.112***	0.109***	0.0967***	0.114***
	(0.0133)	(0.0121)	(0.0153)	(0.0132)	(0.0148)	(0.0134)	(0.0125)	(0.0126)
Labor input	0.230***	0.421***	0.229***	0.436***	0.208***	0.422***	0.222***	0.411***
	(0.0294)	(0.0364)	(0.0325)	(0.0386)	(0.0308)	(0.0435)	(0.0302)	(0.0387)
Green knowledge stock	-0.0180	0.0490	-0.0129	-0.0268	-0.0136	-0.0128	-0.0199	-0.0664
	(0.0161)	(0.0371)	(0.0150)	(0.0697)	(0.0138)	(0.0296)	(0.0131)	(0.0407)
Non-green knowledge stock	0.0555***	0.131**	0.0549***	0.0340	0.0592***	0.151**	0.0581***	0.221***
	(0.0204)	(0.0542)	(0.0200)	(0.0536)	(0.0208)	(0.0686)	(0.0193)	(0.0677)
Year dummies	Yes	No	Yes	No	Yes	No	Yes	No
Observations	6,360	6,625	5,880	6,125	5,760	6,000	6,240	6,500
Regions	265	265	245	245	240	240	260	260

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for FD. Standard errors for CCEMG are based on the nonparametric variance estimators given in Pesaran (2006). For each subgroup there are different regions which have no single patent at all for each year. We exclude those regions for each subgroup, resulting in slightly different sample sizes for each subgroup. Non-green knowledge stocks are always constructed from total patents minus the green subgroup patents.

Table 6. Static production functions: EU15 subsample

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.0758***	0.0654***	0.0871***	0.137***	0.121***
	(0.0267)	(0.00970)	(0.0244)	(0.0152)	(0.0149)
Labor input	0.687***	0.208***	0.320***	0.456***	0.407***
	(0.0585)	(0.0326)	(0.0679)	(0.0430)	(0.0430)
Green knowledge stock	-0.0288	-0.0513**	-0.104**	-0.0314	-0.0242
	(0.0320)	(0.0200)	(0.0411)	(0.0496)	(0.0567)
Non-green knowledge stock	0.0240	0.00767	-0.0124	0.105	0.287***
	(0.0346)	(0.0277)	(0.0709)	(0.0766)	(0.0988)
Year dummies	Yes	Yes	No	Demeaned	No
Order of integration	I(1)	I(0)	I(0)	I(0)	I(0)
CD Test	4.19	6.59	9.68	5.34	2.82
$\bar{\hat{\rho}}$	0.006	0.009	0.013	0.007	0.004
$\lvert\bar{\hat{\rho}}\rvert$	0.445	0.209	0.285	0.206	0.208
$\hat{\alpha}$	0.697	0.702	0.830	0.725	0.563
$\tilde{\alpha}$	0.730	0.721	0.839	0.744	0.560
Observations	5200	4992	5200	5200	5200
Regions	208	208	208	208	208

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. Order of integration refers to the Pesaran (2007) test for unit roots. I(0) refers to the case where the null of a unit root is rejected at 10% level for all lag augmentations until two lags. I(0)/I(1) indicates mixed results, i.e. if null is rejected in some, but not all cases. I(1) refers to the case where the null is never rejected at 10% level. $\hat{\alpha}$ is the bias-corrected version of α given by equation (13) of Bailey et al. (2016b). $\tilde{\alpha}$ refers to the version robust against weak CSD in the errors and autocorrelation of the factors. We use four principal components to construct the estimate robust against weak CSD in the error term. $\bar{\hat{\rho}}$ is the average pairwise correlation coefficient, and $\lvert\bar{\hat{\rho}}\rvert$ the average pairwise absolute correlation coefficient.

Table 7. Static production functions: non-EU15 subsample

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.160***	0.157***	0.130***	0.0972***	0.100***
	(0.0437)	(0.0309)	(0.0322)	(0.0268)	(0.0253)
Labor input	0.490***	0.310***	0.352***	0.453***	0.281***
	(0.103)	(0.0476)	(0.0985)	(0.0990)	(0.0789)
Green knowledge stock	-0.0906*	-0.00654	-0.0379	-0.0984**	-0.149**
	(0.0460)	(0.0375)	(0.0681)	(0.0454)	(0.0665)
Non-green knowledge stock	0.248***	0.0467	0.00362	0.225***	0.116**
	(0.0437)	(0.0354)	(0.0452)	(0.0495)	(0.0570)
Year dummies	Yes	Yes	No	Demeaned	No
Order of integration	I(1)	I(0)	I(0)	I(0)	I(0)
CD Test	-2.91	4.63	0.39	1.05	0.54
$\bar{\hat{\rho}}$	-0.013	0.022	0.002	0.005	0.003
$\lvert\bar{\hat{\rho}}\rvert$	0.426	0.244	0.262	0.242	0.221
$\hat{\alpha}$	0.202	0.626	0.681	0.514	0.442
$\tilde{\alpha}$	0.665	0.641	0.767	0.553	0.500
Observations	1550	1488	1550	1550	1550
Regions	62	62	62	62	62

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. Order of integration refers to the Pesaran (2007) test for unit roots. I(0) refers to the case where the null of a unit root is rejected at 10% level for all lag augmentations until two lags. I(0)/I(1) indicates mixed results, i.e. if null is rejected in some, but not all cases. I(1) refers to the case where the null is never rejected at 10% level. $\hat{\alpha}$ is the bias-corrected version of α given by equation (13) of Bailey et al. (2016b). $\tilde{\alpha}$ refers to the version robust against weak CSD in the errors and autocorrelation of the factors. We use four principal components to construct the estimate robust against weak CSD in the error term. $\bar{\hat{\rho}}$ is the average pairwise correlation coefficient, and $\lvert\bar{\hat{\rho}}\rvert$ the average pairwise absolute correlation coefficient.

