POLITICAL ECONOMY RESEARCH INSTITUTE

PERI

University of Massachusetts Amherst

Inconsistent Definitions of GDP: Implications for Estimates of Decoupling

Gregor Semieniuk

September 2022

WORKINGPAPER SERIES

Number 563

POLITICAL ECONOMY RESEARCH INSTITUTE

Inconsistent Definitions of GDP: Implications for Estimates of Decoupling[1]

Gregor Semieniuk[2]
This version 9/21/22

Abstract:
Efforts to assess the possibilities for decoupling economic growth from negative environmental impacts have examined their historical relationship, with varying and inconclusive results. Part of the problem is ambiguity about definitions of environmental impacts, e.g. whether to use territorial or consumption-based measures of environmental impact. This paper shows that ambiguities arising from definitional changes to GDP are sufficiently large to affect the outcomes. I review the history of structural revisions to GDP using the example of the United States, and on international comparisons of purchasing power parity, compare decoupling results using various historical definitions of GDP on the same environmental indicator, and demonstrate that changing the GDP data vintage does impact decoupling results in qualitatively important ways, with and without purchasing power parity. Inconsistencies in economic measurement introduce an additional layer of ambiguity into historical decoupling evidence and model projection into the future. To advance debate and be clear about scenario assumptions, rigorous reporting of GDP definitions used and the sharing of data vintage for subsequent comparison and replication are urgently needed.

Keywords: decoupling; national accounting; GDP revisions; energy intensity; environmental Kuznets curve

1. Introduction

Understanding the history of the relationship between gross domestic production (GDP) and resource and energy use and pollution is important for thinking about the future. On the question whether GDP growth is compatible with nongrowing or even declining rates of resource use and pollution, 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 or pollution and GDP is often expressed as an intensity with GDP in the denominator. A declining intensity is referred to as decoupling. Patterns of historical decoupling are contested. They influence recommendations about whether countries should “first grow and clean up later” (Dinda, 2004), and have spawned large “growth-environment-nexus” and decomposition literatures; a recent review examined 835 empirical studies of decoupling (Wiedenhofer *et al.*, 2020). The numerator of measures of intensity has been close scrutinized conceptually, leading to competing intensity and decoupling measures to assess environment impact of economic activity, e.g. whether to use footprint or

[1] Acknowledgments: I thank Duncan Foley for introducing me to variations in GDP definitions that inspired this research, and Nancy Folbre, Duncan Foley, Roger Fouquet, Lukas Hardt, Carol Heim, Antoine Missemer, Robert Pollin, Ellis Scharfenaker, Isabella Weber 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.
[2] Political Economy Research Institute and Department of Economics, University of Massachusetts Amherst, gsemieniuk@umass.edu

territorial measures[3] or how to account for primary energy[4]. Meanwhile, to the author's best knowledge, redefinitions of the denominator of any intensity, GDP, has never featured in the discussion of decoupling estimates.

But GDP is an accounting convention. Its measurement depends on social agreement, not on natural constants. In fact, at any moment there is more than one GDP measure available, and 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 (e.g. Studenski, 1958; Kendrick, 1970; 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, Dupuy and McLaughlin, 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 decoupling, e.g. whether or not a country is seen to have decoupled in the past, is vital. In this paper I review how GDP measures regularly undergo revisions, and then empirically analyze their consequences for intensity measures and decoupling.

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 (i.e. the GDP time series published in a certain year in the past, e.g. a vintage published in 1965) from the archive of the *Survey of Current Business*. I also collect vintages for a large number of countries back to the 1990s from the national accounts data in a supplement to successive version of the Penn World Table (PWT). Finally, I collect global GDP from the last 8 editions of the International Energy Agency's (IEA) World Indicators. I combine these times series for GDP with data on primary energy for most countries for the period 1950-2014 to examine changes to 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 what a retrospective analysis would yield. Finally I examine how the IEA's GDP vintage data changes our understanding of the decline of the historical global energy intensity.

The key result is that evidence for decoupling does vary in an economically important way with GDP revisions over time. Some countries decouple or recouple in the same time period, depending on GDP vintage used. Some of the Grossman and Krueger results change qualitatively when a more recent GDP vintage is used. And the IEA's energy intensity accelerates retrospectively as global GDP growth is continually revised upwards. These results contribute to the debate about decoupling by showing that results are, to some extent, contingent on data vintage used. 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

---

[3] Territorial measures allocate resource use to the country where they enter production processes, footprint measures where final products are consumed. Territorial measures typically paint a more favorable picture of high-income country decoupling due to the global division of labor and its history (Peters and Hertwich, 2008; Wiedmann *et al.*, 2015; Jiborn *et al.*, 2018; Tukker, Pollitt and Henkemans, 2020; Akizu-Gardoki *et al.*, 2021; Weber *et al.*, 2021).

[4] 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 (Macknick, 2011; Krey *et al.*, 2014; Koomey *et al.*, 2019; Kraan *et al.*, 2019; Semieniuk and Weber, 2019).

few years difference in the collection of GDP data can impact measurement. And while one can argue that the current version of GDP is the most useful for our moment in time, that also implies that previous GDP vintages are better for earlier years as they were 'current' in their day. 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, while policy targets need to be precise about what decoupling they refer to or avoid reference to aggregate intensities.

The next section reviews GDP revisions both at the national and international level at some length with a view to familiarizing researchers in environmental and ecological economics with them. It also gives examples of how GDP revisions impact its magnitude and rate of change, and reviews related literature. Section 4 introduces the method of analysis and all data sources. Section 5 presents results of the impact of GDP vintage on decoupling and discusses them. Section 6 concludes with 3 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 and energy inputs or unwanted or unintended physical outputs (pollution), 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*.[5] 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 (Van Walbeek, 2006; Fixler, Francisco and Kanal, 2021). Past research documented an upward bias in some such revisions (Glejser and Schavey, 1969; Franses, 2009). 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

[5] 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 dependence of results on GDP revisions.

Croushoure and Stark (2003) terminology, they will be called *structural revisions* here. Section 3.1 illustrates the impact of these revisions using U.S. data. Section 3.3 briefly illustrates the better-known changes to purchasing power parities between countries after explaining the concept. Section 3.4 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.[6] Appendix A reviews each of them in detail. Here I show how structural revisions continually take place and impact GDP measurement, using the US example.

Statistics from the United States National Income and Product Accounts, one of the most detailed and long-lived system 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 the 1990s instead of its close relative, GDP.[7] 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 is long-term growth". This reassurance reveals that the redefinition has changed (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 nonmarket 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 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.[8] Since the UN's SNA serves as a benchmark for internationally comparable national

[6] 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.

[7] 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, and the growth rates of the two measures tend to be very similar.

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

accounts, it was also implemented in the US national accounts in 2013. Subsequently economists continue 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 Figure 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. Growth was cumulatively more than 40% faster as the orange series measured on the right hand axis shows. There was some volatility in the late 1940s, 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 1987it would deliver a strikingly lower GDP growth rate, and hence fewer prospects for decoupling than with current GDP estimates.

