Environmental and Resource Economics
https://doi.org/10.1007/s10640-018-0298-9

Testing for Convergence in Carbon Dioxide Emissions Using a Bayesian Robust Structural Model

Octavio Fernández-Amador[1] · Doris A. Oberdabernig[1] · Patrick Tomberger[1]

Accepted: 26 October 2018

Abstract

We address international convergence in carbon dioxide emissions per capita and per value added derived from emission inventories based on production and consumption patterns. We propose a Bayesian structural model that accounts for heteroscedasticity and endogeneity between emissions and economic growth, and tests for the existence of group-specific convergence via shrinkage priors. We find evidence for country-specific conditional convergence in all emission inventories, implying a half-life of 2.7–3.1 years for production-based emissions and 3.6–4.7 years for consumption-based emissions. When testing for global convergence without allowing for individual-specific convergence paths, the half-life of CO_2 per capita increases to 15–26 years, whereas emission intensities show a half-life of 44–45 years. Our results highlight the current incompatibility between emission targets and economic growth and the need for faster diffusion of green technologies. Moreover, there is no evidence for specific convergence dynamics in the European Union, the OECD, or the countries that are subject to binding emission constraints specified in the Kyoto Protocol. The institutional frameworks implemented in industrialized countries did not induce faster convergence among developed economies.

Keywords CO_2 emissions · Production-based inventories · Carbon footprint · Convergence test · Half-life

JEL Classification F18 · F64 · O44 · Q54 · Q56

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10640-018-0298-9) contains supplementary material, which is available to authorized users.

✉ Octavio Fernández-Amador
octavio.fernandez@wti.org

Doris A. Oberdabernig
doris.oberdabernig@wti.org

Patrick Tomberger
patrick.tomberger@wti.org

1 World Trade Institute, University of Bern, Hallerstrasse 6, 3012 Bern, Switzerland

Published online: 30 November 2018

1 Introduction

Global warming and its consequences are at the center of current policy debate on the sustainability of economic development. The Paris Agreement stipulates holding the global average temperature below 2 °C above pre-industrial levels to bring climate change under control;[1] for this to happen, the 194 countries that signed the agreement are compelled to reach the global peak of greenhouse gas (GHG) emissions as soon as possible (Paris Agreement, Art. 2 and 4). The underlying question is how to make economic growth compatible with limited or decreased pollution, particularly GHG emissions.

The Environmental Kuznets Curve (EKC) predicts that pollution increases with rising income per capita and falls after a peak in emissions has been reached. However, the existence of a turning point in GHG emissions after which emissions start to decrease with economic growth has not been unanimously confirmed by empirical research. Especially for carbon dioxide (CO_2) emissions the existence of such peak has often been rejected.[2] Against this background, it is important to know whether global carbon emissions will eventually reach a limit; only then the growth rate of atmospheric concentrations of CO_2 will stabilize.

The patterns of convergence of CO_2 emissions per capita towards a certain emission level and the height of this level have important implications for the design of the international regulatory framework. Reliable information on whether the steady state of emissions is global or country-specific and on how long it will take for countries to reach this steady state can strengthen the ongoing policy debate. Related to this, the convergence dynamics of carbon emissions derived from both national production and consumption activities should be better understood when revising environmental responsibility, as they characterize the path of emissions associated with further economic development in a globalized context. Increasingly fragmented value chains allow the geographical location of production stages to differ from the place of final consumption. A mere focus on territorial-based emissions neglects the importance of trade in intermediates and carbon leakage, i.e. the shift of highly pollutant industries from countries with stringent environmental regulation to countries with less strict regulation (e.g. Aichele and Felbermayr 2015; Babiker 2005; Fernández-Amador et al. 2016).

In addition, it is relevant to understand to what extent the dynamics of international convergence of emissions per capita is driven by convergence in carbon efficiency worldwide. The adoption of more environmentally friendly technologies will lower carbon intensities, which is particularly relevant for developing countries, as they need to combine remarkable economic growth targets with emission reduction goals.[3] If international technology transfers occur and emerging economies adopt greener production methods, the global production network will eventually become more sustainable, and CO_2 emissions per value added will converge across countries. This will in turn promote convergence in emissions per capita.

[1] See Knutti et al. (2015) for a critical analysis of the 2 °C target.

[2] Empirical studies that investigate the existence of an EKC in CO_2 emissions usually fail to find such a relationship in samples covering a large group of countries (see e.g. Stern 2004, and Stern 2017, for exhaustive surveys, or Fernández-Amador et al. 2017, for a survey of empirical applications). Aslanidis and Iranzo (2009) and Fernández-Amador et al. (2017) provided evidence that the income elasticity of CO_2 emissions decreases as income per capita rises above a threshold level though emissions continue growing, what challenges the sustainability of economic growth.

[3] The Paris Agreement recognizes the need to support developing countries in order to facilitate the effective implementation of the objectives identified in the Agreement (Paris Agreement, Art. 2).

The assessment of convergence in CO_2 emissions has received considerable attention in the empirical literature.[4] Most studies tested for convergence in CO_2 per capita across different groups of countries, but their results remain broadly inconclusive.[5] By contrast, a smaller number of studies investigated convergence in carbon efficiency, pointing invariably towards the existence of convergence across countries.[6] However, all these studies focused on production-based emissions, while cross-country convergence in CO_2 embodied in consumption has not yet been investigated.[7]

We evaluate international convergence in CO_2 emissions per capita and per value added derived from national production- and consumption-based inventories worldwide. We put forward a Bayesian test for β-convergence that is based on the theoretical models by Brock and Taylor (2010) and Ordás Criado et al. (2011). Our model also allows for potential group-specific dynamics of convergence using Bayesian shrinkage priors. Our convergence test is robust to heteroscedasticity and accounts for potential endogeneity between the growth rates of emissions and GDP per capita by means of instrumental variables (IV) estimation.

Our contribution is twofold. First, we assess international convergence in production- and consumption-based carbon emissions for the first time by using a comprehensive dataset on comparable CO_2 emission inventories published, and recently updated, by Fernández-Amador et al. (2016). The dataset covers 178 economies (grouped in 66 countries and 12 composite regions) and extends over 17 years after the ratification of the Kyoto Protocol, a period marked by the implementation of environmental policies against climate change in developed countries. The focus on both inventories allows to account for the increasing detachment between CO_2 per capita generated by production activities and CO_2 embodied in final consumption in a period of rapidly expanding global production networks, which permits cross-border sourcing of carbon in final consumption. In addition, we analyze CO_2 emissions per value added (carbon intensity or efficiency) and draw conclusions on whether the detected patterns are driven by efficiency effects. While CO_2 per capita offers important insights on convergence stemming from the expansion of production or consumption in a country, convergence in CO_2 intensity provides information on whether countries that use more pollutant production methods eventually catch up with environmentally more efficient economies.

Second, our structural model presents some interesting features. It uses a Bayesian stochastic search variable selection prior (SSVS, George and McCulloch 1993) to test for the existence of group-specific convergence dynamics. The groups comprise the European Union (EU), the OECD, and the countries that ratified Annex B to the Kyoto Protocol. The model is robust to cross-sectional heteroscedasticity—it is based on a scale mixture of multivariate normals, where the hyperparameter governing the distribution of the individual-specific

[4] See Pettersson et al. (2014) and Stern (2017) for comprehensive surveys of the literature on convergence in pollution emissions.

[5] The findings of the literature range from evidence for convergence (Strazicich and List 2003; Nguyen 2005; Ezcurra 2007; Romero-Ávila 2008; Lee et al. 2008; Westerlund and Basher 2008; Lee and Chang 2009; Brock and Taylor 2010; Jobert et al. 2010; Huang and Meng 2013; Yavuz and Yilanci 2013; Anjum et al. 2014; Hao et al. 2015; Wu et al. 2016; Zhao et al. 2015) over the existence of convergence clubs (Nguyen 2005; Aldy 2006; Lee and Chang 2008; Panopoulou and Pantelidis 2009; Barassi et al. 2011; Ordás Criado and Grether 2011; Camarero et al. 2013; Herrerias 2013; Wang et al. 2014; Burnett 2016) to no evidence for convergence (Aldy 2007; Barassi et al. 2008; Nourry 2009).

[6] See Anjum et al. (2014), Camarero et al. (2013) and Panopoulou and Pantelidis (2009).

[7] Aldy (2007) investigated convergence of CO_2 emissions across US states. This is the only study so far that also covers consumption inventories. The author did not find evidence for convergence for either CO_2 production or for CO_2 consumption per capita. In contrast to Aldy, our study covers economies at different development states, thus being the first one to evaluate global convergence patterns in CO_2 consumption.

variances is estimated endogenously. Furthermore, it formulates a flexible Cholesky-prior to instrument potentially endogenous regressors. This prior was proposed by Lopes and Polson (2014) in the framework of normal distributions, and, to the best of our knowledge, our study is the first to apply it in the context of scale mixtures of multivariate normals.[8]

Our results point to the existence of country-specific conditional convergence in all four emission inventories. The speed of convergence implies a half-life of 2.7–3.1 years for production inventories and 3.6–4.7 years for CO_2 consumption. Convergence towards global steady states, though conditioned on political and economic structures, is much slower, implying a half-life of 15 and 26 years for emissions from production and consumption per capita, respectively, and 44–45 years for emission intensities. Moreover, we do not find support for the existence of group-specific convergence dynamics for countries belonging to the OECD, the EU, or the Annex B to the Kyoto Protocol. These findings evince the ineffectiveness of environmental policies implemented in developed economies and pose doubts on the feasibility of an effective global action against climate change.

