Faculteit Bedrijfseconomische Wetenschappen

master handelsingenieur in de beleidsinformatica

Masterthesis

Circular Economy and Resource Extraction: A Replication, Robustness Analysis, and Re-estimation

Hans Dekens

Scriptie ingediend tot het behalen van de graad van master handelsingenieur in de beleidsinformatica

PROMOTOR :

Prof. dr. Stephan BRUNS

BEGELEIDER :

De heer Tuan NGUYEN

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

2024

2025

Faculteit Bedrijfseconomische Wetenschappen

master handelsingenieur in de beleidsinformatica

Masterthesis

Circular Economy and Resource Extraction: A Replication, Robustness Analysis, and Re-estimation

Hans Dekens

Scriptie ingediend tot het behalen van de graad van master handelsingenieur in de beleidsinformatica

PROMOTOR :

Prof. dr. Stephan BRUNS

BEGELEIDER :

De heer Tuan NGUYEN

Circular Economy and Resource Extraction: A Replication, Robustness Analysis, and Re-estimation

Hans Dekens

Hasselt University, Faculty of Business Economics, Agoralaan building D, 3590 Diepenbeek, Limburg, Belgium

Abstract. This study revisits and reassesses the empirical findings of Bianchi and Cordella (2023) on the relationship between circular economy (CE) initiatives and domestic resource extraction in Europe. While their results suggest that CE practices minimally reduce domestic virgin resource extraction, a comprehensive four-part analysis involving replication, robustness testing, and two re-estimations using 2SLS with Lewbel residual-based instruments shows that CE initiatives have no impact on domestic resource extraction. Their incorrect finding resulted from methodological flaws in using the System GMM estimator, which undermined the credibility of their findings. In particular, instrument proliferation and questionable exogeneity assumptions played a key role. However, when substituting domestic extraction with raw material consumption, the circular material use rate (a key CE indicator) shows a significant negative relationship. This suggests that the circular material use rate is linked to increased resource exports or reduced resource imports rather than decreased domestic extraction. Additionally, results indicate that GDP has a more substantial effect on domestic resource extraction. These findings challenge the original causal claims but confirm the broader policy conclusion: CE initiatives must ensure structural changes in consumption to minimize resource use meaningfully. This study underscores the importance of replication in exposing methodological weaknesses and advances empirical practice by introducing 2SLS with Lewbel (2012) instruments as an alternative in panel settings characterized by a small number of entities and multiple difficult-to-instrument endogenous variables.

Keywords: replication · robustness analysis · circular economy · sustainable development · circularity indicators · System GMM · instrument proliferation · 2SLS · Lewbel instruments

1 Introduction

Each year, more than 100 billion tonnes of resources are extracted worldwide, with over 90% of these extracted resources ultimately becoming waste [1,2]. To address these and other environmental challenges linked to ongoing resource extraction, circular economy (CE) has emerged as a crucial approach, aiming to decouple economic growth from environmental degradation through the recycling and reuse of materials [3,4,5,6]. However, the extent to which these CE practices can genuinely mitigate resource extraction is still unclear, mainly when economic growth drives increased demand for primary resources [3,7,8].

A recent study by Bianchi and Cordella (2023) [3] provided valuable empirical insights into this matter. Their research examined the relationship between CE initiatives and domestic resource extraction, using panel data from 28 European countries between 2010 and 2019. The results demonstrated that while CE practices contribute to reducing domestic extraction, their effect remains modest compared to the overwhelming influence of economic growth. Specifically, the study reported that for every 1% increase in employment in CE sectors and circular material use rate, the average domestic extraction of resources decreased by 0.34% and 0.11%, respectively. However, the absolute reduction resulting from the

CE policy was offset by a four times greater increase in resource demands caused by the gross domestic product (GDP) growth. This finding suggests that current CE practices are insufficient to counteract economic expansion's environmental pressures [3].

However, their study [3] raises concerns from a credibility and robustness standpoint. Its empirical insights are derived using a System Generalized Method of Moments (GMM) estimation, which reports a Hansen-Sargan test with a p-value of exactly 1.00. While a high Hansen-Sargan p-value is favorable, a value of precisely 1.00 may indicate problems such as instrument proliferation. This occurs when too many instruments are employed, which can overfit the endogenous variables, ultimately undermining the reliability of the empirical results [9]. Moreover, given that the study has quickly become prominent and highly cited since its publication in 2023, the potential impact of unreliable findings is particularly significant. This concern is further amplified as the research is particularly relevant for the European Union, which is transitioning to a circular economy for sustainable growth, and its policymakers, who rely on scientific evidence to inform policy decisions [10,11].

A four-part analysis addresses these concerns and ensures credible and robust results, following the framework outlined by Clemens (2017) [12]. Component 1 involves a full replication of the study, specifically, a verification using the exact data and specifications from the original research to confirm the validity of the reported findings. Component 2, referred to as reanalysis, focuses on conducting robustness tests of the original System GMM. This involves modifying the original model specifications, particularly by addressing instrument proliferation through instrument reduction techniques, while using the original and extended data to account for potential differences in the data-generating process. Component 3 involves a re-estimation of the model using Two-Stage Least Squares (2SLS) with the extended dataset and the instruments derived from the residual-based approach proposed by Lewbel (2012) [13]. This specification retains domestic extraction as the dependent variable, aligning with the original study and component 2. Component 4 builds on component 3 by re-estimating the model with 2SLS, replacing domestic extraction with raw material consumption as the dependent variable. This adjustment considers international resource flows embedded in trade through imports and exports, which can affect the results.

The results of this study challenge the claims made by Bianchi and Cordella (2023) about the role of the circular economy in reducing European domestic resource extraction. While the replication largely verifies their empirical findings, it also uncovers significant methodological issues, most notably the expected instrument proliferation and unjustified exogeneity assumptions, which cast doubt on the reliability of their findings. The reanalysis using a more theoretically sound System GMM approach, designed to correct the identified issues, finds that CE policies do not impact domestic resource extraction. The 2SLS re-estimation confirms this null relationship while relying on a more valid set of instruments. Finally, when raw material consumption is used as the outcome variable in the re-estimation, the results indicate that CE initiatives affect resource trade more than domestic extraction, primarily by decreasing imports or boosting exports. As a result, these initiatives enhance strategic autonomy by lowering net resource imports and thus reducing reliance on imported resources. Additionally, the two re-estimations with 2SLS highlight a stronger influence from GDP than estimated in the original study. Ultimately, these findings go beyond the original study's conclusions by empirically confirming its concern: CE practices such as recycling and reuse do not reduce the European domestic extraction of primary resources. The more pronounced impact of GDP in this study further underlines the importance of tackling the underlying drivers of material use that remain unaffected by CE practices, including demand-side pressures linked to economic growth. Therefore, rather than solely optimizing waste stream management, policy should aim to reshape consumption patterns to reduce overall resource demand [3].

This study contributes to the existing literature by providing a more nuanced understanding of the relationship between CE practices and raw material extraction, incorporating both imports and exports of resources. This broader approach, which distinguishes it from prior research, offers a more comprehensive perspective on the environmental impacts of CE initiatives. Additionally, this study strengthens the credibility of economic research through rigorous replication and robustness checks of the original study's findings, reinforcing the importance of such practices in overcoming methodological limitations and ensuring valid results. From a methodological standpoint, this study introduces 2SLS with Lewbel residual-based instruments as a practical approach for addressing endogeneity in panels with a low number of panel entities and complex instrument identification. This approach resolves specific challenges in panel data analysis and offers a replicable framework for future studies facing similar empirical constraints.

The remainder of the paper is organized as follows: Section 2 provides a detailed explanation of key concepts central to the study. Sections 3 and 4 contain the methodology and results relevant to their respective analyses. Specifically, Section 3 reports the replication and robustness analysis, while Section 4 describes the re-estimation of the study. The replication, robustness analysis, and re-estimation results, along with their limitations, are then discussed. Finally, Section 6 provides a short study conclusion.

2 Background

This section introduces several key concepts necessary for understanding the rest of the paper. Section 2.1 provides a detailed explanation of the circular economy (CE) concept. Section 2.2 presents the European CE monitoring framework, which forms the basis of the data used in the statistical analyses. Section 2.3 explores the issue of burden shifting, a critical challenge in CE assessment. These concepts frame the original study's goals, which are subsequently revisited. The section concludes by outlining the rationale for conducting the replication and robustness analysis and introduces the classification framework proposed by Clemens (2017), which guides the structure of the current study.

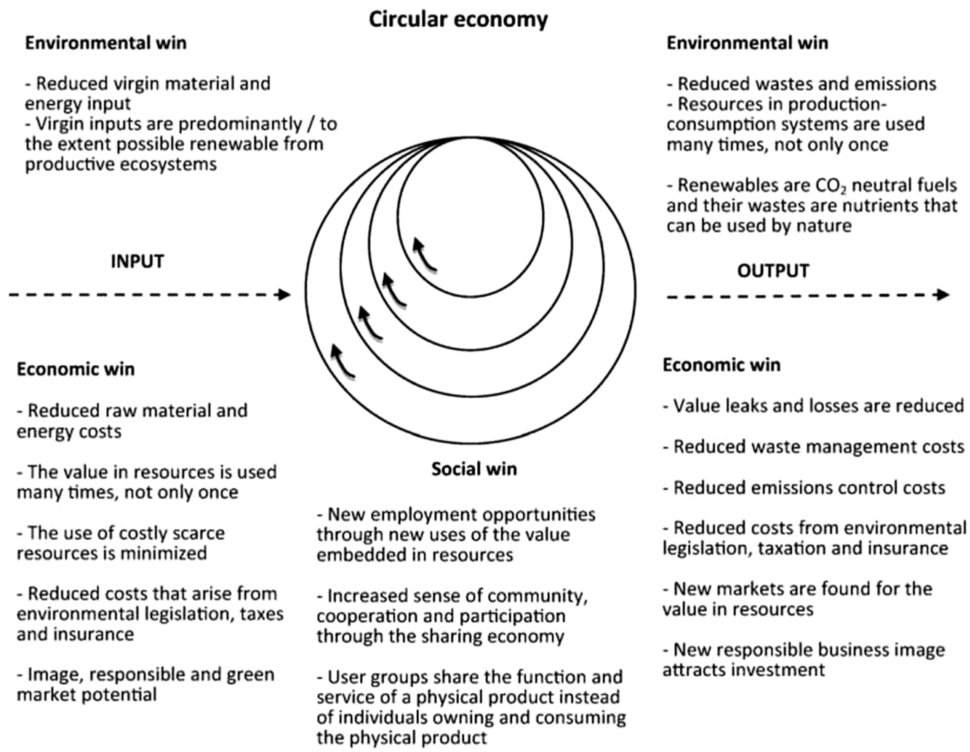
2.1 CE concept

The concept of CE aligns closely with the principles of sustainable development [6]. A critical challenge in sustainability is the current linear resource flow model between the environment and the human economy. This approach depletes resources and generates unsustainable waste, undermining the systems it depends upon [3,5,6]. As the economy grows, the physical capacity of the planet to support human activity diminishes. This is evidenced by expanding deserts, rising sea levels, increasing population, higher per capita consumption, growing livestock production, and accelerating biodiversity loss. These trends represent a looming collision between economic expansion and the planet's shrinking ability to sustain life [6].

