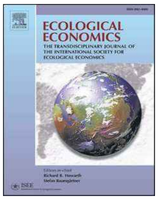
Ecological Economics 215 (2024) 108000

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate/ecolecon



Inconsistent definitions of GDP: Implications for estimates of decoupling

Gregor Semieniuk *

Political Economy Research Institute and Department of Economics, University of Massachusetts Amherst, United States of America

ARTICLE INFO

Keywords:
Decoupling
National accounting
GDP revisions
Energy intensity
Environmental Kuznets curve

ABSTRACT

Efforts to assess the possibilities for decoupling economic growth from resource use and negative environmental impacts have examined their historical relationship, with varying and inconclusive results. This paper shows that ambiguities in the historical measurement arising from definitional changes to GDP are sufficiently large to affect the results. I review the history of structural revisions to GDP using the example of the United States, and on international comparisons of purchasing power parity, and compare decoupling results using GDP vintages reported between 1994 and 2021 for most countries. Between vintages, 10–15% of countries switch between relative decoupling and recoupling from energy or materials on decadal intervals, and up to as many countries as decouple absolutely in an older vintage stop or newly start absolutely decoupling in the newer vintage. GDP vintages also affect environmental Kuznets curve results on absolute decoupling in Grossman and Krueger's seminal paper and accelerate the International Energy Agency's annual global decline in energy intensity by up to −0.2 percentage points. Inconsistencies in economic measurement introduce ambiguity into historical decoupling evidence and model projections into the future. To advance debate, rigorous reporting and sharing of data vintage for subsequent comparison and replication are urgently needed.

1. Introduction

Understanding the history of the relationship between gross domestic product (GDP) and resources and environmental impact is important for thinking about the future. On the question whether GDP growth is feasible with nongrowing or even declining rates of resource use and impacts, such as greenhouse gas emissions, pivots whether the current global mode of social provisioning and reproduction couched around expansion of economic value can continue for the next decades and centuries. The relationship between resources and environmental impacts and GDP is often expressed as an intensity with GDP in the denominator. A declining intensity is referred to as relative decoupling. If, in addition, the numerator falls while GDP grows, one speaks of absolute decoupling. Absolute decoupling is ultimately necessary for continued economic growth on a finite planet.

Given the stakes, these patterns understandably receive immense research attention; a recent review examined 835 empirical studies of decoupling (Wiedenhofer et al., 2020). Yet, despite this intensive empirical research, there is no consensus on what is and isn't feasible (Hickel and Hallegatte, 2022). On one side "green growth" advocates claim that absolute decoupling in certain dimensions by a growing number of countries portends more such absolute decoupling in the immediate future. On the other, many ecological economists retort that since the vast majority of observed country-years and resource and impact dimensions and the world as a whole show no absolute and often not even relative decoupling, the current mode of economic growth is unsustainable. The debate is complicated by competing intensity measures for the same resource or environmental impact, e.g. whether to use territorial, footprint or income measures[1] or how to account for primary energy.[2] While these conceptual arguments all pertain to the numerator of intensities, this paper suggests that additional ambiguity may have arisen from redefinitions of the denominator of any intensity, GDP.

GDP is an accounting convention. Its measurement depends on social agreement, not on natural constants. In fact, at any moment there is

---

* Corresponding author.
E-mail address: gsemieniuk@umass.edu.

[1] Territorial measures allocate resource use to the country where resources enter production processes, footprint measures where final products are consumed, and income measures where the resource use enables factor payments: wages and profits. (Akizu-Gardoki et al., 2021; Jiborn et al., 2018; Marques et al., 2012; Peters and Hertwich, 2008; Tukker et al., 2020; Weber et al., 2022; Wiedmann et al., 2015).

[2] For primary energy, at least three major accounting methods are widely used, leading to inconsistent energy intensities because non-combustion energy sources are assigned varying primary energy values (Koomey et al., 2019; Kraan et al., 2019; Krey et al., 2014; Macknick, 2011; Semieniuk and Weber, 2020).

https://doi.org/10.1016/j.ecolecon.2023.108000
Received 21 September 2022; Received in revised form 14 August 2023; Accepted 19 September 2023

more than one GDP measure available. More importantly, over time these measures are revised and definitions changed and then applied to the whole retrospective GDP data series as national accountants themselves extensively document (Kendrick, 1970; Studenski, 1958; Vanoli, 2005). Rather than scrutinizing how these definitional changes impact the relationship between GDP and environmental impact, the ecological economic debate has instead focused on enlarging GDP with measures of environmental quality or quantity (Costanza et al., 1997; Hanley et al., 2015). Complete alternatives to GDP have been proposed (Hoekstra, 2019), as well as post-growth indicators (Victor, 2008). Yet, since the decoupling question continues to revolve around intensity measures involving plain-vanilla GDP, a better understanding of how changes in its definition impact whether or not a country is seen to have relatively or absolutely decoupled in the past is vital. In this paper I review how GDP measures regularly undergo revisions; empirically analyze consequences of these revisions for past evidence on relative and absolute decoupling; and discuss implications for future analysis.

GDP is revised for a variety of reasons. One cause is the updating of 'base years' to get a handle on inflation, another is a redefinition of what constitutes 'production' suggested by economic theory or historical experience, yet another the availability of new data series. To study the consequences of revisions empirically, I collect vintage data of U.S. GDP going back to the 1960s from the archive of the *Survey of Current Business*. By vintage I mean the then-available GDP time series published in a certain year in the past, e.g. a vintage published in 1965, that reports GDP from some initial year that is typically common to all vintages to the last available year at the time of reporting, which would likely be 1963 or 1964 in a 1965 publication. I also collect vintages for a large number of countries reported between 1994 and 2021 from the national accounts data in a supplement to successive versions of the Penn World Table (PWT). Finally, I collect global GDP from the last 10 editions of the International Energy Agency's (IEA) World Indicators that rely on the World Bank's estimate of the whole world's GDP. I combine these times series for GDP with data on primary energy for most countries for the period 1950–2014 and for material use for 1970–2019 to examine changes to relative and absolute decoupling outcomes for countries contingent on data vintage used. I also re-estimate the random effects model in Grossman and Krueger's (1995) seminal paper on the so-called environmental Kuznets curve in a panel of countries with later GDP vintages to check the robustness of absolute decoupling implied by the environmental Kuznets curve. Finally I examine how the IEA's GDP vintage data changes our understanding of the decline of the historical global energy intensity.

The aim of this research is to make researchers and users aware that relative and absolute decoupling results are, to some extent, contingent on the GDP definition applied, as the results will show. Therefore, studies using different GDP vintages even for the exact same set of countries and years are not directly comparable in their conclusions. Even a few years difference in the collection of GDP data can impact measurement. One strategy is to assume progress over time in national accounting definitions and prefer more recent over older results. However, national accounts are conventions of their times and require value-judgements about what are the most important questions to answer. As such, there are no strong epistemological arguments for preferring only the latest GDP vintage for analysis of economic activity sometimes decades in the past. The existence of different vintages of GDP introduces an unresolvable ambiguity over past patterns of growth and hence the attempt to 'get right' the historical evidence on decoupling. This contributes to the persistence of the disagreement about what kind of growth is possible now or in the future and must be acknowledged when modeling the future based on this evidence. To advance insight, debate and scenario modeling, rigorous reporting of GDP definitions, vintage, and the sharing of data for subsequent comparison and replication is urgently needed. Policy targets need to be either precise about the data vintage against which to measure their decoupling or avoid reference to aggregate intensities and focus directly on absolute resource or impact decline.

The next section reviews GDP revisions both at the national and international level at some length with a view to familiarizing researchers in ecological economics with them. It also gives examples of how GDP revisions impact its magnitude and rate of change, and reviews related literature. Section 3 introduces the method of analysis and all data sources. Section 4 presents results of the impact of GDP vintage on relative and absolute decoupling and section 5 discusses them. Section 6 concludes with three recommendations for research and policy making.

2. GDP revisions and related literature

GDP is part of any analysis of decoupling. Decoupling refers to the relationship between rates of change of resource inputs or environmental impacts, call all of them R, and a measure of economic activity, typically GDP. *Relative decoupling* occurs when the proportional rate of change of GDP over a certain period is greater than the proportional rate of change of R in the same period, and so the intensity, R/GDP, falls. A GDP growth rate of 3% versus an R growth rate of 2% is relative decoupling, and a GDP rate of decline of −3% versus an R rate of decline of −5% is also relative decoupling. *Absolute decoupling* occurs whenever GDP grows and R declines. The complement of these cases is called *recoupling*.[3] The rest of this section examines how GDP for all past years changes over time and how there is more than one such time series in use at any time, thereby influencing intensities and decoupling.

GDP revisions happen all the time. Every quarter and year, statistical agencies first produce preliminary estimates based on incomplete data and projections, which are revised as better and more data become available (Fixler et al., 2021; Van Walbeek, 2006). Past research documented an upward bias in some such revisions (Franses, 2009; Glejser and Schavey, 1969). However, these short-term revisions are not the subject of this paper. Instead, the focus is on structural revisions to the national accounting framework, that do not happen simply because new information becomes available in the months and sometimes years after the first estimate. Such structural revisions involve changes in the accounting conventions used to select and aggregate data. That is, they go beyond mere revisions and completion of the most recent data. In particular, they involve changes in aggregation methods, base years for indices, and definitions of GDP (Croushore and Stark, 2003). This phenomenon has variously been referred to as 'general revision' (Siesto, 1987), alteration of the 'architecture of the national accounts' (Jorgenson, 2009), simply 'changes' or 'improvements' to national accounts (Moulton, 2004) or revision of the system, not just the series (Ruggles, 1990). To avoid confusion with the widely used term 'revision' for successive estimates of the latest data, and following the Croushore and Stark (2003) terminology, they will be called *structural revisions* here. Section 2.1 illustrates the impact of these revisions using U.S. data. Section 2.2 briefly illustrates the better-known changes to purchasing power parities between countries after explaining the concept. Section 2.3 reviews related literature.

2.1. Structural GDP revisions: Example of the US

Structural revisions can be roughly attributed to three causes: reference year changes, redefinitions, and data source changes.[4] Appendix A reviews each of them in detail. Here I show how structural revisions continually take place and impact GDP measurement, using

---

[3] More detailed partitions of the growth rate space are used e.g. by Naqvi and Zwickl (2017), but the simple partition made here is sufficient to illustrate the impact on results of GDP revisions.

[4] One could distinguish more causes. In his magisterial treatise of three centuries of national income estimates, Studenski (1958) already identified altogether eleven reasons for advances and changes in estimates and accounting.

the US example.

Statistics from the United States National Income and Product Accounts, one of the most detailed and long-lived systems for recording the aggregate economy, are published monthly via the *Survey of Current Business*. The August 1965 issue reports gross national product or GNP, used in the United States until 1991 instead of its close relative, GDP.[5] The *Survey* then states in a section titled *Definitional Changes,* that while there is general agreement on how to define GNP, "[d]efinitional revisions continue to suggest themselves as the result of further thought […] and also as the result of improvements in data sources that permit the implementation of more appropriate definitions and concepts." (p. 7). The section goes on to stress that the disagreement about the exact definitions in national accounts resemble debates in social or natural sciences, and quickly adds that it "is reassuring to note that the definitional changes that have been made in this report do not greatly affect our measure of the total size of the national output, [and] of its long-term growth". This reassurance reveals that the redefinition has changed (even if not greatly) both level and rate of change of GDP.

Forty years later, Brent Moulton, the head of the national accounts program at the US Bureau of Economic Analysis, which curates and publishes the US GDP figures, enumerates shortcomings and controversies of the GDP definition (Moulton, 2004). He criticizes not GDP in its 1965 guise, but the United Nations' 1993 System of National Accounts (SNA), an international benchmark for how countries should account for GDP. The 1993 SNA itself included recommendations for substantial redefinitions to GDP compared to the SNAs from 1953 and 1968. Among other things, Moulton criticizes the calculation of return on non-market government investment, the treatment of R&D and of expenditures on military assets as a cost rather than an investment, and certain aspects of measuring financial services (Moulton, 2004). Since GDP growth is a weighted average of its components, changing any component's weight impacts GDP growth, too. Suppose for instance that GDP was revised to feature a larger government activity as a share of GDP, e.g. by imputing a return to non-market government investment (such as into public schools) and imputing it also for all past years for consistency. Then if we further suppose that government activity expanded more slowly than the rest of the economy, GDP will suddenly be found to have grown more slowly in the past.

Another edition of the SNA was released in 2008 and took onboard some of the issues Moulton had raised.[6] Since the UN's SNA serves as a benchmark for internationally comparable national accounts, it was also implemented in the US national accounts in 2013. Subsequently economists have continued to worry about systematic biases downward (Feldstein, 2017) or upward (Tercioglu, 2021) in the U.S. growth rate and the accurate measurement of innovation and intangibles (Corrado et al., 2021). The US example illustrates that GDP gets structurally revised over time, including via conceptual redefinitions. Such revisions have an impact on the growth rate, and there is no end to future revisions in sight.