The next section reviews the literature on convergence. Section 3 describes the data. In Sect. 4, we explain the specification of the convergence test. Section 5 presents the results and Sect. 6 concludes.

2 Literature Review

Convergence tests received considerable attention in the empirical literature evaluating the predictions of the Solow (1956) growth model. Early studies tested whether countries starting out from low income levels experienced higher subsequent growth rates, either conditional or not conditioned on control variables (β-convergence).[9] Later studies suggested that β-convergence could be driven by regression to the mean (see Friedman 1992; Quah 1993) and tested whether the dispersion of income across countries was narrowing over time (σ-convergence).[10] Yet, Sala-i-Martin (1996) pointed out the merits of β-convergence for providing insights into growth dynamics. Although β-convergence is not sufficient for σ-convergence, it is a necessary condition (Sala-i-Martin 1996; Young et al. 2008) and provides valuable information whenever alternative tests for convergence cannot be applied.[11]

Besides cross-sectional convergence tests, also time-series approaches have been developed. Several authors investigated stochastic convergence of income levels via unit root

[8] Salois and Balcombe (2015) proposed a related model in the context of cross-sections, where t-distributed errors in an IV-model are represented by weighted errors of normals. Their modelization shares with ours the use of a scale mixture of normals representation, though the authors do not perform Cholesky-rotation of the system to represent it as a recursive system of equations but condition the weighted errors on each other and use a Wishart prior for the variance–covariance matrix.

[9] Earlier studies focused on unconditional convergence, while more recent studies tested for conditional convergence, i.e. convergence after allowing for heterogeneity across countries by accounting for additional determinants of economic growth. While unconditional convergence was often found for OECD countries, it was generally rejected for samples including non-OECD countries. If countries converge to different steady states, unconditional convergence models might result in biased coefficient estimates because the model used for estimation is miss-specified (Barro and Sala-i-Martin 2004). See for example Baumol (1986), Barro (1991), Barro and Sala-i-Martin (1992), Mankiw et al. (1992) and Barro and Sala-i-Martin (2004).

[10] See e.g. Barro and Sala-i-Martin (1992), Quah (1993), Sala-i-Martin (1996) and Young et al. (2008). Phillips and Sul (2007b) developed a test for identifying club-convergence groups, which corresponds to a test for conditional σ-convergence (see Phillips and Sul 2007b). Phillips and Sul (2007a) provided a short empirical application of the test in the context of economic growth convergence.

[11] See e.g. Ravallion (2003) who applied β-convergence tests to international income inequality.

testing; that is, whether income shocks are of permanent or temporary nature.[12] While these approaches became increasingly popular as more data became available over time, Bernard and Durlauf (1996) pointed out that they are grounded on the assumption that the economies in the sample are near their long-run equilibria. In this sense, the use of time-series tests may be invalid if the data are driven by transition dynamics.[13]

Similar to the Solow model for economic growth, there are theoretical models that predict convergence of pollution emission levels across countries over time (e.g. Brock and Taylor 2010; Ordás Criado et al. 2011). Like the Solow model, these boil down econometrically to an equation of conditional β-convergence.

Empirical studies on convergence in CO_2 per capita derived from production activities led to inconclusive findings. For OECD countries, Strazicich and List (2003), Romero-Ávila (2008), Lee et al. (2008), Lee and Chang (2009), Jobert et al. (2010), and Yavuz and Yilanci (2013) found evidence for convergence. Lee and Chang (2008) and Barassi et al. (2011) reported convergence only for a subgroup of countries, and Barassi et al. (2008) did not detect evidence for convergence.[14]

A growing number of studies included developing countries in their samples. Ezcurra (2007), Westerlund and Basher (2008), Brock and Taylor (2010), and Anjum et al. (2014) provided evidence for convergence across countries of different income status. Panopoulou and Pantelidis (2009), Ordás Criado and Grether (2011), and Herrerias (2013) detected several convergence clubs,[15] and Nguyen (2005) and Aldy (2006) found convergence only in subgroups or clubs of developed economies. Nourry (2009) failed to detect evidence for cross-country convergence.

Some authors focused on convergence across regions in China and states in the US. For China, Huang and Meng (2013) detected overall convergence and Wu et al. (2016) found evidence for club convergence. For the US, Burnett (2016) found a club of 26 converging states, while convergence for the US as a whole was rejected. While all these studies focused on CO_2 production inventories, Aldy (2007) additionally assessed consumption of CO_2 per capita in the US states but did not find convergence in either measure.

The heterogeneous findings of the literature on CO_2 convergence are in line with the inconclusive evidence for the existence of an environmental Kuznets curve (EKC). The EKC hypothesis suggests that as national income levels rise, pollution first increases with income, but after a certain level of income is reached this mechanism is reversed.[16] If income levels are related to CO_2 emissions, the existence of an EKC relationship would ultimately lead to emission convergence (Stern 2017). However, even though empirical studies find a positive relationship between economic growth and CO_2 emissions, the evidence favoring an EKC-type relationship is restricted to time-series or panel studies covering OECD economies.[17]

[12] See e.g. Carlino and Mills (1993), Quah (1993), Bernard and Durlauf (1996) and Evans and Karras (1996).

[13] See also Panopoulou and Pantelidis (2009), Jobert et al. (2010) and Ordás Criado and Grether (2011) for surveys on β-, σ- and stochastic convergence.

[14] Studies for OECD countries focused mainly on stochastic and β-convergence. For more details on the concept of convergence used by the respective studies, see Table A.1 in the Online Appendix.

[15] Panopoulou and Pantelidis (2009) and Herrerias (2013) applied the Phillips and Sul (2007b) test for convergence clubs. Ordás Criado and Grether (2011) found evidence for income-specific and regional convergence clubs especially for the sub-period 1980–2000.

[16] See Dasgupta et al. (2002), Kaika and Zervas (2010), and Stern (2004, 2017) for reviews, and Fernández-Amador et al. (2017) for a summary of the most recent evidence.

[17] Schmalensee et al. (1998) is an exception, finding support for an inverse-U relationship using non-parametric techniques. More recently, Aslanidis and Iranzo (2009) and Fernández-Amador et al. (2017) found

Improvements in carbon efficiencies (i.e. CO_2 per value added) are an important requirement for reaching the turning point postulated by the EKC. High-income countries generally are more carbon efficient than less developed economies (Fernández-Amador et al. 2016). This can be explained by their stronger preferences for a cleaner environment, better access to cleaner technology and potential for carbon leakage. Carbon leakage will impede convergence in carbon emission intensities, as firms with larger emission intensity might relocate to countries with less stringent environmental regulation. However, if the rapid increase in international trade induces transfers of green technology to less developed countries, their carbon efficiency could improve more rapidly (Grossman and Helpman 1995), which would contribute to convergence in carbon intensities. Thus, although most studies focused on CO_2 emissions per capita, evaluating convergence in carbon intensities provides additional insights in the convergence patterns across countries.[18]

Among the existing studies on convergence in the intensity of CO_2 emissions from production activities, Camarero et al. (2013) identified four convergence clubs among 22 OECD countries using the test for club-convergence developed by Phillips and Sul (2007b). Anjum et al. (2014) and Panopoulou and Pantelidis (2009) provided evidence for convergence in a panel of 136 and 128 countries, respectively. Focusing on Chinese regions, the results of Hao et al. (2015) and Zhao et al. (2015) suggested convergence of emission intensity, while Wang et al. (2014) found evidence for club convergence.[19]

3 Data

CO_2 emissions per capita and per value added derived from production and consumption inventories are available from the emissions database constructed by Fernández-Amador et al. (2016). Following Fernández-Amador et al. (2016), we define carbon intensities as carbon per value added rather than per GDP. For production inventories, value added is computed as value added embodied in production, whereas for consumption-based inventories, it is calculated as value added embodied in consumption. Therefore, emission inventories and value added are measured at the same stage of the supply chain. The dataset consists of a balanced panel of national production- and consumption-based carbon dioxide emission inventories from fossil fuel combustion covering 66 countries and 12 composite regions (encompassing a total of 178 economies) over the years 1997, 2001, 2004, 2007, 2011 and 2014 (468 observations).[20] It relies on input–output, trade, and energy data of several releases of the Global Trade Analysis Project (GTAP) database.

To test for the presence of β-convergence, we compute the growth rates of the four emission inventories, which we consecutively use as dependent variables in the empirical analysis. Since the data-points are unequally spaced in time (3–4 years), we calculate the average

Footnote 17 continued

that the income-elasticity of CO_2 emissions decreases slightly after income per capita passes a certain threshold, such that relative decoupling increases with economic growth, though there is no evidence of absolute decoupling and an EKC relationship. Fernández-Amador et al. (2017) also provided evidence for a similar pattern in CO_2 consumption-based inventories.

18 Anjum et al. (2014) reported that the negative correlation between initial emissions and subsequent emission growth is stronger for CO_2 intensity than for CO_2 per capita.

19 All these studies define CO_2 intensity as CO_2 per GDP. In our analysis we refer to CO_2 intensity as CO_2 per value added.

20 A description of the countries included in the composite regions is available in Fernández-Amador et al. (2016). The dataset has been recently extended by the authors to cover the year 2014.

growth rate of emissions between years $t - s$ and t, where s is the number of years between two observations (see Ravallion 2003, for a similar methodology).[21] The resulting average growth rates allow to evaluate convergence in the large-N, small-T panel dataset, for which time-series methods cannot be used.[22] Furthermore, by using average growth rates, we avoid to capture short-term fluctuations in emissions that could result in an upward bias of the estimates of the convergence speed (see Ordás Criado et al. 2011).