A circular economy offers an alternative by replacing this linear flow with a cyclical model. It emphasizes reusing, remanufacturing, refurbishing, repairing, repurposing, and upgrading products and materials. The approach also prioritizes renewable energy sources like solar, wind, biomass, and waste-derived energy across the product lifecycle [3,5,6]. The core idea is straightforward and economically sound: once resources are extracted and converted into products or services of value, they should be utilized repeatedly rather than discarded after a single use. This reduces the need for virgin resource inputs while minimizing waste and emissions [6].

In addition to addressing environmental concerns, the CE concept aims to deliver significant economic and social benefits, creating value and promoting sustainability across multiple dimensions [6]. These benefits are illustrated comprehensively in Figure 1.

Figure 1: CE advantages in environmental, economic, and social dimensions [6]

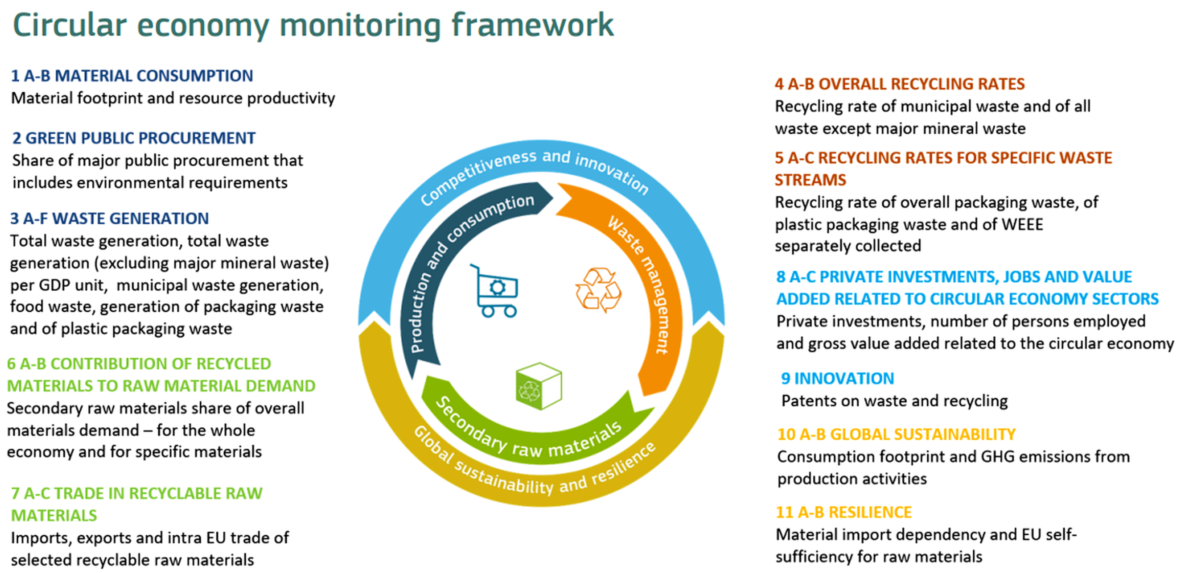


2.2 CE monitoring framework

The European Union strives to become a circular economy to strengthen environmental sustainability and increase competitiveness. To drive this transformation, the European Commission introduced a new CE action plan in 2020 [10,14]. This plan is central to the European Green Deal, outlining Europe's new sustainable growth agenda. In 2018, the European Commission established a monitoring framework (Figure 2) to track progress towards a circular economy. The framework offers a comprehensive approach, measuring the direct and indirect benefits of transitioning to a circular economy. Moreover, it evaluates how CE practices contribute to sustainable living within the planet's boundaries while addressing energy and material supply risks. The framework is structured into five thematic areas, each represented by a different color in Figure 2. It includes 11 key statistical indicators, with some having additional sub-indicators. These indicators were selected to capture the essential elements of a circular economy [14]. This is evident as they align with the key advantages of the circular economy, as depicted in Figure 1.

Most indicators in the monitoring framework are derived from official statistics provided by Eurostat. They have been selected based on strict criteria, ensuring they are relevant, accepted, credible, easy

Figure 2: CE monitoring framework [15]



to use, and robust. They are also designed to make optimal use of existing data whenever available. The framework heavily relies on high-quality statistics available across all EU Member States, drawing primarily from the European Statistical System and contributions from the research community. The new monitoring framework will be managed by Eurostat on its website, with regular updates to ensure the indicators reflect the latest data [16].

In May 2023, the European Commission adopted an updated CE monitoring framework and a detailed staff working document explaining the new indicators [14,15,16]. This revised framework introduced a new focus on global sustainability and resilience. It includes new indicators such as material footprint, resource productivity, consumer footprint, greenhouse gas emissions from manufacturing, and material dependency, providing a more comprehensive understanding of the circular economy's impact on sustainability [16].

2.3 Burden shifting

One key challenge in assessing circularity is the potential risk of burden shifting or problem shifting [3,4,6]. Burden shifting happens when consumption and production are geographically separated, transferring consumption-driven impacts to the countries where production occurs [17]. The environmental impact is thus reduced in one part of the system by shifting the problem to another part of the system [6]. This phenomenon is commonly observed between developed and developing nations [17].

As noted in a 2019 study [4], many metrics for assessing circularity fail to account for burden shifting entirely. One such metric is the original research's dependent variable, domestic extraction (DE). DE serves inherently as a circularity metric as it is a key objective of CE, namely the reduced extraction of virgin materials, as illustrated in Figure 1. Since DE measures progress in only a subset of the CE's defined goals and employs aggregated metrics, it is classified as an assessment indicator [4,6]. Therefore, the possibility of burden shifting exists for two main reasons:

1. DE may decrease in one country at the expense of additional resource extraction in other countries, thereby shifting the burden geographically [3,4].
2. Even though DE may decrease, other CE goals, such as lower emission levels or higher employment, can be negatively affected [4,6].

2.4 Original study

The original study [3] investigates whether CE initiatives reduce domestic extraction of natural resources in the European Union. Domestic extraction was chosen as the study's dependent variable because it explicitly measures the extraction of primary resources within a country. On the other hand, the frequently used indicator domestic material consumption was deemed unsuitable as it combines primary and secondary materials, including those derived from imports and exports, making it less precise to evaluate the extraction of primary resources [3].

The study further incorporates other CE indicators from the EU's monitoring framework (Figure 2) as explanatory variables. The study finds that some CE indicators, such as the circular material use rate and CE-related employment, which involve recycling and reusing materials, are associated with reductions in DE. However, their impact is substantially smaller than the increase in DE driven by economic growth. This highlights the limited capacity of current CE initiatives to offset resource pressures from expanding consumption and suggests a need for systemic, behavioral, and policy changes.

Lastly, while the original study acknowledges the potential for burden shifting, it does not address it empirically. Specifically, it highlights the first type of burden shifting discussed earlier. Namely, reductions in DE within one country may coincide with increased resource extraction elsewhere due to excluding imports and exports from the DE metric [3]. The second type is not addressed in the original study [3], as its scope is limited to evaluating the effect of CE policies on domestic extraction rather than encompassing broader CE goals.

2.5 Why replicating and robustness testing is needed

Building on the findings of the original study [3], their research holds value beyond theoretical insights, offering guidance for policymakers. As with many scientific studies, its impact on the effectiveness of evidence-based policymaking depends on the credibility and robustness of its results [18,19].

Replication studies are essential in this regard. They establish credibility and strengthen confidence in the reliability and integrity of findings by reproducing the original model specifications and using the same data population [19,20,21]. Nevertheless, economic research often faces significant challenges regarding replicability. As highlighted by a 2015 study, only 49% of 67 papers from well-regarded economics journals were replicable [22]. Moreover, research in 2016 found that only two-thirds of 18 economic research papers were replicable [23]. Interestingly, studies published in economics, psychology, and general science journals that fail to replicate are cited more frequently than successfully replicated studies [24]. Those nonreplicable studies present a range of concerns, from minor measurement errors to more severe problems like fraud and ethical considerations in scientific practices [12,25,26]. Such issues arise from human errors or publication bias, driven by the 'publish or perish' culture in academia, where researchers face significant pressure to report statistically significant results to make their work more appealing to editors and readers [18,27]. In sum, replication studies help uncover errors and publication bias by preserving the original sampling distributions and model specifications, thereby strengthening the credibility of scientific research [12,19,20,21].

To ensure robust findings, researchers can conduct robustness analyses, which test how sensitive the original results are to changes in model specifications or data populations. A failed robustness test reveals that the results depend on specific choices and may only generalize across some plausible methods. This does not inherently suggest an error or unethical practice by the original researchers, as methodological decisions are often subjective and open to legitimate debate [12]. Still, robust results are essential for effective evidence-based policymaking [18,19].

This broader context of replications and robustness analyses further highlights the significance of replicating and conducting robustness tests on the original findings, especially considering the concern of instrument proliferation and the prominence of the paper.

2.6 Clemens (2017) replication classification

It is helpful to draw on structured classifications to understand the conceptual distinctions of a replication and robustness analysis. One such framework, proposed by Clemens (2017), begins by distinguishing between replications and robustness analyses and then categorizes replication studies into four types based on data population, sample, and model specifications [12].

Replications are analyses that preserve the original sampling distributions and model specifications. This way, any significant shift in the parameter estimates does not arise from analyzing new populations or applying alternative model specifications. Methods involving such variations fall under robustness analyses rather than strict replications. Replication studies can be either verifications or reproductions. Verifications are performed using the exact data and specifications from the original research, thereby confirming the accuracy of the reported findings. Reproductions, however, use a new sample from the same population while applying the original model specifications to this new dataset. Verifications are often suited to observational research where only one sample may be available, while reproductions are more common in experimental research [12,28].

Compared to replications, robustness analyses can also be divided into two types. First, a reanalysis involves using different model specifications on data from the original population. Second, an extension uses the original specifications on a new population and sample. In time-series research, an extension may involve lengthening the sample period for the same time series, accounting for possible differences in the data-generating process [12,28].

This study conducts a verification replication (component 1) using the original dataset in line with these definitions. The robustness analysis (component 2) entails a reanalysis that includes adjustments to the original System GMM applied to the original data. In addition, the reanalysis will incorporate updated and extended data, which, while typically associated with an extension, remains part of the reanalysis in this study due to the changes in model specification. The following section provides a detailed discussion of these components.

3 Replication & robustness analysis

This section begins with an overview of the underlying methods used in the replication and robustness analysis. Afterward, the verification process is explained with an overview of the methodology[1] employed in the original study [3] and the reanalysis. Lastly, the results are presented.

[1] For a comprehensive account of the methodology used in the original research, see [3].

3.1 Methods

First, the IPAT and STIRPAT frameworks will be explained, as the original study uses these to analyze the circular economies' impact on resource extraction. Secondly, dynamic panel models and System GMM will be discussed shortly as it is a central component of the study, verification, and reanalysis[2]. Finally, Section 3.1.3 discusses two instrument reduction methods.