To get a feeling for the impact of structural revisions on the measurement of U.S. economic growth, consider Fig. 2. Panel (a) shows that the definition of GDP agreed in 1980 indicates that the size of the US economy quadrupled between 1929 and 1986. However, when using the BEA's GDP current as of this research (3rd quarter of 2020), which has seen a structural revision of GDP most recently in 2018, the economy has grown sixfold over the same time period. Growth was cumulatively more than 40% faster as the orange series measured on the right-hand axis shows. There is some volatility in the late 1940s when the US economy ended war planning and price controls (Weber, 2021), but otherwise we see a fairly steady escape of modern GDP from its historical counterpart. The series cannot be compared after 1985. This is due to another structural GDP revision in 1987, which replaced the reporting of the 1980 revision. But if one were to revisit any publication that used GDP data reported before 1987 it would deliver a strikingly lower GDP growth rate, and hence fewer prospects for relative decoupling than with current GDP estimates.

Since all rates of change are positive, the presence or absence of absolute decoupling in any one vintage would carry over to all of them. However, it is notable that growth rates in the 2000s have declined in all vintages relative to previous decades, while the differences in growth rates between vintages persists. If this secular decline were to continue as some predict (Gordon, 2017), it is possible to imagine a situation where changes from vintage to vintage determine whether the country's GDP grows or declines and thus absolutely decouples or not. As we will see below, this flip in the sign of the GDP rate of change can already be observed in the past for a substantial share of those countries reporting absolute decoupling in at least one vintage.

The US national accounts have gone through a total of 15 structural revisions, occurring about every 5 years and starting in 1947. Panel (b) shows how growth rates vary across several of them (all using GNP, since the 1965 revision did not yet report GDP). Growth almost always accelerated retrospectively, from one revision to the next. The differences are typically not very large, year-on-year, but over decades compounded exponential growth adds up to sizeable differences. It is worth stressing that the acceleration often arises from including more and more monetary transactions inside the production boundary whose justification tends to be associated with neoclassical economic theory (Mazzucato, 2018). The latter has, at least since Alfred Marshall's influential analysis (Marshall, 1890 especially bk. 2, ch. 3), endorsed a comprehensive production boundary based on subjective valuation. As such, national accounts tend to interpret any voluntary monetary transaction as adding value (Foley, 2013). Other economic schools, like the classical one, would restrict the production boundary and slow down the growth rate (Basu and Foley, 2013; Shaikh and Tonak, 1994). One important inconsistency in the otherwise comprehensive production boundary however is the exclusion of non-market labor, which includes much of the world's care work (Folbre, 2020). More detail on theoretical reasons for varying the production boundary is in Appendix A.

The relatively good availability of vintage US GDP data (see next section for data sources) makes it convenient to analyze US GDP. However, it is important to realize that similar structural revisions happen in other countries (Bos, 2006; Vanoli, 2005), guided by the same SNA editions, which are in turn informed by economic theory.

As an illustration of how such revisions can have a qualitative impact on relative decoupling, consider the path-breaking work by Schurr et al. (1960, Figure 26a) on US commercial energy intensity, finding an inverted u-curve peaking in 1915, and updated by Schurr (1984).[7] Schurr describes a period of relative decoupling from 1930 to about 1945, followed by stability until 1970 and then again relative decoupling (Fig. 2, black series). However, trends calculated using modern data show more consistent and rapid relative decoupling (Fig. 2, blue series).[8] Moreover, the historical data show 3 five-year intervals with

---

[5] GNP of any country measures what domestic labor and domestically-owned capital earn anywhere in the world. GDP measures earnings on the country's territory, regardless of the earner's nationality. In the US, the switch to GDP occurred to adjust to international convention. For our purposes, and I thank an anonymous reviewer for pointing this out, the territorial boundary of GDP is preferred as the resource or environmental impacts considered here are similarly territorially defined.

[6] Moulton served as a member of the Advisory Expert Group to the 2008 SNA edition.

[7] For even earlier work on the relationship between US energy and economic activity in the 1920s and predating national accounts, see Tryon's work discussed in Missemer and Nadaud (2020).

[8] The blue series assumes Schurr (1984) used the US national accounts estimates incorporating the 1980 revision also for his data pre-1960. These data had already appeared in the earlier 1960 study (Schurr et al., 1960). Both series in Figure 1 use on GNP, the common output measure in the US until 1991.

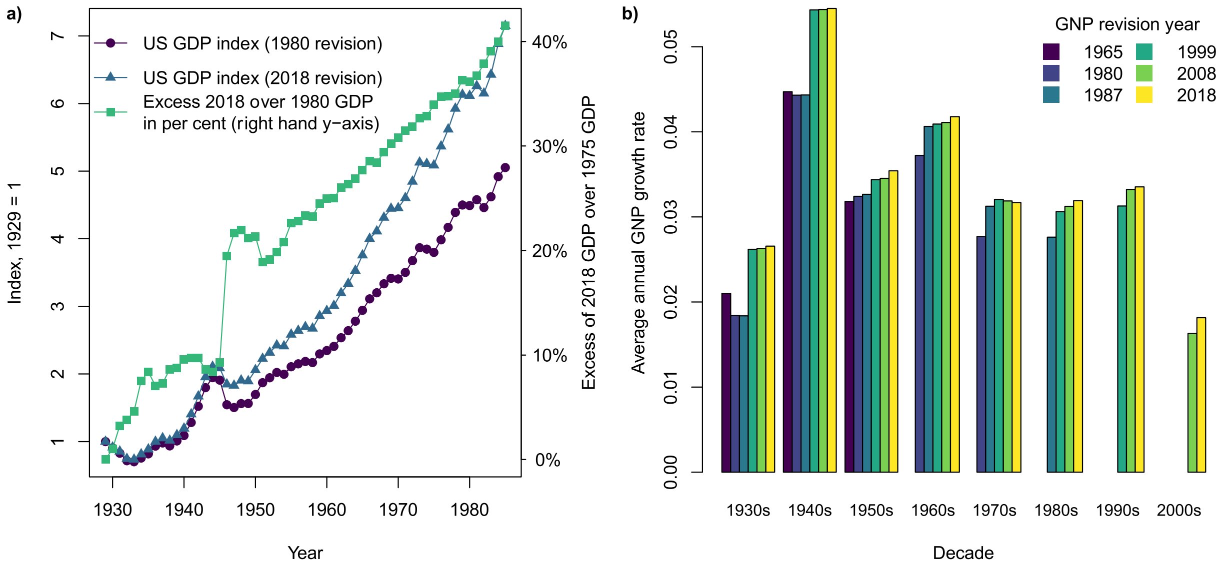


Fig. 1. Changes in growth rates between vintages of US GDP or GNP: (a) GDP indices from 1986 (black) and 2021 (blue) based on 1976 and 2018 structural revisions, and their ratio in orange measured on the right-hand side axis. (b) Annual average growth rate of GNP over decades of selected structural revisions, data from last year before next revision. Sources described in section 3.

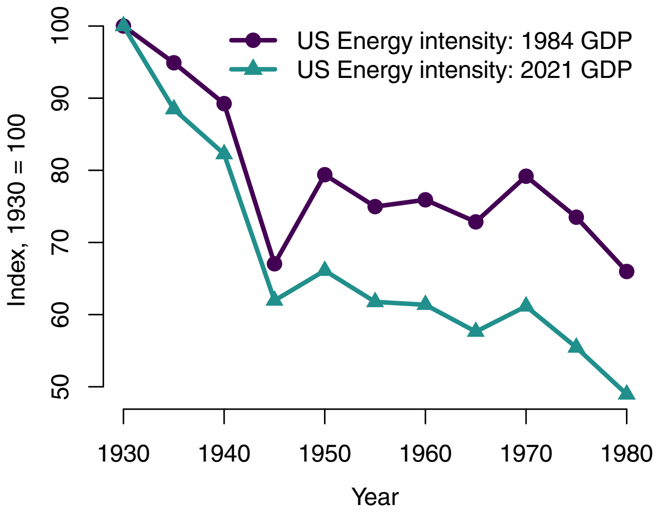


Fig. 2. Impact of US GDP vintage on energy intensity. Primary energy intensity as in Schurr (1984) but adding the 2021 vintage of US GDP. Sources described in section 3.

recoupling, the modern data only 2. Using exactly the same energy data as Schurr and output data collected just a few decades later, energy intensity falls 1.5 times faster over the entire 50-year period. GDP growth rates that are supposed to characterize one and the same economy vary systematically over time.

2.2. GDP revisions for international comparisons: PPPs

So far the discussion has looked at national revisions. Another level of complexity is introduced by international comparisons. The most straightforward approach to comparing countries' GDPs is to use market exchange rates (MER) that can be readily gleaned from currency markets and data repositories. However, economists have long debated over whether this is the appropriate approach (Kravis et al., 1982). In particular, since international comparisons are often made with an aim of assessing the relative standards of living, the question of what one can buy with one's money looms large. GDP converted at market exchange rates gives an incomplete answer to this question because market exchange rates are formed from the demand and supply of traded goods and from international financial market transactions. Many goods and services that determine one's standard of living aren't traded internationally and hence one currency may not buy the same amount of them in another country after exchanging at market exchange rates. Therefore, similar to stripping out inflation to compare a country's growth over time, 'purchasing power parity' (PPP) has attempted to adjust countries' economic activity for varying cross-sectional price levels (see Appendix B for an example).

Measuring PPP exchange rates in practice is difficult and involves many choices. A key problem is that people in different countries don't consume the same goods and so estimations of what prices need to be adjusted are fraught with assumptions (Reddy, 2008). The assumptions made as well as alternative PPP methods have been reviewed, e.g. in Anand and Segal (2008) or Deaton and Heston (2010). One result of these difficulties is that calculating PPPs necessitates the largest global statistical effort, carried out every roughly half decade by the *International Comparison Program*. Good reviews of recent rounds of the program are in Deaton and Aten (2017) and Deaton and Schreyer (2022). Reasons behind revisions over time are discussed in Deaton and Heston (2010) and in Feenstra et al. (2013). Some argue using PPP GDP for cross country comparison is not a good idea in the first place, or at least not the preferred or only measure (Acemoglu et al., 2019; Ghosh, 2018), or should depend on the application (Semieniuk et al., 2023).

The key issue for this paper is that every new international comparison exercise creates a new set of exchange rates, new GDP levels and, to some extent, growth rates. This is not because national accounts have changed their definition but because consumption baskets and prices have changed, and the method of operationalizing PPP has as well – a structural PPP revision, so to speak. In addition, regional and global growth rates are impacted by the change in country weights as GDP levels change relative to each other (the same is true of GDP measured at varying market exchange rates). Section 3 will elaborate as necessary.

To get a sense of magnitudes involved in level changes, Fig. 3 plots per capita income in PPP 'international' dollars for four countries for the year 2000 as measured in 5 versions of the Penn World Table. One can readily see that the level varies considerably. Not only that, but it also changes in idiosyncratic ways for every country. While India sees a steady decline, Mexico sees growth, except in one revision, and China and South Africa depict an undulating movement across versions. Our

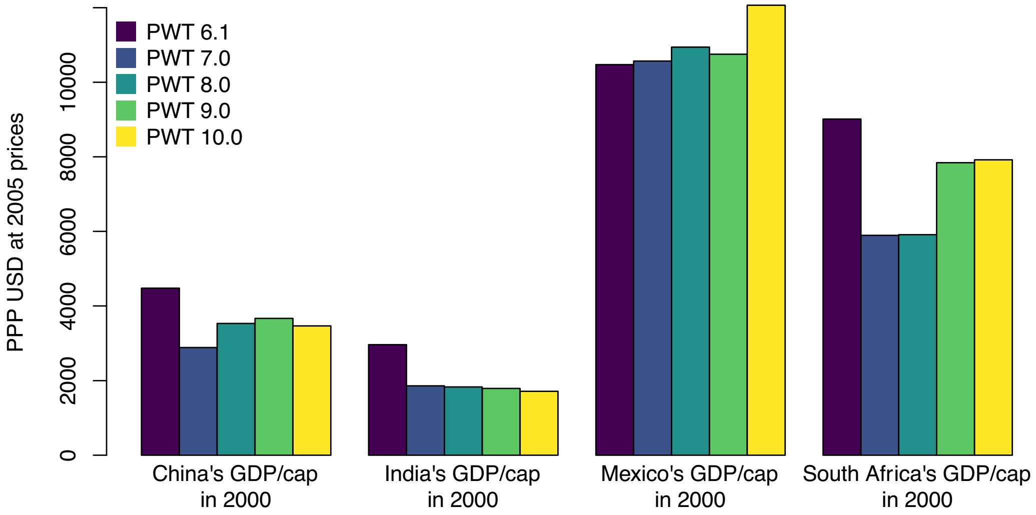


Fig. 3. PPP GDP per capita in 2000 according to successive Penn World Table estimates: PPP GDP per capita in 2000 for four countries at 2005 USD price levels, taken from 5 versions of Penn World Tables published over the period 2002–2020. Sources described in section 3.

focus is on decoupling and rates of change. We cannot examine levels in detail.