References

Alpay, E., Kerkvliet, J., and Buccola, S. (2002). Productivity Growth and Environmental Regulation in Mexican and U.S. Food Manufacturing. *American Journal of Agricultural Economics*, 84(4):887–901.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277–297.

Bailey, N., Holly, S., and Pesaran, M. H. (2016a). A Two-Stage Approach to Spatio-Temporal Analysis with Strong and Weak Cross-Sectional Dependence. *Journal of Applied Econometrics*, 31(1):249–280.

Bailey, N., Kapetanios, G., and Pesaran, M. H. (2016b). Exponent of Cross-Sectional Dependence: Estimation and Inference. *Journal of Applied Econometrics*, 31(6):929–960.

Barbieri, N., Ghisetti, C., Gilli, M., Marin, G., and Nicolli, F. (2016). A survey of the literature on environmental innovation based on main path analysis. *Journal of Economic Surveys*, 30(3):596–623.

Berman, E. and Bui, L. T. M. (2001). Environmental Regulation and Productivity: Evidence from Oil Refineries. *The Review of Economics and Statistics*, 83(3):498–510.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Bond, S., Leblebicioğlu, A., and Schiantarelli, F. (2010). Capital accumulation and growth: a new look at the empirical evidence. *Journal of Applied Econometrics*, 25(7):1073–1099.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, pages 679–741. North-Holland, Amsterdam.

Chudik, A., Mohaddes, K., Pesaran, M. H., and Raissi, M. (2016). Long-Run Effects in Large Heterogeneous Panel Data Models with Cross-Sectionally Correlated Errors. In González-Rivera, G., Hill, R. C., and Lee, T.-H., editors, *Essays in Honor of Aman Ullah*, volume 36 of *Advances in Econometrics*, pages 85–135. Emerald Group Publishing Limited, Bingley.

Chudik, A. and Pesaran, H. (2015a). Large Panel Data Models with Cross-Sectional Dependence: A Survey. In Baltagi, B. H., editor, *The Oxford Handbook of Panel Data*, pages 3–45. Oxford University Press, Oxford.

Chudik, A. and Pesaran, M. H. (2015b). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, 188(2):393–420.

Chudik, A., Pesaran, M. H., and Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels. *Econometrics Journal*, 14(1):C45–C90.

Ciccarelli, C. and Elhorst, J. P. (2018). A dynamic spatial econometric diffusion model with common factors: The rise and spread of cigarette consumption in Italy. *Regional Science and Urban Economics*, 72:131–142.

Coe, D. T. and Helpman, E. (1995). International R&D spillovers. *European Economic Review*, 39(5):859–887.

Costantini, V., Crespi, F., Marin, G., and Paglialunga, E. (2017). Eco-innovation, sustainable supply chains and environmental performance in European industries. *Journal of Cleaner Production*, 155:141–154.

De Vos, I. and Everaert, G. (2021). Bias-Corrected Common Correlated Effects Pooled Estimation in Dynamic Panels. *Journal of Business & Economic Statistics*, 39(1):294–306.

De Vos, I. and Westerlund, J. (2019). On CCE estimation of factor-augmented models when regressors are not linear in the factors. *Economics Letters*, 178:5–7.

Dernis, H. and Khan, M. (2004). Triadic Patent Families Methodology. OECD Science, Technology and Industry Working Papers No. 2004/02, OECD Publishing, Paris.

Ditzen, J. (2019). XTDCCE2: Stata module to estimate heterogeneous coefficient models using common correlated effects in a dynamic panel. Statistical Software Components No. S458204, Boston College Department of Economics.

Dufour, C., Lanoie, P., and Patry, M. (1998). Regulation and Productivity. *Journal of Productivity Analysis*, 9(3):233–247.

Eberhardt, M. (2011). MULTIPURT: Stata module to run 1st and 2nd generation panel unit root tests for multiple variables and lags. Statistical Software Components No. S457239, Boston College Department of Economics.

Eberhardt, M. and Bond, S. (2009). Cross-section dependence in nonstationary panel models: a novel estimator. MPRA Paper No. 17870.

Eberhardt, M., Helmers, C., and Strauss, H. (2013). Do Spillovers Matter When Estimating Private Returns to R&D? *The Review of Economics and Statistics*, 95(2):436–448.

Eberhardt, M. and Teal, F. (2011). Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics. *Journal of Economic Surveys*, 25(1):109–155.

Eberhardt, M. and Teal, F. (2013). No Mangoes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis. *Oxford Bulletin of Economics and Statistics*, 75(6):914–939.

Ellerman, A. D., Marcantonini, C., and Zaklan, A. (2016). The European Union emissions trading system: ten years and counting. *Review of Environmental Economics and Policy*, 10(1):89–107.

Ertur, C. and Musolesi, A. (2017). Weak and Strong Cross-Sectional Dependence: A Panel Data Analysis of International Technology Diffusion. *Journal of Applied Econometrics*, 32(3):477–503.

European Commission (2019). The European Green Deal. Communication from the Commission COM(2019) 640 final, Brussels.

Everaert, G. and De Groote, T. (2016). Common Correlated Effects Estimation of Dynamic Panels with Cross-Sectional Dependence. *Econometric Reviews*, 35(3):428–463.

Frondel, M., Horbach, J., and Rennings, K. (2007). End-of-pipe or cleaner production? An empirical comparison of environmental innovation decisions across OECD countries. *Business Strategy and the Environment*, 16(8):571–584.