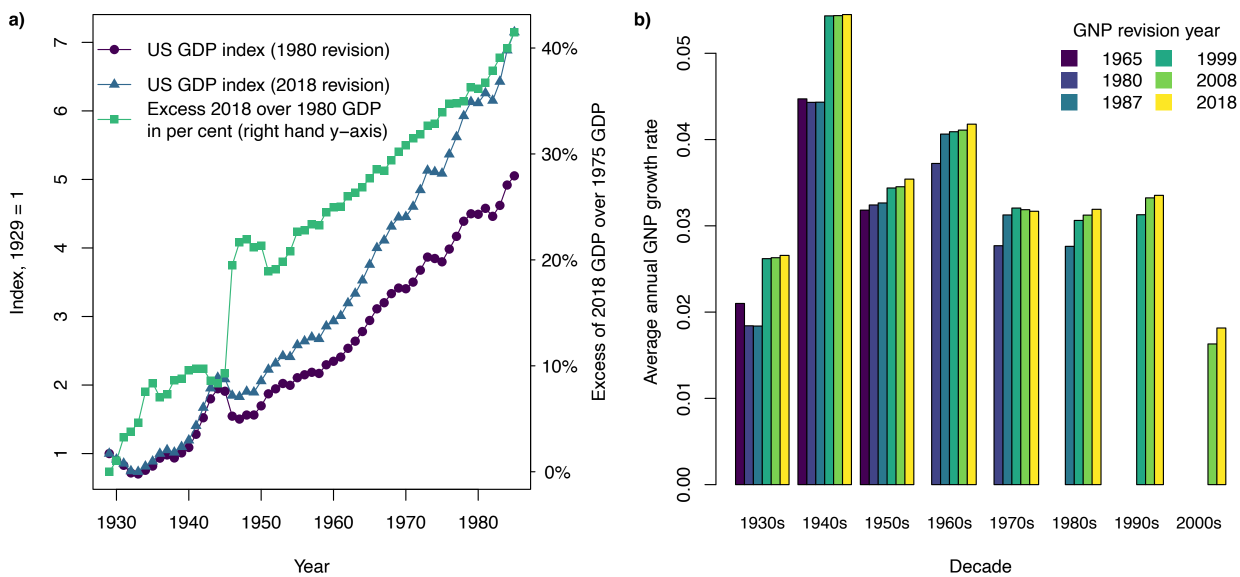

Figure 1 | Changes in growth rates between vintages of US GDP or GNP: (a) GDP indices from 1986 (black) and 2021 (blue) based on 1976 and 2018 structural revisions, and their ratio in orange measured on the right-hand side axis. (b) Annual average growth rate over decades of selected structural revisions,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

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 more systematically 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. The example of Schurr's energy intensity in Figure 1b illustrates how such differences can reverse long-term decoupling results. 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 (Vanoli, 2005; Bos, 2006).

As an illustration of how such revisions can have a qualitative impact on 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 long before association with Simon Kuznets' name, and updated by Schurr (1984).[9] Schurr describes a period of decoupling from 1930 to about 1945, followed by stability until 1970 and then again decoupling (Figure 2, black series). However, trends calculated using modern data show more consistent and rapid decoupling (Figure 2, blue series). [10] Moreover, the historical data show 3 five-year intervals with 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.

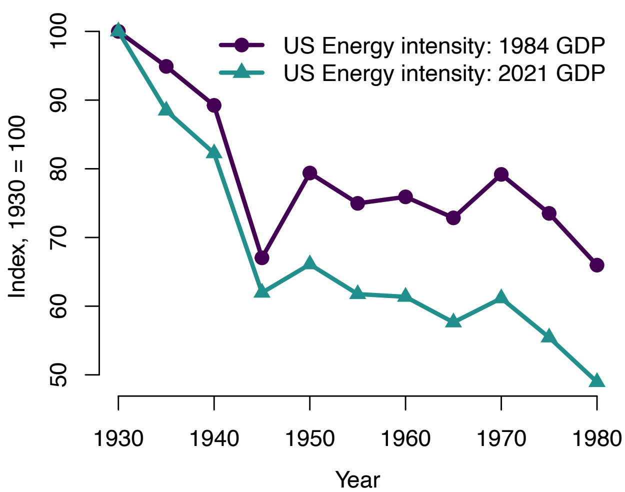

Figure 2 | Impact of US GDP vintage on energy intensity. Primary energy intensity as in Schurr (1984) but adding the 2021 vintage of US GDP. Sources described in section 3.

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 stock markets and data repositories. However, economists have long debated over whether this is the appropriate approach (Kravis, Heston and Summers, 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 it is formed from the demand and supply of traded goods (and 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 even after exchanging at MER. Therefore, similar to stripping out inflation to compare a country's growth over time, 'purchasing power parity' (PPP) has attempted to adjust countries' economies for varying cross-sectional price levels (see Appendix B for an example).

---

[9] 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).

[10] 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 gross national product (GNP), which was used in the US until the 1990s.

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 (2021). Discussions of reasons behind revisions over time are discussed in Deaton and Heston (2010) and in Feenstra, Inklaar and Timmer (2013). Some believe 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, 2022).

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 (the same is true of MER GDP due to varying MER). Section 4 will elaborate as necessary.

To get a sense of magnitudes involved in level changes, Figure 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, it 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 and undulating movement across versions. Our focus is on decoupling and rates of change, and cannot examine levels in detail.

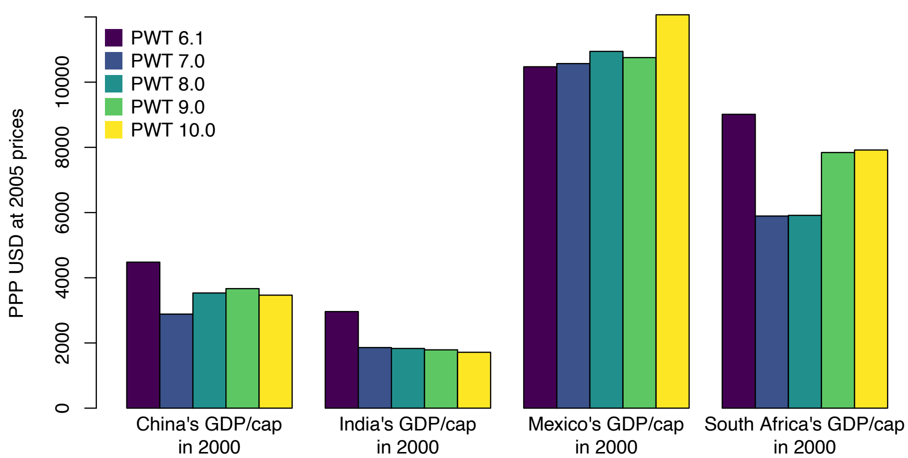

Figure 3 | PPP GDP per capita in 2000 according to successive Penn World Table estimates: PPP GDP per capita in 2000 for four countries at 2005 USD price levels, taken from 5 versions of Penn World Tables published over the period 2002-2020. Sources described in section 3.

To further illustrate how, even at one point in time, PPP exchange rates can introduce an additional GDP measure with consequences for 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 authors indicate their global GDP source 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 momentous implications for the conclusions drawn. If GDP is measured instead at PPP, the entire period is one of relative decoupling. Figure 4 replicates their graph, but also adds the PPP GDP index, according to which resources 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.[11]

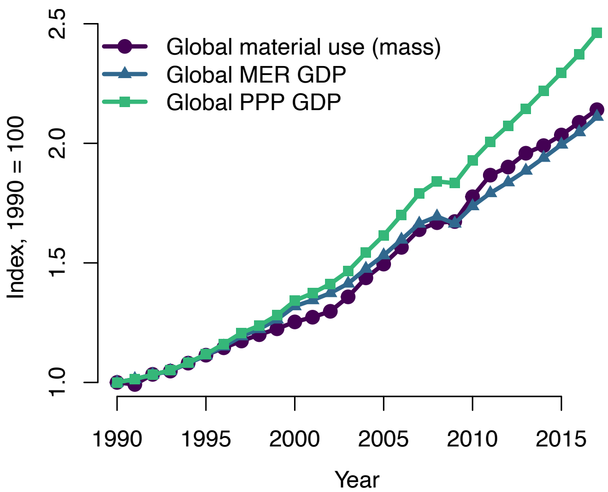

Figure 4 | GDP data definition effect on decoupling: Global material consumption vs GDP indices as in Hickel and Kallis (2020) but also GDP at purchasing power parity.

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 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, Heston and Summers, 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.

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

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

to very high growth rates in developing countries. The report's authors retorted that they were modeling economic activity, not standards of living, among other arguments (Nakicenovic *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 GDP, 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).

Two other papers are worth mentioning. 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. And 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 (the conclusion is 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. In sum, despite the important role of structural GDP revisions, to my best knowledge the question of how this relates to decoupling estimates has not been systematically analyzed. Method and data for doing so are introduced next.

3. Method 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 compare rates of decoupling. I make pairwise comparisons of the same measure using different vintages of data. The aim is to examine changes in degree of decoupling and even of sign over business-cycle length intervals contingent on data vintage used. Since rates of change are dimensionless this method of analysis easily spans every possible combination of definitions of GDP. For n different vintages, the possible combinations of vintages are $n!/(n-2)!/2!$. Given n=11 vintages found 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 11 older vintages.

Another method is to re-estimate important results in the literature with varying GDP to check for the salience of changes in estimated decoupling. I pick the examples of the seminal environmental Kuznets curve study by Grossman and Krueger (1995) and the widely used historical global decoupling estimates for energy by the IEA (2022) to check whether country-level variations matter in the aggregate. Since several countries are involved, purchasing power parities creep in, and dilute the impact of national GDP redefinitions. Nevertheless, since PPPs also form part of GDP redefinitions for international comparisons, the results speak to the problem of GDP vintages, too.

