

HARMONIZED INACCURACY: INTERNATIONAL ECONOMIC STATISTICS IN THE AGE OF GLOBALIZATION¹

Lukas Linsi
University of Amsterdam
l.a.linsi@uva.nl

Daniel K. Mügge
University of Amsterdam
d.k.muegge@uva.nl

Draft version as background for presentation at Leiden University on 21 September 2017.
Please do not cite or circulate without permission. Comments very welcome!

ABSTRACT. Official international economic statistics are generally considered accurate and meaningful indicators of trends in cross-border flows of trade and capital, and most users of these statistics seem confident that the quality of the underlying data keeps improving over time. This article evaluates these common assumptions at hand of the International Monetary Fund's Balance of Payments Statistics. The assessment paints a less rosy picture: Reported key figures are far less accurate than they are typically imagined to be; frequently, they do not correspond to the theoretical concepts with which users associate them and data quality has not improved over time but actually grown worse. In the light of these findings, we ask why these important deficiencies are so little acknowledged in academic research and economic policy discourses. Building on the literatures of quantification and professions, we highlight the stark division between communities of statisticians and the world of research and policymaking as the key driver of these worrying trends: data producers face professional incentives that are misaligned with the interests of data users; the users in turn are not sanctioned for the omission of due diligence procedures – i.e., the use of low quality data with a deceptive pretense of accuracy. The implications of the resulting concept-measurement gap are very consequential for empirical scholarship about the international economy that builds on much weaker data than researchers realize or admit. It highlights the urgent need for a dialogue between statisticians, researchers and policymakers to clarify the content and limitations inherent in economic indicators.

Keywords: Politics of Statistics; Economic Measurement; Balance of Payments; Trade flows; Capital flows; Globalization.

Word Count: 12,581 (incl. references)

¹ We are grateful to Hanna Dose for her research assistance. This research has been supported by the ERC Starting Grant FICKLEFORMS (grant # 637883) and the NWO Vidi project 016.145.395. Details are available on www.fickleformulas.org.

Introduction

The global economy is invisible to the naked eye, and yet we discuss, research and govern it day in, day out. To ‘see’ the economy, we rely on numbers: figures about trade, inflation, economic growth, foreign direct investment (FDI), and so on. What we know about the global economy in the aggregate, we know from spreadsheets.² And citizens, policymakers and of course academics rely on such economic knowledge to such a degree that material defects of such data deserve serious attention.

Inevitably, there are limits to the accuracy of such economic knowledge. Data gathering by tax or customs authorities, or through business surveys, is a complex enterprise with many opportunities for errors to seep into the statistics. In principle, the existence of such errors is both widely recognized and considered manageable, based on the assumption that quality defects are minor, and that advances in data gathering shrink the inaccuracies over time. In other words, international economic data is not too bad to begin with, and it is getting better.³

We argue that both assumptions are unwarranted. A series of analyses of error margins in international economic statistics, which we present below, and insights gained from interviews with high-level statistical officers reveal that the measurement uncertainty in such data is much larger than we initially expected. In addition, there is a gap between the concepts we want to capture with international economic data, for example trade flows or FDI, on the one hand, and the things actually measured by the numbers in official databases on the other. What we call the concept-measurement gap (CMG) is big – and it is growing rather than decreasing. To our mind, the stark illustrations of its size that we offer below call into question the appropriateness of international economic data for many analytical and political purposes for which they are frequently used.

That said, we not only want to demonstrate *that* the gap is increasing, but also understand *why*. Our answer to this question comes in two parts: what causes the CMG to widen, and why do we not close it with better data? Our analysis suggests that the worsening of measurement problems in international economic statistics are primarily driven by the inertia of statistical concepts in the face of fundamental structural economic changes, such as the growing cross-border fragmentation of economic production chains, the rise of the services sector, and the deepening knowledge-intensity of production. Due to a failure to meaningfully adapt statistical measurements to such changes, many of the categories structuring the collection of international economic data today map poorly onto present-day economic realities. Furthermore, because globalization is a dynamic process, the magnitude, incidence and direction of measurement problems change over time, such that there is no easy way to control for biases in the data, and the principal problem often is not that data gathering is deficient, but that it is fundamentally unclear

² Karabell 2014; Hirschman and Berman 2014.

³ Kerner (2014) is an exception in his meticulous and critical discussion of FDI data.

how complex transactions could be translated into statistical categories such as FDI or portfolio investment without stretching them beyond recognition.⁴

If these problems increase and are not easily fixed, why is that not reflected in more determined efforts to find better measurements, or at least a growing caution with which for example the IMF or the OECD present their data? First of all, both producers and users of data often have incentives not to rock the boat and stick with established measures, not least for fear that fundamental criticisms would devalue large swathes of (historical) data and analyses.

Second, over the past decades international statistics have become increasingly governed by conceptual and theoretical considerations. The ambition of national accounting has moved beyond the systematic collection of economic data – proverbial bean-counting – towards the integration of the latter within an intellectual system of its own. In line with this ethos, most energy is invested in adjusting the different elements of the BOP statistics to each other, such that a consistent overall picture emerges, rather than making sure that the individual elements correspond closely to economic concepts as we commonly understand them as well as to 21st century economic realities.

In short: structural economic changes have gradually undermined the validity of key international economic indicators. The logic of national accounting, which systematically prioritizes *comparability*, *consistency*, *coherence* and *continuity* over the *accuracy* of individual indicators, has blocked the road to the introduction of new, globalization-compatible indicators, while also deflecting demands to advertise measurement problems more prominently. Data users, in turn, seem unaware or willingly ignoring of measurement problems and put much greater confidence in the accuracy of international economic statistics for purposes of causal inference than it is warranted. In combination, we argue, these factors have fomented a problematic fetish of statistical precision where none exists.

Empirically, our investigation centers on the evolution of the IMF's Balance of Payments Manuals (BPM) and the key economic indicators defined therein (in particular trade, FDI and portfolio capital flow statistics)⁵. To develop our argument, the paper takes three steps: first, it demonstrates empirically that measurements of central economic quantities in the global economy are worryingly inaccurate. We do so using a range of measures that highlight internal inconsistencies in BOP data and their persistence over time. This descriptive contribution is important for those of us, both in- and outside academia, who use statistical data as analytical raw material. Second, we show how economic globalization has aggravated the concept-measurement gap to the point where official statistics are gross misrepresentations of underlying

⁴ Our analysis resonates with that of Dick Bryan, who in 2001 already noted the contradiction between globalization on the one hand and an increasing importance attached to national economic performance on the other. Bryan 2001.

⁵ The focus on the BPM is justified in that it constitutes the global benchmark setting out the rules and principles according to which most national governments collect statistics on cross-border flows of trade, foreign direct and portfolio investments; and, at the same time, the statistics published on that basis in the IMF's Balance of Payments Yearbooks constitute the most authoritative corpus of data that economists preferably rely upon for their analyses.

economic relationships and dynamics. Third, we outline key dynamics that can explain why these data problems have been insufficiently addressed by the epistemic communities in charge.

As a whole, our analysis indicates that much of our empirical knowledge about the international economy rests on shakier (statistical) foundations than scholarly analyses typically acknowledge. In conclusion, we suggest that much could be gained from a more honest engagement with the limitations that are inherent in the data we use.

Economic statistics as social artefacts

The observation that there is more to quantitative data than meets the eye has a long history. Adam Smith disparaged the *political arithmetick* that was fashionable in the 18th century: data quality was too poor to allow solid conclusions, and the putative hardness of numbers belied behind-the-scenes politicking about how they were put together.⁶ In the 1940s, Simon Kuznets warned against reading too much into the national income indicator that he himself had developed,⁷ and reputed Princeton economist Oskar Morgenstern meticulously outlined the many shortcomings and limitations of popular macroeconomic measurements in his 1950 monograph *On the Accuracy of Economic Observations*.⁸ In vain. Macroeconomic indicators have gained an ever stronger foothold and become indispensable tools in economic policy and research.

In recent years, the ever tighter embrace of indicators has spawned a flourishing literature critically assessing their role in social life.⁹ The insufficiencies of GDP as a gauge for national welfare has attracted most attention.¹⁰ Our analysis takes inspiration from that work. But GDP critics have primarily highlighted gaps between casual, commonsense understandings of the measure and the more narrow, economic dynamics GDP figures actually capture (i.e. the issues in box I in Figure 1).¹¹ These criticisms have merit, but they frequently stem from careless *interpretation* of data, not problems in the data itself. Statisticians are often the first ones to admit that GDP is a production measure and may say little about welfare in a society, let alone wellbeing.¹² Instead, we shift attention to box II of Figure 1: the gaps between the official definition of a concept and what actually ends up in the reported data – the correspondence between what it says on the box and what is inside. We refer to these tensions as the *concept-measurement gap* (CMG).

⁶ Dimand 1995.

⁷ Coyle 2014; Fioramonti 2014. [page numbers tbc]

⁸ Morgenstern 1963 [1950].

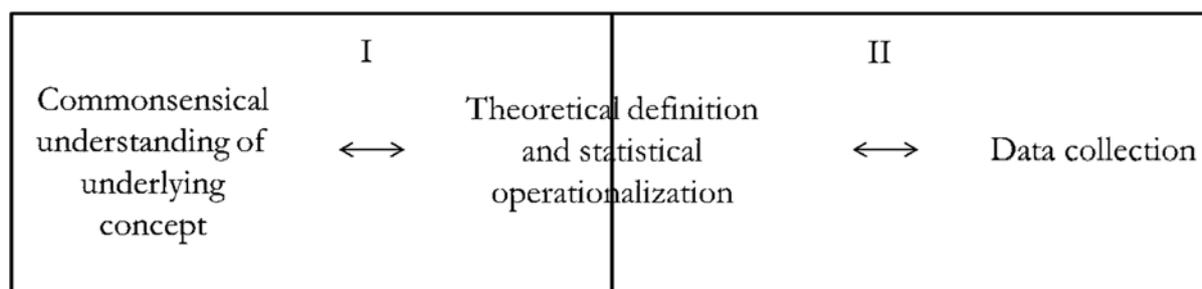
⁹ Schmelzer 2016; Lepenies 2013; Philipsen 2015; Kelley and Simmons 2015; Broome and Quirk 2015; Snyder and Cooley 2015; Fioramonti 2014.