3.1.1 IPAT & STIRPAT

The original study [3] utilizes the STIRPAT model, an econometric extension of the IPAT (Impact = Population × Affluence × Technology) framework, which describes how human activities drive ecological impact (I). First introduced in a study [29] from 1972, IPAT highlights the multiplicative relationship between population (P), affluence (A), and technology (T). The formula in Eq. (1) underscores the need to address all three factors simultaneously to mitigate human impact on ecosystems. This necessity arises from the multiplicative specification of IPAT and highlights that changes in one driving force are scaled by the others, emphasizing their interdependence. Consequently, no single factor can be solely blamed for environmental impacts [29,30].

$$I = P \times A \times T \quad (\text{Impact} = \text{Population} \times \text{Affluence} \times \text{Technology}) \tag{1}$$

Where, according to [30], the variables are often defined as follows:

$$P = \text{Total population}, \quad A = \frac{\text{GDP}}{\text{Capita}}, \quad T = \frac{I}{\text{GDP}}$$

In 1994, a reformulation of the model to stochastic form was proposed in [31], making it suitable for empirical testing with conventional statistical methods. This reformulation, later known as STIRPAT (Stochastic Impacts by Regression Population, Affluence, and Technology), is presented in Eq. (2) [30].

$$I_i = aP_i^b A_i^c T_i^d e_i \tag{2}$$

The model is scaled by the constant a, with b, c, and d as the estimable exponents of P, A, and T, respectively. The error term is represented by e. By adding the residual (error) term, the model becomes stochastic. The stochastic version preserves the original model as the original model is simply a case in which a = b = c = d = e = 1. The subscript i indicates that I, P, A, T, and e may vary across different observations (units, periods, or both) [30,31].

$$log(I_i) = a + b \cdot log(P_i) + c \cdot log(A_i) + e_i \tag{3}$$

In 2003, an additive regression model in logarithmic form (Eq. (3)) was suggested [30]. The new model simplified estimation and hypothesis testing using traditional statistical techniques, such as regression analysis, with cross-sectional, time-series, or panel data. In this log-transformed version, the coefficients represent elasticities, indicating the percentage change in environmental impact (I) in response to a 1% change in either population (P) or affluence (A) while holding other factors constant [30].

[2] The discussion here focuses on the intuition and application of System GMM within the context of this study. For a more detailed explanation of the underlying assumptions and technical derivations, please refer to the sources cited in their respective section.

The technology variable (T) is not a single factor but a composite of multiple influences on environmental impacts. There are three approaches to handling T in the STIRPAT model. First, T can be treated as a residual term, encompassing all factors that influence environmental impact beyond population (P) and affluence (A) [31,30]. Second, T can be disaggregated by introducing additional variables into the model hypothesized to affect environmental impact per production unit. In this case, the residual term represents T after accounting for the effects of the newly added variables. Third, T can be analyzed using non-linear specifications and elasticity measures (b and c) to explore the interactions between T, population, and affluence. This method helps assess whether changes in population or affluence affect environmental impacts proportionally or disproportionately, providing insights into how T may amplify or moderate these effects [30].

The log-transformed version of STIRPAT, as shown in Eq. (3), allows for hypothesis testing and exploring the non-linear effects of the driving factors on environmental impacts. Furthermore, the model can incorporate a wide range of relevant variables. By evaluating impacts in terms of percentage changes (elasticities), the model provides valuable insights into how population and affluence influence environmental outcomes. This makes it an effective tool for evaluating sustainability and informing policy decisions [30]. Because of these advantages, it is well-suited for the original study [3].

3.1.2 Dynamic panel models & System GMM

Dynamic panel models are commonly used to study processes where the current outcome depends on past realizations [32]. A general form of a dynamic panel model is:

$$y_{it} = \alpha y_{i,t-1} + x'_{it}\beta + \mu_i + \varepsilon_{it} \tag{4}$$

where y_{it} is the dependent variable for unit i at time t, x_{it} is a vector of explanatory variables, μ_i represents unobserved time-invariant individual effects, ε_{it} is the idiosyncratic error term and $y_{i,t-1}$ is the lagged dependent variable. Including this last term makes the model dynamic but also introduces endogeneity. This is because $y_{i,t-1}$ is correlated with μ_i [32].

To eliminate the fixed effects μ_i, the model can be first-differenced:

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta x'_{it}\beta + \Delta\varepsilon_{it} \tag{5}$$

However, in Eq. (5), $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ is correlated with $\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{i,t-1}$, since $y_{i,t-1}$ is itself a function of $\varepsilon_{i,t-1}$. Thus, endogeneity is again introduced[3] [32].

In 1991, Arellano and Bond introduced what is now known as the Difference GMM to solve this endogeneity issue. They proposed using lagged levels of y_{it} as instruments for the endogenous term $\Delta y_{i,t-1}$. Levels of $y_{i,t-2}$ and earlier are valid instruments for Δy_{it-1}. Consequently, the earliest period for which a valid instrument ($y_{i,t-2}$) becomes available is T = 3. This forms the basis of the Difference GMM estimator, which estimates Eq. (5) using GMM with internal instruments [32,33]. Afterward, in 1995, Arellano and Bover extended the Difference GMM framework by proposing the use of lagged differences of y_{it} as instruments for the endogenous term y_{it-1} for the level equation (Eq. (4)). Thereby addressing some of the weaknesses of the original Difference GMM estimator [32,34]. Finally, in 1998, Blundell and Bond formalized the System GMM estimator [32,35], which combines:

[3] This issue of endogeneity also arises when using the within estimator in dynamic panel models to remove fixed effects μ_i. See [32] for a detailed discussion.

1. The first-differenced equation Eq. (5), where Δy_{it-1} is instrumented with lagged levels of y. Valid instruments for $\Delta y_{i,t-1}$ include $y_{i,t-2}, y_{i,t-3}, \ldots, y_{i,t-k}$ provided that $t \geq 3$ and $t-k \geq 1$.
2. The level equation Eq. (4), where y_{it-1} is instrumented with lagged differences of y. Valid instruments for $y_{i,t-1}$ include $\Delta y_{i,t-1}, \Delta y_{i,t-2}, \ldots, \Delta y_{i,t-k}$ provided that $t \geq 3$ and $t-k \geq 1$.

This system of equations increases efficiency, in contrast to the Difference GMM, by using additional moment conditions. The efficiency gains can be substantial, especially when the number of periods T is moderately small or when the time series exhibits high persistence [32,35]. Furthermore, as evident from the equations, the number of valid instruments increases with the number of periods T. This is due to the greater availability of lagged values of the endogenous variable [9]. Finally, it is essential to note that the Difference and System GMM are designed for panel data where the number of individuals N is large and the number of periods T is small. As $N \to \infty$, both estimators are consistent under standard assumptions [33,35].

3.1.3 Instrument reduction methods

There are two widely used techniques for instrument reduction in GMM estimation. The first method involves limiting the number of lags used as instruments [9]. Specifically, by restricting the instruments to lags 2 and 3 of the dependent variable y, the set of available instruments I for each period T in the difference equation changes from:

$$\begin{array}{c|ccccccc} & I_1 & I_2 & I_3 & I_4 & I_5 & I_6 & \cdots \\ \hline T_2 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ T_3 & y_{i1} & 0 & 0 & 0 & 0 & 0 & \cdots \\ T_4 & 0 & y_{i2} & y_{i1} & 0 & 0 & 0 & \cdots \\ T_5 & 0 & 0 & 0 & y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \tag{6}$$

to:

$$\begin{array}{c|cccccc} & I_1 & I_2 & I_3 & I_4 & I_5 & \cdots \\ \hline T_2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ T_3 & y_{i1} & 0 & 0 & 0 & 0 & \cdots \\ T_4 & 0 & y_{i2} & y_{i1} & 0 & 0 & \cdots \\ T_5 & 0 & 0 & 0 & y_{i3} & y_{i2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \tag{7}$$

The second method is to 'collapse' the instruments horizontally so the instruments of (6) become [9]:

$$\begin{array}{c|cccc} & I_1 & I_2 & I_3 & \cdots \\ \hline T_2 & 0 & 0 & 0 & \cdots \\ T_3 & y_{i1} & 0 & 0 & \cdots \\ T_4 & y_{i2} & y_{i1} & 0 & \cdots \\ T_5 & y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \tag{8}$$

3.2 Research procedure

This section outlines the research procedures for the verification (replication) and reanalysis (robustness analysis).

3.2.1 Verification

The analysis of the original research was based on 28 EU countries over the period 2010-2019. The authors were not contacted for data, as it was publicly available on the ScienceDirect publication page [3]. The verification is conducted using R and RStudio, with the 'plm' package employed to perform the panel data analyses, consistent with the original authors' use of these tools [3,36].

Table 1: Variables overview [3]

Variables	**Definition**	**Eurostat code**	**Unit of measure**
Dependent variable			
DE	Domestic Extraction	env_ac_mfa	1000 tonnes
Socioeconomic and structural variables			
GDP	Gross Domestic Product	name_10_gdp	million PPS
POP	Population	demo_gind	million inhabitants
DE/DMC	Economy reliance on domestic extraction	env_ac_mfa	% of DMC
CONST/GDP	Gross value added by construction sector	nama_10_a10	% of total gross value added
CE-related variables			
MWAS	Municipal waste per capita	cei_pc031	kilograms per capita
RECW	Recycling rate of municipal waste	cei_wm011	%
CMU	Circular Material Use rate	cei_srm030	%
CE_EMP	Employment in CE sectors	cei_cie011*	% of total employment
CE_INV	Gross private investment in tangible goods in CE sectors	cei_cie012*	% of GDP
CE_VA	Gross value added by CE sectors	cei_cie012*	% of GDP
CE_PAT	Number of patents related to recycling and secondary raw materials	cei_cie020	nr. per million inhabitants*

Notes: * indicates update/correction of the original paper [3]

Table 1 provides an overview of the variables categorized into macro-socioeconomic and CE-related factors. The macro-socioeconomic variables include the classical STIRPAT variables, DE/DMC, and CONST/GDP, which were included to capture structural economic influences on DE. The DE/DMC ratio reflects a country's dependence on domestic resources versus imports, with values above 100% indicating a resource-rich economy, while values below 100% suggest greater reliance on imported materials. According to the author, this ratio was the best approach to approximate the effects of imports and exports on resource use, as DE only accounts for domestic extraction. Other trade data were not measured by raw material equivalents but by weight, which does not accurately capture the physical quantities of resources extracted. The CONST/GDP variable accounts for the domestic demand for construction materials, which typically comprise a significant portion of DE and are rarely traded over long distances due to their low value-to-weight ratio. This makes the construction sector a key driver of domestic resource extraction [3]. The CE-related variables included in the analysis are summarized in Table 1. For additional details about these variables, please refer to the explanation in Appendix A.

$$\ln(\mathrm{DE}_{it}) = \beta_1 \ln(\mathrm{GDP}_{it}) + \beta_2 \ln(\mathrm{POP}_{it}) + \beta_3 \ln(\mathrm{DE/DMC}_{it}) + \beta_4 \ln(\mathrm{CONST/GDP}_{it}) + \gamma\, x\, \ln(\mathrm{CE}_{it}) + \mu_i + \epsilon_{it} \tag{9}$$

The structure of the panel data analysis is presented in Eq. (9). The dependent and explanatory variables vary across time (t) and countries (i). Where t refers to a specific year (t = 1,...,9) and i to a particular European country (i = 1,...,28). The term CE denotes a vector containing the selected CE variables, as detailed in Table 1. The parameter μ_i accounts for unobserved individual effects or heterogeneity, while ϵ_{it} represents the idiosyncratic error. Since all variables are expressed in a natural logarithmic (ln) scale, the regression coefficients β_x and the vector γ_x measure the elasticities between the explanatory and dependent variables. With the STIRPAT in logarithmic form as the underlying modeling approach, these coefficients indicate the percentage change in DE resulting from a 1% shift in the corresponding explanatory variable, assuming all other factors remain constant [3].

