EURONA

EUROSTAT REVIEW ON NATIONAL ACCOUNTS AND MACROECONOMIC INDICATORS

2/2017

eurostat

EUROSTAT REVIEW ON NATIONAL ACCOUNTS AND MACROECONOMIC INDICATORS

2/2017

Printed by *Imprimerie Centrale* in Luxembourg
Manuscript completed in January 2018

Luxembourg: Publications Office of the European Union, 2018

Print	ISSN 2443-7832	Cat.: KS-GP-17-002-EN-C
PDF	ISSN 1977-978X	Cat.: KS-GP-17-002-EN-N

Contents

Editorial 5

Financial Stability Analysis: What are the Data Needs?
Robert Heath and Evrim Bese Goksu 7

Explaining nowcast errors
Henriette Druba, Jennifer L.Castle and David F. Hendry 57

Advancements in measuring intangibles for European economies
Carol Corrado, Jonathan Haskel, Massimiliano Iommi, Cecilia Jona-Lasinio,
Matilde Mas and Mary O'Mahony 89

A focused strategy for the estimation of Italian quarterly external trade
by geographical area, open issues and solutions
Giancarlo Lutero and Paolo Forestieri 107

Aims and scope

EURONA is an open access, peer-reviewed, scholarly journal dedicated to National Accounts and Macroeconomic Indicators. EURONA aims at providing a platform for researchers, scholars, producers and users of macroeconomic statistics to exchange their research findings, thereby facilitating and promoting the advancement of National Accounts and Macroeconomic Indicators.

EURONA publishes empirical and theoretical articles within the scope of National Accounts and Macroeconomic Indicators, as well as articles on important policy uses of these statistics. They may relate to both users' and producers' interests, present subjects of general relevance or investigate specific topics.

EURONA is non-partisan and applies the highest standards to its content, by emphasising research integrity, high ethical standards, validity of the findings and cutting edge results. EURONA gives room to all viewpoints.

The articles published in EURONA do not necessarily reflect the views or policies of the European Commission.

Website: http://ec.europa.eu/eurostat/web/national-accounts/publications/eurona

Contact: ESTAT-EURONA@ec.europa.eu

Editors

Editorial board

(1) Paulus.Konijn@ec.europa.eu
(2) sartore@unive.it

Editorial

That the national accounts domain covers a wide variety of different subjects, purposes, techniques and methodologies is once again demonstrated in this second issue of EURONA of 2017, which includes articles related to the financial markets, to nowcasting techniques, to the measurement of intangible assets and to the estimation of data on foreign trade.

In the first article, Robert Heath and Evrim Bese Goksu discuss the efforts undertaken by the statistical community to close the data gaps that were uncovered during and after the global financial crisis, in particular as regards data needed for the monitoring and analysis of financial stability. They also identify the work that still needs to be done and suggest a number of areas where the statistical manuals could be further developed.

Henriette Druba, Jennifer Castle and David Hendry present a taxonomy for nowcast errors in the second article of this issue. This taxonomy, and the simulation exercise on nowcasting consumption data that was carried out by the authors, provide a number of insights that can help to determine the main sources for error and thus to improve nowcasts.

Intangible assets play a crucial role as a driver for economic growth in today's global and digital economy. In the third article, Carol Corrado, Jonathan Haskel, Massimiliano Iommi, Cecilia Jona-Lasinio, Matilde Mas and Mary O'Mahony provide an overview of recent advancements in measuring investment in intangible assets, including those that are not (yet?) part of the standard national accounts framework. Estimates for those have been developed in a series of research projects, for both private and public sectors.

Finally, Giancarlo Lutero and Paolo Forestieri present the methods developed by the Italian statistical office to estimate quarterly imports and exports data by geographical area, taking into account the new standards set by ESA 2010 as regards the treatment of good sent for processing and merchanting. Several models have been empirically tested to determine the best way forward.

To close, I would like to thank the team of Hendyplan - François Libeau, Domenico Sartore, Agathe Guelen, Christiane Gonzalez and Yves Robinson Kruse-Becher - for the support provided during the first four years of EURONA. Hendyplan provided both editorial and technical services for the production of this journal. Their dedication and professionalism contributed significantly to the success of EURONA that has been achieved so far.

Paul Konijn

Editor of EURONA

1 Financial stability analysis: what are the data needs?

ROBERT HEATH [1] AND EVRIM BESE GOKSU [2] [3]

Abstract: The growing incidences of financial crises and their damage to the economy has led policy makers to sharpen the focus on financial stability analysis (FSA), crisis prevention and management over the past 10-15 years. The statistical world has reacted with a number of initiatives, but does more need to be done? Taking a holistic view, based on a review of experiences of policy makers and analysts, this paper identifies common international threads in the data needed for FSA and suggests ways to address these. While there has been an encouragingly constructive response by statisticians, not least through the G-20 Data Gaps Initiative, more work is needed, including with regard to shadow banking, capital flows, corporate borrowing, and granular data. Further, to support FSA, the paper identifies potential enhancements to the conceptual advice in statistical manuals including with regard to foreign currency and remaining maturity.

JEL codes: E44, F65, G1, G2

Keywords: financial stability, data gaps, financial interconnections, spillovers, financial sector, credit, debt, global financial crisis, macro-prudential analysis, macroeconomic statistical manuals, stress testing

[1] Former Deputy Director of the IMF Statistics Department.
[2] IMF Statistics Department.
[3] The views expressed in this paper are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management. This paper draws heavily on the previous research published as IMF Working Paper, WP/17/153.

1. Introduction

The first chapter of Charles Kindleberger's 1978 seminal work on financial crises 'Manias, Panics and Crashes' is entitled 'Financial Crisis: A Hardy Perennial.' But the chapter starts by pointing out the relative lack of such crises, particularly in advanced economies, during the several decades after World War II, the period when the core economic and financial statistical manuals of national accounts and balance of payments were developed to help support macroeconomic policy making.

Recent decades have witnessed the return of significant financial crises, notably, but not only, the global financial crisis (GFC) of 2007/8 that have resulted in significant losses to the real economy (4) including years of under-performance in economic growth. The growing incidences of financial crises and their damage to the economy and the well-being of the population, together with the increased scale and interconnectedness of financial transactions, and their complexity, has led policy makers to give a greater focus to financial stability analysis (FSA) (5), financial system resilience, crisis prevention, and management over the past 10–15 years.

Consequential to the focus on financial stability, the desired composition of the policy makers' tool boxes and the nature of the data needed to support policy has changed. Indeed, the greater policy focus on financial stability has resulted in a global regulatory reform agenda endorsed by the Group of 20 (G-20) leaders and a significant demand for financial and economic data to support the monitoring of the risks and vulnerabilities in the system. Even though not the cause of the crisis, a lack of data hampered such monitoring in advance of the GFC. While the statistical world has subsequently reacted with a number of initiatives, including the G-20 Data Gaps Initiative (DGI) and others (6), the inevitable question arises as to what more needs to be done.

Against this background, the paper has two main aims. First, based on the experience of policy makers and analysts, the paper takes a holistic view of data actually used for FSA, drawing out the common international threads in analysis. To our knowledge this is the first paper to take such an approach at the international level. Second, the paper sets out the data gaps identified by policy makers that remain to be filled particularly with regard to shadow banking, capital flows, corporate borrowing, as well as a demand for more granular data, and more broadly how official statisticians can adapt their conceptual advice to better meet the needs of FSA including with regard to remaining maturity and foreign currency data.

Inevitably, before embarking on any data enhancements statisticians would need to address the challenges, costs, and trade-offs of implementation, as well as discuss priorities with users. This paper does not specifically address these issues as it is mainly focused on setting out user needs and exploring how they could be addressed. Further, the paper recognizes that users' data needs and the priorities attached to these needs will differ based on the country circumstances.

The paper starts with a discussion of what is financial stability and how it is analyzed, identifies the datasets typically used in FSA, addresses the data gaps that have emerged, and sets out

(4) Laeven and Valencia estimated median output losses of banking crises between 1970 and 2011 at 23.2 percent of GDP. They updated the work in 2012, finding that for the crises that started in 2007 onwards, the median output loss reached 25 percent for those countries that experienced a banking crisis. See Laeven and Valencia (2013).
(5) There has also been a consequential increase in resources allocated to FSA.
(6) For instance, see Heath (2015).

proposals for how economic and financial (macroeconomic) statistical manuals can better meet financial stability data needs without undermining their conceptual framework [7].

2. Financial stability policy and analysis

The August 2016 International Monetary Fund (IMF)/Financial Stability Board (FSB)/Bank for International Settlements (BIS) report to the G-20 stated that macro-prudential policy is aimed at avoiding 'the risk of widespread disruption to the provision of financial services that is caused by an impairment of all or parts of the financial system, and which can cause serious negative consequences for the real economy.' [8]

While the above definition and observation might be considered as applying to financial stability policy more broadly, the latter appears to have a wider remit than macro-prudential policy. As the then Head of Financial Stability at the Reserve Bank of Australia observed in 2013, macro-prudential policy is subsumed in the broader financial stability policy framework-prudential supervision, market conduct regulation, consumer protection, land supply, tax system, and exchange rate regime [9]. To this list could be added the functioning of market infrastructure such as clearing houses, and corporate governance and investor protection. [10]

Indeed, beyond avoiding financial crises, it would appear that financial stability policy is concerned not only with the risk of widespread disruption to the provision of financial services, but with the efficiency of those services on an on-going basis, helping to identify where policy actions might improve efficiency and reduce systemic risk. For instance, the Bank of Korea considers that financial stability is 'a condition in which the financial system works smoothly with all of its key components satisfactorily performing their roles: financial institutions carrying out their financial intermediary functions, market participants maintaining a high level of confidence in their financial market, and the financial infrastructure being well developed.' [11] Indeed, while financial stability does not have a universally accepted definition, there seems to be a broad consensus that financial stability refers to the smooth functioning of the key elements that make up the financial system. [12] [13]

What has developed, particularly since the GFC, has been a greater focus on strengthening financial stability policy and analysis, with consequential data demands. Further, new governance arrangements have been established bringing together macro-prudential analysis, micro prudential analysis, and other aspects of FSA, in a holistic approach to FSA.

[7] The major statistical manuals such as the System of National Accounts (SNA) and the Balance of Payments Manual (BPM) are updated around every 16 years, with a long preparatory period in order to debate and agree how best to address the new issues arising. The SNA and BPM were last updated in 2008 and 2009, respectively.
[8] International Monetary Fund et al (2016).
[9] See Ellis (2013).
[10] See International Monetary Fund (2016d).
[11] See Bank of Korea (2016).
[12] Garry Schinasi (2004) defined financial stability as follows: 'A financial system is in a range of stability whenever it is capable of facilitating (rather than impeding) the performance of an economy, and of dissipating financial imbalances that arise endogenously or as a result of significant adverse and unanticipated events.'
[13] Osterloo and de Haan (2004).

Relationship between micro-prudential, macro-prudential and macro-economic policy and analysis

Micro-prudential supervision ([14]) and market conduct regulation, etc., existed well before the GFC and as indeed, in some economies, did macro-prudential policy. But the GFC demonstrated that micro-prudential policy is necessary but not sufficient to ensure financial stability; hence the emergence of macro-prudential policy to complement micro-prudential policy with a more systemic perspective. As a former First Deputy Managing Director of the IMF observed, 'macro-prudential analysis looks at the intersection of the real economy and the financial sector, providing a bird's eye view of the entire system instead of focusing on individual instruments and individual institutions.' ([15])

Macroeconomic developments and policy are directly relevant for financial stability policy as economic developments have an impact on the financial stability risks facing an economy, and vice versa. For instance, a rapidly growing economy might encourage excessive credit growth, while an economy with a weak economic position might have increasing levels of nonperforming loans. Therefore, traditional macroeconomic indicators are relevant for FSA. Nonetheless, as the August 2016 IMF/FSB/BIS report to G-20 observes, the difference between macroeconomic and macro-prudential policy is that 'rather than managing the level and composition of aggregate demand or the business cycle, macro-prudential policy aims to strengthen the financial system's defenses in the face of economic and financial shocks.'

The holistic approach of FSA helps 'straddle the gap' between micro and macro analysis, as it became increasingly obvious that micro-prudential analysts need macro data and macro-prudential analysts need granular micro data, and they potentially benefit from each other's insights in order to identify emerging systemic risks and vulnerabilities. ([16]) As the IMF Financial Surveillance Strategy published in 2012 ([17]) emphasized: there is a need to understand 'the interactions between macro-prudential, macroeconomic, and micro-prudential policies, as well as potential costs and side effects.' In a similar vein, there have been calls for a breaking of the silos between macro-economists and financial sector specialists. ([18])

Governance arrangements that have emerged for financial stability assessment

There has been a significant growth of governance arrangements around financial stability policy in recent years.

First, there have been enhanced institutional arrangements at both the domestic and international levels: the allocation of financial stability responsibility within domestic economies, often to the central bank; the creation of the FSB and the convening of G-20 leaders annually in support of economic and financial cooperation at the international level; ([19]) and enhanced regulation, notably of banks, both nationally and internationally.

([14]) Micro-prudential supervision has focused particularly, but far from exclusively, on banking supervision, and developed in earnest from the mid-1970s with the Basel Committee on Banking Supervision established in 1975 to strengthen the regulation, supervision, and practices of individual banks worldwide.
([15]) See Lipsky (2013).
([16]) The regulatory reforms that increase transparency also help increase the amount of data available for FSA.
([17]) See International Monetary Fund (2012b).
([18]) See for example, Frécaut (2016).
([19]) The FSB was established in 2009, as the successor to the Financial Stability Forum, itself established in 1999.

Second, central banks, in addition to their traditional focus on monetary policy and price stability, have increased their focus on financial stability. In some countries, financial stability committees have been established, perhaps involving multiple agencies including those with fiscal and regulatory policy responsibilities, to keep an ever-watchful eye on these risks. The allocation of financial stability responsibilities has facilitated the publication of financial stability reports on a regular, usually semi-annual, frequency to inform the public on the risks to the financial system and economy more broadly. At the international level, the IMF produces a semi-annual Global Financial Stability Report (GFSR) as a contribution to global financial stability and sustained economic growth of member countries. ([20]) Clearly, meaningful data are an essential feature of financial stability reports at both national and international levels.

Further, at the national level, financial stability departments have been created and strengthened to support this enhanced analysis. At the international level, the FSB coordinates the work of national financial authorities and international standard setting bodies in order to develop and promote the implementation of effective regulatory, supervisory and other financial sector policies; ([21]) the IMF mandates financial system stability assessments under the Financial Sector Assessment Program (FSAP) every five years for economies with globally systemically important financial sectors (See Box 1); ([22]) while the BIS's Financial Stability Institute assists supervisors around the world in improving and strengthening their financial systems. ([23])

Stress testing

Stress tests have increasingly become integral to FSA as a method of testing the resilience of the financial sector. As noted in the foreword to the IMF book *A Guide to IMF Stress Testing: Methods and Models*, ([24]) 'the GFC has placed a spotlight on the stress testing of financial systems.' These tests typically take extreme but plausible stress scenarios and test the extent to which different elements of the financial system would be able to cope and continue to provide financial services. Many central banks and/or regulatory agencies run stress tests, and some publish the results.

The balance sheet approach ([25]) is a common approach to stress testing, drawing on balance sheet data of deposit-takers and other financial institutions ([26]). A second approach is the market price-based approach that uses market data and statistical techniques to capture interlinkages between institutions, markets, or sources of risk. ([27]) Stress tests can be top-down, conducted by the national authorities or IMF staff (typically in FSAPs) using bank-by-bank data and applying a consistent methodology and assumptions, or bottom-up, conducted by individual financial institutions using their own internal data and models based on a common scenario. ([28]) As stress tests are data intensive, particularly for the balance-sheet approach, with granular data often

([20]) The GFSR was first published in March 2002.
([21]) In addition to its role in coordinating regulatory developments in the financial sector, the FSB has a focus on promoting a more resilient global financial system.
([22]) Presently 29 economies are classified as having systemically important financial sectors. See https://www.imf.org/external/np/fsap/mandatoryfsap.htm.
([23]) The Financial Stability Institute was established in 1999 and has an annual program of activities. See https://www.bis.org/fsi/activities.htm
([24]) Li L. Ong (editor) (2014).
([25]) The terminology 'balance sheet approach' used for stress testing means something different to the 'Balance Sheet Approach' terminology used in economic analysis, as the former is more focused on individual institutions while the latter covers sectors within the whole economy.
([26]) See Schneider and Schumacher (2014).
([27]) For information on liquidity stress testing approaches, see Jobst et al (2017).
([28]) See International Monetary Fund (2012c) for seven 'best practice' principles for stress testing. Box 1 discusses and compares the balance sheet- and market price-based approaches.

Box 1: Financial Sector Assessment Program (FSAP)

The most comprehensive international approach to assessing the financial sector in individual economies is the IMF's FSAP. The goal of FSAP assessments is twofold: to gauge the stability and soundness of the financial sector, and to assess its potential contribution to growth and development. ([a])

This assessment examines three key components: ([b])

- **the soundness of banks and other financial institutions, including stress tests;**
- **the quality of financial market oversight in banking and, if appropriate, insurance and securities; and**
- **the ability of supervisors, policy makers, and financial safety nets to respond effectively in case of a crisis.**

The data requirements are focused on the first bullet — as the second and third bullets are clearly of a qualitative and judgmental nature, and include Financial Soundness Indicators (FSIs). ([c])

As mentioned above, the IMF has identified economies that from a global perspective have systemically important financial sectors. Such economies are required to have mandatory FSAPs every five years with the intent of better safeguarding global financial stability. This identification is based on an assessment of the size and interconnectedness of an economy's financial sector, using Gross Domestic Product (GDP) data on a purchasing power parity basis, and data from the BIS's Locational International Banking Statistics (IBS) and the IMF's Coordinated Portfolio Investment Survey (CPIS). ([d])

The FSAP reports for individual economies are available at https://www.imf.org/external/np/fsap/fssa.aspx.

([a]) See International Monetary Fund (2016e).
([b]) See International Monetary Fund (2014a).
([c]) Typically, an FSAP report will also include a table of key macroeconomic indicators as macroeconomic developments are highly relevant for FSA.
([d]) Equity return data from a private data provider are used to construct cross-country equity return correlations.

needed, it is important to have well maintained and consistent databases with appropriate access available for those conducting the stress tests. ([29])

There remains room for improvement to use stress testing as a tool for macro-prudential risk assessment going beyond its traditional use for micro-prudential supervision. Demekas points out that very few stress testing models focus on, and measure correctly, the resilience of the financial system as a whole and its ability to continue providing financial intermediation services under stress in a way that makes the results readily actionable for individual banks and their supervisors. ([30])

([29]) For instance, see the Opening remarks of IMF Deputy Managing Director, Tao Zhang at the IMF-LSE Symposium on 'Macroprudential Stress Tests and Policies: A Framework,' December 15, 2016. (Zhang (2016).
([30]) See Demekas (2015).

Financial stability policy ([31])

The policy applications of FSA, and particularly the tools used to meet financial stability needs, are still developing. ([32]) While there is long experience of the use of prudential regulation and micro-prudential policy tools, the same is not true for macro-prudential policy tools. Nonetheless based on international experience, the August 2016 BIS/FSB/IMF G-20 report discussed the various tools in use and their application. These include:

- broad-based capital tools (e.g., dynamic provisioning, countercyclical capital buffers, and time-varying leverage ratio caps);
- sectoral capital and asset-side tools (e.g., foreign currency loans to corporates, caps on loan-to-value (LTV), debt-service-to-income (DSTI), or loan-to-income (LTI) ratios); and
- liquidity-related tools (e.g., liquidity coverage ratio (potentially calibrated by currency)), as well as tools to contain maturity mismatch (such as core funding ratios), price-based tools (such as a levy on volatile funding), and caps such as on the loan-to-deposit ratio.

Among other policy tools have been capital surcharges for global and domestic systemically important banks (G-SIBs and D-SIBs) and, scheduled from 2022, for global systemically important insurers (G-SIIs), and increases in risk-weights and large exposure/concentration limits. Further, interbank exposure limits, and foreign and domestic currency reserve requirements are being used as policy tools to lower macro-prudential risks. ([33])

Policies have also been developed to address potential financial stability risks arising from non-bank activities, such as central clearing of over-the-counter derivatives, ([34]) and in market infrastructure, such as ensuring the resilience of central counter-parties (e.g., margining requirements and liquidity resources). ([35])

Given that there is relatively little experience with financial stability policy tools, data are needed to assess the impact of the use of these tools, and indeed of financial regulation more broadly, on economic and financial activities. This 'ex-post' policy evaluation is an important field of FSA. ([36])

It is also important to note that beyond micro, and macro-prudential policies, other types of policies can affect financial stability, such as the tax system with incentives for debt finance, ([37]) and and housing policies, affecting the supply and demand for real estate, and consumer protection affecting lending standards. Further there is an on-going debate as to whether interest rates could be used to meet financial stability policy needs ([38]), given that monetary

([31]) Since 2008, the G-20 has promoted a comprehensive program of regulatory reforms designed to increase the resilience of the global financial system while preserving its open and integrated structure. These reforms and their implementation are presented in the annual FSB report on *Implementation and Effects of the G20 Financial Regulatory Reforms*. Financial Stability Board (2016b).
([32]) For an example of work to develop a macro-prudential policy framework see the European Systemic Risk Board (2016)
([33]) For instance, see Cerutti et al (2016).
([34]) In Europe, the EU has introduced the European Market Infrastructure Regulation on derivatives, central counterparties and trade repositories, which imposes requirements to improve transparency and reduce the risks associated with the derivatives market.
([35]) For instance, see Financial Stability Board et al (2015).
([36]) A description of the usage of macro-prudential policies for a large, diverse sample of 119 countries over the 2000-13 period, and the relationships between the use of these policies and developments in credit and housing markets is set out Cerutti et al (2015). See also Macroeconomic Assessment Group (established by the Financial Stability Board and the Basel Committee on Banking Supervision) (2010), and Financial Stability Board (2017).
([37]) For instance, see International Monetary Fund (2016f).
([38]) For instance, see Bernanke (2015).

policy and financial stability objectives are interrelated, ([39]) and regarding the relationship between price and financial stability. ([40]) Also, movements in the exchange rate can have domestic financial stability implications. ([41]) Indeed, the objectives of capital flow measures — designed to limit capital flows by influencing their size or composition, can overlap with macro-prudential policies, if the latter are designed to limit systemic risks by limiting capital flows. ([42])

To promote information sharing, the IMF, in consultation with the FSB and the BIS, is compiling a publicly available macro-prudential policy database. ([43])

3. Data used in financial stability policy and analysis

In drafting this paper, the authors examined a cross-section of financial stability reports and IMFs FSAP reports to identify the datasets used for FSA. ([44]) This section sets out the main 'story' lines that emerge from this research. It is important to realize that FSA is constrained to available data and this has led policy makers to make a number of requests to official statisticians to expand available information. These requests are discussed in the next section. A more detailed discussion of the data used in FSA is provided in Appendix 1.

The complexity of modern economies is such that the potential risks and vulnerabilities are many and varied. They can also differ according to the nature of the economy, its financial system, and over time. Consequently, FSA has a very large demand for, and access to, meaningful data. ([45]) Having said this, it is important to recognize that not all aspects of FSA involve data as issues such as the strength of the regulatory framework and of the 'safety net' also arise.