¹⁰ Stiglitz, Sen, and Fitoussi 2010; Fioramonti 2014; Coyle 2014; Lepenies 2013; Philipsen 2015.

¹¹ For example: the exclusion of unpaid labour from GDP, its ignorance of environmental destruction, or its inability to capture people's 'happiness'.

¹² E.g. Lequiller and Blades 2014.

Figure 1. The production chain of economic indicators



Although our empirical analysis below focuses on international economic statistics, the CMG is a general feature of social statistics. Consider the best-known macroeconomic statistic, GDP. Annual GDP measures final output of goods and services within a country’s borders during one year.¹³ In principle, it is limited to goods and services that are produced for money. Yet for large chunks of our economies, we cannot observe any market prices: government services – education, administration, defense, sometimes healthcare – are often heavily subsidized or provided free of direct charge to consumers and financed through taxes or collective contributions instead. So how should the American National Park Service or an elementary school teacher show up in GDP statistics? By convention, the output of the public sector is equated with its cost.¹⁴ But needless to say, that may vastly overstate or understate actual output. In short, even when we closely heed the actual definition of an indicator such as GDP, the number we pull from spreadsheets can be seriously misleading.

While the CMG can be viewed statically, as in the preceding example, its most intriguing aspects are dynamic. The constant evolution of economic structures and processes poses an important challenge to economic indicators, which are supposed to keep track of the same phenomenon in an ever changing and unforeseeable environment.¹⁵ Our investigation below will highlight the specific stress globalization puts on BOP statistics. But the general question for social scientists is how public authorities address these challenges. The obvious answer would be for authorities simply to adapt indicators so that measurements still correspond to the respective concepts. Or, should that be impossible, to explain in data reporting how structural economic changes have dented data quality, and substitute plausible data ranges for point estimates.

Surprisingly, such approaches are uncommon for macroeconomic data, even while we do find them in other publicly reported data. For example, many governments have adapted ethnicity categories for citizens over the decades.¹⁶ And the efforts of international organizations such as the World Bank and United

¹³ For a more elaborate definition, and overview of what exactly GDP captures, see Ibid.

¹⁴ Ibid.

¹⁵ Beckert 2016.

¹⁶ Marquardt and Herrera 2015; Petersen 1987.

Nations to measure ‘development’ in quantitative terms have evolved substantially over time¹⁷. Such categorizations and indices, if one deems them appropriate at all, leave much to be desired. But they demonstrate how measurement systems *can* adapt in the face of social and political change. Reporting of data ranges, in turn, is common in particular in forecasting, for example for climate change.¹⁸ The use of data ranges instead of point estimates might seem more intuitive for future projections than current or past developments.¹⁹ But as we will show, uncertainties about past economic developments are substantial enough to justify similar caution there, as well.

In this light, the relative reluctance to confront measurement problems (either by introducing new concepts or by making inherent measurement uncertainties more explicit) in the case of international economic statistics is not only problematic for practical reasons, but also somewhat puzzling theoretically. The following paragraphs will elaborate the measurement problems affecting BOP statistics in some more detail. Subsequently, we will return to the question why statistical authorities and data users have resisted pressures to overhaul international economic measurements or to engage more explicitly with the uncertainty inherent in the measurements they produce and use.

Globalization and the quality of international economic data

Many key economic indicators derive from balance of payments (BOP) statistics, which chronicle countries’ economic cross-border exchanges and complement the Systems of National Accounts (SNA). Obsessed with trade balances in the mercantilist era, many countries had started to monitor exports and imports in the 16th and 17th centuries.²⁰ These efforts intensified in the subsequent centuries as governments expanded their economic records. Simon Kuznets’ National Income estimates for the United States in the 1930s are generally portrayed as a watershed in this regard.²¹ Even though conceptually they stood in a long tradition of national income estimates,²² the American effort was more systematic and comprehensive, and above all politically more consequential than what others had tried before.²³

The key innovation for our purposes arrives in the 1940s: an ambition to build an integrated statistical edifice that offers a coherent, comprehensive and consistent picture of the economy. John Maynard Keynes had formalized macroeconomics and systematized the idea of a “national economy” with, at least conceptually, a boundary with the outside world and theoretically unambiguous relationships between

¹⁷ Finnemore 1996.

¹⁸ Randall and Wood 2007; Maslin and Austin 2012.

¹⁹ Two examples for databases that include confidence intervals for the presentation of data about the past are the Comparative Manifestoes Project, and Frederick Solt’s Standardized World Income Inequality Database.

²⁰ Lipsey 2006; Morgenstern 1963; McCormick 2009; Studenski 1958.

²¹ Fioramonti 2014; Coyle 2014.

²² Especially after the first World War, interest in these kinds of statistics grows significantly. Studenski 1958.

²³ Lepenies 2013.

major economic variables²⁴, such as investment, imports and exports, government and consumption expenditures, and so on.²⁵ Macroeconomic equations that related these quantities to each other would allow economists to impute missing data.²⁶

The BOP framework, an integral component of the SNA system (the ‘Rest of the World’ account from the perspective of domestic economic activity), is therefore not just a collection of relevant tables that are of interest to policymakers in their own right. Instead, it contains its own internal logic and rules, such as the principles of double-entry book keeping, which systematically relate the equations in the system.

The earliest efforts to collect balance of payments statistics at the international level were conducted under the auspices of the League of Nations in the 1920s and 1930s. In the aftermath of the Second World War, the responsibility shifted to the International Monetary Fund (IMF).²⁷ As the guardian of international financial and economic stability in the Bretton Woods era, the IMF was tasked to identify unsustainable imbalances in global financial flows.²⁸ As a result, it strove to establish international conventions on how member countries should collect data on cross-border payments. The first *Balance of Payments Manual* (BPM1) issued in 1948²⁹ offered standardized template tables to be filled out by member countries on an annual basis. A slightly expanded version, with more detail about what to include and exclude, followed three years later.³⁰

Since then, the IMF’s BOP Statistics enterprise has grown substantially in size and ambition. The Polak model, developed by the director of the IMF’s research department from 1958 to 1979, Jacques Polak, related key domestic macroeconomic variables such as GNP growth and domestic credit of the banking system, to cross-border economic variables such as foreign exchange reserves and trade.³¹ Later theoretical refinements formalized the relationship between the balance of payments, changes in the domestic money supply and developments in the real economy.³² Rather than isolated macroeconomic quantities to be

²⁴ Mitchell 2002; Blaug 1996, vol. 5th.

²⁵ Keynes emphasized in particular the importance of monetary variables, such as interest rates, inflation and exchange rates, as determinants of national output, rather than as their consequence. For macroeconomic management, it therefore became imperative to monitor those well, including their cross-border dimension. It is also interesting to note that the two originators of the 1940 national income estimate of the UK, James Meade and Richard Stone, were both early Keynesian disciples (Perlman 1987.); the latter of the two later lead-authored the first international standard for national income accounting, the 1953 version of the SNA. Lепенies 2013.

²⁶ Cf. Vanoli 2005.

²⁷ Ref to IMF Archive Document [REF TBC]

²⁸ See International Monetary Fund 1948, 1.

²⁹ International Monetary Fund 1948.

³⁰ International Monetary Fund 1950.

³¹ Polak 1997. On the centrality of the Polak model to IMF policy, see Woods 2006.

³² International Monetary Fund 1977b.

observed individually, the constituent elements of the balance of payments came to be seen as building blocks of a larger integrated whole.³³

As a result of these developments, the most recent, sixth edition of the BPM published in 2009 contrasts starkly with the original editions in both substance and style.³⁴ It does not merely provide practical templates, but is organized as a highly didactical volume³⁵ that strongly emphasizes the theoretical underpinnings and rules of the BOP system. Accordingly, while the concise BPM1 covered less than fifty pages, the most recent version has grown into a carefully crafted authoritative document of close to four hundred pages, accompanied by a separate six hundred-page *Compilation Guide*.³⁶

The complexity of the underlying data collection process

In stark contrast to the tidy appearance of international economic statistics in spreadsheets, the processes through which they are created are highly complex and messy. First, the compilation of balance of payments statistics is a massive *logistical* challenge because it requires information from a large and heterogeneous pool of data sources. In each one of the more than two hundred participating countries, BOP data collection is not done by the statistical office alone, but involves a large part of the government apparatus, including the work of customs authorities (for information on merchandise and some services trade), port authorities (for transportation services), immigration authorities (for the number of short-term and long-term migrants), tourism authorities (for expenditures of foreign visitors and residents travelling abroad), insurance and banking regulators (for cross-border lending and securities flows), external affairs departments (for operations of embassies), tax authorities (for residents' external assets and liabilities and income from abroad), the compiler of the international transactions reporting system (for financial flows),³⁷ the central bank (for banking sector activities and reserve assets), and many more.³⁸

³³ A good example of this new way of thinking is that a measured net surplus or deficit of cross-border flows must immediately imply a depletion or increase in net foreign reserves, such that the latter could in fact be imputed from knowledge of cross-border financial flows.

³⁴ International Monetary Fund 2009.

³⁵ As the authors of BPM5 proudly declare, “[t]his edition ... not only defines and describes the content of the categories employed but also attempts to explain their rationale. In particular, the introductory sections have been elaborated to describe more fully the basic concepts and accounting principles that are common to the balance of payments and the other social accounts. With these amendments, the Manual has become as much an introduction to the principles of balance of payments accounting as a guide to reporting.” International Monetary Fund 1993, 1.

³⁶ International Monetary Fund 2014.

³⁷ An International transactions reporting system (ITRS) measures individual balance of payments cash transactions passing through the domestic banks and foreign bank accounts of enterprises, and noncash transactions and stock positions. Statistics are compiled from forms submitted by domestic banks to the compilers and from forms submitted by enterprises to the compiler. See OECD 2017a.

³⁸ This list is drawn from International Monetary Fund 2014.

Second, the aggregation of these heterogeneous data series involve immense *technical* challenges because the actors involved, including national statistical offices, do not necessarily follow BOP-conform accounting practices. Often, different legitimate ways of recording and valuing transaction exist, which can make data compilation into supposedly ‘comparable’ units a nigh impossible task. The four most obvious sources for accounting-technical disharmonies are differing practices for the *timing of recording, currency conversions, valuation adjustments* and *classification* decisions³⁹.

