

Global Energy Use: Decoupling or Convergence?

Zsuzsanna Csereklyei^a

David I. Stern^b

August 24, 2015

Abstract

We examine the key factors driving change in energy use globally over the past four decades. We test for both strong decoupling where economic growth has less effect on energy use as income increases, and weak decoupling where energy use declines over time in richer countries, *ceteris paribus*. Our econometric approach is robust to the presence of unit roots, unobserved time effects, and spatial effects. Our key findings are that the growth of per capita energy use has been primarily driven by economic growth, convergence in energy intensity, and weak decoupling. There is no sign of strong decoupling.

JEL Classification: Q43, O13

Keywords: Energy consumption, convergence, decoupling

1. Introduction

Though global energy use continues to increase (BP, 2015), energy intensity has declined faster in some developed countries such as the United States (-1.94% annually from 1971 to 2010) and the United Kingdom (-2.63%) than in the world as a whole (-1.08%).¹ Does this mean that economic growth has less of an effect on growth in energy use in richer countries - a decoupling of energy and growth - or is this due to convergence of formerly more energy intensive countries towards the global mean? Here, we show that: economic growth has a similar effect on the growth of energy use across the full income continuum from less developed to highly developed countries; convergence is very important in explaining the evolution of energy use; but that, *ceteris paribus*, energy use per capita declines autonomously (not associated with growth) in high income countries.

The rise in energy consumption of rapidly growing developing countries, especially from China and India, has accounted for the vast majority of the global increase in energy use in recent years.

^aLudwig Maximilian University of Munich, Oettingenstrasse 67, 80538-Munich, Germany [✉] Zsuzsanna.Csereklyei@gsi.uni-muenchen.de.

^bThe Australian National University, 132 Lennox Crossing, Acton, ACT 2601, Australia [✉] david.stern@anu.edu.au.

¹Data sources are provided in the Appendix.

Non-OECD countries currently account for approximately 60 percent of global energy demand, which is predicted to rise to 70% by 2040 (International Energy Agency, 2014). This increasing energy use exacerbates environmental problems including global climate change due to greenhouse gas emissions and local environmental problems such as the recent episodes of extreme air pollution in Beijing and other Chinese cities. Besides its environmental impacts, increasing energy use also raises questions of national energy supply security. As the share of world energy use consumed in developing countries increases, it is increasingly important to understand how energy use evolves across the full income continuum from less developed to highly developed countries (Ruijven et al., 2009).

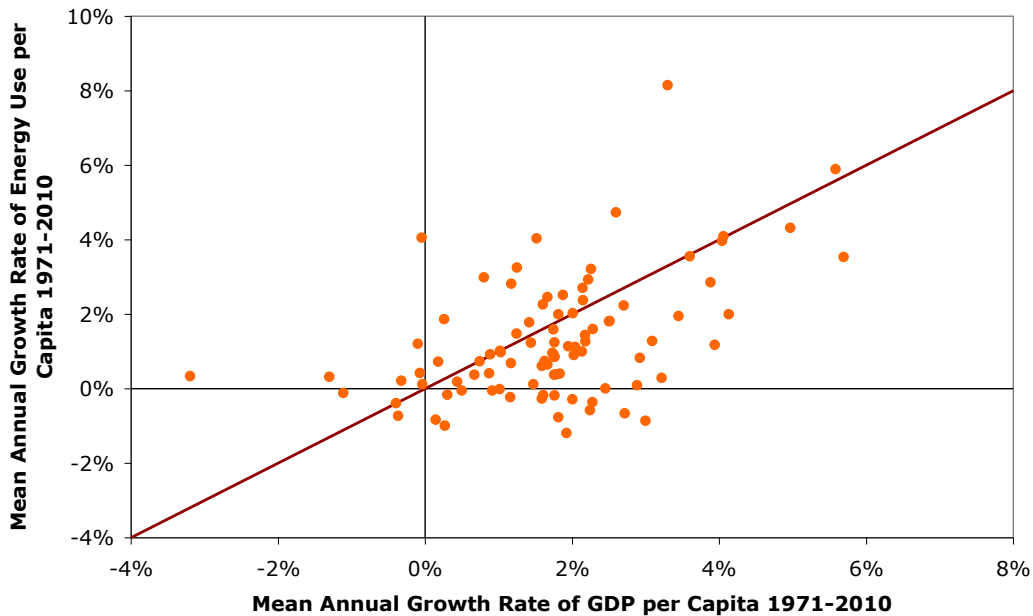


Figure 1: Growth rate of energy use and GDP per capita 1971-2010

Though the growth rate of per capita energy use is correlated with the growth rate of GDP per capita (Figure 1) there is clearly much variation in growth rates that might be explained by decoupling, convergence, or other factors. However, the two main hypotheses explaining this dynamics, the decoupling (e.g. Jakob et al., 2012; Lescaroux, 2010; Medlock and Soligo, 2001) and convergence hypotheses (e.g. Ezcurra, 2007; Le Pen and Sévi, 2010; Liddle, 2010; Mulder and de Groot, 2012) have mostly been tested independently of each other, when, in fact, they may both be involved in driving changes in energy use. Csereklyei et al. (2016) find that over the last forty years there has been a stable cross-sectional relationship between energy use per capita and

income with an elasticity of energy use with respect to income of less than unity. This implies that energy intensity has tended to decrease in countries that have become richer, but not in others. But, in the long run, per capita energy use tends to rise with no sign of decoupling at higher income levels. These results contrast with Jakob et al. (2012) who find decoupling between energy use and growth at higher income levels. Csereklyei et al. (2016) also find that over the last two centuries there has been convergence in energy intensity towards the current distribution of energy intensity and income per capita. This contradicts some (e.g. Le Pen and Sévi, 2010) but not other (e.g. Liddle, 2010) previous convergence studies.

In this paper, we examine the relationship between the average growth rate of energy use per capita and the average growth rate of real GDP per capita over a forty-year period in 93 countries, testing for the effects of convergence, decoupling, and other potential determinants of the growth rate of per capita energy use in a simple single equation framework. Our main contribution is to allow both these hypotheses to be tested in a single, econometrically robust model.

We do this by estimating a model in long-run growth rates rather than in levels of the variables. This allows a simple test of beta convergence, where the growth rate of energy use per capita depends on the initial level of energy intensity, as well as on the growth rate of GDP per capita. We also test for strong decoupling by including an interaction term between the rate of economic growth and the (log) level of income. If the coefficient of this variable is negative then growth will increase energy use less at higher income levels and potentially there could be a turning point beyond which further growth reduces per capita energy use. We also consider weak decoupling, where the growth rate of energy use declines as income increases though this is unrelated to economic growth. We test this hypothesis by including the (log) level of income in the regression. We also include a number of control variables, which may affect the growth of energy use.

Rather than using first differences we use long-run growth rates to estimate our model. This avoids many of the known econometric pitfalls that can affect panel data and cross-country studies. First, energy consumption and GDP have both been found to be non-stationary in numerous studies (Apergis and Payne, 2009; Csereklyei et al., 2016; Stern, 2000). Differencing the data removes unit roots and, therefore, any concerns about spurious regressions or issues involved in modeling non-linear functions of unit root variables (Wagner, 2008). Second, using long-run differences rather than first differences, focuses attention on the long-run behavior of the time series (Chirinko et al., 2011). Third, we only estimate the average size of the time effect across the sample, avoiding the problems of explicitly modeling unobserved time effects (Vollebergh et al., 2009). Fourth, our method also reduces the main problem associated with the between estimator proposed by Stern (2010) – that omitted variables correlated with the levels of the explanatory variables may result in biased estimates. In our approach, the means of these variables are removed by differencing.