At the core, the datasets used for FSA appear to be those that have the purpose of:

- monitoring the soundness and efficiency of the financial system (institutions and markets), and the growth of credit to and indebtedness of non-financial sectors;
- identifying pockets of vulnerability emerging within the financial system;

([39]) As noted in the International Monetary Fund (2015c), the GFC was a reminder that price stability is not sufficient for financial stability. Further the paper considered that generally monetary policy should not be altered to contain financial stability risks but that 'the door should remain open as our knowledge of the relationship between monetary policy and financial risks evolves and circumstances change.'

([40]) For instance, see White (2006).

([41]) For instance, Philip Turner gives the example of movements in the exchange rate being relevant for financial stability because they have wealth effects and affect risk-taking, both by banks and in capital markets. See Turner (2016).

([42]) An example of where a macro-prudential measure might have impacted capital flows was in Korea. In 2010 the Korea authorities placed a cap on the ratio of foreign exchange (FX) derivatives positions to curb banks' building up of excessive FX derivatives positions which tended to be financed by short-term borrowing. According to the FSAP report on Korea (2014, paragraph 31) (International Monetary Fund (2014b) this measure appeared to have contributed to a shift away from short-term FX funding and may have caused interbank capital flows into Korea to become less sensitive to global financial conditions.

([43]) The development of this database was welcomed by the G-20 Finance Ministers and Central Bank Governors in their March 2017 Communique. See G-20 Finance Ministers and Central Bank Governors (2017).

([44]) The authors undertook a review of 23 recent IMF FSAPs: Algeria, Argentina, Barbados, Belarus, Canada, Comoros, Republic of Congo, Denmark, El Salvador, Germany, Hong Kong SAR, Ireland, Montenegro, Morocco, Peru, Republic of Korea, Russia, Samoa, South Africa, Switzerland, Tajikistan, United Kingdom, and United States, and financial stability reports of Australia, People's Republic of China, ECB, IMF, Republic of Korea, Japan, Mexico, United Kingdom, and United States. The detailed list of the datasets identified by this review is available on request from the authors.

([45]) As noted in Office of Financial Research (OFR) (2015) financial data must have three attributes to be useful: (1) sufficient scope (comprehensive and granular), (2) high quality (comprehensive, accurate, timely), and (3) accessible (shared and secured). Further, identification of data gaps begins by deciding on the most important questions related to potential vulnerabilities, the analytical framework to answer them, and the data needed to quantify that framework.

- assessing the sustainability and vulnerability of the non-financial sectors debt positions; the potential impact on FSA of the growth in asset prices; and the financial links within and across economies that might cause shocks to permeate within the domestic economy; and
- testing for potential vulnerabilities in the system through stress tests and assessing the impact of regulatory changes and policy actions.

Against this background, the research reveals both a common frame of analysis to address the first three bullets above and cross-cutting issues regarding the use of time series/cross sectional data and residence-based/cross-border consolidated data. The rest of this section explores these topics.

Also from the research undertaken, the degree of sophistication and depth of markets, the range and number of financial institutions, and the extent of interconnectedness both domestically and cross-border, impacts the scope of data monitored for FSA by countries of different economic development. But the main impression arising from the research was of the similarities of analysis and commonalities of data monitored (e.g., the structure of the financial sector, the relevance of credit and debt statistics, the need to monitor asset prices, etc.). Some datasets that are particularly relevant for developing economies are highlighted below.

Framework of analysis

Macro-economic analysis is focused on economic behavior among resident entities and between resident entities and nonresident entities, within well-defined frameworks of analysis, such as the national accounts framework, and with well-established indicators of economic performance, such as growth, inflation, employment, etc. On the other hand, FSA is focused on potential risks and vulnerabilities to the system without a firmly established theoretical framework. Nonetheless, a common frame of analysis emerges from the research the authors have undertaken broadly consistent with the three interlocking objectives set out in the August 2016 IMF/FSB/BIS report to G-20. ([46]) These objectives were:

(1) increasing the resilience of the financial system to aggregate shocks by building and releasing buffers that help maintain the ability of the financial system to function effectively, even under adverse conditions;

(2) containing the build-up of systemic vulnerabilities over time by reducing pro-cyclical feedback between asset prices and credit and containing unsustainable increases in leverage, debt stocks, and volatile funding; and

(3) controlling structural vulnerabilities within the financial system that arise through interlinkages, common exposures, and the critical role of individual intermediaries in key markets that can render individual institutions 'too-big-to-fail.'

Figure 1 provides a schematic overview of the key data needs that emerged from the authors' research.

Resilience of the financial system: Data are used to undertake a holistic review of the financial system that is common to financial stability reports. Such reviews encompass not just deposit-takers, but also other financial institutions, and the relationships between them; the structure of the system and concentration measures — not least for assessing the potential impact

([46]) The paper recognizes that these objectives are inter-related, or 'inter-locking' as described in the IMF/FSB/BIS report to G-20. For instance, the growth of credit impacts the soundness of the banking system, while borrowing through debt securities abroad affects financial interconnectedness between domestic sectors and the rest of the world.

on competition; the markets in which these institutions, (47) and other debtors and creditors, operate; the infrastructure of the financial system, such as clearing houses; and, particularly for developing countries, financial inclusion.

For deposit-takers, data collected and compiled to support prudential supervision of individual banking institutions remain essential.

Figure 1: Key Data used for FSA

1) Increasing the resilience of the financial system		
Key data	**Financial institutions** • Deposit-takers • Non-bank financial institutions (With domestic/foreign and state/ private breakdowns)	**Financial markets:** • Money markets, securities and derivatives markets • Interest rates and FX rates • Infrastructure and financial inclusion

2) Containing the build-up of systemic vulnerabilities: credit, debt and asset prices			
Key data	**Credit-related:** • Credit growth • Type of credit • Connected, concentrated, directed, lending	**Debt-related:** • Outstanding debt levels • Leverage ratios • Non-financial corporations Households and Governments	**Asset prices:** • Real estate • Equity and bond price

3) Structural vulnerabilities within the financial system: financial interconnections and spillovers	
Key data	**Domestic & cross-border:** • Financial interconnections within economies and across borders • Capital flows • Spillover analysis • International environment

Systemic vulnerabilities arising from credit and debt, and asset prices (including leverage, currency, and liquidity): Data on credit and debt are generally considered central to FSA, as research suggests that fast growth in credit can be an early warning indicator of financial crisis, (48) while liquidity and solvency problems can arise with high levels of debt relative to income and wealth. As customers of the financial sector, data on non-financial corporations (NFC) and households (HH) are used to identify potential problems in these sectors that might cause problems for the financial sector, and vice versa. Due to the inherent risks, data on connected, concentrated and/ or government directed lending are monitored, while the growth of credit through the FinTech industry is beginning to be assessed where relevant.

Among asset prices, real estate prices, both for residential and commercial property, equity and bond prices, as well as for land are closely monitored because fluctuations in prices affect their use as collateral, directly impact financial wealth and, indirectly impact the economy through the effect on consumer and corporate confidence. (49) There is also growing interest in volatility measures so as to understand better the uncertainties/risk in financial markets.

(47) Including market liquidity indicators.
(48) For instance, see Drehmann (2013).
(49) When households can borrow against their residential real estate collateral, rising real estate prices can lead to higher borrowing and associated consumption.

Structural vulnerabilities within the financial system arising from financial interconnections and spillovers, both domestic and cross-border: There is an increasing use of data that supports an understanding of financial interconnections and spillovers among individual financial institutions, ([50]) the financial and domestic non-financial sectors, and, between each sector and the rest of the world. This is often the most complex area of FSA in that financial connections between different sectors are complicated by second or third round inter-linkages — who lends to the entity funding my position, and by common exposures — I have no relationship to you except the fact that we are both exposed to the same kind of risks. In other words, negative spillovers can arise from both direct and indirect exposures. Indeed, vulnerabilities can arise from the complexity of increased interconnectedness as well as from the use of complex, and often opaque, financial instruments.

Also, policies of major economies can potentially have spillover implications for the domestic economy, perhaps through unexpected channels. ([51]) Given this, national FSA typically monitors data that helps assess developments in the international environment and the potential impact of capital flows.

Cross-sectional and time-series data

The literature suggests that it is important to distinguish between the cross-sectional and time dimension aspects of FSA. The August 2016 IMF/FSB/BIS report to G-20 picks up on this distinction in noting that 'systemic risk is generally recognized as having two dimensions: vulnerabilities related to the build-up of risks over time ('time dimension'), and vulnerabilities from interconnectedness and the associated distribution of risk within the financial system at any given point in time ('cross-sectional' or 'structural' dimension).' This has an important implication for statistical work in that traditionally economic and financial statistics have been focused on the time dimension rather than cross-sectoral dimension, although the increasing analytical focus and use of position data is beginning to give more emphasis to the latter. ([52]) ([53])

Cross-border consolidated- and residence-based data

The primary interest of the authorities when analyzing financial stability is on the impact on residents ([54]) and the domestic economy, as the ultimate goal of domestic policy makers is to protect the domestic economy. Therefore, the majority of the datasets used by domestic policy makers for FSA are residence-based. This is primarily true for data on credit and debt, financial markets, interest rates, financial market infrastructure and inclusion, and domestic and cross-border inter-connectedness.

Nonetheless, risks to financial stability may come from the activities of domestically-owned individual institutions in foreign markets - the involvement of European banks in the sub-prime market prior to the GFC being a prime example. This implies that cross-border consolidated

([50]) For instance, see Alexandrova-Kabadjova et al (2012).
([51]) For instance, see International Monetary Fund (2012).
([52]) Micro-prudential banking supervision data are collected and analyzed by individual institution often involving the inclusion of foreign affiliates and activities that are ancillary to banking business, focusing on the latest end-period. As macroeconomic statistics have a residence and traditionally time-series focus, supervisory data have developed, and remain largely, separate from, macroeconomic statistics.
([53]) See Heath (2015), page 18.
([54]) Residents cover those entities and households that are located — have their center of economic interest, in the domestic economy. This is regardless of ownership or nationality. So, a foreign-owned deposit-taker located in the domestic economy is resident of that economy.

statistics of domestically-owned individual institutions (incorporating foreign branches and subsidiaries) located in an economy are also relevant for FSA.

Indeed, data for deposit-takers is typically analyzed both for micro- and macro-prudential purposes on a cross-border consolidated basis. ([55]) For instance, the scope of application under the Basel standards for banking supervision provides that to the greatest extent possible, all banking and other relevant financial activities (both regulated and unregulated) conducted within a group containing an internationally active bank will be captured through consolidation. ([56]) Similarly data for FSIs for deposit-takers are typically compiled and analyzed using one of a range of consolidation approaches including those based on the Basel standards.

In this context, there has been the longstanding use of BIS IBS data on a cross-border consolidated basis that captures the nationality of international banking activities, including where the ultimate risk lies. This is because, as noted by Tissot, 'the IBS consolidated data yield a comprehensive picture of the national lenders' risk exposures, in particular to country risks,' ([57]) and so can help identify potential risks and vulnerabilities to the domestic economy arising from the foreign activities of domestically headquartered international banks.

Also, the activities of domestic subsidiaries and branches of foreign deposit-takers can be significant in the host market but relatively small within the context of the consolidated foreign banking group. In such circumstances, the behavior of these foreign affiliates can be affected as much, if not more, by activity, and decisions made, outside as inside the host market — for instance a funding shock to the parent bank or economy. In addition, vulnerabilities of subsidiaries in foreign markets may not be apparent in the home country's residence-based data. These insights were one reason why the recent enhancements to the locational BIS IBS datasets included more granular information by nationality of the reporting bank. ([58]) As McGuire and von Peter noted, 'in any particular host country, a long or short net cross-border position in a particular currency booked by the offices of foreign banks there may be offset or hedged elsewhere on those banks' global balance sheet.' ([59])

Nonetheless, deposit-takers residence-based data are used for FSA, not least in terms of analyzing domestic interconnectedness and the relationship between the domestic lending and funding sides of the balance sheet. Indeed, for foreign-owned deposit-takers, the extent to which domestic lending is matched by domestic retail deposits, provides insights into the stability of their lending activity within the economy. ([60])

Data for non-bank financial institutions (NBFI) might be analyzed on a cross-border consolidated basis, if the relevant data are available. However, residence-based data are often the only data available. For instance, the FSB annual global shadow banking monitoring report draws heavily on national financial accounts data although it also includes estimates of shadow banking that excludes NBFI that are part of a regulated banking group.

Further, while residence-based data are the basis of analyzing debt and credit, FSA is also increasingly interested in data on borrowing by subsidiaries of resident entities located abroad. ([61]) As was seen in the GFC, many countries, particularly emerging market economies (EME), found that borrowing by foreign subsidiaries of domestic NFCs came onto the domestic

([55]) The concept of consolidation is not as clear as that of residence. Indeed, there are many 'varieties' of consolidation. See Inter-Agency Group on Economic and Financial Statistics (2015).
([56]) See Basel Committee on Banking Supervision (2006) .
([57]) Tissot (2016).
([58]) See also Fender and McGuire (2010) and Cecchetti *et al* (2010).
([59]) See McGuire and von Peter (2009).
([60]) For instance, see McGuire and von Peter (2016).
([61]) See the chart 'Credit to non-banks including offshore issuance' on page 21 of Tissot (2016).

balance sheet in the crisis. Even outside of a crisis, significant recent U.S dollar borrowing by foreign subsidiaries of emerging market NFCs, often through issuance of debt securities in foreign markets (offshore borrowing), has raised questions of the extent to which they are facing foreign currency risks that might in turn affect the domestic parent (see also the next section under 'corporate borrowing'). ([62])

4. What data gaps have emerged?

The previous section discussed the data used by national and international authorities in their FSA. While acknowledging the progress that statisticians have made in closing data gaps (see Appendix 2), policy makers at the national and international level have continued to draw attention to specific gaps that they consider need to be addressed. Drawing on these calls, this section sets out the most significant of these needs and suggests a way forward for each.

Before addressing the specific gaps to be filled, some more general observations about the data needed for FSA can be made.

First, from a review of the data used (see Appendix 1) it is evident that many of the needed data are already available to the financial institutions and authorities, although coverage varies across countries. Since the GFC, statisticians have taken a number of initiatives to expand the availability of data for FSA as circumstances have demanded. There have been increased efforts in several fora and significant progress has been made in closing the gaps identified, notably at the international level through the G-20 DGI ([63]) and the IMF's Special Data Dissemination Standard Plus (SDDS Plus). ([64]) While these international initiatives are not relevant for all economies, implementation by countries for which they are relevant would support FSA in a significant way. In an interconnected global economy, the benefits of implementing such initiatives not only accrue for the implementing economy but also for the broader international community.

In particular, the G-20 DGI has promoted work to close gaps and strengthen datasets that support both the monitoring of risk in the financial sector (for example, FSIs, securities issuance, and credit default swaps data), and the monitoring of domestic and cross border risks and vulnerabilities (such as through sectoral balance sheets and the major cross-border surveys of the BIS and IMF) (see Appendix 2 for more details).

Despite these improvements there is often a need for further steps (i.e., enhancements of several datasets in terms of coverage, scope, quality, consistency) including some new data collection initiatives (e.g., collection of new data such as the granular dataset on globally systemically important financial institutions).

In addition, there does remain a question as to whether all the available official data are being fully used to meet user needs. Official statisticians may need to do a better job in communicating to policy makers the possibilities of available data. The user may not be aware that the data they need are available either directly, or indirectly through manipulating available data, or that

([62]) Gruić and Wooldridge (2014).
([63]) For an overview of the progress made under the G-20 DGI, see Annex 3 of, Financial Stability Board and International Monetary Fund (2015).
([64]) The Special Data Dissemination Standard (SDDS) and enhanced-General Data Dissemination System (e-GDDS) can also support the work of FSA. See http://dsbb.imf.org.

available data have an informational content that is of relevance to FSA. [65] One attempt to address this 'publicity' issue has been through the Principal Global Indicators (PGI) website set up by the Inter-Agency Group on Economic and Financial Statistics (IAG). [66]

Finally, financial stability policy makers and analysts increasingly use market/private sector data as well as official statistics in their work. This is particularly true for market-related data and high-frequency data. Private sector data can be more timely if less comprehensive. But policy makers often want early indications of emerging risks and vulnerabilities. In this regard, there is also growing interest in big data as they can provide timely data at a high speed. [67] In other words, official statistics do not, and do not need to try to, meet all the FSA data needs.

Specific data gaps

SHADOW BANKING

While the banking sector has traditionally been at the heart of the financial sector, the GFC demonstrated the key role shadow banking financial institutions and markets play in credit and maturity transformation, performing bank-like activities. [68] However, unlike deposit takers, these institutions are usually not strictly regulated and supervised, and have no access to deposit insurance, to the rediscount operations, or to the last resort credit lines of central banks. [69]

As has been emphasized by the FSB, the IMF, and other international and national authorities, there is a need for data that identifies and estimates the scale of shadow banking activity, provides a better understanding of both the entities involved and the risks they are facing, and can indicate potential vulnerabilities to the financial system arising from their activities. In doing so, the relationship with the banking industry can be assessed along with the risks to financial stability arising from shadow banking activities.

Experience during the GFC has shown that risks to financial stability may emerge in these institutions and markets from high leverage, maturity mismatches, and/or illiquidity, materialization of which could spread through the whole financial system. An example was the experience of money market funds (MMFs). [70] While not typically leveraged institutions, the GFC illustrated how rapidly the risks and vulnerabilities of MMFs can be transmitted to the rest of the financial system when investors start withdrawing their funds on a significant scale: MMFs liquidated financial assets so helping to depress market prices and scaled back their wholesale funding of deposit-takers, particularly to European banks. [71]

This has led policy makers to adopt stricter regulatory oversight on shadow banking institutions and markets, including, greater disclosure on asset valuations and collateral haircuts, reforms of governance and ownership, as well as stricter oversight, regulation and limitations on collateral

[65] For example, the build-up of European investments in U.S. securities during the middle years of the 2000s was clearly evident in the Coordinated Portfolio Investment Survey data but was not picked up by analysts. See Rodriguez (2008).
[66] See Inter-Agency Group on Economic and Financial Statistics (2017). The members of the IAG are senior statisticians from the BIS, the European Central Bank (ECB), Eurostat, the IMF, the Organization for Economic Cooperation and Development (OECD), the United Nations (UN), and the World Bank.
[67] For instance, see Hammer *et al* (2017).
[68] Measures of shadow banking activity include by type of non-bank financial, and/or through the value of activity through securitization, repos, money market funds, bankers' acceptances and commercial paper. The former is more typically used but see also the 2014 Canada FSAP, page 10 (International Monetary Fund (2014c).
[69] See Farhi and Macedo Cintra (2009).
[70] MMFs invest in short-term assets and offer redemptions on demand. So traditionally MMFs have been considered an alternative to bank deposits.
[71] See International Monetary Fund (2014e).

lending. (72) The FSB has led the work at the international level, producing an annual monitoring report using available data (as well as addressing the regulatory aspects of shadow banking). (73)

Unlike the detailed information available to the supervisors and central banks for the monitoring of the banking sector, data on shadow banking has generally been lacking due to the heterogeneous nature of the institutions, lack of regulatory oversight, a previous lack of recognition of the systemic importance of shadow banking, and a lack of a consistent definition of shadow banking.

In its May 2016 Financial Stability Review (FSR), (74) the ECB pointed out the limited availability of disaggregated data needed for FSA on assets, liabilities, capital, and profitability of financial institutions other than deposit-takers and insurance companies.

Eichner, Kohn and Palumbo (75) pointed out that the growth of maturity transformation outside the traditional banking sector contributed to the severity of the financial crisis but was not conveyed in aggregate financial statistics for the U.S. economy. (76)

In addition, there is a lack of data with regard to securities financing activities for FSA considering the reliance of shadow banking institutions on wholesale funding (such as through repo and securities lending markets). For instance, the importance of closing the data gaps in securities financing markets was pointed out by the U.S. Financial Stability Oversight Committee (FSOC) (77) (emphasizing that data are needed to assist policy makers' understanding of (1) how the repo market operates; (2) the interdependencies of institutions and participants; and (3) changes in risk characteristics, such as collateral and haircuts. (78)

At present, existing balance sheet and other relevant data are collected, in most cases, under jurisdictions' existing statistical (and regulatory) reporting requirements, with the level of granularity and frequency of reporting varying across entity types within and across jurisdictions. (79) Data gaps are particularly prominent for non-regulated entities for whom the national authorities' data collection powers often do not extend.

While the national accounts-based sectoral balance sheet and flow of funds data provide a good initial basis for assessment of the shadow banking risks, there are a number of limitations that require addressing either through methodological developments (see next section) or new data collections. In particular, data by economic function, with more granular information on maturity and liquidity transformation, and foreign currency exposures, is needed to support the risk metrics used for assessing the extent of shadow banking risks. To have a full picture of the risks and vulnerabilities associated with NBFIs, including a thorough analysis of their cross-border linkages, cross-border consolidated data on a nationality basis are needed to complement the currently available residency-based data.

(72) See Financial Stability Board (2013).
(73) At the 2010 Summit meeting in Seoul, the G-20 Leaders tasked the FSB to undertake a major initiative to monitor the activity, and review the regulation, of shadow banks. See G-20 Leaders Summit (2010).
(74) See European Central Bank (2016).
(75) See Eichner *et al* (2010).
(76) The IMF, in GFSR October 2016, also highlights the increasing importance of non-bank financial institutions in the transmission of monetary policy as the risk-taking channel of monetary policy has gained importance and asset managers have assumed a greater role in financial intermediation. In this context, IMF points out that given the growth of the non-bank financial sector, the information contained in the balance sheets of non-banks is potentially at least as useful as traditional measures of monetary aggregates.
(77) US Financial Stability Oversight Committee (2016).
(78) In 2015, the European Union adopted a Securities Financing Transaction (SFT) regulation that is intended to increase the transparency of these markets. Entities engaged in this activity will need to provide details of their transactions to an approved EU trade repository.
(79) See Financial Stability Board (2016a).

The FSB sees the need for granular data on shadow banking entities on an economic functions basis, inter alia covering leverage, liquidity, and maturity transformation activities, currency mismatches, and credit intermediation activities. (80) In addition, the FSB is working to develop a regular flow of data on securities financing markets at the national and global levels by end-2018, that will shed light into the size, composition, pricing, and risk profile of these markets. (81)

Suggested way forward

To contribute to the global efforts to improve the availability of data on the shadow banking sector, it is important for statisticians, regulators, and other users to share experiences in compiling and analyzing shadow banking data, including on ways to ensure comprehensive coverage and avoid duplication of effort. In addition, frequent dissemination of data would facilitate the timely assessment of the shadow banking system and its linkages with the rest of the financial system, and hence provide a better assessment of the systemic risks associated with shadow banking institutions and markets.

The FSB's efforts to improve the availability of information are key (82) and to this end the FSB has set up a Shadow Banking Experts Group that shares national and regional experiences in compiling and analyzing shadow banking data in the context of the FSB's annual global shadow banking monitoring report. Further, the FSB-led work on securities financing markets will provide important information on markets in which shadow banking institutions operate. Also, the IAG Working Group on Institutional Sector Accounts is currently working on better capturing shadow banking activity using macro-economic based data, by exploring the possibility of capturing more granular sub-sectoral breakdowns and instruments (see next section) for the non-bank financial sector. All this work is endorsed by the second phase of the DGI (DGI-2) recommendation 5 on shadow banking.