Normally no one way of handling such accounting-technical issues is obviously correct and practices thus naturally differ across countries and across agencies. Very much aware of such complexities, the IMF and other international organizations (especially the United Nations and OECD) have worked tirelessly over the past few decades in order to harmonize national practices in such regards to the extent that this is feasible. Nonetheless, despite all these efforts and all the goodwill statisticians from around the world have shown towards such initiatives, national statistical offices (and the many other agencies providing necessary data for the compilation of national accounts⁴⁰) ultimately remain free to adopt the methods that are most convenient for their own purposes and they may thus at times, often for good reasons, differ from the ‘best practice’ advocated in the BPM.

In short, the compilation of BoP statistics is far more complicated and error-prone than most users realize. Despite all its sophistication, some international transactions may simply be missed during data collection; the sampling of households and enterprises from which important estimates are derived may not be perfectly random and thus induce biases throughout the accounts; civil servants and statisticians from different countries (as well as those from various agencies within a country) may openly disagree about what the most appropriate way to value a certain asset is; or they may merely unwittingly follow different practices in such regards.⁴¹

In light of such uncertainties, the idea that one number could accurately capture a country’s trade and investment activities becomes untenable⁴², calling for a much more serious consideration of the error margins that surround the estimates presented in spreadsheets. But our point is not merely to emphasize

³⁹ Ibid. Some examples: should an import be recorded on the day the purchasing party made the order, on the day she settles the bill, on the day the good crosses the border or on the day of delivery? Should compilers use the exchange rate prevalent at a specific point of the transaction process or a fictive annualized average to convert the value of transactions recorded in local currency to USD? What is the value today of a foreign investment in an unlisted domestic company that was made years ago? Is a resident’s purchase of an e-book from an online bookseller domiciled abroad to be recorded as an import of a good or a service? (The latter is a particularly contentious issue because the BOP has followed conventional notions of residency in chronicling trade in services whereas online service provision frequently fail to map onto these categories. See Ward 2004.)

⁴⁰ Cf. Ibid.; Vanoli 2005.

⁴¹ For an illustration of differing practices in the case of FDI statistics, see International Monetary Fund and Organisation for Economic Co-operation and Development 2003. For an overview of accounting-technical differences among EU member countries in the compilation of trade data, see Eurostat 2004.

⁴² In the words of Enrico Giovannini, former President of the Italian Statistical Office, the ‘right’ numbers “are impossible to get”. UNSD and Eurostat 2011, 3.

that there are holes in the nets with which BOP data is collected, but also that these holes have been growing bigger over time and that, as a result, the fishes that are ultimately pulled out correspond ever less to the preferred species on the minds of their consumers.

The challenge of globalization

Discourses about economic statistics frequently imply that data is becoming more and more accurate. And there are indeed good reasons to expect this to be the case. Impressive technological advances have supercharged our abilities to collect, store and process ever greater amounts of data electronically⁴³; as part of broader ‘Big Data’ initiatives, many governments have moved data collection up their priority (and funding) list.⁴⁴ Around the world, initiatives to harmonize and synchronize data collection procedures have multiplied through the sustained efforts of international statistical organizations.⁴⁵ However, we suggest that, at least in the case of BOP statistics, structural economic changes stemming from globalization have counteracted such real progress.

The national accounting framework builds on the economic statism prevalent in the post-War period, seeing the world economy primarily as a system of more or less autonomous national units exchanging (relatively modest amounts of) goods and capital.⁴⁶ This perspective is deeply ingrained in the BOP framework, originally designed for a world in which companies have clear national identities, production chains are clustered within national units, and most cross-national transactions involve simple bilateral relationships. The internationalization of economic processes over the past decades has invalidated many of these assumptions, raising tensions between how goods, services and capital are traded cross-nationally and how BOP statistics account for them⁴⁷.

Five forces in particular have posed severe challenges for a smooth operation of the BOP framework: Regimes of *economic production* have been gradually de-nationalized as companies – fueled by technological advances and political support for economic liberalization - built integrated value chains spanning national borders.⁴⁸ As a result, national ownership relations and the localization of value-adding inputs in traded goods have become increasingly opaque. For example, ‘processing abroad’ - companies sending a semi-finished product to another country for further processing – or ‘merchanting’ activities – in which goods are moved between two economies by a trader who is resident in a third country – account for ever larger

⁴³ Mayer-Schonberger and Cukier 2013; Data Revolution Group 2014.

⁴⁴ Dunleavy and Margetts 2010.

⁴⁵ Clegg 2010; Ward 2004.

⁴⁶ Jessop 2002; Frieden 2006.

⁴⁷ Cf. Research interview 20170425.

⁴⁸ Dicken et al. 2001; Gereffi, Humphrey, and Sturgeon 2005.

shares of global trading activities, but it is not obvious how to allocate such flows in a national economy-framework.⁴⁹ At the same time, the growing importance of the service sector and knowledge-intensive production have meant that ever larger shares of economic value-creation are captured by *intangible assets* such as trademarks, patents and other forms of intellectual property.⁵⁰ These elusive assets complicate the task of national accountants because of the impossibility to determine their precise economic value. When knowledge, or legal entitlement to its use (as in patents), are considered a production factor, how should we allocate such production geographically?

The *de-nationalization of firm structures* has exacerbated these problems, in particular the growing popularity of the usage of special purpose entities in low-tax offshore financial centers (OFCs).⁵¹ More and more companies (and individuals) use chain ownership structures as part of transfer-pricing strategies and other tax-reducing purposes.⁵² National accountants therefore struggle to identify the ultimate owners and beneficiaries of cross-border transactions. And corporate balance sheets, a key source for BOP statistics, become less reliable as information sources when companies artificially move their declared income to low-tax OFCs.⁵³ Capital market deregulation combined with *financial innovation* has unleashed cross-border capital flows, often bundled into complex derivative packages that obscure effective ownership.

To make matters worse, while all these developments were simultaneously contributing to a dazzling complication of cross-border economic activities in the 1990s and 2000s, many governments were at the same time *rescinding* some of the *administrative apparatuses* they had traditionally relied upon for BOP data collection. Financial liberalization implied the removal of rigid exchange control systems⁵⁴; the establishment of custom unions meant that many goods could now cross borders without inspection⁵⁵; and widespread budget cuts, austerity measures and broader efforts to push back state institutions unavoidably also affected governments' capacity to collect data.⁵⁶

Taken together, these factors have create substantial problems for BOP statistics⁵⁷. As we illustrate in the following sections, it has substantially increased measurement errors and, at the same time, further widened the gap between the theoretical concepts behind key indicators and their statistical operationalization.

⁴⁹ UNECE, Eurostat, and OECD 2011. 'Processing abroad' leads to distortion of trade figures as it can lead to double-counting of products as they move along the production chains; in contrast, 'merchanting' activity is difficult to identify and is therefore excluded from most trade statistics. Currently, only the EU, Switzerland and Turkey are collecting merchanting statistics; the data suggests that merchanting activity in these countries increased from 28.5 billion euro in 2004 to 65.2 billion in 2008. Ibid.

⁵⁰ Picciotto 2012.

⁵¹ Palan 2002; Shaxon 2011; Picciotto 2012.

⁵² Seabrooke and Wigan 2017.

⁵³ Lipsey 2007; Lipsey 2010.

⁵⁴ Cf. International Monetary Fund 1992, 7.

⁵⁵ Cf. UNECE, Eurostat, and OECD 2011.

⁵⁶ Cf. International Monetary Fund 1987, 1. For more recent examples of the impact of budget cuts on the collection of statistics, see CBS 2017 or Moulton 2017.

⁵⁷ See in particular UNECE, Eurostat, and OECD 2011.

Measurement problems

To assess the evolution of measurement errors in BOP statistics, we focus on two axes of ‘balance’ in BOP statistics. First, following the rules of double-entry book keeping, the sum of credits that a country reports in its BOP should in principle equal the sum of debits. Second, because the world economy as a whole is a closed system in BOP terms, every outflow from one economy should theoretically be recorded as an inflow in another and, as a corollary, flows between partner countries as well as the total global sum of inflows and outflows should in principle match. Discrepancies among the figures that should mirror each other highlight measurement problems.

Aggregate discrepancies will understate actual measurement problems, because excess credits and debits in different sub-components offset each other. Indeed, large discrepancies irk national accountants, because they suggest poor data quality. In consequence, they have repeatedly instigated revisions and reform proposals to eliminate or at least reduce them⁵⁸; a goal that at times even seems to take precedence over the accuracy of measurements *per se*. For instance, the report of a IMF Working Group explicitly states that “the results given in this report do not purport the ‘true’ level of income accruing on cross-border assets, but are limited to deriving adjustments that should be made to eliminate discrepancies in the statistics as they are now published.”⁵⁹

Against this background, it seems fair to assume that discrepancies between what should be equalities in the BOP statistics are highly conservative estimates of data problems and thus useful benchmarks against which to assess the hypothesis that BOP data have been ‘getting better’ over time.

Discrepancies Within Countries: Net Errors and Omissions

Figure 2 first plots the evolution of the total absolute values of net errors and omissions (E&O)⁶⁰ of all countries reporting that data (bars), as well as the median⁶¹ country-specific ratio of net E&O as a share of total merchandise imports, separately for OECD and non-OECD countries (lines). It is important to

⁵⁸ For example, in the late 1980s and early 1990s, the IMF created two Working Parties to address the challenges from economic globalization for BOP statistics. The resulting reports (International Monetary Fund 1987; International Monetary Fund 1992) pay remarkably little attention to questions surrounding the conceptual ‘accuracy’ of measurements and instead emphasize the reductions in errors and omissions and discrepancies in mirror statistics.

⁵⁹ International Monetary Fund 1987. See also International Monetary Fund 1992.

⁶⁰ The concept of net E&O was introduced in International Monetary Fund 1950.