Ghisetti, C. and Quatraro, F. (2017). Green Technologies and Environmental Productivity: A Cross-sectoral Analysis of Direct and Indirect Effects in Italian Regions. *Ecological Economics*, 132:1–13.

Ghisetti, C. and Rennings, K. (2014). Environmental innovations and profitability: how does it pay to be green? An empirical analysis on the German innovation survey. *Journal of Cleaner Production*, 75:106–117.

Gollop, F. M. and Roberts, M. J. (1983). Environmental Regulations and Productivity Growth: The Case of Fossil-fueled Electric Power Generation. *Journal of Political Economy*, 91(4):654–674.

Gray, W. B. (1987). The Cost of Regulation: OSHA, EPA and the Productivity Slowdown. *The American Economic Review*, 77(5):998–1006.

Gray, W. B. and Shadbegian, R. J. (2003). Plant vintage, technology, and environmental regulation. *Journal of Environmental Economics and Management*, 46(3):384–402.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10(1):92–116.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4):1661–1707.

Haščič, I. and Migotto, M. (2015). Measuring environmental innovation using patent data. OECD Environment Working Papers No. 89, OECD Publishing, Paris.

Iammarino, S., Rodriguez-Pose, A., and Storper, M. (2019). Regional inequality in Europe: evidence, theory and policy implications. *Journal of Economic Geography*, 19(2):273–298.

Inklaar, R. and Timmer, M. P. (2013). Capital, Labor and TFP in PWT 8.0. Technical Report, University of Groningen, Groningen Growth and Development Centre.

Johnstone, N., Haščič, I., and Popp, D. (2010). Renewable Energy Policies and Technological Innovation: Evidence Based on Patent Counts. *Environmental and Resource Economics*, 45(1):133–155.

Jones, C. I. (1995). Time Series Tests of Endogenous Growth Models. *The Quarterly Journal of Economics*, 110(2):495–525.

Kapetanios, G., Pesaran, M. H., and Yamagata, T. (2011). Panels with non-stationary multifactor error structures. *Journal of Econometrics*, 160(2):326–348.

Keller, W. (2004). International Technology Diffusion. *Journal of Economic Literature*, 42(3):752–782.

Kemp, R. (2010). Eco-Innovation: definition, measurement and open research issues. *Economia politica*, 27(3):397–420.

Kiviet, J. F. (2020). Microeconometric dynamic panel data methods: Model specification and selection issues. *Econometrics and Statistics*, 13:16–45.

Klarl, T. (2016). Pollution externalities, endogenous health and the speed of convergence in an endogenous growth model. *Journal of Macroeconomics*, 50:98–113.

Kruse, J. and Wetzel, H. (2016). Energy Prices, Technological Knowledge, and Innovation in Green Energy Technologies: a Dynamic Panel Analysis of European Patent Data. *Cesifo Economic Studies*, 62(3):397–425.

Lanjouw, J. O. and Mody, A. (1996). Innovation and the international diffusion of environmentally responsive technology. *Research Policy*, 25(4):549–571.

Lanoie, P., Patry, M., and Lajeunesse, R. (2008). Environmental regulation and productivity: testing the porter hypothesis. *Journal of Productivity Analysis*, 30(2):121–128.

Lewandowski, P. (2007). PESCADF: Stata module to perform Pesaran's CADF panel unit root test in presence of cross section dependence. Statistical Software Components No. S456732, Boston College Department of Economics.

Álvarez, I. C., Barbero, J., and Zofío, J. L. (2017). A Panel Data Toolbox for MATLAB. *Journal of Statistical Software*, 76(6):1–27.

Maddala, G. S., Trost, R. P., Li, H., and Joutz, F. (1997). Estimation of Short-Run and Long-Run Elasticities of Energy Demand From Panel Data Using Shrinkage Estimators. *Journal of Business & Economic Statistics*, 15(1):90–100.

Marin, G. and Lotti, F. (2017). Productivity effects of eco-innovations using data on eco-patents. *Industrial and Corporate Change*, 26(1):125–148.

McCann, P. and Ortega-Argilés, R. (2015). Smart Specialization, Regional Growth and Applications to European Union Cohesion Policy. *Regional Studies*, 49(8):1291–1302.

Merryman, S. (2005). XTFISHER: Stata module to compute Fisher type unit root test for panel data. Statistical Software Components No. S448201, Boston College Department of Economics.

Millo, G. (2019). Private returns to R&D in the presence of spillovers, revisited. *Journal of Applied Econometrics*, 34(1):155–159.

Mitze, T., Naveed, A., and Ahmad, N. (2016). International, intersectoral, or unobservable? Measuring R&D spillovers under weak and strong cross-sectional dependence. *Journal of Macroeconomics*, 50:259–272.

Murty, M. N. and Kumar, S. (2003). Win–win opportunities and environmental regulation: testing of porter hypothesis for Indian manufacturing industries. *Journal of Environmental Management*, 67(2):139–144.

O'Mahony, M. and Timmer, M. P. (2009). Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database. *The Economic Journal*, 119(538):F374–F403.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446.

Pesaran, M. H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels. CESifo Working Paper Series No. 1229. IZA Discussion Paper No. 1240.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2):265–312.

Pesaran, M. H. (2015a). Testing Weak Cross-Sectional Dependence in Large Panels. *Econometric Reviews*, 34(6-10):1089–1117.

Pesaran, M. H. (2015b). *Time series and panel data econometrics*. Oxford University Press, Oxford.

Pesaran, M. H., Shin, Y., and Smith, R. P. (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association*, 94(446):621–634.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.