Our baseline control variables are derived from the theoretical model by Ordás Criado et al. (2011). They comprise the lagged level of CO_2 emissions, which should capture potential convergence forces, the growth rate of purchasing-power parity (ppp) adjusted real GDP per capita over the period considered, which should capture the scale effect of economic growth on emissions, and the lagged level of ppp-adjusted GDP per capita, as a proxy for capital per efficient labor (see Ordás Criado et al. 2011, for details). GDP per capita is sourced from the World Development Indicators (WDI) dataset.

To limit potential omitted variable bias (see Barro and Sala-i-Martin 2004), we add a large set of additional control variables capturing economic, structural, and institutional characteristics of the individuals in the sample, and include individual- and time-dummies (see Table A.2 in the Online Appendix for details on the variable definitions and data sources). We derive trade flows as a share of GDP and value added shares of different sectors of the economy (agriculture, energy, light manufacturing, heavy manufacturing, textiles, water services, construction, trade and transport, and remaining services) from the GTAP database.[23] Data on population density, the share of fossil fuels and nuclear energy in total electricity production, and rents from fossil fuel production as a share of GDP are available from the WDI database. A political regime index, which may channel citizens' preferences for a cleaner environment, is sourced from the Polity IV database (see Farzin and Bond 2006). Finally, in order to investigate group-specific convergence patterns, we generate dummy variables for members of the EU, OECD, and Annex B to the Kyoto Protocol.

4 Econometric Model

We develop a Bayesian test for β-convergence as an extension of the model proposed by Ordás Criado et al. (2011).[24] The model specification is a dynamic panel that enables to test

[21] This corresponds to calculating average annual growth rates. For a similar method see Ravallion (2003), who accounts for the unequal spacing in time between measures of income inequality for large-N, small-T panel data by regressing the difference in inequality between time t and the initial period t_1 on a constant and initial inequality at time t_1, both multiplied by a time-trend $(t - 1)$. In contrast to Ravallion's data, our panel is balanced in the sense that for every individual we observe all variables at the same points in time. Thus, we can also exploit the variation of the data across time and use initial emissions in year $t - s$ instead of in year t_1 as a regressor.

[22] Bernard and Durlauf (1996) pointed out that the power of time-series tests may be weak when the dynamics do not occur near the steady state. In this sense, time-series approaches to test for stochastic convergence may not be particularly suitable in our context, since data on CO_2 emissions covering a global sample of countries are very likely to be driven by transition dynamics rather than being near the steady state.

[23] Detailed information on the sector aggregation from the original GTAP sectoral disaggregation is available from the authors upon request.

[24] Ordás Criado et al. (2011) tested for convergence in sulfur oxides and nitrogen oxides. Their theoretical model assumes optimal control of pollution emissions at the national level, making it particularly suited for applications to local air pollutants. Nevertheless, the structure of the empirical model they specify is compatible with the green Solow model by Brock and Taylor (2010), which the authors applied to CO_2 emissions. Ordás Criado et al. (2011) regressed the average growth rates of emissions over the period $t - 5$ to t on the level of

for the existence of specific convergence groups, accounts for endogeneity of the regressors, and allows for cross-sectional heteroscedasticity of the error terms.

Let E_{it} be, alternatively, the natural logarithm of CO_2 emissions per capita or per value added in country i at time t, where $i \subseteq [1, \ldots, N]$ and $t \subseteq [1, \ldots, T]$, and let $G_{i,t,s} = (E_{i,t} - E_{i,t-s})/s$ be the average growth rate of E_i over the period $t - s$ and t. The test for convergence is defined by the following recursive structural model with instrumental equation:

$$G_{i,t,s} = \beta E_{i,t-s} + \pi_0 g_{i,t,s} + \pi_1 Y_{i,t-s} \sum_r \left[\theta_r z_{r,i,t-s}\right] + \\ + \delta_t + \alpha_i + \sum_j \left[\beta_j d_j E_{i,t-s}\right] + \epsilon_{1,it} \quad (1)$$

$$g_{i,t,s} = \alpha_{iv} + \beta_{iv} L(g_{i,t,s}) + \epsilon_{2,it} \quad (2)$$

$$(\epsilon_{1,it}, \epsilon_{2,it}) \sim N(0, \Sigma\omega_i) \quad (3)$$

The (average) growth rate of emissions ($G_{i,t,s}$) over $t - s$ and t depends on the logarithm of the level of emissions in country i at period $t - s$ ($E_{i,t-s}$), the (average) growth rate of real GDP per capita over the period $t - s$ and t ($g_{i,t,s}$), the logarithm of real GDP per capita of country i in $t - s$ ($Y_{i,t-s}$), a set of control variables as described in the data section ($z_{r,i,t-s}$), time-effects (δ_t), and individual-dummies (α_i). The d_j's are dummy variables for group membership in the EU, OECD, and Annex B to the Kyoto Protocol. The parameter associated with $E_{i,t-s}$ is the parameter of interest; in particular, $\beta < 0$ provides evidence for conditional β-convergence.

The relationship between the growth rate of emissions and the growth rate of GDP per capita is potentially endogenous. Thus, we follow Barro and Sala-i-Martin (1992) and instrument the growth rate of GDP per capita with its growth rate in the previous period, denoted $L(g_{i,t,s})$, as described in Eq. (2), where $L(\cdot)$ is the lag operator.[25]

The priors for the parameters in (1) and (2) are collected in the following set of equations:

$$\beta \sim N(0, \tau) \quad (4)$$

$$\pi_0, \pi_1, \theta_r, \delta_t, \alpha_{iv}, \beta_{iv} \sim N(0, \phi) \quad (5)$$

$$\alpha_i \sim N(0, \psi) \quad (6)$$

$$\beta_j \sim (1 - \gamma_j) N(0, \kappa_0) + \gamma_j N(0, \kappa_1) \quad (7)$$

$$\gamma_j \sim Bernoulli(p) \quad (8)$$

$$\omega_i^{-1} \sim \Gamma(\nu/2, \nu/2) \quad (9)$$

The prior of β follows a normal distribution with zero mean and precision τ, where $\tau = (2/3)^2$, such that the case for unit root in the original dynamic model of emissions

Footnote 24 continued
emissions at the initial period of the growth rate ($t - 5$), the growth rate of GDP over $t - 5$ and t, GDP in $t - 5$, and time- and individual-dummies using OLS and a non-parametric model. The authors also addressed endogeneity between emissions and GDP by instrumenting GDP and its growth rate with their lagged values (following Barro and Sala-i-Martin 1992). Brock and Taylor (2010) developed a theoretical model that also predicts conditional β-convergence, which the authors applied to CO_2 emissions. Although Brock and Taylor's model is applicable to global pollutants, in our empirical approach we follow Ordás Criado et al. (2011) since their empirical analysis makes use of the panel structure of the data and accounts for the potential endogeneity of GDP per capita.

[25] For the first period in our sample, 1997–2001, we use the average growth rate for a period of the same length, 1993–1997, as instrument.

is not ruled out.[26] The priors of the parameters π_0, π_1, θ_r, δ_t, α_{iv}, and β_{iv} follow a normal distribution with zero mean and precision $\phi = 0.2$.[27] We estimate the individual-effects using the dummy variables approach, where α_i is normally distributed with precision $\psi = 0.5$.[28] An intercept of the model can be retrieved as $\overline{\alpha} = \frac{1}{N}\sum_i^N \alpha_i$.

Equations (7) and (8) characterize a hierarchical SSVS shrinkage prior (George and McCulloch 1993) that grants flexibility for the data to discriminate among models including group-specific convergence dynamics (for EU, OECD, and Annex B membership). Each group-specific prior on β_j is modeled as a mixture of two normals with different precisions κ_0 and κ_1. $\kappa_0 > \kappa_1$ so that when $\gamma_j = 0$, β_j is restricted to be estimated around 0, whereas when $\gamma_j = 1$, β_j remains unrestricted. We set $\kappa_0 = 10$ and $\kappa_1 = 1$. To reflect the absence of prior beliefs about the existence of specific group convergence we set $p = 0.5$.

The prior elicited in Eq. (9) defines the distribution of the variances of the individual-specific error terms $(\epsilon_{1,it}, \epsilon_{2,it})$. Each individual-specific variance parameter takes the form $\Sigma\omega_i$, such that the model exhibits cross-sectional heteroscedasticity. The equation is defined in terms of precisions (inverse of the variances). The gamma prior for the precisions in Eq. (9) is equivalent to a $\chi^2(\nu)/\nu$ and characterizes the model as a scale mixture of normals, where the weights are individual-specific.[29]

The hyperparameter ν is estimated endogenously with prior

$$\nu = \lfloor u \rfloor \tag{10}$$

$$u \sim Exp(1/\lambda), u \in [3, 60], \tag{11}$$

where the function $\lfloor \cdot \rfloor$ rounds the values of u to the nearest integer. Exp in Eq. (11) stands for an exponential distribution where the rate parameter λ is set to 25, such that the density function of ν is centered at a mean of 25, giving substantial prior weight both to fat-tailed error distributions ($\nu < 10$) and error distributions which are effectively Normal ($\nu > 40$). The estimation of ν renders the specification in (1)–(3) rather flexible. Small values of ν will yield heteroscedasticity-robust parameter estimates, while as ν increases the errors' distribution will approach normality (homoscedasticity). We truncate the prior for ν such that it is contained in the interval [3, 60].[30]

[26] In the original dynamic model of emissions, once we undo the average growth rate, the relevant parameter for the case of regular sampling every period is $(1 + \beta)$. The precision elicited ensures that the hypothesis of an unit root in our autoregressive model with explanatory variables is not an extreme event in our prior for β.