Working with a cross-sectional dataset raises the question of spatial dependence – changes to a variable in one country may be correlated with changes in the same or other variables in neighboring countries. Most of the research on energy consumption in the past has been in a time-series or panel setting and, with the exception of Jiang et al. (2014), has not explicitly

addressed the issue of spatial dependence. To deal with the problem of spatial dependence, we apply spatial filtering (Tiefelsdorf and Griffith, 2007), rendering the remaining spatial dependence in the residuals statistically insignificant, and therefore, reducing the potential bias of the estimators. As our models include the growth rate of GDP as a regressor, reverse causality from the growth rate of energy use to that of GDP could result in simultaneity bias and hinder a causal interpretation of our regression results. Bruns et al. (2014) show that the only robust result in the very large literature on causality between energy and economic output is that GDP causes energy use (when energy prices are controlled for). This justifies including the GDP growth rate on the right hand side of our regression model. However we also give an approximation of the magnitude of the possible bias in the parameter estimates. We use income per capita data that is adjusted for purchasing power parity and the IEA primary energy use data that we use includes the use of traditional biomass, which are recommended choices for comparing developed and developing countries (Csereklyei et al., 2016; Ruijven et al., 2009).

Our paper extends the econometric method of Anjum et al. (2014), by considering spatial dependence and by applying the approach to the evolution of energy use rather than pollution emissions. We also extend the investigation in Csereklyei et al. (2016) by integrating the different factors they consider into a single econometric model, which allows us to assess the contribution of each factor to the growth in energy use. Our key findings are that over the period examined, the growth of per capita energy use has been primarily driven by economic growth, weak decoupling, and convergence effects. There is no sign of strong decoupling. We find that resource endowments and climate also significantly affect the growth rate of per capita energy use. These findings need to be taken into account in projections and forecasts of future energy use.

The next section of the paper introduces the data and methodology used, Section 3 discusses the econometric results, while in Section 4 conclusions are presented. Data sources and the choice of spatial weights are discussed in the Appendix.

2. Methods

2.1. Hypotheses and Models

We work with a balanced dataset covering 93 countries between 1971 and 2010. Our basic model is:

$$g(E/P)_i = \alpha + (\beta_1 + \beta_2(Y/P)_i) * g(Y/P)_i + \epsilon_i \quad (1)$$

where $g(E/P)$ indicates the long-run growth rate of per capita energy use computed as described in the Appendix, $g(Y/P)$ is the long-run growth rate of income per capita, and $(Y/P)_i$ is the mean of the natural logarithm of income per capita over our forty-year sample period. The interaction term tests the hypothesis that energy use and economic growth decouple as income increases. When β_2 is negative, energy use first increases and then decreases as income increases above a given turning point, so that decoupling takes place. We call this the "strong decoupling"

hypothesis. As $(Y/P)_i$ is demeaned, that is the sample mean is subtracted from the variable, β_1 is the income elasticity of energy use at the sample mean log income level. As the variables are expressed in growth rates, the intercept term, α is the average time effect in the sample—the growth rate of energy use when the economic growth rate is zero. ϵ is a random error term.

Model 2 also tests the independent effect of the level of income on the growth rate of energy use, or a weak decoupling hypothesis:

$$g(E/P)_i = \alpha + (\beta_1 + \beta_2(Y/P)_i) * g(Y/P)_i + \beta_3(Y/P)_i + \epsilon_i \quad (2)$$

If β_3 is significantly less than zero, then controlling for the rate of economic growth, the growth rate of energy use is lower in richer countries than in poorer countries.

We can test for beta convergence in energy intensity by adding the log of energy intensity at the beginning of the sample period (1971) to Model 2:

$$g(E/P)_i = \alpha + (\beta_1 + \beta_2(Y/P)_i) * g(Y/P)_i + \beta_3(Y/P)_i + \beta_4(E/Y)_{i,0} + \epsilon_i \quad (3)$$

where $(E/Y)_{i,0}$ is the log of energy intensity in country i in the first year of the sample period, in our case 1971. If β_4 is significantly less than zero, then energy intensity grows more slowly in countries that were energy intensive in 1971. and *vice versa* so that there is convergence in energy intensity over time. We also test the effects of a number of additional control variables in Models 4 to 6. The general form of these models is:

$$g(E/P)_i = \alpha + (\beta_1 + \beta_2(Y/P)_i) * g(Y/P)_i + \beta_3(Y/P)_i + \beta_4(E/Y)_{i,0} + \sum_j \eta_j X_{j,i} + \epsilon_i \quad (4)$$

where X is a vector of exogenous control variables. A large number of variables may affect per capita energy use and energy intensity (Csereklyei and Humer, 2012; Stern, 2012). At the same time, numerous potential control variables will be clearly influenced or driven by the economic development and growth process. While these variables may be important in shaping energy use, they will subtract from measuring the full effect of economic growth on energy use. Therefore, we only include explanatory variables that are in our judgment not directly affected by economic growth or the development level of a country. We include the following variables that Anjum et al. (2014) found had significant effects on the growth rate of carbon dioxide emissions: summer and winter average temperatures and fossil fuel endowments (Norman, 2009). We also test water resource endowments (H_2O) to reflect the potential for hydroelectric power (Burke, 2010). We include a dummy for whether a country was centrally planned (CPE) during at least part of the sample period. Stern (2012) found that this variable was statistically significant in explaining the distance of countries from the energy intensity frontier, though Anjum et al. (2014) did not find it had a significant effect on the growth rate of CO_2 emissions. We also test for the effects of democracy as measured by the Polity2 variable (Marshall et al., 2014).² While Stern (2012) finds

²We estimated models including the interaction of polity and resources, to examine whether good governance

the level of democracy is not significant in determining the level of energy efficiency, the question of whether policies in more democratic countries, such as lower subsidies and stronger energy efficiency programs, reduce the rate of growth of energy use, remains open. Comin and Hobijn (2004) note that democracies seem more successful in maintaining the property rights necessary for investment and in deterring pressure groups from preventing the adoption new technologies. Similarly, Fredriksson and Neumayer (2013) find that a country’s level of democracy capital has significant impact on its climate change policies, while Cheon et al. (2013) find that countries with weaker institutions are found to be more prone to providing gasoline subsidies.

All control variables, with the exception of the centrally planned economy dummy are demeaned, so that the intercept term in the regression can be interpreted as the mean time effect at the sample mean for a non-centrally planned economy. The Appendix provides an exact description of the data sources and transformations. The energy use and GDP variables are from the same sources used by Csereklyei et al. (2016). In the current study we drop six countries from the sample due to lack of data relating to freshwater reserves, fossil fuel endowments, democracy data, or a combination of these. As climate and resource endowments as well as the degree of democracy may affect a country’s rate of economic growth, in Models 4-6 the effect of economic growth on the growth rate of energy use must be interpreted as the effect of those components of the growth rate of GDP that are unaffected by these additional explanatory variables.

2.2. Reverse Causality

We estimate the six models using OLS with heteroskedasticity robust standard errors. Obviously, it is possible that there is reverse causation from the growth in energy use to the growth of GDP. Regression models need not be interpreted causally (Chen and Pearl, 2013). In this interpretation the model simply shows how the growth rate of energy use varies with the other variables. There is a very large and inconclusive literature that tests for Granger causality between the levels of energy use and GDP (Bruns et al., 2014). Bruns et al. (2014) find that the most robust result in that literature is that GDP causes energy use, when models control for energy prices. As energy is an input to production, it is still reasonable to think that exogenous changes to energy use have an effect on output, though Granger causality tests may have too low power to detect these effects in the usual small samples used. We provide some intuition on the likely size of the bias induced by reverse causality if we wish to interpret our models structurally. For simplicity, we consider a model that is simpler even than our Model 1:

$$g(E/P)_i = \beta_1 + \beta_2 g(Y/P)_i + \epsilon_{Ei} \tag{5}$$

We can write a model for the effect of energy on GDP thus:

$$g(Y/P)_i = \alpha_1 + \alpha_2 g(E/P)_i + \gamma^T X_i + \epsilon_{Yi} \tag{6}$$

influences the impact of resources on energy consumption. The relevant regression coefficients were statistically insignificant.

where X is a vector of additional explanatory variables. We now give the parameters a structural interpretation. If we estimate (5) using OLS, then the limit in probability of β_2 is given by:

$$plim\beta_2^{OLS} = \beta_2 + \frac{cov(\epsilon_E, g(E/P))}{var(g(Y/P))} \quad (7)$$

and it can be shown that:

$$cov(\epsilon_E, g(E/P)) = \frac{\alpha_2 var(\epsilon_E)}{1 - \alpha_2 \beta_2} \quad (8)$$