ASSESSMENT OF CAPITAL FLOWS

During the past years, there has been an increased policy interest in the financial stability policy implications of large swings in international capital flows. While the freer flow of capital is considered to have significant benefits for domestic economies including by enhancing efficiency, promoting financial sector competitiveness, and facilitating productive investment and consumption smoothing, the potential risks associated with the swings in capital flows need to be closely assessed as financial interconnectedness associated with greater capital flows can exacerbate the transmission and spillover of shocks between economies. (83)

In 2016, the BIS and IMF reported to the G-20 their assessments of the effects of capital flow volatility with a particular emphasis on data needs. (84) Both the IMF and the BIS recognized the Balance of Payments (BoP) as a key source of information on cross-border capital flows but

(80) See Financial Stability Board (2015a).

(81) In the EU, in 2016 a regulation requiring the reporting of securities transactions to a trade repository (TR) came into effect to improve the transparency of these markets. In the US, the OFR and Federal Reserve launched a pilot project to fill gaps in data on bilateral repo and securities markets. A pilot project on securities financing activities was also conducted in Japan.

(82) The March 18, 2017 communique of the G-20 Finance Ministers and Central Bank Governors asked 'the FSB to present by the Leaders Summit in July 2017 its assessment of the adequacy of the monitoring and policy tools available to address such risks from shadow banking.' See G-20 Finance Ministers and Central Bank Governors (2017).

(83) International Monetary Fund (2016g). Inter alia, the note points out that during 2009 and 2015, four main capital flow episodes were observed with changes in net capital flows of about 3–5 percent of the GDP on average.

(84) The IMF presented notes to the G-20, in February and June 2016. The BIS was presented in August 2016. Both notes are available at http://www.g20chn.org/English/Documents/Current/index.html. The March 18, 2017 communique of the G-20 Finance Ministers supported 'continuing to enhance the monitoring of capital flows and management of risks stemming from excessive capital flow volatility.' See G-20 Finance Ministers and Central Bank Governors (2017).

identified data gaps that need to be addressed to order to obtain a detailed picture of capital flows. These included:

- More timely BoP data (shorter reporting lag) with a higher frequency of indicators to assess capital flows.
- Identification of the direction of flows between individual countries or groups of countries, e.g., capital flows to advanced economies both from other advanced economies and emerging market and developing economies (and similarly for capital flows to emerging market and developing economies).
- Separation of the flows associated with non-financial corporate activity from those of the financial sector in the BoP. In DGI-2 the possibility of separate identification of NFCs is being investigated.
- Need for an increase in the number of countries disseminating the breakdown of direct investment data by geographical location, sector and currency. In DGI-2, inward and outward investment by country is promoted through the IMF Coordinated Direct Investment Survey (CDIS).
- Need for an increase in the number of countries disseminating the breakdown portfolio investment asset and liability data by the geographical location of debtors/creditors and by currency. Under DGI-2 sector breakdowns within the CPIS are being promoted, with a move to quarterly reporting by 2019.
- External balance sheet data on currency composition, remaining maturity of debt, and off-balance sheet items such as contingent assets and liabilities, guarantees and lines of credit, and hedging using financial derivatives. ([85])

Further, the G-20 International Financial Architecture (IFA) Working Group in their 2016 Final Report underlined the importance of enhancing capital flows and stocks data collection to better identify currency and maturity mismatches, while also explicitly supporting the recommendations in the G-20 DGI that support capital flow analysis. ([86]) Also, in late 2016 the IMF published a paper on 'Capital Flows — Review of Experience with the Institutional View' that considered improving capital flows data a priority with a focus on the timeliness, scope and granularity of balance of payments data. ([87]) Also highlighted was the importance of more detailed balance sheet (by sector and foreign currency exposure) and off-balance sheet data (such as contingent liabilities and derivative transactions). ([88])

Suggested way forward

There is considerable data available on cross-border positions and flows. A holistic review of cross-border exposures data could be undertaken by the IMF Committee on Balance of Payments Statistics (BOPCOM) to see how these data could be leveraged to best meet policy makers' needs.

([85]) Complex measurement issues are raised when attempting to measure the hedging of exposures through financial derivatives, not least because the market value of financial derivatives does not equate to exposures, and for multinational companies hedges may be undertaken at the group level.
([86]) See G-20 International Financial Architecture Working Group (2016).
([87]) The paper also called for a reduction in reporting lags and an increase in frequency. See International Monetary Fund (2016h).
([88]) With regard to contingent liabilities, Chapter 9 of the 2013 External Debt Statistics Guide (Task Force on Finance Statistics (2014)) discusses the statistical measurement of such liabilities, and includes presentation tables that present inward and outward risk transfers through contingent liabilities.

In addition, data from the G-SIBs common data template that covers these institutions exposures to national markets and sectors (see Appendix 2) could be aggregated to provide an early indication of cross-border capital flows from the largest global banks. Further the template could be used by a broader range of national authorities to collect granular information on national banking systems exposures and funding dependencies. Such data would shed light on flows to and among EMEs.

CORPORATE BORROWING

Since the GFC borrowing by NFCs, particularly in EMEs, has increased significantly, as highlighted by BIS research that has drawn on a BIS database of total credit to NFCs. ([89]) ([90]) These data show that NFC debt in the major EMEs increased from less than 60 percent of GDP in 2006 to 110 percent at end-2015. Further, the BIS research points out that any analysis of the vulnerability of EME debtors to foreign currency exposures must take account of leverage, debt maturity, and the external/domestic distinction of debt. Against this background, specific data gaps for NFCs include (i) foreign currency borrowing, particularly through off-shore affiliates; and (ii) information on corporates' risk exposures, such as maturity mismatches and foreign currency exposures (including hedging activities). ([91])

Regarding NFCs' foreign currency borrowing, the BIS international debt securities (IDS) database provides comprehensive information on total issuance of international debt securities, with currency and maturity breakdowns. But there is a lack of data on NFCs offshore foreign currency borrowing from deposit-takers as noted in the August 2016 BIS note to the G-20 IFA Working Group.

The BIS IDS database highlights the scale of off-shore borrowing in debt securities. As of September 2015, offshore borrowing accounted for a significant amount of total (including offshore) borrowing through international debt securities by Chinese (93 percent), Brazilian (53 percent), and Russian (45 percent) nationality NFCs. As noted by the Bank of England (BoE) ([92]) this offshore borrowing by NFCs with a global presence cuts across traditional residence-based data either, as BIS explains, not showing up in residence-based external debt statistics (when proceeds are not repatriated) or classified as foreign direct investment (FDI) flows (when repatriated). In either case, residence-based measures could paint an overly benign picture of vulnerabilities ([93]) and does not capture all the potential financial stability risks facing a country.

In addition to cross-border foreign currency borrowing by NFCs in international debt securities, domestic foreign currency borrowing, e.g., from domestic deposit-takers, also needs to be assessed as this form of borrowing also exposes NFCs, and through the NFCs, the domestic deposit-takers, to foreign exchange risks.

Consistent information on off-balance sheet activities, such as contingent assets and liabilities, guarantees and lines of credit, and hedging using financial derivatives also remain data gaps. While countries, at the national level, generate some information based on different data sources, through surveys or through information from derivatives exchanges, lack of consistency in the coverage and definitions used across jurisdictions does not allow for meaningful aggregation at an international level. ([94])

([89]) See Chui *et al* (2016).
([90]) Credit is measured as loan, debt security, and currency and deposit liabilities to domestic banks, all domestic sectors, and nonresidents. See Dembiermont *et al* (2013).
([91]) See page 15 in Tarashev, *et al* (2016).
([92]) See Hoggarth *et al* (2016).
([93]) See Gruić *et al* (2014).
([94]) See Bank for International Settlements and Financial Stability Board (2014).

Also, the August 2016 BIS note to the G-20 IFA Working Group points out that there is no international database on NFCs financial assets including currency and maturity composition as well as on the country and sector of their counterparty debtors.

According to the BIS, the lack of information contributes to the uncertainty about NFCs volume of foreign currency exposures, the links with the banking system, and the degree to which hedging reduces systemic risk.

Suggested way forward

Recommendation 14 of DGI-2 asks the IAG to improve the consistency and dissemination of data on NFCs' cross-border exposures, including those through foreign affiliates and intra-group funding, to better analyze the risks and vulnerabilities arising from such exposures including foreign currency mismatches.

The BIS note to the G-20 IFA Working Group suggested that in the short term combining the residence-based BoP data with the BIS IBS and IDS could shed more light on NFCs cross-border exposures and their evolution. The IAG document produced for the DGI also sets out some ideas for further work with regard to capturing NFCs cross-border exposures. ([95])

As also emphasized by the BIS, enhanced disclosures of financial hedges and derivatives positions (including detailed currency and maturity information on financial hedges and their underlying positions) on a timely basis through improved accounting standards could also contribute to the availability of consistent information on the risk exposures of NFC.

GRANULAR AND MICRO-DATA

With the nature of financial stability risks changing over time, FSA needs to be sufficiently flexible to address shifting vulnerabilities. As Jenkinson and Leonova ([96]) emphasized, given the increasing focus of financial stability on the risks to the financial system as a whole, new approaches to financial data based on the uniform representation and standardization of its key elements is becoming more important to allow for flexible data aggregation to support multiple policy objectives.

To this end, detailed and granular information is increasingly being requested to contribute to the flexibility of FSA tools. ([97]) Several statistical initiatives mentioned in this paper aim to increase the granularity of available information (e.g., the common data template on G-SIBs', data on repo and securities financing transaction, enhanced BIS IBS data, sectoral balance sheets, and the enhanced IMF CPIS). There is particular emphasis on the sector, country, and currency dimensions of both creditor and debtor positions, all of which are important to FSA.

If shared, granular data would allow statistical compilers to identify, and resolve inconsistencies between data compiled in different institutions and in different countries, while possibly reducing the burden for the data reporters. In addition, as the policy makers' data needs change, through the availability of granular data, statistical agencies could compile aggregates in ways that meet these changing needs without sending data requests to data reporters. ([98])

([95]) See Inter-Agency Group on Economic and Financial Statistics (2015).
([96]) See Jenkinson and Leonova (2013).
([97]) For instance, in Deutsche Bundesbank (2015) the importance was stressed of granular loan-level data to facilitate risk analysis, including allowing a better assessment of systemic risks stemming from residential property loans, and, where appropriate, for macro-prudential instrument calibration purposes.
([98]) Bean (2016) underlines the benefits of granular data from the perspective of UK statistics.

To meet the need for increased availability of granular data not only could the collection of more granular data be considered but more use could be made of existing micro data (data that are collected for supervisory or micro-prudential purposes). In this context, the development of principles for effective risk data aggregation and risk reporting by the Basel Committee on Banking Supervision (BCBS), ([99]) the creation of a common data template for G-SIBs to include bi-lateral exposures and exposures to countries/sectors/instruments, and the development of a legal entity identifier system (LEI), ([100]) to identify unique parties to financial transactions are all relevant. Further, these international initiatives help to promote harmonized reporting across reporters, an important issue to address if aggregate micro data are to be used for more macro analysis.

Other initiatives to strengthen financial institutions' risk reporting practices include data reporting requirements arising from the implementation of Basel III ([101]) and the Solvency II rules; the development of recovery and resolution plans by national banking groups; and the efforts to enhance international financial reporting standards. ([102]) In addition to contributing to financial institutions' own risk managements, the improvements in regulatory reporting can contribute to the quality of the more aggregate macro-prudential data for the assessment of system-wide financial stability risks at the national, regional, and international levels.

However, the use of micro data for macro financial assessment has its challenges, the most important being the strict confidentiality requirements associated with the use of micro data. Such requirements typically limit data sharing among statistical and supervisory agencies, and with users. But also granular information brings data quality and consistency issues that need to be dealt with to be able to draw appropriate conclusions for macro-prudential analysis. Tissot points out the importance of being able to aggregate micro information so it can be analyzed, and communicated to policy makers while on the other hand the 'macro' picture on its own can be misleading, as it may mask micro fragilities that have system-wide implications. ([103])

Macro-stress testing is a key tool to assess the resilience of financial institutions and sectors to shocks and would benefit from more detailed information particularly for the top-down stress tests.

Suggested way forward

Work is ongoing as part of Recommendation 20 of DGI-2 to promote the sharing of data within jurisdictions and with other national authorities. ([104]) However, given the differences in legal, statistical structures and cultural backgrounds across jurisdictions, enhancing data sharing is a challenging task and cannot be accomplished overnight. Going forward, international organizations (IOs) should continue their facilitator role by creating platforms to exchange experiences and to help the building of trust.

([99]) Basel Committee on Banking Supervision (2012).
([100]) For more information on the Legal Entity Identifier (LEI) see http://www.fsb.org/what-we-do/policy-development/additional-policy-areas/legalentityidentifier/.
([101]) Under Basel III, Pillar III aims to promote market discipline through enhanced regulatory disclosure requirements. See Basel Committee on Banking Supervision (2015).
([102]) See Basel Committee on Banking Supervision (2012).
([103]) See Tissot (2015b).
([104]) In their March 18, 2017 communiqué, the G-20 Finance Ministers stated that they 'welcome the recommendations of the IAG for sharing and accessibility of granular data' (see G-20 Finance Ministers and Central Bank Governors (2017)). The recommendations are set out in the IAG data sharing report on the PGI website (IAG documents). See 'Update on the Data Gaps Initiative and the Outcome of the Workshop on Data Sharing.'

REAL ESTATE MARKETS

Considering the potential direct and indirect effects on the stability of the financial system, as demonstrated during the GFC, national and regional authorities are placing increasing emphasis on the monitoring of real estate markets.

Significant improvements have been made by national authorities since the GFC in both the scope and coverage of data on residential and commercial real estate markets. This improvement has been promoted in particular through the support of the DGI, the BIS public property price statistics database, and the inclusion of residential property price index (RPPI) among the core FSIs. Residential real estate prices are also an item in the SDDS Plus.

Despite the progress in the number of economies disseminating real estate price indices, the datasets on residential and commercial property prices vary in terms of quality and coverage. For the residential property prices, given the availability of conceptual guidance, (105) the situation is relatively well covered by the BIS database, with wide country coverage, some consistency of data, and, for several advanced and emerging economies, with long-time series data. But the geographical and type of property coverage varies significantly among countries. As regards commercial property prices, their coverage in the BIS public property price statistics database has been expanding significantly since 2016 in the context of the DGI-2, although data are currently available from only a few number of countries with differing frequencies and scopes (e.g., in terms of type of property, area covered, compilation method).

Against this background, the need to improve the quality and availability of data on real estate markets has been emphasized in the Financial Stability Reviews of many economies. (106)

In November 2016, the European Systemic Risk Board (ESRB) published a recommendation. (107) on closing data gaps for residential and commercial real estate markets, underscoring the significance of developments in the real estate sector for financial stability and the considerable data gaps that continue to exist in this area. The aim is to establish a more harmonized framework for monitoring developments in real estate markets in the European Union. The recommendation sets out a common set of indicators that national macro-prudential authorities are recommended to monitor along with working definitions of these indicators.

Finally, in addition to the price indices, there is an FSA need for additional housing-related indicators to complement the price indices.

Suggested way forward

At the international level, guidance has been provided on the compilation of RPPI, while for Commercial Property Prices Indices (CPPI), conceptual guidance is in early stages of development. (108) Going forward, national efforts are key to ensuring the availability of consistent data on property prices, and other indicators of the property market.

(105) See Inter-Secretariat Working Group on Price Statistics (2013).
(106) European Central Bank (2016) focused on the limited coverage of existing price indicators focusing on prime commercial property in large cities. The Central Bank of Ireland, in its December 2016 Macro-Financial Review, also indicated the need for comprehensive and independent information on the commercial real estate sector including data on stock, sales, leases and planning (Central Bank of Ireland (2016)) . Australian, German and Canadian authorities have highlighted the need for improved information on the housing market.
(107) https://www.esrb.europa.eu/news/pr/date/2016/html/pr161128.en.html.
(108) The ISWGPS is expected to provide an overview of methods for compiling CPPI-data with the publication of 'Commercial Property Price Indicators: Sources, Methods and Issues.' See https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2016/Eurostat_commercial_property_price.pdf.

Under Recommendation 17 of DGI-2, the Inter-Secretariat Working Group on Price Statistics (ISWGPS), led by the OECD, and in collaboration with the IAG, is developing a list of other housing- related indicators, such as price-to-rent and price to income ratios.

INSURANCE COMPANIES

As emphasized in the IMF April 2016 GFSR, [109] before the GFC insurance companies were not thought to pose significant systemic risks having longer-term liabilities, greater diversification of assets, and less extensive interconnections with the rest of the financial system than deposit-takers. However, the near-collapse of the AIG in 2008 revealed the potential systemic risks that could be associated with large insurance companies. As a consequence, the International Association of Insurance Supervisors (IAIS) has identified G-SIIs whose distress or disorderly failure would cause significant disruption to the global financial system and for whom additional capital surcharges are scheduled to be applied starting in 2022. [110]

While there is more comprehensive data on insurance companies available from micro and supervisory data sources compared to other non-bank financial institutions, data gaps (such as information on liability structures) still remain, addressing of which would allow for more complete risk assessments. [111] In this context, the April 2016 GFSR emphasizes that while progress is being made on the micro side, there needs to be a greater macro-prudential focus.

Enhancements to insurance sector data would include better data on common exposures, on interconnections with other financial institutions including cross-border, on the duration gap between assets and liabilities, and on the structure of liabilities including for life insurance companies the relative size of minimum guaranteed products [112] and variable annuities within total liabilities.

Suggested way forward

Under Recommendation 4 of DGI-2 the FSB, in close consultation with the IMF and IAIS, is to consider the possibility of a common data template for G-SIIs. As with the G-SIBs common data template, developing such a template and the subsequent collection of systematic granular information could be challenging, although the work would benefit from the experiences with G-SIBs. Depending upon the outcome of this initiative, in the long-term collection of granular information using the template could be considered including more widely by the regulators for domestic and non-systemic insurers.

HOUSEHOLDS

Another area where better data are needed to assess financial stability risks is related to the monitoring of the household sector. [113] Such data include comprehensive information on the composition of assets and liabilities, and household income and debt service payments. [114]

(109) See International Monetary Fund (2016c).
(110) See IAIS 2016 list of G-SIIs, http://www.fsb.org/2016/11/fsb-publishes-2016-g-sii-list/ and 'IAIS Announces Systemic Risk Assessment and Policy Workplan,' February 2017.
(111) In recent Financial Stability Reviews, the ECB, U.S., and Australia place emphasis on the need for additional data on insurance companies.
(112) Such as investment-oriented life insurance policies.
(113) The importance of the development of a comprehensive dataset on the assets and liabilities of the household sector was emphasized by the IMF in Argentina's 2016 FSAP.
(114) The BIS publishes a database of debt servicing ratios, defined as interest payments plus amortizations (repayments of principal) to income, for the total private non-financial sector for 32 countries from 1999 onwards. http://www.bis.org/publ/qtrpdf/r_qt1509h.htm.

Further, the growing interest of policy makers in the inequality gap (i.e., of consumption, saving, income and wealth) has led to a demand for distributional information.

Suggested way forward

Countries could share their experiences in the compilation of household data including as part of their sectoral accounts statistics. ([115]) While household surveys are key data sources to provide structured information, they are costly to conduct therefore could be complemented with administrative data to the extent that the confidentiality restrictions allow. ([116])

As part of Recommendation 9 of DGI-2, the OECD, in cooperation with Eurostat and the ECB, is working with G-20 economies to encourage the production and dissemination of distributional information on income, consumption, saving, and wealth, for the household sector based on the sectoral accounts framework.

Figure 2: Key data gaps and the suggested way to close them

Key data gaps	Suggested ways forward
Shadow banking	
• Lack of data due to heterogeneous nature of the institutions, lack of regulatory oversight, a previously lack of recognition of the systemic importance of the shadow banking system, and a lack of consistent definition. • Sectoral balance sheets and flow of funds data provide a good initial basis but a number of limitations remain. • Data by economic function, with more granular information on maturity and liquidity transformation, and currency mismatches is needed.	• FSB's efforts to improve availability of information as part of its annual monitoring report is key. • The IAG Working Group on Institutional Sector Accounts is currently working on better capturing shadow banking activity using macro-economic based data, by exploring the possibility of capturing more granular sub-sectoral breakdowns and instruments (for the non-bank financial sector. • It is important for statisticians, regulators, and other users to share experiences in compiling and analyzing shadow banking data, including on ways to ensure comprehensive coverage and avoid duplication of effort. • Frequent dissemination of data would facilitate the timely assessment of the shadow banking system and its linkages with the rest of the financial system.
Capital flows	
BIS, IMF and IFA Group assessments of the effects of capital flow volatility presented to the G-20 in 2016 include: • BoP a key source of information but for a detailed picture of capital flows, particularly regarding country and currency dimensions, some data gaps need to be addressed. • Provide more timely BoP data with a higher frequency of indicators to assess capital flows. • Separating the flows associated with non-financial corporate activity from those of the financial sector in the BoP. • Identify the direction of flows between individual countries or groups of countries. • Increasing the number of countries disseminating the breakdown of direct investment data by geographical location, sector and currency, as well as the breakdown of portfolio investment asset and liability data by	• A holistic review of cross-border exposures could be undertaken by the IMF BOPCOM to see how these data could be leveraged to best meet policy makers' needs. • Data from the G-SIBs common data template that covers these institutions exposures to national markets and sectors could be aggregated to provide an early indication of cross-border capital flows from the largest global banks. • Further the G-SIBs template could be used by a broader range of national authorities to collect granular information on national banking systems exposures and funding dependencies.

([115]) A comprehensive set of sectoral balance sheets would also cover data for the nonfinancial corporate sector, another important gap for some economies.
([116]) The ECB has established the Household Finance and Consumption Network that conducts the Eurosystem's Household Finance and Consumption Survey. See https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html.

Figure 2: Key data gaps and the suggested way to close them (Cont.)

geographical location, sector and currency, as well as the breakdown of portfolio investment asset and liability data by geographical location. • Provide external balance sheet data on currency composition and remaining maturity of capital flows.	
Corporate borrowing	
• Significant increase in offshore borrowing by NFCs, particularly in EMEs. Cuts across resident-based data. • Specific data gaps for NFCs include: (i) foreign currency borrowing, particularly through off-shore affiliates; and (ii) information on corporates' risk exposures (including hedging activities). • Analysis of the vulnerability of EMEs debtors to foreign currency exposures must take account of leverage, debt maturity, and the external/domestic distinction of debt. • Regarding NFCs' foreign currency borrowing, BIS international debt securities database provides comprehensive information on total issuance, with currency and maturity breakdowns. But there is a lack of data on NFCs off-shore foreign currency borrowing from deposit-takers, while domestic foreign currency borrowing also needs to be assessed.	• The BIS suggestion to combine the residence-based BoP data with the BIS IBS and IDS to shed more light on NFCs cross-border exposures and their evolution. • The IAG reference document on 'Consolidation and corporate groups: an overview of methodological and practical issues' (2015) provides some ideas for further work with regard to capturing NFCs cross-border exposures. • Enhanced disclosures of financial hedges and derivatives positions (including detailed currency and maturity information on financial hedges and their underlying positions) through improved accounting standards could also contribute to the availability of consistent information on the risk exposures of NFCs.
Granular data	
• Detailed and granular information is increasingly being requested to contribute to the flexibility of FSA tools. • The use of micro data for macro financial assessment has its challenges, including strict confidentiality requirements, and data quality and consistency issues. • Macro-stress testing is a key tool to assess the resilience of financial institutions and sectors to shocks and would benefit from more detailed information. • Availability and sharing of granular and micro data would allow compilers to identify and resolve inconsistencies.	• Given the differences in legal, statistical structures and cultural backgrounds across jurisdictions, enhancing data sharing is a challenging task and cannot be accomplished overnight. • Going forward, international organizations should continue their facilitator role by creating platforms to exchange experiences and to help the building of trust.
Other gaps	**Ways forward**
Real estate markets	
• The datasets on residential and commercial property prices vary in terms of quality and coverage.	• National efforts are key to ensuring the availability of consistent data on property prices, and other indicators of the property market. Work under the recommendation 17 of DGI-2.
Insurance companies	
• Enhancements to insurance sector data would include better data on common exposures, on interconnections with other financial institutions including cross-border, on the duration gap between assets and liabilities, and the structure of liabilities including for life insurance companies the relative size of minimum guaranteed products and variable annuities within total liabilities.	• Under recommendation 4 of DGI-2 the FSB, in close consultation with the IMF and IAIS, is to consider the possibility of a common data template for G-SIIs.
Households	
• Such data include comprehensive information on the composition of assets and liabilities, and household income and debt service payments.	• Recommendations 8 (sectoral accounts statistics) and 9 (distributional information) of DGI-2.