[27] The precision is defined as the inverse of the variance. A precision of 0.2 implies a variance of 5.

[28] Note that the precision of the individual-dummies is larger than the precision of the rest of the parameters. A uniform prior on the individual fixed effects would lead to improper posterior distributions for the parameters of interest, while very diffuse priors would lead to very slow convergence of the MCMC algorithm used for inference (see e.g. Lancaster 2008, Chap. 7).

[29] Scale mixture of normals with the weights specified as in (9) are equivalent to a t-student distribution (see e.g. Andrews and Mallows 1974; West 1987; Ding 2016). The degrees of freedom of the t-student are equal to the hyperparameter governing the distribution of the weights ω_i, ν. With growing ν the distribution converges to a normal distribution, as less probability mass is concentrated at the tails of the distribution. The prior for the weights in the scale mixture of normals, ω_i, together with the prior for the components of the variance matrix Σ that we will define below, imply a form of cross-sectional heteroscedasticity of the gamma type (Andrews and Mallows 1974; Geweke 1993; Koop 2003, Chap. 6; Lancaster 2008, Chap. 3). There are two main advantages of modeling the problem in terms of scale mixture of normals rather than as a t-student distribution. The first one is that the type of heteroscedasticity, cross-sectional in our case, can be explicitly stated. The second is that it is less computational demanding for the numerical algorithm to estimate the posterior distributions of the parameters.

[30] We regard the priors for the parameters of interest $(\beta, \pi_0, \pi_1, \{\theta_r\}, \{\delta_t\}, \alpha_{iv}, \beta_{iv}, \{\alpha_i\}, \{\beta_j\})$ as informative. Geweke (1993) shows that under informative (normal) priors for the slopes, both the first and the second moments of the slopes exist. When the priors of the slopes are uninformative, $\nu > 2$ ensures existence of the

In order to complete the prior for the covariance matrix in (3), we propose a Cholesky-based prior for Σ. Lopes and Polson (2014) have shown the better performance of this type of prior compared to the more widely used approach of specifying an inverted Wishart prior for Σ in IV-models in the context of normal-distributed errors.[31] More specifically, the components of the error vector are modeled based on the recursive conditional regressions arising from the Cholesky decomposition of $\Sigma = ADA'$, such that $D = diag(\Sigma_{1|2}, \Sigma_{22})$ and A is an upper triangular matrix with ones in the main diagonal and upper triangular component $a_{12} = \Sigma_{12}/\Sigma_{22}$. However, the specific modelization of heteroscedasticity by means of scale mixture of normals requires taking into consideration the effect of the Cholesky-rotation in the individual-specific term of the variance (see Ding 2016).[32] Thus, Eq. (3) can be re-written in recursive conditional form, using the specification of the conditionals of a multivariate scale mixture of normals.

$$\epsilon_{1|2,it} \sim N\left(a_{12}\epsilon_{2,it}, \Sigma_{1|2}\omega_{1|2,i}\right) \tag{12}$$

$$\epsilon_{2,it} \sim N\left(0, \Sigma_{22}\omega_{2,i}\right), \tag{13}$$

where $\Sigma_{11} = \Sigma_{1|2} + \Sigma_{12}^2/\Sigma_{22}$. We must specify priors for Σ_{22}, the conditional variance $\Sigma_{1|2}$, the parameter a_{12}, which calibrates the strength of the correlation between $\epsilon_{1,it}$ and $\epsilon_{2,it}$, as well as the weights in the instrumental equation, $\omega_{2,i}$, and in the main equation conditional on the instrumental equation, $\omega_{1|2,i}$. We assign Σ_{22}^{-1} and $\Sigma_{1|2}^{-1}$ a gamma prior with shape and scale parameters $a, b = 0.001$ so that we remain uninformative about the precision of the model. a_{12} follows a normal prior centered at zero and with precision $\tau = 0.2$. Finally, the priors for $\omega_{2,i}^{-1}$ and $\omega_{1|2,i}^{-1}$ also follow a gamma distribution with hyperparameter ν as defined in Eq. (10) and, where $d_m = \epsilon_{2i}'\Sigma_{22}^{-1}\epsilon_{2i}$ is the square Mahalanobis distance in the instrumental equation, on which the main equation is conditioned. That is,

$$\Sigma_{22}^{-1}, \Sigma_{1|2}^{-1} \sim \Gamma(a, b) \tag{14}$$

$$a_{12} \sim N(0, \tau) \tag{15}$$

$$\omega_{2,i}^{-1} \sim \Gamma(\nu/2, \nu/2) \tag{16}$$

$$\omega_{1|2,i}^{-1} \sim \Gamma\left((\nu+1)/2, (\nu+d_m)/2\right), \tag{17}$$

The relationship between the location and the rate parameters of the gamma priors in (16) and (17) deserves special attention: the location parameter of $\omega_{1|2,i}^{-1}$ has increased by 1/2 as compared to the location parameter governing $\omega_{2,i}^{-1}$, what reduces the heavy-tailedness of the (conditional) main equation. The rate parameter in (17), $(\nu + d_m)/2$, will be larger in

Footnote 30 continued
first moments, while $\nu > 4$ ensures existence of the second moments. Thus, the truncation defined contains roughly 80% of the density around the mean of the prior, while ensuring existence of first moments even in the case of noninformative priors for the parameters of interest.

[31] We explain the derivation of the IV-prior in terms of covariance matrices because this is common in the literature, though the specification of the priors is in terms of precisions, as carried out in the software. Alternatively, we could use an inverted Wishart prior for Σ, $\Sigma \sim IW(v_0, \Sigma_0)$, with parameters v_0 and Σ_0. Priors for covariance matrices and variances have usually been addressed by means of inverted Wishart and inverted Gamma distributions, respectively, while Wishart or Gamma distributions have been used as priors for precision matrices and precisions. Wishart priors have been extensively used in the framework of Bayesian instrumental variable models under normal-distributed errors (see e.g. Kleibergen and Zivot 2003; Lancaster 2008, Chap. 8; Rossi et al. 2005).

[32] Ding (2016) used the representation of a multivariate t-student distribution as a scale mixture of multivariate normals to derive the conditional distribution of the multivariate t-student, which can be represented by the conditional normal distribution times the conditional distribution of the weights.

comparison with (16), since d_m is typically larger than one, what increases the dispersion of the distribution of individuals' variances in the (conditional) main equation. That is, the more extreme the values of the endogenous variable are, the more dispersive is the conditional distribution of the explained variable in the (conditional) main equation.

A Markov Chain Monte Carlo (MCMC) algorithm is used to carry out Bayesian inference. Gibbs sampling can accommodate all priors specified, including the SSVS prior, Eqs. (7) and (8), the individual-specific weights and the Cholesky-based priors for covariance of the error terms, Eqs. (14)–(17), and the degrees of freedom parameter, Eqs. (10) and (11).[33] The vector of parameters to estimate is $P = (\beta, \pi_0, \pi_1, \{\theta_r\}, \{\delta_t\}, \{\alpha_i\}, \alpha_{iv}, \beta_{iv}, \{\beta_j, \gamma_j\}, \nu, \{\omega_{2,i}^{-1}, \omega_{1|2,i}^{-1}\}, \Sigma_{22}^{-1}, \Sigma_{1|2}^{-1}, a_{12})$. We implement three Markov chains from which, after a burn-in of 7.5×10^5 draws, we retain a posterior sample of 7.5×10^5 draws each.[34] We apply a thinning of 3, ending up with a mixed posterior sample of 7.5×10^5 draws. We average across the posterior sample to calculate the posterior means, standard errors and quantiles of the coefficients, and the posterior inclusion probabilities (PIP) of the coefficients associated with specific group convergence. The PIPs of the coefficients for group convergence show the posterior probability of observing specific dynamics associated with those groups.

The model proposed is a dynamic panel model. Nickell (1981) showed that incidental parameters yield inconsistent OLS or maximum likelihood (ML) estimates in dynamic panels with short time dimension. The phenomenon is a consequence of having a limited number of observations from which each incidental (individual-specific) parameter is estimated, which in turn contaminates the estimation of the common parameters and, in particular, of the dynamic (autoregressive) parameter.[35] The literature has proposed alternative estimators with the aim to correct Nickell (1981) bias such as IV estimators, generalized method of moments (GMM) estimators, analytical corrections for the least squares dummy variable (LSDV) estimator, and bias-corrected estimators based on iterative bootstrapping (see Everaert and Pozzi 2007, for a review of these estimators). Maddala and Hu (1996) and Hsiao et al. (1999) showed that the Bayesian approach performs fairly well in the context of dynamic panels when T is small, in comparison with some classical estimators.

Our posterior inference is based on the mean of the mixed posterior sample resulting from the Gibbs sampler after thinning. In addition, it is based on informative priors. Therefore, we expect that our posterior estimates do not suffer from considerable bias.[36] A simulation exercise under homoscedasticity and heteroscedasticity confirmed that the Bayesian estimator resulted in substantial bias reduction. The performance of the Bayesian estimator was comparable to, and sometimes outperformed, the performance of difference-GMM (Arellano and Bond 1991), system-GMM (Blundell and Bond 1998), an extension of Kiviet's (1995) bias-corrected estimator (see Bruno 2005), and De Vos et al.'s (2015) bootstrap-based bias correction.[37]

[33] See George and McCulloch (1993) for details on the Gibbs sampler for the SSVS prior, and Lopes and Polson (2014) for the details of the Gibbs sampling for IV-estimation in the context of the normal distribution.