In order to determine the potential size of the bias, we assess the likely magnitudes of the parameters and moments in equations (7) and (8). In our data $var(g(Y/P))=0.0002$ and the OLS estimate of $var(\epsilon_E)$ is also 0.0002. Therefore, the ratio of these two moments is unity and the bias is approximately $\frac{\alpha_2}{(1-\alpha_2\beta_2)}$. Next, we assess the likely size of α_2 , which is the elasticity of GDP with respect to energy in equation (6). The cost share of energy is likely to be in the range of 0.05 to 0.1. When other materials are taken into account the cost share is likely to be near the lower end of this range (Frondel and Schmidt, 2002). Using a traditional growth accounting approach for GDP and under the assumptions made by Stern and Kander (2012) the elasticity of GDP with respect to energy is given by

$$\frac{\partial \ln Y}{\partial \ln E} = \left(1 - \frac{\sigma - 1}{\sigma}\right) \frac{\partial \ln G}{\partial \ln E} \quad (9)$$

where σ is the elasticity of substitution between energy and other inputs, and G is gross output. This is probably an upper limit on the size of this elasticity.³ Assuming that the magnitude of β_2 , the income elasticity of energy in equation (5), is near to unity, this means that the OLS estimate of β_2 is likely to be biased upwards in the region of 0.05.⁴

We also considered addressing the endogeneity problem using instrumental variable techniques. Finding a suitable instrument for the long-term growth rate of GDP per capita is a very challenging task. We tried several alternatives including human capital growth, or the initial value of human capital at the beginning of the period, based on the Barro and Lee (2010) database, trade openness, as well as the initial value of GDP per capita, and 5-year growth rates of GDP per capita before the start of the period. Additionally we tested climate variables such as summer or winter temperatures and average rainfall. We also considered the growth rates of each country's export partners as used by Acemoglu et al. (2008) and the modified version suggested by Burke (2012). Other variables relating to long-term growth rates such as the number of hospital beds, measuring health expenditures and thus driving growth (Rivera and Currais, 2003), were not available for the entire period, while information pertaining to intellectual property rights, such as the Ginarte-Park Index was not available for a large number of countries.

³The dynamic effects of growth in energy use might be larger if increased energy availability stimulates capital accumulation (Stern and Kander, 2012) or induces technological change.

⁴The presence of the interaction term in our model considerably complicates this exposition. The structural equation for the interaction term involves interactions of (Y/P) and the regression parameters and error terms as it is simply the structural equation for $g(Y/P)$ multiplied by (Y/P) . The denominators of the reduced form for both $(Y/P)_i$ and the interaction term involve a function of (Y/P) .

Though some of these seemed acceptable as instruments for the growth rate of GDP, the main difficulty is finding an instrument for $(Y/P) * g(Y/P)$. The usual approach is to interact the instrument for $g(Y/P)$ with the actual data for (Y/P) to obtain a second instrument. However, usually this instrument is highly correlated with (Y/P) , which is also included in the regression. As a result, the instrument provides little if any additional explanatory power in the first stage regression and so fails to pass the weak instrument test of Stock and Yogo (2005). For example, using the export partners' growth rate instrument the correlation between the interacted instrument and (Y/P) is 0.975 but the correlation with $(Y/P)*g(Y/P)$, which is actually targeted, is only 0.64. Therefore, in the end we abandoned the quest for appropriate instruments.

2.3. Spatial Filtering

Recent research using panel data emphasizes the importance of taking cross-sectional dependence into account (Wagner, 2008). Similarly, spatial dependence often arises in macroeconomic cross-sectional data (LeSage and Pace, 2009; Tiefelsdorf and Griffith, 2007). In particular, spatially autocorrelated omitted variables might induce spatial autocorrelation in the error term. If these omitted variables are correlated with the explanatory variables included in the model, then the parameter estimates will be biased and inconsistent, unless spatially lagged dependent and explanatory variables are added to the model (LeSage and Pace, 2009), or spatial filtering is used (Crespo Cuaresma and Feldkircher, 2013; Tiefelsdorf and Griffith, 2007). We used Moran's I (Moran, 1950) to test the null of spatial independence in both our dependent and explanatory variables and in the OLS model residuals. We conclude that both the variables and the residuals of the OLS models are spatially dependent.

We assume that the spatial autocorrelation in the residuals is due to omitted variables that are spatially autocorrelated and may be correlated with the included regressors. Though this means that the data could be represented by a spatial lag model (LeSage and Pace, 2009), we do not think that the lagged values of the dependent variable actually cause variation in the dependent variable and so we are not interested in estimating the spatial autocorrelation coefficient. Rather, we think it is more appropriate to remove the spatial autocorrelation using the spatial filtering approach, thus rendering the coefficients of the independent variables unbiased.

Spatial filtering aims to remove the residual spatial autocorrelation by adding additional spatial control variables to the regression equation (Crespo Cuaresma and Feldkircher, 2013). Spatial filtering can also reduce the correlation among the explanatory variables by accounting for their common spatial patterns (Crespo Cuaresma and Feldkircher, 2013). Here, we follow the approach of Griffith (2000) and Tiefelsdorf and Griffith (2007) in using a non-parametric spatial filtering approach, in which a parsimonious subset of the eigenvectors of a transformed spatial contiguity matrix, W , is used to capture dependencies among the disturbances of a spatial regression model.⁵

Tiefelsdorf and Griffith (2007) argue that a linear combination of these eigenvectors should be

⁵As a robustness check we also estimate the model with some alternative weighting matrices. The preferred and alternative spatial matrices are described in the Appendix, which also provides some econometric results for the alternative matrices.

able to capture the spatial patterns of the stochastic component of the model, and to remove this pattern from the residuals. The transformed matrix is given by $M_1 \frac{1}{2}(W + W^T)M_1$, where $M_1 = I - i(i^T i)^{-1}i^T$ and i is the unity vector. We then estimate the regression with a subset of m of the eigenvectors $\vec{e}_1, \dots, \vec{e}_m$ of $M_1 \frac{1}{2}(W + W^T)M_1$:

$$g(E/P)_i = \sum_{i=1}^m \gamma_i \vec{e}_i + \chi^T Z_i + u_{Ei} \quad (10)$$

where Z is a vector of explanatory variables, and the u_{Ei} are the disturbances of the filtered model. The key challenge is to choose a unique and parsimonious set of eigenvectors to be added to the regression equation, thus adequately reducing the level of spatial autocorrelation in the residuals with the least number of eigenvectors. We use the spatial filtering procedure of Tiefelsdorf and Griffith (2007), setting the threshold level of tolerated autocorrelation

$$|z[I(\widehat{u}_E)]| < \delta, \delta = 1 \quad (11)$$

where $|z[I(\widehat{u}_E)]|$ is the standardized Moran's I statistic, for the estimated regression residuals \widehat{u}_E . While Tiefelsdorf and Griffith (2007) set the threshold at 0.1, they note that for smaller panels, with less than 50 spatial observations $\delta = 1$ could be used. We set the tolerance level z at 1, which is the same tolerance level used by stepwise regression procedures that maximize adjusted R squared.⁶

3. Results

3.1. Estimates without Spatial Filtering

The long-run growth rate of income per capita is highly statistically significant in all models. The elasticity mostly increases when additional variables are added (Table 1). In Model 6 the elasticity is 0.80. Figure 1 shows the mean annual growth rate of energy use and the mean annual growth rate of GDP per capita. The intercept term is not statistically significantly different from zero, implying that in the absence of economic growth or other effects there is no particular tendency for per capita energy use to change. Given the results presented by Csereklyei et al. (2016) neither of these findings comes as a surprise. But we also find a significantly negative coefficient for the level of the mean log income per capita for four out of the five models where we include it. This means that as countries get richer, their rate of growth of energy use per capita will decline, *ceteris paribus*. This validates the weak decoupling story.