I use various 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 was a revision 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.[12]

For multi-country analysis, PWT GDP is from the website of the Groningen Growth and Development Centre (Feenstra, Inklaar and Timmer, 2015).[13] GDP for a large number of countries is reported in altogether eleven vintages starting with data in 1950 and running until a few years before the release (see 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 $CHKON_t$, $GKON_t$, $IKON_t$, $IMPK_t$ and subtracting $EXPK_t$. I calculated average annual growth rates, *g*, at from time *t* over *s* years as $g_{t,t+s}=(GDP_{t+s} / GDP_t)$^(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.[14]

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	2000
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

Global and regional GDP data are from the IEA World Indicators. The IEA make 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 purchasing power parity GDP, from 1990 onwards. Prior to that the IEA converts its market exchange rates "based on the PPP conversion factor (GDP) to market exchange rate

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

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

[14] 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.

ratio" (IEA 2020, p. 26). World Bank data from the World Development Indicators downloaded in 2019 also underlies Figure 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 are from the IEA and the United Nations and the dataset description in Semieniuk et al. (2021). They are used to form energy intensities. The figure 2 example uses U.S. energy data from the Energy Information Agency website.[15] Material data as in Hickel and Kallis is from the International Resource Panel.[16] Pollution 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 intensity rates of changes, Figure 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. The Pearson correlation coefficient is only 0.86 for plot (a) which compares PWT10.0 with the oldest available vintage, PWT 5.6. The data cloud is also centered away from the origin, documenting an upward translation in revising GDP growth rates. Consequently, more decoupling takes place in PWT10.0 simply by GDP revision. Since all later vintage plots (b-d) are both roughly centered around zero when compared with PWT10.0 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 Figure 5 plots (b-d) most of the observations remain in the corridor of +/-2 percentage points difference. Still, the mean difference in annual growth rates is 0.8 percentage points in plot (b), and 0.6 and 0.5 in plots (c) and (d) respectively. For these averages, compound growth leads differences of about 8, 6 and 5 percent in the estimated energy intensity over a decade. Moreover, a remarkable number of observations lie far below the corridor, even between PWT 8 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 (van Benthem, 2015; Liddle and Huntington, 2021).

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

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

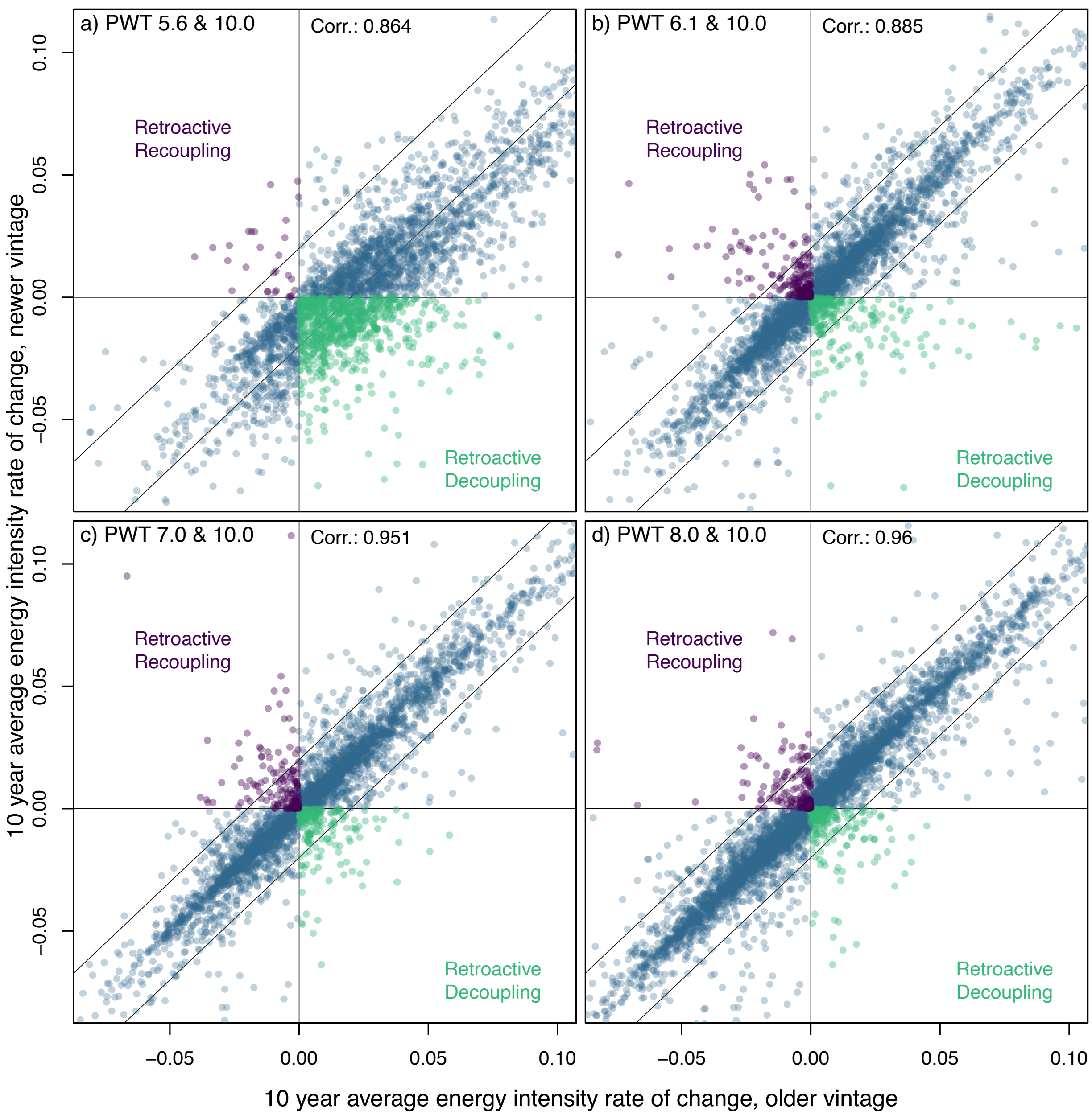

Figure 5 | Variance in energy intensity rates of change continent on GDP vintage: In each plot, the x-axis reports the 10-year average energy intensity rate of change in an older PWT version, the y-axis for PWT 10.0, noted in the plot's top left. The diagonal corridor includes observations with less than 2 percentage point difference in *annual* growth rate. Green and violet observations see retrospective switches from decoupling to recoupling and vice versa.

But 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 decouple and vice versa. Retrospective decoupling is highlighted green and recoupling dark blue in Figure 5. Figure 6 plots these observations as a share of total 10-year growth rate observations. The number of countries reporting a switch is quantitatively important. Researchers studying the problem some 20 years apart (PWT 6.1 vs 10.0 in plot (a)) would find roughly 15% of countries switching sign in any 10-year period starting in the mid 1960s. 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 8.0 and 10.0 (Figure 5b), only 8 years apart, some more than 5 percent of countries flip sign on average, again with a slight bias towards retrospective decoupling. In other words, depending on vintage comparison, about 10 to 30 countries would be found to have the opposite behavior over any given 10 year interval purely due to structural GDP revisions from the 1960s onward.