Pesaran, M. H. and Tosetti, E. (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics*, 161(2):182–202.

Popp, D. (2010). Innovation and Climate Policy. *Annual Review of Resource Economics*, 2(1):275–298.

Popp, D., Hascic, I., and Medhi, N. (2011). Technology and the diffusion of renewable energy. *Energy Economics*, 33(4):648–662.

Porter, M. E. and van der Linde, C. (1995). Toward a New Conception of the Environment-Competitiveness Relationship. *Journal of Economic Perspectives*, 9(4):97–118.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. URL http://www.R-project.org/.

Rebelo, S. (1991). Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy*, 99(3):500–521.

Rexhäuser, S. and Rammer, C. (2014). Environmental Innovations and Firm Profitability: Unmasking the Porter Hypothesis. *Environmental and Resource Economics*, 57(1):145–167.

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy*, 94(5):1002–1037.

Roodman, D. (2009a). How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal*, 9(1):86–136.

Roodman, D. (2009b). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Sarafidis, V. and Robertson, D. (2009). On the impact of error cross-sectional dependence in short dynamic panel estimation. *The Econometrics Journal*, 12(1):62–81.

Sarafidis, V. and Wansbeek, T. (2012). Cross-Sectional Dependence in Panel Data Analysis. *Econometric Reviews*, 31(5):483–531.

Schmoch, U., Laville, F., Patel, P., and Frietsch, R. (2003). Linking Technology Areas to Industrial Sectors. Final Report to the European Commission, DG Research.

Simon, C. P. and Blume, L. (1994). *Mathematics for economists*. W. W. Norton & Company, New York.

Smith, J. B. and Sims, W. A. (1985). The Impact of Pollution Charges on Productivity Growth in Canadian Brewing. *The RAND Journal of Economics*, 16(3):410–423.

Soltmann, C., Stucki, T., and Woerter, M. (2015). The Impact of Environmentally Friendly Innovations on Value Added. *Environmental and Resource Economics*, 62(3):457–479.

Stucki, T. and Woerter, M. (2019). The private returns to knowledge: A comparison of ICT, biotechnologies, nanotechnologies, and green technologies. *Technological Forecasting and Social Change*, 145:62–81.

Tennekes, M. (2018). tmap: Thematic Maps in R. *Journal of Statistical Software*, 84(6):1–39.

Turnovsky, S. J. and Monteiro, G. (2007). Consumption externalities, production externalities, and efficient capital accumulation under time non-separable preferences. *European Economic Review*, 51(2):479–504.

Van Leeuwen, G. and Mohnen, P. (2017). Revisiting the Porter hypothesis: an empirical analysis of Green innovation for the Netherlands. *Economics of Innovation and New Technology*, 26(1-2):63–77.

Verdolini, E. and Galeotti, M. (2011). At home and abroad: An empirical analysis of innovation and diffusion in energy technologies. *Journal of Environmental Economics and Management*, 61(2):119–134.

Wendler, T. (2019). About the Relationship Between Green Technology and Material Usage. *Environmental and Resource Economics*, 74(3):1383–1423.

Westerlund, J. (2018). CCE in panels with general unknown factors. *The Econometrics Journal*, 21(3):264–276.

Wurlod, J.-D. and Noailly, J. (2018). The impact of green innovation on energy intensity: An empirical analysis for 14 industrial sectors in OECD countries. *Energy Economics*, 71:47–61.

Xepapadeas, A. and de Zeeuw, A. (1999). Environmental Policy and Competitiveness: The Porter Hypothesis and the Composition of Capital. *Journal of Environmental Economics and Management*, 37(2):165–182.

Appendix

Part A: Proofs and Theory

Proof of Proposition 1

Linearizing (2.20)-(2.22) around their steady states, we may express the linearized core dynamics of our model in the form

$$\begin{bmatrix} \dot{c} \\ \dot{k}_p \\ \dot{k}_g \end{bmatrix} = \Pi \begin{bmatrix} c - \tilde{c} \\ k_p - \tilde{k}_p \\ k_g - \tilde{k}_g \end{bmatrix},$$

with

$$\Pi \equiv \begin{bmatrix} 0 & (\sigma_p - 1)\left[\delta + \rho + n\left(1 - \theta + \frac{\sigma_n \theta}{1-\tilde{\sigma}}\right)\right]\tilde{k}_p^{-1} & \sigma_g\left[\delta + \rho + n\left(1 - \theta + \frac{\sigma_n \theta}{1-\tilde{\sigma}}\right)\right]\tilde{k}_g^{-1} \\ -(1-\tau) & \pi_{22} & (1-\tau)\frac{\sigma_g}{\kappa\alpha(1+\xi_g)}\left[\delta + \rho + n\left(1 - \theta + \frac{\sigma_n \theta}{1-\tilde{\sigma}}\right)\right] \\ -\tau & \tau\frac{\sigma_p}{(1-\kappa)\alpha(1-\xi_p)}\left[\delta + \rho + n\left(1 - \theta + \frac{\sigma_n \theta}{1-\tilde{\sigma}}\right)\right] & \pi_{33} \end{bmatrix},$$

and

$$\pi_{22} \equiv \frac{(1-\tau)\sigma_p\left[n\left(\frac{\theta\sigma_n}{1-\tilde{\sigma}} - \theta + 1\right) + \delta + \rho\right]}{\alpha(1-\kappa)\left(1-\xi_p\right)} - \frac{n\sigma_n}{1-\tilde{\sigma}} - \delta,$$

$$\pi_{33} \equiv \frac{\tau\sigma_g\left[n\left(\frac{\theta\sigma_n}{1-\tilde{\sigma}} - \theta + 1\right) + \delta + \rho\right]}{\alpha\kappa\left(\xi_g + 1\right)} - \frac{n\sigma_n}{1-\tilde{\sigma}} - \delta.$$