The interaction term of the income per capita growth rate and the log income per capita

⁶ In the time series context, it is possible to select the appropriate number of lags using either serial correlation tests or goodness of fit statistics such as the Akaike Information Criterion or likelihood ratios. In the spatial context, goodness of fit statistics are not appropriate because in the presence of autocorrelation the parameters will be estimated inconsistently, which is not the case in the time series setting. Therefore, we do not use goodness of fit statistics to select the number of eigenvectors to include in the model and instead use a spatial autocorrelation statistic.

level tests the strong decoupling hypothesis that per capita GDP growth has reduced or even had negative effects on energy use growth at higher income levels. For Models 1-2 this term is statistically insignificant, while for models 3-6 it is positive and significant. We find no evidence in the results in Table 1 in support of the strong decoupling story. Even for Model 1, which does not control for any other variables, the coefficient on the interaction term is positive though not statistically significant. This contrasts strongly with the results of researchers such as Jakob et al. (2012) who claim that the strong decoupling story holds, yet it is in line with Csereklyei et al. (2016) who also do not find evidence of decoupling at higher income levels. These strong differences appear to arise from the different sample of countries examined by Jakob et al. (2012) and Csereklyei et al. (2016). Using the method of Jakob et al. (2012) as shown in equation (2) of their paper and our sample of 93 countries we estimate that the elasticity of energy use growth with respect to GDP growth is 0.37 in non-OECD countries and 0.48 in OECD countries. Jakob et al. (2012) estimate these elasticities as 0.63 and -0.35 (for PPP data) for samples of 30 non-OECD and 21 OECD countries, respectively.

The beta convergence effect coefficient is remarkably consistent in value across models and highly significant. These results confirm the findings of Liddle (2010) and Csereklyei et al. (2016) but contrast with the results of convergence tests using other approaches, such as Le Pen and Sévi (2010).

Democracy values only have a significant coefficient at the 5% level in one out of three of the models that we include them in; the effect is negative. As expected, fossil fuel endowments are found to be highly significant and affect the growth rate of energy use positively. Higher fossil fuel reserves are often coupled with lower prices and higher subsidies. However, water resources are not statistically significant. Summer temperatures have a positive effect on the energy use growth rate, which might be explained by the increased spread of air conditioning over time. Winter temperatures have a negative effect, though this effect is smaller in absolute value than the effect of higher summer temperatures. This means that countries with colder winters saw a faster rate of growth of energy use, *ceteris paribus*. This is somewhat surprising as we would have expected that countries with cold winters, which are generally developed countries, heated homes and workspaces significantly already at the beginning of our sample period and have also improved their energy efficiency over time, so that cold winters would not boost the rate of growth of energy use. It is, of course, possible that these climate variables, which are highly spatially autocorrelated, are picking up the effect of other omitted spatially autocorrelated variables. Centrally planned status has a negative and significant effect. The majority of these countries, most remarkably China, reformed their economies around the middle of our sample period, and reduced their energy intensity very significantly as a result (Ma and Stern, 2008; Stern, 2012).

3.2. Estimates with Spatial Filters

The spatially filtered estimates, shown in Table 2, indicate that the unfiltered estimate of the effect of the growth rate of GDP on the growth of energy use is biased upward in the first four

Table 1: Model Estimates without Spatial Filtering

	Mod 1	Mod 2	Mod 3	Mod 4	Mod 5	Mod 6
(Intercept)	0.0025	0.0018	-0.0018	-0.0014	-0.0019	-0.0012
	0.002	0.003	0.002	0.002	0.002	0.002
g(Y/P)	0.5750 ***	0.6083 ***	0.7937 ***	0.7725 ***	0.7954 ***	0.7958 ***
	0.118	0.145	0.110	0.097	0.086	0.078
(Y/P)*g(Y/P)	0.0691	0.1041	0.1948 ***	0.1967 ***	0.2246 ***	0.2308 ***
	0.067	0.092	0.067	0.061	0.053	0.046
Y/P		-0.0010	-0.0071 ***	-0.0049 ***	-0.0071 ***	-0.0077 ***
		0.002	0.002	0.002	0.002	0.001
E/Y_{1971}			-0.0171 ***	-0.0163 ***	-0.0171 ***	-0.0154 ***
			0.003	0.002	0.002	0.002
Democracy				-0.0006 **	-0.0003	-0.0002
				0.000	0.000	0.000
FFE					0.0010 ***	0.0010 ***
					0.000	0.000
H_2O						-0.0004
						0.001
SumT						0.0008 **
						0.000
WinT						-0.0004 **
						0.000
CPE						-0.0086 **
						0.003
N	93	93	93	93	93	93
adj. R-sq	0.23	0.22	0.62	0.65	0.69	0.74
Moran's I (p)	0.00	0.00	0.00	0.00	0.00	0.02

*Note*₁: OLS estimates with robust standard errors. *p<0.1; **p<0.05; ***p<0.01.

*Note*₂: E/Y_{1971} measures the energy intensity convergence, Democracy denotes the Polity values, FFE the fossil fuel endowment of a country, H_2O accounts for the freshwater reserves, SumT and WinT denote summer and winter temperatures respectively, while CPE stands for centrally planned economies.

*Note*₃: Moran's I (p) gives the p-value of Moran's I statistic for the residual of the estimated models.

of our models. The interaction of the income per capita growth rate and its level now has a statistically significant positive effect in all our models, further supporting our claim that there is no strong decoupling between energy consumption growth and income per capita growth at higher income levels. For the last two models, however, the unfiltered estimates tend to underestimate the elasticity of income per capita growth.

While the freshwater resources variable, or the potential for hydroelectric power, was not significant in the unfiltered results, it is now significant at the 5% level, with a positive sign, meaning that in common with fossil fuel resource endowments, the presence of hydroelectric potential tends to increase the growth rate of per capita energy use over the period. Also, democracy is now significant at the 10% level in Model 6, indicating that countries with more democratic governments experience slower rates of growth in per capita energy use, *ceteris paribus*. All other coefficients retain the same sign and remain significant or not, and are generally corrected downwards in these estimates compared to the unfiltered estimates. The spatial autocorrelation in the residuals has been reduced to an acceptable level, and the null hypothesis of Moran's I test can no longer be rejected.