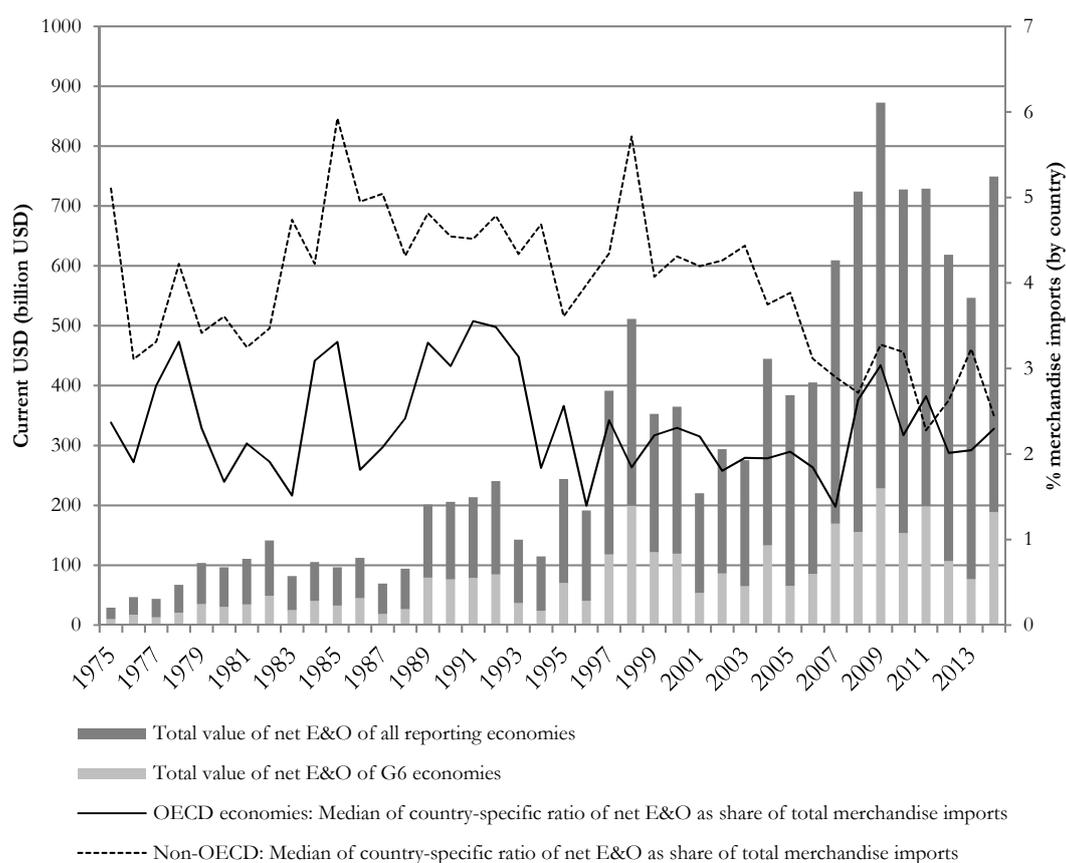
⁶¹ We choose the median rather than the mean value to reduce the influence of outliers (the observed values of net E&O as share of total merchandise imports are extreme for some countries under exceptional circumstances - for example a staggering 314% for Kuwait in 1991, or 182% for Iceland in 2009)

underline that the shown absolute values indicate the total *reported* values, not the global totals (which is an unobserved quantity because country coverage is not universal). This implies that, due to the gradually growing number of reporting economies, parts of the observed increases in total reported E&O over time are simply due to the addition of new reporters; taking these dynamics into account, we separately also track the evolution of the total value of reported net E&Os of six G7 economies (excluding Japan⁶²), which have been reporting this data consistently throughout the period of observation (light grey bars).

The graph suggests some interesting patterns: Hinting at a series of elusive victories on behalf of national accountants, sharp increases in net E&Os (such as in the early 1990s, late 1990s and late 2000s) tend to be met by substantial corrections, only to be followed by a resurgence of even larger E&Os in subsequent years (this is the case for aggregate values of all reporting economies as well as for the consistent G6-sample separately). In terms of volume, despite statisticians' sustained efforts to reduce net E&Os, they have continued to grow at similar rates as global merchandise trade has increased during the same period, reaching hundreds of billions USD in most recent years. On the positive side, the country-specific median ratios suggest that the problem has indeed become somewhat smaller for non-OECD economies since the early 2000s, which may be a sign of a partial success of the many international harmonization and statistical capacity building efforts. However, no similar improvements are evident among OECD economies, where ratios stubbornly resisted efforts to get them under control. In short, despite statisticians' efforts to reduce them, net errors and omissions persist and there are no reasons to believe that they have become significantly smaller over time.

⁶² Japanese data is only available from the 1990s.

Figure 2. The evolution of net errors and omissions over time, 1975-2014



DATA SOURCE: IMF BOPS (E&O) and IMF DOTS (Merchandise trade), made available through World Bank Databank. NOTE: The number of reporting economies increases over time, especially in early period (1975: 54 reporting countries; 1980: 118; 1985: 129; 1990: 132; 1995: 144; 2000: 146; 2005: 173; 2010: 180; 2014: 163).

Discrepancies Across Countries I: Global aggregate mirror statistics

Next we analyze the discrepancies in aggregate total reported in- and outflows of several key BOP categories: Merchandise trade, services trade, foreign direct investments (FDI) and portfolio investments (PFI). Although the combined value of trade and capital out- and inflows reported by all economies should in principle be identical, this is not the case. Figure 3 shows the differences between the total global assets and liabilities reported in the IMF BOP Yearbooks for the various categories in absolute levels (bars), as well as a percentage of the total reported liabilities of the relevant category.⁶³ As before, these differences only capture the ‘extra’ discrepancies on top of those offsetting each other, and they have been a subject of

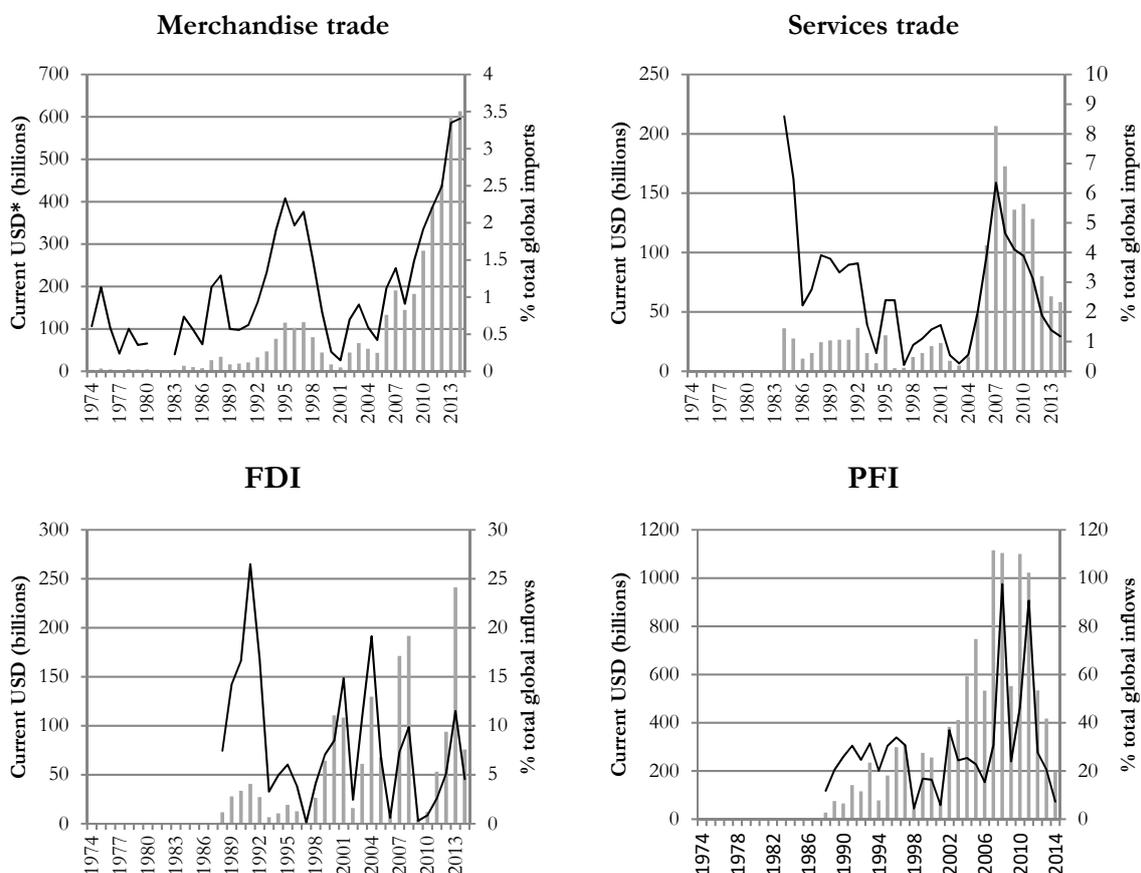
⁶³ The IMF started collecting global aggregates on an experimental basis from 1967 onwards. See Doc DM/67/76 in IMF Archives, 3.

heated debate among statisticians, who have devised strategies to actively reduce them.⁶⁴ Yet, against these efforts, we in fact observe fairly dramatic *increases* in discrepancies in the early 2000s in all sub-categories, both in absolute terms as well as relative to the value of total reported inflows (with the partial exception of FDI where no clear pattern emerges in relative terms).

A comparison of discrepancies among various types of flows is equally insightful: As one would expect, although being substantial, they are smaller in relative terms for trade than capital flows (which are harder to measure), and with the exception of the most recent years, somewhat smaller for merchandise than services trade statistics. The size of relative discrepancies for FDI and PFI flows are generally large and truly remarkable in some years. Bearing in mind that these discrepancies in all likelihood only reflect a small part of the total underlying cross-national measurement differences, the observation that they regularly exceed ten percent in the case of FDI and can even reach close to one hundred percent (!) in the case of PFI poses indeed very serious questions about the extent to which we should be relying on this kind of data for the purpose of economic analysis.

⁶⁴ See International Monetary Fund 1987; International Monetary Fund 1992.

Figure 3. Evolution of statistical discrepancies in reported total global inflows and outflows of trade and capital over time for four BOP sub-items



SOURCE: Print version of IMF Balance of Payment Statistics Yearbooks 1981, 1990-1992, 1994-2007, 2011-2015.