To further illustrate how, even at one point in time, PPP exchange rates can introduce an additional GDP measure with consequences for relative decoupling, consider Hickel and Kallis' (2020) question "Is Green Growth Possible?". They answer with a forceful "no", using selected quantitative evidence in their argument. A dramatic piece of this evidence is a graph with indices of global material use and GDP from 1990 to 2017. The material use index grows faster than GDP and therefore the world has recoupled precisely in that period where more attention was devoted to decoupling than perhaps ever before. The global GDP source is reported as "World Bank". The authors apparently rely on MER GDP rather than the World Bank's PPP rates, without explaining their choice. While it may not be clear which choice is the best one, it has qualitative implications for the conclusions drawn. If GDP is measured instead at PPP, the entire period is one of relative decoupling. Fig. 4 replicates their graph, but also adds the PPP GDP index, according to which material intensity has declined to about 80% of its 1990 value. Since both types of GDP have their reasons for being used (and PPP GDP is now much more widely used for global analysis), it is just not unambiguously possible to claim that there has been recent recoupling with material use.[9] Of course, relative decoupling with PPP GDP is no confirmation of 'green growth', which would additionally require a demonstration of eventual absolute decoupling at the global level.

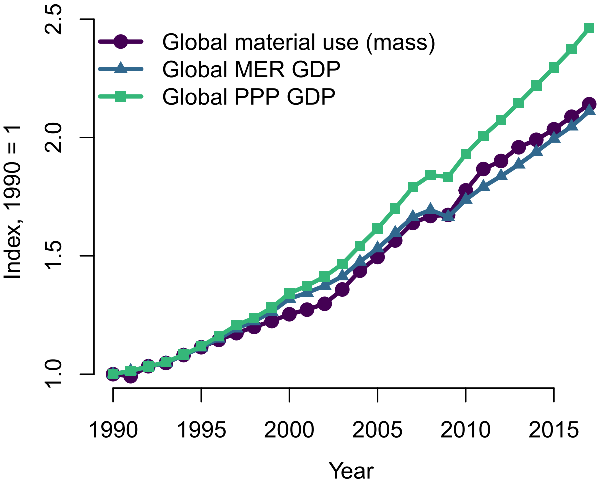


Fig. 4. GDP data definition effect on relative decoupling: Global material consumption vs GDP indices reproducing Hickel and Kallis (2020) but also showing GDP at purchasing power parity. Sources described in section 3.

2.3. Related studies

A few studies with environmental concerns have analyzed certain features of sectoral price indices. Kander (2005) highlights the importance of accounting for sectoral output either in real or nominal prices. At nominal prices, services attain a larger and larger share in output. Since services tend to have lower energy intensities, this drives overall relative decoupling. However, since the prices of produced goods do not rise as fast as those for services or even fall due to productivity gains, the volume of services (measured at sectorally deflated prices) does not necessarily rise as a share of output (see also Kravis et al., 1983; Tregenna, 2009). Therefore expecting a low energy intensity simply because of a nominally measured large service economy is misguided. Kander provides long-run evidence for Sweden, and Henriques and Kander (2010) show that service sector transition led to only modest declines in energy intensity for a larger set of countries. Witt and Gross (2020) document a similar result for Germany.

Looking at aggregate, not sectoral data, Stern (2017) compares the performance of past forecasts of global energy intensity from a series of IEA *World Economic Outlooks* with the historical times series of that intensity available in 2017 and finds that they overestimate the decline in energy intensity. To the extent that the GDP differs between the historical forecasts and the time of analysis, this may affect Stern's evaluation of the forecasts.

The climate change modeling community has debated the impact on growth rates of the use of MER or PPP GDP. In the early 2000s, Castles and Henderson (2003) attacked the IPCC's *Special Report on Emissions Scenarios* (Nakićenović et al., 2000) as showing "technically unsound" scenarios due to their use of MER GDP. Their most salient criticism for present purposes centered on the combination of MER GDP with a convergence assumption, i.e., less affluent countries grow to 'catch up' with richer ones in GDP per capita terms. Since the gap between GDP per capita in rich and in developing countries was larger with MER GDP, this led to very high growth rates in developing countries. The report's

---

[9] The same caveat applies e.g. to the results shown in Fig. 1 in Wiedmann et al. (2020).

authors retorted that they were modeling economic activity, not standards of living, among other arguments (Nakićenović et al., 2003). The debate went on for several years and its intensity is showcased by Nordhaus (2007) who argued in favor of using PPP, calling the use of MER "fundamentally wrong" as understating the income of developing countries, and the refutation of his work by Pant and Fisher (2007) based on the argument that higher market prices in rich countries may include the funding of more abundant public goods. The debate and references are reviewed in Pitcher (2009) who is the only one to my knowledge to note that the revision of historical PPP levels (but not growth rates) had an impact on subsequent modeling. Ultimately, PPP GDP became the measure of choice. The 5th Assessment Report in 2014 already used PPP GDP to calculate intensities, reporting MER GDP only in the online databases. With retrospect, there are two ironies to this debate. First, while the debate was kindled by differences in MER and PPP growth rates, the SSP scenarios used in the current IPCC assessment cycle appear to assume MER and PPP GDP growth rates to be equal (Leimbach et al., 2017). Second, the problem of incredibly fast growth rates really only occurred due to assumptions about convergence, for which there is little evidence in the historical growth record (Johnson and Papageorgiou, 2020).

Finally, there is the burgeoning literature on the environmental Kuznets curve, which started out with publications in the early 1990s that claimed to find an inverted u-shape (with GDP/capita on the x-axis and pollution concentration on the y-axis). Due to its policy message that growth itself is the solution it has generated a lot of scrutiny by researchers, which continues to this day (Carson, 2010; Dinda, 2004; Özcan and Öztürk, 2019; Sarkodie and Strezov, 2019; Stern, 2004). The seminal paper that brought the concept into the academic mainstream is Grossman and Krueger (1995).[10] It has been subjected to considerable scrutiny. For instance, Harbaugh et al. (2002) showed that varying the extent of the dataset or the pollution measure definition invalidates the results. Torras and Boyce (1998) showed that adding co-variates on power inequality to the original dataset tempers the inverted u-curve relationship. However, despite the considerable controversy and debate about data sources (e.g. in Stern, 2004), there has never seemed to be a question about the GDP data itself. Kacprzyk and Kuchta (2020) use an 'alternative' GDP to re-estimate environmental Kuznets curve relationships. Rather than comparing vintages of GDP, however, they estimate GDP based on night-time lighting. In sum, despite the important role of structural GDP revisions, to my best knowledge the question of how this relates to either relative or absolute decoupling estimates has not been systematically analyzed. Methods and data for doing so are introduced next.

3. Methods and data

The most straightforward way to examine the impact of GDP revisions on decoupling is to change the GDP data, all else equal, and analyze the impact on rates of change of intensity or GDP. I make pairwise comparisons of the same measure over business-cycle length (10 year) intervals using different vintages of data. Since rates of change are dimensionless this method of analysis easily spans every possible combination of GDP definitions. To examine relative decoupling, I simply scatter plot intensity rates of change in earlier and later vintages; data in quadrants 2 and 4 switch from and to relative decoupling retrospectively. To study absolute decoupling, I identify all countries that absolutely decouple in the earlier vintage and then compare with the set of countries that absolutely decouple in the later vintage. Some countries may no more absolutely decouple, others newly so: this latter analysis is entirely contingent on GDP growth rates changing sign over 10-year periods. For n different vintages, the possible combinations of vintages are $n(n-1)/2$. Given $n = 11$ vintages of GDP in the Penn World Table (PWT) and 55 possible combinations, I reduce dimensions by only comparing with the most recent vintage (PWT 10.0) that would be used by a researcher collecting data today and by only picking a sample of the 10 older vintages that covers each of the other four ICP rounds.

Another method to analyze the robustness of decoupling results is to re-estimate existing studies' results with varying GDP vintages. For absolute decoupling, I pick the seminal environmental Kuznets curve study by Grossman and Krueger (1995), which did more than any other paper to solidify the idea of falling pollutant concentrations as countries get richer. I reestimate their curves using their Stata programs and scaling all GDP vintages to 2017 US-dollars. I replace their GDP (rgdpch) data from the Penn World Table Mark 5 (Summers and Heston, 1991) with the same years taken from the 7.0 and current 10.0 vintages, without making any other modification to their data. One problem is that the rgdpch measure has been superseded by different GDP measures after PWT 7 and isn't available in PWT 10.0. PWT 10.0 instead reports the measures rgdpe, rgdpo and rgdpna (where the "r" always refers to "real" or inflation-adjusted GDP). Both rgdpe and rgdpo use measurements from all ICP rounds to establish GDP in those years (benchmark years) and interpolate between them, an innovation from PWT 8.0. GDP growth rates are therefore determined, in part, by the different measurement conventions used in different ICP rounds. Rgdpe measures GDP from the expenditure side, rgdpo from the output side. Rgdpna extrapolates from the 2017 benchmarks using national accounts growth rates. While the PWT creators do not recommend using rgdna for both cross-country and time comparison, it is closest to rgdpch, so I report it for comparison (Feenstra et al., 2015, p. 3165). As a result, I reestimate the results with all 3 measures from PWT 10.0.

For relative decoupling, I compare different versions of the widely used historical global relative decoupling estimates for energy by the IEA (2022). Since several countries are involved in both examples, both MER and PPP GDP are candidate measures. For the Grossman and Krueger results, I follow them in using PPP GDP. For the IEA, I analyze both as both time series are available.

PWT GDP is from the website of the Groningen Growth and Development Centre (Feenstra et al., 2015).[11] It reports GDP for a large number of countries in altogether eleven vintages each starting with data in 1950 and running until a few years before the release (see

Table 1
Overview over PWT vintages.

Version	Release year	Last data year	Price year (ICP round)
PWT 5.6	1994	1992	1985
PWT 6.1	2003	2000	1996
PWT 6.2	2007	2004	1996
PWT 6.3	2009	2007	2005
PWT 7.0	2011	2009	2005
PWT 7.1	2012	2010	2005
PWT 8.0	2013	2011	2005
PWT 8.1	2015	2011	2005
PWT 9.0	2016	2014	2011
PWT 9.1	2018	2017	2011
PWT 10.0	2021	2019	2017

Note: OECD countries can have other price years. In PWT 6.2 for instance price data for 30 OECD countries is from 2002. ICP = International Comparison Project.

---

[10] The paper had over 9000 citations on Google Scholar in August 2022, and is the authors' second/first-most cited paper, followed by their more preliminary study on the same topic (Grossman and Krueger, 1991). The Google Scholar profiles are available at: https://scholar.google.com/citations?user=f46No0UAAAAJ&hl=de&oi=sra for Grossman and https://scholar.google.com/citations?user=5fY6_jMAAAAJ&hl=de&oi=sra for Krueger.

[11] https://www.rug.nl/ggdc/productivity/pwt/earlier-releases

Table 1 for an overview). To derive national accounts GDP growth rates, not those of GDP at PPP, I used the accompanying national accounts data. In particular, I calculated GDP_t for every year by summing $CHKON_t$, $GKON_t$, $IKON_t$, $EXPK_t$ and subtracting $IMPK_t$. I calculated average annual growth rates, g, from time t over s years as $g_{t,t+s} = (GDP_{t+s} / GDP_t)\hat{}(1/s)$-1. This allows focusing on the impact on national growth rates of the revision to GDP by national accountants, rather than the variation in relative GDP levels due to international price comparisons.[12]

Besides the PWT, I use various other data sources. The US GNP data shown previously are from various issues of the *Survey of Current Business* (*SCB*). Some issues describe a structural revision (called comprehensive update) that was just completed. These issues then report revised GDP data series back to 1929. The issues just before that will report the last GDP in the old version, and all the way back to when it was first reported. By joining first and last reports of a particular comprehensive update, it is possible to construct complete series of GDP of one vintage. For instance, there were revisions of GDP in 1965 and in 1970. Thus, GDP data were collected from the *SCB* August 1965 issue for the years 1929–64 and from the *SCB* July 1969 issue for 1965–68. Data were extracted from *SCB* pdfs on the BEA website, read into Excel using Adobe's text recognition software and checked and brought into a table format. From 2003, vintage GDP data are available readily in Excel format on the BEA website. In total this gives 15 GDP series, one for each structural revision.[13]

Global GDP data is from the IEA World Indicators using both MER and PPP. The IEA makes vintage datasets available back to 2015, and I retrieved the 2013 vintage from earlier work. 2013 reflects the ICP round 2005, the others the ICP rounds from 2011 and 2017. The IEA uses the World Bank's PPP GDP from 1990 onwards for most countries. Prior to that the IEA converts its market exchange rates "based on the PPP conversion factor (GDP) to market exchange rate ratio" (IEA, 2020, p. 26). World Bank data from the World Development Indicators downloaded in 2019 also underlies Fig. 1a and so uses ICP data from 2011, like the data in Hickel and Kallis (2020). The World Bank updated its data to ICP 2017 after 2019.