Table 2: Model Estimates with Spatial Filtering

Model 1	Estimate	Std. Error		Model 2	Estimate	Std. Error	
(Intercept)	0.0034	0.002	**	(Intercept)	0.0021	0.002	
g(Y/P)	0.5047	0.083	***	g(Y/P)	0.5648	0.113	***
(Y/P)*g(Y/P)	0.1525	0.058	**	(Y/P)*g(Y/P)	0.2026	0.067	***
fitted(lagcol)vec6	0.0517	0.012	***	Y/P	-0.0017	0.001	
fitted(lagcol)vec5	0.0408	0.010	***	fitted(lagcol2)vec6	0.0526	0.012	***
fitted(lagcol)vec15	-0.0369	0.014	**	fitted(lagcol2)vec5	0.0418	0.010	***
fitted(lagcol)vec2	0.0421	0.010	***	fitted(lagcol2)vec15	-0.0389	0.014	***
fitted(lagcol)vec11	0.0270	0.012	**	fitted(lagcol2)vec2	0.0395	0.012	***
fitted(lagcol)vec14	0.0277	0.011	**	fitted(lagcol2)vec11	0.0259	0.011	**
Adjusted R2		0.554		fitted(lagcol2)vec14	0.0271	0.010	***
Moran's I stat. (p)		0.253		Adjusted R2		0.556	
				Moran's I stat. (p)		0.208	
Model 3	Estimate	Std. Error		Model 4	Estimate	Std. Error	
(Intercept)	-0.0014	0.001		(Intercept)	-0.0008	0.001	
g(Y/P)	0.7644	0.073	***	g(Y/P)	0.7324	0.075	***
(Y/P)*g(Y/P)	0.2154	0.053	***	(Y/P)*g(Y/P)	0.2178	0.062	***
Y/P	-0.0047	0.001	***	Y/P	-0.0024	0.001	**
E/Y ₁₉₇₁	-0.0157	0.002	***	E/Y ₁₉₇₁	-0.0173	0.002	***
fitted(lagcol3)vec8	0.0279	0.008	***	Democracy	-0.0006	0.000	***
fitted(lagcol3)vec14	0.0297	0.006	***	fitted(lagcol4)vec14	0.0298	0.005	***
fitted(lagcol3)vec23	-0.0279	0.011	**	fitted(lagcol4)vec8	0.0341	0.009	***
fitted(lagcol3)vec15	-0.0210	0.006	***	fitted(lagcol4)vec21	-0.0258	0.005	***
fitted(lagcol3)vec5	0.0194	0.007	***	fitted(lagcol4)vec23	-0.0275	0.011	**
fitted(lagcol3)vec6	0.0241	0.006	***	fitted(lagcol4)vec2	0.0313	0.007	***
fitted(lagcol3)vec11	0.0181	0.007	**	fitted(lagcol4)vec9	-0.0163	0.008	**
fitted(lagcol3)vec2	0.0276	0.007	***	fitted(lagcol4)vec11	0.0165	0.007	**
fitted(lagcol3)vec9	-0.0163	0.007	**	fitted(lagcol4)vec6	0.0180	0.007	**
fitted(lagcol3)vec21	-0.0185	0.005	***	Adjusted R2		0.805	
Adjusted R2		0.805		Moran's I stat. (p)		0.205	
Moran's I stat. (p)		0.243					
Model 5	Estimate	Std. Error		Model 6	Estimate	Std. Error	
(Intercept)	-0.0024	0.001	.	(Intercept)	-0.0028	0.001	*
g(Y/P)	0.8264	0.058	***	g(Y/P)	0.8635	0.061	***
(Y/P)*g(Y/P)	0.1996	0.049	***	(Y/P)*g(Y/P)	0.2274	0.044	***
Y/P	-0.0065	0.002	***	Y/P	-0.0070	0.002	***
E/Y ₁₉₇₁	-0.0191	0.002	***	E/Y ₁₉₇₁	-0.0191	0.002	***
FFE	0.0012	0.000	***	FFE	0.0007	0.000	***
Democracy	-0.0003	0.000		Democracy	-0.0003	0.000	*
fitted(lagcol5)vec14	0.0323	0.006	***	H ₂ O	0.0010	0.000	**
fitted(lagcol5)vec29	0.0277	0.009	***	SumT	0.0008	0.000	***
fitted(lagcol5)vec21	-0.0251	0.005	***	WinT	-0.0004	0.000	***
fitted(lagcol5)vec11	0.0163	0.007	**	CPE	-0.0033	0.003	
fitted(lagcol5)vec24	-0.0164	0.005	***	fitted(lagcol6)vec14	0.0293	0.006	***
fitted(lagcol5)vec8	0.0186	0.013		fitted(lagcol6)vec21	-0.0283	0.006	***
Adjusted R2		0.812		fitted(lagcol6)vec23	-0.0272	0.008	***
Moran's I stat. (p)		0.162		fitted(lagcol6)vec29	0.0214	0.009	**
				fitted(lagcol6)vec9	0.0226	0.010	**
				fitted(lagcol6)vec8	-0.0144	0.006	*
				Adjusted R2		0.832	
				Moran's I stat. (p)		0.285	

Note₁: OLS estimates with eigenvector filtering, & robust standard errors. *p<0.1; **p<0.05; ***p<0.01.

Note₂: E/Y₁₉₇₁ measures the energy intensity convergence, Democracy denotes the Polity values, FFE the fossil fuel endowment of a country, H₂O accounts for the freshwater reserves, SumT and WinT denote summer and winter temperatures respectively, while CPE stands for centrally planned economies.

Note₃: Moran's I (p) gives the p-value of Moran's I statistic for the residual of the estimated models.

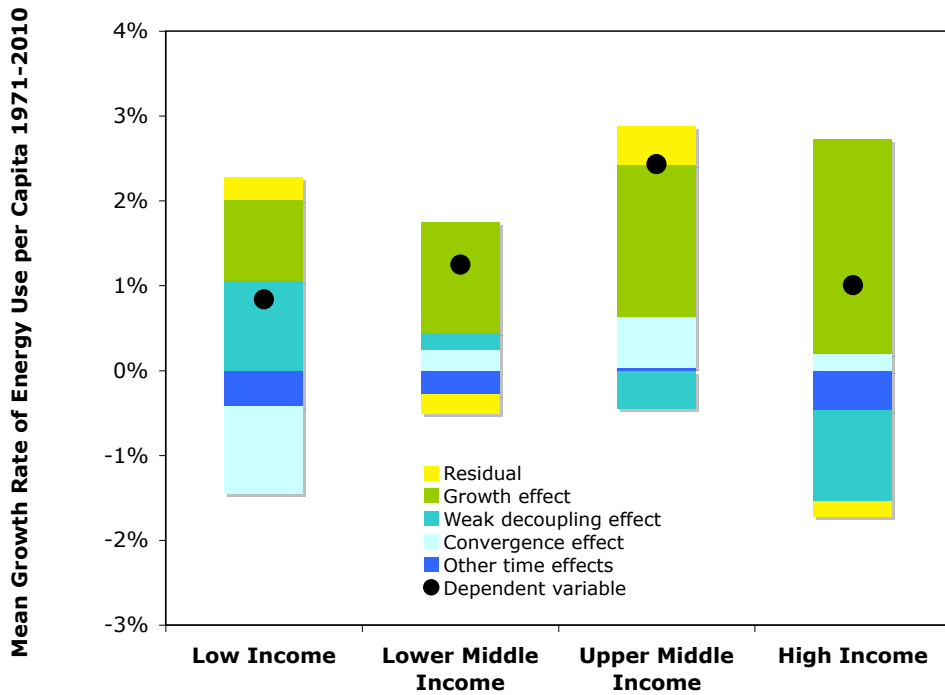


Figure 2: Decomposition of Energy Use per Capita Growth into Growth and Other Effects

Figure 2 illustrates these results using the parameter estimates for Model 6 in Table 2 and our dataset. We divided our 93 countries into four income groups using the World Bank’s classification in 1990 (The World Bank, 1992), which is near the midpoint of our sample. We then computed the mean of each of the variables apart from the eigenvector variables in Model 6 and multiplied these values by the relevant regression coefficients. Summing the effect of the growth rate of per capita GDP and its interaction term, we arrive at the “growth effect” in Figure 2. The weak decoupling effect is the effect of the level of GDP per capita, the convergence effect is the effect of the initial energy intensity, and the time effect aggregates the remaining effects. The unexplained remainder is the “residual”. We see that the growth effect was in fact largest in the high income countries, while the time effect and weak decoupling effect are negative and counteract the effect of growth. The convergence effect is positive but quite small - most developed countries had relatively low energy intensity in 1971 but a few, including the United States and Canada, were very energy intensive. The picture for upper-middle income countries is very different, as the weak decoupling effect is much smaller, the time effect is positive but very small, and the convergence effect is larger. As a result, average energy use per capita growth rates rise with income up to the upper

middle income level, but were lower in the high income countries. In low income countries the weak decoupling effect results in positive growth of energy use rather than decoupling and the time and convergence effects are both quite large and negative. Many of the low income countries, which include China here, were very energy intensive in 1971.

3.3. Robustness Checks

As robustness checks, we estimate our models using different time periods and also a different spatial matrix. We estimate Model 6 for four different sample periods. As it is of particular interest whether the relationship changes as we approach the present, we chose the periods 1971–2010, 1981–2010, 1991–2010 and 2001–2010. The results can be found in Table 3. Interestingly, the coefficient on the growth rate of GDP per capita declines while the coefficient on the interaction term increases. This implies that the elasticity at the sample mean has declined over time as the average income at the sample mean has increased. On the other hand, the difference in elasticities for two countries at given different income levels has increased over time. But none of these changes alters our main results.

The alternative spatial weighting matrix uses a distance-based rule to select neighbors rather than a contiguity rule. The matrix design is described in the Appendix. We present the results using the alternative spatial weighting matrix in Table 4. The differences in coefficient estimates between those using the original contiguity matrix and this "minimum distance neighbour" matrix are minimal, although the impact of water resources and democracy are not found to be significant in these alternative estimates. Based on these results, the coefficient estimates of our original filtered results are robust.

4. Conclusions

In this paper, we examined the determinants of the growth rate of per capita energy use between 1971–2010 for a representative panel of developing and developed countries using a long-run growth rates model, taking into account spatial dependence. We significantly extend the econometric method of Anjum et al. (2014), by correcting for spatial dependence and by applying the approach to the evolution of energy use rather than pollution emissions. We also extend the investigation in Csereklyei et al. (2016) by integrating the different factors they consider into a single model. The novelty of our approach is thus that it allows us to test the relative importance of growth, decoupling, convergence, and other time related effects in a single equation framework, that is robust to the presence of spatial autocorrelation.