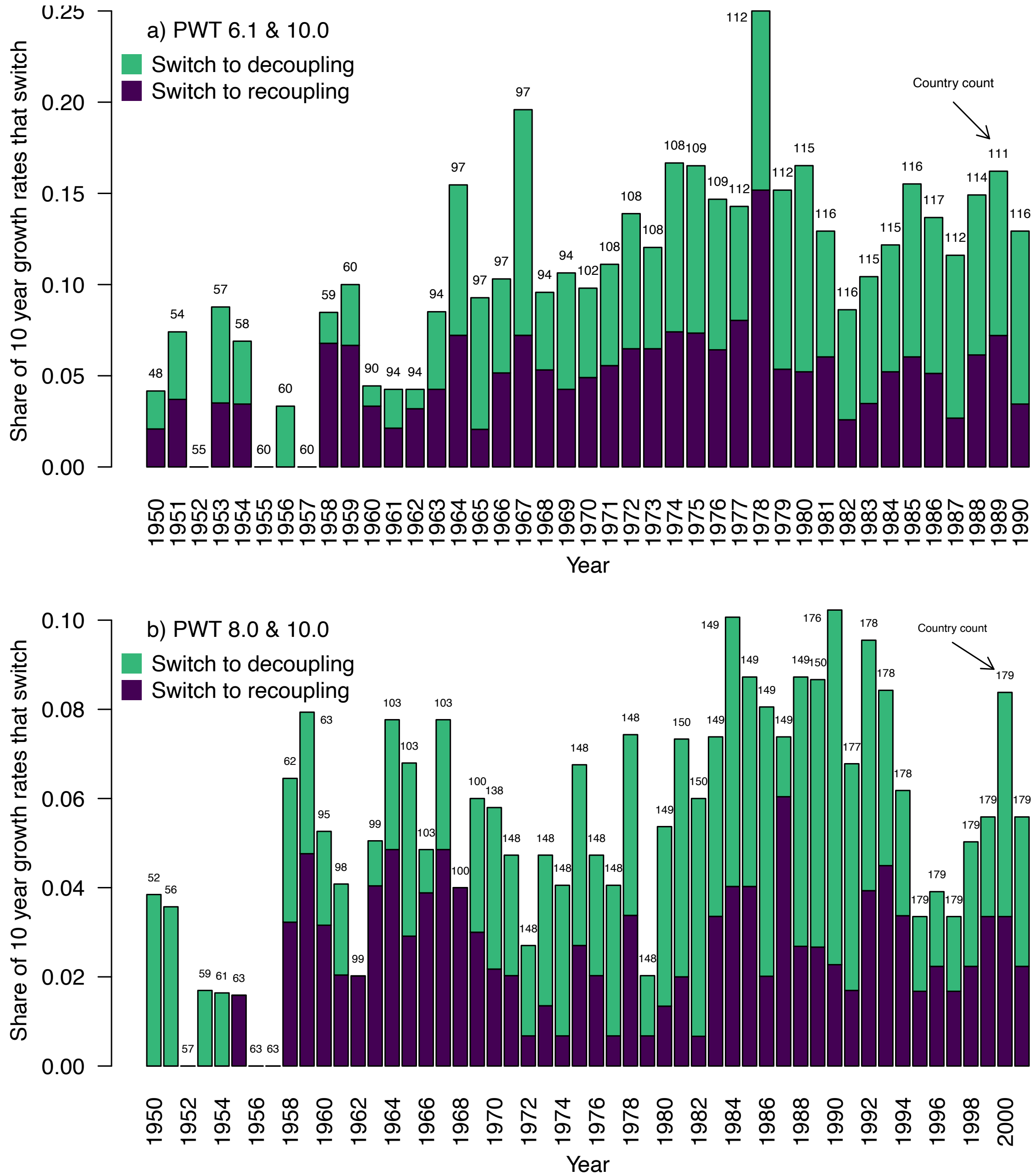

Figure 6 | Share and number of countries with retrospective decoupling or recoupling over 10 year periods.

Even more remarkably, a number of countries switch the sign of their GDP growth over periods of 10 years, introducing the potential to switch from/to absolute decoupling. Thus, when switching from PWT6.1 to PWT10.0, 3 to 6 countries over any given 10-year interval are found to have been in a long-term depression rather than a decade of positive economic growth (Figure 7a). Fewer cases go the other way. From PWT8.0 to 10.0, about the same numbers of countries switch in both directions (Figure 7b). In other words, even absolute decoupling is to some extent conventional.

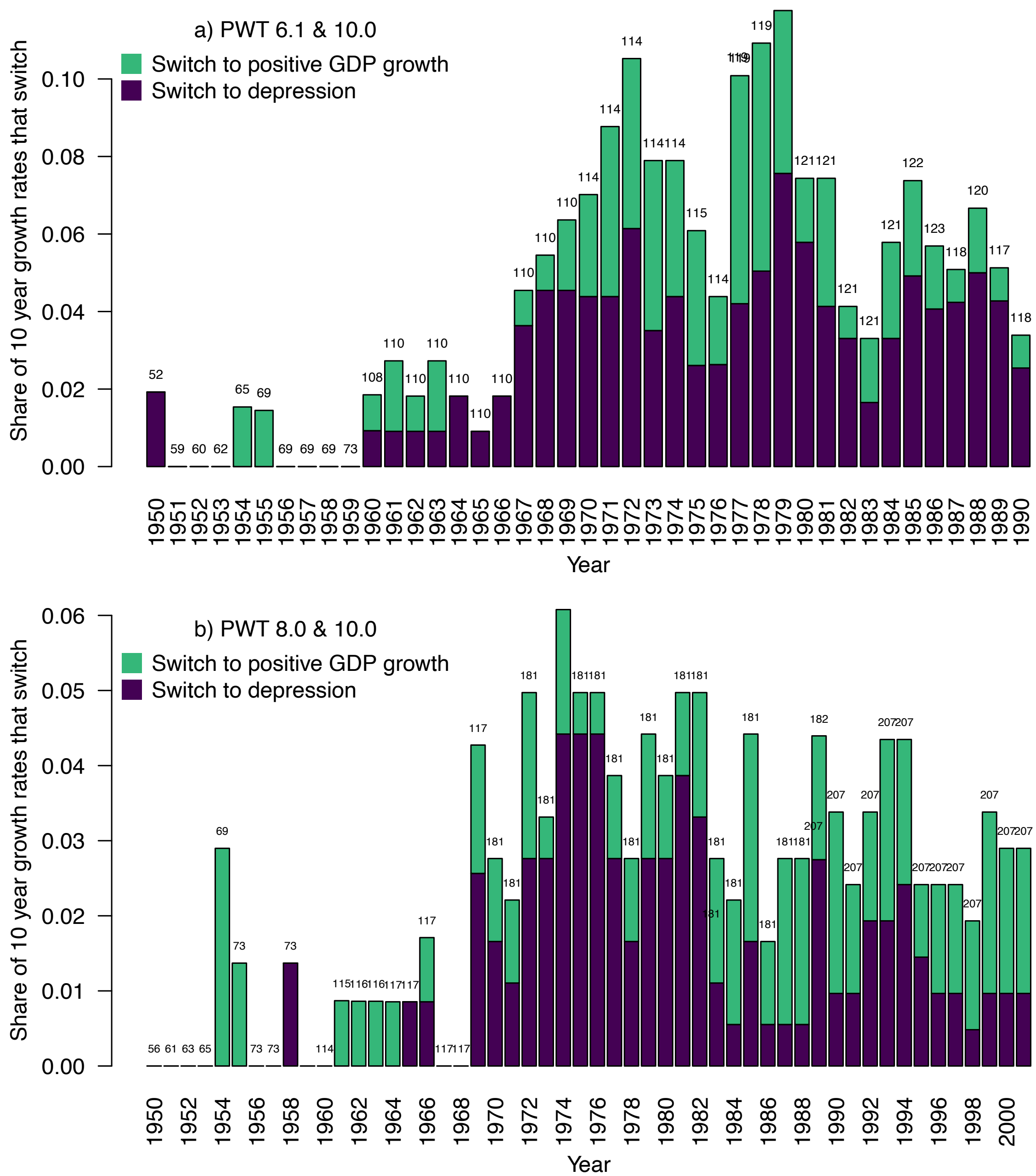

Figure 7 | Share and number of countries with retrospective switch in the sign of GDP rate of change over 10 years.

4.2. Impact on environmental Kuznets curve estimates

Two examples illustrate the relevance of these changes for research results. First consider pollution intensities, the magnitude of interest for the environmental Kuznets curve. The seminal paper by Gene Grossman and Alan Krueger (1995) had over 9,000 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).[17] In the 1995 paper, 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" (p. 353). This inverted u-shape (with GDP/capita on the x-axis) fuelled the take-off of the environmental Kuznets curve literature, which continues – with ongoing controversy, due to its policy message – to this day (Dinda, 2004; Stern, 2004; Carson, 2010; Özcan and Öztürk, 2019; Sarkodie and Strezov, 2019).

The Grossman and Krueger (1995) results have subsequently been subjected to scrutiny. Harbaugh, Levinson and Wilson (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. 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. Hence, 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.