5. How can FSA data needs be addressed in economic and financial statistical manuals? [117]

While the FSA data needs identified cover a wide range of data series, and the previous section discussed the data gaps that are requested be filled, the question arises as to whether there are common themes in the data needed for FSA that could be met through adaptions of the System of National Accounts (SNA), BoP and related manuals (macroeconomic statistical manuals). [118] The authors believe that such common themes do exist and so advocates a discussion on how the national accounts framework might be best developed to help meet the needs of FSA in the upcoming review of the manuals, likely to start later this decade.

The paper makes this call for three main reasons:

- Since the last update round in the 2000s there has been a much-increased policy focus on financial stability, and it is the purpose of each update round to take account of economic and financial developments, and the consequential needs of policy makers, that have inevitably occurred since the last round;
- the macroeconomic statistical manuals have a central role in the production of economic and financial statistics at national and international statistical offices, with the SNA covering the whole economy; and
- to support an integrated approach to the use of datasets for different policy purposes avoiding duplication of data collection. [119]

As background, the underlying conceptual framework is grounded in sound theoretical economic concepts with the consequence that it has remained largely unchanged over many decades. The periodic updates of the core manuals have thus focused on enhancements that: (1) address new economic and financial developments, and new and emerging policy needs; (2) provide a fuller exposition of existing conceptual advice; and (3) further integrate conceptual advice across the various macroeconomic statistical manuals.

To contribute to the discussion this section sets out some suggestions for enhancements to the macroeconomic statistical manuals with regard to credit quality, financial derivatives, remaining maturity and foreign currency, the sub-sector breakdown of NBFI, and distinctions by size. In choosing these items, two considerations were taken into account:

(1) That the proposed enhancements are consistent with, but avoid as far as possible overwhelming, the SNA framework. This is an important consideration because the conceptual framework of these macroeconomic statistical manuals is designed primarily for national authorities to collect, compile and disseminate data to help support macroeconomic policy making; and

(117) This section focuses on macroeconomic statistical manuals but, as noted earlier, there are other data needs, such as with regard to micro data and data by nationality, and these go beyond the present macroeconomic framework.

(118) These manuals include the Government Finance Statistics Manual, the Monetary and Financial Statistics Manual and Compilation Guide, and the External Debt Statistics and Public Debt Statistics Guides. The manuals are primarily intended to meet the needs of macroeconomic policy making and analysis. Their implementation promotes standardization across countries and consistency over time. These manuals are available at http://www.imf.org/en/data.

(119) The integration of conceptual advice across the manuals is one of the strengths of this family of manuals, allowing cross-sector interconnections to be identified. See for instance Shrestha and Mink (2011).

(2) The items are aligned with financial stability policy needs and tools based on identified user needs that have been commonly expressed and referenced earlier in the paper.

When referring ahead to including the enhancements in the 'central framework,' this means a proposal to include the enhancements in the core statistical accounts, not as supplementary or as memorandum items.

Nonperforming loan (NPL) and provisions

While the macroeconomic statistical manuals provide guidance to compile credit data, and the System of National Accounts (SNA) recommends supplementary items on contingent items such as loan commitments, letters of credit, guarantees, etc. (2008 SNA paragraph 11.24), there is a lack of information on credit quality for non-traded instruments in the central framework. ([120]) Yet credit quality information is important to FSA as it is an indicator of problems borrowers are having, with implications for creditors.

The SNA does recommend memorandum items on NPLs, at nominal and market value, for the government and financial corporations sectors and if significant, as supplementary items for the other sectors, including the rest of the world (2008 SNA, paragraph 13.67). Provisions are in the central framework of the Monetary and Financial Statistics Manual and Compilation Guide (MFSMCG), as they are taken into account when determining the capital of deposit-takers by being included in other accounts payable [Monetary Statistics [MS]] (see MFSMCG paragraph 2.32 and Figure 2.2).

Two possibilities exist for bringing some measure of creditworthiness into the central framework of the macroeconomic statistical manuals. First, NPLs at nominal value could be introduced into the central framework for all sectors with data from creditors providing information on the counterpart borrower sector. ([121]) Flows for NPLs would be recorded as other changes in volume of assets (OCVA).

Second, provisions for losses on assets that are valued at nominal value could be brought into the central framework as provisions affect economic activity, both through the impact they have on the profitability of credit extension and, for deposit-takers, on capital through regulatory provisioning practices. Further, as credit quality worsens and provisions increase deposit-takers typically become more cautious in their lending activity. The flows would be recorded as OCVA given provisions are not an exchange between parties, allowing the outstanding value of loans to be calculated more closely reflecting their market value. ([122])

This paper considers including provisions in the central framework rather than NPLs to be the more robust approach for the reasons given below, while there is already compilation experience through monetary data. ([123])

In the national accounts, if credit quality deteriorates, for traded instruments, the market price changes, resulting in lower net worth of the creditor or a decrease in the market price of the equity liability. However, for instruments valued at nominal value, such as loans, a deterioration in credit quality is not reflected in the value of the instrument but because it might well feed through to a lower market price of equity liabilities of creditors, is likely to

([120]) In the SNA, non-traded instruments are valued at nominal value, unlike traded instruments which are valued at market value.
([121]) A standard definition of NPLs was introduced in the FSI Compilation Guide and subsequently in the 2008 SNA.
([122]) The market value of loans is also impacted by changes in interest rates.
([123]) While aggregate data on NPLs provide useful information for FSA, aggregate data would not deliver enough granular data, such as NPL by credit quality of borrower.

be reflected in an increase in net worth as measured in the national accounts system. The latter arises because net worth is the balancing item of the national accounts balance sheet. So, the present approach reduces the analytical value of the accounts because it does not reflect economic developments in, or attribute them to, the relevant instruments, and disguises signs of worsening creditworthiness among debtors.

Unlike debt securities which the debtor can buy back in the market, for non-traded instruments there is not that opportunity, so from the debtor, and indeed creditor perspective, the value of the debt obligation remains the full contractual amount. So, it can be argued that the value of non-traded instruments valued at nominal value should remain the amount owed without adjustment for provisions, as indeed is the approach in the MFSMCG and FSI Compilation Guide, with provisions added as a separate line item in the accounts. This approach would have the advantage of not only ensuring that the amount owed continues to be recorded but that provisions and write-offs would be separate line items in OCVA, because write-offs, unlike provisions, reduce the amount owed and hence the outstanding value of the instrument.

Including provisions in the central framework of the macroeconomic statistical manuals would affect the timing of the transfer of value within the system as value would transfer when the provisions are made rather than when write-offs occur. ([124]) But as indicated above such timing more accurately reflects the profitability and net worth of deposit-takers, and avoids disguising a deterioration in the creditworthiness of debtors.

Notional value of derivatives

Financial derivatives were introduced into the central framework in the 1990s as these markets begun to flourish. The data are compiled at market value consistent with the principles of the SNA. However, financial derivatives are not debt instruments through which economic agents finance imbalances in consumption and production but rather instruments through which risk is transferred around the system. Recording only market value misses the extent of risk exposures and transfers, and it is these risk exposures and transfers around the system that interest FSA.

So, to gain a fuller picture, not least to measure foreign currency exposures and leverage more broadly, data on notional value (in addition to market value) are needed. Indeed, the Balance of Payments and International Investment Position Manual, sixth edition (BPM6) includes the notional value of foreign currency derivatives in its memorandum table on foreign currency, while the BIS publishes notional (and market) values in its six-monthly survey of over-the-counter derivatives (on a cross-border consolidated basis). But despite these data sets, important as they are, there lacks a residence-based economy-wide picture of financial derivative positions by risk category by sector and sub-sector.

So, while recognizing that notional value does not fit the conceptual framework of the macroeconomic statistical manuals, but to provide a more comprehensive view of the risks underlying the economic and financial system, and how they change over time, the full range of financial derivative positions held, by type of risk category, by counterpart sectors, at notional value could be added as memorandum items.

([124]) Value created and transferred stays within the system. In this way gains and losses in value arising from the policy actions arising from financial crises can be tracked through the national accounts. See Frécaut (2004) and Frécaut (2016).

Remaining maturity

Original maturity of debt assets and liabilities is the standard approach to maturity in the macroeconomic statistical manuals, with a distinction between short-term (up to one-year) and long-term (over one-year). While data on an original maturity basis is of interest to FSA in that it provides information on borrower's access to the short and long markets, there is greater FSA interest in remaining maturity as it informs on debt falling due in the near-term. Remaining maturity data helps indicate the amount of debt that needs to be refinanced, the liquidity of debtors and creditors, and the extent of maturity mismatches between assets and liabilities. ([125])

A number of manuals including the BPM6, MFSMCG, and Public Debt Statistics and External Debt Statistics Guides, have already introduced remaining maturity as a memorandum or supplementary item to position data: long-term original maturity data is broken down into up-to-one year due and over-one year due; and by adding the up-to-one year due data to short-term original maturity data, short-term maturity on a remaining maturity basis can be calculated without undermining the concept of original maturity. Bringing this distinction into the position data in the central framework of the macroeconomic statistical manuals would help meet the needs of FSA. ([126])

Foreign currency

Policy makers have clearly indicated through the G-20 a need for more information on foreign currency exposures. The MFSMCG includes a domestic and foreign currency breakdown through its sectoral balance sheet by instrument and counterpart; BPM6 includes a memorandum table with a foreign currency breakdown of positions by sector, major currency, that also takes account of financial derivatives; and the Public Debt Statistics and External Debt Statistics Guides have domestic/foreign currency splits in their presentational tables.

Introducing a foreign currency/domestic currency breakdown into the central frameworks of both the SNA and the BPM, combined with the introduction of a remaining maturity breakdown in the position data would immensely improve understanding of the foreign currency risks facing the economy. Supplementary or memorandum items breaking down foreign currency data by major currency could also be considered.

Sub-sector breakdown of NBFI

The 2008 SNA introduced a new breakdown of NBFIs ([127]) with seven sub-sectors. ([128]) The composition of the seven sub-sectors is logical and well-considered but depending on countries experience with compiling and analyzing data for the seven sub-sectors, the subsections of the NBFI could be reviewed to determine if the sub-sectoring of NBFIs should be modified to meet the analytical needs of FSA.

([125]) Remaining maturity is also highly relevant for bank-level liquidity stress tests.

([126]) A further remaining maturity breakdown within up-to-one year, at three months would also be of benefit for FSA but would add to the degree of disaggregation beyond the existing macroeconomic statistical framework.

([127]) NBFI is not a subsection recognized by the SNA but it is often referred to as covering all financial corporations except deposit-takers, including the central bank.

([128]) The seven sub-sectors are money market funds (MMFs), non-MMF investment funds, other financial intermediaries, except insurance corporations and pension funds, financial auxiliaries, captive financial institutions and money lenders, insurance corporations, and pension funds.

The FSB has developed an analysis of shadow banking using five economic functions. ([129]) While such a characterization might not be appropriate for the SNA, given the interest of policy makers in shadow banking activity and entities, the work of the FSB could inform any SNA review of NBFI sub-sectoring ([130]).

Distinctions based on size

The SNA and BPM6 frameworks make no allowance for size of entity in advising on the compilation of sector and sub-sector aggregate data. Nonetheless, there is considerable interest among financial stability analysts in activity by households by income level, by non-bank financial corporate by sales, assets, and/or employment, and deposit-takers by assets.

While such data are not typically disseminated by national statisticians, the raw data they receive from reporters often allows for such data to be compiled. Therefore, the possibility of compiling and disseminating data by size, probably as memorandum items, could be investigated. The data for households could be informed by the work under the DGI on household distributional information.

6. Concluding remarks

The past 10 -15 years have seen a major change in policy makers' attitude to analyzing financial stability. Particularly since the GFC there has been an emergence of FSA governance arrangements alongside an increased focus on macro-prudential analysis. With FSA firmly established this paper has undertaken a holistic stock-take of the types of data series used. Our understanding is that this is the first review of its kind at the international level. The paper has found that while the specific datasets used can differ across country and over time, common patterns of data use emerge.

Where does this leave statisticians? Overall there has been an encouragingly constructive response to this increased policy focus on FSA, not least through the G-20 DGI. But more work is required to meet FSA data needs, not least in implementing the initiatives underway. This includes data relating to shadow banking, capital flows, and corporate borrowing, as well as the increased demand for granular data. Further, with the start of the update round of the SNA and BPM expected later this decade, this paper has identified enhancements such as adding provisions data and including remaining maturity and foreign currency breakdowns in the central framework to support FSA without undermining the conceptual framework of the manuals.

([129]) These five functions are management of collective investment vehicles with features that make them susceptible to runs; lending dependent on short-term funding; market intermediation dependent on short-term funding or secured funding of client assets; facilitating credit creation; and securitization-based intermediation.

([130]) An OECD Working Party on Financial Statistics' survey of shadow banking using national accounts-based data also saw the possibility of distinguishing financial institutions and financial instruments 'according to the criteria established by the FSB,' see Organization for Economic Co-operation and Development (2015).

Appendix 1: Data used in FSA

This appendix discusses the data used for FSA based on the research conducted by the authors and set out using the framework of analysis provided in Section 3. The detailed list of the datasets identified by this review is available on request from the authors.

Increasing the resilience of the financial system

The data used covers a very broad range of activities, starting with the size and structure of the financial system including relative to GDP. Such data compiled over time provides not only a cross-sectoral view of the relative size of the financial sector and its components, but also picks up shifts in the structure and size over time, including those that could arise from changes in regulations.

FINANCIAL INSTITUTIONS

For deposit-takers, data collected and compiled to support prudential supervision of individual banking institutions remain essential. Further, at the aggregate level, balance sheet data provide an overview of funding (including retail/wholesale split), type of assets owned by instrument, sector, maturity and currency breakdowns, ([131]) and of capital available, inter alia, allowing calculation of the equity capitalization to book value ratio. Further, data based on the IMF list of FSIs covering capital adequacy, credit worthiness, profitability, and liquidity indicators are monitored. ([132])

Separate indicators by domestic private, domestic state- and foreign-owned deposit-takers (preferably disclosing branches and subsidiaries separately) are analyzed given the differing nature of the capital (and potential liquidity) support. Concentration measures and structure, such as in terms of the types of deposit-takers, are also used, not least for assessing the potential impact on competition.

Whereas for deposit-takers there is often a significant amount of information available to financial stability analysts, with some exceptions, at this time such depth of information is lacking for NBFIs. ([133]) Nonetheless there is increasing interest in data on non-bank financials, particularly broken down by type of activity in order to assess the scale and type of risks to which they are exposed.

Typically, the types of data analyzed are those covering balance sheets, assets and liabilities, with granular instrument, maturity, currency, breakdowns for insurance companies, pension funds, MMF, investment funds (such as hedge and bond funds), mutual credit institutions, and leasing companies, etc. Leverage, liquidity, and various profitability indicators are also monitored. Beyond these datasets the significant heterogeneity among these entities means that different types of datasets are used depending upon the type of NBFI; for example, solvency measures

([131]) For example, International Monetary Fund (2017) points out that foreign-currency denominated lending is a potential risk factor in many low-income developing countries, given the significance of foreign currency-denominated assets and liabilities on bank balance sheets. The paper notes that as seen in many more developed economies, the quality of foreign currency loans to unhedged domestic borrowers can be quickly impaired by significant depreciation of the domestic currency (paragraph 44).

([132]) See International Monetary Fund (2013) Basel III introduced two new liquidity requirements for deposit-takers - Liquidity Coverage Ratio (LCR) and net stable funding ratio (NSFR), (see http://www.bis.org/publ/bcbs189.pdf) and they have been adopted in the revised list of FSIs. With regard to deposit-taker liquidity see also Basel Committee for Banking Supervision (2008).

([133]) See for instance, International Monetary Fund (2015a).

for insurance companies and actuarial liabilities of defined pension funds are dataset very specific to those types of NBFI.

FINANCIAL MARKETS ([134])

Monitoring activity in financial markets is increasingly important for FSA. These markets include the money markets, other short-term borrowing markets (such as repurchase and security lending), debt security, equity, derivatives and foreign exchange markets. In addition to their role in allocating savings and supplying short-term finance to financial institutions, financial markets provide real-time price signals, both for the market as a whole and for individual institutions: including from interest rates and yield curves; spreads, such as between domestic government bonds and international benchmarks; credit default swap rates; and exchange rates.

Experience suggests that too often in financial crisis the lack of liquidity, previously deep and so creating an illusion of continued availability, has been a cause of severe financial difficulty as investors all try to exit at the same time. So, an important aspect of financial sector resilience is the depth of domestic financial markets. Also, deeper and more liquid domestic financial markets can reduce the incentive for domestic corporations and government to borrow in foreign markets and so reduce the vulnerability to foreign induced shocks, in particular vulnerability to foreign exchange risk. ([135]) ([136]) So liquidity indicators are monitored, including data on inventories (and scaled by trading volume and including hedges and other offsetting positions) of dealers who intermediate in these markets due to their important role in providing liquidity.

Further, FSA financial market monitoring includes data on market capitalization, the volume (turnover) of activity; the type of investors in the market, and their interrelationships, particularly relevant if financial stability conditions become fragile; the extent to which collateral is used and reused and with what 'haircut' - with changes in the level of the 'haircut' providing information on market sentiment; where relevant, gross short positions by type of participant; and with regard to derivatives markets, the scale of on- and over-the-counter market activity.

FINANCIAL MARKET INFRASTRUCTURE AND INCLUSION

The effective operation of financial market infrastructure such as payments systems and clearing houses, including Central Counter-parties (CCP), is a crucial aspect of financial stability, as any failure of such infrastructure can cause significant losses on the financial sector and their customers, as well as undermine trust in the financial system. In broad terms data covering scale of activity (including scale and timing of intra-day settlements), margining requirements, and capitalization are monitored.

Further, financial inclusion impacts FSA, in that beyond the economic benefits arising to households from financial inclusion, broader participation in the formal financial system adds to liquidity and the spreading of risks. ([137]) Various indicators including access to banks, Automated

([134]) The resilience of the financial system can be looked at from the viewpoint of both a holistic view of institutions by type of institution using data on the range of their activities and positions, and by the activities of all institutions transacting in a specific market - institution-or market-based monitoring.

([135]) However, domestic markets could still be exposed to spillover risks from external capital flows.

([136]) The original list of FSIs included two market liquidity indicators — average bid-ask spread in the securities market; and average daily turnover ratio in the securities market — but as these data are readily available from commercial sources in most countries, they were dropped in the revised list.

([137]) For a discussion on the trade-offs between the benefits and risks of greater financial inclusion see Mehrortra and Yetman (2015).

Teller Machines (ATM) and Mobile Money Accounts are used to monitor the supply of financial services to the household sector, along with indicators that assess the associated risks.

Containing the build-up of systemic vulnerabilities: Credit, debt and asset prices

CREDIT-RELATED

As noted in a 2012 IMF paper, prolonged credit booms are a harbinger of financial crises and have real costs while the optimal macro-prudential policy response, as well as the optimal policy mix, will likely have to depend on the type of credit boom. ([138]) So the demands from financial stability analysts are for increasingly detailed breakdowns of this indicator. To this end, there is a growing interest in collecting granular data from deposit-takers to allow for the compilation of various types of specific dimensions by the compiling agency. Measures of credit monitored can include borrowing in foreign markets, such as in foreign debt security markets, and from foreign parents of domestic entities. ([139]) Further, the growth of credit through the FinTech industry, such as peer-to-peer lending, is beginning to be monitored where relevant. ([140])

In short, among the data used for FSA are total credit, its growth and size relative to GDP, its currency composition, maturity profile, and the interest rates charged; the sector, industry and regional distribution of credit; the provision of credit by type (loan, securities, trade credit etc.) and ownership of institution (private domestic, domestic state-owned, and foreign-owned); the type of household credit (mortgage, ([141]) credit card, student loans etc.); commercial real estate credit (including real estate companies); collateralized and non-collateralized loans distinguished; and measures of asset quality such as nonperforming loans and provisions data. And various combinations of these variables.

For developing economies, financial deepening in terms of the involvement of the non-bank private sector is an important aspect of FSA. Such data that may be drawn upon are credit to the private sector as percentage of GDP, in addition to the financial inclusion indicators described above. ([142])

Further, due to the inherent risks, data on deposit-takers connected lending — banks lending to an entity that has ownership connections with the deposit-taker providing the credit, concentrated lending — banks lending to a common lender, and directed lending — government directing deposit-takers to whom to lend, as percentage of total loans are monitored.

DEBT-RELATED

The relevance of debt to FSA lies in monitoring its total size, not least relative to GDP, an indicator considered highly relevant for FSA, its composition in terms of instruments, currencies, and

([138]) See Dell'Ariccia *et al* (2012).
([139]) The BIS definition of credit data published for 43 countries covers loan, debt security, and currency and deposit liabilities of the private nonfinancial sector to domestic banks, all domestic sectors, and nonresidents. See https://www.bis.org/publ/qtrpdf/r_qt1609c.htm.
([140]) Peer-to-peer lending brings together individual lenders and borrowers outside the traditional deposit-taking system.
([141]) 'Mortgage' is a common word to use for residential real estate loans, but more precisely the interest is in household debt collateralized by residential real estate.
([142]) For instance, see the International Monetary Fund (2016b).

maturity, and in the ability (or lack of ability) of debtors to service their debt. ([143]) ([144]) Further liquidity and solvency problems can arise with high levels of debt relative to income and wealth. Detailed information on government debt (at both nominal and market values), is particularly important given its central role in financial markets and as an indicator of sovereign risk. ([145]) Also, data on the sector of the depositor and the scale and composition of wholesale borrowing are also used to better understand the diversification of funding sources for deposit-takers and to assess the concentration risks of the financial institutions. ([146])

Further, currency and maturity mismatches between assets and liabilities can raise potential financial stability risks regardless of sector, hence the use of remaining maturity and foreign currency data.