[34] That was sufficient for the chains to show mixing and the estimates of the coefficients to show convergence to their ergodic distribution.

[35] The concept of incidental parameters and the problem of limited information to estimate incidental parameters was first defined by Neyman and Scott (1948). Lancaster (2000) and Moon et al. (2015) offer rigurous treatments of the incidental parameters problem.

[36] Several authors have shown the connection between the non-parametric bootstrap, the parametric (Bayesian) bootstrap and MCMC, respectively (see e.g. Rubin 1981; Efron 1982, 2011; Newton and Raftery 1994; Hastie et al. 2009, Chap. 8).

[37] The results from our simulations are available from the authors upon request. These simulations did not include the null of endogeneity between explained and explanatory variables and thus do not introduce an IV

Finally, as a robustness check, we also estimate an alternative (homoscedastic) model where the priors in Eqs. (3), (12), and (13) are replaced, respectively, by

$$\left(\epsilon_{1,it}, \epsilon_{2,it}\right) \sim N(0, \Sigma) \tag{18}$$

$$\epsilon_{1|2,it} \sim N\left(a_{12}\epsilon_{2,it}, \Sigma_{1|2}\right) \tag{19}$$

$$\epsilon_{2,it} \sim N\left(0, \Sigma_{22}\right) \tag{20}$$

where again $a_{12} = \Sigma_{12}/\Sigma_{22}$ and Eqs. (9), (16) and (17), for the priors of the individual-specific weights in the variance parameter, and Eqs. (10) and (11) for the prior of the degrees of freedom ν are eliminated. Therefore, the model collapses to the Bayesian IV-model proposed by Lopes and Polson (2014). The Gibbs sampling algorithm for estimating this model's posterior is simplified by deleting the steps corresponding to the parameters $\left\{\omega_{2,i}^{-1}, \omega_{1|2,i}^{-1}\right\}_{i=1,\ldots,N}$ and ν.

5 Results

We implement two types of IV models with errors distributed as a scale mixture of normals that differ in the inclusion or exclusion of individual-specific dummy variables (DV). The *DV-conditional heteroscedastic model* includes a set of economic, political and structural controls, and individual-specific dummy variables. It constitutes a test for (fully) conditional convergence. The *conditional heteroscedastic model* does not include individual-specific effects and is only conditioned on economic, political and structural variables. This model provides evidence on a stronger assumption about convergence than the *DV-conditional heteroscedastic model*, as it reflects the concept of global convergence.

Table 1 summarizes the results of the *DV-conditional heteroscedastic model*. The results of the *conditional heteroscedastic model* (without individual-dummies) are available in Table 2.[38] The four columns of the tables report the posterior means of the parameter estimates from the outcome (upper panel) and the instrumental equations (middle panel) together with the R^2, the Deviance Information Criterion (DIC), the PIPs of the regressors associated with specific-group convergence, the half-life derived from the convergence estimates, the posterior the hyperparameter governing the weights associated with the country-specific variances ν, and the number of observations of the regressions for the four CO_2 inventories (CO_2 per capita and per value added for production and consumption inventories; lower panel).[39] The asterisks next to the parameter estimates indicate whether the parameter is different from zero at the 99%, 95% or 90% (equal-tailed) credible intervals (CI).

The estimated ν turn out to be very low (between 4 and 5), pointing to the existence of heteroscedasticity for each of the four inventories, for the specifications with and without individual fixed effects (Tables 1 and 2, respectively). The R^2 are relatively high throughout, indicating that the included regressors explain a large part of the variation in the growth

Footnote 37 continued
structure in the model. A more detailed simulation-based analysis of the Bayesian estimator in comparison with alternative dynamic panel estimators can be found in Fernández-Amador and Oberdabernig (2018).

38 Furthermore, we report the results of the *DV-conditional* and *conditional homoscedastic models* (normal-distributed errors with and without individual-dummies) in Tables A.4 and A.5 in the Online Appendix.

39 The underlying data cover 468 observations. Because we use the growth rate of emissions as dependent variable, our final sample includes 390 observations.

rates of all four emission inventories.[40] In particular, the *DV-conditional model* explains 73–74% of the variation in growth of CO_2 emissions per capita and per value added embodied in production activities, while it accounts for 61–65% of the variation in the growth rate of emissions per capita and per value added embodied in consumption. The explanatory power of the *conditional model* decreases to 26–28% for production inventories, and respectively 31% and 47% for CO_2 consumption intensities and CO_2 consumption per capita. The difference between the R^2 of the specifications underlines the importance of country-specific steady states. Additionally, the inspection of the DIC across specifications lends support to both *DV-conditional* and *conditional* models as a representation of the dynamics of carbon emissions per capita and per value added.

In all specifications we instrument the growth rate of income per capita in order to account for potential reverse causality (see Barro and Sala-i-Martin 1992). The coefficient of lagged income per capita growth, which we use as an instrument, is positive with a CI of 99% in each specification, indicating a high relevance of this variable. At the same time it is exogenous, as emission growth cannot affect lagged growth rates of income per capita. The estimate for a_{iv}, the strength of the correlation between the errors of the instrumental and the outcome equations is insignificant at the 90% CI for all emission inventories but for carbon emissions per capita from consumption in the *DV-conditional model*.

5.1 DV-Conditional Convergence

For the *DV-conditional convergence model* in Table 1, the posterior mean of the parameter connected to lagged emissions (the convergence parameter, β) reveals a negative effect of lagged emissions on the average growth rate of all four emission inventories, at a CI of 99%. This provides strong evidence for convergence in all four CO_2 emission inventories. The magnitudes of the posterior mean of the convergence parameter are larger in absolute value for production inventories than for consumption inventories.

Given the size of the convergence parameters, it is possible to calculate the time needed for countries to halve their emission gap towards their country-specific steady states. Assuming that the average emission trajectories observed in the sample remain unchanged, the half-life of emissions amounts to 3.1 (3.6) years for CO_2 per capita production (consumption) and 2.7 (4.7) years for CO_2 production (consumption) per unit of value added.[41] These rather fast convergence rates implied by our estimates are in line with the findings of Westerlund and Basher (2008) and Jobert et al. (2010) for CO_2 per capita from production activities. Westerlund and Basher (2008) reported a half-life of CO_2 emissions per capita between 3.1 and 6.1 years in a sample of developed and developing countries.[42] Jobert et al. (2010) found the half-life of CO_2 emissions to be between 2.2 and 3.4 years for various OECD countries.[43]

[40] It should be noted that Bayesian estimation does not aim at minimizing the sum of square residuals and thus, it does not maximize the R^2. However, we consider it together with the DIC when assessing how well our models fit the data and whether they can be regarded as consistent with our data. The DIC penalizes the number of parameters and is often regarded as a better measure of fit in the Bayesian context than the R^2.

[41] The half-life provides an indication of the speed of convergence. It is defined as the time required to eliminate half of the initial gap between actual emissions levels and the steady state. The half-life is calculated as $\frac{-ln(0.5)}{-ln(1+\beta)}$ (see Allington and McCombie 2007, p. 206).

[42] The half-life in their sample of developed countries was estimated to lie between 4.2 and 6.2 years; this is longer than the half-life estimated in their pooled sample including developing countries.

[43] These figures correspond to estimates of conditional convergence. For unconditional convergence the authors reported a half-life between 4 and 8.5 years.

Table 1 Results scale-mixture of normals, *DV-conditional heteroscedastic model*

	(1) CO_2 pc prod.	(2) CO_2 pc cons.	(3) CO_2 va prod.	(4) CO_2 va cons.
Outcome equation				
Constant	− 0.4089***	− 0.4012***	0.2011	0.1240
Ln(emissions)	− 0.2009***	− 0.1741***	− 0.2269***	− 0.1372***
Ln(emissions)·EU	− 0.0001	− 0.0001	− 0.0005	− 0.0002
Ln(emissions)·OECD	− 0.0004	− 0.0004	0.0066	0.0004
Ln(emissions)·Annex B	0.0000	0.0000	− 0.0008	− 0.0124
Ln(Income pc)	0.0932***	0.0769***	− 0.0187	− 0.0020
Income pc growth	0.8159***	1.5034***	− 0.4024	− 0.1206
Ln(pop. density)	− 0.0750***	− 0.0423	− 0.0189	− 0.0261
Fossil rents	0.0029**	0.0019	0.0025*	0.0015
Nuclear (%)	0.0007*	0.0008*	0.0005	0.0001
Fossil fuels (%)	0.0007**	0.0007**	0.0007*	0.0002
Openness	− 0.0001	0.0000	0.0000	0.0000
Political regime	− 0.0017**	− 0.0009	− 0.0020**	− 0.0015*
VA energy (%)	0.0003	0.0004	− 0.0004	− 0.0005
VA light manufacturing (%)	0.0008	0.0002	0.0006	0.0007
VA heavy manufacturing (%)	− 0.0004	0.0002	− 0.0018*	− 0.0020*
VA textiles (%)	0.0026	− 0.0005	0.0070***	0.0016
VA water services (%)	0.0146*	0.0083	0.0075	0.0038
VA construction (%)	− 0.0020*	0.0007	− 0.0039***	− 0.0002
VA trade and transport (%)	0.0004	0.0001	0.0001	− 0.0004
VA other services (%)	0.0006	− 0.0007	0.0002	0.0004
2004	0.0244***	0.0291***	− 0.0058	− 0.0135**
2007	0.0064	0.0171***	− 0.0244***	− 0.0338***
2011	0.0119*	0.0199***	− 0.0248***	− 0.0383***
2014	− 0.0002	0.0102	− 0.0485***	− 0.0554***
Individual-dummies	Yes	Yes	Yes	Yes
R^2	0.7272	0.6513	0.7398	0.6055
DIC	− 3283.5	− 3212.1	− 3165.0	− 3141.1
Instrumental equation for income pc growth				
Constant	0.0147***	0.0147***	0.0147***	0.0147***
Income pc growth, lagged	0.3597***	0.3618***	0.3596***	0.3597***
a_{iv}	− 0.1678	− 0.5926**	0.1366	− 0.1570
R^2	0.5328	0.5334	0.5328	0.5322
PIP EU	0.0100	0.0225	0.0192	0.0223
PIP OECD	0.0062	0.0185	0.1885	0.0192
PIP Annex B	0.0229	0.0032	0.0309	0.5115
Half-life	3.0907	3.6237	2.6934	4.6970
ν	4.6581	4.8148	4.6797	4.4001
N	390	390	390	390