The first Eigenvalue of matrix Π is given by $\frac{\delta(\tilde{\sigma}-1)-n\sigma_n}{1-\tilde{\sigma}}$ which is negative provided that $\tilde{\sigma} < 1$. As the Determinant of Π, $det(\Pi)$ is positive for $\tilde{\sigma} < 1$, the second and third Eigenvalue must be of opposite sign as the Determinant of a symmetric matrix is also equal to the product of its Eigenvalues (Simon and Blume, 1994, Theorem 23.9). As a consequence, we have two negative and one positive Eigenvalue for $\tilde{\sigma} < 1$. Thus, the stable manifold, which is the hyperplane generated by the associated Eigenvectors, has dimension two (see Simon and Blume, 1994). In a nutshell, since our system features two state variables, k_g and k_p, and one jump variable, c, the equilibrium yields a unique stable saddle path.

For our calibrated example, we obtain the following real Eigenvalues: $\lambda_1 = 0.253534, \lambda_2 = -0.152932, \lambda_3 = -0.0410526$, and $det(\Pi) = 0.00159175$.

Impulse Responses for the Environmental Policy Shock

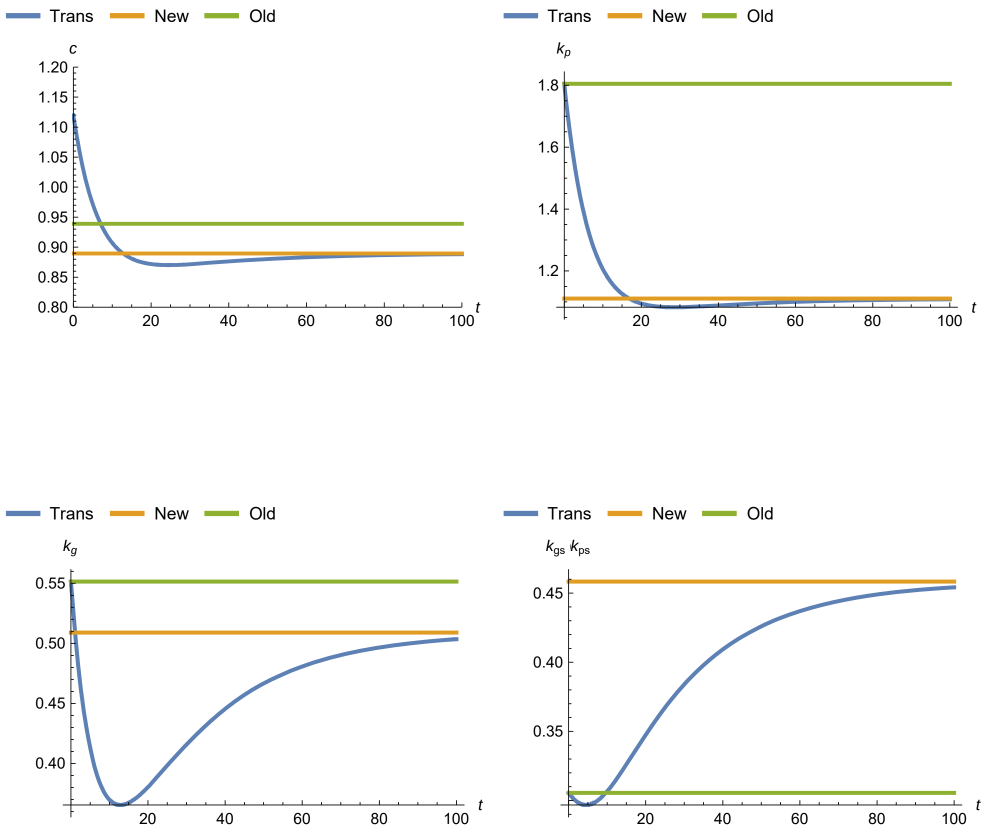


Figure A1. Transitional dynamics for moderate environmental externalities ($0 < \sigma_p < 1$ and $\sigma_g < 1$). *Note:* For the majority of the parameter space, we rely on standard values from the literature: $\theta \to 2.5, n \to 0.02, \rho \to 0.04, \sigma_n \to 0.8, \delta \to 0.02, \alpha \to 0.2, \beta \to 0.1, \kappa \to 0.2, \xi_g \to 0.1, \xi_p \to 0.1, \mu \to 0.3, \phi_p \to 0.3, \phi_g \to 0.1$. Thus, this calibration indicates slight increasing returns to scale: $\sigma_n + \tilde{\sigma} = 1.04$. Moreover, we have $\sigma_p = 0.15$ and $\sigma_g = 0.09$ showing a moderate influence of non-internalized, environmental externalties.

Part B: Time-Series Properties and Descriptive Statistics

Correlations

Table B1. Pairwise correlation matrix: total variation

	y_{it}	$k_{k,it}$	l_{it}	$k_{g,it}$	$k_{p,it}$	$\bar{y}_t$	$\bar{k}_{k,t}$	$\bar{l}_t$	$\bar{k}_{g,t}$	$\bar{k}_{p,t}$
y_{it}	1									
$k_{k,it}$	.9335	1								
l_{it}	.6921	.6628	1							
$k_{g,it}$	.7955	.7236	.3121	1						
$k_{p,it}$	.8198	.7518	.3142	.9723	1					
$\bar{y}_t$	.1342	.1166	-.0650	.2760	.2586	1				
$\bar{k}_{k,t}$	.1291	.1211	-.0558	.2605	.2457	.9626	1			
$\bar{l}_t$	-.1110	-.0860	.0785	-.2686	-.2461	-.8276	-.7100	1		
$\bar{k}_{g,t}$	.1290	.1099	-.0735	.2871	.2667	.9616	.9074	-.9358	1	
$\bar{k}_{p,t}$	.1299	.1114	-.0724	.2867	.2670	.9685	.9200	-.9218	.9989	1

Note: Pairwise unconditional correlation coefficients. Cross-sectional averages are included.