Grossman and Krueger estimate cubic polynomials of GDP and its lags and plot their results which gives a powerful impression of the inverted u shape. I reestimate their curves using their Stata programs and scaling all GDP vintages to 2017 US-dollars. The rgdpch measure isn't available in PWT 10.0 and I report instead all three alternative measures of GDP. rgdpe measures standards of living while interpolating between benchmark years, i.e., the information from previous international price comparisons is used, an innovation from PWT 8.0. rgdpo does the same but includes price indices from exports and imports, and so calculates a country's productive capacity rather than the standard of living (which excludes export and import prices). rgdpna extrapolates from the 2017 benchmarks using national accounts growth rates (like we use above). While the PWT creators do not recommend using rgdna for both cross-country and time comparison, it is closest to the previous measures, so I report it for comparison (Feenstra, Inklaar and Timmer, 2015, p. 3157).

[17] 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.

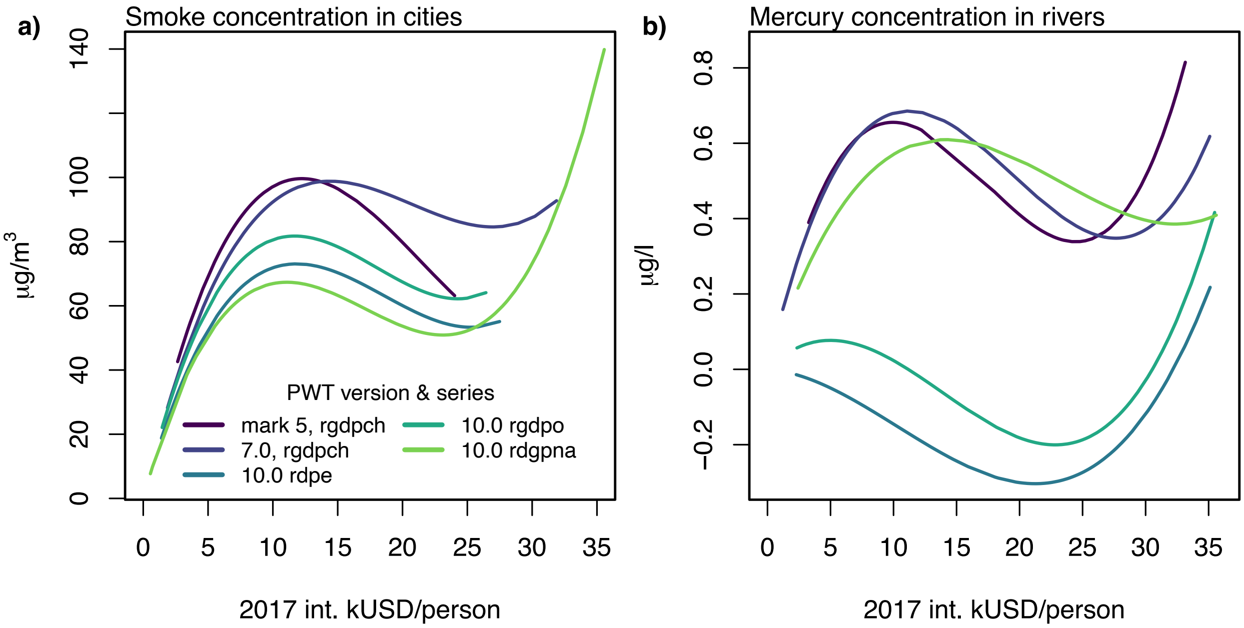

Figure 8 | Grossman and Krueger re-estimates with different PWT vintages for selected pollutants.

Figure 8 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 but drops very little 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 C 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 as GDP per capita grows.

4.3. Impact on global energy intensity estimates

As a second example, consider the IEA's estimate of changes in global historical energy intensity. These data are e.g. used as a historical benchmark for assessing decoupling assumptions in future energy and climate scenarios (see e.g. IPCC 2018, SPM1). The IEA publishes a new vintage of its dataset every year, which updates its energy and GDP data, including an estimate for the world. Figure 9a shows an index of the world energy intensity for eight vintages. Until around 1990, not much difference can be detected but then the time series start to fan out. Newer vintages tend to show a faster decline. Over the period 1971-2010, the vintages from 2013 and 2015 had a compound annual decline of -0.8%, compared with -1.0% in the 2021 vintage. As the Figure 9b main window shows this led the 2021 vintage to report an 8% lower energy intensity by 2010 than the 2013 vintage. The insets in Fig 9b further show that this downward bias is due entirely to GDP accounting, not that for primary energy. The energy demand estimates fluctuate around the 2021 estimate with usually less than 0.5% deviation in any year and with some mean reversion. In contrast, all GDP estimates fall below the 2021 estimate after 1990 and bar the 2020 one diverge by 5% and more. It is rather interesting that the divergence only starts after 1990 since pre-1990 the IEA uses national accounts growth rates which appear to have varied less in the aggregate (see also Data section).

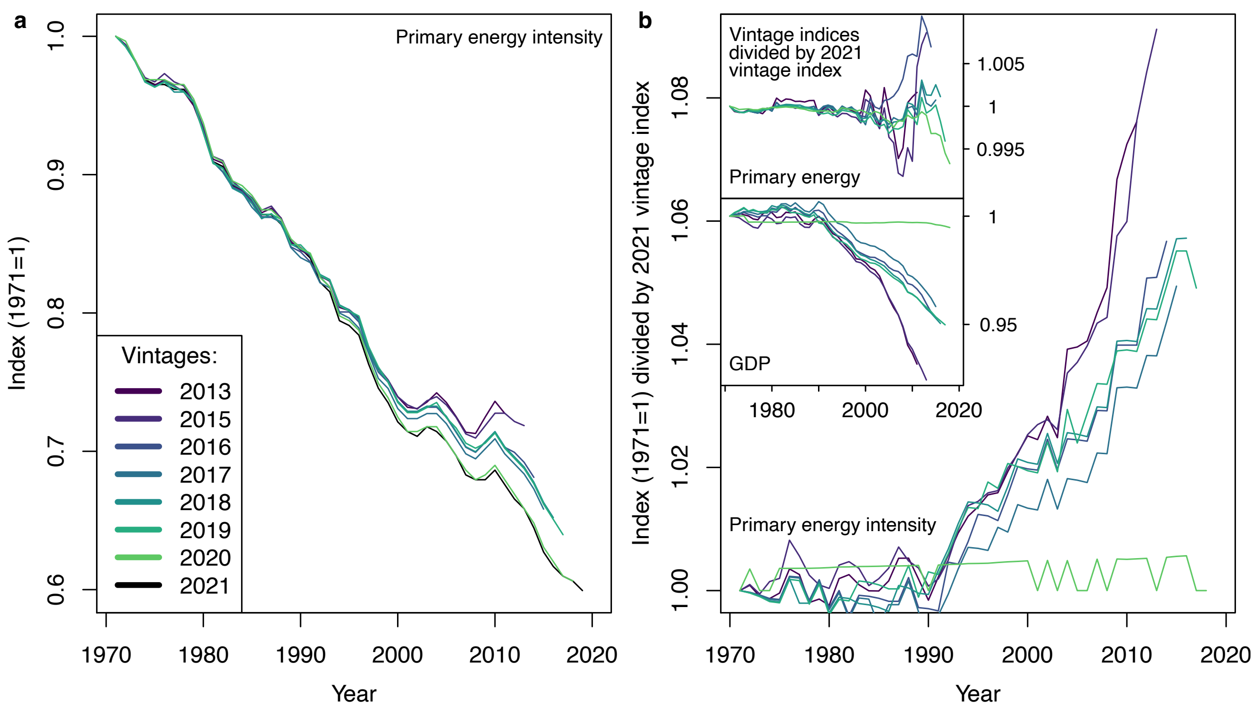

Figure 9 | Global energy intensity for vintages of the IEA World Indicators, using PPP GDP.

Such divergence over decades matters greatly for mitigation scenarios. Suppose modelers extrapolated the intensity trend over the next 40 years and updated with the faster intensity decline. 8% less energy demand for a given GDP in 2060 implies 8% less energy demand and less required mitigation. In principle, faster energy intensity decline should only reflect a faster growing GDP so the amount of energy demanded should not change. In practice, it is unclear to what extent these updates are synchronized. For instance that modelers who have exogenous GDP growth as an input 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 growth is ‘slow’, leading to an overall lower pressure on energy demand as an accounting artifact.