Primary energy data for 1950–2014 for most countries is from the IEA and the United Nations and the dataset description is in Semieniuk et al. (2021). This data is used for energy intensities. The fig. 2 example uses U.S. energy data from the Energy Information Agency website.[14] Material flow data is from the International Resource Panel.[15] To replicate Hickel and Kallis above, the previous, 2018 edition is used. For additional analysis below the current 2021 edition is used (West et al., 2021). Pollution concentration data for 14 pollutants is from Grossman and Krueger (1995).

4. Results

4.1. Country level switches to and from decoupling

To understand changes in energy and material intensity rates of changes, Fig. 5 plots national accounts (not PPP) growth rates for the same country-time couple under older vintages on the x-axis and the current PWT 10.0 vintage of GDP on the y-axis. 10-year-average annual rates of change are measured to avoid short-run fluctuations driving results. Observations come from a rolling 10-year window over each country's time series for all years available in both vintages. Clearly, the data are organized along the 45-degree line which indicates continuity in GDP growth rate measurement across vintages. However, there is considerable scattering around it. Growth rates for the same period vary across GDP vintages.

Larger time differences between vintages result in lower correlation. Comparing PWT 10.0 with the oldest available vintage, PWT 5.6, the Pearson correlation coefficient is only 0.87 and 0.67 for energy and materials respectively (Fig. 5a & e). The data cloud is also centered away from the origin, documenting an upward translation in revised GDP growth rates. Consequently, more decoupling takes place in PWT 10.0 simply by GDP revision. Since all later vintage plots (Fig. 5b-d & f-g) are both roughly centered around zero when compared with PWT 10.0 and from years after the 1993 SNA publication (published after PWT 5.6), it is likely that the implementation of that SNA revision led to an upward revision of growth rates on average (see also Assa and Kvangraven, 2021).

In the remaining plots of Fig. 5 most of the observations remain in the corridor of ±2 %age points difference. Still, the mean absolute difference in annual growth rates is 0.8, and 0.7 and 0.5%age points in the second, third and fourth rows of plots. For these averages, compound growth leads to differences of between about 4 and 5% in the estimated energy or material intensity over a decade. Moreover, a remarkable number of observations lie far below the corridor, even between PWT 9 and PWT 10. Most of these observations represent African countries, including in the most recent years of data. Knowledge about decoupling patterns is weakest precisely for those countries in which scenarios plant the highest hopes for "leap-frogging" over past, resource and pollution intensive phases of development (Semieniuk et al., 2021). This ambiguous evidence puts in perspective contradictory claims about energy leapfrogging based on – among other differences – different data vintages and sources (Liddle and Huntington, 2021; van Benthem, 2015).

The most spectacular result is a retrospective flip of the direction of change. That is, a country that recouples according to the older vintage, is in retrospect shown to relatively decouple and vice versa. Retrospective relative decoupling is highlighted green and recoupling violet in Fig. 5. The number of countries reporting a switch is quantitatively important. For instance, researchers studying energy intensity some 20 years apart (PWT 6.1 vs 10.0 in plot (b)) would find roughly 12% of countries switching sign in any 10-year period and even 14% if the analysis started only in 1971, as most studies do. While the flips in one versus the other direction are first roughly balanced, there is a bias towards decoupling after around 1980. Even between PWT 9.0 and 10.0, only 5 years apart, some 5 percent of countries flip sign on average both for energy and materials, again with a slight bias towards retrospective relative decoupling. Importantly, these substantial changes to long-term GDP rates of change, when occurring in slow-growing countries or those in depression, can lead to changes in the sign of long-term growth and thus evidence on absolute decoupling, too.

In fact, a large share of countries seen to absolutely decouple in one vintage switches the sign of cumulative GDP growth over periods of 10 years in a subsequent vintage. When moving from PWT5.6 to PWT10.0, half or even as many countries as show absolute decoupling from energy or material in the early vintage either stop or newly start absolutely decoupling when moving to the later vintage (Fig. 6a & e). That is, while resource use declines, the countries in question flip their GDP sign from depression to long-term growth or vice versa.[16] Since most countries flip towards positive growth, that effectively increases the number of countries that are supposed to absolutely decouple. Between more proximate vintages, the switch to and from absolute decoupling is more even and for energy the shares of countries that switch declines after

---

[12] Version 5.6 uses RGDPL (same as RGDP in PWT Mark 5, Summers and Heston, 1991). For versions 8.0 and above, I took national accounts growth rates (not levels) directly from the $RGDP^{NA}$ variable (Feenstra et al., 2015). In 6.x versions CKON instead of CHKON is reported.

[13] I thank Karl Rohrer from the BEA for pointing out comprehensive updates that an initial literature search hadn't unearthed.

[14] Appendix D1 https://www.eia.gov/totalenergy/data/monthly/

[15] https://www.resourcepanel.org/data-resources

[16] The number of countries switching growth rates when including also positive energy or material footprint growth over the same period is, of course, even higher.

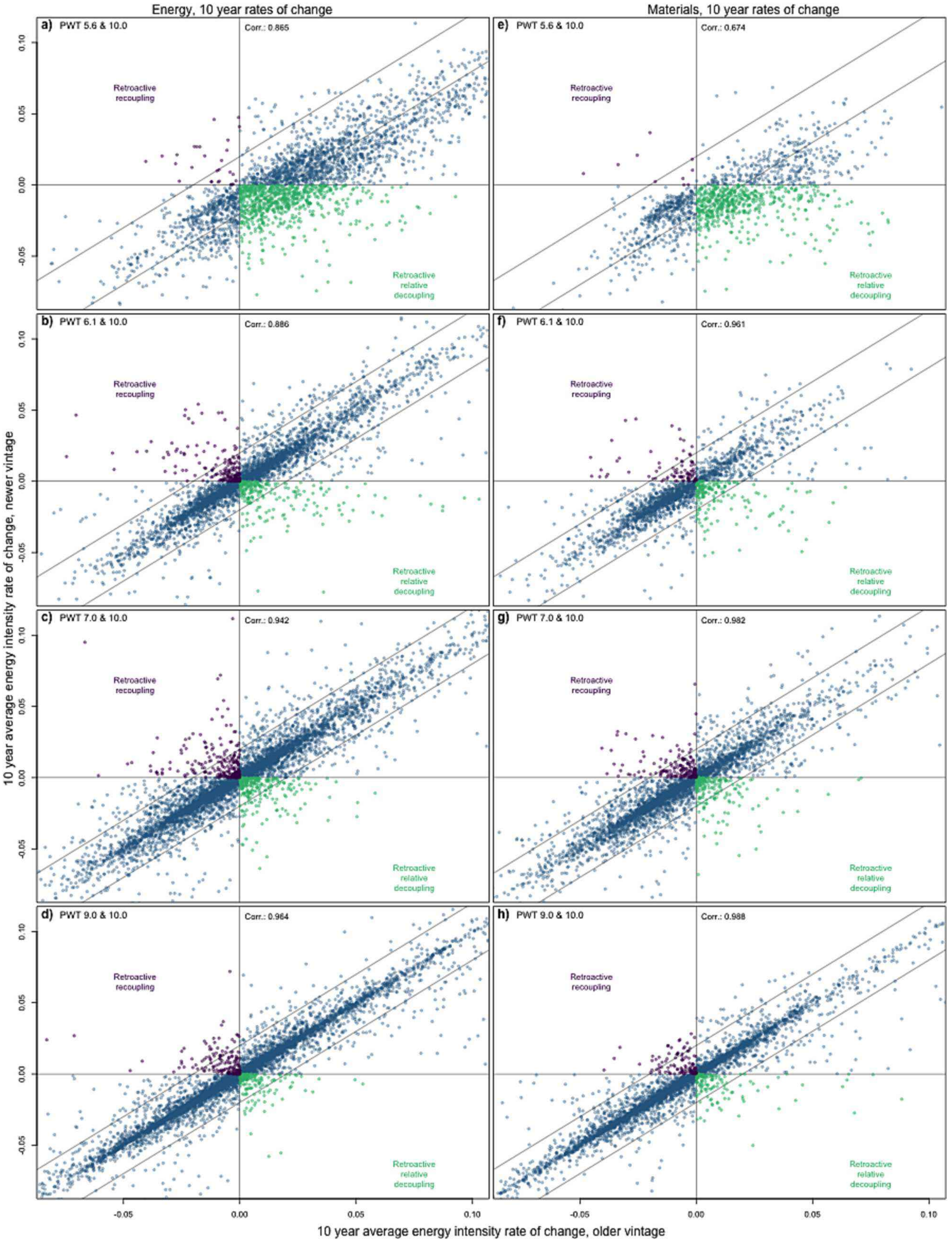


Fig. 5. Difference in energy and material intensity rates of change contingent on GDP vintage: The x-axis reports the 10-year average annual energy intensity rate of change in an older PWT version noted in each plot's top left, the y-axis for PWT 10.0. The diagonal corridor includes observations with less than 2 % age point difference in *annual* growth rates. Green (quadrant 2) and violet (quadrant 4) observations see retrospective switches from relative decoupling to recoupling and vice versa.

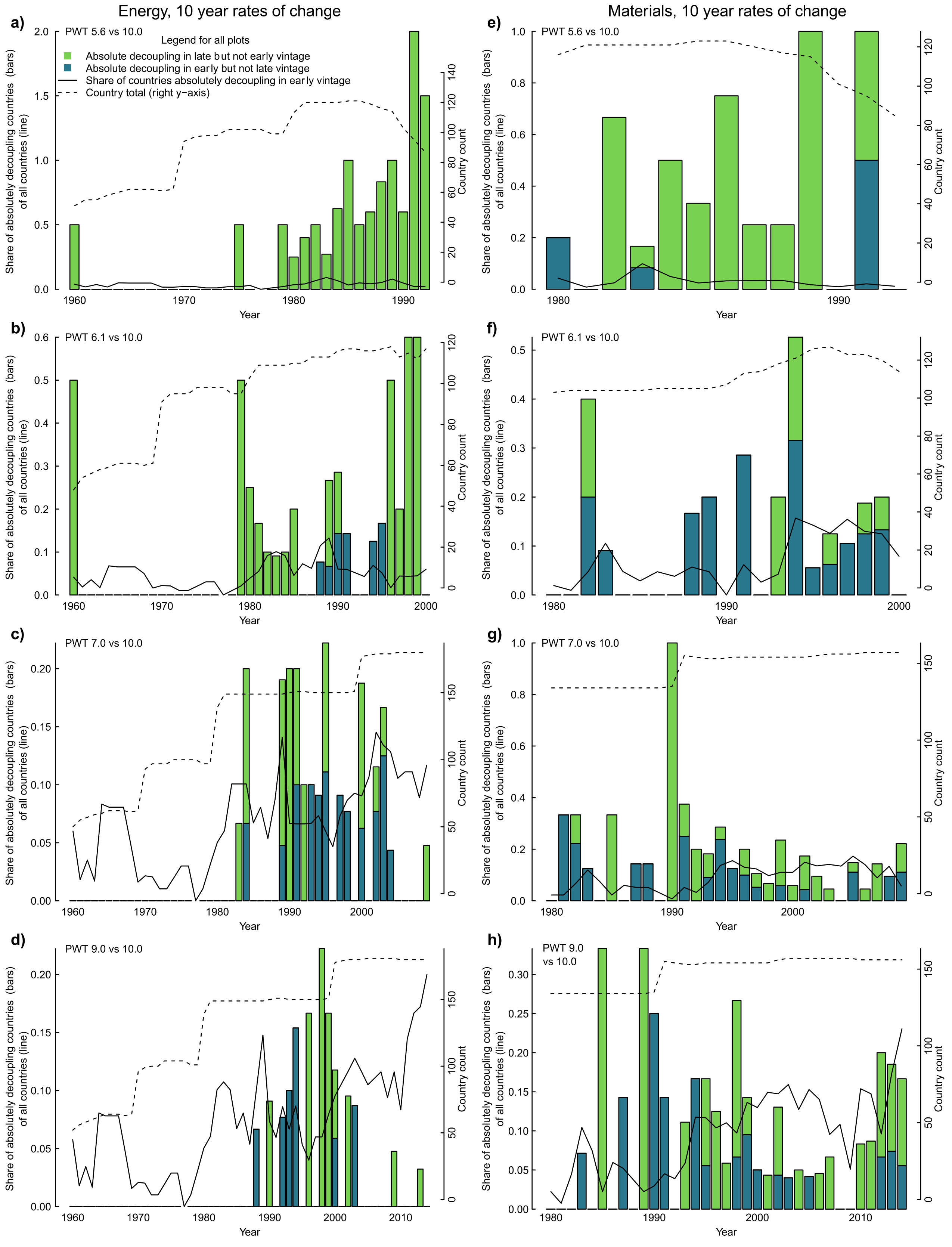


Fig. 6. Retrospective absolute decoupling or recoupling over 10-year periods for energy (left) and materials (right) over time. Retrospective changes (bars) to (light part) and away from (dark part) absolute decoupling in the newer vintage shown as share of all countries that absolutely decoupled in the older vintage. Observations refer to rates of change over the preceding ten years. Solid line shows share of countries absolutely decoupling in earlier vintage, dashed lines show sample size.