*CI 90%, **CI 95%, ***CI 99%. All variables but group dummies and income pc growth enter in lagged values. The half-life is calculated as $-ln(0.5)/-ln(1+\beta)$ (see Allington and McCombie 2007). The Bayesian R^2 is the mean of the R^2 computed for each draw q of the Markov chain (MC), R^2_q where $R^2_q = \sum_{i,t} \hat{y}_{it} / \left(\sum_{i,t} \hat{y}_{it} + \sum_{i,t} \epsilon_{it}\right)$, where $\hat{y}_{it}$ is the estimate of y_{it} implied by the model and $\epsilon_{it} = y_{it} - \hat{y}_{it}$ (see Gelman et al. 2017). The Deviance Information Criterion (DIC) is computed as $DIC = \hat{D_q} + Var(D_q)/2$, where D_q is the deviance measure associated with draw q in the MC (see Spiegelhalter et al. 2002; Gelman et al. 2004, Chap. 7)

Thus, our results confirm that the findings of earlier studies covering a smaller number of countries also hold for a sample of countries comprising the whole world.

There is no strong evidence for the existence of specific convergence dynamics for the EU, OECD or Annex B members. The PIPs of the group-specific regressors are usually smaller than 10%, with the exception of the group of OECD countries in the model for CO_2 production per value added and the group of Annex B countries in the model for CO_2 consumption per value added, with PIPs of 19% and 51%, respectively. A low PIP implies that the estimation algorithm tends to exclude group-specific dynamics. Consequently, the slope estimates of the group-specific regressors are very low in magnitude and not different from zero at any of the CIs considered. Also the group-specific convergence terms with a higher PIP fail to be different from zero at any of the specified CIs.

Some of the control variables capturing economic and institutional characteristics have significant effects on emission growth. Higher per capita income and a higher growth rate of per capita income are associated with higher growth rates of CO_2 per capita, while CO_2 intensities are not significantly affected by these variables. This highlights the role of energy—and thus energy-derived CO_2 emissions—as a necessary input for production and consumption patterns. Population density has a negative effect on the growth rate of CO_2 per capita from production inventories. The opposite is true for the share of rents from fossil fuel production in GDP, which has a positive effect on the growth rate of both CO_2 production inventories. With respect to the variables related to the composition of electricity production in an economy, a higher share of fossil fuels or nuclear sources in total electricity production is connected to a larger growth rate of CO_2 emissions per capita, and in the case of fossil fuels also to a higher growth rate of CO_2 production intensity. Noteworthy, trade openness does not affect emission growth for any of the inventories considered. More democratic regimes are connected to lower growth rates of CO_2 for all inventories but CO_2 consumption per capita, suggesting that democracy may be a channel through which citizens' preferences are revealed (see Farzin and Bond 2006).[44]

Regarding the sectoral shares in value added, only four sectors are relevant at a CI of at least 90%. These are heavy manufacturing, which tends to reduce the growth rate of CO_2 intensity, textiles, which increases the growth of CO_2 production intensity, water services, which are connected to higher emission growth rates for CO_2 per capita from production, and the construction sector, which lowers the growth rate of both production-based emission inventories.[45] The time-dummies are different from zero at the selected CI in most cases. For carbon emissions per capita, they point towards a significant increase in emissions growth in 2001–2004 worldwide, followed by a slight decrease afterwards. For CO_2 per value added, by contrast, the results indicate a global decrease in emission intensities over time.

5.2 Conditional Convergence Without Individual-Dummies

The *DV-conditional model* analyzed above includes individual-specific effects and is thus concerned with convergence towards individual-specific steady states. A stronger assumption

[44] The negative effect of democracy on emissions growth is not robust to using alternative measures of democracy, such as the democracy measure sourced from the FSD1289 Measures of Democracy 1810–2014 database (see Finnish Social Science Data Archive 2018) or the average of the Freedom House indices of political rights and civil liberties (see Freedom House 2018). The main results are not sensitive to these alternative specifications and are also robust to the exclusion of the democracy variable.

[45] The negative impact of the construction sector may be related to the low carbon intensity of this sector during the period analyzed. We take the value added share of agriculture as the benchmark sector and exclude it from the specifications in order to avoid multicollinearity.

Table 2 Results scale-mixture of normals, *conditional heteroscedastic model*

	(1) CO_2 pc prod.	(2) CO_2 pc cons.	(3) CO_2 va prod.	(4) CO_2 va cons.
Outcome equation				
Constant	−0.3611***	−0.4721***	−0.0938**	−0.0112
Ln(emissions)	−0.0266***	−0.0460***	−0.0154***	−0.0158***
Ln(emissions)·EU	0.0001	0.0000	0.0003	0.0002
Ln(emissions)·OECD	0.0002	0.0002	−0.0002	−0.0002
Ln(emissions)·Annex B	−0.0002	−0.0003	0.0006	0.0002
Ln(income pc)	0.0336***	0.0525***	0.0105**	0.0011
Income pc growth	0.7263***	1.0017***	−0.1722	0.0095
Ln(pop. density)	−0.0033*	−0.0047***	−0.0034	−0.0040**
Fossil rents	−0.0006	−0.0009	0.0002	0.0005
Nuclear (%)	−0.0001	0.0001	0.0001	0.0001
Fossil fuels (%)	0.0002*	0.0002*	0.0001	0.0001
Openness	0.0000	0.0000	−0.0001	0.0000
Political regime	−0.0011**	−0.0008	−0.0008	−0.0004
VA energy (%)	0.0013*	0.0014**	0.0001	−0.0002
VA light manufacturing (%)	0.0004	−0.0005	0.0001	−0.0001
VA heavy manufacturing (%)	0.0005	0.0003	−0.0003	−0.0001
VA textiles (%)	0.0042***	0.0024**	0.0052***	0.0027**
VA water services (%)	−0.0020	−0.0046	0.0062	−0.0051
VA construction (%)	−0.0004	−0.0008	−0.0015	0.0002
VA trade and transport (%)	0.0008	0.0006	−0.0001	0.0001
VA other services (%)	0.0008	0.0004	−0.0003	−0.0001
2004	0.0335***	0.0305***	0.0140*	0.0037
2007	0.0063	0.0035	0.0042	−0.0079
2011	0.0115*	0.0083	0.0117	−0.0039
2014	−0.0017	−0.0089	−0.0013	−0.0148**
Individual-dummies	no	no	no	no
R^2	0.2757	0.4659	0.2554	0.3117
DIC	−3230.3	−3259.2	−3018.1	−3223.5
Instrumental equation for income pc growth				
Constant	0.0145***	0.0147***	0.0146***	0.0147***
Income pc growth, lagged	0.3620***	0.3594***	0.3603***	0.3589***
a_{iv}	−0.0714	0.0114	−0.2114	−0.3918
R^2	0.5314	0.5324	0.5319	0.5326
PIP EU	0.0029	0.0020	0.0075	0.0055
PIP OECD	0.0038	0.0095	0.0083	0.0061
PIP Annex B	0.0023	0.0056	0.0322	0.0050
Half-life	25.7100	14.7191	44.6621	43.5226
ν	3.8105	4.5663	4.2396	4.6691
N	390	390	390	390

*CI 90%, **CI 95%, ***CI 99%. All variables but group dummies and income pc growth enter in lagged values. The half-life is calculated as $-ln(0.5)/-ln(1+\beta)$ (see Allington and McCombie 2007). The Bayesian R^2 is the mean of the R^2 computed for each draw q of the Markov chain (MC), R^2_q where $R^2_q = \sum_{i,t}\hat{y}_{it}/\left(\sum_{i,t}\hat{y}_{it}+\sum_{i,t}\epsilon_{it}\right)$, where $\hat{y}_{it}$ is the estimate of y_{it} implied by the model and $\epsilon_{it} = y_{it}-\hat{y}_{it}$ (Gelman et al. 2017, see). The Deviance Information Criterion (DIC) is computed as $DIC = \hat{D_q} + Var(D_q)/2$, where D_q is the deviance measure associated with draw q in the MC (see Spiegelhalter et al. 2002; Gelman et al. 2004, Chap. 7) is that convergence occurs towards a common steady state that is determined by economic and political factors. In order to test for international convergence towards a common level of emissions per capita or per value added, we also estimate models without individual-specific effects. The results from the *conditional models* without individual-dummies, displayed in Table 2, show a slightly different pattern of convergence. The convergence parameters (of lagged emissions) are still relevant at the 99% CI for all inventories, but are substantially smaller in absolute values than in the *DV-conditional models*; they indicate a half-life of 26 (15) years for CO_2 per capita production (consumption) and of 45 (44) years for CO_2 production (consumption) intensities. Group-specific convergence patterns remain unimportant, with PIPs that are even lower than for the *DV-conditional models* (in most cases below 1%).