Table B2. Pairwise correlation matrix: within dimension

	y_{it}	$k_{k,it}$	l_{it}	$k_{g,it}$	$k_{p,it}$	$\bar{y}_t$	$\bar{k}_{k,t}$	$\bar{l}_t$	$\bar{k}_{g,t}$	$\bar{k}_{p,t}$
y_{it}	1									
$k_{k,it}$	.4937	1								
l_{it}	-.0267	-.0837	1							
$k_{g,it}$	.6017	.3420	-.4472	1						
$k_{p,it}$	.6892	.4019	-.4416	.9161	1					
$\bar{y}_t$	.6633	.3889	-.4651	.8560	.8641	1				
$\bar{k}_{k,t}$	.6385	.4040	-.3990	.8078	.8209	.9626	1			
$\bar{l}_t$	-.5490	-.2869	.5620	-.8331	-.8224	-.8276	-.7100	1		
$\bar{k}_{g,t}$	.6378	.3666	-.5259	.8902	.8912	.9616	.9074	-.9358	1	
$\bar{k}_{p,t}$	.6424	.3717	-.5180	.8892	.8922	.9685	.9200	-.9218	.9989	1

Note: Pairwise unconditional correlation coefficients for the within-dimension. Cross-sectional averages are included.

Table B3. Correlation between different knowledge stocks: total variation

	$k^a_{g,it}$	$k^a_{p,it}$	$k^b_{g,it}$	$k^b_{p,it}$
$k^a_{g,it}$	1			
$k^a_{p,it}$	.9723	1		
$k^b_{g,it}$	.9948	.9746	1	
$k^b_{p,it}$	.9629	.9969	.9709	1

Note: Pairwise unconditional correlation coefficients. Compared are knowledge stocks based on our main approach ($k^a_{g,it}$) with those computed with the perpetual inventory method with a depreciation rate of 10 % ($k^b_{g,it}$).

Table B4. Correlation between different knowledge stocks: within dimension

	$k^a_{g,it}$	$k^a_{p,it}$	$k^b_{g,it}$	$k^b_{p,it}$
$k^a_{g,it}$	1			
$k^a_{p,it}$	.9161	1		
$k^b_{g,it}$	.9613	.9030	1	
$k^b_{p,it}$	.8433	.9726	.8701	1

Note: Pairwise unconditional correlation coefficients for the within dimension. Compared are knowledge stocks based on our main approach ($k^a_{g,it}$) with those computed with the perpetual inventory method with a depreciation rate of 10 % ($k^b_{g,it}$).

PURTs

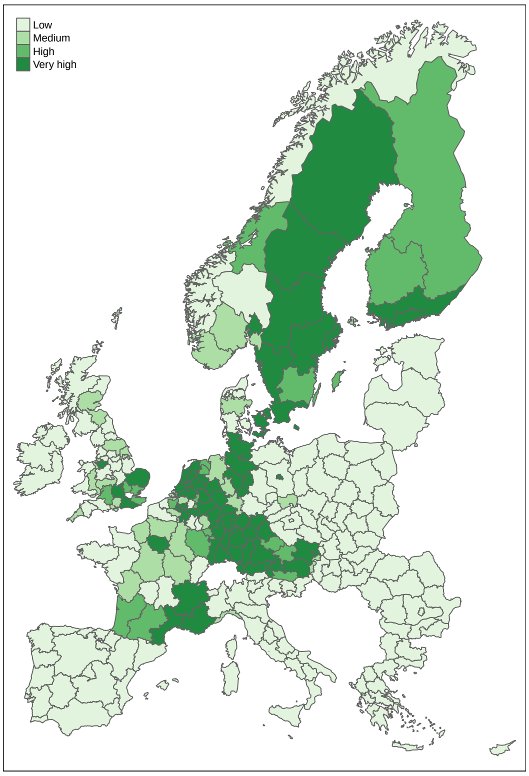
Table B5. Panel unit root tests: constant

lags	y_{it}	l_{it}	$k_{k,it}$	$k_{g,it}$	$k_{p,it}$
0	-1.82(0.03)	-3.16(0.00)	-8.95(0.00)	-0.84(0.20)	-4.90(0.00)
1	-0.63(0.27)	-6.37(0.00)	-0.97(0.17)	-4.31(0.00)	-1.55(0.06)
2	0.44(0.67)	-2.35(0.01)	2.00(0.98)	-3.07(0.00)	0.20(0.58)
3	0.85(0.80)	1.14(0.87)	6.50(1.00)	0.21(0.59)	-0.35(0.36)