Also, related to debt is the concept of leverage — the relative size of debt to equity in funding the asset side of the balance sheet. What might be considered excessive leverage caused problems for some financial institutions in the GFC as sharp declines in asset prices meant that their debt positions exceeded the value of their assets, wiping out their capital base. Leverage is measured through balance sheet data, but the use of off-balance sheet instruments such as financial derivatives — whereby exposures entered into greatly exceed the initial investment, mean that leverage measures used also take account of off-balance sheet positions. ([147]) These leverage measures are relevant for all economic sectors, although those for households and government differ from corporations in that to all intents and purposes these sectors do not have equity capital, although they have assets against which debt can be set — so called 'net debt' measures.

Non-financial corporate (NFC), household (HH) and government sectors

As customers of the financial sector, data on NFC and HH are monitored as problems in these sectors can cause problems for the financial sector, and vice versa. For both sectors, comprehensive balance sheet data is the starting point for FSA, including assets and liabilities (in domestic and foreign currencies), equity (for NFC), and net worth.

For NFC, there is interest in profitability (e.g., return on assets), and various income and debt measures (including interest coverage ratio, ([148]) debt service ratio, and interest exposure (fixed and variable) to assess their creditworthiness. Further to identify 'risky companies,' and so gauge the amount of credit out of total credit to such companies, measures such as high interest coverage ratios are used. And monitoring activity by industry and size (large, and small and medium sized enterprises (SME)) is becoming important.

For HH, various income (e.g., disposable income) and debt measures (e.g., debt service to disposable income) are monitored, including by income distribution. 'Risky households' are

([143]) For instance, Reinhart and Rogoff (2009).
([144]) Debt covers a broad range of instruments — those that require payments of interest and/or principal to be made, including currency and deposits, debt securities, and loans as well as instruments such as trade credit and contractual pension liabilities. The scope of debt instruments is described in Task Force on Finance Statistics (2014).
([145]) The IMF's Debt Sustainability Analysis framework assesses a country's current debt situation and identifies vulnerabilities in the debt structure covering analysis of the sustainability of total public debt and that of total external debt. Two types of frameworks have been designed: those for market-access countries and those tailored for low-income countries. See http://www.imf.org/external/pubs/ft/dsa/ Typically debt sustainability analysis is conducted at nominal value - see chapter 14 of Task Force on Finance Statistics (2016).
([146]) More specifically, International Monetary Fund (2017), points out that over reliance on public sector funding could be a risk factor in several countries, particularly oil exporters, as the fiscal conditions could easily lead to drawing of public deposits hence resulting in funding strains in financial institutions.
([147]) Under Basel III, the calculation of leverage takes account of the market value and an 'add- on' to take of the potential future exposure in the remaining life of the financial derivative contract. The 'add-on' is calculated by applying an add-on factor to the notional principal amount of the financial derivative.
([148]) The interest coverage ratio is usually defined as earnings before interest and tax (EBIT) to interest.

identified based on debt service ratios and negative net financial assets. But also relevant is information on household assets, whether they are in financial or non-financial assets, whether liquid or not.

For government, in addition to the potential impact of debt and deficits on financial stability, public sector arrears to private sector suppliers can be a transmission channel through which fiscal strains undermine non-financial corporate and financial sector health. ([149]) Further, contingent liabilities can be a potential source of financial risk. ([150])

ASSET PRICES

Among asset prices, real estate prices, both for residential and commercial property, as well as for land are closely monitored. Financial institutions lend against the collateral of real estate, while for households' mortgage loans are invariably the largest liability they take out, and if prices fall below the value of the loan (so-called negative equity) they can face financial difficulties. Consequently, real estate prices are closely watched albeit the problem of homogeneity of properties and infrequency of transactions of the same property make measurement a conceptual challenge. ([151]) There is also interest in other housing-related indicators such as the volume of transactions, price-to-rent and price-to-income ratios, including by region, and in commercial real estate indicators such as commercial property yields and vacancy rates.

Financial market asset prices, such as equity and bond prices, and its components of equity prices by type of sector - for instance, commodity-related, financial-related, foreign-demand related, etc., are also monitored for FSA, not least when used as collateral. Such prices directly impact financial wealth and can indirectly impact the economy through the effect on consumer and corporate confidence. There is also considerable interest in volatility measures, both intra-day and over time. A popular measure is the VIX indices from the Chicago Board Options Exchange such as on equity prices, interest rates, exchange-traded funds, and currency-related, as volatility is telling something about uncertainties/risk in the markets.

Structural vulnerabilities within the financial system: Financial interconnections and spillovers

One message that clearly emerged from the GFC was the importance of monitoring financial interconnections both within economies and across border. ([152]) The data sets used to meet these requests include sectoral balance sheets, to-whom from-whom data, and the major cross-border internationally coordinated surveys such the BIS IBS and the IMF's CPIS and CDIS.

More specifically, an increasing important aspect of the work on interconnectedness relates to debt both from the debtor and creditor perspective, at the national and sectoral level and in terms of to-whom is the debt is owed. Who owes what to-whom is as important for the soundness of the creditors as it is for the creditworthiness of the borrowers. Further, inter-

([149]) See International Monetary Fund (2017).
([150]) For example, Commonwealth of Australia (2016) and New Zealand (2016, page 82) annually assess the potential fiscal risks of contingent liabilities.
([151]) A 2011 IMF working paper investigated the question: what kind of indicators should trigger policy intervention to stop or slow down a real estate boom. See Crowe *et al* (2011).
([152]) The IMF's 2014 Triennial Surveillance Review (TSR) noted the importance of an in-depth view of how global and domestic shocks affect individual economies and propagate across sectors (paragraph 18 of the Overview paper, International Monetary Fund (2014d)). For this purpose, the TSR called for the adoption of the Balance Sheet Approach (BSA) and the development of a global flow of funds. See 'Global Flow of Funds: Mapping Bilateral Geographic Flows' ISI conference, Hong Kong 2013 for a description of the global flow of funds. See www.statistics.gov.hk/wsc/STS083-P1-S.pdf.

connections that arise through ownership structures among financial institutions, and between financial institutions and non-financial corporations, need to be monitored as they can have systemic implications for the efficiency, and perhaps the stability of, the financial system.

The cross-border interconnections of the deposit-taking sector are particularly important because it is through these institutions that financial stability issues in foreign economies can be transmitted into domestic economy. So, data on cross-border assets and liabilities, preferably on a granular basis (including by individual deposit-taker), are used, with instrument, country, sector, and currency breakdowns. This is relevant on both residence- and cross-border consolidated (nationality) based approaches. ([153]) ([154])

The GFC demonstrated that some institutions are so big that when they get into difficulty it has global consequences. For these Global Systemically Important Financial Institutions, and particularly G-SIBs, ([155]) extra capital charges have been imposed to offset what might be seen as a too-big-to-fail benefit. ([156]) Under the G-20 DGI data are being collected in a common template on the relationships among G-SIBs and their exposure to national sectors and markets. ([157]) Inter alia, these data allow identification of common exposures. At the domestic level, similar considerations can arise for Domestic Systemically Important Financial Institutions.

To inform on the reliance of residents on different sources of external finance that could dry up if there was a period of financial stress in foreign economies data are used on residents borrowing: (1) directly from abroad; (2) from domestic branches of foreign deposit-takers; and (3) through foreign subsidiaries of resident entities. Indeed, evidence has shown that increased reliance by deposit-takers on wholesale borrowing from abroad may be an early indication that credit growth in the economy is becoming unsustainable. ([158])

Swings in scale and direction of cross-border capital flows can have financial stability consequences. ([159]) In response the IMF has, since 2011, been publishing an annual Spillover Report that initially focused on the major economies whose policies have spillover implications. ([160]) The IMF has also developed an institutional view on capital flow measures. ([161]) More recently the G-20 IFA working group has also focused attention on capital flows and crisis prevention. ([162]) Balance of payments data are a key source of information but for spillover analysis also relevant are bond yields to assess the correlation of yields across national financial markets; GDP and industrial production to assess the impact of cross-border spillovers on real activity; and, international investment position (IIP) data, and external debt data, of which short-term, government debt, and corporate debt securities by currency and by country of creditor, to identify potential vulnerabilities to changes in capital flows.

([153]) See Tissot (2016).
([154]) The BIS consolidated international banking statistics differ from the SNA-based concepts due to the use of a consolidated (nationality) rather than residence-based approach with regard to the bank reporters. These data were developed in the 1970-80s to capture international banking business not covered by the resident-based data.
([155]) G-SIBs are identified through an indicator-based measurement approach that take into consideration size, interconnectedness, substitutability, complexity and cross-jurisdiction activity. For more information see www.fsb.org/2015/11/fsb-publishes-the-2015-update-of-the-g-sib-list/ and www.bis.org/publ/bcbs255.htm.
([156]) Similar work has been undertaken for G-SIIs with additional capital charges scheduled to be applied starting in 2022. See http://www.fsb.org/wp-content/uploads/FSB-publishes-2016-G-SII-list.pdf.
([157]) See Appendix 2 for information on the collection of data from G-SIBs.
([158]) For instance, see Borio (2011).
([159]) For instance, see International Monetary Fund (2016a). Also, it is argued that open capital markets create a 'financial trilemma' in that only two of (1) national control over financial policies; (2) financial integration with the global market; and (3) financial stability, can be enjoyed simultaneously. See Obstfeld (2015).
([160]) See International Monetary Fund (2015b).
([161]) See International Monetary Fund (2012d).
([162]) The working group was reactivated in 2016 in response of a G-20 call to strengthen the international financial architecture. The 2016 Final Report is available at www.g20.utoronto.ca/2016/P020160815362591309719.pdf

In the context of spillover analysis, the G-20 in particular has been interested in global liquidity and its drivers, as the global liquidity environment can have implications for domestic macro prudential policy. ([163]) The BIS publishes global liquidity indicators that draw on bank credit data from the IMF, BIS locational banking statistics, and BIS debt securities data. ([164])

Finally, national FSAs typically consider the international environment, monitoring developments in foreign economies using real and financial indicators, and where relevant commodity prices such as oil.

Appendix 2: What has been the response of statisticians to the growing interest in FSA?

The single most important initiative of the statistical community in meeting the needs of FSA since the GFC has been the G-20 DGI. ([165])

Data gaps initiative

The DGI, now in its second phase, was launched as a response to data gaps identified by policy makers and analysts in the immediate aftermath of the GFC. Consequently, the DGI is primarily focused on financial stability policy needs. ([166]) The DGI consists of 20 recommendations that both strengthen and enhance existing statistical initiatives and, in some instances, initiate new statistical collections. ([167])

Among the data sets promoted under the DGI are:

- *Financial Soundness Indicators* (*FSIs*): Developed to meet the needs of FSA following the Asian crisis of the late 1990s. In response to the DGI, the list of FSIs was updated in 2013 following a global consultation. ([168]) Focused primarily on deposit-takers, FSIs take a macro look at supervisory-type data covering capital, profitability, asset quality, liquidity, and market risk sensitivity. The data are largely drawn from supervisory balance sheet data and income statements, compiled for deposit-takers. In addition, the updated FSI list includes the size and investments of NBFI, solvency and liquidity indicators for NFCs, debt to disposable income ratio for HH, and as a core indicator, residential real estate prices, which some countries are reporting. The number of countries reporting FSI data to the IMF increased to over 120 by 2016. A table of FSIs is typically presented in FSAP assessments, and included in the GFSR.
- *Concentration and distribution measures (CDM)* can identify vulnerabilities developing within the deposit-taking sector that aggregate data may disguise. To this end, under the DGI, the IMF undertook a pilot project on CDMs with member countries on a voluntary basis based

([163]) For instance, see Bank for International Settlements (2016).
([164]) Global liquidity indicators are published quarterly in the BIS Statistical Bulletin .
([165]) The G-20 DGI was launched in 2009 at the request of the G-20 Finance Ministers and Central Bank Governors, to whom the IMF and FSB Secretariat provide a progress report each year, and by the IMF's International Monetary and Financial Committee (IMFC). The DGI is overseen by the IAG.
([166]) See Heath (2013) and Heath and Bese Goksu (2016).
([167]) Specific action plans for the implementation of the DGI-2 recommendations are included in Financial Stability Board and International Monetary Fund (2016).
([168]) See International Monetary Fund (2013).

on a selected number of FSIs. ([169]) DGI-2 will investigate whether to collect such data on a regular basis.

- *Debt Statistics*: The early 2000s witnessed the development of conceptual advice on external debt statistics, with public sector debt statistics following a few years later. ([170]) The two guides, published under the auspices of the Task Force on Finance Statistics (TFFS), ([171]) provide internationally agreed guidance for the measurement of debt, incorporating foreign currency, debt service, remaining maturity, and, for external debt, ultimate risk concepts. The DGI recommended the reporting of public sector debt data to the World Bank (host)/IMF/OECD quarterly public sector debt statistics hub. By 2016 over 70 economies reported data, while over 120 countries report data consistent with the *External Debt Statistics Guide* to the World Bank (host)/IMF quarterly joint external debt hub. ([172])
- *Securities statistics*: To promote improved reporting of security statistics, in 2015 the Working Group on Securities Databases (WGSD) published a *Handbook on Security Statistics (Handbook)*. ([173]) The *Handbook* provides high-level and detailed presentation tables that assist in the compilation and dissemination of securities statistics, along with classifications for different possible breakdowns. ([174]) DGI-2 sets the reporting of debt securities issuance data to the BIS on a quarterly frequency, starting with sector, currency, type of interest rate, and original maturity as an objective for G-20 economies.
- *Real estate prices:* The DGI recommended a two-pronged approach: the BIS and its member central banks to disseminate available real estate prices, both residential and commercial — by 2016, 58 countries reported data; ([175]) and the ISWGPS to publish robust methodology — a *Handbook on Residential Property Price Indices* was published in 2013, ([176]) and prepare guidance on compiling commercial real estate price indices.
- *Derivative and repo statistics:* Following the GFC, the BIS enhanced data on credit derivatives. ([177]) Subsequently, the FSB created two separate expert groups with the objective of aggregating micro-data on derivative markets ([178]) and on repo and security lending markets ([179]) ([180]) to support analysis of the financial stability risks in these markets.
- *Shadow banking:* The FSB at the request of the G-20 has been producing an annual report on shadow banking since 2011 primarily using national accounts-based data from national sources to gain at least a broad estimate of the size of this sector of the economy. For FSA, this is necessary but not sufficient as there is a need for more refined disaggregation along the lines of economic functionality and a greater focus on risk analysis than is possible solely

([169]) See Crowley *et al* (2014).
([170]) The most recent editions of these guides are, External Debt Statistics Guide, 2013 and Public Sector Debt Statistics Guide, 2011.
([171]) The member agencies of the TFFS are the BIS, the Commonwealth Secretariat, the ECB, Eurostat, the IMF, OECD, the United Nations Conference on Trade and Development (UNCTAD) and the World Bank.
([172]) See http://datatopics.worldbank.org/debt/.
([173]) Working Group on Securities Databases (2015). The members of the WGSD are the BIS, the ECB, the IMF, and the World Bank.
([174]) See www.imf.org/external/np/sta/wgsd/pdf/hss.pdf.
([175]) See www.bis.org/statistics/pp.htm.
([176]) The Handbook was co-ordinated by Eurostat under the joint responsibility of the International Labour Organization (ILO), the IMF, OECD, Eurostat, the United Nations Economic Commission for Europe (UNECE), and the World Bank. See Inter-Secretariat Working Group on Price Statistics (2013).
([177]) See Committee on the Global Financial System (2009). Also, for further information on derivatives data published by the BIS, see Tissot (2015).
([178]) See Financial Stability Board (2014).
([179]) See Financial Stability Board (2013).
([180]) See Financial Stability Board (2015b).

with national accounts-based data. With the help of member countries, the FSB is refining its estimates to meet FSA needs, supported by DGI-2.

- *Sectoral accounts*: Comprehensive sectoral balance sheet data are essential for FSA to provide an overview of developments across the whole economy as well as allowing the compilation of many relevant ratios, such as debt to equity (leverage), financial assets and liabilities of individual sectors to total assets and liabilities, etc. In addition, to-whom from-whom data compiled under sectoral accounts provide valuable information on domestic interconnections. For these reasons alone, the dissemination of a comprehensive set of sectoral accounts is probably the single most important contribution national account statisticians could make towards supporting FSA. ([181])

  Further, the work under the DGI by the OECD in close cooperation with Eurostat and the ECB on distributional data on household income, consumption and wealth can help identify inequalities that could impact FSA. ([182])

- *International investment position*: The DGI promoted quarterly IIP data, which became a required item in the Special Data Dissemination Standard (SDDS) in September 2014. ([183]) More broadly, the G-20 DGI is supporting work to improve the availability of foreign currency data ([184]) including by promoting the foreign currency and remaining maturity enhancements to the IIP included in BPM6.

- *International Banking Statistics:* The BIS IBS data have existed since the 1960s. Over time as analytical needs have emerged, the data have been enhanced. In 2012, the Committee on the Global Financial System (CGFS), which oversees the collection of the BIS IBS, approved a major set of enhancements to close gaps in the information available to monitor and respond to financial stability risks. The BIS and its central banks started publishing data with the new enhancements by 2015 and intend to disseminate more data depending on the progress made by reporting countries. ([185])

- *Global Flow of Funds*: While the DGI does not have a specific recommendation to develop a global flow of funds, it does cover most of the data sets needed to compile such a matrix. This includes recommending more frequent compilation of the CPIS with sector breakdowns and, in DGI-2, compilation of both inward and outward CDIS. While a global flow of funds concept has been developed that will allow analysis of financial interconnections within and across border, the work is still in its embryonic stage.

- *Global Systemically Important Banks*: In response to two recommendations in the DGI, two unique datasets based on common templates have been developed that allow monitoring of both the bilateral institutional links of GSIBs and also their exposures to national markets and sectors. ([186]) The work is led by the FSB, in close consultation with the IMF, with a data hub established at the BIS to collect and process the national data in a confidential setting. In DGI-2, the possibility of a common data template for global systemically important NBFI starting with insurance companies is being investigated.

([181]) An extensive sector balance sheet template was internationally agreed in 2012 as a longer-term objective for IMF and OECD economies to work towards, http://www.imf.org/external/np/sta/templates/sectacct/index.htm.
([182]) For instance, see a presentation on this work at https://www.oecd.org/els/soc/Session4-3-Work-of-EG-DNA.pdf.
([183]) The SDDS is a data transparency standard to which IMF member countries voluntary subscribe. See http://dsbb.imf.org/Pages/SDDS/home.aspx.
([184]) See G-20 Finance Ministers and Central Bank Governors (2014). See http://www.fin.gc.ca/n14/14-054-eng.asp.
([185]) The enhancements include expanding the coverage of banks' balance sheets to include their domestic positions, as well as their international activities, and provide more information about the sector of banks' counterparties, in particular banks' exposures to and reliance on funding from non-bank financial counterparts. Please see Avdjiev et al (2015).
([186]) See www.fsb.org/2014/05/r_140506/.

- *Granular data*: To help identify emerging risks to financial stability there is an increasing demand among policy makers for granular data, such as for loans by deposit-takers. Granular data not only allows closer scrutiny of activity, but also allows for the compilation of many different dis-aggregations of data depending upon the needs of policy makers. It might also reduce the costs for reporters by reducing the need to compile different types of disaggregations as policy needs change. However, there are clearly confidentiality issues involved. DGI-2 includes a recommendation to investigate the possibilities, including sharing data across border.

Special Data Dissemination Standard Plus (SDDS Plus)

The SDDS Plus was established by the IMF in 2014 as the third tier of the IMF's data dissemination standards initiative. ([187]) It draws heavily on the DGI datasets with an objective to 'guide member countries on the provision of economic and financial data to the public in support of domestic and international financial stability.' There are nine datasets that provide information on the real and government sectors, but most extensively on the financial and external sectors. ([188])

Data Sets At-A-glance

The IMF paper 2012 Review of Data Provision to the Fund for Surveillance Purposes ([189]) contained a list of data sources (with web addresses) both from the IMF and other IOs in Appendix 8. Many of these data sources are relevant for FSA.

Acknowledgements

The authors would like to thank Emma Angulo, Jose Maria Cartas, Antonio Galicia-Escotto, Alicia Hierro, Deniz Igan, Nigel Jenkinson, Phousnith Khay, Alfredo Leone, Gianmatteo Piazza, Gabriel Quirós, Roberto Rosales, Mike Seiferling, Florina Tanase, Bruno Tissot, and all those who attended the brown bag seminar in the IMF Statistics Department on February 21, 2017.

([187]) See http://dsbb.imf.org.
([188]) The nine datasets are: sectoral balance sheets, general government operations and debt, other financial corporations survey, FSIs, debt securities, Coordinated Portfolio and Direct investment surveys, and the currency composition of official foreign exchange reserves (COFER) survey.
([189]) See International Monetary Fund (2012e)

Acronyms

ATM	Automated Teller Machines
BCBS	Basel Committee on Banking Supervision
BIS	Bank for International Settlements
BoE	Bank of England
BoP	Balance of Payments
BOPCOM	Committee on Balance of Payments Statistics
BPM	Balance of Payments Manual
BPM6	Balance of Payments and International Investment Position Manual, sixth edition
BSA	Balance Sheet Approach
CCP	Central Counterparties
CDIS	Coordinated Direct Investment Survey
CDM	Concentration and Distribution Measures
CGFS	Committee on Global Financial System
COFER	Currency Composition of Official Foreign Exchange Reserves Survey
CPIS	Coordinated Portfolio Investment Survey
CPPI	Commercial Property Price Indices
DGI	Data Gaps Initiative
DGI-2	Second Phase of the Data Gaps Initiative
D-SIB	Domestically Systemically Important Bank
DSTI	Debt-service-to-income
EBIT	Earnings before Interest and Tax
ECB	European Central Bank
e-GDDS	Enhanced General Data Dissemination System
EME	Emerging Market Economies
ESRB	European Systemic Risk Board
FDI	Foreign Direct Investment
FSA	Financial Stability Analysis
FSAP	Financial Sector Assessment Program
FSB	Financial Stability Board
FSIs	Financial Soundness Indicators
FSOC	US Financial Stability Oversight Committee
FSR	Financial Stability Review
FX	Foreign Exchange
G-20	Group of 20
GDP	Gross Domestic Product
GFC	Global Financial Crisis
GFSR	Global Financial Stability Report
G-SIBs	Global Systemically Important Banks
G-SIIs	Global Systemically Important Insurers
HH	Households
IAG	Inter-Agency Group on Economic and Financial Statistics
IAIS	International Association of Insurance Supervisors
IBS	International Banking Statistics
IDS	International Debt Securities
IFA	International Financial Architecture
IIP	International Investment Position

ILO	International Labour Organization
IMF	International Monetary Fund
IMFC	International Monetary and Financial Committee
IO	International Organization
ISWGPS	Inter-Secretariat Working Group on Price Statistics
LCR	Liquidity Coverage Ratio
LDCs	Less Developed Countries
LEI	Legal Entity Identifier
LTI	Loan-to-income
LTV	Loan-to-value
MFSMCG	Monetary and Financial Statistics Manual and Compilation Guide
MMFs	Money Market Funds
MS	Monetary Statistics
NBFIs	Non-bank Financial Institutions
NFCs	Non-financial Corporations
NSFR	Net Stable Funding Ratio
NPLs	Non-performing loans
OECD	Organization for Economic Co-operation and Development
OFR	US Office of Financial Research
OCVA	Other Changes in Volume of Assets
PGI	Principal Global Indicators
RPPI	Residential Property Price Indices
SDDS	Special Data Dissemination Standard
SDDS Plus	Special Data Dissemination Standard Plus
SME	Small and Medium Sized Enterprises
SNA	System of National Accounts
TFFS	Task Force on Finance Statistics
TRs	Trade Repositories
TSR	Triennial Surveillance Review
UN	United Nations
UNCTAD	United Nations Conference on Trade and Development
UNECE	United Nations Economic Commission for Europe
US	United States
WGSD	Working Group on Securities Databases

References

Alexandrova-Kabadjova, B., B. Bravo-Benitez, S. Martinez-Jaramillo, and J. P. Solórzano-Margain (2012), 'An Empirical Study of the Mexican Banking System's Network and its Implications for Systemic Risk', Banco de Mexico, *Working Papers*, No. 2012-07, http://catedrabde.uji.es/naim/images/presentations/banxico_wp.pdf.