Our key findings are that, over the examined period, the most robust driver of the growth of energy use has been economic growth. There is no sign of decoupling of economic growth and the growth of energy use at higher income levels. In fact the elasticity of energy with respect to economic growth is greater in richer countries. However, there is a negative time effect in high income countries and a positive time effect in low income countries. So we can accept the "weak

Table 3: Model Estimates with Spatial Filtering for Different Time Periods

Model 6: 1971-2010	Estimate	Std. Error		Model 6: 1981-2010	Estimate	Std. Error	
(Intercept)	-0.0028	0.001	*	(Intercept)	-0.0029	0.001	**
g(Y/P)	0.8635	0.061	***	g(Y/P)	0.8486	0.063	***
(Y/P)*g(Y/P)	0.2274	0.044	***	(Y/P)*g(Y/P)	0.2590	0.045	***
Y/P	-0.0070	0.002	***	Y/P	-0.0079	0.001	***
E/Y_{1971}	-0.0191	0.002	***	E/Y_{1971}	-0.0175	0.003	***
FFE	0.0007	0.000	***	FFE	0.0009	0.000	***
Democracy	-0.0003	0.000	*	Democracy	-0.0006	0.000	**
H_2O	0.0010	0.000	**	H_2O	0.0013	0.001	**
SumT	0.0008	0.000	***	SumT	0.0007	0.000	***
WinT	-0.0004	0.000	***	WinT	-0.0003	0.000	**
CPE	-0.0033	0.003		CPE	-0.0088	0.004	**
fitted(lagcol6)vec14	0.0293	0.006	***	fitted(lagcol6)vec7	-0.0197	0.011	*
fitted(lagcol6)vec21	-0.0283	0.006	***	fitted(lagcol6)vec14	0.0310	0.008	***
fitted(lagcol6)vec23	-0.0272	0.008	***	fitted(lagcol6)vec23	-0.0281	0.013	**
fitted(lagcol6)vec29	0.0214	0.009	**	fitted(lagcol6)vec21	-0.0268	0.007	***
fitted(lagcol6)vec9	0.0226	0.010	**	fitted(lagcol6)vec29	0.0288	0.011	**
fitted(lagcol6)vec8	-0.0144	0.006	*				
Adjusted R2		0.832		Adjusted R2		0.739	
Moran's I stat. (p)		0.285		Moran's I stat. (p)		0.167	

Model 6: 1991-2010	Estimate	Std. Error		Model 6: 2001-2010	Estimate	Std. Error	
(Intercept)	-0.0040	0.002	**	(Intercept)	-0.0013	0.002	
g(Y/P)	0.7835	0.096	***	g(Y/P)	0.6676	0.083	***
(Y/P)*g(Y/P)	0.2560	0.056	***	(Y/P)*g(Y/P)	0.3028	0.068	***
Y/P	-0.0070	0.002	***	Y/P	-0.0107	0.004	***
E/Y_{1971}	-0.0099	0.002	***	E/Y_{1971}	-0.0117	0.004	***
FFE	0.0004	0.000		FFE	0.0013	0.000	***
Democracy	-0.0005	0.000		Democracy	-0.0008	0.001	
H_2O	0.0011	0.001		H_2O	0.0017	0.001	*
SumT	0.0007	0.000	**	SumT	0.0005	0.000	
WinT	-0.0001	0.000		WinT	-0.0002	0.000	
CPE	-0.0121	0.006	**	CPE	-0.0124	0.008	
fitted(lagcol6)vec7	-0.0304	0.011	**	fitted(lagcol6)	-0.0624	0.017	***
fitted(lagcol6)vec35	0.0378	0.011	***				
fitted(lagcol6)vec11	0.0254	0.003	***				
Adjusted R2		0.612		Adjusted R2		0.512	
Moran's I stat. (p)		0.185		Moran's I stat. (p)		0.181	

*Note*₁: OLS estimates with eigenvector filtering 4 periods, & robust standard errors. *p<0.1; **p<0.05; ***p<0.01.

*Note*₂: E/Y_{1971} measures the energy intensity convergence, Democracy denotes the Polity values, FFE the fossil fuel endowment of a country, H_2O accounts for the freshwater reserves, SumT and WinT denote summer and winter temperatures respectively, while CPE stands for centrally planned economies.

*Note*₃: Moran's I (p) gives the p-value of Moran's I statistic for the residual of the estimated models.

Table 4: Model Estimates with Spatial Filtering with 1.4 Minimum Distance Weights

Model 1	Estimate	Std. Error		Model 2	Estimate	Std. Error	
(Intercept)	0.003	0.002	*	(Intercept)	0.002	0.002	
g(Y/P)	0.548	0.090	***	g(Y/P)	0.622	0.105	***
(Y/P)*g(Y/P)	0.032	0.057		(Y/P)*g(Y/P)	0.106	0.073	
fitted(lagcol)vec20	0.047	0.022	**	dm_ypmean	-0.002	0.001	*
fitted(lagcol)vec12	0.038	0.009	***	fitted(lagcol2)vec20	0.050	0.021	**
fitted(lagcol)vec2	-0.045	0.009	***	fitted(lagcol2)vec12	0.039	0.009	***
Adjusted R2		0.446		fitted(lagcol2)vec2	-0.045	0.009	***
Moran's I stat. (p value)		0.339		Adjusted R2		0.453	
				Moran's I stat. (p value)		0.260	
Model 3	Estimate	Std. Error		Model 4	Estimate	Std. Error	
(Intercept)	-0.001	0.001		(Intercept)	-0.002	0.001	
g(Y/P)	0.766	0.067	***	g(Y/P)	0.801	0.063	***
(Y/P)*g(Y/P)	0.152	0.039	***	(Y/P)*g(Y/P)	0.191	0.046	***
Y/P	-0.008	0.001	***	Y/P	-0.008	0.001	***
E/Y_{1971}	-0.016	0.002	***	E/Y_{1971}	-0.017	0.002	***
fitted(lagcol3)vec20	0.043	0.009	***	dm_avpolity	0.000	0.000	
fitted(lagcol3)vec17	-0.041	0.010	***	fitted(lagcol4)vec17	-0.041	0.010	***
fitted(lagcol3)vec2	-0.023	0.007	***	fitted(lagcol4)vec3	-0.019	0.008	**
fitted(lagcol3)vec3	-0.018	0.008	**	fitted(lagcol4)vec20	0.037	0.011	***
fitted(lagcol3)vec12	0.013	0.006	**	fitted(lagcol4)vec2	-0.021	0.007	***
fitted(lagcol3)vec35	-0.020	0.008	**	fitted(lagcol4)vec12	0.011	0.006	*
Adjusted R2		0.792		fitted(lagcol4)vec19	0.020	0.007	***
Moran's I stat. (p value)		0.231		Adjusted R2		0.792	
				Moran's I stat. (p value)		0.205	
Model 5	Estimate	Std. Error		Model 6	Estimate	Std. Error	
(Intercept)	-0.003	0.001	*	(Intercept)	-0.002	0.001	
g(Y/P)	0.837	0.064	***	g(Y/P)	0.852	0.057	***
(Y/P)*g(Y/P)	0.218	0.038	***	(Y/P)*g(Y/P)	0.254	0.050	***
Y/P	-0.009	0.001	***	Y/P	-0.007	0.001	***
E/Y_{1971}	-0.016	0.002	***	E/Y_{1971}	-0.015	0.002	***
FFE	0.001	0.000	***	FFE	0.001	0.000	***
Democracy	0.000	0.000		Democracy	0.000	0.000	
fitted(lagcol5)vec17	-0.027	0.009	***	H_2O	0.000	0.001	
fitted(lagcol5)vec20	0.037	0.011	***	SumT	0.001	0.000	***
fitted(lagcol5)vec2	-0.024	0.008	***	WinT	-0.001	0.000	***
fitted(lagcol5)vec12	0.016	0.006	***	CPE	-0.012	0.003	***
fitted(lagcol5)vec8	-0.018	0.002	***	fitted(lagcol6)vec21	0.028	0.009	***
fitted(lagcol5)vec21	0.016	0.006	**	fitted(lagcol6)vec27	0.020	0.008	**
Adjusted R2		0.799		fitted(lagcol6)vec23	0.018	0.007	**
Moran's I stat. (p value)		0.235		Adjusted R2		0.784	
				Moran's I stat. (p value)		0.192	

Note₁: OLS estimates with eigenvector filtering, & robust standard errors. *p<0.1; **p<0.05; ***p<0.01.