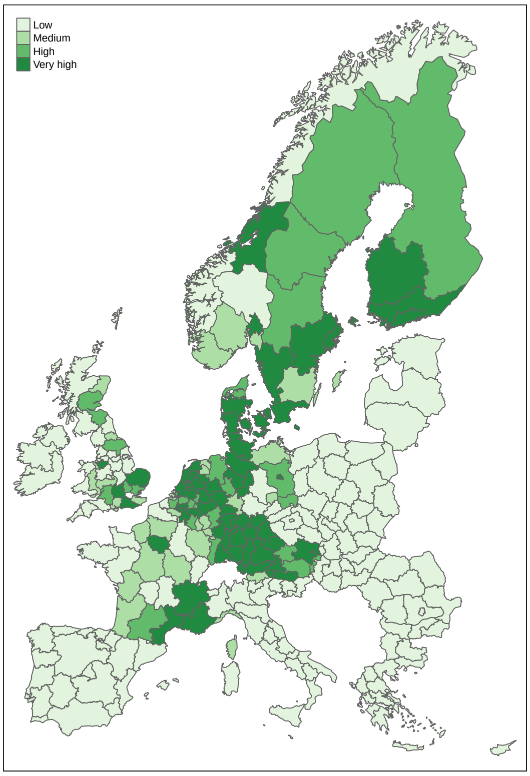
Note: Panel unit root test of the second generation of Pesaran (2007). Constant added, no trend. Reported are Z statistics and p-values in brackets. All individual groups are integrated of order 1 under the null hypothesis. Implemented in STATA with the *multipurt* routine written by Eberhardt (2011), making use of the *pescadf* command by Lewandowski (2007) and the *xtfisher* routine by Merryman (2005).

Sample Overview

The following maps are created in R (R Core Team, 2020) with the *tmap* package (Tennekes, 2018), additionally complemented mainly with the package *sf* (Pebesma, 2018). The categories on which the coloring is based on are adopted from Iammarino et al. (2019) and correspond to: "Very high": 150% of the average (over all regions) or greater; "High": 120-149 % of the average; "Medium": 75-119 % of the average; "Low": less than 75% of the average.

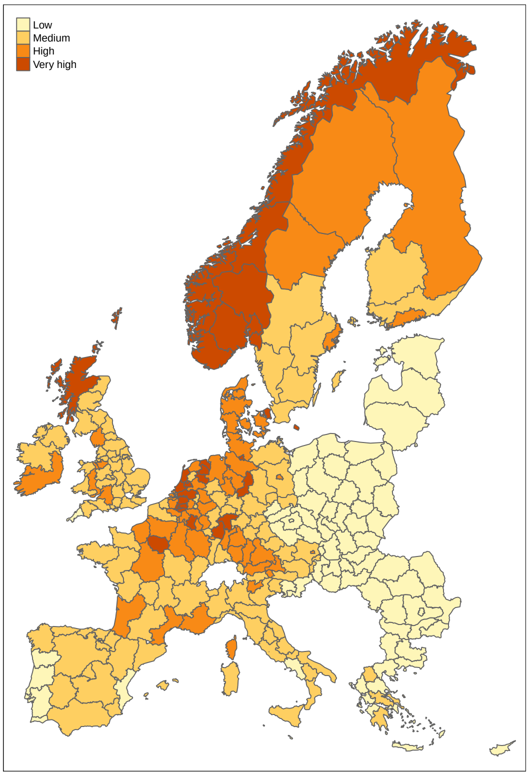


(a) 1991-2003

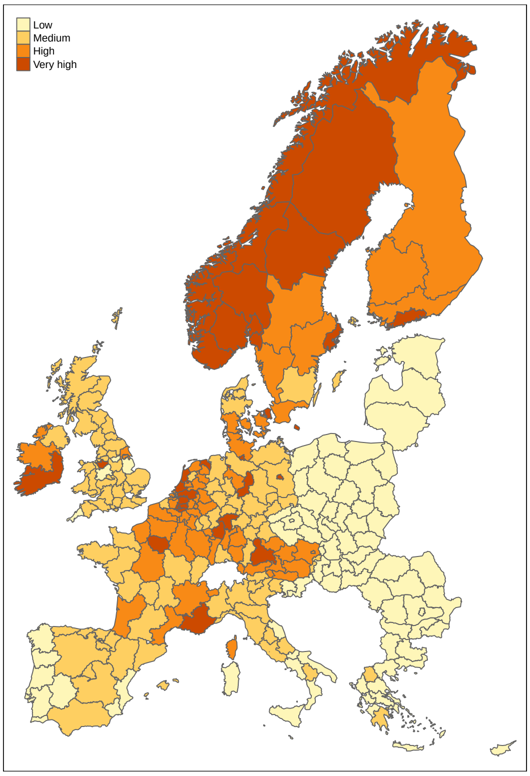


(b) 2004-2015

Figure B1. Labor deflated green knowledge stocks. *Note:* Green knowledge stocks divided by labor input averaged for the first and the second period of the sample. For greater clarity, we excluded some regions (FRA2, FRA3, FRA4, ES70, PT20, PT30), which are geographically far away. Shapefiles are obtained from Eurostat, ©EuroGeographics for the administrative boundaries.



(a) 1991-2003



(b) 2004-2015

Figure B2. Labor productivity. *Note:* Value added divided by labor input averaged for the first and the second period of the sample. For greater clarity, we excluded some regions (FRA2, FRA3, FRA4, ES70, PT20, PT30), which are geographically far away. Shapefiles are obtained from Eurostat, ©EuroGeographics for the administrative boundaries.

Part C: Further Robustness Results

Table C1. Static production functions: lagged inputs

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.114***	-0.000344	0.0216	0.0847***	0.00363
	(0.0218)	(0.00740)	(0.0147)	(0.0122)	(0.0107)
Labor input	0.530***	0.110***	0.271***	0.304***	0.120***
	(0.0595)	(0.0194)	(0.0520)	(0.0355)	(0.0351)
Green knowledge stock	-0.105***	-0.0437**	-0.0411	-0.0364	-0.0579
	(0.0307)	(0.0178)	(0.0415)	(0.0387)	(0.0457)
Non-green knowledge stock	0.228***	0.0891***	0.0661*	0.200***	0.0806
	(0.0244)	(0.0195)	(0.0357)	(0.0450)	(0.0517)
Year dummies	Yes	Yes	No	Demeaned	No
Observations	6480	6210	6480	6480	6480
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. All production inputs are lagged by one period.