4.4. Discussion

The foregoing results show that evidence for decoupling varies in an economically important way with GDP revisions over time. It follows that the entrenched debate about whether EKCs exist or not, and the extent to which decoupling can be expected to lead to growing economy that sheds its current environmental impact, is marred by an ambiguity 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 between two structural GDP revisions a large share of countries would switch from positive to negative long-term GDP growth or vice versa and provide a completely different account about the potential for absolute decoupling. However, the results about the frequent switches in relative decoupling suggest that plenty of room for arguing about whether we are seeing a tendency toward or away from decoupling exists. And the results on the EKC do not seem to come to any agreement on a particular shape, at least for some pollutants. It is also important to note that GDP is but one more indicator that can change, compounding rather than introducing ambiguity into the results.

Another objection could follow national accountants to claim that GDP revisions render a superior picture of the economy as it is today. Therefore, results arrived at with later vintages should be privileged over older ones.[18] The problem introduced by this claim is that surely previous GDP vintages were 'superior' in their day. Is the GDP vintage in 2020 really better for understanding the 1980 economy than the 1980 vintage that was created by people living through the problems of the day? This line of reasoning with its emphasis on the current economy would suggest it is not. If this claim was made, moreover, it does not seem possible to use past trends arrived at with today's vintage to make confident projections into future. Surely future accountants will have a different view about what is important in the economy and will revise GDP accordingly.

If, with several 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 in the carbon emission 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 modelling 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.[19] All in all, measurement changes in GDP can have important real-world policy implications that are currently underappreciated in the environmental policy debate.

5. Conclusion

This paper has traced structural revisions in how GDP is accounted for and shown that these revisions impact measures of decoupling in a quantitatively and qualitatively important way. Some countries switch from decoupling to recoupling and vice versa, environmental Kuznets curve estimates are sensitive to the GDP vintage used and the IEA has been reporting

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

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

accelerating GDP growth rates for the world, leading to a faster historically estimate of energy intensity decline. The dilemma is that there is no particular reason why one vintage is better than the other. Decoupling analysts must recognize an ambiguity built into the denominator of their intensity measures, just as they have become accustomed to problems such as different patterns of de- and recoupling for territorial vs footprint measures of the numerator. This 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.

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, is 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 being commendable exceptions). And 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), so that later analysist have the ability to discriminate between models also along this dimension.

Second, the unreliability of some 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 stay 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, rather than variation over vintages being a conundrum, the variation serves as an opportunity for a robustness check on the susceptibility of current GDP estimates to political preoccupations of the day.

Bibliography

Acemoglu, D. *et al.* (2019) 'Democracy Does Cause Growth', *Journal of political Economy*, 127(1), pp. 47–100.
Akizu-Gardoki, O. *et al.* (2021) 'Hidden Energy Flow indicator to reflect the outsourced energy requirements of countries', *Journal of Cleaner Production*, 278, p. 123827. doi: https://doi.org/10.1016/j.jclepro.2020.123827.
Anand, S. and Segal, P. (2008) 'What Do We Know about Global Income Inequality?', *Journal of Economic Literature*, 46(1), pp. 57–94. doi: 10.1257/jel.46.1.57.
Assa, J. (2017) *The Financialization of GDP*. Oxford and New York: Routledge.
Assa, J. and 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 Economy. Taylor & Francis, 26(6), pp. 985–1014. doi: 10.1080/13563467.2020.1865899.
Basu, D. and Foley, D. K. (2013) ‘Dynamics of output and employment in the US economy’, *Cambridge Journal of Economics*, 37, pp. 1077–1106. Available at: file:///Users/gs299/Documents/Library.papers3/Articles/2013/Basu/Basu_2013_Dynamics of output and employment in the US economy.pdf.
van Benthem, A. A. (2015) ‘Energy Leapfrogging’, *Journal of the Association of Environmental and Resource Economists*, 2(1), pp. 93–132. doi: 10.1086/680317.
Bos, F. (2006) ‘The development of the Dutch national accounts as a tool for analysis and policy’, *Statistica Neerlandica*. John Wiley & Sons, Ltd, 60(2), pp. 225–258. doi: 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’, *The American Economic Review*. American Economic Association, 70(1), pp. 64–77. Available at: http://www.jstor.org/stable/1814738.
Brynjolfsson, E. and McAfee, A. (2014) *The Second machine age : work, progress, and prosperity in a time of brilliant technologies*.
Carson, R. T. (2010) ‘The Environmental Kuznets Curve: Seeking Empirical Regularity and Theoretical Structure’, *Review of Environmental Economics and Policy*. The University of Chicago Press, 4(1), pp. 3–23. doi: 10.1093/reep/rep021.
Castles, I. and Henderson, D. (2003) ‘The IPCC Emission Scenarios: An Economic-Statistical Critique’, *Energy & Environment*. SAGE Publications Ltd STM, 14(2–3), pp. 159–185. doi: 10.1260/095830503765184583.
Christophers, B. (2011) ‘Making finance productive’, *Economy and Society*. Available at: http://www.tandfonline.com/doi/abs/10.1080/03085147.2011.529337.
Corrado, C. *et al.* (2021) *Measuring and Accounting for Innovation in the Twenty-First Century*. University of Chicago Press. Available at: https://www.nber.org/books-and-chapters/measuring-and-accounting-innovation-twenty-first-century.
Costanza, R. *et al.* (1997) ‘The value of the world’s ecosystem services and natural capital’, *Nature*, 387(6630), pp. 253–260. doi: 10.1038/387253a0.
Coyle, D. (2014) *GDP: A short but affectionate history*. Princeton and Oxford: Princeton University Press.
Croushore, D. and Stark, T. (2003) ‘A real-time data set for macroeconomists: Does the data vintage matter?’, *Review of Economics and Statistics*, 85(3), pp. 605–617. doi: 10.1162/003465303322369759.
Deaton, A. and Aten, B. (2017) ‘Trying to understand the PPPs in ICP 2011: Why are the results so different’, *American Economic Journal: Macroeconomics*, 9(1), pp. 243–64. doi: 10.1257/mac.20150153.
Deaton, A. and Heston, A. (2010) ‘Understanding PPPs and PPP-based National Accounts’, 2(4), pp. 1–35.
Deaton, A. and Schreyer, P. (2021) ‘GDP, Wellbeing, and Health: Thoughts on the 2017 Round of the International Comparison Program’, *Review of Income and Wealth*. John Wiley & Sons, Ltd, n/a(n/a). doi: https://doi.org/10.1111/roiw.12520.
Dellink, R. *et al.* (2017) ‘Long-term economic growth projections in the Shared Socioeconomic Pathways’, *Global Environmental Change*. Elsevier Ltd, 42, pp. 200–214.
Desrosières, A. (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
Dinda, S. (2004) ‘Environmental Kuznets Curve Hypothesis: A Survey’, *Ecological Economics*, 49(4), pp. 431–455. doi: https://doi.org/10.1016/j.ecolecon.2004.02.011.
Eisner, R. and Nebhut, D. H. (1981) ‘AN EXTENDED MEASURE OF GOVERNMENT PRODUCT: PRELIMINARY RESULTS FOR THE UNITED STATES, 1946–76’, *Review of Income and Wealth*. John Wiley & Sons, Ltd, 27(1), pp. 33–64. doi:

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*. Cheltenham, UK, and Northampton, MA: Elgar.
European Commission *et al.* (2009) *System of National Accounts 2008*. New York: United Nations.
Feenstra, R. C., Inklaar, R. and Timmer, M. P. (2013) ‘PWT 8.0 - a user guide’.
Feenstra, R. C., Inklaar, R. and Timmer, M. P. (2015) ‘The Next Generation of the Penn World Table’, *American Economic Review*, 105(10), pp. 3150–3182. doi: 10.1257/aer.20130954.
Feldstein, M. (2017) ‘Underestimating the real growth of GDP, personal income, and productivity’, *Journal of Economic Perspectives*, 31(2), pp. 145–164. doi: 10.1257/jep.31.2.145.
Fixler, B. D. J., Francisco, E. De and Kanal, D. (2021) ‘The Revisions to Gross Domestic Product, Gross Domestic Income, and Their Major Components’, *Survey of Current Business*, 101(1).
Fixler, D. J., Greenaway-Mcgrevy, R. and Grimm, B. T. (2014) ‘The Revisions to GDP, GDI, and Their Major Components’, *Survey of Current Business*, 94(8), pp. 1–23.
Foley, D. K. (2013) ‘Rethinking Financial Capitalism and the “Information” Economy’, *Review of Radical Political Economics*, 45(3), pp. 257–268. doi: 10.1177/0486613413487154.
Franses, P. H. (2009) ‘Why is GDP typically revised upwards?’, *Statistica Neerlandica*, 63(2), pp. 125–130. doi: https://doi.org/10.1111/j.1467-9574.2008.00410.x.
Ghosh, J. (2020) ‘Global inequality in a time of pandemic’, *Real-world Economics Review*, 92, pp. 22–32.
Ghosh, J. (2022) ‘What Do We Really Know about Productivity Differentials across Countries?’, *Review of Radical Political Economics*. SAGE Publications Inc, p. 04866134221099507. doi: 10.1177/04866134221099507.
Glejser, H. and Schavey, P. (1969) ‘An Analysis of Revisions of National Accounts Data for 40 Countries’, *Review of Income and Wealth*, 20(3), pp. 317–332.
Grossman, G. M. and Krueger, A. B. (1991) ‘Environmental Impacts of a North American Free Trade Agreement’, *NBER Working Paper*, 3914.
Grossman, G. M. and Krueger, A. B. (1995) ‘Economic Growth and the Environment*’, *The Quarterly Journal of Economics*, 110(2), pp. 353–377. doi: 10.2307/2118443.
Hanley, N., Dupuy, L. and McLaughlin, E. (2015) ‘GENUINE SAVINGS AND SUSTAINABILITY’, *Journal of Economic Surveys*. John Wiley & Sons, Ltd, 29(4), pp. 779–806. doi: https://doi.org/10.1111/joes.12120.
Harbaugh, W. T., Levinson, A. and Wilson, D. M. (2002) ‘Reexamining the empirical evidence for an environmental Kuznets curve’, *Review of Economics and Statistics*, 84(3), pp. 541–551. doi: 10.1162/003465302320259538.
Hickel, J. and Kallis, G. (2020) ‘Is Green Growth Possible?’, *New Political Economy*. Taylor & Francis, 25(4), pp. 469–486. doi: 10.1080/13563467.2019.1598964.
Hoekstra, R. (2019) *Replacing GDP by 2030: Towards a Common Language for the Well-being and Sustainability Community*. Cambridge: Cambridge University Press. doi: DOI: 10.1017/9781108608558.
Huppmann, D. *et al.* (2018) ‘A new scenario resource for integrated 1.5 °C research’, *Nature Climate Change*, 8(12), pp. 1027–1030. doi: 10.1038/s41558-018-0317-4.
IEA (2022) *World Energy Indicators*. Paris: International Energy Agency.
Jerven, M. (2013) ‘Briefing for richer, for poorer: Gdp revisions and africa’s statistical tragedy’, *African Affairs*, 112(446), pp. 138–147. doi: 10.1093/afraf/ads063.
Jiborn, M. *et al.* (2018) ‘Decoupling or delusion? Measuring emissions displacement in foreign trade’, *Global Environmental Change*. Elsevier Ltd, 49(November 2017), pp. 27–34. doi: 10.1016/j.gloenvcha.2017.12.006.
Johnson, P. and Papageorgiou, C. (2020) ‘What Remains of Cross-Country Convergence ?’, *Journal of Economic Literature*, 58, pp. 129–175.

Jorgenson, D. W. (2009) 'A New Architecture for the U.S. National Accounts', 55(1), pp. 199–219.
Kacprzyk, A. and Kuchta, Z. (2020) 'Shining a new light on the environmental Kuznets curve for CO2 emissions', *Energy Economics*. Elsevier B.V., 87, p. 104704. doi: 10.1016/j.eneco.2020.104704.
Kendrick, J. (1970) 'The Historical Development of National-Income Accounts', *History of Political Economy*, 2(2), pp. 284–315. Available at: file:///Users/gs299/Documents/Library.papers3/Articles/1970/Kendrick/Kendrick_1970_The Historical Development of National-Income Accounts.pdf.
Koomey, J. *et al.* (2019) 'Inside the Black Box : Understanding key drivers of global emission scenarios', *Environmental Modelling and Software*. Elsevier, 111(August 2018), pp. 268–281. doi: 10.1016/j.envsoft.2018.08.019.
Kraan, O. *et al.* (2019) 'The influence of the energy transition on the significance of key energy metrics', *Renewable and Sustainable Energy Reviews*, 111(August 2018), pp. 215–223. doi: 10.1016/j.rser.2019.04.032.
Kravis, I. B., Heston, A. and Summers, R. (1982) *World Product and Income: International Comparisons of Real Gross Product*. Baltimore and London: John Hopkins Press.
Kravis, I. B., Heston, A. and Summers, R. (1983) 'The Share of Services in Economic Growth', in *Global Econometrics: Essays in Honor of Lawrence R. Klein*. MIT Press.
Krey, V. *et al.* (2014) 'Annex II: Metrics and Methodology', in Edenhofer, O. et al. (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*.
Kuznets, S. (1948) 'National Income: A New Version', *The Review of Economics and Statistics*. The MIT Press, 30(3), pp. 151–179. doi: 10.2307/1926746.
Landefeld, J. S., Moulton, B. R. and Vojtech, C. M. (2003) 'Chained-dollar indexes: Issues, tips on their use, and upcoming changes', *Survey of Current Business*, 83(11), pp. 8–16. Available at: https://www.bea.gov/scb/pdf/2003/11November/1103 Chain-dollar.pdf.
Leimbach, M. *et al.* (2017) 'Future growth patterns of world regions – A GDP scenario approach', *Global Environmental Change*. Elsevier Ltd, 42, pp. 215–225. doi: 10.1016/j.gloenvcha.2015.02.005.
Liddle, B. and Huntington, H. (2021) 'There's Technology Improvement, but is there Economy-wide Energy Leapfrogging? A Country Panel Analysis', *World Development*, 140, p. 105259. doi: https://doi.org/10.1016/j.worlddev.2020.105259.
Macknick, J. (2011) 'Energy and CO2 emission data uncertainties', *Carbon Management*, 2(2), pp. 189–205. doi: 10.4155/cmt.11.10.
Mazzucato, M. (2018) *The Value of Everything*. Allen Lane.
Missemer, A. and Nadaud, F. (2020) 'Energy as a factor of production: Historical roots in the American institutionalist context', *Energy Economics*, 86, p. 104706. doi: 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', *Review of Social Economy*. Routledge, 79(1), pp. 51-75n. doi: 10.1080/00346764.2019.1703031.
Moulton, B. R. (2004) 'The system of national accounts for the new economy: What should change?', *Review of Income and Wealth*, 50(2 SPEC. ISS.), pp. 261–278. doi: 10.1111/j.0034-6586.2004.00125.x.
Nagaraj, R., Sapre, A. and 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, India. Available at: https://econpapers.repec.org/RePEc:ind:igiwpp:2021-006.
Nakicenovic, N. *et al.* (2003) 'IPCC Sres Revisited: A Response', *Energy & Environment*. SAGE Publications Ltd STM, 14(2–3), pp. 187–214. doi: 10.1260/095830503765184592.