With regard to the control variables, some turn irrelevant for explaining emissions growth—fossil rents, the share of nuclear sources in electricity production, and value added shares of heavy manufacturing, water services and construction for all inventories, as well as political regime and the share of fossil fuels in electricity production for CO_2 intensities. Some others gain relevance, namely population density for consumption inventories, the value added share of the energy sector for per capita emissions, and the value added share of the textile sector for all inventories. Furthermore, income per capita is now relevant for all inventories but CO_2 consumption intensity.[46]

To sum up, our findings provide strong evidence for rather fast rates of convergence towards country-specific steady states for all four CO_2 inventories (*DV-conditional model*). International convergence towards global steady states determined by political and economic structures proceeds at a much slower pace (*conditional model*). Although some previous studies have found evidence for group-specific convergence patterns for OECD and EU members (e.g. Aldy 2006; Nguyen 2005; Ordás Criado and Grether 2011; Panopoulou and Pantelidis 2009; Westerlund and Basher 2008), none of our models provides evidence for differences in convergence dynamics implied by membership in the OECD, EU, or Annex B to the Kyoto Protocol. Therefore, climate change policies of industrialized countries such as the OECD or the EU have not been effective in accelerating emission convergence among developed economies (see also Westerlund and Basher 2008, who found slower convergence for OECD countries). Furthermore, the binding commitments of the Kyoto Protocol have been largely ineffective in accelerating emission convergence among Annex B countries (see also Ordás Criado and Grether 2011).[47]

6 Conclusion and Discussion

We tested for international convergence of CO_2 per capita and per value added derived from production and consumption patterns across a global sample of countries during 1997–2014. In so doing, we put forward a Bayesian test for convergence that is robust to cross-sectional heteroscedasticity, accounts for endogeneity between the growth rate of CO_2 emissions and economic growth, and allows for the existence of group-specific convergence among members of the EU, the OECD, and the Annex B to the Kyoto Protocol.

[46] The results of the models without individual-dummies could be affected by omitted variables and should be taken with care (see Barro and Sala-i-Martin 2004).

[47] Tables A.4 and A.5 in the Online Appendix report the results for the models with homoscedastic errors. The main results do not change qualitatively. The only important qualitative change is that income growth becomes significant for emissions per value added in the *DV-conditional models*. The convergence coefficient only changes slightly, though this change is amplified in the half-lives, decreasing them in the specifications for emissions per value added in the *conditional models*.

Our findings suggest that all four emission inventories converge towards country-specific steady states. The short half-lives calculated show that emissions per capita as well as emission intensities are close to their country-specific steady states. Production-based inventories show a shorter half-life (2.7–3.1 years) than consumption-based inventories (3.6–4.7 years). The finding that for production inventories CO_2 intensity converges faster than CO_2 per capita is in line with the results of Anjum et al. (2014). For CO_2 consumption inventories the reverse is true, what provides a first indication that, internationally, CO_2 intensities converge more slowly than CO_2 per capita.

In fact, a much slower pace of global CO_2 convergence across countries—towards global steady states that are determined by economic and political structures—is detected for all four emission inventories. Emissions per capita embodied in consumption show substantially faster global conditional convergence (with a half-life of 14.7 years) than emissions per capita embodied in production (25.7 years). This is consistent with converging consumption bundles across countries as a result of increasing globalization and the homogenization of consumer tastes, while the slower cross-country convergence of production-based emissions can be related to the long-run nature of structural transformations of production patterns. Emission intensities converge towards global steady states at an even slower pace, which is similar for production and consumption inventories (implying half-lives of 44–45 years). The very slow pace of global convergence of emission intensities indicates that technology transfers across countries have been rather limited and cannot explain the faster rate of convergence of emissions per capita. Thus, the faster convergence of emissions per capita is likely to be driven by some of the conditioning variables, possibly income growth. Various empirical studies have shown that income and CO_2 emissions are positively related (e.g. Azomahou et al. 2006; Holtz-Eakin and Selden 1995; Fernández-Amador et al. 2017). To the extent that the economic crisis of 2008 has led to a decline in economic growth in high-income countries, whereas countries with lower income levels were able to catch up, this catch-up process could have sped up the international convergence of per capita emissions (see also Brock and Taylor 2010).[48]

Actual levels of carbon emissions have proven to be unsustainable. Our results indicate that higher income growth is related to a higher growth rate of emissions per capita, while it is not connected to a decrease in emission intensities. The evidence for convergence does not automatically imply convergence towards steady states that are sustainable, as highlighted by Brock and Taylor (2010) and Ordás Criado et al. (2011). Moreover, the evidence for fast country-specific convergence towards not sustainable steady states underlines the current incompatibility between economic growth and the 2 °C target, and the need for further abatement and mitigation policies to keep global warming under control while maintaining reasonable economic growth rates.

The historical responsibility for atmospheric CO_2 concentrations corresponds to developed economies. However, those economies, represented in our sample by three groups—OECD, EU, and the countries that ratified the Annex B to the Kyoto Protocol—have not experienced faster group convergence. This lack of specific patterns of convergence among developed economies, despite the environmental policies implemented in these countries during the period of analysis after the Kyoto Protocol, shows the difficulties in achieving effective agreements and policies to take action against global warming.

The slow pace of international convergence of emissions and the limited extent of international technology diffusion pose doubts on the feasibility of the agreed sustainability targets, unless a significant change in the international institutional framework takes place, and new,

48 A more detailed analysis of this mechanism is out of the scope of this paper.

stronger abatement policies are implemented. Also, much faster transfers of green technologies will be necessary.

The lack of a stabilization of emissions in industrialized economies at sustainable emission levels may discourage developing economies to accept a cap on emissions. In addition, the evidence found for country-specific steady states in emissions points to significant transaction costs connected to the design of multilateral policy frameworks aimed at global emission reduction. Even though there is an urgency for multilateral approaches to fight climate change that encompass developed and developing countries, developed economies should foster further national environmental policies to promote carbon efficiency and less polluting sources of energy in order to reinforce the international action against global warming.

Acknowledgements The authors thank Carlos Ordás Criado for valuable insights and the participants of the SWSR for inspiring discussions. The authors acknowledge support of the NRP 73 project Switzerland's Sustainability Footprint: Economic and Legal Challenges, grant No. 407340-172437, University of Bern, supported by the Swiss National Science Foundation (SNSF) within the framework of the National Research Programme "Sustainable Economy: resource-friendly, future-oriented, innovative" (NRP 73).

References

Aichele R, Felbermayr G (2015) Kyoto and carbon leakage: an empirical analysis of the carbon content of bilateral trade. Rev Econ Stat 97(1):104–115

Aldy JE (2006) Per capita carbon dioxide emissions: convergence or divergence? Environ Resour Econ 33:533–555

Aldy JE (2007) Divergence in state-level per capita carbon dioxide emissions. Land Econ 83:353–369

Allington NF, McCombie JS (2007) Economic growth: new directions in theory and policy. Edward Elgar Publishing Limited, Ch. Economic growth and beta-convergence in the East European Transition Economies, pp 200–222

Andrews D, Mallows C (1974) Scale mixtures of normal distributions. J R Stat Soc Ser B 36:99–102

Anjum Z, Burke P, Gerlagh R, Stern D (2014) Modeling the emissions-income relationship using long-run growth rates. CCEP working paper 1403

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58(2):277–297

Aslanidis A, Iranzo S (2009) Environment and development: is there a Kuznets curve for CO_2 emissions? Appl Econ 41(6):803–810

Azomahou T, Laisney F, Van PN (2006) Economic development and CO_2 emissions: a nonparametric panel approach. J Public Econ 90(6):1347–1363

Babiker MH (2005) Climate change policy, market structure, and carbon leakage. J Int Econ 65:421–445

Barassi M, Cole MA, Elliott RJ (2008) Stochastic divergence or convergence of per capita carbon dioxide emissions: re-examining the evidence. Environ Resour Econ 40:121–137

Barassi MR, Cole MA, Elliott RJ (2011) The stochastic convergence of CO_2 emissions: a long memory approach. Environ Resour Econ 49:367–385

Barro R (1991) Economic growth in a cross section of countries. Q J Econ 106(2):407–443

Barro J, Sala-i-Martin X (1992) Convergence. J Polit Econ 100(2):223–251

Barro R, Sala-i-Martin X (2004) Economic growth, 2nd edn. The MIT Press, Cambridge

Baumol W (1986) Productivity growth, convergence, and welfare: what the long-run data show. Am Econ Rev 76(5):1072–1085

Bernard A, Durlauf S (1996) Interpreting tests of the convergence hypothesis. J Econom 71(1–2):161–173

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

Brock W, Taylor M (2010) The green Solow model. J Econ Growth 15:127–153

Bruno GS (2005) Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. Econ Lett 87(3):361–366

Burnett J (2016) Club convergence and clustering of US energy-related CO_2 emissions. Resour Energy Econ 46:62–84

Camarero M, Castillo J, Picazo AJ, Tamarit C (2013) Eco-efficiency and convergence in OECD countries. Environ Resour Econ 55:87–106

Carlino G, Mills L (1993) Are US regional incomes converging? a time series analysis. J Monet Econ 32(2):335–346