NOTE: *Figures before 1984 are in billion SDRs, all others in billion USD. The reported numbers are those corresponding to the figures of the YBs in our sample published closest to the year of observation.

Discrepancies Across Countries II: Mirror statistics in bilateral merchandise trade

Lastly, to get a better sense of the size of these cross-national measurement discrepancies, we also take a closer look at the discrepancies in bilateral trade statistics; that is, we evaluate the difference between the exports from A to B reported by B with A's imports from B reported by A, etc. While such bilateral mirror statistics can still hide offsetting discrepancies at the level of industrial categories (which NSOs add up to get at 'total' trade figures), such effects should generally be much smaller in this setting than they are at the global aggregate level.

For our analysis we rely on the OECD's STAN bilateral merchandise trade statistics database (rev. 3) and examine discrepancies in mirror statistics in reported trade flows among the United States, United Kingdom, Germany, China and India from 1990 to 2011. Merchandise trade data is generally considered to be one of the most readily observable (and thus easiest to measure) of all BOP components. In the words of the OECD, it involves "package[s] crossing the customs frontier with accompanying documentation showing an internationally recognized commodity code; a description of the contents; information on quantity, origin, and destination; an invoice; and an administrative system based on customs duty collection which facilitates data compilation"⁶⁵. And as the preceding analysis of global aggregate mirror statistics has shown, discrepancies in merchandise trade data – moving between 0.5 and 2.5 per cent of the total value of reported imports – do indeed tend to be notably smaller than they are for capital flow statistics.

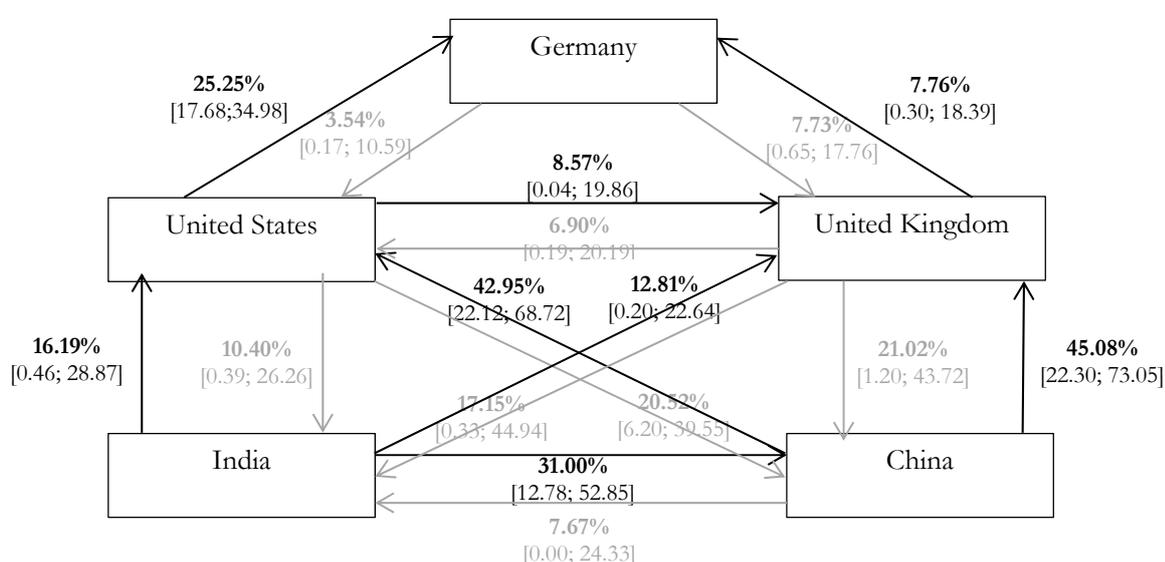
However, our analysis of discrepancies at the *bilateral* level, which can uncover some of the measurement problems offsetting each other at the global aggregate level suggests that they are in fact surprisingly large. Drawing from the authoritative OECD STAN bilateral merchandise trade statistics database, Figure 4 illustrates the size of the discrepancy in bilateral mirror discrepancy among a diverse group of five large economies. The values shown indicate the size of the discrepancy as a percentage of the respective value reported by the importing economy. Bold numbers represent the average discrepancy over the entire 22-year period (spanning from 1990 to 2011); numbers in parentheses mark the minimum and maximum values reached during the observed period. The figure shows that bilateral discrepancies are a pervasive feature of all analyzed trading relationships and that their size is remarkable in many cases. Reported mirror statistics between the two closely integrated EU-members Germany and the United Kingdom, for example, fluctuate notably. While they are nearly identical in some years, they diverge as much as 18 percent of the value of reported imports in others and the 1990-2011 period-average value of discrepancies hovers around 7.7 percent in both directions. Not surprisingly, discrepancies can be much larger for other, less closely integrated dyads. For example, period-average discrepancies for merchandise trade flows from China to the United Kingdom and China to the United States lie above 40 percent of the total value of imports. This is a serious issue. If we are interested in the US trade deficit with China and Germany, for instance, we arrive at very different estimates depending on which data source we use: The US trade deficit with China [Germany] in 2015 amounted to 388 [78] billion USD according the US sources, but 'only' 260 [58] billion USD according to Chinese [German] figures. And *a priori* there is no reason to believe that one number is more accurate than the other.

While the BPM recommends valuing all trade transaction at 'free on board' (f.o.b.) value (that is, the value of goods excluding international transportation costs), the OECD database reports only export at f.o.b. while import data also includes cost of insurance and freight (c.i.f.). The implication is that exports from A to B should be lower than imports to B from A, because the import figure contains the net value of exports *plus* the cost of freight and insurance. This factor can explain a small part of the observed discrepancies,

⁶⁵ OECD 2001, 4.

but it cannot account for their size or degree of fluctuation. Contrary to this logic, reported (f.o.b.) export values regularly exceed corresponding (c.i.f.) import figures. Furthermore, discrepancies can also change wildly over time, belying the idea that trade statistics gaps would arise from a rather stable gap between export and import figures for the same set of transaction. The main reason for these differences are real measurement problems rather than deliberate mis-reporting by any particular economy or cost for freight and insurance that importers pay but exporters don't receive as revenue.⁶⁶

Figure 4. Discrepancies in OECD STAN bilateral merchandise trade mirror statistics as a percentage of totals reported by importing economy, average values 1990-2011



SOURCE: Own calculations based on OECD STAN 2012 rev. 3 bilateral trade statistics database. NOTE: Bold values indicate average values of all relevant dyad-year observations between 1990 and 2011. Values in parentheses indicate the minimum and maximum values observed during that period.

Furthermore, as in the case of net E&Os, there are indications that harmonization efforts have had some effect to reduce measurement discrepancies between OECD and non-OECD economies, but that these relative progresses have been undermined by simultaneous increases in discrepancies among OECD economies: While the decade-average discrepancy of trade flows between OECD and non-OECD members (that is, US/UK/DEU ↔ IND/CHN) decreased from 25.6 per cent in the 1990s to 19.8 per

⁶⁶ One curious exception is Chinese data, which systematically reports lower values of Chinese exports than trading partner report Chinese imports. Re-exporting via Hong Kong may be one reason for this pattern. Zhu, Yamano, and Cimper 2011, 27–29. Excluding China from the analysis, in the remaining 262 directed dyad-year observations, importing economies report higher values than respective exporters partners in 170 cases, while exporters – against the trend of f.o.b. vs. c.i.f. valuation - report higher values than importers in 92 cases.

cent in the 2000s, the discrepancy for trade among the three OECD countries in the sample *increased* from 5.6 to 8.8 percent in the same period.

Bearing in mind that merchandise trade indicators are generally considered to be among the most reliable information that can be derived from BOP statistics, the magnitude of these discrepancies at the bilateral level is worrisome, to say the least. Discrepancies in bilateral merchandise trade statistics are several times as large as those at the global aggregate level. If this difference could be extrapolated to capital flow statistics – where global aggregate differences are already very substantial (cf. Figure 3) – these observations would suggest that bilateral capital flow figures should be treated with even more caution.

At the same time, these severe measurement problems are only one part of the international data problems we wish to highlight. The other, in some ways even more problematic, dimension in the gradual deterioration of BOP data quality is the widening gap between the statistical operationalizations of key quantities of interest and the concepts that users have in mind when referring to these statistics.

The widening concept-measurement gap

When commentators, journalists, policymakers, or academic researchers talk about trade, FDI or PFI flows, they typically refer to transactions that involve a good or capital flow that i) crosses a national border and ii) changes hands from a domestic to a foreign owner (or the reverse): ‘exports’ stand for products made in the home country that are moved abroad to be sold to foreign customers; ‘FDI inflows’ are short-hand for foreign companies moving into a country to create a new company there (or to buy an existing one); ‘PFI outflows’ refer to investors putting their capital at work outside their home economy, and so on. Analysts commonly build their insights about these phenomena on BOP statistics. However, the statistical operationalization of key concepts in international economic affairs frequently do not correspond to these interpretations. Indeed, although the existence of this concept-measurement gap is nothing new, it has become significantly worse in recent years.

In the case of trade, for instance, the de-nationalization of economic production and the growing importance of services have boosted ‘trading’ activities in statistical terms that do not correspond to the traditional notion of trade as a product made in country A moving across the border to country B for consumption there. Many products cross national borders numerous times throughout their production process. While the WTO and OECD’s *Trade in Value Added* (TiVA) joint initiative aims to take such dynamics into account and some changes into that direction were also introduced in the BPM5 and BPM6 guidelines, administrative data collection by customs authorities still tend to count the full value of exports and imports each time that a good crosses a border – irrespectively of whether it was entirely produced

there, processed, or merely passed through without any modification (such as in the case of re-exporting) – inflating and distorting trade statistics in the process.⁶⁷

At the same time, the growing importance of service activities in overall trade statistics challenge traditional notions of trade on a more fundamental level: while some types of services trade, such as those provided by jet-setting consultants, do involve a physical crossing of borders by the service provider, most do not. In the cases of tourism and university tuition fees (two sizable components of total trade in services⁶⁸), for example, it is the consumption patterns of travelers and international students consuming locally provided services, which are the source of ‘trade’. And in the case of intellectual property products (IPPs) (e.g. franchising, licensing, etc.), the most important as well as most problematic component of services trade, ‘exporting’ activity essentially refers to legal and accounting procedures rather than a product’s physical crossing of national borders.