Avdjiev, S., P. McGuire, and P. Wooldridge (2015), 'Enhanced data to analyse international banking,' *BIS Quarterly Review*, http://www.bis.org/publ/qtrpdf/r_qt1509f.htm.

Bank for International Settlements and Financial Stability Board (2014), 'Summary: Joint CGFS – FSB-SCAV workshop on risks from currency mismatches and leverage on corporate balance sheets', http://www.fsb.org/wp-content/uploads/r_140923a.pdf.

Bank for International Settlements (2016), 'Economic resilience: a financial perspective, Note submitted to the G20', http://www.g20.utoronto.ca/2017/2017-Germany-BIS-economic-resilience.pdf.

Bank of Korea (2016), *Financial Stability Bulletin*, http://www.bok.or.kr/broadcast.action?menuNaviId=2578.

Basel Committee on Banking Supervision (2006), 'Scope of Application, International Convergence of Capital Measurement and Capital Standards', www.bis.org/publ/bcbs128.htm.

Basel Committee on Banking Supervision (2008), 'Principles for Sound Liquidity Risk Management and Supervision', http://www.bis.org/publ/bcbs144.htm.

Basel Committee on Banking Supervision (2012), 'Principles for effective risk data aggregation and risk reporting', http://www.bis.org/publ/bcbs239.pdf.

Basel Committee on Banking Supervision (2015), 'Revised Pillar 3 disclosure requirements', http://www.bis.org/bcbs/publ/d309.pdf.

Bean, C. (2016), *Independent review of UK economic and financial statistics*, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/507081/2904936_Bean_Review_Web_Accessible.pdf.

Bernanke, B. (2015), 'Should monetary policy take into account risks to financial stability?' Brookings Institute, https://www.brookings.edu/blog/ben-bernanke/2015/04/07/should-monetary-policy-take-into-account-risks-to-financial-stability/.

Borio, C., R. McCauley, and P. McGuire (2011), 'Global credit and domestic credit booms', *BIS Quarterly Review*, http://www.bis.org/publ/qtrpdf/r_qt1109f.pdf.

Cecchetti S. G., I. Fender, and P. McGuire (2010), 'Toward a global risk map' pp. 5, *BIS Working Papers*, No. 309, http://www.bis.org/publ/work309.htm

Central Bank of Ireland (2016), *Macro-Financial Review*, http://www.centralbank.ie/stability/Pages/Introduction.aspx.

Cerutti, E., S. Claessens and L. Laeven (2015), 'The Use and Effectiveness of Macroprudential Policies: New Evidence', *IMF Working Paper*, 15/61, https://www.imf.org/external/pubs/ft/wp/2015/wp1561.pdf.

Cerutti, E., R. Correa, E. Fiorentino and E. Segalla (2016), 'Changes in Prudential Policy Instruments–A New Cross-Country Database', *IMF Working Paper,* 16/110. https://www.imf.org/external/pubs/cat/longres.aspx?sk=43945.0.

Committee on the Global Financial System (2009), '*Credit risk transfer statistics*', No. 35, www.bis.org/publ/cgfs35.htm.

Commonwealth of Australia (2016), *Budget Strategy and Outlook Budget Paper,* No. 1, 2016-17, http://www.budget.gov.au/2016-17/content/bp1/download/bp1.pdf.

Chui, M., E. Kuruc, and P. Turner (2016), 'A new dimension to currency mismatches in the emerging markets: nonfinancial companies', *BIS Working Papers*, No. 550, http://www.bis.org/publ/work550.pdf.

Cihak, M. (2006), 'How do Central Banks Write on Financial Stability', *IMF Working Paper* 06/163, www.imf.org/external/pubs/ft/wp/2006/wp06163.pdf.

Crowe, C., G. Dell'Ariccia, D. Igan and P. Rabanal (2011), 'How to Deal with Real Estate Booms: Lessons from Country Experiences', *IMF Working Paper,* 11/91, https://www.imf.org/external/pubs/ft/wp/2011/wp1191.pdf.

Crowley, J., P. Koukpamou, E. Loukoianova, and A. Mialou (2016), 'Concentration and Distribution Measures for a Selected Set of Financial Soundness Indicators', *IMF Working Paper,* 16/26, http://www.imf.org/external/pubs/cat/longres.aspx?sk=43706.0.

Dell'Ariccia, G., D. Igan, L. Laeven, and H. Tong with B. Bakker and J. Vandenbussche (2012), 'Policies for Macrofinancial Stability: How to Deal with Credit Booms', IMF Staff Discussion Note, https://www.imf.org/external/pubs/ft/sdn/2012/sdn1206.pdf.

Dembiermont C., M. Drehmann and S. Muksakunratana (2013), 'How much does the private sector really borrow - a new database for total credit to the private non-financial sector', *BIS Quarterly Review*, http://www.bis.org/publ/qtrpdf/r_qt1303h.pdf.

Demekas D. G. (2015), 'Designing Effective Macroprudential Stress Tests: Progress So Far and the Way Forward', *IMF Working Paper,* 15/146, http://www.imf.org/external/pubs/cat/longres.aspx?sk=43043.0.

Deutsche Bundesbank (2015), *Financial Stability Review*, https://www.bundesbank.de/Redaktion/EN/Downloads/Publications/Financial_Stability_Review/2015_financial_stability_review.pdf?__blob=publicationFile.

Drehmann, M. (2013), 'Total credit as an early warning indicator for systemic banking crises', *BIS Quarterly Review*, www.bis.org/publ/qtrpdf/r_qt1306f.htm.

Eichner M., D. Kohn and M. Palumbo (2010), 'Financial Statistics for the United States and the Crisis: What did they get right, what did they miss and how should they change?', Federal Reserve Board, Finance and Economics Discussion Series, https://www.federalreserve.gov/pubs/feds/2010/201020/201020pap.pdf.

Ellis, L. (2013), 'Macroprudential policy, what have we learned', Centre for Central Banking Studies, Workshop for Heads of Financial Stability, www.bankofengland.co.uk/research/Documents/ccbs/Workshop2013/Paper_Ellis.pdf.

Errico, L., R. Walton, A. Hierro, H. AbuShanab, and G. Amidzic (2013), 'Global Flow of Funds: Mapping Bilateral Geographic Flows', ISI conference, Hong Kong 2013. www.statistics.gov.hk/wsc/STS083-P1-S.pdf.

European Central Bank (2016), *Financial Stability Review*, May 2016, https://www.ecb.europa.eu/pub/pdf/other/financialstabilityreview201605.en.pdf?e1a2bef79ed901d8ceaee004f2fcecdd.

European Systemic Risk Board (2014), 'Flagship Report on Macro-prudential Policy in the Banking Sector', https://www.esrb.europa.eu/pub/pdf/reports/140303_flagship_report.pdf.

Farhi, M. and Macedo Cintra, M. A. (2009), 'The Financial Crisis and the Global Shadow Banking System', https://regulation.revues.org/7473?lang=en#ftn1.

Fender, I. and P. McGuire (2010), 'Bank structure, funding risk and the transmission of shocks across countries: concepts and measurement', *BIS Quarterly Review*, http://www.bis.org/publ/qtrpdf/r_qt1009h.pdf.

Frécaut, O. (2004), 'Indonesia's Banking Crisis: A New Perspective on $50 Billion of Losses', *Economic Studies*, Vol. 40, Issue 1.

Frécaut, O. (2016), 'A National Wealth Approach to Banking Crises and Financial Stability', *IMF Working Paper,* 16/128, https://www.imf.org/external/pubs/ft/wp/2016/wp16128.pdf.

Financial Stability Board (2013), 'Strengthening Oversight and Regulation of Shadow Banking: Policy Framework for Addressing Shadow Banking Risks in Securities Lending and Repos', http://www.fsb.org/wp-content/uploads/r_130829b.pdf.

Financial Stability Board (2014), 'Feasibility study on approaches to aggregate OTC derivatives data', http://www.fsb.org/wp-content/uploads/r_140919.pdf.

Financial Stability Board (2015a), 'Global Shadow Banking Monitoring Report', http://www.fsb.org/wp-content/uploads/global-shadow-banking-monitoring-report-2015.pdf.

Financial Stability Board (2015b), 'Standards and Processes for Global Securities Financing Data Collection and Aggregation', http://www.fsb.org/wp-content/uploads/FSB-Standards-for-Global-Securities-Financing-Data-Collection.pdf.

Financial Stability Board (2016a), 'Thematic Review on the Implementation of the FSB Policy Framework for Shadow Banking Entities', http://www.fsb.org/wp-content/uploads/Shadow-banking-peer-review.pdf.

Financial Stability Board (2016b), 'Implementation and Effects of the G20 Financial Regulatory Reforms', http://www.fsb.org/2016/08/implementation-and-effects-of-the-g20-financial-regulatory-reforms-2/.

Financial Stability Board (2017), 'Framework for Post-Implementation Evaluation of the Effects of the G20 Financial Regulatory Reforms', http://www.fsb.org/wp-content/uploads/P030717-4.pdf.

Financial Stability Board, Basel Committee on Banking Supervision, Basel Committee on Payments and Market Infrastructure, and International Organization of Securities Commission (2015), '2015 CCP Work Plan', http://www.fsb.org/2015/09/2015-ccp-workplan.

Financial Stability Board and International Monetary Fund (2015), 'Sixth Progress Report on the G-20 Data Gaps Initiative', http://www.imf.org/external/np/g20/pdf/2015/6thprogressrep.pdf.

Financial Stability Board and International Monetary Fund (2016), 'First Progress Report on the Second Phase of the G-20 Data Gaps Initiative', http://www.imf.org/external/np/g20/pdf/2016/090216.pdf.

G-20 Finance Ministers and Central Bank Governors (2014), Communiqué, http://www.fin.gc.ca/n14/14-054-eng.asp.

G-20 Finance Ministers and Central Bank Governors (2017), Communiqué, http://www.g20.utoronto.ca/2017/170318-finance-en.html.

G-20 International Financial Architecture Working Group (2016), '2016 Final Report', http:// www.g20.utoronto.ca/2016/P020160815362591309719.pdf.

G-20 Leaders Summit (2010), 'The Seoul Document', http://www.g20.utoronto.ca/2010/g20seoul-doc.html.

Gruić, B, C. Upper and A. Villar (2014), 'What does the sectoral classification of offshore affiliates tell us about risks?' , Box in http://www.bis.org/publ/qtrpdf/r_qt1412b.pdf.

Gruić, B. and Ph. Wooldridge (2015), 'BIS debt securities statistics: a comparison of nationality data with external debt statistics', *Irving Fisher Committee Bulletin*, No. 39, https://www.bis.org/ifc/publ/ifcb39w_rh.pdf.

Hammer, C. L., D. C. Kostroch, G. Quirós, and STA Internal Group (2017), 'Big Data: Potential, Challenges, and Statistical Implications', IMF Staff Discussion Note, SDN/ 17/06, https://www.imf.org/~/media/Files/Publications/SDN/2017/sdn1706-bigdata.ashx.

Heath, R. (2013), 'Why are the G-20 Data Gaps Initiative and the SDDS Plus relevant for financial stability analysis?', *IMF Working Paper,* 13/6, www.imf.org/external/pubs/ft/wp/2013/wp1306.pdf.

Heath, R. (2015), 'What has Capital Flow Liberalization meant for Economic and Financial Statistics', *IMF Working Paper,* 15/88, http://www.imf.org/external/pubs/ft/wp/2015/wp1588.pdf.

Heath, R. and Bese Goksu, E. (2016), 'G-20 Data Gaps Initiative II: Meeting the Policy Challenge', *IMF Working Paper,* 16/43, https://www.imf.org/external/pubs/cat/longres.aspx?sk=43760.0.

Hoggarth, G., C. Jung and D. Reinhardt (2016), 'Capital inflows — the good, the bad and the bubbly', *Bank of England Financial Stability Paper,* 40, http://www.bankofengland.co.uk/financialstability/Documents/fpc/fspapers/fs_paper40.pdf.

International Monetary Fund (2012a), 'Enhancing Surveillance: Interconnectedness and Clusters', https://www.imf.org/external/np/pp/eng/2012/031512.pdf.

International Monetary Fund (2012b), 'The IMF's Financial Surveillance Strategy', https://www.imf.org/external/np/pp/eng/2012/082812.pdf.

International Monetary Fund (2012c), 'Macrofinancial Stress Testing — Principles and Practices', http://www.imf.org/external/np/pp/eng/2012/082212.pdf.

International Monetary Fund (2012d), 'The Liberalization and Management of Capital Flows: An Institutional View', www.imf.org/external/np/pp/eng/2012/111412.pdf.

International Monetary Fund (2012e), '2012 Review of Data Provision to the Fund for Surveillance Purposes', https://www.imf.org/external/np/pp/eng/2012/082812a.pdf.

International Monetary Fund (2013), 'Modifications to the Current List of Financial Soundness Indicators', https://www.imf.org/external/np/pp/eng/2013/111313.pdf.

International Monetary Fund (2014a), 'IMF Broadens Financial Surveillance: Financial Sector Assessment Program', IMF Survey, http://www.imf.org/external/pubs/ft/survey/so/2014/POL011314A.htm.

International Monetary Fund(2014b), 'Republic of Korea Financial Sector Assessment Program', https://www.imf.org/external/pubs/ft/scr/2014/cr14126.pdf.

International Monetary Fund (2014c), 'Canada Financial Sector Assessment Program', https://www.imf.org/external/pubs/ft/scr/2014/cr1429.pdf.

International Monetary Fund (2014d), 'Triennial Surveillance Review', Overview paper, http://www.imf.org/external/np/pp/eng/2014/073014.pdf.

International Monetary Fund (2014e), Global Financial Stability Report, October 2014, Chapter 2: 'Shadow Banking Around the Globe: How Large, and How Risky?', https://www.imf.org/external/pubs/ft/gfsr/2014/02/pdf/c2.pdf.

International Monetary Fund (2015a), Global Financial Stability Report, April 2015, Chapter 3: 'The Asset Management Industry and Financial Stability', http://www.imf.org/External/Pubs/FT/GFSR/2015/01/index.htm.

International Monetary Fund (2015b), 'Spillover Report', http://www.imf.org/external/np/pp/eng/2015/060815.pdf.

International Monetary Fund (2015c), 'Monetary Policy and Financial Stability', IMF Policy Paper, http://www.imf.org/external/np/pp/eng/2015/082815a.pdf.

International Monetary Fund (2016a), 'Strengthening the International Monetary System — A stocktaking', http://www.imf.org/external/np/pp/eng/2016/022216b.pdf.

International Monetary Fund (2016b), 'Financial System Stability Assessment for the Central African Economic and Monetary Community', http://www.imf.org/external/pubs/cat/longres.aspx?sk=43874.0.

International Monetary Fund (2016c), 'Global Financial Stability Report, April 2016', Chapter 3: 'The Insurance Sector — Trends and Systemic Risk Implications', http://www.imf.org/external/pubs/ft/gfsr/2016/01/.

International Monetary Fund (2016d), 'Global Financial Stability Report, October 2016,' Chapter 3: 'Corporate Governance, Investor Protection, and Financial Stability in Emerging Markets', http://www.imf.org/external/pubs/ft/gfsr/2016/02/.

International Monetary Fund (2016e), 'The Financial Sector Assessment Program (FSAP), Factsheet', https://www.imf.org/en/About/Factsheets/Sheets/2016/08/01/16/14/Financial-Sector-Assessment-Program.

International Monetary Fund (2016f), 'Tax policy, leverage and macroeconomic stability', IMF Policy Paper, http://www.imf.org/external/np/pp/eng/2016/100716.pdf.

International Monetary Fund (2016g), Note to the G-20, April 2016: 'Recent Experiences in Managing Capital Flows', http://g20chn.org/English/Documents/Current/201608/t20160811_3127.html.

International Monetary Fund (2016h), 'Capital Flows — Review of Experience with the Institutional View', IMF Policy Paper, https://www.imf.org/external/np/pp/eng/2016/110416a.pdf.

International Monetary Fund (2017), 'Macroeconomic Developments and Prospects in Low-Income Developing Countries — 2016', IMF Policy Paper, http://www.imf.org/external/np/pp/eng/2016/112316.pdf.

International Monetary Fund, Financial Stability Board and Bank for International Settlements (2016), 'Elements of Effective Macroprudential Policies: Lessons from International Experience', https://www.imf.org/external/np/g20/pdf/2016/083116.pdf.

Inter-Agency Group on Economic and Financial Statistics (2015), 'Consolidation and corporate groups: an overview of methodological and practical issues', http://www.bis.org/ifc/publ/iagrefdoc-oct15.pdf.

Inter-Agency Group on Economic and Financial Statistics (2017), 'Update on the Data Gaps Initiative and the Outcome of the Workshop on Data Sharing', http://www.principalglobalindicators.org/?sk=E30FAADE-77D0-4F8E-953C-C48DD9D14735.

Inter-Secretariat Working Group on Price Statistics (2013), 'Handbook on Residential Property Price Indices', ec.europa.eu/eurostat/documents/3859598/5925925/KS-RA-12-022-EN.PDF.

Jenkinson, N. and I. Leonova (2013), 'The importance of data quality for effective financial stability policies, Legal entity identifier: a first step towards necessary financial data reforms', *Financial Stability Review*, Banque De France, No. 17, https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=2538229.

Jobst, A., L. L. Ong, and C. Schmieder (2017), 'Macroprudential Liquidity Stress Testing in FSAPs for Systemically Important Financial Systems,' *IMF Working Paper,* 17/102 https://www.imf.org/en/Publications/WP/Issues/2017/05/01/Macroprudential-Liquidity-Stress-Testing-in-FSAPs-for-Systemically-Important-Financial-44873.

Kindleberger, C. (1978), *Manias, Panics and Crashes, A History of Financial Crises*.

Laeven, L. and F. Valencia (2013), 'Systemic Banking Crisis Database: An Update', *IMF Working Paper*, 12/163, https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Systemic-Banking-Crises-Database-An-Update-26015.

Lipsky, J. (2011), 'Macroprudential policy — What it is and how to do it' http://www.bfg.pl/sites/default/files/dokumenty/4_2_47_2012.pdf.

Macroeconomic Assessment Group (2010) (established by the Financial Stability Board and the Basel Committee on Banking Supervision), 'Assessing the macroeconomic impact of the transition to stronger capital and liquidity requirement', https://www.bis.org/publ/othp10.pdf.

McGuire, P. and von Peter, G. (2009), 'The US dollar shortage in global banking and the international policy response', *BIS Working Papers,* No. 291, http://www.bis.org/publ/work291.htm.

McGuire, P. and von Peter, G. (2016), 'The resilience of banks' international operations', *BIS Quarterly Review*, http://www.bis.org/publ/qtrpdf/r_qt1603g.pdf.

Mehrortra, A and Yetman, J (2015), 'Financial inclusion - issues for central banks', *BIS Quarterly Review*, http://www.bis.org/publ/qtrpdf/r_qt1503h.pdf.

New Zealand Treasury (2016), 'Budget Economic and Fiscal Update 2016', http://www.treasury.govt.nz/budget/forecasts/befu2016.

Obstfeld, M. (2015), 'Trilemmas and trade-offs: living with financial globalization', *BIS Working Papers,* No. 480, http://www.bis.org/publ/work480.pdf.

Office of Financial Research (2015), 'Financial Stability Report', https://www.financialresearch.gov/financial-stability-reports/files/OFR_2015-Financial-Stability-Report_12-15-2015.pdf.

Ong, L. L. (editor) (2014), 'A Guide to IMF Stress Testing: Methods and Models', IMF, December 2014. https://www.imf.org/external/pubs/cat/longres.aspx?sk=41117.0.

Organization for Economic Co-operation and Development (2015), 'Results of the Survey on Shadow Banking', http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=COM/STD/DAF(2015)1&docLanguage=En.

Osterloo, S and J. de Haan (2004), 'Central Banks and Financial Stability: A Survey', *Journal of Financial Stabilit*y, Vol 1, Issue 2, pages 257–273.

Reinhart, C. M. and K. S. Rogoff (2009), 'This time is different: A panoramic view of eight centuries of financial crises', Princeton University Press.

Rodriguez, P. (2008), 'Data Provision to the IMF for Surveillance Purposes', IMF Committee on Balance of Payments Statistics (BOPCOM-08/28), http://www.imf.org/external/pubs/ft/bop/2008/08-28.pdf.

Schinasi, G. (2004), 'Defining Financial Stability', *IMF Working Paper*, 04/187, www.imf.org/en/Publications/WP/Issues/2016/12/31/Defining-Financial-Stability-17740.

Schneider, C. and L. Schumacher (2014) , 'Introduction to the Balance Sheet-Based Approach to Stress Testing', chapter 2 of *A Guide to IMF Stress Testing: Methods and Models*, http://www.elibrary.imf.org/staticfiles/misc/toolkit/pdf/chap2.pdf.

Shrestha, M. and R. Mink (2011), 'An Integrated Framework for Financial Flows and Positions on a From-Whom-To-Whom Basis', http://www.imf.org/external/np/seminars/eng/2011/sta/pdf/whom.pdf.

Tarashev, N., S. Avdjiev, and B. Cohen (2016), 'International capital flows and financial vulnerabilities in emerging market economies: analysis and data gaps,' BIS note submitted to the G-20 International Financial Architecture Working Group, http://www.bis.org/publ/othp25.pdf.

Task Force on Finance Statistics (2014), External Debt Statistics Guide, www.tffs.org/pdf/edsg/ft2014.pdf.

Tissot, B. (2015a), 'Derivatives statistics: the BIS contribution', ISI conference, Rio de Janeiro, https://www.bis.org/ifc/events/ifc_isi_2015/089_tissot_paper.pdf.

Tissot, B. (2015b), 'Closing information gaps at the global level — what micro data can bring' https://www.bis.org/ifc/events/ws_micro_macro/tissot_paper.pdf.

Tissot, B. (2016), 'Globalisation and financial stability risks: is the residency-based approach of the national accounts old-fashioned?', *BIS Working Papers*, No. 587, http://www.bis.org/publ/work587.htm.

Turner, P. (2016), 'Macroprudential policies, the long-term interest rate and the exchange rate', *BIS Working Papers*, No. 588, http://www.bis.org/publ/work588.pdf.

US Financial Stability Oversight Committee 2016 Annual Report, https://www.treasury.gov/initiatives/fsoc/studies-reports/Pages/2016-Annual-Report.aspx.

White, W. (2006), 'Is price stability enough?', *BIS Working Paper,* No. 205, http://www.bis.org/publ/work205.pdf.

Working Group on Securities Databases (2015), 'Handbook on Securities Statistics', http://www.imf.org/external/np/sta/wgsd/pdf/hss.pdf.