Note₂: E/Y_{1971} measures energy intensity convergence, Democracy denotes the Polity values, FFE the fossil fuel endowment of a country, H_2O accounts for the freshwater reserves, SumT and WinT denote summer and winter temperatures respectively, while CPE stands for centrally planned economies.

Note₃: Moran's I (p) gives the p-value of Moran's I statistic for the residual of the estimated models.

decoupling hypothesis” but not the ”strong decoupling hypothesis”. The beta convergence effect is also very robust across our different model specifications. It has a net negative effect on the growth of energy use in low income countries because some of these, such as China, were the most energy intensive countries in our sample in 1971. Convergence actually has a small positive effect on average in high income countries. But this varies a lot across countries. In Canada, Luxembourg, and the United States, which were all very energy intensive in 1971, convergence contributed -0.9%, -1.6%, and -1.0% p.a. to the growth rate of energy use according to our model. We find that resource endowments, climate, and centrally planned status were all also significant in shaping the dynamics of per capita energy use. Our results finding no apparent sign of decoupling between economic growth and the growth of per capita energy use are in contrast with the results of, among others, Jakob et al. (e.g. 2012), Lescaroux (2010), and Medlock and Soligo (2001), but in line with the findings of Csereklyei et al. (2016). The empirical evidence on the significance of beta convergence of energy intensity supports the findings of Liddle (2010), and Mulder and de Groot (2012), but not of Le Pen and Sévi (2010).

Projections and forecasts of future energy use should not, therefore, assume that economic growth will be associated with decreased energy use in the future. Instead, the scale effect seems to be alive and well. On the other hand, there appear to be improvements in energy efficiency across high income countries irrespective of their growth rates or their initial level of energy intensity. These would tend to moderate the growth in energy use as countries get richer at the upper end of the income continuum. At the lower end of the income continuum the same effects serve to raise energy intensity. But, some of the major reductions in energy intensity in countries, such as the United States and China, have probably been the result of convergence towards the global mean, and so are unlikely to be reproduced in the future.

Acknowledgements

The authors thank Stefan Humer wholeheartedly for his help, support and suggestions, and Stephan Bruns, Paul Burke, Corrado Di Maria, Jesus Crespo Cuaresma, Biliana Yontcheva, and two anonymous referees for useful suggestions.

References

- Acemoglu, D., S. Johnson, J.A. Robinson, and P. Yared (2008). “Income and democracy”. In: *American Economic Review* 98.3, pp. 808–842.
- Anjum, Z., P.J. Burke, R. Gerlagh, and D.I. Stern (2014). “Modeling the emissions-income relationship using long-run growth rates”. In: CCEP Working Paper 1403.
- Apergis, N. and J.E. Payne (2009). “Energy consumption and economic growth in Central America: Evidence from a panel cointegration and error correction model”. In: *Energy Economics* 31, pp. 211–216.

- Barro, R. and J.W. Lee (2010). “A new data set of educational attainment in the World, 1950-2010”. In: *Journal of Development Economics* 104, pp. 184–198.
- BP (2015). *Statistical Review of World Energy 2015*. British Petrol.
- Bruns, S., C. Gross, and D.I. Stern (2014). “Is there really Granger causality between energy use and output?” In: *The Energy Journal* 35.4, pp. 101–134.
- Burke, P.J. (2010). “Income, resources and the electricity mix”. In: *Energy Economics* 32, pp. 616–626.
- (2012). “Economic growth and political survival”. In: *B.E. Journal of Macroeconomics* 12.1, Article 5.
- Chen, B. and J. Pearl (2013). “Regression and causation: a critical examination of econometrics textbooks”. In: *Real-World Economics Review* 65, pp. 2–20.
- Cheon, A., J. Urpelainen, and M. Lackner (2013). “Why do governments subsidize gasoline consumption? An empirical analysis of global gasoline prices, 2002-2009”. In: *Energy Policy* 56, pp. 382–390.
- Chirinko, R.S., S.M. Fazzari, and A.P. Meyer (2011). “A new approach to estimating production function parameters: The elusive capital-labor substitution elasticity”. In: *Journal of Business & Economic Statistics* 29.4, pp. 587–594.
- Comin, D. and B. Hobijn (2004). “Cross-country technology adoption: making the theories face the facts”. In: *Journal of Monetary Economics* 51, pp. 39–83.
- Crespo Cuaresma, J. and M. Feldkircher (2013). “Spatial filtering, model uncertainty and the speed of income convergence in Europe”. In: *Journal of Applied Econometrics* 28, pp. 720–741.
- Csereklyei, Z. and S. Humer (2012). “Modelling primary energy consumption under model uncertainty”. In: *WU Working Papers in Economics* 147.
- Csereklyei, Z., M.d.M. Rubio-Varas, and D.I. Stern (2016). “Energy and economic growth: the stylized facts”. In: *Energy Journal* 37.2, pp. 223–255.
- Ezcurra, R. (2007). “Distribution dynamics of energy intensities: A cross-country analysis”. In: *Energy Policy* 35, pp. 5254–5259.
- Fredriksson, P.G. and E. Neumayer (2013). “Democracy and climate change policies: Is history important?” In: *Ecological Economics* 95, pp. 11–19.
- Frondel, M. and C.M. Schmidt (2002). “The capital-energy controversy: An artifact of cost shares?” In: *The Energy Journal* 23.3, pp. 53–79.
- Griffith, D.A. (2000). “EA linear regression solution to the spatial autocorrelation problem”. In: *Journal of Geographical Systems* 2, pp. 141–156.
- Heston, A., R. Summers, and B. Aten (2012). “Penn World Table Version 7.1”. In: *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania*.
- International Energy Agency (2014). *World Energy Outlook, ISBN: 978-92-64-20805-6*. Tech. rep.
- Jakob, M., M. Haller, and R. Marschinski (2012). “Will history repeat itself? Economic convergence and convergence in energy use patterns”. In: *Energy Economics* 34, pp. 95–104.
- Jiang, L., H. Folmer, and M. Ji (2014). “The drivers of energy intensity in China A spatial panel data approach”. In: *China Economic Review* 31, pp. 351–360.

- Le Pen, Y. and B. Sévi (2010). “On the non-convergence of energy intensities: Evidence from a pair-wise econometric approach”. In: *Ecological Economics* 69, pp. 641–650.
- LeSage, J. and R.K. Pace (2009). *Introduction to Spatial Econometrics*. CRC Press.
- Lescaroux, F. (2010). “Dynamics of final sectoral energy demand and aggregate energy intensity”. In: *Energy Policy* 39.1, pp. 66–82.
- Liddle, B. (2010). “Revisiting world energy intensity convergence for regional differences”. In: *Applied Energy* 87, pp. 3218–3225.
- Ma, C. and D.I. Stern (2008). “China’s changing energy intensity trend: A decomposition analysis”. In: *Energy Economics* 30.3, pp. 1037–1053.
- Marshall, M.G., T.R. Gurr, and K. Jagers (2014). “Polity IV Project, Political Regime Characteristics and Transitions, 1800-2013: Dataset Users’ Manual”. In: *Center for Systemic Peace*.
- Medlock, K B. and R. Soligo (2001). “Economic development and end-use energy demand”. In: *Energy Journal* 22.2, pp. 77–105.
- Mitchell, T.D., M. Hulme, and M. New (2002). “Climate data for political areas”. In: *Area* 34, pp. 109–112.
- Moran, P.A.P. (1950). “Notes on continuous stochastic phenomena”. In: *Biometrika* 37, pp. 17–23.
- Mulder, P. and H.L.F. de Groot (2012). “Structural change and convergence of energy intensity across OECD countries, 1970-2005”. In: *Energy Economics* 34, pp. 1910–1921.
- Norman, C.S. (2009). “Rule of law and the resource curse: Abundance versus intensity”. In: *Environmental and Resource Economics* 43, pp. 183–207.
- Rivera, B. and L. Currais (2003). “The effect of health investment on growth: A causality analysis”. In: *International Advances in Economic Research* 9.4, pp. 312–323.
- Ruijven, B. van, F. Urban, R.M.J. Benders, H.C. Moll, JP. van der Sluijs, B. de Vries, and D.P. van Vuuren (2009). “Modeling energy and development: An evaluation of models and concepts”. In: *World Development* 36.12, pp. 2801–2821.
- Stern, D.I. (2000). “A multivariate cointegration analysis of the role of energy in the US macroeconomy”. In: *Energy Economics* 22, pp. 267–283.
- (2010). “Between estimates of the emissions-income elasticity”. In: *Ecological Economics* 69, pp. 2173–2182.
- (2012). “Modeling international trends in energy efficiency”. In: *Energy Economics* 34, pp. 2200–2208.
- Stern, D.I. and A. Kander (2012). “The Role of Energy in the Industrial Revolution and Modern Economic Growth”. In: *The Energy Journal* 33.3, pp. 125–152.
- Stock, J.H. and M. Yogo (2005). “Testing for weak instruments in linear IV regression”. In: *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Ed. by D.W.K Andrews. New York: Cambridge University Press.
- The World Bank (1992). *World Development Report 1992*. Oxford University Press.
- Tiefelsdorf, M. and D.A. Griffith (2007). “Semiparametric filtering of spatial autocorrelation: the eigenvector approach”. In: *Environment and Planning A* 37, pp. 1193–1221.