PWT6.1. But for materials, the share stays remarkably high. Even between PWT9.0 to 10.0 (Fig. 6 h), in several decadal intervals 15% and more of the number of initially decoupling countries change the sign of their long-term GDP growth as their material use declines. In other words, even absolute decoupling is to some extent conventional.

4.2. Impact on environmental Kuznets curve estimates

The lack of robustness in absolute decoupling evidence spills over into estimates of environmental Kuznets curves. The seminal paper by Grossman and Krueger (1995) considers the correlation between GDP per capita and pollution concentrations. If the relationship is negative, absolute decoupling is implied. Grossman and Krueger carefully regress pollution intensity for 14 pollutants on GDP per capita, its square and cube, as well as on the same powers of the average of the previous three years of GDP. They find that the GDP coefficients tend to be jointly significant, and from there derive the conclusion that "for most indicators, economic growth brings an initial phase of deterioration followed by a subsequent phase of improvement [i.e. absolute decoupling]" (p. 353). Here I ask what a researcher today or a decade ago would find, using contemporary data on GDP for the periods studied by Grossman and Krueger.

Fig. 7 plots the resulting predicted curves for two of their 14 pollutants: smoke in cities, which has a beautifully inverted u shape in their paper, and mercury concentration in rivers, whose inverted u-shape is blemished by an uptick for very high incomes but for which the null hypothesis of joint statistical insignificance of the model's six GDP parameters cannot be rejected. Yet, Grossman and Krueger plot it without further comment, presumably adding to the evidence for the inverted u for most observations. Superimposing the alternative estimates shows that there is a variety of shapes, muddying an inverted u-shape message. All later smoke estimates would suggest smoke rises quickly and hardly drops after the peak only to then rise again. For mercury, the 'peak' is anywhere from USD0 to USD15,000. The PWT 10.0 rgdpe and rgdpe estimates have what can be called an uninverted u-shape. The graphs are also scattered vertically. This is due to the widely scattered pollution data (there are observations above 200 micrograms/m^3 for smoke concentration) and the 6-parameter fit of the polynomial that is sensitive to small variations in the data. As Appendix Table 1 reports systematically for all 14 pollutants, not all results are equally dispersed, but enough have qualitative changes to question whether researchers with later GDP vintages would have been able to write with the same conviction about the initial deterioration and then improvement in environmental quality with absolute decoupling as GDP per capita grows.

4.3. Impact on global energy intensity estimates

For an important example involving relative decoupling, consider the IEA's estimate of changes in global historical energy intensity. These data are often used as a historical benchmark for assessing decoupling assumptions in future energy and climate scenarios, e.g. in the *Global Warming of 1.5 °C* report of the IPCC (Foster et al., 2018, p. 2A22). The IEA publishes a new vintage of its database every year, which updates its energy and GDP data, including an estimate for the world. Table 2 columns I-III show the ratio of primary energy as well as GDP reported for the world in 2010 over that in 1971 in 10 database vintages. While the primary energy ratio stays fairly constant across vintages, the GDP ratio varies more, expressed by a standard deviation that is an order of magnitude higher than that of primary energy. In other words, the main driver of any energy intensity changes is a change in accounting for GDP, not energy, across vintages. 2010 energy intensity rates as a share of those in 1970 are reported in columns IV-V. In the extreme case of the market exchange rate and 2013 and 2021 vintages, the difference implies a 0.2% faster annual decline using 2021 data (columns VI-VII). This compounds to an 8% lower energy intensity over 40 years. Interestingly, the variation in PPP GDP is smaller than that in MER. And the negative rate of change for MER intensity accelerates almost monotonely over time, while that for PPP slows down after the 2020 vintage. Yet, both intensities' annual rates of change vary by more than 0.1% in the course of just a few years difference and the oldest vintage has the slowest decline.

5. Discussion

The foregoing results show that evidence for relative and absolute decoupling varies in an economically important way with GDP revisions over time. It follows that the entrenched debate about whether environmental Kuznets curves exist or not, and the extent to which a growing economy can limit its environmental impact, is marred by an ambiguity

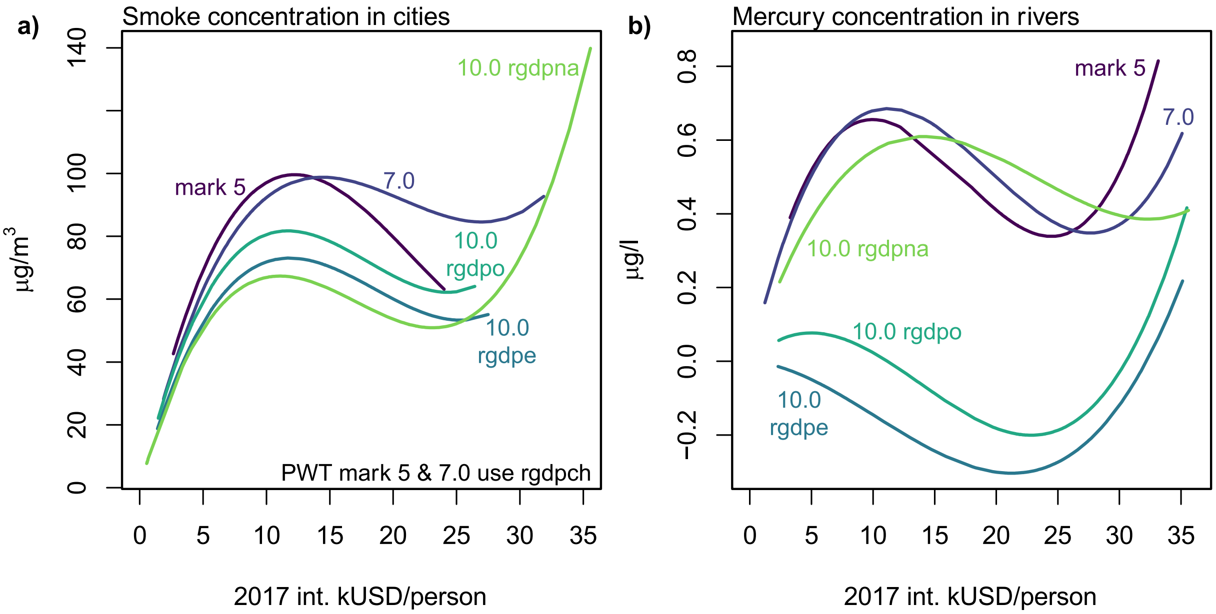


Fig. 7. Grossman and Krueger re-estimates with different PWT vintages for selected pollutants. The original estimate was with mark 5, rgdpch data, all other data newly estimated by the author.

Table 2
Ratio of 2010 to 1971 observations at the world level from the IEA World Indicators for various vintages, and implied compound annual growth rates.

Vintage	Primary energy (PE)	MER GDP	PPP GDP	PE intensity MER	PE intensity PPP	CAGR PE intensity MER	CAGR PE intensity PPP
Column:	I	II	III	IV=I/II	V=I/III	VI=IV$^{(1/40)}$-1	VII=V$^{(1/40)}$-1
2013	2.333	3.170	3.597	0.736	0.649	−0.76%	−1.08%
2015	2.316	3.176	3.753	0.728	0.618	−0.79%	−1.20%
2016	2.345	3.277	3.776	0.714	0.619	−0.84%	−1.19%
2017	2.331	3.290	3.785	0.709	0.614	−0.86%	−1.21%
2018	2.331	3.263	3.748	0.714	0.621	−0.84%	−1.19%
2019	2.328	3.262	3.736	0.713	0.623	−0.84%	−1.17%
2020	2.328	3.374	3.776	0.690	0.615	−0.92%	−1.21%
2021	2.331	3.384	3.734	0.686	0.624	−0.94%	−1.17%
2022	2.331	3.388	3.640	0.688	0.640	−0.93%	−1.11%
2023	2.331	3.388	3.640	0.688	0.640	−0.93%	−1.11%
Standard deviation	0.007	0.084	0.067	0.018	0.012		

Notes: MER = market exchange rate, PPP = purchasing power parity, CAGR = compound annual growth rate.

that hasn't previously been acknowledged. Evidence from different sources can only be directly compared if both sources use the same GDP definition. If they do not, it is unclear whether one is more appropriate than the other. If social reality changes over time as expressed by structural revisions in GDP, it throws a spanner in the works of establishing a truth by accumulating a growing body of evidence. These news may be vexing to the environmental scientists who contribute a good part of the evidence on this subject, but it is necessary to acknowledge the role of social reality that intensity indicators with GDP in the denominator embody.

One objection to this claim of ambiguity is that while GDP may change on the margin, it does not make a difference for the qualitative results in meta-analyses. Of course it is unlikely that in any one structural GDP revision a large share of countries would switch from positive to negative long-term GDP growth or vice versa. However, the results about sensitivity of absolute decoupling to vintage in Fig. 6 show that precisely those countries that are shown to achieve absolute decoupling in one vintage have reported close to zero GDP growth leading to sign flips in a surprisingly large share of them. And the results on the environmental Kuznets curve, another measure of absolute decoupling or lack thereof, do not seem to come to any agreement on a particular shape, at least for some pollutants. Finally, GDP is but one more indicator that can change, compounding ambiguity rather than newly introducing it into the results.

Another objection could follow from a progress of science and measurement perspective: results arrived at with later vintages should be privileged over older ones.[17] For instance, the SNA 2008 explains that "methodological research over the past decade or so had resulted in improved methods of measuring some of the more difficult components of the accounts" (European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank, 2009, p. xlvii). These measurement improvements certainly support a progress perspective. However, in the same sentence the SNA also reports that the main reason for an overhaul of the accounting framework is "that the economic environment in many countries had evolved significantly since the early 1990s when the 1993 SNA had been developed" (ibid.) The national accountants decided that changed economies required different accounting conventions, but for consistency then impose them over the entire historical period including the time for which the earlier version of the SNA had been developed. More generally, where to draw production boundaries in national accounts is an analytical problem whose solution depends on the questions that the national accounts are supposed to answer (Shaikh and Tonak, 1994). For instance, turning military weapons into assets that provide capital services and whose manufacture therefore counts as investment and thus adds to GDP since SNA 2008 can hardly be called an improvement in the understanding of how economies add value. Rather, deciding which questions merit answering involves a value judgment (Assa and Kvangraven, 2021). While there is thus a particular GDP definition imposed at any time, it does not follow that it is the correct or best one in some universal sense (Hoekstra, 2019). Although preferring the latest revision is one strategy, there are no strong epistemological arguments for why it provides a superior picture of past economic activity on the scale of decades. But that is exactly the time scale that matters for decoupling.

If, with some economists cited earlier, the opposite position is taken instead that GDP and its revisions do not necessarily provide a good or improving description of the economy, the consequences are even more powerful. If GDP does not capture the actual functioning of an economy well, then the usefulness of intensity indicators deteriorates. For instance, if one believes that growth is understated in rich economies due to an underestimation of the value of innovations, then the decoupling potential in these economies may remain unrecognized, leading to unduly timid policy goals. The opposite problem is more worrisome. If growth rates in developing countries are overstated due to GDP revisions but also purchasing power parity in models of the economy and climate change, this could inspire overconfidence e.g. in the carbon emissions mitigation potential in these economies according to models using such GDP rates. Consequently, it may turn out to be harder to reduce emissions in these countries than the modeling effort suggested, frustrating ambitions. It may also lead to calls by rich countries for developing countries to take on a larger share of the mitigation burden because of the apparent ease with which they decouple.[18] All in all, measurement changes in GDP can have important real-world policy implications that are currently underappreciated in the environmental policy debate.