The selected CE variables differ in availability across both years and countries. Including all these variables simultaneously in Eq. (9) would have significantly reduced the panel size. Therefore, the socioeconomic and structural variables served as a consistent foundation for testing each CE variable individually. This approach was examined using pooled OLS, which assumes no individual or time effects, and fixed (FE) and random effects (RE) regression. The Breusch-Pagan heteroskedasticity test was conducted to assess the presence of heteroskedasticity, which would necessitate robust standard errors in all models. An F-test and the Lagrange Multiplier test were conducted to determine whether a pooled OLS, FE, or RE model was more suitable. Additionally, the Hausman specification test was applied to evaluate whether individual effects were uncorrelated with the explanatory variables, guiding the decision on whether the RE model was more appropriate than the FE specification [3].

Building on the previous steps, three fixed effects (FE) models were constructed for 2010–2018, 2010–2015, and 2013–2018, incorporating all significant CE variables. Excluding the lagged dependent variable, Eq. (10) presents their model specification, following [3].

$$\ln(\mathrm{DE}_{it}) = \mathrm{lag}(\ln(\mathrm{DE}_{it})) + \beta_1 \ln(\mathrm{GDP}_{it}) + \beta_2 \ln(\mathrm{POP}_{it}) + \beta_3 \ln(\mathrm{DE/DMC}_{it}) + \beta_4 \ln(\mathrm{CONST/GDP}_{it}) + \gamma\, x\, \ln(\mathrm{CE}_{it}) + \mu_i + \epsilon_{it} \tag{10}$$

Subsequently, Eq. (10) represents the System GMM, which was used for inference. Given that current levels of resource use are likely to be highly dependent on past consumption levels, a lagged value of DE was included as an explanatory variable to address endogeneity. Specifically, $\mathrm{lag}(\ln(\mathrm{DE}_{it}))$ denotes the first lag of the natural logarithm of DE. The first (AR1) and second-order (AR2) residual autocorrelation tests by Arellano-Bond were employed to evaluate the model. Additionally, the Sargan test was used to assess the validity of the instruments [3].

Finally, variables with significant effects on DE were used to assess the absolute impact on DE. Therefore, each significant coefficient (elasticity) was multiplied by its compound annual growth rate (CAGR) over its available timeframe. This product was then multiplied by the total DE in the final year of the panel across all entities, yielding the absolute contribution of each variable to DE [3].

3.2.2 Reanalysis

This section outlines the model specifications used in the reanalysis, which form the basis of the robustness analysis. The reanalysis is conducted first using the original data (2010-2019) and then with the updated and extended dataset (2010-2022) to test the sensitivity and validity of the original System GMM results. The updated and extended data was drawn from the European Commission's CE monitoring framework, explained in Section 2.2, using the R package 'eurostat' [37].

It is essential to clarify the terminology surrounding the Sargan and Hansen tests. The Hansen J-test, which is central to the analysis in Roodman's (2009) paper [9], is also referred to as the Sargan test in the context of System GMM in R's pgmm package, which labels the test as the Hansen-Sargan J-test [38]. For simplicity, this study will refer to it as the Sargan test throughout.

As noted in the introduction, there are signs of excessive instrument use in the original model. Most notably, the p-value of the Sargan test, which is exactly 1.00, is an outcome that raises concerns. Even lower p-values above 0.25 may indicate instrument proliferation, which would overfit the variables they intended to instrument. In such cases, it compromises the instruments' ability to eliminate endogeneity. Moreover, it weakens the Sargan test's power to detect invalid overidentifying restrictions, leading it to reject the null hypothesis too infrequently. To address this, the approach proposed by Roodman (2009) is followed: systematically reduce the instrument count and observe the robustness of the coefficient estimates and Sargan test. This approach also restores the power of the Sargan test without introducing additional bias [9]. Two instrument reduction techniques are employed to follow the approach: restricting instruments to the second and third lags and collapsing the instrument set. These methods are applied both individually and in combination.

A related subsequent change concerns the lower-bound value of the original lag structure, set at 1. This specification allows the first lag of DE to be used as an instrument for itself, which introduces endogeneity in the difference equation. As discussed in Section 3.1.2, only second and higher-order lags of DE are valid instruments for the difference equation. Additionally, when printing the instruments for the difference equation, two instruments appear for T3 in Eq. (6), further confirming this specification issue. To address this, the lower bound of the lag structure is now corrected to start at 2.

The third adjustment in specification involves replacing the two-step System GMM estimation with the one-step version. While this may seem counterintuitive given the consensus that two-step System GMM yields more efficient estimates, the decision is motivated by the small sample size in this study, with only 27 or 28 individuals [39]. In such contexts, two-step estimators may perform poorly, and even the use of GMM itself can be questionable (a point discussed later). Research supports this adjustment: a study in 2009 recommends the one-step estimator for small samples, noting its comparable accuracy and efficiency to the two-step approach [40]. A 2018 study further argues that efficiency gains from the two-step method may not hold in finite samples and that relying on the two-step procedure without careful consideration can do more harm than good [41]. Lastly, a study of 2020 also opted for the one-step estimation due to their low number of individuals [39].

Finally, a change is made to the original author's treatment of covariates in the System GMM specification. All covariates except for $DE_{i,t-1}$ were treated as exogenous in the original setup. This assumption is revised to reflect a broader endogeneity concern. One may argue that unobserved factors, such as national or regional policies or the political orientation of the ruling party, may simultaneously influence CE indicators and DE levels. For instance, a country's government reflects societal attitudes and priorities, which in turn may influence DE levels and key CE indicators like the municipal recycling rate (RECW), the circular material use rate (CMU), and employment in CE sectors (CE_EMP). The omission of such factors renders these covariates endogenous due to omitted variable bias. Moreover, GDP and the domestic extraction to domestic material consumption ratio (DE/DMC) are also considered endogenous, given the possibility of reverse causality. For example, higher DE may directly contribute to increased GDP and DE/DMC. In contrast, population (POP) and the share of construction in total GDP (CON-

ST/GDP) are assumed to be exogenous. Although CONST/GDP is a key driver of DE, it is treated as exogenous under the assumption that it primarily reflects the demand-driven structure of the economy, not the supply of resources. As such, reverse causality from DE to CONST/GDP is considered unlikely. This revised specification will hereafter be referred to as the endogeneity adjustment.

Table 2: Reanalysis - table overview

Specification changes	**Table 6**	**Table 11**	**Table 7**
Instrument reduction	yes	yes	yes
Lower-bound lag structure change	yes	yes	yes
One- or two-step estimation	one	two	both
Endogeneity adjustment	yes	yes	no

Notes: All tables also contain a re-estimation on the extended and updated dataset.

In total, the reanalysis implements a maximum of the four discussed specification changes, presented across three tables as depicted in Table 2. Table 6 serves as the primary table, while Table 11 is designed for robustness checks and is included in Appendix B. Additionally, to bring the models closer to the original authors' specification, a third table (7) reversing the above endogeneity assumption is also included in the results.

3.3 Results

The descriptive statistics in Table 3 were largely unverifiable. In particular, none of the reported compound annual growth rates (CAGR) could be verified, including for variables such as MWAS, MRECW, and CMU, despite other descriptive statistics for these variables aligning with the original study. The CAGR was calculated individually for each EU member state and averaged across countries. An alternative approach was tested for CMU by first aggregating CMU data for the entire European Union per year, resulting in a CAGR of 1.91%. However, this value still did not match the figure reported in the original paper of 1.23%.

Table 3: Descriptive statistics

Variable	**Obs**	**Mean**	**CV**	**Min**	**Max**	**CAGR**
DE	280	204227	1.14	1520	1046259	0.20
GDP	280	518158	1.40	9100	3209111	3.44
POP	280	18.16	1.27	0.41	83.09	0.24
DE/DMC	280	0.85	0.24	0.14	1.39	-0.35
CONST/GDP	280	4.87	0.27	1.00	8.20	-0.52
MWAS	273	484	0.26	247	862	0.42
RECW	273	34.23	0.45	4.00	67.20	6.00
CMU	280	8.75	0.72	1.20	30.00	2.50
CE_EMP	196	1.79	0.23	1.10	2.89	0.67
CE_INV	178	0.14	0.42	0.02	0.35	-1.55
CE_VA	199	0.95	0.21	0.35	1.56	0.26
CE_PAT	196	0.80	1.52	0.00	12.10	6.23

Notes: CV refers to the coefficient of variation. CAGR refers to the compound annual growth rate.

Table 4 presents the verification of Table 2 from the original study, where each CE-related variable was regressed individually alongside the socioeconomic and structural variables to assess individual significance. Notably, the CE_INV variable was significant in the verification, diverging from the original study's findings. This result was also associated with differing R^2, adjusted R^2, and panel structure values. In addition, the F-statistics differed consistently from the original paper. Nonetheless, the coefficients and standard errors were exactly matched across the entire table. Heteroskedasticity-consistent standard errors were computed using the Arellano (1987) method with the HC0 setting in R [42].

Table 4: Fixed-effects exploratory models results

Coefficient	**(1:1)** MWAS	**(1:2)** RECW	**(1:3)** CMU	**(1:4)** CE_EMP	**(1:5)** CE_INV	**(1:6)** CE_VA	**(1:7)** CE_PAT
GDP	**0.335** **	**0.500** ***	**0.410** ***	**0.476** ***	**0.396** ***	**0.417** ***	**0.344** **
	(0.140)	(0.142)	(0.102)	(0.128)	(0.146)	(0.134)	(0.163)
POP	**-0.963** *	**-1.434** **	**-1.214** ***	**-2.182** ***	**-1.826** ***	**-2.087** ***	**-1.413** *
	(0.495)	(0.623)	(0.445)	(0.561)	(0.627)	(0.640)	(0.820)
DE/DMC	**0.343** **	**0.342** ***	**0.330** **	**0.979** ***	**0.915** ***	**0.939** ***	**0.316** **
	(0.133)	(0.126)	(0.134)	(0.224)	(0.233)	(0.222)	(0.154)
CONST/GDP	**0.370** ***	**0.357** ***	**0.349** ***	**0.447** ***	**0.467** ***	**0.498** ***	**0.398** ***
	(0.100)	(0.096)	(0.101)	(0.099)	(0.116)	(0.120)	(0.144)
MWAS	0.101						
	(0.144)						
RECW		**-0.090** *					
		(0.049)					
CMU			**-0.101** **				
			(0.045)				
CE_EMP				**-0.486** ***			
				(0.163)			
CE_INV					**0.069** *		
					(0.040)		
CE_VA						-0.084	
						(0.104)	
CE_PAT							**0.039** *
							(0.023)
R2	0.4154	0.4427	0.4585	0.5181	0.4904	0.4739	0.3730
R2 adjusted	0.3374	0.3684	0.3884	0.4373	0.3946	0.3872	0.2499
F-statistic	34.1022	38.1288	41.8305	35.9078	28.6734	30.6232	19.3962
Panel structure	N = 28, T = 10, n = 273	N = 28, T = 10, n = 273	N = 28, T = 10, n = 280	N = 24, T = 9, n = 196	N = 24, T = 9, n = 178	N = 24, T = 9, n = 199	N = 28, T = 7, n = 196

Notes: Standard errors in parentheses. ******* $p < 0.01$; ****** $p < 0.05$; ***** $p < 0.1$.