Dasgupta S, Laplante B, Wang H, Wheeler D (2002) Confronting the environmental Kuznets curve. J Econ Perspect 16:147–168

De Vos I, Everaert G, Ruyssen I (2015) Bootstrap-based bias correction and inference for dynamic panels with fixed effects. Stata J 15(4):986–1018

Ding P (2016) On the conditional distribution of the multivariate t distribution. Am Stat 70:293–295

Efron B (1982) The Jacknife, the bootstrap and other resampling plans. Vol. 38 of CBMS-NSF regional conference series in applied mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia

Efron B (2011) The bootstrap and Markov Chain Monte Carlo. J Biopharm Stat 21:152–162

Evans P, Karras G (1996) Convergence revisited. J Monet Econ 37(2):249–265

Everaert G, Pozzi L (2007) Bootstrap-based bias correction for dynamic panel. J Econ Dyn Control 31:1160–1184

Ezcurra R (2007) Is there cross-country convergence in carbon dioxide emissions? Energy Policy 35:1363–1372

Farzin Y, Bond C (2006) Democracy and environmental quality. J Dev Econ 81:213–235

Fernández-Amador O, Oberdabernig D (2018) Bayesian estimation of dynamic panels. Mimeo

Fernández-Amador O, Francois JF, Tomberger P (2016) Carbon dioxide emissions and international trade at the turn of the millennium. Ecolog Econ 125:14–26

Fernández-Amador O, Francois JF, Oberdabernig DA, Tomberger P (2017) Carbon dioxide emissions and economic growth: an assessment based on production and consumption emission inventories. Ecolog Econ 135:269–279

Finnish Social Science Data Archive (2018) Measures of democracy 1810–2018

Freedom House (2018) Country and territory ratings and statuses, 1973–2018

Friedman M (1992) Do old fallacies ever die? J Econ Lit 30(4):2129–2132

Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis. Texts in statistical science, 2nd edn. CRC Press, Boca Raton

Gelman A, Goodrich B, Gabry J, Ali I (2017) R-squared for Bayesian regression models. Unpublished manuscript, Department of Statistics, Columbia University

George EI, McCulloch RE (1993) Variable selection via Gibbs sampling. J Am Stat Assoc 88:881–889

Geweke J (1993) Bayesian treatment of the independent Student-t linear model. J Appl Econom 8:S19–S40

Grossman G, Helpman E (1995) Handbook of international economics. North-Holland, Amsterdam (Chap. 25: Technology and trade)

Hao Y, Liao H, Wei Y-M (2015) Is China's carbon reduction target allocation reasonable? An analysis based on carbon intensity convergence. Appl Energy 142:229–239

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Series in statistics, 2nd edn. Springer, Berlin

Herrerias M (2013) The environmental convergence hypothesis: carbon dioxide emissions according to the source of energy. Energy Policy 61:1140–1150

Holtz-Eakin D, Selden TM (1995) Stoking the fires? CO_2 emissions and economic growth. J Public Econ 1995:85–101

Hsiao C, Pesaran H, Tahmiscioglu K (1999) Analysis of panels and limited dependent variable models. Cambridge University Press, Cambridge (Chap. 11: Bayes estimation of short-run coefficients in dynamic panel data models, pp 268–296)

Huang B, Meng L (2013) Convergence of per capita carbon dioxide emissions in urban China: a spatio-tomporal perspective. Appl Geogr 40:21–29

Jobert T, Karanfil F, Tykhonenko A (2010) Convergence of per capita carbon dioxide emissions in the EU: legend or reality? Energy Econ 32:1364–1373

Kaika D, Zervas E (2010) The environmental kuznets curve (EKC) theory—part A: concept, causes and the CO_2 emissions case. Energy Policy 62:1392–1402

Kiviet JF (1995) On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. J Econom 68(1):53–78

Kleibergen F, Zivot E (2003) Bayesian and classical approaches to instrumental variable regression. J Econom 114:29–72

Knutti R, Rogelj J, Sedlace J, Fischer E (2015) A scientific critique of the two-degree climate change target. Nat Geosci 9:13–19

Koop G (2003) Bayesian econometrics. Wiley, London

Lancaster T (2000) The incidental parameter problem since 1948. J Econom 95:391–413

Lancaster T (2008) An introduction to modern Bayesian econometrics. Blackwell, Oxford

Lee C-C, Chang C-P (2008) New evidence on the convergence of per capita carbon dioxide emissions from panel seemingly unrelated regressions augmented Dickey–Fuller tests. Energy 33:1468–1475
Lee C-C, Chang C-P (2009) Stochastic convergence of per capita carbon dioxide emissions and multiple structural breaks in OECD countries. Econ Model 26:1375–1381
Lee C-C, Chang C-P, Chen P-F (2008) Do CO_2 emission levels converge among 21 OECD countries? new evidence from unit root structural break tests. Appl Econ Lett 15:551–556
Lopes H, Polson N (2014) Bayesian instrumental variables: priors and likelihoods. Econom Rev 33:100–121
Maddala G, Hu W (1996) The econometrics of panel data: a handbook of theory with applications, 2nd edn. Kluwer, Boston (Chap. 13: The pooling problem, pp 307–322)
Mankiw N, Romer D, Weil D (1992) A contribution to the empirics of economic growth. Q J Econ 107(2):407–437
Moon H, Perron B, Phillips P (2015) The Oxford Handbook of Panel Data. Oxford University Press, Oxford, pp 111–148 (Chap. 4: Incidental parameters and dynamic panel modeling)
Newton M, Raftery A (1994) Approximate Bayesian inference with the weighted likelihood bootstrap. J R Stat Assoc Ser B 56:3–48 (With discussion and reply by the authors)
Neyman J, Scott E (1948) Consistent estimation from partially consistent observations. Econometrica 16:1–32
Nguyen PV (2005) Distribution dynamics of CO_2 emissions. Environ Resour Econ 32:495–508
Nickell S (1981) Biases in dynamic models with fixed effects. Econometrica 49:1417–1426
Nourry M (2009) Re-examining the empirical evidence for stochastic convergence of two air pollutants with a pair-wise approach. Environ Resour Econ 44:555–570
Ordás Criado C, Grether J-M (2011) Convergence in per capita CO_2 emissions: a robust distributional approach. Resour Energy Econ 33:637–665
Ordás Criado C, Valente S, Stengos T (2011) Growth and pollution convergence: theroy and evidence. J Environ Econ Manag 62:199–214
Panopoulou E, Pantelidis T (2009) Club convergence in carbon dioxide emissions. Environ Resour Econ 44:47–70
Pettersson F, Maddison D, Acar S, Söderholm P (2014) Convergence of carbon dioxide emissions: a review of the literature. Int Rev Environ Resour Econ 7:141–178
Phillips P, Sul D (2007a) Some empirics on economic growth under heterogenous technology. J Macroecon 29:455–469
Phillips P, Sul D (2007b) Transition modelling and econometric convergence tests. Econometrica 75(6):1771–1855
Quah D (1993) Empirical cross-section dynamics in economic growth. Eur Econ Rev 37(2–3):426–434
Quah D (1996) Empirics for economic growth and convergence. Eur Econ Rev 40:1353–1375
Ravallion M (2003) Inequality convergence. Econ Lett 80:351–356
Romero-Ávila D (2008) Convergence in carbon dioxide emissions among industrialised countries revisited. Energy Econ 30:2265–2282
Rossi PE, Greg A, McCulloch R (2005) Bayesian statistics and marketing. Wiley series in probability and statistics. Wiley, Chichester
Rubin D (1981) The Bayesian bootstrap. Ann Stat 9:130–134
Sala-i-Martin X (1996) Regional cohesion: evidence and theories of regional growth and convergence. Eur Econ Rev 40(6):1325–1352
Salois M, Balcombe K (2015) A generalized Bayesian instrumental variable approach under student t-distributed errors with application. Manch Sch 83:499–522
Schmalensee R, Stoker T, Judson R (1998) World carbon dioxide emissions: 1950–2050. Rev Econ Stat 80(1):15–27
Solow R (1956) A contribution to the theory of economic growth. Q J Econ 70(1):65–94
Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc Ser B 64:583–639 (with discussion)
Stern D (2004) The rise and fall of the environmental Kuznets curve. World Dev 32:1419–1439
Stern D (2017) The environmental Kuznets curve after 25 years. J Bioecon 19:1–22
Strazicich MC, List JA (2003) Are CO_2 emission levels converging among industrial countries? Environ Resour Econ 24:263–271
Wang Y, Zhang P, Huang D, Cai C (2014) Convergence behavior of carbon dioxide emissions in China. Econ Model 43:75–80
West M (1987) On scale mixtures of normal distributions. Biometrika 74:646–648
Westerlund J, Basher SA (2008) Testing for convergence in carbon dioxide emissions using a century of panel data. Environ Resour Econ 40:109–120
Wu J, Wu Y, Guo X, Cheong TS (2016) Convergence of carbon dioxide emissions in Chinese cities: a continuous dynamic distribution approach. Energy Policy 91:207–219

Yavuz N, Yilanci V (2013) Convergence in per capita carbon dioxide emissions among G7 countries: a TAR panel root approach. Environ Resour Econ 54:283–291
Young A, Higgins M, Levy D (2008) Sigma convergence vs beta convergence: evidence from US country-level data. J Money Credit Bank 40(5):1083–1093
Zhao X, Burnett W, Lacombe D (2015) Province-level convergence of China's carbon dioxide emissions. Appl Energy 150:286–295