Table C2. Static production functions: physical capital stock

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.538***	0.255***	-0.0166	0.400***	0.223**
	(0.0440)	(0.0721)	(0.115)	(0.0592)	(0.110)
Labor input	0.447***	0.256***	0.465***	0.518***	0.504***
	(0.0456)	(0.0314)	(0.0643)	(0.0368)	(0.0396)
Green knowledge stock	-0.0763***	-0.0500**	-0.0748**	-0.0589	-0.0289
	(0.0236)	(0.0209)	(0.0319)	(0.0412)	(0.0552)
Non-green knowledge stock	0.142***	0.0511**	0.0741**	0.160***	0.129
	(0.0198)	(0.0231)	(0.0300)	(0.0572)	(0.0827)
Year dummies	Yes	Yes	No	Demeaned	No
Observations	6750	6480	6750	6750	6750
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. Physical capital input refers now to a capital stock computed with the perpetual inventory method from gross fixed capital formation.

Table C3. Static production functions: no physical capital

	2FE	FD	CCEP	MG	CCEMG
Labor input	0.638***	0.265***	0.462***	0.596***	0.516***
	(0.0574)	(0.0314)	(0.0576)	(0.0433)	(0.0376)
Green knowledge stock	-0.121***	-0.0508**	-0.0599*	-0.00420	0.00327
	(0.0304)	(0.0210)	(0.0334)	(0.0415)	(0.0461)
Non-green knowledge stock	0.246***	0.0846***	0.0819***	0.168***	0.144**
	(0.0257)	(0.0220)	(0.0298)	(0.0474)	(0.0569)
Year dummies	Yes	Yes	No	Demeaned	No
Observations	6750	6480	6750	6750	6750
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively.

Table C4. Static production functions: diffusion stocks and applicant

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.157***	0.0997***	0.113***	0.126***	0.104***
	(0.0223)	(0.0129)	(0.0192)	(0.0137)	(0.0134)
Labor input	0.576***	0.236***	0.365***	0.463***	0.427***
	(0.0608)	(0.0299)	(0.0476)	(0.0407)	(0.0420)
Green knowledge stock	-0.0544**	-0.0152	-0.0159	0.0320	-0.0265
	(0.0239)	(0.0165)	(0.0261)	(0.0321)	(0.0446)
Non-green knowledge stock	0.150***	0.0287	0.0130	0.149***	0.127**
	(0.0244)	(0.0202)	(0.0291)	(0.0366)	(0.0577)
Year dummies	Yes	Yes	No	Demeaned	No
Observations	6750	6480	6750	6750	6750
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. Knowledge stocks are computed as in the main model, but patents are assigned to regions based on the address of the applicant of the patent.

Table C5. Static production functions: perpetual inventory stocks and inventor

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.126***	0.0990***	0.115***	0.133***	0.126***
	(0.0206)	(0.0129)	(0.0194)	(0.0145)	(0.0133)
Labor input	0.592***	0.237***	0.474***	0.488***	0.480***
	(0.0577)	(0.0301)	(0.0566)	(0.0384)	(0.0347)
Green knowledge stock	-0.0388*	-0.00266	-0.0255	0.0162	0.0256
	(0.0216)	(0.00933)	(0.0173)	(0.0189)	(0.0230)
Non-green knowledge stock	0.175***	0.0321***	0.0201	0.110***	0.119***
	(0.0180)	(0.0113)	(0.0214)	(0.0266)	(0.0344)
Year dummies	Yes	Yes	No	Demeaned	No
Observations	6750	6480	6750	6750	6750
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. Knowledge stocks are computed with the perpetual inventory method. Patent assignment is based on the address of the inventor, as in the main approach.

Table C6. Static production functions: U-shaped relationship of green knowledge

	2FE	FD	CCEP	MG	CCEMG
Physical capital input	0.131***	0.0990***	0.116***	0.115***	0.134***
	(0.0217)	(0.0129)	(0.0211)	(0.0125)	(0.0147)
Labor input	0.597***	0.238***	0.416***	0.451***	0.401***
	(0.0581)	(0.0300)	(0.0570)	(0.0376)	(0.0390)
Green knowledge stock	-0.0648	-0.0392	-0.164***	-0.0115	0.0235
	(0.0403)	(0.0285)	(0.0599)	(0.0911)	(0.194)
Non-green knowledge stock	0.196***	0.0771***	0.104***	0.220***	0.138**
	(0.0262)	(0.0212)	(0.0306)	(0.0512)	(0.0641)
(Green knowledge stock ×	-0.00623	-0.00194	0.0303**	-0.00888	-0.0374
Green knowledge stock)	(0.00465)	(0.00373)	(0.0120)	(0.0140)	(0.0833)
Year dummies	Yes	Yes	No	Demeaned	No
Observations	6,750	6,480	6,750	6,750	6,750
Regions	270	270	270	270	270

Note: Asterisks indicate significance at * 10%, ** 5%, *** 1%. Standard errors in parentheses are of heteroskedasticity-robust sandwich type for 2FE and FD. Standard errors for CCEP, CCEMG and MG are based on the nonparametric variance estimators given in Pesaran (2006) and Pesaran et al. (1999), respectively. To allow for nonlinearities, the procedure described by De Vos and Westerlund (2019) is used. Specifically, we augment the regression by cross-sectional averages of the linear regressors only, excluding cross-sectional averages of the dependent variable and the squared term.