Zhang, T. (2016), Opening Remarks at the IMF-LSE Symposium on 'Macroprudential Stress Tests and Policies: A Framework', http://www.imf.org/en/News/Articles/2016/12/15/SP121516-IMF-LSE-Symposium-on-Macroprudential-Stress-Tests-Policies-A-Framework.

2 Explaining nowcast errors

HENRIETTE DRUBA [1], JENNIFER L. CASTLE [2] AND DAVID F. HENDRY [3]

Abstract: Preliminary estimates of macroeconomic aggregates only become available with a time lag and are often unreliable and must be revised, more so during periods of structural change such as the financial crisis and subsequent recession. This article takes the forecast error taxonomy in Hendry and Mizon (2012), and adapts it to nowcasting. The taxonomy provides a framework for thinking about potential problems facing the nowcaster in achieving useful nowcasts. It considers a wide array of sources of nowcast errors, from estimation uncertainty to model mis-specification. Importantly, the taxonomy incorporates unforeseen changes in parameters, and thus allows for a formal analysis into the consequences of structural breaks for nowcasting. Additionally, the taxonomy is applied to evaluate the impact of the different error sources on the expected nowcast error. This evaluation yields seven insights into the sources of nowcast errors, and into sources of nowcast failure. An empirical example nowcasting consumption expenditure illustrates the insights from the taxonomy.

JEL codes: C52, C53, E21

Keywords: Nowcasting, bias correction, model selection, structural breaks, consumption

(1) Department of Economics, University of Oxford.
(2) Magdalen College, and Institute for New Economic Thinking, Oxford Martin School, Un iversity of Oxford.
(3) Department of Economics, and Institute for New Economic Thinking, Oxford Martin School, University of Oxford.

1. Introduction

The current state of an economy is never known with any certainty, as even preliminary estimates of most macroeconomic aggregates only become available with a time lag, and usually are revised later as new information accrues. Periods of rapid or sudden structural change, like the financial crisis and subsequent recession, exacerbate such problems. Unfortunately, incomplete knowledge of the economy's current state can lead to inappropriate economic policies and inaccurate forecasts of its future performance. Lacking accurate measurements, statistical agencies and policy makers face the challenge of 'forecasting' the contemporaneous state of the economy, a process known as nowcasting. The key difference between a nowcast and a forecast is that the former seeks to ascertain what has actually happened, but as yet is not fully observed, whereas the latter concerns what might happen by a future date. This paper outlines an approach to nowcasting that uses statistical forecasting models to exploit high-frequency, real-time information, disaggregated data, and leading indicators published at higher frequencies to provide 'contemporaneous forecasts' of economic activity.

There are two distinct forms of nowcasting; either predicting the macroeconomic aggregate (e.g., GDP) directly using available cognate information, or using the released data on components and supplementing with predictions for the components with no statistical data on the contemporaneous values. For examples of the former see Giannone *et al.* (2008) and Giannone *et al.* (2009). In this paper we propose following the latter approach, with the aim of assisting statistical agencies to construct timely preliminary estimates of the aggregate series. The aggregate series to be nowcast comprises of many disaggregates, some of which are 'known', i.e., the statistical agencies have reported data (which may be revised) on the components contemporaneously, and some are unknown, so data is missing at the current time. The proposed procedure involves a bridge-equation framework, first producing accurate estimates of the disaggregates and then nowcasts of the aggregate are calculated from the disaggregates.

The nowcast problem faces a 'ragged edge' at the nowcast origin, where some disaggregates have statistical releases for the current time, and some do not. The unknown disaggregates are 'forecast', taking account of the data on already reported disaggregates, and other higher-frequency indicators of the state of the economy that are usually available in a timely manner. The higher-frequency data are transformed to remove any unit-root non-stationarity, and to match the frequency of the aggregate. Given the large number of variables involved, automatic model selection offers a viable approach: see Doornik (2009) and Hendry and Doornik (2014). The general approach is flexible to allow for missing data on components to vary over time. Every disaggregate is 'forecast', including those that are already reported, as the contemporaneous forecast errors from the known disaggregates are informative for adjusting the forecasts of the unknown disaggregates. The bridge equation provides the nowcast of the aggregate using these transformed series.

As with forecasting, producing an accurate nowcast is difficult, more so in turbulent times. Despite the key difference noted above, many of the problems that confront forecasting also impinge on nowcasting: see Castle *et al.* (2017). There are a number of taxonomies of the sources of forecast errors: see for example, Clements and Hendry (1998), Clements and Hendry (2006), Hendry and Hubrich (2011) and Hendry and Mizon (2012). This article takes the more general forecast-error taxonomy for open systems of equations in Hendry and Mizon (2012), and adapts it to nowcasting, with extensions that reflect the differences between nowcasts and forecasts, especially contemporaneous information and the ragged edge problem for missing disaggregates. Our taxonomy is designed to provide a framework for thinking about potential problems facing the nowcaster in achieving useful nowcasts. It therefore considers a wide array of sources of nowcast errors, from estimation uncertainty to model mis-specification.

Importantly, the taxonomy incorporates unforeseen changes in parameters, and thus allows for a formal analysis into the consequences for nowcasting of structural breaks. Additionally, the taxonomy is applied to evaluate the impact of the different error sources on the expected nowcast error. This evaluation delivers seven insights into the sources of nowcast errors we discuss below. In particular, the analysis is focused on isolating sources of nowcast errors that cause nowcast failure. Nowcast failure occurs when nowcasts are significantly different from the eventually measured outcome, examples of which are shown in Ericsson (2017). We also record the variance components of the various error sources, but focus on the expected values of the 28 possible errors as these yield seven insights which we believe may be helpful to agencies producing nowcasts.

Section 2 describes the derivation of the nowcast-error taxonomy, and presents the resulting taxonomy table. Section 3 presents seven insights into potential sources of nowcast failure stemming from the taxonomy. Subsequently, Section 4 presents evidence from Monte Carlo simulations for a simplified setting designed to clarify the analysis, illustrated by an empirical example in Section 5. We conclude in Section 6.

2. Nowcast error taxonomy

This section derives the nowcast error taxonomy. In the most general set-up, the variable to be nowcast is a function of disaggregates and other exogenous information that are contemporaneously available, and those that are missing at the nowcast origin. Due to asynchronous release dates of economic data, nowcasting often involves unbalanced panels, here referred to as 'ragged edges'. The missing disaggregates must be forecast in order to avoid ragged edges. The way of handling missing information may vary between different approaches to nowcasting. The derivation of the nowcast error taxonomy is based on a data-generating process (DGP) that is a function of two strongly exogenous vectors, x_t and z_t, which are stationary processes in sample of dimensions $(N_1 \times 1)$ and $(N_2 \times 1)$ respectively [4]. While x_t is contemporaneously available, z_t includes variables that are missing at the nowcast origin and have to be filled in. Since different variables may be missing at different nowcast origins, the dimensions N_1 and N_2 should be interpreted as time-variant. The DGP thus takes the following form in sample for $t = 1, \ldots, T$, where lags have been omitted for clarity:

$$(1) \quad y_t = \tau + \lambda_1' x_t + \lambda_2' z_t + \epsilon_t = \phi + \lambda_1' (x_t - \rho_1) + \lambda_2' (z_t - \rho_2) + \epsilon_t ,$$

where $\epsilon_t \sim N[0, \sigma_\epsilon^2]$ is the innovation error and $E[\epsilon_t | x_1, \ldots, x_t; z_1, \ldots, z_t] = 0$ is assumed. Further, it holds that $E[y_t] = \phi$, $E[x_t] = \rho_1$ and $E[z_t] = \rho_2$. This gives the relationship

$$(2) \quad \phi = (\tau + \lambda_1' \rho_1 + \lambda_2' \rho_2).$$

Since the DGP is unknown in practice, the researcher may end up falsely including irrelevant variables in the nowcasting model, for example due to their retention in model selection. Thus, suppose equation (1) has been selected in sample over $t = 1, \ldots, T$, starting from a general unrestricted model (GUM) that also includes the vector w_t of irrelevant and strongly exogenous

(4) The assumption of strong exogeneity is introduced to limit dependencies between mean zero results in the nowcast error taxonomy.

variables, which are assumed to be uncorrelated with ϵ_t. The vector is retained, though its true population parameter, λ_3, equals a vector of zeros. This model is then estimated in-sample, and used to nowcast *T+1*. Since z_{T+1} is missing at the nowcast origin it has to be forecast. The nowcasted value of y_t may be written as

$$(3)\quad \hat{y}_{T+1|T+1} = \hat{\phi} + \hat{\lambda}_1^{'}(x_{T+1} - \hat{\rho}_1) + \hat{\lambda}_2^{'}(\hat{z}_{T+1|T} - \hat{\rho}_2) + \hat{\lambda}_3^{'}(w_{T+1} - \hat{\rho}_3).$$

We consider a situation in which there has been an unanticipated and permanent shift in the DGP of y_t between the nowcast origin, *T*, and period *T*+1. Allowing for shifts in all terms, the post-shift DGP is:

$$y_{T+h} = \phi^* + \hat{\lambda}_1^{*'}(x_{T+h} - \rho_1^*) + \hat{\lambda}_2^{*'}(z_{T+h} - \rho_2^*) + \epsilon_{T+h} \quad \text{with } h\geq 1.$$

While mean shifts in the irrelevant variables are incorporated, $\rho_3^* \neq \rho_3$, it is assumed that the vector w_t remains irrelevant following the unanticipated shift, so $\lambda_3^* = \lambda_3 = 0$.

In the nowcast error taxonomy, we also account for estimation uncertainty and for the effect of model selection on expected parameter estimates labelled search bias. Further, we allow for model mis-specification, for example due to omitted variables or in-sample location shifts that were not modelled. The first two factors lead to biased parameter estimates, while the latter two in general entail biased and inconsistent parameter estimates. This is captured by indexing the expected values of the estimators with *e*, and incorporating that the expected values may not equal the true parameter values provided in the DGP (1), e.g. $\phi \neq \phi^e$.

The expression of the nowcast error can be derived by subtracting (3) from (1), and rearranging terms:

$$\hat{\epsilon}_{T+1|T+1} = y_{T+1} - \hat{y}_{T+1|T+1} = \epsilon_{T+1} + (\phi^* - \hat{\phi}) + \lambda_1^{*'}(x_{T+1} - \rho_1^*) - \hat{\lambda}_1^{'}(x_{T+1} - \hat{\rho}_1)$$

$$+ \lambda_2^{*'}(z_{T+1} - \rho_2^*) - \hat{\lambda}_2^{'}(\hat{z}_{T+1|T} - \hat{\rho}_2) - \hat{\lambda}_3^{'}(w_{T+1} - \hat{\rho}_3)$$

The key expansion to arrive at the full nowcast error taxonomy is the following:

$$\phi^* - \hat{\phi} = (\phi^* - \phi) + (\phi - \phi^e) + (\phi^e - \hat{\phi}).$$

This expansion is used for all parameters to re-write the nowcast error in terms of separate components for shifts, mis-specification and estimation uncertainty to yield the full nowcast error displayed in Table 1. Because of the large number of terms in Table 1, we will gradually introduce complications starting with the simplest case in equation (4). As noted above, the explanation here focuses on the expected values of the mistakes and shifts, although direct variance terms of individual components are included in the table ([5]).

([5]) For reasons of parsimony we abstract from any covariance terms between the components.

Table 1: Nowcast error taxonomy

	$\hat{\epsilon}_{T+1\|T+1}=$	Description	Expectation	Variance
(i)	ϵ_{T+1}	Innovation error	0	σ_{ϵ}^{2}
(iia)	$+(\phi^{*}-\phi)$	Equation mean shift	$+(\phi^{*}-\phi)$	0
(iib)	$+(\phi-\phi^{e})$	Equation mis-specification	$+(\phi-\phi^{e})$	0
(iic)	$+(\phi^{e}-\hat{\phi})$	Equation mean mis-estimation	0	$O_p(T^{-1})$
(iiia)	$-\lambda_1'(\rho_1^{*}-\rho_1)$	Mean shift	$-\lambda_1'(\rho_1^{*}-\rho_1)$	0
(iiib)	$+(\lambda_1^{*}-\lambda_1)'(x_{T+1}-\rho_1^{*})$	Slope shift	0	$(\lambda_1^{*}-\lambda_1)'V[x_{T+1}](\lambda_1^{*}-\lambda_1)$
(iiic)	$+\lambda_1^{e\prime}(\rho_1^{e}-\rho_1)$	Mean mis-specification	$\lambda_1^{e\prime}(\rho_1^{e}-\rho_1)$	0
(iiid)	$+(\lambda_1-\lambda_1^{e})'(x_{T+1}-\rho_1^{*})$	Slope mis-specification	0	$(\lambda_1-\lambda_1^{e})'V[x_{T+1}](\lambda_1-\lambda_1^{e})$
(iiie)	$-\lambda_1^{e\prime}(\rho_1^{e}-\hat{\rho}_1)$	Mean mis-estimation	0	$O_p(T^{-1})$
(iiif)	$+(\lambda_1^{e}-\hat{\lambda}_1)'(x_{T+1}-\rho_1^{*})$	Slope mis-estimation	$O_p(T^{-1/2})$	$O_p(T^{-1/2})$
(iiig)	$+(\lambda_1-\hat{\lambda}_1)'(\rho_1^{*}-\rho_1)$	Mean shift covariance	$(\lambda_1-\lambda_1^{e})'(\rho_1^{*}-\rho_1)$	$O_p(T^{-1})$
(iiih)	$+(\hat{\lambda}_1-\lambda_1^{e})'(\hat{\rho}_1-\rho_1)$	Estimation covariance	$O_p(T^{-1})$	$O_p(T^{-1})$
(iva)	$-\lambda_2'(\rho_2^{*}-\rho_2)$	Mean shift	$-\lambda_2'(\rho_2^{*}-\rho_2)$	0
(ivb)	$+(\lambda_2^{*}-\lambda_2)'(z_{T+1}-\rho_2^{*})$	Slope shift	0	$(\lambda_2^{*}-\lambda_2)'V[z_{T+1}](\lambda_2^{*}-\lambda_2)$
(ivc)	$+\lambda_2^{e\prime}(\rho_2^{e}-\rho_2)$	Mean mis-specification	$\lambda_2^{e\prime}(\rho_2^{e}-\rho_2)$	0
(ivd)	$+(\lambda_2-\hat{\lambda}_2)'(z_{T+1}-\rho_2^{*})$	Slope mis-specification	$O_p(T^{-1/2})$	$V[(\lambda_2-\hat{\lambda}_2)'(z_{T+1}-\rho_2^{*})]$
(ive)	$-\lambda_2'(z_{T+1}-E_{T+1}(\hat{z}_{T+1\|T}))$	Mean mis-forecast	$\lambda_2'(\rho_2^{*}-E_{T+1}[\hat{z}_{T+1\|T}])$	$(\lambda_2)'V[z_{T+1}](\lambda_2)$
(ivf)	$-\lambda_2^{e\prime}(\rho_2^{e}-\hat{\rho}_2)$	Mean mis-estimation	0	$O_p(T^{-1})$
(ivg)	$+(\lambda_2-\hat{\lambda}_2)'(\rho_2^{*}-\rho_2)$	Mean shift covariance	$(\lambda_2-\lambda_2^{e})'(\rho_2^{*}-\rho_2)$	$O_p(T^{-1})$
(ivh)	$-\lambda_2'(E_{T+1}(\hat{z}_{T+1\|T})-\hat{z}_{T+1\|T})$	Mis-forecast	0	$(\lambda_2)'[\hat{z}_{T+1\|T}](\lambda_2)$
(ivi)	$+(\hat{\lambda}_2-\lambda_2)'(z_{T+1}-\hat{z}_{T+1\|T})$	Mis-forecast covariance	$E[(\hat{\lambda}_2-\lambda_2)'(z_{T+1}-\hat{z}_{T+1\|T})]$	$V[(\hat{\lambda}_2-\lambda_2)'(z_{T+1}-\hat{z}_{T+1\|T})]$
(ivj)	$+(\hat{\lambda}_2-\lambda_2^{e})'(\hat{\rho}_2-\rho_2)$	Estimation covariance	$O_p(T^{-1})$	$O_p(T^{-1})$
(va)	$+\lambda_3^{e\prime}(\rho_3^{e}-\rho_3)$	Mean mis-specification	$\lambda_3^{e\prime}(\rho_3^{e}-\rho_3)$	0
(vb)	$-\lambda_3^{e\prime}(w_{T+1}-\rho_3^{*})$	Slope mis-specification	0	$(\lambda_3^{e})'V[w_{T+1}](\lambda_3^{e})$
(vc)	$-\lambda_3^{e\prime}(\rho_3^{e}-\hat{\rho}_3)$	Mean mis-estimation	0	$O_p(T^{-1})$
(vd)	$-\hat{\lambda}_3'(\rho_3^{*}-\rho_3)$	Mean shift covariance	$-\lambda_3^{e\prime}(\rho_3^{*}-\rho_3)$	$O_p(T^{-1})$
(ve)	$+(\lambda_3^{e}-\hat{\lambda}_3)'(w_{T+1}-\rho_3^{*})$	Slope mis-estimation	$O_p(T^{-1/2})$	$O_p(T^{-1/2})$
(vf)	$+(\hat{\lambda}_3-\lambda_3^{e})'(\hat{\rho}_3-\rho_3)$	Estimation covariance	$O_p(T^{-1})$	$O_p(T^{-1})$

3. Sources of nowcast errors

From this nowcast error taxonomy seven insights into sources for nowcast failure can be derived.

3.1. The exogenous vector: insights 1 and 2

We first focus on the exogenous vector that is contemporaneously available, x_t, and assume that the DGP is a function of this vector only. We therefore effectively set $\lambda_2=\lambda_2^*=0$, and ignore the terms (iva)-(vf) in Table 1. Hence, the nowcast error includes the components (i)-(iiih). If we additionally abstract from search bias and mis-specification so that $\lambda_1^e=\lambda_1$ and $\rho_1^e=\rho_1$, and ignore estimation covariances ([6]), the terms (i), (iia), (iic), (iiia,b) in Table 1 are the only sources of the nowcast error. Using definition (2), and adding and subtracting the product $\lambda_1'\rho_1^*$, the expected nowcast error thus reduces to:

(4) $E[\hat{\epsilon}_{T+1|T+1}] \approx (\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*+\lambda_1'(\rho_1^*-\rho_1)-\lambda_1'(\rho_1^*-\rho_1)=(\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*.$

If the strongly exogenous variable is omitted the nowcast error becomes:

$\tilde{\epsilon}_{T+1|T+1} \approx (i),(iia),(iic)+\lambda_1^{*\prime}(x_{T+1}-\rho_1^*),$

with expectation:

(5) $E[\tilde{\epsilon}_{T+1|T+1}] \approx (\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*+\lambda_1'(\rho_1^*-\rho_1).$

The *first insight* is that a change in dynamics of the unmodelled exogenous variable, $\lambda_1\neq\lambda_1^*$, alone induces nowcast failure as long as its mean, ρ_1, is different from 0. This nowcast failure is reflected in the expectation of term (iia). Comparing $E[\hat{\epsilon}_{T+1|T+1}]$ with $E[\tilde{\epsilon}_{T+1|T+1}]$ in (4) and (5) gives rise to the *second insight* : incorrectly omitting x_t does not lead to or augment nowcast failure if its mean remains constant, so $\rho_1^*=\rho_1$. Together these first two insights imply that if the mean of the exogenous variable remains constant at 0, the size of the nowcast failure as given by the expectation of the nowcast error is independent of whether or not x_{T+1} is correctly modelled ([7]). By comparing the above two expectations it can also be seen that if a mean shift $\rho_1^*\neq\rho_1$ takes place, then the expected nowcast error is minimised by correctly including x_{T+1}.

3.2. Filling in ragged edges: insights 3 and 4

With these insights in mind, we include exogenous variables in z_{T+1}, and work with the full DGP (1). Thus, we consider terms (i)-(ivj) in Table 1 as potential sources of the nowcast error. Absent mis-specification and estimation covariances, we can focus attention on terms (i),(iia), (iic),(iiia,b),(iva,b),(ive), and the expectation of the nowcast error equals:

(6) $E[\hat{\epsilon}_{T+1|T+1}] \approx (\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*+(\lambda_2^*-\lambda_2)'\rho_2^*+\lambda_2'(\rho_2^*-E_{T+1}[\hat{z}_{T+1|T}]).$

Since we are being agnostic to the method of infilling ragged edges, we allow for the case $\rho_2^*\neq E_{T+1}[\hat{z}_{T+1|T}]$. If z_{T+1} were to be omitted the following nowcast error results:

$\tilde{\epsilon}_{T+1|T+1} \approx (i),(iia),(iic),(iiia)+\lambda_2^{*\prime}(z_{T+1}-\rho_2^*)$

([6]) The latter can be justified with the argument that in a congruent model, the estimation uncertainty is of order $T^{-1/2}$, and at a sample size of eg. $T=100$ may be ignored relative to other sources of bias and variance.

([7]) The squared nowcast error, however, is increased by the $O_p(1)$ term $(\lambda_1^*)'V[x_{T+1}](\lambda_1^*)$ if x_{T+1} is omitted.

with expectation

(7) $E[\tilde{\epsilon}_{T+1|T+1}] = (\tau^* - \tau) + (\lambda_1^* - \lambda_1)'\rho_1^* + (\lambda_2^* - \lambda_2)'\rho_2^* + \lambda'_2(\rho_2^* - \rho_2)$.

A comparison of the two expectations in (6) and (7) yields the *third insight* : If there are no mean shifts in the exogenous vector z_{T+1}, so $\rho_2^* = \rho_2$, there is no gain in mean nowcast accuracy from accurately forecasting the exogenous variable. Equally, accurately forecasting z_{T+1} does not reduce the nowcast error if the mean remains constant but the slope shifts.

The *fourth insight* results by considering a mean shift in z_{T+1}: If there is a location shift in the missing variable, the forecast of z_{T+1} has to be closer to the new mean, ρ_2^*, than the old mean, ρ_2, in order to reduce the expected nowcast error relative to omitting the missing variable.

$$E[\tilde{\epsilon}_{T+1|T+1}] - E[\hat{\epsilon}_{T+1|T+1}] \approx \lambda'_2(\rho_2^* - \rho_2) - \lambda'_2(\rho_2^* - E_{T+1}[\hat{z}_{T+1|T}]).$$

This highlights how forecast failure at the stage of infilling missing variables at the nowcast origin may affect the final nowcast. This also stresses the importance of updating forecasts of missing disaggregates during times of structural breaks, which may imply an unexpected location shift, and hence systematic forecast failure.

The approximation signs reflect that mis-specification, model selection, and estimation covariances were ignored in the preceding analysis. Mis-specification impacts on the expected nowcast error through (iiic) and (ivc), and can only be avoided when omitting z_{T+1} if x_{T+1} and z_{T+1} are orthogonal. Both mis-specification and estimation covariance are functions of all regressors in the model, and their magnitudes relative to unexpected shifts are difficult to compare analytically. The simulations, however, show that if a congruent nowcasting model capturing in-sample shifts has been selected, mis-specification, and estimation covariance seem to be less detrimental for nowcasting accuracy than unexpected shifts.

3.3. Adding model selection: insights 5, 6, and 7

So far the discussion has ignored model selection. In practice, however, this is an important step, and we therefore consider the full nowcast error taxonomy to evaluate the impact of model selection on nowcasting. As statistical estimation entails non-degenerate null distributions, there is a non-zero probability of retaining irrelevant variables. This is labelled a cost of search.