A last subtle point about modeling of future scenarios must be mentioned. The divergence of IEA energy intensity estimates can matter greatly for a climate change mitigation scenario. Suppose modelers extrapolated the historical intensity trend over the next 40 years and updated it with the faster intensity decline in the later vintage but kept the existing GDP growth rate. In that case, a 4 or 8% lower energy intensity in 2060 implies 4 or 8% less modeled energy demand and less required mitigation via other measures, such as investments into renewable energy. In principle, the faster energy intensity decline should only reflect a faster growing GDP so the amount of energy demanded should not change in the model. In practice, it is unclear to what extent these updates are synchronized. For instance, modelers who input GDP growth as a parameter into their model could take the GDP projections based on 2012 World Bank PPP GDP and provided by Dellink et al. (2017) which are promoted as a standard for the *Shared Socioeconomic Pathways*, but calculate energy intensity projections from more updated IEA data. In this case, decoupling would seem easier while GDP

---

[17] I thank Tiago Domingos for making this argument at the ISEE conference.

[18] I thank Jayati Ghosh for alerting me to this possibility.

growth is ‘slow’, leading to an overall lower pressure on energy demand, which turns out, however, to be merely a misleading accounting artifact. The converse applies when a new vintage reports a slower decline in intensity.

6. Conclusion

This paper has traced structural revisions in how GDP is accounted for and shown that these revisions impact measures of decoupling in both quantitatively and qualitatively important ways. Between vintages, 10–15% of countries switch between relative decoupling and recoupling from energy or materials on decadal intervals, and up to as many countries as decouple absolutely in an older vintage stop or newly start absolutely decoupling in the newer vintage. Some of the largest swings in decoupling estimates occur in African countries, suggesting that estimates of ‘leapfrogging’ in these countries are particularly dependent on vintage used. Some of the environmental Kuznets curve results on absolute decoupling in Grossman and Krueger’s seminal paper are not robust to switching the GDP data vintage. And the IEA’s annual global decline in energy intensity accelerates by up to −0.2 percentage points, affecting the baselines of current integrated assessment models.

One response is to only use results from the latest vintage, due to its better measurement of the economy. But aside from discarding past evidence, there is no strong epistemological argument that the newest vintage is better for longer-term economic measurement. Decoupling analysts must recognize an ambiguity built into the denominator of their intensity measures which can even affect the sign of GDP growth and thus absolute decoupling, just as they have become accustomed to numerator problems such as different patterns of de- and recoupling for territorial vs footprint measures. The ambiguity is particularly important for modeling long-term economic and environmental change, where historical correlations between GDP and other measures are used both for model calibration and validation. And if efforts to absolutely decouple intensify while adverse effects of climate change and the breaching of other planetary boundaries slow down growth, the variation in GDP across vintages around that slower trend might bring more crossing of the zero-growth mark. This could introduce even more ambiguity about absolute decoupling, the ultimate test of economic growth for an indefinite future period.

I draw three conclusions. First, to advance insight, debate and scenario modeling, rigorous reporting of GDP definitions, vintage, and the sharing of data for subsequent comparison and replication in empirical analyses, are urgently needed. The data sharing is particularly important because older vintages of GDP or other macro data are not normally available in the usual repositories (the PWT and recent vintages at the BEA and IEA being commendable exceptions). If feasible or if there is a concern about robustness, analysts of longer-term descriptive studies could additionally employ ranges from the upper to lower bound, like reported in Table 2. Analysts of inferential studies could reestimate their statistical models using different GDP vintages like for the environmental Kuznets analysis above. Of course, when data for the most recent years is analyzed, this method suffers from the problem of unavailability of earlier vintages for the most recent years or quarters. It may be more suited for retrospective analyses. For projections, repositories of scenarios of the future stemming from different models likely using various underlying historical time series should require modeling teams to add information on the vintage of these times series for each component of these ‘ensembles of opportunity’ (Huppmann et al., 2018). In this way, IPCC summary analyses have the ability to correct for possible differences in baselines, and later analysts can differentiate between models also in this dimension.

Second, the unreliability of some relative and even absolute decoupling estimates highlights the limitations of using evidence for or against historical decoupling in the debate about the feasibility of continued economic growth under successful measures to halt and reverse environmental degradation (Pollin, 2019; Schor and Jorgenson, 2019). One alternative is to focus directly on the indicators that need to decline (e.g. CO_2 emissions) or remain within ‘planetary boundaries’. Since these are often concentrated in certain activities or sectors (e.g. emissions from fossil-fuel production or use in certain applications) it could be more effective to focus on sectoral growth or degrowth (Pollin, 2018) instead of reasoning in terms of the whole economy. It also follows that policy targets formulated in terms of aggregate intensities should be specific about the GDP definition used or use absolute emissions/resource figures rather than intensities to avoid ambiguity.

Third, beyond the epistemological barriers to understanding decoupling presented here, the work by Desrosières (1998) reminds us that the revision of GDP series itself may influence how the possibility of decoupling is perceived. This political element was recently examined and found to influence indicators of ecological impact (Requena-i-Mora and Brockington, 2022). Seen from this perspective, the variation over vintages ceases to be a conundrum. The variation serves instead as an opportunity for a robustness check on the susceptibility of current GDP estimates to political preoccupations of the day.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data and code for generating all figures is available at Zenodo at https://doi.org/10.5281/zenodo.8381368.

Acknowledgments

I thank Duncan Foley for introducing me to variations in GDP definitions that inspired this research, and Shouvik Chakraborty, Nancy Folbre, Duncan Foley, Roger Fouquet, Lukas Hardt, Carol Heim, Antoine Missemer, Robert Pollin, Michael Ross, Ellis Scharfenaker, Isabella Weber, three anonymous reviewers and participants at the 2021 ISEE Conference and 2022 Berlin RSERC workshop for feedback on earlier drafts. I thank Alex Smith and Calista Amoah for research assistance. I also thank Ricardo Salas for helping me get started with Stata so I could replicate the Grossman and Krueger results, for which the authors helpfully provided Stata commands. All remaining errors are mine.

Appendix A. Reasons for structural revisions

The first reason for structural revisions is the reference year change. It presents an index number problem. Since GDP is measured at current prices every year, but the magnitude of interest is often ‘real’ growth, national accountants subtract inflation from economic growth and so attempt to recover the growth in the actual quantity of goods and services. Traditionally, in the US, Laspeyres quantity indices were used that compare quantities in the reference year measured at reference year prices with quantities in other years but also measured at reference year prices (constant dollars).

Fixing prices, and so the weight by which goods and services entered into the GDP aggregate, however, makes growth rates contingent on the reference year. In particular, it causes substitution bias (Braithwait, 1980), which tends to overestimate growth rates after the reference period and underestimate growth rates before the reference period. The problem is that consumption tends to shift towards sectors with relatively low price increases or even declines (think of solar PV panels). The U.S. therefore switched to Fisher or chained-dollar quantity indices in 1996 that instead use information about prices from both periods. Changing the reference period has no impact on their measured growth rate (for a detailed discussion and examples see Landefeld et al., 2003). This may be one reason for the relatively large jump in the pre-reference year growth rates between the 1987 and 1999 structural revisions depicted in fig. 2b above. One drawback of chained-price measures is that GDP cannot anymore be partitioned exactly into its components like major expenditure categories, such as consumption and investment (Landefeld et al., 2003). The rebasing of the reference year itself is a technical problem. It has a political component however as base-year and choice of deflation method affect the reporting of the economy's past performance.

The second and more controversial because less technical driver of structural changes are redefinitions of what counts as part of GDP and how. National accountants themselves stress that GDP and other aggregates need to "meet a wide range of analytical purposes" (European Commission et al., 2009, p. 6), therefore they must "provide a relevant and accurate picture of the evolving U.S. [or any other] economy" (D. J. Fixler et al., 2014, p. 1). To retain this usefulness, it follows, the accounts must evolve with the economy. Thus, Fixler and colleagues (ibid, footnote 1) point out that investment in software was negligible in the 1950s but grew to 1.7% of GDP by 2012, implying that not including it as investment (but as intermediate consumption and hence cost to final consumers, netted out), would make the 'picture' of the economy less relevant. They also note that its inclusion in the 1999 comprehensive revision raised level and growth rates of the economy. This is a case where a new component was added, that grew faster than existing components historically, thus raising past growth rates. There is a wide variety of redefinitions, and they range from large (changing the treatment of financial sector or government) to more subtle such as hedonic pricing to account for product quality (Coyle, 2014). The main point is that these revisions do impact GDP growth rates as growth rates are changed either directly or through the alteration of component weights.

It is difficult to exactly trace the causes of redefinitions, but safe to say that national accountants are critically accompanied by economists. Many critiques by economists are motivated by value theory. Simon Kuznets famously prepared the first U.S. national income estimates in 1931 but disagreed with the national accounting framework settled on after the Second World War. Kuznets was convinced that the national income should reflect welfare, not economic activity, bringing it closer to the value theory based on classic utilitarianism propounded in Pigou (1920). From this stance, Kuznets argued that many 'final expenditures' of households adding to GDP should really be "business costs" (Kuznets, 1948, p. 157). This refers to employees' personal expenditures enabling them to do their work (e.g. a public transport ticket to get to work). If Kuznets was moved by neoclassical economics to shrink GDP, more recent critiques motivated by neoclassical economics tend to argue for enlarging GDP. Recently proposed changes often revolve around better measurement of intangibles and innovation (Corrado et al., 2021; Coyle, 2014; Jorgenson, 2009) or how best to account for digital services (Brynjolfsson and McAfee, 2014), and would align national accounts more with measuring growth drivers identified in the recent endogenous growth literature.

Economists taking other than a neoclassical lens have other critiques. Feminist economists have shown the impact of the treatment of unpaid (care) work on GDP growth rates (Wagman and Folbre, 1996), a concern that overlaps with problems of measuring the informal economy, the largest economic sector in many countries (Ghosh, 2020). Reich (2001) has elaborated inconsistencies between neoclassical value theory and national accounts and shows that in important respects the accounts' construction is more compatible with classical political economy. Marxist economists that distinguish a sphere of production and one of exchange (that does not however add value) show how such more restricted or redefined measures of output correlate better with other macroeconomic measures of interest such as investment, (un)employment or inequality (Assa, 2017; Basu and Foley, 2013; Shaikh and Tonak, 1994; Tercioglu, 2021; Wolff, 1987).[19] One likely reason for the good correlation is that these alternative measures tend to reduce what is an increasing share of imputed value added in national accounts (Foley, 2013). A recent slate of contributions investigates the political motivations behind national accounts definitions and (lack of) revisions. Christophers (2011) recounts the political process of 'making finance productive' in the national accounts, which accompanies the broader trend of financialization of economies (Epstein, 2005), and Mazzucato (2018) argues that national accounts may undervalue government activity (see also Eisner and Nebhut, 1981). Assa (2017) consequently labels GDP as "statistical rhetoric with political goals" (p. 22). All of this is to say, that economic theory has and will continue to exert pressure on redefinitions of GDP.[20]

A third important cause for revisions has to do with the use of new datasets, even apart from conceptual novelties. To return to the 1965 *Survey of Current Business*, some changes in the US GDP definition then were due to new incorporation of company censuses that revised historical data. In 2010, Ghana's GDP was revised 60% upwards mainly due to the use of new data (Jerven, 2013), and in 2012 Nigeria's GDP was revised upwards by 100% (Feenstra et al., 2015). Of course, these changes also affect past GDP and only have an impact on growth rates to the extent that the new data shows other trends over time. The controversy over India's recent GDP growth being in good part not about completely new data but about which of several existing datasets (and assumptions about the informal sector) to use also highlights the political nature of this type of revision (Nagaraj et al., 2021). Finally, it may also be that countries' governments intervene in the publications of figures for political reasons, so the revision is more about withholding than releasing data (Seltzer, 1994). In sum, there are many reasons why GDP definitions are changed, and these will persist into the future.

Appendix B. Purchasing power parity example

The basic idea of purchasing power parity is simple: suppose that after exchanging a certain amount of currency from country A for that of country B at MER, one can buy twice the amount of goods and services in country B that one could have with that money in country A. Think of being able to purchase two haircuts instead of one. Country A has higher prices. Expressed in the currency of country A, the 'real' GDP of country B should be twice as big as the MER would suggest. Country B's prices are lower for the same goods and so need to be 'inflated' for comparison. Empirical PPP estimates show that price levels in rich countries tend to be higher (so-called Penn effect) and therefore the GDP of developing countries needs to be inflated for comparison. For instance, India's MER GDP in 2020 was $2.7 trillion but its PPP GDP according to the World Bank was $9.0 trillion.[21]

---

[19] Marxist-feminist social reproduction theory instead uses an expanded measurement base (Moos, 2021).

[20] Naturally there are also debates about revisions of components of GDP, such as investment and savings (Pollin, 1997).

[21] Data as of January 13, 2022. Series NY.GDP.MKTP.CD and NY.GDP.MKTP.PP.CD on https://data.worldbank.org/

Appendix C. Additional Grossman Krueger calculations

Table A1 reports my reestimation of Grossman and Krueger's internal maxima and minima, for the three PWT vintages for each of their pollutants. Variation of the GDP/person level at which the "EKC turning point" occurs can be large, e.g. for lead or sulfur dioxide, even when all estimates are statistically significant. Changes in the sign of the cubic polynomial are reported with a ˆ. This change is particularly powerful when the min lies to the left of the maximum, i.e. the EKC turning point but is above zero. NA means the polynomial declines monotonically. Changes in statistical significance are reported in the right column. As the smoke example shows, these qualitative change indicators do not exhaust the possible variations in levels and shape, which could be gleaned from a look at the plot.