At the same time, the intangibility of IPPs⁶⁹ make them a favorite object of MNCs’ transfer pricing strategies. As the IMF BOP Compilation Guidelines note: “Between companies in a direct investment relationship, transactions may occur wherein values shown in the books of transactors are significantly distorted from market values. A company may sell goods to a related company for prices unrelated to the cost of production or the acquisition cost of the goods. Such a sale might be made, for example, to transfer profits from one economy to another for tax reasons or because the economy of the direct investment enterprise (DIENT) imposes restrictions on the repatriation of income.”⁷⁰ Such dynamics can lead to the artificial creation of large amounts of “phantom international [trade] flows”.⁷¹ For instance, there are strong indications that US companies book large parts of their US sales as services ‘exported’ from subsidiaries located in low-tax jurisdictions.⁷² In other words, geographically purely domestic transactions are in this way technically recorded as international trading activity in BOP statistics, leading to a situation in which “measures of the current balance and national income and output begin to lose their meaning”⁷³.

The state of affairs is no better for capital flow statistics. Similar to the case of trade, many of the activities that BOP statistics captured as FDI do not involve the actual crossing of a border, while the growing complexity of corporate ownership relations has raised more fundamental questions about the ability of BOP data collection procedures to measure FDI or to distinguish FDI flows from other types of capital

⁶⁷ An analysis undertaken by the Dallas FED suggests, for instance, that correcting trade balances for value added reduces the US trade deficit with China in 2009 by 33 percent, from USD 189 to USD 126 billion. Sposi and Koech 2013, 3.

⁶⁸ Lipsey 2006.

⁶⁹ “The intangible nature of IPPs means that they can easily be registered as the property of a unit in one country when they are used in production by an enterprise located in another”. UNECE, Eurostat, and OECD 2011.

⁷⁰ International Monetary Fund 2014.

⁷¹ Lipsey 2006:46

⁷² “If the location of production is ambiguous, the distinction between home production and imports of services is correspondingly ambiguous, as is the distinction between home production and exports”. Lipsey 2009, 29.

⁷³ Lipsey 2006: 50.

flows. Commonsensical understandings of FDI portray them as long-term investments ‘coming in’ from abroad, be it to build a new company or to buy an existing one. The truth is, however, that a large share of ‘statistical’ FDI are not a net addition to a national economy’s capital stock, but are rather generated locally, either through the income that previously established foreign affiliates produce (i.e. re-invested earnings) or by foreign investors raising capital directly in local markets – a practice that has been facilitated significantly by financial liberalization⁷⁴. In either case, *de facto* domestic activities by foreign-owned firms are construed as cross-border economic transactions even though no factual crossing of any physical border is involved.

At the same time, ever larger shares of cross-border capital flows are channeled through Special Purpose Entities (SPEs; ‘letterbox companies’ in colloquial parlance) located in low-tax offshore jurisdictions. BOP statistics, which classify any cross-border investment involving an ownership stake of 10 percent or more as FDI and allocate flows according to the countries through which transactions are made rather than their ultimate origin or destination, are unable to identify the purpose or actual national ownership relationships underlying such flows. Given that up to half of FDI outflows leaving the USA – one of the few countries publishing such data – are being channeled through such conduit structures in most recent years⁷⁵, this is a problem of considerable importance. Recent studies have shown, for example, that patterns in ‘BOP-FDI’ data are much more akin to the behavior of short-term capital than one would expect⁷⁶ and that the distribution of their book value in comparison to indicators of actual economic activity is heavily distorted towards tax haven countries.⁷⁷ In other words, what gets reported as FDI may have little to do with traditional notions of long-term productive investment, and for substantial shares of reported FDI flows it is simply impossible to determine ultimate origin or destination.

Data on PFI flows are plagued by similar issues. For instance, it is often not clear whether registered in- or outflows of securities or loans really enter or leave the country in question. An analysis by the US Treasury found, for example, that nearly two thirds (!) of total registered portfolio equity ‘outflows’ from the US in the 1990s were in fact accounted for by stock swaps resulting from foreign takeovers of US firms.⁷⁸ In other words, more than half of the money that in commonsensical interpretations of PFI statistics were considered having been invested abroad by US citizens did not in fact leave the US economy, but were merely registered as such because of a change in legal nationality of the US-based companies in which these investments were held.

⁷⁴ A recent study indicates that as much as two thirds of the debt of US foreign affiliates abroad is raised locally. Bilir, Chor, and Manova 2017, Table 2.

⁷⁵ Ibarra-Caton and Mataloni 2014, 7–8.

⁷⁶ Blanchard and Acalin 2016.

⁷⁷ Lipsey 2007; Kerner 2014.

⁷⁸ Griever, Lee, and Warncock 2001, 641.

But, as in the case of trade and FDI, probably the greatest challenge for the validity of BOP-PFI statistics lies in the offshorization of global finance⁷⁹.

Banks and investment funds are among the most avid users of transnational holding company structures. For instance, two recent studies suggest that about 60 percent of the total assets of the global hedge fund sector (worth more than USD 3 trillion⁸⁰) are legally domiciled in Cayman Islands⁸¹, three tiny Caribbean islands with about 58,000 inhabitants. Given that BOP statistics are collected according to the jurisdictions through which transactions are made rather than the nationality of the ultimate owners of these capital flows, such dynamics appear like a truly daunting prospect for the suitability of BOP statistics to offer a meaningful picture of trends in cross-border capital flows.

The 4Cs of national accounting

The growing CMG that we observe is thus primarily the result from a combination of two forces: profound structural transformations on the one hand, and a reluctance to adapt statistical concepts to these changes on the other. Although it is clear that the adaptation of statistical concepts to profound and complex economic evolutions is anything but straightforward, the resistance to reform data collection and aggregation procedures more comprehensively is nonetheless remarkable and somewhat puzzling in theoretical terms (as we have argued above). After all, statisticians are keenly aware of these issues and do consider them as serious problems⁸². The observation that statisticians uphold concepts that they themselves describe as ‘outdated’ and ‘out of sync’⁸³ is not trivial. It calls for explanation. We argue that it derives primarily from four dynamics that are inherent to the enterprise of national accounting itself – the aspiration to comparability, continuity, consistency and coherence.

Comparability

A key ambition of the national accounting enterprise is to provide metrics that are comparable across countries. As we have elaborated above, the harmonization of accounting-technical standards that this aspiration involves is an extremely challenging and costly undertaking. As a result, the wish for *comparability* retards change. Even though there is no shortage of good ideas among statistical communities about how to reduce the CMG, a key difficulty is to reach international agreement on new standards that all parties

⁷⁹ Cf. International Monetary Fund 1992.

⁸⁰ Fichtner 2016, 1051.

⁸¹ Fichtner 2016.

⁸² Interview 20170425

⁸³ Interview 20170425

find acceptable and implementable. For instance, to achieve a more fine-grained picture of different types of cross-border investment flows, the OECD's 4th edition of the Benchmark Definition of FDI (developed over several years in coordination with the IMF and UN) has asked countries to compile separate figures for greenfield, M&A and SPE IFDI flows since 2008⁸⁴. Yet, despite the very significant improvement that such a distinction would entail, progress has been very slow⁸⁵. And as long as only few countries implement the new standards, they cannot be included in cross-national databases. In other words, the daunting coordination problems that the implementation of new statistical methods across close to two hundred national bureaucracies involve, makes change very difficult and inevitably means that statistical standards will lag significantly behind economic developments.

Continuity

A closely related dynamic is the aspiration for *continuity*. One of the great attractions of statistics is their ability to capture macrosocial or -economic developments over time. If we adapt indicators along the way and have no way of adjusting past measurements retrospectively, diachronic comparability is inevitably lost. The answer to this problem often lies in the development of parallel statistics: starting to build a new, more appropriate series while continuing the increasingly obsolete one for the time being. For example, although it is generally accepted that the figures from the WTO-OECD TiVA initiative provide more accurate indications of cross-country trade than conventional statistics, these figures will not be directly integrated into the BOP system to avoid a break in the series. Continuity over time will be more important for one time series than for another, and more relevant for some users than for others. And the break in the series that a conceptual change would generate will vary from case to case. What is clear, however, is that the need to adapt indicators continually to changing economic realities diminishes the comparability of data over time, such that authorities may well face incentives to stick with indicators even if they are becoming increasingly obsolete.

Consistency

The ambition for *consistency* means that measurement instruments should give the same reading on repeated application and not depend on subjective judgment or invite gross manipulation. The intuition is obvious and in line with good statistical practice. It does, however, come at a cost. Even where informed estimates might generate the best data, they may be eschewed for measurement procedures that rely on harder, external figures. FDI statistics are again a good example. A key challenge for capital flow statistics is to

⁸⁴ See OECD 2017b.

⁸⁵ Interview 20170530

determine whether an investment is predominantly ‘financial’ (i.e. PFI) or ‘productive’(i.e. FDI). While earlier versions of the Balance of Payments Manual explicitly called for national accountants’ qualitative judgement in these regards⁸⁶, the IMF subsequently abandoned this approach in favour of the establishment of a clear threshold rule which determines that all foreign investments that involve at least ten percent of a company’s voting stock are to be counted as FDI. The arbitrariness and conceptual shortcomings of such a rule are clear. But it can be uniformly applied across governments, and it is consistent in the methodological sense: having a different person repeat the same procedure would yield rather similar results. But consistent rules are less flexible than qualitative judgments and risk ignoring changing circumstances. In a trade-off familiar to social scientists, reliability comes at the cost of validity.

Coherence

Coherence as an aim means that individual economic measures should fit into a larger, coherent whole that depicts the entire economy. The BOP systems doesn’t stand on its own, but is itself an integrated component of the System of National Accounts (SNA), and “linkage of the international investment position and balance of payments accounts to the rest of the world account in the System of National Accounts (SNA) is strengthened and harmonized to the maximum extent possible”.⁸⁷ The different sectoral accounts “hang together”, such that a change in one account, , must be accounted for somewhere else, as well – i.e. a trade deficit comes with a capital account surplus, and export revenues in the BOP are also someone’s income in the SNA. . Thinking of national economic and BOP statistics as one integrated whole also means that statistical concepts are often deductively defined. The conceptual coherence of accounts tempts us to impute values for concepts that are not directly observable. If in an equation such as $a + b = X$ we have measures for a and X , we might impute b and then report this as a known quantity. But in the process all kinds of measurement problems with a and X disappear.