Nakićenović, N. *et al.* (2000) *Special Report on Emissions Scenarios*. Cambridge: Cambridge University Press.
Naqvi, A. and Zwickl, K. (2017) ‘Fifty shades of green: Revisiting decoupling by economic sectors and air pollutants’, *Ecological Economics*. Elsevier B.V., 133, pp. 111–126. doi: 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 Economics*, 29(3), pp. 349–372.
Özcan, B. and Öztürk, I. (2019) *Environmental Kuznets Curve (EKC): A Manual*. Elsevier Science. Available at: https://books.google.com/books?id=TrWXDwAAQBAJ.
Pant, H. M. and Fisher, B. S. (2007) ‘Alternative measures of output in global economic-environmental models: Purchasing power parity or market exchange rates? - Comment’, *Energy Economics*, 29(3), pp. 375–389.
Peters, G. P. and Hertwich, E. G. (2008) ‘CO2 Embodied in International Trade with Implications for Global Climate Policy’, *Environmental Science & Technology*. American Chemical Society, 42(5), pp. 1401–1407. doi: 10.1021/es072023k.
Pigou, A. C. (1920) *The Economic of Welfare*. London: Macmillan.
Pitcher, H. M. (2009) ‘Measuring income and projecting energy use’, *Climatic Change*, 97(1), pp. 49–58. doi: 10.1007/s10584-009-9696-x.
Pollin, R. (1997) *The macroeconomics of saving, finance, and investment*. Ann Arbor: University of Michigan Press.
Pollin, R. (2018) ‘De-Growth vs A Green New Deal’, *New Left Review*, 112, pp. 5–25.
Pollin, R. (2019) ‘Advancing a Viable Global Climate Stabilization Project: Degrowth versus the Green New Deal’, *Review of Radical Political Economics*. SAGE Publications Inc, 51(2), pp. 311–319. doi: 10.1177/0486613419833518.
Reddy, S. (2008) ‘ The World Bank’s New Poverty Estimates: Digging Deeper into a Hole ’, *Challenge*, 51(6), pp. 105–112. doi: 10.2753/0577-5132510607.
Reich, U. P. (2001) *National Accounts and Economic Value*. Basingstoke: Palgrave.
Requena-i-Mora, M. and Brockington, D. (2022) ‘Seeing environmental injustices: the mechanics, devices and assumptions of environmental sustainability indices and indicators’, *Journal of Political Ecology*, 28.
Ruggles, R. (1990) ‘A Note on the Revision of the United Nations System of National Accounts’, *Review of Income and Wealth*, 36(4), pp. 413–419. doi: 10.1111/j.1475-4991.1990.tb00323.x.
Sarkodie, S. A. and Strezov, V. (2019) ‘A review on Environmental Kuznets Curve hypothesis using bibliometric and meta-analysis’, *Science of The Total Environment*, 649, pp. 128–145. doi: https://doi.org/10.1016/j.scitotenv.2018.08.276.
Schor, J. B. and Jorgenson, A. K. (2019) ‘Is it Too Late for Growth?’, *Review of Radical Political Economics*. SAGE Publications Inc, 51(2), pp. 320–329. doi: 10.1177/0486613419831109.
Schurr, S. (1984) ‘Energy Use, Technological Change, and Productive Efficiency: An Economic-Historical Interpretation’, *Annual Review of Energy*, 9, pp. 409–425.
Schurr, S. H. *et al.* (1960) *Energy in the American economy : 1850-1975, an economic study of its history and prospects, by Sam H. Schurr and Bruce C. Netschert, with Vera F. Eliasberg, Joseph Lerner, Hans H. Landsberg.* Baltimore: John Hopkins Press.
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. *et al.* (2021) ‘Plausible energy demand patterns in a growing global economy with climate policy’, *Nature Climate Change*, 11(4), pp. 313–318.
Semieniuk, G. and Weber, I. M. (2019) ‘Inequality in Energy Consumption: Statistical Equilibrium or a Question of Accounting Conventions?’, *European Physics Journal ST*, forthcomin.
Shaikh, A. and Tonak, E. A. (1994) *Measuring the Wealth of Nations*. Cambridge University

Press.
Siesto, V. (1987) ‘Concepts and Methods Involved in the Last Revision of Italy’s GDP’, *LABOUR*, 1(3), pp. 181–200. doi: 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 Development*, 32(8), pp. 1419–1439. doi: 10.1016/j.worlddev.2004.03.004.
Stern, D. I. (2017) ‘How accurate are energy intensity projections?’, *Climatic Change*. Climatic Change, 143(3–4), pp. 537–545. doi: 10.1007/s10584-017-2003-3.
Studenski, P. (1958) ‘The Income of Nations’. New York: New York University Press.
Summers, R. and Heston, A. (1991) ‘The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988’, *The Quarterly Journal of Economics*, 106(2), p. 327. doi: 10.2307/2937941.
Tercioglu, R. B. (2021) ‘Rethinking growth and inequality in the US : What is the role of measurement of GDP?’, *International Review of Applied Economics*. Routledge, 00(February), pp. 1–26. doi: 10.1080/02692171.2021.1892038.
Torras, M. and Boyce, J. K. (1998) ‘Income, inequality, and pollution: A reassessment of the environmental Kuznets curve’, *Ecological Economics*, 25(2), pp. 147–160. doi: 10.1016/S0921-8009(97)00177-8.
Tregenna, F. (2009) ‘Characterising deindustrialisation: An analysis of changes in manufacturing employment and output internationally’, *Cambridge Journal of Economics*, 33(3), pp. 433–466. doi: 10.1093/cje/ben032.
Tukker, A., Pollitt, H. and Henkemans, M. (2020) ‘Consumption-based carbon accounting : sense and sensibility’, *Climate Policy*. Taylor & Francis, In press, pp. 1–13. doi: 10.1080/14693062.2020.1728208.
Vanoli, A. (2005) *A History of National Accounting*. Amsterdam and Oxford: IOS Press.
Victor, P. A. (2008) ‘Managing without Growth’. Routledge.
Wagman, B. and Folbre, N. (1996) ‘Household services and economic growth in the United States, 1870-1930’, *Feminist Economics*, 2(1), pp. 43–66. doi: 10.1080/738552685.
Van Walbeek, C. (2006) ‘Official revisions to south African national accounts data: Magnitudes and implications’, *South African Journal of Economics*, 74(4), pp. 745–765. doi: 10.1111/j.1813-6982.2006.00094.x.
Weber, I. M. *et al.* (2021) ‘What You Exported Matters: Persistence in Productive Capabilities across Two Eras of Globalization’, *UMass Economics Working Paper Series*, 299.
Wiedenhofer, D. *et al.* (2020) ‘A systematic review of the evidence on decoupling of GDP, resource use and GHG emissions, part I: bibliometric and conceptual mapping’, *Environmental Research Letters*, 15, p. 063002. doi: 10.1088/1748-9326/ab8429.
Wiedmann, T. *et al.* (2020) ‘Scientists’ warning on affluence’, *Nature Communications*, 11(1), p. 3107. doi: 10.1038/s41467-020-16941-y.
Wiedmann, T. O. *et al.* (2015) ‘The material footprint of nations’, *Proceedings of the National Academy of Sciences*, 112(20), pp. 6271–6276. doi: 10.1073/pnas.1220362110.
Witt, U. and Gross, C. (2020) ‘The rise of the “service economy” in the second half of the twentieth century and its energetic contingencies’, *Journal of Evolutionary Economics*. Journal of Evolutionary Economics, 30(2), pp. 231–246. doi: 10.1007/s00191-019-00649-4.
Wolff, E. N. (1987) *Growth, Accumulation, and Unproductive Activity: An Analysis of the Postwar U.S. Economy*. New York: Cambridge University Press.

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, 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 dollars, 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 (think of solar PV panels), thereby overweighting the output of periods prior to the reference period and underestimating the change in real GDP. Conversely inflation is overestimated. The U.S. therefore switched to Fisher or chained-dollar quantity indices that instead use information about prices from both periods in 1996. Changing the reference period has no impact on their measured growth rate (for a detailed discussion and examples see Landefeld, Moulton and Vojtech, 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 figure 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, Moulton and Vojtech, 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" (Fixler, Greenaway-Mcgrevy and Grimm, 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, so 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 (Jorgenson, 2009; Coyle, 2014; Corrado *et al.*, 2021) 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 (Wolff, 1987; Shaikh and Tonak, 1994; Basu and Foley, 2013; Assa, 2017; Tercioglu, 2021).[20] 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.[21]

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, Inklaar and Timmer, 2015).[22] And 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, Sapre and Sengupta, 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.

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

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

[22] 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.

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 2 instead of 1 haircuts in country B. 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 was $9.0 trillion.[23]

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 ^. 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.

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

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	^15.4	14.2	13.2	^16.8	PWT 10.0 GDP coefficients jointly insignificant
	Min	74.8	^-70.1	89.8	52.2	^-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	^19.1	19.5	18.2	21.3	Coefficients always jointly insignificant
	Min	-79.7	^277.2	-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	^5.0	-5.7	-45.9	6.7	
	Min	5.3	^14.0	8.5	8.5	17.3	
Fecal coliform	Max	15.7	^5.3	14.5	14.0	^10.4	
	Min	0.6	^34.1	-5.1	-12.4	^43.2	
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	^25.8	^25.5	NA	Coefficients always jointly insignificant
	Min	29.0	NA	^11.4	^11.6	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	

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