Table A1
Grossmann and Krueger cubic polynomial with internal max (=EKC turning point) and min in thousands 2017$ per capita with varying GDP vintages and measures.

		PWT5	PWT7.0	PWT10.0			
Pollutant		rgdpch	rgdpch	rgdpe	rgdpo	rgdpna	Joint significance[*]
Arsenic	Max	9.7	11.5	10.1	10.3	12.0	
	Min	28.9	36.4	29.7	30.3	34.7	
BOD	Max	15.1	ˆ15.4	14.2	13.2	ˆ16.8	PWT 10.0 GDP coefficients jointly insignificant
	Min	74.8	ˆ-70.1	89.8	52.2	ˆ-22.6	
Cadmium	Max	22.9	23.5	22.1	23.2	NA	rgdpe & rgdpo coefficients jointly insignificant
	Min	9.0	12.5	8.0	8.5	NA	
COD	Max	15.5	ˆ19.1	19.5	18.2	21.3	Coefficients always jointly insignificant
	Min	−79.7	ˆ277.2	−0.2	−6.7	−6.9	
Coliform	Max	6.0	6.0	10.0	11.2	7.3	
	Min	16.2	20.6	32.8	45.1	21.2	
Dissolved oxygen[**]	Max	−31.9	ˆ5.0	−5.7	−45.9	6.7	
	Min	5.3	ˆ14.0	8.5	8.5	17.3	
Fecal coliform	Max	15.7	ˆ5.3	14.5	14.0	ˆ10.4	
	Min	0.6	ˆ34.1	−5.1	−12.4	ˆ43.2	
Lead	Max	3.7	10.8	7.98	7.85	11.73	
	Min	28.1	36.3	30.2	30.4	39.7	
Mercury	Max	10.0	11.2	−0.7	4.97	14.1	Coefficients always jointly insignificant
	Min	24.5	27.6	21.2	22.8	32.4	
Nickel	Max	8.2	NA	ˆ25.8	ˆ25.5	NA	Coefficients always jointly insignificant
	Min	29.0	NA	ˆ11.4	ˆ11.6	NA	
Nitrate	Max	20.8	19.0	20.6	20.6	25.8	PWT 7.0 GDP coefficients jointly insignificant
	Min	3.2	−4.4	0.6	−1.0	6.2	
Smoke	Max	12.23	14.4	11.8	11.7	11.0	
	Min	30.5	27.0	25.3	24.3	23.1	
Sulfur dioxide	Max	7.9	5.8	0.9	2.1	10.0	
	Min	26.7	32.8	26.9	28.2	44.7	
Suspended particles	Max	NA	23.3	NA	52.0	138.0	
	Min	NA	13.7	NA	38.5	21.6	

NA monotone slope.
[*] Empty cells imply all estimates are jointly significant.
[**] Not a pollutant, u-shape expected.
[ˆ] Order of local max and min reversed.

References

Acemoglu, D., Naidu, S., Restrepo, P., Robinson, J.A., 2019. Democracy does cause growth. J. Polit. Econ. 127 (1), 47–100.

Akizu-Gardoki, O., Wakiyama, T., Wiedmann, T., Bueno, G., Arto, I., Lenzen, M., Lopez-Guede, J.M., 2021. Hidden energy flow indicator to reflect the outsourced energy requirements of countries. J. Clean. Prod. 278, 123827. https://doi.org/10.1016/j.jclepro.2020.123827.

Anand, S., Segal, P., 2008. What do we know about global income inequality? J. Econ. Lit. 46 (1), 57–94. https://doi.org/10.1257/jel.46.1.57.

Assa, J., 2017. The Financialization of GDP. Routledge, Oxford and New York.

Assa, J., Kvangraven, I.H., 2021. Imputing away the ladder: implications of changes in GDP measurement for convergence debates and the political economy of development. New Political Econ. 26 (6), 985–1014. https://doi.org/10.1080/13563467.2020.1865899.

Basu, D., Foley, D.K., 2013. Dynamics of output and employment in the US economy. Camb. J. Econ. 37, 1077–1106. https://doi.org/10.1093/cje/bes088.

Bos, F., 2006. The development of the Dutch national accounts as a tool for analysis and policy. Statistica Neerlandica 60 (2), 225–258. https://doi.org/10.1111/j.1467-9574.2006.00314.x.

Braithwait, S.D., 1980. The substitution bias of the laspeyres price index: an analysis using estimated cost-of-living indexes. Am. Econ. Rev. 70 (1), 64–77.

Brynjolfsson, E., McAfee, A., 2014. The Second Machine Age : Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.

Carson, R.T., 2010. The environmental Kuznets curve: seeking empirical regularity and theoretical structure. Rev. Environ. Econ. Policy 4 (1), 3–23. https://doi.org/10.1093/reep/rep021.

Castles, I., Henderson, D., 2003. The IPCC emission scenarios: an economic-statistical critique. Energy Environ. 14 (2–3), 159–185. https://doi.org/10.1260/095830503765184583.

Christophers, B., 2011. Making finance productive. Econ. Soc. 40 (1), 112–140. https://doi.org/10.1080/03085147.2011.529337.

Corrado, C., Haskel, J., Miranda, J., Sichel, D., 2021. Measuring and Accounting for Innovation in the Twenty-First Century. University of Chicago Press, Chicago.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., van den Belt, M., 1997. The value of the world's ecosystem services and natural capital. Nature 387 (6630), 253–260. https://doi.org/10.1038/387253a0.

Coyle, D., 2014. GDP: A Short but Affectionate History. Princeton University Press, Princeton and Oxford.

Croushore, D., Stark, T., 2003. A real-time data set for macroeconomists: does the data vintage matter? Rev. Econ. Stat. 85 (3), 605–617. https://doi.org/10.1162/003465303322369759.

Deaton, A., Aten, B., 2017. Trying to understand the PPPs in ICP 2011: why are the results so different. Am. Econ. J. Macroecon. 9 (1), 243–264. https://doi.org/10.1257/mac.20150153.

Deaton, A., Heston, A., 2010. Understanding PPPs and PPP-Based National Accounts. Understanding PPPs and PPP-Based National Accounts 2 (4), 1–35.

Deaton, A., Schreyer, P., 2022. GDP, wellbeing, and health: thoughts on the 2017 round of the international comparison program. Rev. Income Wealth 68 (1), 1–15. https://doi.org/10.1111/roiw.12520.

Dellink, R., Chateau, J., Lanzi, E., Magné, B., 2017. Long-term economic growth projections in the Shared Socioeconomic Pathways. Glob. Environ. Chang. 42, 200–214.

Desrosières, A., 1998. The Politics of Large Numbers: A History of Statistical Reasoning. Harvard University Press, Cambridge, MA.

Dinda, S., 2004. Environmental Kuznets curve hypothesis: a survey. Ecol. Econ. 49 (4), 431–455. https://doi.org/10.1016/j.ecolecon.2004.02.011.

Eisner, R., Nebhut, D.H., 1981. An extended measure of government product: Preliminary results for the United States, 1946–76. Rev. Income Wealth 27 (1), 33–64. https://doi.org/10.1111/j.1475-4991.1981.tb00191.x.

Epstein, G.A., 2005. Introduction: financialization and the world economy. In: Epstein, G.A. (Ed.), Financialization and the World Economy. Elgar, Cheltenham, UK, and Northampton, MA.

European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, & World Bank, 2009. System of National Accounts 2008. United Nations, New York.

Feenstra, R.C., Inklaar, R., Timmer, M.P., 2013. PWT 8.0 - a user guide. Retrieved from. https://www.rug.nl/ggdc/docs/pwt_80_user_guide.pdf.

Feenstra, R.C., Inklaar, R., Timmer, M.P., 2015. The next generation of the Penn World Table. Am. Econ. Rev. 105 (10), 3150–3182. https://doi.org/10.1257/aer.20130954.

Feldstein, M., 2017. Underestimating the real growth of GDP, personal income, and productivity. J. Econ. Perspect. 31 (2), 145–164. https://doi.org/10.1257/jep.31.2.145.

Fixler, D.J., Greenaway-Mcgrevy, R., Grimm, B.T., 2014. The revisions to GDP, GDI, and their major components. Surv. Curr. Bus. 94 (8), 1–23.

Fixler, B.D.J., De Francisco, E., Kanal, D., 2021. The revisions to Gross domestic product, gross domestic income, and their major components. Surv. Curr. Bus. 101 (1).

Folbre, N., 2020. Chapter 6: Informal employment and non-market work. In: Charmes, J. (Ed.), Research Handbook on Development and the Informal Economy. Edward Elgar, Cheltenham, UK.

Foley, D.K., 2013. Rethinking financial capitalism and the “information” economy. Rev. Rad. Polit. Econ. 45 (3), 257–268. https://doi.org/10.1177/0486613413487154.

Foster, P., Huppmann, D., Kriegler, E., Mundaca, L., Smith, C., Rogelj, J., Séférian, R., 2018. 2.SM Mitigation pathways compatible with 1.5°C in the context of sustainable development – Supplementary Material. In: Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P.R., Waterfield, T. (Eds.), Global warming of 1.5°C. Cambridge University Press, Cambridge, UK and New York, NY.

Franses, P.H., 2009. Why is GDP typically revised upwards? Statistica Neerlandica 63 (2), 125–130. https://doi.org/10.1111/j.1467-9574.2008.00410.x.

Ghosh, J., 2018. A note on estimating income inequality across countries using PPP exchange rates. Econ. Labour Relat. Rev. 29 (1), 24–37. https://doi.org/10.1177/1035304618756263.

Ghosh, J., 2020. Global inequality in a time of pandemic. Real-World Econ. Rev. 92, 22–32.

Glejser, H., Schavey, P., 1969. An analysis of revisions of national accounts data for 40 countries. Rev. Income Wealth 20 (3), 317–332.

Gordon, R.J., 2017. The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War. Princeton University Press, Princeton.

Grossman, G.M., Krueger, A.B., 1991. Environmental Impacts of a North American Free Trade Agreement. *NBER Working Paper, 3914*.

Grossman, G.M., Krueger, A.B., 1995. Economic growth and the environment*. Q. J. Econ. 110 (2), 353–377. https://doi.org/10.2307/2118443.

Hanley, N., Dupuy, L., McLaughlin, E., 2015. Genuine savings and sustainability. J. Econ. Surv. 29 (4), 779–806. https://doi.org/10.1111/joes.12120.

Harbaugh, W.T., Levinson, A., Wilson, D.M., 2002. Reexamining the empirical evidence for an environmental Kuznets curve. Rev. Econ. Stat. 84 (3), 541–551. https://doi.org/10.1162/003465302320259538.

Henriques, S.T., Kander, A., 2010. The modest environmental relief resulting from the transition to a service economy. Ecol. Econ. 70 (2), 271–282. https://doi.org/10.1016/j.ecolecon.2010.08.010.

Hickel, J., Hallegatte, S., 2022. Can we live within environmental limits and still reduce poverty? Degrowth or decoupling? Dev. Policy Rev. 40 (1), e12584 https://doi.org/10.1111/dpr.12584.

Hickel, J., Kallis, G., 2020. Is green growth possible? New Political Econ. 25 (4), 469–486. https://doi.org/10.1080/13563467.2019.1598964.

Hoekstra, R., 2019. Replacing GDP by 2030: Towards a Common Language for the Well-Being and Sustainability Community. Cambridge University Press, Cambridge.

Huppmann, D., Rogelj, J., Kriegler, E., Krey, V., Riahi, K., 2018. A new scenario resource for integrated 1.5 °C research. Nat. Clim. Chang. 8 (12), 1027–1030. https://doi.org/10.1038/s41558-018-0317-4.

IEA, 2020. World Energy Balances 2020 Edition: Database Documentation. International Energy Agency, Paris.

IEA, 2022. World Energy Indicators. International Energy Agency, Paris.

Jerven, M., 2013. Briefing for richer, for poorer: Gdp revisions and africa’s statistical tragedy. Afr. Aff. 112 (446), 138–147. https://doi.org/10.1093/afraf/ads063.

Jiborn, M., Kander, A., Kulionis, V., Nielsen, H., Moran, D.D., 2018. Decoupling or delusion? Measuring emissions displacement in foreign trade. Glob. Environ. Chang. 49, 27–34. https://doi.org/10.1016/j.gloenvcha.2017.12.006.

Johnson, P., Papageorgiou, C., 2020. What remains of cross-country convergence ? J. Econ. Lit. 58, 129–175.