Table 5 replicates the final model results, including the System GMM. Again, the F-statistic could not be verified. The fixed effects models used the same robust standard error computation previously described. All coefficients and standard errors were again precisely verified, except for the standard error for GDP in the System GMM model. Instead, in the original paper, the standard error for GDP matched that of the lagged DE coefficient.

Table 5: Fixed-effects and GMM full models results

Coefficient	**(1)** FE 2010–2018	**(1a)** FE 2010–2015	**(1b)** FE 2013–2018	**(2)** GMM 2010–2019
GDP	**0.673** ***	**0.872** ***	**0.506** ***	**0.307** ***
	(0.127)	(0.219)	(0.071)	(0.105)
POP	**-2.815** ***	**-4.126** ***	**-0.911** *	**0.105** **
	(0.544)	(0.902)	(0.507)	(0.042)
DE/DMC	**1.006** ***	**0.761** ***	**0.804** ***	**1.117** ***
	(0.187)	(0.190)	(0.129)	(0.319)
CONST/GDP	**0.431** ***	**0.433** ***	**0.218** **	**0.183** ***
	(0.084)	(0.110)	(0.087)	(0.056)
RECW	**-0.088** ***	**-0.156** ***	0.034	0.033
	(0.030)	(0.053)	(0.031)	(0.045)
CMU	-0.060	-0.055	**-0.141** ***	**-0.112** **
	(0.039)	(0.046)	(0.043)	(0.050)
CE_EMP	**-0.415** ***	-0.341	0.039	**-0.338** ***
	(0.138)	(0.260)	(0.152)	(0.099)
Lag 1 DE				**0.545** ***
				(0.142)
R2	0.5854	0.5991	0.5027	
R2 adjusted	0.5095	0.4751	0.3550	
F-statistic	33.0748	20.7088	14.5836	
AR (1) p-value				0.0109
AR (2) p-value				0.1140
Sargan test p-value				1.0000
Wald test p-value				0.0000
Panel structure	N = 24, T = 9, n = 195	N = 24, T = 6, n = 128	N = 24, T = 6, n = 132	N = 28, T = 10, n = 321

Notes: Standard errors in parentheses. ******* $p < 0.01$; ****** $p < 0.05$; ***** $p < 0.1$.

After extensive model specification testing, the original System GMM specification was found (R-code is included in Appendix B). The specification employs two-step estimation without time dummies. The lag structure (1:99) consists of all available lags of DE from the first onward as instruments [43]. The transformation option 'ld' confirms using both level and difference equations, corresponding to the System GMM approach. With the above specification, all variables (except lag(log(DE), 1) are assumed to be exogenous. Lastly, the 'robust = TRUE' setting enables the Windmeijer (2005) correction for finite sample bias in two-step GMM estimations [44].

The random effects model (Table 9) and the specification tests (Table 10) were successfully verified, with one exception: the coefficient for CE_EMP in the random effects model was missing a minus sign. Both outputs are included in Appendix B.

Concerning the reanalysis, Table 6 summarizes the one-step System GMM estimation under each robustness specification. The original System GMM, model (1), now reports the total number of instruments: 56 for each of the 28 countries. In model (2), despite restricting lag depth, the instrument count rises sharply due to the endogeneity adjustment, resulting in the Sargan test p-value remaining at 1.00 and the loss of statistical significance in the effect of CE_EMP. Combining lag restrictions with collapsed instruments in model (4) sharply cuts the instrument count, which consequently removes the significant impact of CMU and lowers the Sargan p-value enough to reject the null hypothesis, suggesting model misspecification or the correlation of one or more instruments with the error term [45]. Re-estimating with updated and extended data in model (5) similarly fails the Sargan test and confirms that the effects of CMU and CE_EMP remain statistically insignificant.

This pattern is also observed in the two-step estimation (Table 11 in Appendix B) except that model (4) passes the Sargan test ($p = 0.1559$) when estimated on the original data. However, when the data is updated and extended, the Sargan test p-value falls to 0.0339, again rejecting instrument validity. Across all one- and two-step model specifications, the AR(2) tests confirm no second-order autocorrelation.

Table 6: One-step System GMM estimation results under alternative specifications

Coefficient	**(1)** Original System GMM 2010–2019	**(2)** Second-third lag instruments 2010–2019	**(3)** Collapsed full instruments 2010–2019	**(4)** Collapsed second-third lag instruments 2010–2019	**(5)** Extended & updated data 2010–2022
GDP	**0.307** ***	0.076	**0.463** ***	**0.462** **	**0.120** **
	(0.105)	(0.048)	(0.151)	(0.210)	(0.049)
POP	**0.105** **	0.016	-0.066	-0.102	0.037
	(0.042)	(0.029)	(0.101)	(0.127)	(0.076)
DE/DMC	**1.117** ***	**0.305** **	**1.126** **	**1.075** *	0.177
	(0.319)	(0.141)	(0.502)	(0.604)	(0.170)
CONST/GDP	**0.183** ***	**0.057** ***	**0.173** ***	**0.141** **	0.094
	(0.056)	(0.021)	(0.061)	(0.066)	(0.060)
RECW	0.033	0.013	-0.010	-0.062	-0.060
	(0.045)	(0.015)	(0.058)	(0.090)	(0.054)
CMU	**-0.112** **	**-0.035** *	**-0.124** **	-0.081	-0.019
	(0.050)	(0.019)	(0.051)	(0.072)	(0.040)
CE_EMP	**-0.338** ***	-0.062	-0.137	-0.019	-0.022
	(0.099)	(0.058)	(0.219)	(0.277)	(0.112)
Lag 1 DE	**0.545** ***	**0.900** ***	**0.620** ***	**0.677** ***	**0.835** ***
	(0.142)	(0.042)	(0.112)	(0.138)	(0.131)
One- or two-step	two-step	one-step	one-step	one-step	one-step
Endogeneity adjustment	no	yes	yes	yes	yes
Lag structure	1:99	2:3	$2{:}99^{collapsed}$	$2{:}3^{collapsed}$	$2{:}3^{collapsed}$
AR (1) p-value	0.0109	0.0016	0.0055	0.0061	0.0002
AR (2) p-value	0.1140	0.1746	0.1435	0.1399	0.0646
Sargan test p-value	1.0000	1.0000	0.9993	0.0477	0.0231
Wald test p-value	0.0000	0.0000	0.0000	0.0000	0.0000
Number of instruments	56	122	52	22	22
Panel structure	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 27, T = 13, n = 609

Notes: Standard errors in parentheses. ******* $p < 0.01$; ****** $p < 0.05$; ***** $p < 0.1$. The lag structure is defined with respect to period t. Robust standard errors are employed. No time effects included.

Additionally, Table 7 is designed to approximate the original authors' specifications more closely by removing the endogeneity adjustment. The effect of CE_EMP remains insignificant under the reduced-instrument settings. In contrast to the models with endogeneity adjustment, the significant effect of CMU on DE persists even after reducing the number of instruments. However, the models are uniformly invalid across all resulting specifications, both one-step and two-step estimators, using original and updated datasets. The Sargan test rejects the null hypothesis of valid overidentifying restrictions at the 5% significance level in three instances and at the 10% level in one instance.

Table 7: One and two-step System GMM estimation results without endogeneity adjustment

Coefficient	**(1)** One-step 2010–2019	**(2)** Two-step 2010–2019	**(3)** One-step, extended & updated data 2010–2022	**(4)** Two-step, extended & updated data 2010–2022
GDP	**0.262 *****	**0.264 ****	**0.165 ****	**0.153 ***
	(0.073)	(0.104)	(0.070)	(0.078)
POP	0.079	0.085	0.074	0.077
	(0.055)	(0.083)	(0.052)	(0.066)
DE/DMC	**0.942 *****	**0.860 *****	**0.402 *****	**0.393 ****
	(0.292)	(0.314)	(0.144)	(0.182)
CONST/GDP	**0.150 *****	**0.149 ****	**0.130 ****	**0.132 ***
	(0.050)	(0.067)	(0.065)	(0.069)
RECW	0.006	0.049	0.006	0.010
	(0.047)	(0.069)	(0.029)	(0.038)
CMU	**-0.078 ****	**-0.110 ***	**-0.070 ****	**-0.076 ***
	(0.037)	(0.057)	(0.036)	(0.039)
CE_EMP	**-0.274 ****	-0.266	-0.063	-0.052
	(0.118)	(0.182)	(0.064)	(0.051)
Lag 1 DE	**0.630 *****	**0.609 *****	**0.728 *****	**0.736 *****
	(0.097)	(0.142)	(0.114)	(0.136)
One- or two-step	one-step	two-step	one-step	two-step
Endogeneity adjustment	no	no	no	no
Lag structure	$2{:}3^{collapsed}$	$2{:}3^{collapsed}$	$2{:}3^{collapsed}$	$2{:}3^{collapsed}$
AR(1) p-value	0.0049	0.0058	0.0005	0.0014
AR(2) p-value	0.1324	0.1627	0.0889	0.0783
Sargan test p-value	0.0243	0.0761	0.0142	0.0225
Wald test p-value	0.0000	0.0000	0.0000	0.0000
Number of instruments	17	17	17	17
Panel structure	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 27, T = 13, n = 609	N = 27, T = 13, n = 609

Notes: Standard errors in parentheses. ******* $p < 0.01$; ****** $p < 0.05$; ***** $p < 0.1$. The lag structure is defined with respect to period t. Robust standard errors are employed for one-step estimation. The finite sample correction for standard errors proposed by Windmeijer (2005) is applied for two-step estimation [44]. No time effects included.

4 Re-estimation

As demonstrated in the reanalysis, the System GMM estimation revealed no significant effect of CE initiatives on domestic extraction when implementing the endogeneity adjustment or lowering the number of instruments. However, the Sargan tests indicated that endogeneity persisted through invalid instruments. Therefore, an alternative estimation method is used to corroborate the apparent lack of an effect of CE policies on domestic extraction. This section first explains the rationale behind the chosen method and outlines its underlying mechanics, followed by a presentation of the corresponding results.

4.1 Methods & research procedure

As discussed in Section 3.2.2, correcting for endogeneity in several key variables is necessary. Therefore, the primary requirements for this new approach were its ability to handle panel data and adequately address endogeneity. In this context, 2SLS was selected as the preferred method, as it allows instrumental variables to correct for endogeneity while also applying to panel data settings [45]. However, one consequence of using 2SLS is excluding the lagged dependent variable from the model. In 2SLS, using lagged variables as instruments would have required dropping initial periods, reducing the sample size [9]. Given the limited time dimension of the panel, it was decided to omit the lagged variable.