- Vollebergh, H.R.J., B. Melenberg, and E. Dijkgraaf (2009). "Identifying reduced-form relations with panel data: The case of pollution and income". In: *Journal of Environmental Economics and Management* 58.1, pp. 27–42.
- Wagner, M. (2008). "The carbon Kuznets curve: A cloudy picture emitted by bad econometrics". In: *Resource and Energy Economics* 30, pp. 388–408.

A. Appendix

A.1. Data Sources

Our analysis is based on a balanced panel dataset for 93 countries covering the period 1971 to 2010. The countries included in our sample are Albania, Algeria, Angola, Argentina, Australia, Austria, Bahrain, Bangladesh, Belgium, Benin, Bolivia, Brazil, Bulgaria, Cameroon, Canada, Chile, China, Colombia, Congo, Costa Rica, Cote d'Ivoire, Cuba, Cyprus, Democratic Republic of Congo, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Finland, France, Gabon, Ghana, Greece, Guatemala, Haiti, Honduras, Hungary, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Korea, Lebanon, Luxembourg, Malaysia, Mexico, Morocco, Mozambique, Nepal, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Senegal, Singapore, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, United Kingdom, United States, Uruguay, Venezuela, Vietnam, Zambia, and Zimbabwe.

Population and real income per capita adjusted for purchasing power parity between 1971 and 2010 were sourced from the Penn World Table, Version 7.1. (Heston et al., 2012). To ensure working with a balanced panel, we excluded a number of Eastern-European and Middle-Eastern countries, where data was not available for the entire period. However these geopolitical regions are still represented by a few countries, such as Albania, Bulgaria, Hungary, Poland and Romania in Eastern Europe; and by Algeria, Bahrain, Iran, Iraq and Oman in the Middle-Eastern and North-African region. Primary energy consumption data originate from the International Energy Agency database and is measured in TJ, and includes coal, oil, natural gas, primary electricity, and biomass.

We calculate the long-run growth rates as interval differences, using the natural logarithm of the level variable in 2010 less the natural logarithm of the level variable in 1971 divided by T-1, in our case 39. The long-run growth rate of per capita energy use was gained thus by:

$$g(E/P) = \frac{\ln E/P_{2010} - \ln E/P_{1971}}{39} \quad (12)$$

We applied the same method to the growth rate of real income per capita "g(Y/P)".

Freshwater resources (H_2O) per capita are sourced from the World Development Indicators. Data is available at five-year intervals over the period from 1972 to 2007. We average the available data over time, add one and log it for the regressions. Average monthly temperatures from 1960 to 1990 were available from Mitchell et al. (2002). Summer temperatures are gained by averaging the values for the three summer months from June to August in the Northern Hemisphere and December to February in the Southern Hemisphere. Resource endowments (FFE) are calculated as in Anjum et al. (2014) by multiplying Norman's (2009) ratio of the value of fossil fuel reserves to GDP in 1971 by GDP per capita at market exchange rates in 1971 (World Bank) to gain the value of per capita fossil fuel endowments in 1971. We use the Polity2 variable from the Polity

IV database (Marshall et al., 2014) as an indicator of the level of democracy. This variable scores regimes from 0 to 10 on a democracy scale and 0 to 10 on an autocracy scale and then subtracts the autocracy score from the democracy score.

We demeaned most of the levels variables in the regressions, including the natural log of energy intensity in 1971, summer and winter temperatures, resource endowments, average democracy values, and the mean of the log of real GDP per capita. The growth rates of the variables are not demeaned, including when they are interacted with the demeaned level variables. To construct the instruments discussed in Section 2, we use the data assembled by Paul Burke (Burke, 2012) data on the shares of a country’s exports going to each export partner.

A.2. Spatial Weighting Matrices

We constructed the contiguity matrix used in the original spatial analysis as follows: First we build a contiguity matrix with zero entries for pairs of countries that are not neighbors, and ones for pairs of countries that are neighbors. In our definition all countries that share land borders are neighbors. These land borders also include borders on lakes such as between DR Congo and Tanzania. For all countries without a land border with any country in our dataset we judged which country is its nearest neighbor. As an example, Australia’s nearest neighbor is Indonesia. New Zealand’s nearest neighbor is Australia. As a result, Australia has two neighboring countries. New Zealand has only one. Indonesia also shares a land border with Malaysia. As Singapore is nearer to Malaysia than to Indonesia and is connected by road to Malaysia, Singapore is not deemed a neighbor of Indonesia. The United Kingdom and Ireland share a land border and so each only has one neighbor. The diagonal entries are all zero and this matrix is symmetric. We then row-normalized the matrix so that each row sums to unity.

The choice of the spatial weighting matrix might influence the emergence and identification of common spatial patterns. Therefore, as a robustness check, we also created an alternative weighting matrix. For this matrix, we use a distance-based rule to select neighbors. All countries whose centroid lies within a given distance of a country are considered to be neighbors of the latter country. However, if all countries must have at least one neighbor, and we use a uniform distance to select neighbors then the cutoff distance has to be over 4000km in order to allow New Zealand to have a neighbor. But this would mean that, for example, that Switzerland is a neighbor to all countries as far away as Jordan. Therefore, our rule allows the distance to vary by country size.

We first computed the distances between all countries using the Spherical Law of Cosines. The distance between two countries i and j is given by:

$$d_{ij} = acos(sin(\phi_i)sin(\phi_j) + cos(\phi_i)cos(\lambda_j - \lambda_i))r \quad (13)$$

where ϕ is latitude and λ is longitude in radians and r is the radius of the Earth. Latitudes and longitudes are the country centroids taken from the CIA World Factbook (<https://www.cia.gov/library/publications/the-world-factbook/fields/2011.html>). Next we found each country i ’s nearest neighbor in terms of the distance between the two country centroids. The

distance between these two countries is δ_i . All country centroids that fall within $1.4\delta_i$ of country i 's centroid are then considered neighbors. We chose the factor 1.4 so that the number of countries neighboring each other was similar to that in the matrix in the main analysis. There are still many differences between the two matrices, including 21 countries with fewer, and 45 countries with more spatial links. For example, countries in Latin American countries such as Brazil, Peru, Bolivia, Argentina, Chile, or Colombia have considerably fewer neighbors while many countries in Africa and the Middle East have moderately more neighbors. In the extreme, Brazil and Peru have 5 and 4 fewer neighbors, respectively, and Egypt and Lebanon have 3 and 5 more neighbors, respectively. The contiguity matrix has on average 2.4 neighbors per country. The alternative "minimum distance neighbor matrix" with a 1.4 distance factor has 2.7 neighbors per country on average. The mean absolute difference in number of neighbors for each country is 1.1.