Jorgenson, D.W., 2009. A new architecture for the U.S. National Accounts. 55 (1), 199–219.

Kacprzyk, A., Kuchta, Z., 2020. Shining a new light on the environmental Kuznets curve for CO2 emissions. Energy Econ. 87, 104704. https://doi.org/10.1016/j.eneco.2020.104704.

Kander, A., 2005. Baumol’s disease and dematerialization of the economy. Ecol. Econ. 55 (1), 119–130. https://doi.org/10.1016/j.ecolecon.2004.10.008.

Kendrick, J., 1970. The historical development of national-income accounts. History Polit. Econ. 2 (2), 284–315.

Koomey, J., Schmidt, Z., Hummel, H., Weyant, J., 2019. Inside the Black Box : understanding key drivers of global emission scenarios. Environ. Model. Softw. 111 (August 2018), 268–281. https://doi.org/10.1016/j.envsoft.2018.08.019.

Kraan, O., Chappin, E., Kramer, G.J., Nikolic, I., 2019. The influence of the energy transition on the significance of key energy metrics. Renew. Sust. Energ. Rev. 111, 215–223. https://doi.org/10.1016/j.rser.2019.04.032.

Kravis, I.B., Heston, A., Summers, R., 1982. World Product and Income: International Comparisons of Real Gross Product. John Hopkins Press, Baltimore and London.

Kravis, I.B., Heston, A., Summers, R., 1983. The share of services in economic growth. In: Global Econometrics: Essays in Honor of Lawrence R. Klein. MIT Press, Cambridge, MA.

Krey, V., Masera, O., Blanford, G., Bruckner, T., Cooke, R., Fisher-Vanden, K., Zwickel, T., 2014. Annex II: metrics and methodology. In: Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Minx, J.C. (Eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Kuznets, S., 1948. National income: a new version. Rev. Econ. Stat. 30 (3), 151–179. https://doi.org/10.2307/1926746.

Landefeld, J.S., Moulton, B.R., Vojtech, C.M., 2003. Chained-dollar indexes: issues, tips on their use, and upcoming changes. Surv. Curr. Bus. 83 (11), 8–16. Retrieved from. https://www.bea.gov/scb/pdf/2003/11November/1103Chain-dollar.pdf.

Leimbach, M., Kriegler, E., Roming, N., Schwanitz, J., 2017. Future growth patterns of world regions – a GDP scenario approach. Glob. Environ. Chang. 42, 215–225. https://doi.org/10.1016/j.gloenvcha.2015.02.005.

Liddle, B., Huntington, H., 2021. There’s technology improvement, but is there economy-wide energy leapfrogging? A country panel analysis. World Dev. 140, 105259. https://doi.org/10.1016/j.worlddev.2020.105259.

Macknick, J., 2011. Energy and CO2 emission data uncertainties. Carbon Manag. 2 (2), 189–205. https://doi.org/10.4155/cmt.11.10.

Marques, A., Rodrigues, J., Lenzen, M., Domingos, T., 2012. Income-based environmental responsibility. Ecol. Econ. 84, 57–65. https://doi.org/10.1016/j.ecolecon.2012.09.010.

Marshall, A., 1890. Principles of Economics. Macmillan, London.

Mazzucato, M., 2018. The Value of Everything. Allen Lane.

Missemer, A., Nadaud, F., 2020. Energy as a factor of production: historical roots in the American institutionalist context. Energy Econ. 86, 104706. https://doi.org/10.1016/j.eneco.2020.104706.

Moos, K.A., 2021. The historical evolution of the cost of social reproduction in the United States, 1959–2012. Rev. Soc. Econ. 79 (1) https://doi.org/10.1080/00346764.2019.1703031, 51-75n.

Moulton, B.R., 2004. The system of national accounts for the new economy: what should change? Rev. Income Wealth 50 (2), 261–278. https://doi.org/10.1111/j.0034-6586.2004.00125.x.

Nagaraj, R., Sapre, A., Sengupta, R., 2021. Four Years after the Base-Year Revision: Taking Stock of the Debate Surrounding India’s National Accounts Estimates. Indira Gandhi Institute of Development Research. Mumbai Working Papers, 2021–006.

Nakićenović, N., Alcamo, J., Davis, G., de Vries, H.J.M., Fenhann, J., Gaffin, S., Dadi, Z., 2000. Special Report on Emissions Scenarios. Cambridge University Press, Cambridge.

Nakićenović, N., Grübler, A., Gaffin, S., Jung, T.T., Kram, T., Morita, T., Victor, N., 2003. IPCC Sres revisited: a response. Energy Environ. 14 (2–3), 187–214. https://doi.org/10.1260/095830503765184592.

Naqvi, A., Zwickl, K., 2017. Fifty shades of green: revisiting decoupling by economic sectors and air pollutants. Ecol. Econ. 133, 111–126. https://doi.org/10.1016/j.ecolecon.2016.09.017.

Nordhaus, W.D., 2007. Alternative measures of output in global economic-environmental models: purchasing power parity or market exchange rates? Energy Econ. 29 (3), 349–372.

Özcan, B., Öztürk, I., 2019. Environmental Kuznets Curve (EKC): A Manual. Elsevier, London.

Pant, H.M., Fisher, B.S., 2007. Alternative measures of output in global economic-environmental models: purchasing power parity or market exchange rates? - comment. Energy Econ. 29 (3), 375–389.

Peters, G.P., Hertwich, E.G., 2008. CO2 embodied in international trade with implications for global climate policy. Environ. Sci. Technol. 42 (5), 1401–1407. https://doi.org/10.1021/es072023k.

Pigou, A.C., 1920. The Economic of Welfare. Macmillan, London.

Pitcher, H.M., 2009. Measuring income and projecting energy use. Clim. Chang. 97 (1), 49–58. https://doi.org/10.1007/s10584-009-9696-x.

Pollin, R., 1997. The Macroeconomics of Saving, Finance, and Investment. University of Michigan Press, Ann Arbor.

Pollin, R., 2018. De-growth vs A green new deal. New Left Rev 112, 5–25.

Pollin, R., 2019. Advancing a viable global climate stabilization project: degrowth versus the green new deal. Rev. Rad. Polit. Econ. 51 (2), 311–319. https://doi.org/10.1177/0486613419833518.

Reddy, S., 2008. The World Bank’s new poverty estimates: digging deeper into a hole. Challenge 51 (6), 105–112. https://doi.org/10.2753/0577-5132510607.

Reich, U.P., 2001. National Accounts and Economic Value. Palgrave, Basingstoke.

Requena-i-Mora, M., Brockington, D., 2022. Seeing environmental injustices: the mechanics, devices and assumptions of environmental sustainability indices and indicators. J. Polit. Ecol. 28 (1–28) https://doi.org/10.2458/jpe.4765.

Ruggles, R., 1990. A note on the revision of the United Nations System of National Accounts. Rev. Income Wealth 36 (4), 413–419. https://doi.org/10.1111/j.1475-4991.1990.tb00323.x.

Sarkodie, S.A., Strezov, V., 2019. A review on Environmental Kuznets Curve hypothesis using bibliometric and meta-analysis. Sci. Total Environ. 649, 128–145. https://doi.org/10.1016/j.scitotenv.2018.08.276.

Schor, J.B., Jorgenson, A.K., 2019. Is it too late for growth? Rev. Rad. Polit. Econ. 51 (2), 320–329. https://doi.org/10.1177/0486613419831109.

Schurr, S., 1984. Energy use, technological change, and productive efficiency: an economic-historical interpretation. Annu. Rev. Energy 9, 409–425.

Schurr, S.H., Netschert, B.C., Eliasberg, V.F., Lerner, J., Landsberg, H.H., 1960. Energy in the American Economy : 1850–1975. Johns Hopkins Press, Baltimore.

Seltzer, W., 1994. Politics and statistics: Independence, dependence or interaction? UN Department for Economics and Social Information and Policy Analysis WP Series, 6.

Semieniuk, G., Weber, I., 2020. Inequality in energy consumption : statistical equilibrium or a question of accounting conventions? Eur. Phys. J. Spec. Top. 229, 1705–1714. https://doi.org/10.1140/epjst/e2020-900125-5.

Semieniuk, G., Taylor, L., Rezai, A., Foley, D.K., 2021. Plausible energy demand patterns in a growing global economy with climate policy. Nat. Clim. Chang. 11 (4), 313–318. https://doi.org/10.1038/s41558-020-00975-7.

Semieniuk, G., Ghosh, J., Folbre, N., 2023. Technical comment on "Fairness considerations in global mitigation investments". Science *380* (6646). https://doi.org/10.1126/science.adg5893 eadg5893.

Shaikh, A., Tonak, E.A., 1994. Measuring the Wealth of Nations. Cambridge University Press, Cambridge, England and New York.

Siesto, V., 1987. Concepts and methods involved in the last revision of Italy's GDP. LABOUR 1 (3), 181–200. https://doi.org/10.1111/j.1467-9914.1987.tb00125.x.

Stern, D.I., 2004. The rise and fall of the environmental Kuznets curve. World Dev. 32 (8), 1419–1439. https://doi.org/10.1016/j.worlddev.2004.03.004.

Stern, D.I., 2017. How accurate are energy intensity projections? Clim. Chang. 143 (3–4), 537–545. https://doi.org/10.1007/s10584-017-2003-3.

Studenski, P., 1958. The Income of Nations. New York University Press, New York.

Summers, R., Heston, A., 1991. The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988. Q. J. Econ. 106 (2), 327. https://doi.org/10.2307/2937941.

Tercioglu, R.B., 2021. Rethinking growth and inequality in the US : what is the role of measurement of GDP? Int. Rev. Appl. Econ. 35 (3–4), 551–576. https://doi.org/10.1080/02692171.2021.1892038.

Torras, M., Boyce, J.K., 1998. Income, inequality, and pollution: a reassessment of the environmental Kuznets curve. Ecol. Econ. 25 (2), 147–160. https://doi.org/10.1016/S0921-8009(97)00177-8.

Tregenna, F., 2009. Characterising deindustrialisation: an analysis of changes in manufacturing employment and output internationally. Camb. J. Econ. 33 (3), 433–466. https://doi.org/10.1093/cje/ben032.

Tukker, A., Pollitt, H., Henkemans, M., 2020. Consumption-based carbon accounting : sense and sensibility. Clim. Pol. *20* (Sup1) https://doi.org/10.1080/14693062.2020.1728208. S1–S13.

van Benthem, A.A., 2015. Energy leapfrogging. J. Assoc. Environ. Resour. Econ. 2 (1), 93–132. https://doi.org/10.1086/680317.

Van Walbeek, C., 2006. Official revisions to south African national accounts data: magnitudes and implications. South African J. Econ. 74 (4), 745–765. https://doi.org/10.1111/j.1813-6982.2006.00094.x.

Vanoli, A., 2005. A History of National Accounting. IOS Press, Amsterdam and Oxford.

Victor, P.A., 2008. Managing without Growth. Routledge, Cheltenham, UK, and Northampton, MA.

Wagman, B., Folbre, N., 1996. Household services and economic growth in the United States, 1870-1930. Fem. Econ. 2 (1), 43–66. https://doi.org/10.1080/738552685.

Weber, I.M., 2021. How China escaped shock therapy: The market reform debate. Routledge, New York.

Weber, I., Semieniuk, G., Liang, J., Westland, T., 2022. Persistence in World Export Patterns and Productive Capabilities Across Two Globalizations. *UMass Economics Working Paper Series, 328*. https://doi.org/10.7275/t1w3-w223.

West, J., Lieber, M., Wang, H., 2021. Technical Annex for Global Material Flows Database - 2021 Edition. UNEP International Resource Panel.

Wiedenhofer, D., Virág, D., Kalt, G., Plank, B., Streeck, J., Pichler, M., Haberl, H., 2020. A systematic review of the evidence on decoupling of GDP, resource use and GHG emissions, part I: bibliometric and conceptual mapping. Environ. Res. Lett. 15, 063002 https://doi.org/10.1088/1748-9326/ab8429.

Wiedmann, T., Schandl, H., Lenzen, M., Moran, D., Suh, S., West, J., Kanemoto, K., 2015. The material footprint of nations. Proc. Natl. Acad. Sci. 112 (20), 6271–6276. https://doi.org/10.1073/pnas.1220362110.

Wiedmann, T., Lenzen, M., Keyßer, L.T., Steinberger, J.K., 2020. Scientists' warning on affluence. Nat. Commun. 11 (1), 3107. https://doi.org/10.1038/s41467-020-16941-y.

Witt, U., Gross, C., 2020. The rise of the "service economy" in the second half of the twentieth century and its energetic contingencies. J. Evol. Econ. 30 (2), 231–246. https://doi.org/10.1007/s00191-019-00649-4.

Wolff, E.N., 1987. Growth, Accumulation, and Unproductive Activity: An Analysis of the Postwar U.S. Economy. Cambridge University Press, New York.