Moreover, identifying suitable instruments, particularly for RECW, CMU, CE_EMP, and DE/DMC, is especially challenging. Given the panel structure of the data, valid instruments must have repeated observations across countries and years. For example, in the case of CMU, one would need an instrument that varies by country and year and influences only the exogenous component of CMU without directly affecting the dependent variable. Finding such instruments for CMU and the other endogenous variables mentioned was ultimately not feasible for the scope of this study. The only exception was GDP, for which total employment[4] (TOTAL_EMP) was considered a valid instrument. This is based on the assumption that employment levels strongly correlate with GDP, while their direct effect on domestic extraction is minimal or operates primarily through GDP. Eurostat supports this assumption, showing that the mining and quarrying sector accounted for only about 0.2% of total employment in 2022 [46]. The Lewbel residual-based instruments provide a solution for this challenge of identifying appropriate instruments, as they are constructed from the data at hand. A simplified version of the instrument creation process for this study is outlined below[5].

As per Theorem 1 in [13], instruments are constructed separately for each endogenous variable (GDP, DE/DMC, RECW, CMU, and CE_EMP) by regressing them on the exogenous variables, as shown in Eq. (11). Where y_{it} denotes the endogenous variable for unit i at time t, x_{it} is the vector of exogenous variables, and ε_{it} is the idiosyncratic error term, with μ_i removed via the within transformation[6]. The exogenous variables include POP, CONST/GDP, and time dummies. This differs from the System GMM approach, where time dummies were excluded as they resulted in the creation of additional instruments.

$$y_{it} = x_{it}'\beta + \varepsilon_{it} \tag{11}$$

Next, the instruments are generated using the product of the residuals and the centered exogenous variables. For time dummies, the instruments take the form $(Z - \overline{Z})\varepsilon_{it}$ while for the remaining exogenous variables, the log-transformed version $(\ln(Z) - \ln(\overline{Z}))\varepsilon_{it}$ was used, reflecting the log-log specification of the

[4] Eurostat code: `lfsi_emp_a`, measured in thousands of persons aged 20–64.

[5] For full details, refer to [13].

[6] For simplicity, the notation abstracts from demeaning and the log-log model structure.

model. Here, Z denotes a subset of x'_{it}. These instruments, individually or in combination, were included on the right-hand side of Eq. (11) to form the first-stage regression for each endogenous variable in the 2SLS estimation. The following criteria guided the final selection of instruments:

1. The total number of instruments was minimized. Therefore, a maximum of two instruments per endogenous variable was used.
2. The instrument set with the highest first-stage F-statistic, exceeding the 10 threshold to exclude weak instruments, was selected for each regression [47].
3. To enable the Sargan test in the second stage, at least one of the first-stage regressions must be overidentified. Ideally, the regression with the weakest F-statistic, provided that this adjustment also improved the F-statistic.

The standard 2SLS procedure was then applied. The fitted values from the first-stage regressions and the exogenous variables were used in the second stage [45]. Time dummies were included this time, and country-clustered standard errors were employed. Wald and Sargan tests were conducted to assess the strength and validity of the instruments.

As discussed in Section 2.2, the CE monitoring framework was updated with new indicators. One key addition was raw material consumption[7] (RMC), which addresses potential burden shifting (see Section 2.3) in the dependent variable DE. Unlike DE, RMC accounts not only for the raw materials extracted domestically but also incorporates the imports and exports of raw materials. RMC is expressed in the raw material equivalent (RME) unit, which conceptually aligns with DE. The RME of a product represents the total extraction of materials required throughout the entire production chain to manufacture that product, regardless of whether the extraction occurred domestically or abroad [48]. RMC is thus a burden-based indicator and is computed as follows [4,48]:

$$RMC = RME_{\text{import}} + DE - RME_{\text{export}} \tag{12}$$

The second-stage regression was therefore also run using RMC as the dependent variable, in addition to DE, to assess the potential for burden shifting. The same instrument sets were applied.

4.2 Results

Table 8 presents the results from the 2SLS estimation using Lewbel residual-based instruments. Based on the instrument selection criteria, the first-stage regression for RECW could not be overidentified, as none of the available instrument sets met the minimum F-statistic threshold. Similarly, attempts to overidentify CMU led to a substantial decline in the F-statistic, down to approximately 70. As a result, GDP was the only variable intentionally overidentified, as its instrument set maintained strength and included an external instrument (TOTAL_EMP), which was selected based on its relevance. This configuration allowed the Sargan test to be applied in the second stage. Consequently, a total of six instruments were used across five endogenous variables. The Sargan test failed to reject the null hypothesis of valid overidentifying restrictions in both second-stage models, indicating that no evidence was found that the instruments employed are correlated with the error term. Additionally, the Wald test p-values are below 1% in both models, confirming the joint statistical significance of the included regressors.

In the second-stage regression with DE as the dependent variable, a 1% increase in GDP is associated with a 1.095% increase in domestic extraction[8]. A 1% increase in DE/DMC corresponds to a 0.364%

[7] Eurostat code: `env_ac_rme`, measured in thousands of tonnes.
[8] All reported effects represent average effects, conditional on the other variables in the model being held constant.

increase in DE. The CE indicators, including CMU, CE_EMP, and MRECW, do not show statistically significant effects in this model.

Table 8: 2SLS with Lewbel residual-based instruments

Coefficient	Stage 1: GDP	Stage 1: DE/DMC	Stage 1: RECW	Stage 1: CMU	Stage 1: CE_EMP	Stage 2: DE	Stage 2: RMC
POP	**-0.291 ****	**-0.502 *****	**-3.234 *****	-0.519	**-0.261 ****	-0.519	**-1.584 *****
	(0.117)	(0.084)	(0.371)	(0.437)	(0.112)	(0.536)	(0.551)
CONST/GDP	**0.140 *****	0.007	**0.195 ****	-0.160	-0.016	0.109	**0.242 ****
	(0.022)	(0.021)	(0.095)	(0.112)	(0.029)	(0.107)	(0.095)
TOTAL_EMP^i	**0.802 *****						
	(0.088)						
res(year_2016)^i	**-3.915 *****						
	(0.370)						
res(POP)^i		**-0.326 *****					
		(0.018)					
res(CONST/GDP)^i			**-0.799 *****				
			(0.231)				
res(POP)^i				**-0.306 *****			
				(0.033)			
res(POP)^i					**-0.419 *****		
					(0.019)		
E(GDP)						**1.095 *****	**1.028 *****
						(0.268)	(0.233)
E(DE/DMC)						**0.364 ***	-0.046
						(0.204)	(0.108)
E(RECW)						0.101	-0.106
						(0.097)	(0.082)
E(CMU)						-0.055	**-0.100 ***
						(0.067)	(0.055)
E(CE_EMP)						0.056	**0.158 ****
						(0.069)	(0.069)
F-statistic	145.7314	359.7810	12.9951	95.5678	505.4896		
Wald test p-value						0.0007	0.0001
Sargan test p-value						0.1649	0.1669
Number of instruments	2	1	1	1	1	6	6
Panel structure	N = 27, T = 13, n = 345	...	...	...	...	...	...

Notes: Standard errors in parentheses. Stage 1 reports uncorrected standard errors; Stage 2 reports country-clustered standard errors. ******* $p < 0.01$; ****** $p < 0.05$; ***** $p < 0.1$. ^{i} represents instruments, res represents the Lewbel (2012) instruments, E represents the estimate of the coefficient.

More statistically significant effects are observed in the second-stage regression with RMC as the dependent variable, which incorporates imports and exports of raw materials. First, a 1% increase in GDP continues to be significantly associated with RMC, corresponding to a 1.028% increase. The coefficient on DE/DMC is not statistically significant in this specification. A 1% rise in population is associated with a 1.584% decrease in RMC, while a 1% increase in CMU corresponds to a 0.100% reduction. Lastly, CONST/GDP and CE_EMP are both positively and significantly associated with RMC, with a 1% increase in each corresponding to a 0.242% and 0.158% increase in RMC, respectively.

5 Discussion

This study aimed to replicate, conduct robustness tests on, and re-estimate the study titled 'Does circular economy mitigate the extraction of natural resources? Empirical evidence based on analysis of 28 European economies over the past decade' published in Ecological Economics by Bianchi and Cordella (2023) [3]. Overall, after addressing methodological flaws in the robustness analysis and performing a re-estimation, no support was found for Bianchi and Cordella's claim that circular economy (CE) initiatives reduce domestic resource extraction.

Concerning the verification of the original study's findings, the replication process largely succeeded in verifying the main results despite some discrepancies. Minor deviations in descriptive statistics are likely due to continuous Eurostat updates. The potential manual construction of original tables introduced minor inconsistencies in standard errors, signs, and panel structure. More notably, excluding the circular economy sector investments from final models appears inconsistent with its prior significance and the stated methodology. The verification process also proved unnecessarily complex due to a lack of methodological detail, particularly concerning the calculation of CAGR, the type of robust standard errors used, and the precise specification of the GMM estimator. Nevertheless, the core findings could still be reasonably verified.

As the model specification of the original System GMM became clear in the verification, it revealed critical methodological flaws, which fundamentally undermined the reliability of its results. Most importantly, the model included 56 instruments for a panel of only 28 countries, illustrating instrument proliferation. Secondly, using the first lag of the dependent variable as an instrument in the differenced equation introduced, by definition, endogeneity. Lastly, the assumption that all explanatory variables are exogenous is insufficiently justified and subject to reasonable doubt.

The reanalysis aimed to address these methodological flaws by modifying the System GMM specification. The number of instruments was systematically reduced to evaluate the impact of instrument proliferation on the coefficient estimates and Sargan test results. In terms of coefficient estimates, the effect of employment in the CE sector was rendered insignificant, even under the original exogeneity assumption used by the authors. However, under this same exogeneity assumption, there was still an indication of a statistically significant impact of the circular material use rate on domestic extraction. In contrast, when more explanatory variables were treated as endogenous, there was strong evidence that no CE practices significantly impacted domestic extraction. Nevertheless, by restoring the Sargan test's power in the reduced-instrument setting, the tests rejected the null hypothesis of instrument validity in all but one case, indicating that some endogeneity remained in the models. Moreover, the panel structures (N = 28/27, T = 10/13) are poorly suited for GMM estimation, which requires many entities N[9]. Therefore, a re-estimation using an alternative estimation method was necessary to validate the apparent lack of an effect of CE policies on domestic resource extraction.

This re-estimation, using 2SLS, yielded quantitatively similar results to the findings of the reanalysis while eliminating all endogeneity. Neither approach found any significant effect of CE practices on domestic resource extraction, regardless of how those practices are measured. The consistency across these

[9] In contrast to settings with larger N, where additional lags and time effects may be justifiable (as suggested by Roodman (2009) [49]), such modifications only worsen the overidentification problem here.

two independent estimation strategies, each treating more explanatory variables as endogenous, suggests that the original study likely wrongly assessed the impact of the circular economy on domestic resource extraction. In addition, the effect of GDP on domestic resource extraction was found to be at least three times larger than reported in the original study. This result holds when using either the original CAGR or the alternative specifications presented in Table 3. At the same time, the estimated impact of a country's dependence on domestic resources became less pronounced.