Imprecision, lost in translation: Accounting for the neglect of measurement problems

The dynamics just outlined, we argue, all present incentives to produce data that suffer from potentially growing concept-measurement gaps. Considering these trade-offs in data production, some of these gaps may well be inevitable. But given the seriousness of these issues, which have the potential to substantially distort economic analyses, it remains an open question why they receive so little acknowledgement in the design of empirical investigations or the presentation of results that build on this data.

⁸⁶ See for instance: International Monetary Fund 1961, 120; International Monetary Fund 1977a, 138.

⁸⁷ International Monetary Fund 1993, 3.

While a full-fledged investigation of these issues is beyond the ambit of this article, our research sheds some light on the dynamics that lead to this outcome. In particular, we observe a continual dilution of measurement imprecision at each stage of the production chain of economic statistics. The main problem, we find, lies in the loss of nuance and caution as data passes from one entity to the next. National Statistical Offices frequently highlight measurement difficulties or uncertainties when they forward the statistical tables that international organizations collect from them and at times provide several estimates for the same figure. If possible, IO staff discuss such issues with representatives from national agencies and may revise the original figures in coordination with them. But at other times such quality checks have to be omitted due to resource constraints.⁸⁸ In either case, once international organizations (such as the IMF, OECD or UN) assemble the collected data into cross-national databases, only one – seemingly precise – figure ends up there. Frequently IOs add disclaimers to the published data that mention some of these issues. Yet, these reservations typically do not figure very prominently. A meta-data link in an online database or a footnote somewhere in a Statistical Yearbook might mention that values have been imputed, but even (the few) users who read it and pause to consider its implications will find it impossible to gauge just how much uncertainty this imputation masks. In all likelihood, they pay no further attention and simply work with the values provided in the databases as solid estimates.

Indeed, numbers are so attractive that hardly anyone using them has an incentive to play up their imprecision. Journalists, press officers, academics, professional analysts all thrive on offering information. An acknowledgement that what looks like a hard number may just be noise, or that a regression analysis is shaky because the underlying data contains big margins of error, goes against the professional incentives all of these actors face⁸⁹. Without any bad intentions, then, nuances and health warnings for the users of the data get lost as it is passed along the food chain. The semblance of precision increases even as in fact, it withers.

These professional incentives are underpinned by a growing demand for quantitative data. In an attempt to reach objectively sound and hence defensible decisions, policymakers increasingly resort to scientific – which often means quantitative – analysis.⁹⁰ Impact assessments and cost-benefit analyses have become integral to economic policy. They do not rhyme well, of course, with a decrease in the quality of important kinds of data. Indeed, the abundance of information that is produced and available for analysis through digital technology buttresses the belief that data and its analysis are the key to better policy.⁹¹ Once administrative apparatuses and routines have embraced that mode of operation, professional incentives discourage acknowledging the full extent of data imprecision.

⁸⁸ Interviews20170425 and 20170530.

⁸⁹ Decades of serious work invested in the development of statistical models that account for error models have received remarkably little attention, for instance. See Fuller 1987.

⁹⁰ Marcussen 2005 has detailed a parallel dynamic for central banks.

⁹¹ Cf. Data Revolution Group 2014.

Conclusions

BOP data is our window on the global economy. It is what we use to gauge and analyze trends in global economic affairs – trade, international financial flows, and cross-border investment patterns. We normally take this data for granted, and analysts of the global economy rarely spend much time dissecting the data itself.

In this article, we have offered a different view of BOP data. To begin with, the reliability of the data is remarkably low, and it has decreased over the past two decades rather than increased. These data defects are mostly caused by the globalization of economic activity. Patterns of production, trade, and financial flows no longer conform to textbook imagery in which country A sends a domestically produced good to country B and in return receives a payment that can be traced to consumers in that country. Multinational enterprises, obscure special purpose entities, highly fragmented production chains and complex patterns of debts and credits have proliferated. National accounting templates that presume simplistic economic relationships capture our economic realities less and less well.

This growing gap between what the data aspires to represent and what it actually captures has potentially worrying implications. After all, credit rating agencies, investors and international organizations rely on macroeconomic data in their country assessment and surveillance, often with material consequences for those countries. Such data can also inform international judicial deliberations (such as in WTO arbitration panels) and can thus carry hard-wired legal consequences. It feeds research and analysis on determinants and consequences of international economic flows undertaken by academics, policy analysts and journalists, nurturing construction of broader narratives about macroeconomic trends and development trajectories. And every now and then, these numbers themselves become directly politicized, for example in spats about American trade relations with China, Mexico or Germany, or the trading position of Germany within the European Union.

If the gaps in the data are big and growing, it is all the more remarkable that they are not mirrored in an institutionalized data-skepticism. The statisticians we interviewed are well aware of the problems we discuss in this article. Yet unless you pay attention to footnotes in statistical yearbooks or other details, the published numbers continue to betray an accuracy that sits uneasily with the ambiguities in the data. Headline data takes relatively little heed of the growing problems, for example by overhauling definitions or replacing outdated concepts. And data users – analysts, policymakers, journalists, etc. – frequently use the data as if nothing was wrong.

To round off our analysis, we have therefore asked why data reporting and treatment does not move in sync with these new challenges. Two clusters of answers emerge: first, in the production of data, statisticians prioritize a range of data qualities that are actually at odds with accommodating messy economic realities. They prioritize consistency, coherence, continuity and comparability – all of which insert a conservative bias into measurement systems, and favor a more deductive, textbook style approach to data categorization

and measurement. Second, because there is a high demand for quantitative data, neither data producers nor data users have incentives to highlight the defects of data. As data is handed from one individual or organization to the next, awareness of potential ambiguities or inaccuracies is likely to get downplayed, ignored or lost. To be sure, international organizations such as the OECD pro-actively seek to develop alternatives to increasingly obsolete indicators – the Trade in Value-Added (TiVA) project is a good example. It also highlights, however, how hard it is both to gather data that is sufficiently robust and to have such new measures replace better-known and widely entrenched metrics such as the standard trade statistics offered by the OECD and the IMF.

Our analysis has a range of implications that go beyond the scholarly contribution itself. It shows that even data as authoritative as that based on the BPM should be used with care. Users should appreciate the significant margins of error in the data, which may make them inappropriate bases for the formulation of strong causal inferential claims, or at least require the use of probability intervals rather than point estimates in statistical analyses. And even when the errors seem manageable, users should understand what the different metrics, for example those covering different categories of investment, actually measure. Without such awareness, it remains a real possibility that our analyses and the real international economy “out there” become more and more detached.

References

- Beckert, Jens. 2016. *Imagined futures: fictional expectations and capitalist dynamics*. Cambridge, Massachusetts: Harvard University Press.
- Bilir, Kamran, Davin Chor, and Kalina Manova. 2017. *Host-Country Financial Development and Multinational Activity*. NBER Working Paper Series. Available at <<http://www.nber.org/papers/w20046.pdf>>.
- Blanchard, Olivier, and Julien Acalin. 2016. *What Does Measured FDI Actually Measure?* Washington, D.C. Available at <<https://piie.com/system/files/documents/pb16-17.pdf>>.
- Blaug, Mark. 1996. *Economic theory in retrospect*. Vol. 5th. Cambridge: Cambridge University Press.
- Broome, André, and Joel Quirk. 2015. The politics of numbers: the normative agendas of global benchmarking. *Review of International Studies* 41 (5): 813–818. Available at <http://www.journals.cambridge.org/abstract_S0260210515000339>.
- Bryan, Dick. 2001. Global accumulation and accounting for national economic identity. *Review of Radical Political Economics* 33 (1): 57–77.
- CBS. 2017. Budget cuts at Statistics Netherlands: state of play. Available at <<https://www.cbs.nl/en-gb/about-us/organisation/operational-management/archief/budget-cuts-at-statistics-netherlands-state-of-play>>. Accessed 2 June 2017.
- Clegg, Liam. 2010. Our Dream is a World Full of Poverty Indicators: The US, the World Bank, and the Power of Numbers. *New Political Economy* 15 (4): 473–492. Available at <<http://www.tandfonline.com/doi/abs/10.1080/13563461003763170>>.
- Coyle, Diane. 2014. *GDP: a brief but affectionate history*. Princeton, N.J.: Princeton University Press.
- Data Revolution Group. 2014. *A World That Counts: Mobilising the Data Revolution for Sustainable Development*. Available at <<http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>>.
- Dicken, Peter, Philip F. Kelly, Kris Olds, and Henry Wai-Chung Yeung. 2001. Chains and networks, territories and scales: towards a relational framework for analysing the global economy. *Global Networks* 1 (2): 89–112.
- Dimand, Robert. 1995. ‘I have no great faith in political arithmetick’ Adam Smith and quantitative political economy. In *Measurement, Quantification and Economic Analysis. Numeracy in Economics*, edited by Ingrid Rima, 22–30. London: Routledge.
- Dunleavy, Patrick, and Helen Margetts. 2010. The Second Wave of Digital Era Governance. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371 (1987). Available at

<http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1643850%5Cn<http://eprints.lse.ac.uk/27684/>>.

Eurostat. 2004. *Differences between Balance of Payments and Foreign Trade Statistics*. Luxembourg. Available at <<http://ec.europa.eu/eurostat/documents/3888793/5829933/KS-DB-04-001-EN.PDF/9948afe5-bcf6-45b2-85d6-2f580d6ec759>>.

Fichtner, Jan. 2016. The anatomy of the Cayman Islands offshore financial center: Anglo-America, Japan, and the role of hedge funds. *Review of International Political Economy* 23 (6): 1034–1063.

Finnemore, Martha. 1996. *National interests in international society*. Ithaca, N.Y.: Cornell University Press.

Fioramonti, Lorenzo. 2014. *How numbers rule the world: the use and abuse of statistics in global politics*. London: Zed Books.