Lastly, raw material consumption was used as the dependent variable instead of domestic resource extraction to evaluate the findings in a context that accounts for the import and export of raw materials (resources). In this case, the 2SLS results revealed notable differences. In particular, the circular material use rate was significantly associated with lower net trade in raw materials, whether through reduced imports or increased exports, suggesting that its primary effect operates through trade adjustments rather than cuts in domestic extraction. Concerning reduced imports specifically, this highlights the dual goals of the circular material use rate as a policy instrument: to advance the circular economy and to strengthen strategic autonomy [50]. Moreover, CE employment was found to positively impact raw material consumption, which indicates that countries with higher CE employment import more or export fewer resources, potentially due to limited domestic resource availability. Given the opposing effects of CE indicators on raw material consumption, no definitive conclusion can be made regarding the presence or absence of burden shifting. In addition, the greater impact of GDP on resource extraction was reaffirmed. Other factors, such as the GDP share in construction, also exhibited significant effects. The material-intensive nature of the construction sector likely increases raw material imports or limits exports. Meanwhile, the negative coefficient for population suggested that more populous countries depend less on raw material imports or are better positioned to increase exports. Overall, the 2SLS estimation results using raw material consumption as the dependent variable underscore the importance of considering international trade in assessments of the circular economy.

In conclusion, this study refutes the original claims by Bianchi and Cordella (2023) regarding the circular economy's role in mitigating European domestic resource extraction. While the verification broadly confirmed their original findings, several methodological inconsistencies raised concerns about the reliability of the estimates. Most notably, the instrument proliferation and questionable assumptions of exogeneity. Reanalysis using a more theoretically grounded System GMM specification to address these flaws found no effect of CE policies on domestic resource extraction. However, due to instrument invalidity, a 2SLS re-estimation was conducted. This re-estimation confirmed the absence of any significant link between CE practices and domestic extraction while showing GDP to have a much stronger impact than originally reported. Finally, replacing domestic extraction with raw material consumption as the outcome variable in the re-estimation revealed that CE initiatives influence resource trade more than extraction itself, particularly through reduced imports or increased exports. This highlights the role of trade mechanisms and the influence of CE policies on strategic autonomy. Together, these findings suggest that the original study overestimated the environmental impact of CE initiatives on domestic extraction. These findings extend beyond those of Bianchi and Cordella (2023), yet they ultimately support their policy recommendations. CE initiatives must go beyond recycling and waste reuse to reduce resource extraction effectively, as they do little to counteract the growing demand for resources [3]. The more substantial impact of GDP observed in this study only reinforces the urgency of addressing the drivers of material use, including demand-side pressures tied to economic expansion. Accordingly, policy should focus not only on optimizing waste stream management but also on actively reshaping consumption patterns to reduce overall resource demand.

This study contributes to the literature by advancing the understanding of the relationship between raw material extraction and CE by considering both imports and exports of resources, which sets it apart from prior research and offers a more comprehensive perspective. Additionally, this study enhances the credibility of economic research by replicating and performing robustness tests on the original study's findings, demonstrating the value of these practices in validating results. Methodologically, this study introduces 2SLS with Lewbel residual-based instruments as a practical solution for dealing with hard-to-instrument endogenous variables in panel data with a small number of entities. This provides a replicable strategy for similar empirical settings.

This study has several limitations. First, the 2SLS model no longer employs a dynamic panel specification, as the lagged dependent variable is excluded as a regressor. While this exclusion is necessary, it remains a notable limitation, given that the lagged dependent variable consistently exhibits a significant positive effect in all System GMM specifications. Second, the re-estimation using 2SLS is limited to the variables included in the System GMM of the original study. While some indicators, such as the number of CE patents, are excluded in the original study due to limited availability over a longer time frame, updated data made their inclusion possible. Additionally, the 2023 revision of the CE monitoring framework introduced new indicators that could enhance explanatory power. However, aside from RMC, these variables are not incorporated due to endogeneity concerns, as discussed in Section 3.2.2. Including them will require additional instruments, making the 2SLS specification more restrictive and potentially undermining the feasibility of estimating a complete second stage. Finally, the use of Lewbel residual-based instruments introduces some uncertainty. Unlike a System GMM, which relies on clearly defined lag structures, Lewbel instruments are constructed from model residuals and heteroskedasticity, which may raise concerns about their interpretability and relevance. However, this does not affect the instruments' strength or the estimates' statistical validity.

6 Conclusion

Through rigorous replication, robustness analyses, and re-estimation using 2SLS with Lewbel residual-based instruments, it is shown that CE initiatives do not reduce domestic resource extraction as originally claimed by Bianchi and Cordella (2023). Methodological flaws are identified as key factors undermining the original findings. However, when shifting the focus to raw material consumption, CE indicators like circular material use rate show significant effects through trade adjustments. While CE policies may not directly curb domestic extraction, they reduce net resource imports and thereby enhance strategic autonomy.

Acknowledgments

I want to thank Prof. Dr. Stephan Bruns, my promotor, for proposing the paper that formed the basis of this thesis. Thanks also to Mr. Tuan Nguyen, my supervisor, for being available for day-to-day questions, reviewing my work, and providing guidance on structure and academic writing. I am particularly grateful to Dr. Robin Clerckx for his support with instrumental variable estimation. Specifically, the implementation of 2SLS using Lewbel residual-based instruments. His input also helped improve the academic structure and clarity of the paper.

References

1. European Investment Bank: Circular economy Overview 2024 (2024)
2. United Nations Environment Programme, I.R.P.: Global Resources Outlook 2024 - Bend the trend: Pathways to a Liveable Planet as Resource Use Spikes (2024)
3. Bianchi, M., Cordella, M.: Does circular economy mitigate the extraction of natural resources? Empirical evidence based on analysis of 28 European economies over the past decade. Ecological Economics **203**, 107607 (2023)
4. Corona, B., Shen, L., Reike, D., Rosales Carreón, J., Worrell, E.: Towards sustainable development through the circular economy—A review and critical assessment on current circularity metrics. Resources, conservation and recycling **151**(Journal Article), 104498 (2019)
5. De Abreu, V.H.S., Da Costa, M.G., Da Costa, V.X., De Assis, T.F., Santos, A.S., D'Agosto, M.d.A.: The Role of the Circular Economy in Road Transport to Mitigate Climate Change and Reduce Resource Depletion. Sustainability **14**(14), 8951 (2022)
6. Korhonen, J., Honkasalo, A., Seppälä, J.: Circular Economy: The Concept and its Limitations. Ecological economics **143**(Journal Article), 37–46 (2018)
7. Kirchherr, J.: Circular economy and growth: A critical review of "post-growth" circularity and a plea for a circular economy that grows. Resources, Conservation and Recycling **179**, 106033 (2022)
8. Zink, T., Geyer, R.: Circular Economy Rebound. Journal of Industrial Ecology **21**(3), 593–602 (2017)
9. Roodman, D.: A Note on the Theme of Too Many Instruments*. Oxford Bulletin of Economics and Statistics **71**(1), 135–158 (2009)
10. European Commission: Circular economy action plan - European Commission (2023), `https://environment.ec.europa.eu/strategy/circular-economy-action-plan`
11. European Commission: How decisions are made - European Commission (2024), `https://commission.europa.eu/strategy-and-policy/how-decisions-are-made`
12. Clemens, M.A.: The Meaning of Failed Replications: A Review and Proposal. Journal of Economic Surveys **31**(1), 326–342 (2017)
13. Lewbel, A.: Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. Journal of Business & Economic Statistics **30**(1), 67–80 (2012)
14. Eurostat: Information on data - Eurostat (2023), `https://ec.europa.eu/eurostat/web/circular-economy/information-data`
15. European Commission: Directorate-General for Environment: Circular economy – New tool for measuring progress (2023), `https://circulareconomy.europa.eu/platform/en/measuring-circular-economy/circular-economy-new-tool-measuring-progress`
16. European Commission: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a revised monitoring framework for the circular economy (2023), `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2023:306:FIN`
17. International Resource Panel: Glossary (2017), `https://www.resourcepanel.org/glossary`
18. Brown, A.N., Cameron, D.B., Wood, B.D.K.: Quality evidence for policymaking: I'll believe it when I see the replication. Journal of Development Effectiveness **6**(3), 215–235 (2014)
19. Leuz, C.: Evidence-based policymaking: promise, challenges and opportunities for accounting and financial markets research**. Accounting and Business Research **48**(5), 582–608 (2018)
20. Anderson, R.G., Kichkha, A.: Replication, Meta-Analysis, and Research Synthesis in Economics. The American economic review **107**(5), 56–59 (2017)
21. Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., Buck, S., Chambers, C.D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D.P., Hesse, B., Humphreys, M., Ishiyama, J., Karlan, D., Kraut, A., Lupia, A., Mabry, P., Madon, T., Malhotra, N., Mayo-Wilson, E., McNutt, M., Miguel, E., Paluck, E.L., Simonsohn, U., Soderberg, C.,

Spellman, B.A., Turitto, J., VandenBos, G., Vazire, S., Wagenmakers, E.J., Wilson, R., Yarkoni, T.: Promoting an open research culture. Science (American Association for the Advancement of Science) **348**(6242), 1422–1425 (2015)
22. Chang, A.C., Li, P.: Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not". Finance and Economics Discussion Series **2015.0**(83), 1–26 (2015)
23. Camerer, C.F., Dreber, A., Forsell, E., Ho, T.H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., Wu, H.: Evaluating replicability of laboratory experiments in economics. Science **351**(6280), 1433–1436 (2016)
24. Serra-Garcia, M., Gneezy, U.: Nonreplicable publications are cited more than replicable ones. Science advances **7**(21) (2021)
25. Furman, J.L., Jensen, K., Murray, F.: Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. Research policy **41**(2), 276–290 (2012)
26. Trikalinos, N.A., Evangelou, E., Ioannidis, J.P.A.: Falsified papers in high-impact journals were slow to retract and indistinguishable from nonfraudulent papers. Journal of clinical epidemiology **61**(5), 464–470 (2008)
27. Kiri, B., Lacetera, N., Zirulia, L.: Above a swamp: A theory of high-quality scientific production. Research Policy **47**(5), 827–839 (2018)
28. Bruns, S.B., König, J., Stern, D.I.: Replication and robustness analysis of 'energy and economic growth in the USA: A multivariate approach'. Energy economics **82**(Journal Article), 100–113 (2019)
29. Ehrlich, P.R., Holdren, J.P.: A bulletin dialogue on "The Closing Circle" : critique: one-dimensional ecology. Bulletin of the Atomic Scientists **28**(5), 16–27 (1972)
30. York, R., Rosa, E.A., Dietz, T.: STIRPAT, IPAT and ImPACT: analytic tools for unpacking the driving forces of environmental impacts. Ecological Economics **46**(3), 351–365 (2003)
31. Dietz, T., Rosa, E.A.: Rethinking the Environmental Impacts of Population, Affluence and Technology. Human Ecology Review **1**(2), 277–300 (1994)
32. Baltagi, B.H.: Econometric Analysis of Panel Data. No. Book, Whole, Springer International Publishing AG, Cham, 6th edn. (2021)
33. Arellano, M., Bond, S.: Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. The Review of Economic Studies **58**(2), 277–297 (1991)
34. Arellano, M., Bover, O.: Another look at the instrumental variable estimation of error-components models. Journal of Econometrics **68**(1), 29–51 (1995)
35. Blundell, R., Bond, S.: Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics **87**(1), 115–143 (1998)
36. Croissant, Y., Millo, G.: Panel Data Econometrics in R: The plm Package. Journal of Statistical Software **27**(2), 1 – 43 (2008)
37. Lahti, L., Huovari, J., Kainu, M., Biecek, P.: Retrieval and Analysis of Eurostat Open Data with the eurostat Package. The R Journal **9**(1), 385–392 (2017)
38. Croissant, Y.: Hansen-Sargan Test of Overidentifying Restrictions (2025), `https://search.r-project.org/CRAN/refmans/plm/html/sargan.html`
39. Berk, I., Kasman, A., Kılınç, D.: Towards a common renewable future: The System-GMM approach to assess the convergence in renewable energy consumption of EU countries. Energy Economics **87**, 103922 (2020)
40. Soto, M.: System GMM Estimation With A Small Sample. Tech. rep., Federal Reserve Bank of St. Louis, St. Louis (2009)
41. Hwang, J., Sun, Y.: Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework. Journal of econometrics **207**(2), 381–405 (2018)
42. Arellano, M.: PRACTITIONERS' CORNER: Computing Robust Standard Errors for Within-groups Estimators. Oxford bulletin of economics and statistics **49**(4), 431–434 (1987)
43. Croissant, Y.: Generalized Method of Moments (GMM) Estimation for Panel Data (2025), `https://search.r-project.org/CRAN/refmans/plm/html/pgmm.html`
44. Windmeijer, F.: A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of econometrics **126**(1), 25–51 (2005)

45. Davidson, R., MacKinnon, J.G.: Econometric Theory and Methods. Oxford university press, New York (N.Y.) (2004)
46. Eurostat: Businesses in the mining and quarrying sector (2025), https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Businesses_in_the_mining_and_quarrying_sector
47. Staiger, D., Stock, J.H.: Instrumental Variables Regression with Weak Instruments. Econometrica **65**(3), 557–586 (1997)
48. Eurostat: Handbook for estimating raw material equivalents (2024), https://ec.europa.eu/eurostat/documents/1798247/6874172/Handbook-country-RME-tool
49. Roodman, D.: How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal **9**(1), 86–136 (2009)
50. European Environment Agency: Circular material use rate in Europe (2025), https://www.eea.europa.eu/en/analysis/indicators/circular-material-use-rate-in-europe

Appendix A

The following definitions are sourced directly from the Eurostat website. To access them, insert the Eurostat codes provided in Table 1 into the search bar, then navigate to the 'Dataset information' section. When additional details were needed to define the variables below, supplementary information was obtained from the metadata.

- Municipal waste per capita (MWAS)

  The indicator measures the waste collected by or on behalf of municipal authorities and disposed of through the waste management system. It consists to a large extent of waste generated by households, though similar wastes from sources such as commerce, offices and public institutions may be included. Reducing municipal waste generation is an indication of the effectiveness of waste prevention measures and changing patterns of consumption on the part of the citizens. Concentrating on municipal waste rather than on industrial waste has the advantage that it reflects the consumption side and is not affected by the presence or lack of strong manufacturing sectors in a country.
  This indicator focuses on municipal waste. Even though municipal waste only represents about 10% of the total waste generated or about 30% of the generated amount of waste excluding major mineral waste, following up on its evolution can give a good indication of changing consumption patterns and of Member States' waste prevention performance and where citizens' actions and involvement is most relevant. For the amount of municipal waste generated, the data refer to the handover over the waste to the waste collector or to a disposal site.

- Recycling rate of municipal waste (RECW)

  The indicator measures the share of recycled municipal waste in the total municipal waste generation. Recycling includes material recycling, composting and anaerobic digestion. The ratio is expressed in percent (%) as both terms are measured in the same unit, namely tonnes.

- Circular Material Use rate (CMU)

  The indicator measures the share of material recycled and fed back into the economy - thus saving extraction of primary raw materials - in overall material use. The circular material use, also known as circularity rate is defined as the ratio of the circular use of materials to the overall material use. The overall material use is measured by summing up the aggregate domestic material consumption (DMC) and the circular use of materials. DMC is defined in economy-wide material flow accounts. The circular use of materials is approximated by the amount of waste recycled in domestic recovery plants minus imported waste destined for recovery plus exported waste destined for recovery abroad. Waste recycled in domestic recovery plants comprises the recovery operations R2 to R11 - as defined in the Waste Framework Directive 75/442/EEC. The imports and exports of waste destined for recycling - i.e. the amount of imported and exported waste bound for recovery – are approximated from the European statistics on international trade in goods. A higher circularity rate value means that more secondary materials substitute for primary raw materials thus reducing the environmental impacts of extracting primary material.

- Employment in CE sectors (CE_EMP)

  The indicator measures 'Number of persons employed' in the following three sectors: the recycling sector, repair and reuse sector and rental and leasing sector. Jobs are expressed in number of persons employed and as a percentage of total employment. Number of persons employed is defined as the total number of persons who work in the observation unit, i.e. the firm (inclusive of working proprietors, partners working regularly in the unit and unpaid family workers), as well as persons who work outside the unit who belong to it and are paid by it - e.g. sales representatives, delivery personnel, repair and maintenance teams. It excludes manpower supplied to the unit by other enterprises, persons carrying out repair and maintenance work in the enquiry unit on behalf of other enterprises, as well as those on compulsory military service.

- Gross private investment in tangible goods in CE sectors (CE_INV)

  The indicator measures the gross private investment in tangible goods. Gross investment in tangible goods is defined as investment during the reference year in all tangible goods. Included are new and existing tangible capital goods, whether bought from third parties or produced for own use (i.e. capitalised production of tangible capital goods), having a useful life of more than one year including non-produced tangible goods such as land. Investments in intangible and financial assets are excluded.

- Gross value added by CE sectors (CE_VA)

  The indicator measures the gross value added by CE sectors. Value added at factor costs is the gross income from operating activities after adjusting for operating subsidies and indirect taxes. It can be calculated as the sum of turnover, capitalized production, other operating income, increases minus decreases of stocks, and deducting the following items: purchases of goods and services, other taxes on products which are linked to turnover but not deductible, duties and taxes linked to production. Value adjustments (such as depreciation) are not subtracted.

- Number of patents related to recycling and secondary raw materials (CE_PAT)

  The indicator measures the number of patents related to recycling and secondary raw materials. The attribution to recycling and secondary raw materials was done using the relevant codes in the Cooperative Patent Classification (CPC) (list of CPC codes selected). The term 'patents' refers to patent families, which include all documents relevant to a distinct invention (e.g. applications to multiple authorities), thus preventing multiple counting. A fraction of the family is allocated to each applicant and relevant technology.

Appendix B

Listing 1.1: R-code for original System GMM Estimation

```
gmm <- pgmm(log(DE) ~ lag(log(DE), 1) + log(GDP) + log(POP) + log(CONST_GDP) + log(DE_DMC) +
    log(RECW) + log(CMU) + log(CE_EMP)
                | lag(log(DE), 1:99),
                data = "data",
                effect = "individual",
                model = "twosteps",
                transformation = "ld")
summary(gmm, robust = TRUE)
```

Table 9: Random-effect model results

Coefficient	RE **(1b)**
GDP	**0.545** ***
	(0.076)
POP	**0.333** ***
	(0.086)
DE/DMC	**1.049** ***
	(0.104)
CONST/GDP	**0.160** **
	(0.075)
RECW	**0.051** *
	(0.029)
CMU	**-0.169** ***
	(0.043)
CE_EMP	-0.007
	(0.155)
intercept	-0.332
	(0.839)
R2	0.8927
R2 adjusted	0.8866
Panel structure	N = 24, T = 6, n = 132

Notes: Standard errors in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 10: Specification tests

Type of test	**H0**	**(1:1)** MWAS	**(1:2)** RECW	**(1:3)** CMU	**(1:4)** CE_EMP	**(1:5)** CE_INV	**(1:6)** CE_VA	**(1:7)** CE_PAT	**(1)** Full	**(1a)** Full	**(1b)** Full
F Test for individual and/or time effects based on the comparison of fixed and pooled effects models	No significant time and/or individual effects	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lagrange FF multiplier tests for individual and/or time effects based on the results of the pooling model (type Breusch-Pagan)	No significant time and/or individual effects	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Hausman model specification test: based on the comparison of random and fixed effects models	No correlation between the unique errors and the regressors	0.00	0.00	0.05	0.00	0.00	0.02	0.00	0.00	0.00	0.66
Breusch-Pagan Heteroskedasticity test	Presence of homoskedasticity	0.07	0.05	0.00	0.05	0.00	0.00	0.01	0.00	0.00	0.00

Notes: Reported values are p-values.

Table 11: Two-step System GMM estimation results under alternative specifications

Coefficient	**(1)** Original System GMM 2010–2019	**(2)** Second-third lag instruments 2010–2019	**(3)** Collapsed full instruments 2010–2019	**(4)** Collapsed second-third lag instruments 2010–2019	**(5)** Extended & updated data 2010–2022
GDP	**0.307** ***	0.348	**0.656** **	**0.559** *	0.110
	(0.105)	(0.366)	(0.269)	(0.323)	(0.079)
POP	**0.105** **	0.015	-0.173	-0.173	0.027
	(0.042)	(0.274)	(0.199)	(0.180)	(0.098)
DE/DMC	**1.117** ***	0.778	**1.381** *	**1.186** **	0.141
	(0.319)	(0.824)	(0.726)	(0.597)	(0.210)
CONST/GDP	**0.183** ***	0.171	**0.217** **	0.134	0.082
	(0.056)	(0.171)	(0.097)	(0.129)	(0.077)
RECW	0.033	-0.064	-0.069	-0.090	-0.060
	(0.045)	(0.243)	(0.166)	(0.119)	(0.067)
CMU	**-0.112** **	-0.053	-0.131	-0.075	-0.017
	(0.050)	(0.091)	(0.090)	(0.135)	(0.061)
CE_EMP	**-0.338** ***	-0.143	-0.017	0.091	-0.008
	(0.099)	(0.607)	(0.465)	(0.364)	(0.137)
Lag 1 DE	**0.545** ***	**0.632** **	**0.571** ***	**0.676** ***	**0.861** ***
	(0.142)	(0.291)	(0.122)	(0.215)	(0.166)
One- or two-step	two-step	two-step	two-step	two-step	two-step
Endogeneity adjustment	no	yes	yes	yes	yes
Lag structure	1:99	2:3	$2{:}99^{collapsed}$	$2{:}3^{collapsed}$	$2{:}3^{collapsed}$
AR (1) p-value	0.0109	0.0306	0.0091	0.0091	0.0018
AR (2) p-value	0.1140	0.1610	0.1913	0.1474	0.0896
Sargan test p-value	1.0000	1.0000	1.0000	0.1559	0.0339
Wald test p-value	0.0000	1.0000	0.0000	0.0000	0.0000
Number of instruments	56	122	52	22	22
Panel structure	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 28, T = 10, n = 321	N = 27, T = 13, n = 609

Notes: Standard errors in parentheses. ***** $p < 0.01$; ****** $p < 0.05$; ***** $p < 0.1$. The lag structure is defined with respect to period t. The finite sample correction for standard errors proposed by Windmeijer (2005) is applied [44]. No time effects included.