Frieden, Jeffrey A. 2006. *Global capitalism: its fall and rise in the twentieth century*. 1st ed. New York: Norton.

Fuller, Wayne A. 1987. *Measurement Error Models*. New York: John Wiley & Sons.

Gereffi, Gary, John Humphrey, and Timothy J. Sturgeon. 2005. The governance of global value chains. *Review of International Political Economy* 12 (1): 78–104.

Griever, William L., Gary A. Lee, and Francis E. Warncock. 2001. The U.S. System for Measuring Cross-Border Investment in Securities: A Primer with a Discussion of Recent Developments. *Federal Reserve Bulletin* (October).

Hirschman, Daniel, and Elizabeth Popp Berman. 2014. Do economists make policies? On the political effects of economics. *Socio-Economic Review* 12 (4): 779–811.

Ibarra-Caton, Marilyn, and Raymond J. Jr. Mataloni. 2014. *Direct Investment Position for 2013*. Washington, D.C. Available at <https://bea.gov/scb/pdf/2014/07July/0714_direct_investment_positions.pdf>.

International Monetary Fund. 2009. *Balance of Payments and international Investment Position Manual*. Washington, D.C.

International Monetary Fund. 1993. *Balance of Payments Manual - Fifth Edition*. Washington, D.C.

International Monetary Fund. 1948. *Balance of Payments Manual - First Edition*. Washington, D.C.

International Monetary Fund. 1977a. *Balance of Payments Manual - Fourth Edition*. Washington, D.C.

International Monetary Fund. 1950. *Balance of Payments Manual - Second Edition*. Washington, D.C.

International Monetary Fund. 1961. *Balance of Payments Manual - Third Edition*. Washington, D.C.

International Monetary Fund. 2014. *BPM6 Compilation Guide: Companion Document to the Sixth Edition of the*

Balance of Payments and International Investment Position Manual.

- International Monetary Fund. 1992. *Report on the Measurement of International Capital Flows*. Washington, D.C.
- International Monetary Fund. 1987. *Report on the World Current Account Discrepancy*. Washington, D.C.
- International Monetary Fund. 1977b. *The Monetary Approach to the Balance of Payments*. Washington DC.
- International Monetary Fund, and Organisation for Economic Co-operation and Development. 2003. *Foreign Direct Investment Statistics: How Countries Measure FDI 2001*. Washington, D.C. Available at <<https://www.imf.org/external/pubs/ft/fdis/2003/fdistat.pdf>>.
- Jessop, Bob. 2002. *The Future of the Capitalist State*. Oxford: Polity Press.
- Karabell, Zachary. 2014. *The leading indicators: a short history of the numbers that rule our world*. New York: Simon & Schuster.
- Kelley, Judith G., and Beth A. Simmons. 2015. Politics by number: Indicators as social pressure in international relations. *American Journal of Political Science* 59 (1): 55–70.
- Kerner, Andrew. 2014. What We Talk About When We Talk About Foreign Direct Investment. *International Studies Quarterly* 58 (4): 804–815.
- Lepenies, Philipp. 2013. *Die Macht der einen Zahl: eine politische Geschichte des Bruttoinlandsprodukts. Politische Geschichte des Bruttoinlandsprodukts*. Erste Aufl. Berlin: Suhrkamp.
- Lequiller, Francois, and Derek Blades. 2014. *Understanding National Accounts*. Paris.
- Lipsey, Robert E. 2006. Measuring International Trade in Services. *National Bureau of Economic Research Working Paper Series* No. 12271 (May). Available at <<http://www.nber.org/papers/w12271>>.
- Lipsey, Robert E. 2007. *Defining and Measuring the Location of FDI Output*. NBER Working Paper Series. Available at <<http://www.nber.org/papers/w12996>>.
- Lipsey, Robert E. 2009. Measuring International Trade in Services. In *International Trade in Services and Intangibles in the Era of Globalization*, edited by Marshall Reinsdorf and Matthew J. Slaughter, 27–70. Chicago: University of Chicago Press. Available at <<http://www.nber.org/chapters/c11605.pdf>>.
- Lipsey, Robert E. 2010. Measuring the location of production in a world of intangible productive assets, fdi, and intrafirm trade. *Review of Income and Wealth* 56 (SUPPL. 1).
- Marcussen, Martin. 2005. Transcending Politics: A New Stage in Management Transformation? The Scientization of Central Banking as Case. In *SCANDOR/SOG workshop*. Stanford University.
- Marquardt, Kyle L., and Yoshiko M. Herrera. 2015. Ethnicity as a Variable: An Assessment of Measures

- and Data Sets of Ethnicity and Related Identities*. *Social Science Quarterly* 96 (3): 689–716.
- Maslin, Mark, and Patrick Austin. 2012. Uncertainty: Climate models at their limit? *Nature* 486 (7402): 183–184. Available at <<http://www.nature.com/doi/10.1038/486183a>>.
- Mayer-Schonberger, Viktor, and Kenneth Cukier. 2013. *Big data: A revolution that will change how we live, work and think*. London: John Murray.
- McCormick, Ted. 2009. *William Petty and the Ambitions of Political Arithmetic. William Petty: And the Ambitions of Political Arithmetic*. Oxford: Oxford University Press.
- Mitchell, Timothy. 2002. *Rule of experts: Egypt, techno-politics, modernity*. Berkeley: University of California Press.
- Morgenstern, Oskar. 1963. *On the accuracy of economic observations*. Second edi. Princeton, N.J.: Princeton University Press.
- Moulton, Brent. 2017. Reading between the lines: What Trump’s Budget Blueprint may mean for funding statistics. Available at <<https://politicalarithmetick.com/2017/03/16/reading-between-the-lines-what-trumps-budget-blueprint-may-mean-for-funding-statistics/>>. Accessed 2 June 2017.
- OECD. 2017a. Glossary of Statistical Terms. Available at <<https://stats.oecd.org/glossary/detail.asp?ID=1406>>. Accessed 1 June 2017.
- OECD. 2017b. OECD Benchmark Definition of Foreign Direct Investment - 4th Edition. Available at <<http://www.oecd.org/investment/fdibenchmarkdefinition.htm>>. Accessed 2 June 2017.
- OECD. 2001. *Statistics Brief - Trade in Goods and Services: Statistical Trends and Measurement Challenges*. Paris. Available at <<https://www.oecd.org/trade/its/2539563.pdf>>.
- Palan, Ronen. 2002. Tax Havens and the Commercialization of State Sovereignty. *International Organization* 56 (1): 151–176.
- Perlman, Mark. 1987. Political Purpose and the National Accounts. In *The Politics of Numbers*, edited by William Alonso and Paul Starr, 133–152. New York: Russell Sage Foundation.
- Petersen, William. 1987. Politics and the Measurement of Ethnicity. In *The Politics of Numbers*, edited by William Alonso and Paul Starr, 187–233. New York: Russell Sage Foundation.
- Philipsen, Dirk. 2015. *The little big number: how GDP came to rule the world and what to do about it*. Princeton, N.J.: Princeton University Press.
- Picciotto, Sol. 2012. *Regulating Global Corporate Capitalism*. Cambridge: Cambridge University Press.
- Polak, Jacques J. 1997. *The IMF Monetary Model at Forty*. IMF Working Paper. Washington, D.C. Available at <<http://www.imf.org/external/pubs/ft/wp/wp9749.pdf>>.

- Randall, David A., and Richard A. Wood. 2007. Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*: 590–662. Available at <http://www.geog.ucsb.edu/~joel/g280_s09/model_eval/AR4WG1_Pub_Ch08.pdf>.
- Schmelzer, Matthias. 2016. *The hegemony of growth: the OECD and the making of the economic growth paradigm*. Cambridge, U.K.: Cambridge University Press.
- Seabrooke, Leonard, and Duncan Wigan. 2017. The governance of global wealth chains. *Review of International Political Economy* 24 (1): 1–29. Available at <<https://www.tandfonline.com/doi/full/10.1080/09692290.2016.1268189>>.
- Shaxon, Nicholas. 2011. *Treasure Islands: Tax Havens and the men who stole the world*. London: Palgrave Macmillan.
- Snyder, Jack L, and Alexander Cooley, eds. 2015. *Ranking the world: grading states as a tool of global governance*. Cambridge, U.: Cambridge University Press.
- Sposi, Michael, and Janet Koech. 2013. Value-Added Data Recast the U.S.-China Trade Deficit. *Dallas FED: Economic Letter*. Dallas. Available at <<http://www.dallasfed.org/assets/documents/research/ecllett/2013/el1305.pdf>>.
- Stiglitz, Joseph E, Amartya Sen, and Jean-Paul Fitoussi. 2010. *Mismeasuring our lives: why GDP doesn't add up*. New York, London: The New Press.
- Studenski, Paul. 1958. *The Income of Nations: Theory, Measurement and Analysis Past and Present*. New York: New York University Press.
- UNECE, Eurostat, and OECD. 2011. *The Impact of Globalization on National Accounts*. New York and Geneva: United Nations.
- UNSD, and Eurostat. 2011. *Measuring Global Trade - Do we have the right numbers?* Geneva. Available at <[https://unstats.un.org/unsd/trade/events/2014/mexico/documents/general/Global Forum on Trade Statistics - Report.pdf](https://unstats.un.org/unsd/trade/events/2014/mexico/documents/general/Global_Forum_on_Trade_Statistics_Report.pdf)>.
- Vanoli, André. 2005. *Une histoire de la comptabilité nationale*. Paris: Editions La Découverte.
- Ward, Michael. 2004. *Quantifying the world: UN ideas and statistics*. Bloomington: Indiana University Press.
- Woods, Ngaire. 2006. *The Globalizers. The IMF, the World Bank, and Their Borrowers*. Ithaca: Cornell University Press.
- Zhu, Shiguang, Norihiko Yamano, and Agnès Cimper. 2011. *Compilation of Bilateral Trade Database by Industry and End-Use Category*. OECD Science, Technology and Industry Working papers. Paris. Available at <<http://www.oecd-ilibrary.org/docserver/download/5k9h6vx2z07f-en.pdf?expires=1496392737&id=id&accname=guest&checksum=79C64B040D3858D1A9A184AA>>

089E1C64>.