The *fifth insight* refers to the fact that absent location shifts, falsely retaining w_{T+1} only impacts on the expected nowcast error to the extent that the expected values of the estimated parameters differ from the true values of 0 in (va). Again, the bias introduced by mis-specification is difficult to quantify as it depends on all regressors, and may be negligible. A high false retention rate, however, is not without cost. Component (vb) directly increases the squared nowcast error.

Variables are frequently subject to shifts. This motivates the *sixth insight* that location shifts in any retained irrelevant variables lead to systematic nowcast failure through (vb). Of course, variable selection may result in the omission of relevant variables, too. The impact of omitting x_{T+1} and z_{T+1} has already been presented.

Model selection also affects parameter estimates. All statistics for selecting variables to be kept in the final model have interdependent distributions, which differ under the null and the alternative, and are affected by each modelling decision. Thus, model selection impacts on the expected values of parameter estimates. This effect is termed search bias. Hendry and Krolzig (2005) show that search bias is negligible for highly relevant variables, and positive for the squared magnitude of parameter estimates of irrelevant variables. This provides the *seventh* *insight* that correcting parameter estimates for search bias, for example by using the two-step procedure introduced in Hendry and Krolzig (2005), has the strongest impact on the expected nowcast error through the reduction of term (vd): By eliminating the bias in the parameter estimate of w_{T+1}, correcting for search bias reduces systematic nowcast failure following mean shifts in the falsely retained variable. In addition, it may reduce the expected nowcast error through the terms (iib), (iiic,g), (ivc,g) and (va).

In summary, it may be stated that:

- Absent structural breaks, omitting relevant exogenous vectors, x_{T+1} and z_{T+1}, does not cause nowcast failure in expectation, while retained irrelevant variables in w_{T+1} do so to the extent that there is mean mis-estimation.
- With location shifts, retention of relevant exogenous variables minimises the expected nowcast error, while forecasts of missing disaggregates must reflect any location shifts in real time in order to improve nowcast accuracy. Further, with non-zero means of x_{T+1} and z_{T+1}, slope shifts lead to nowcast failure, which is not attenuated by correctly including the relevant vectors. Equally, retained irrelevant variables may lead to nowcast failure if they shift out of sample.
- Overall, the impact of model selection, which may lead to omission of relevant or retention of irrelevant variables, on the expected nowcast error is most pronounced in the face of structural breaks. As a result, the most important mean accuracy gain due to search bias correction stems from reducing parameter estimates of irrelevant variables, which are likely to be marginally significant, if these variables break out of sample.

4. Simulations

The nowcast error taxonomy highlights the costs associated with various sources of nowcast errors. The simulations complement the taxonomy by consecutively adding ragged edges, model selection, and structural breaks to the simulation design. The simulations abstract from the practical issue of the mismatch between frequencies of the variable to be nowcast and potential explanatory variables. We assume that unit roots in the data have been removed. Across simulations, *Autometrics* (see Doornik (2009) and Hendry and Doornik (2014) for a description of the algorithm) is used for variable selection at the significance level $\alpha = 0.05$, and the intercept term is always retained.

First consider the 'first-best' scenario, in which the researcher has perfect knowledge on the DGPs of the variable to be nowcast, y_t, and all relevant disaggregates are contemporaneously available. Thus, we use the DGP of y_t as the nowcasting model, and exclude ragged edges. Subsequently, we introduce ragged edges, which are filled in using the DGP of missing disaggregates as the forecasting model. Consequently, we are ignoring estimation uncertainty concerning the first stage of the bridge equation framework. This simplication has been introduced since the focus of the taxonomy lies on the second stage of the bridge equation framework.

Building on this, we consider the impact of model selection on nowcast accuracy. First, we introduce model selection starting from the DGP of y_t as the GUM including an intercept term. This isolates omission of relevant variables, which we refer to as costs of inference, from retention of relevant variables. Second, we add irrelevant variables to variable selection, but we force the retention of relevant variables. This means that we select over irrelevant variables only, and consequently consider separately costs of search. Third, we search jointly over all available variables. This replicates that the distinction between relevant and irrelevant variables is not known in practice while it does not allow to differentiate between costs of search and costs of inference.

Last, we add structural breaks in two designs. In the first design all variables are affected by identical breaking dynamics, and in the second, only irrelevant variables break. Table 2 gives an overview of the different simulation designs.

Table 2: Overview of simulation designs

Label	DGP	Ragged edges	GUM 1	GUM 2	GUM 3
Features	**Variable Selection**				
	No	No	Yes	Yes	Yes
	The DGP of y_t (equation 9) is used as the nowcasting model	The DGP of y_t is used as the nowcasting model	Variable selection starts from the DGP of y_t.	Variable selection starts from GUM including x_1-x_7 x_1-x_3 are forced into the final nowcasting model.	Variable selection starts from GUM including x_1-x_7 We select over x_1-x_7
	Missing disaggregates				
	No	Yes			
	All disaggregates are available	Ragged edges in disaggregates according to Table 4. DGP (equation 10) of disaggregates used as forecasting models.			
	Structural breaks				
	a) No location shifts b) x_1-x_7 shift simultaneously and identically c) Irrelevant variables x_4-x_7 shift simultaneously and identically.				

To evaluate the predictive performance of the nowcasts, we consider two statistics: the mean forecast error (MFE), and the root mean squared forecast error (RMSFE) based on the nowcast error in the nowcasting equation of y_t. The MFE is defined as:

$$\frac{1}{M}\sum_{i=1}^{M}\frac{1}{H}\sum_{h=1}^{H}(y_{T+h}-\hat{y}_{T+h|T+h}).$$

The MFE averages the nowcast errors over the nowcast periods, $h = 1, \ldots, H$, and over the M replications performed. If nowcasts are not systematically biased, the MFE should not be statistically significantly different from 0. The RMSFE is defined as:

$$(8)\quad \frac{1}{M}\sum_{i=1}^{M}\sqrt{\left[\frac{1}{H}\sum_{h=1}^{H}(y_{T+h}-\hat{y}_{T+h|T+h})^2\right]}.$$

It measures the variation of the nowcasted values around the true values averaged over the M replications performed, so that smaller RMSFEs are preferred.

4.1. Formulating the data-generating processes

Here, we introduce the simple DGP for y_t, which may represent an economic aggregate such as GDP. We suppose there is data on seven time series, $x_{i,t}$, $i = 1, \ldots, 7$ available. Out of the monthly time series, the first three are relevant for y_t, and enter the DGP of the aggregate with equal weights. Variables $x_{4,t} - x_{7,t}$ are noise, which are introduced to make variable selection relevant. While the number of irrelevant variables remains small, the design implies that there are more irrelevant than relevant variables. Given the detailed data available to statistical offices this is deemed a realistic set-up. The time series may be interpreted as monthly disaggregated data or leading indicators for y_t. The DGP of y_t takes the following form:

(9) $y_t = 0.5x_{1,t} + 0.5x_{2,t} + 0.5x_{3,t} + \epsilon_t$ with $\epsilon_t \sim IN[0, 1]$ for $t = 0, \ldots, T + h$.

T refers to the number of in-sample periods, which are used to fit the model to the data. h specifies the number of nowcasting periods.

The DGP for the disaggregated time series is specified as a VAR

(10) $x_t = \pi_0 + \pi_1 x_{t-1} + \delta 1_{7,t>T} + v_t$ with $v_t \sim IN_7[0, \Omega_v]$ for $t = 0, \ldots, T + h$

where $x_t = [x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}, x_{5,t}, x_{6,t}, x_{7,t}]'$. π_1 is a (7×7) matrix, and set to equal a diagonal matrix $\pi_1 = \pi I_7$. We assume that $|\pi| < 1$. The intercept π_0 and error term v_t are (7×1) vectors. $1_{7,t>T}$ denotes an indicator variable that enters the DGP for the periods $t > T$ and hence represents an identical location shift across the 7 disaggregates for the out-of-sample periods, $h = 1, 2, 3$. The (7×1) vector δ specifies which disaggregates are affected by the location shift. In the simulations, the initial values for the disaggregates are set equal to zero, $x_0 = 0$. To reduce the dependence on the initial values of the simulated data, we discard the first 20 simulated observations.

Table 3: Parameter values

Number of simulations M	10 000
In-sample period T	75
Nowcasting horizon h	3
Non-centralities of $x_1 - x_3$ in DGP of y_t	3.8
Slope parameter π	0.6
Intercept π_0	0.1
Correlation ρ	0.6
Break coefficient: All δ	(1, 1, 1, 1, 1, 1, 1)
Break coefficient: Irrelevant δ	(0, 0, 0, 1, 1, 1, 1)
Magnitude of break	$4\sigma_x = \frac{4}{\sqrt{(1-0.6^2)}} = 5$

The non-centralities of x_1 - x_3 were computed by averaging their t-statistics computed in the DGP of y_t over all replications.
Source: Authors' calculations.

The correlation structure between the disaggregated time series $x_{i,t}$, $i = 1,\ldots,7$ must be specified. The 7×7 variance-covariance matrix between the regressors, Σ_x, is induced through the variance-covariance matrix of the error term, Ω_v. The covariance between the error terms entering the DGP of all regressors, ρ, is symmetric, and non-zero. We can derive the correlation structure between regressors based on Ω_v. Since π_1 is a diagonal and symmetric matrix, Σ_X, simplifies to (8):

$$\Sigma_X = \frac{1}{1-\pi^2}\Omega_v .$$

From the unit variance of the error terms it follows that contemporaneous correlations equal the covariances in this example. Table 3 summarizes the numerical parameter values in the simulations.

4.2. Formulating the general unrestricted models

We consider an in-sample period of length $T = 75$, which is equivalent to over six years of monthly data, and focus on three nowcasting horizons $h = 3$. If ragged edges are incorporated, data on one of the disaggregates is treated as missing in each nowcasting horizon, introducing a ragged-edge dataset as presented in Table 4.

Table 4: **Ragged-edge structure**

Nowcasting period	x1	x2	x3
$h = 1$	X	✓	✓
$h = 2$	✓	X	✓
$h = 3$	✓	✓	X

Note: ✓ indicates contemporaneous availability. X stands for missing values.

The forecasting models are given by the DGP of the disaggregates (equation 10). We consider three GUMs for model selection. The GUM refers to the most general specification that is the starting point of selection of the final nowcasting model, and should summarize the actual DGP of the variable to be modelled in the space of variables under consideration. The GUMs are chosen to illustrate the different costs associated with model selection on nowcast accuracy. First, we start model selection from the DGP of y_t, equation (9), including an intercept term. We refer to this set-up as 'GUM 1'. Subsequently, we add irrelevant variables. In the set-ups 'GUM 2' and 'GUM 3', the starting point for model selections is:

$$y_t = \beta_0 + \sum_{i=1}^{7} \beta_i x_{i,t} + U_t \qquad \text{for } t = 1, \ldots, T.$$

In scenario GUM 2, the nowcasting model is selected using automatic model selection over disaggregates $x_4 - x_7$ for the in-sample period, $t = 1, \ldots, T$. The nowcast of y_t is computed from the selected model, where $\hat{\beta}_i$ denotes the parameter estimates obtained in-sample, and k_1 is the number of retained variables after variable selection.

$$\hat{y}_{T+h|T+h} = \hat{\beta}_0 + \hat{\beta}_1 x_{1,T+h} + \hat{\beta}_2 x_{2,T+h} + \hat{\beta}_3 x_{3,T+h} + \sum_{k_1} \hat{\beta}_i x_{i,T+h} \qquad \text{for } h = 1, 2, 3 \text{ and } i = 4, \ldots, 7$$

In GUM 3, automatic model selection considers all disaggregates $x_1 - x_7$ over the in-sample period, $t = 1, \ldots, T$. The nowcast of y_t is computed from the selected model with k_2 denoting the number of retained variables after considering all variables for selection:

(8) This assumes that the disaggregates are *I(0)* or that Π_1 has all its eigenvalues within the unit circle.

$$\hat{y}_{T+h|T+h} = \hat{\beta}_0 + \sum_{k_2} \hat{\beta}_i x_{i,T+h} \qquad \text{for } h = 1, 2, 3 \text{ and } i = 1, \ldots, 7$$

4.3. Simulation results

The discussion of the results is structured into four parts, of which the first three refer to the diferent breaking dynamics in Table 2. We begin by evaluating simulations without location shifts. We then discuss simulations with location shifts affecting all disaggregates, and end with an analysis of simulations with location shifts occurring to irrelevant disaggregates $x_{4,t} - x_{7,t}$ only. Additionally, we consider the impact of correcting parameter estimates for search in the last subsection. The numerical simulation results can be found in Table 5, where we distinguish between simulations with or without bias correction. They are illustrated in barplots throughout the analysis.

Table 5: Simulation Results

	DGP	Ragged edges	GUM 1		GUM 2		GUM 3	
Bias correction	**No**	**No**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**
				No location shifts				
MFE	0.004	–0.0003	–0.003	0.017	–0.003	0.008	–0.003	0.018
RMSFE	0.922	1.033	1.066	1.091	1.064	1.073	1.077	1.097
				Location shifts in $x_1 - x_7$				
MFE	0.004	2.500	2.545	3.085	2.493	2.769	2.510	3.043
RMSFE	0.922	2.667	2.862	3.392	2.786	3.041	2.863	3.371
				Location shifts in $x_4 - x_7$				
MFE	0.004	–0.0003	–0.0035	0.017	–0.007	0.006	–0.153	–0.118
RMSFE	0.922	1.033	1.066	1.091	1.365	1.217	1.477	1.330

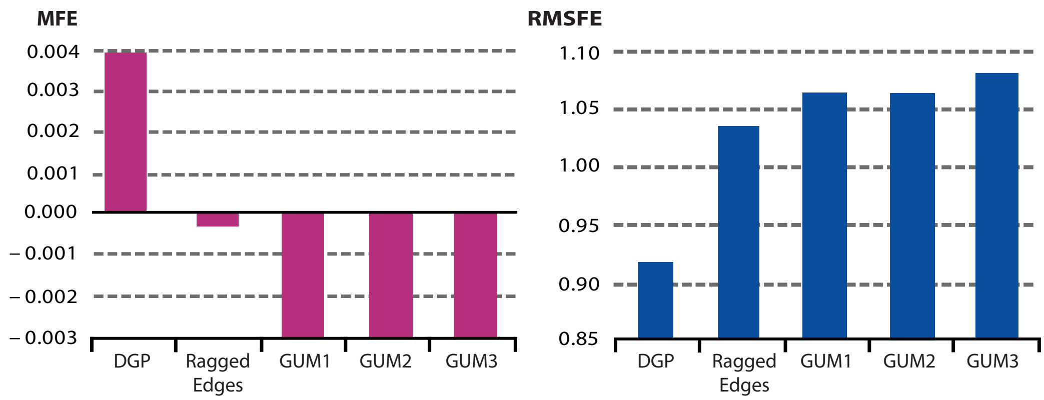
The line 'Bias correction' indicates whether parameters have (Yes) or have not (No) been corrected for search.
Source: Authors' calculations.

4.3.1. WITHOUT STRUCTURAL BREAKS

Across all simulations, the MFE averages to a number close to 0, confirming that there is no systematic nowcast failure as seen in Figure 1. Since their DGP is used to fill in missing disaggregates, and since they do not shift out of sample, their forecasts are correct on average, and the incorporation of ragged edges does not worsen but slightly improves the fit of nowcasts. Equally, starting model selection from GUM 1, and hence introducing costs of inference as well as estimation uncertainty, does not significantly impact on mean nowcast accuracy. This confirms insights 1 and 3: absent location shifts, omitting relevant variables does not cause nowcast failure. Selecting variables in GUM 2 illustrates insight 5, and suggests that costs of search are low if there are no location shifts out of sample. The joint cost of inference, search, and estimation uncertainty as summarised by scenario GUM 3 remains low.

The RMSFEs in Figure 1 show that the magnitude of nowcast errors increases once ragged edges are included, and underline the comparatively small costs related to search and inference. Even if the researcher were able to know the true DGPs for both stages of the bridge equation framework, the fact that there is missing data that have to be forecast increases the RMSFE by around 12 %. Adding costs of inference by introducing model selection starting from GUM 1

Figure 1: MFE and RMSFE: No structural breaks



Source: Authors' calculations.

has a small impact on RMSFEs, which increase by an additional 3 %. Equally, costs of search have negligible effects on the RMSFE, both individually or jointly with costs of inference.

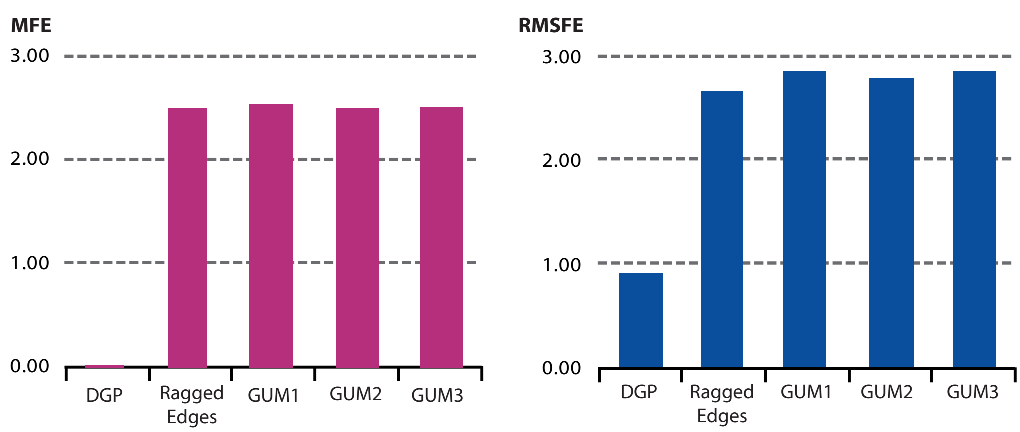
4.3.2. WITH STRUCTURAL BREAKS IN ALL DISAGGREGATES

The initial magnitude of the induced location shift is $\delta_i = 5$ for all disaggregates $i = 1, \ldots, 7$, see Table 3. Since the location shift is permanent, this implies an increase in the long-run mean from $E[x_t] = 0.1/(1 - 0.6) = 0.25$ for $t = 1, \ldots, 75$ *to* $E[x_t] = 5.1/(1 - 0.6) = 12.75$ out of sample. Since the in-sample DGP (equation 10) is used as the forecasting model, it is possible to derive that the expected mean forecast error of missing disaggregates equals the size of the location shift:

$$E[\hat{v}_{i,T+h|T+h-1}] = 0.1 + \delta_i + 0.6E[x_{i,T+h-1}] - (0.1 + 0.6E[x_{i,T+h-1}]) = 5. \quad (11)$$

From the DGP of y_t it then follows that the location shift in the aggregate variable amounts to the weighted average of the shifts in the relevant disaggregates, where the weights are

Figure 2: MFE and RMSFE: Location shifts in x_1 –x_7



Source: Authors' calculations.

given by the parameters of 0.5 in equation (9). Once ragged edges are included, the expected nowcast error of y_t equals $0.5 \times \delta_i = 2.5$ as a result of systematic failure in forecasting missing disaggregates. This systematic bias introduced by the location shift is indeed visible in the MFE and RMSFE in Figure 2, and is close to the analytical value of 2.5. This underlines insight 4: The forecasts of missing disaggregates has to be close to the post-shift mean to avoid nowcast failure.

As insights 1 and 2 suggest, it is optimal to retain relevant regressors if they shift out of sample. With respect to model selection, this suggests that in the face of structural breaks costs of inference, so the omission of relevant variables, should be particularly high. In this example, all three relevant variables have non-centralities of 3.8 as shown in Table 3. Given this high non-centrality, the retention of the relevant variables at $\alpha = 0.05$ is probable as can be gathered from Table 6, so that costs of inference remain low. Nevertheless, MFE and RMSFE increase in GUM 1 and GUM 3, which involve selection over relevant variables. Both statistics in GUM 2 are close to those in ragged edges in magnitude. As we are forcing all relevant variables, and the irrelevant disaggregates shift by the same amount, costs of search and estimation uncertainty remain of negligible importance for nowcast accuracy.

Overall, simulations with shifts in all disaggregates reveal the detrimental effect of location shifts for nowcast accuracy relative to costs of search, inference or estimation uncertainty.

Table 6: Rejection probabilities as a function of non-centralities, ψ, and selection significance levels, α, for T = 75

ψ	0	2	3	4	5
$\alpha = 0.001$	0.001	0.09	0.35	0.71	0.94
$\alpha = 0.01$	0.01	0.27	0.64	0.91	0.99
$\alpha = 0.05$	0.05	0.51	0.84	0.98	1.00
$\alpha = 0.16$	0.16	0.72	0.94	0.99	1.00

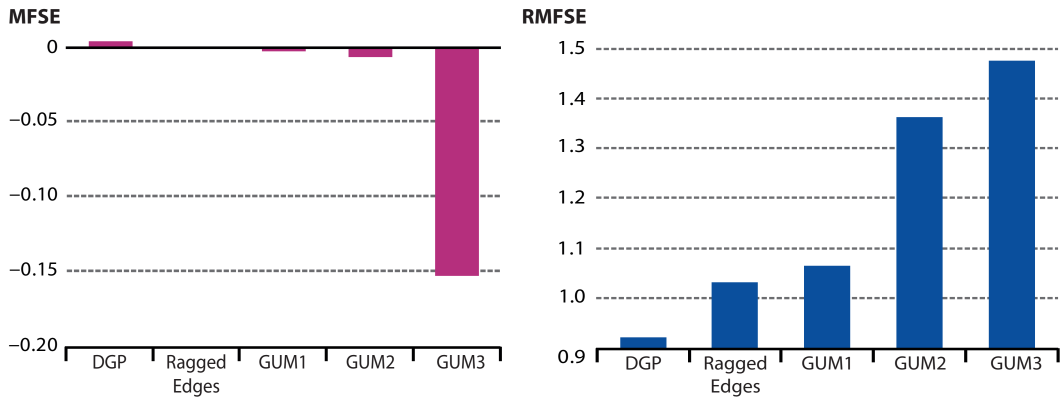
Source: Authors' calculations.

4.3.3. WITH STRUCTURAL BREAKS IN IRRELEVANT VARIABLES

So far, the cost of retaining irrelevant variables has been low. By introducing out-of-sample shifts in irrelevant variables, the distinction between relevant and irrelevant variables in model selection receives more practical importance, and costs of search are increased. Note that the simulated data for y_t, and disaggregates $x_1 - x_3$ are unchanged compared with simulations without location shifts. Consequently, simulation results of DGP, Ragged Edges, and GUM 1, which do not involve irrelevant variables, are identical to those in the first subsection. Figure 3 shows the relevant MFEs and RMSFEs.

Both GUM 2 and GUM 3 involve selection over irrelevant variables. In GUM 2, all relevant variables are forced into the final nowcasting model, and consequently nowcast accuracy does not show a systematic bias. The RMSFE, however, increases substantially, and indicates that magnitudes of nowcast errors have risen. In GUM 3, the example that is closest to model selection in practice, nowcast failure in terms of mean nowcast accuracy becomes apparent. Even in a set-up with a small number of irrelevant variables, so that almost no irrelevant variables will ever be retained in any replication, we can confirm insight 6 that a location shift in irrelevant variables may lead to systematic nowcast failure.

Figure 3: MFE and RMSFE: Location shifts in $x_4 - x_7$



Source: Authors' calculations.

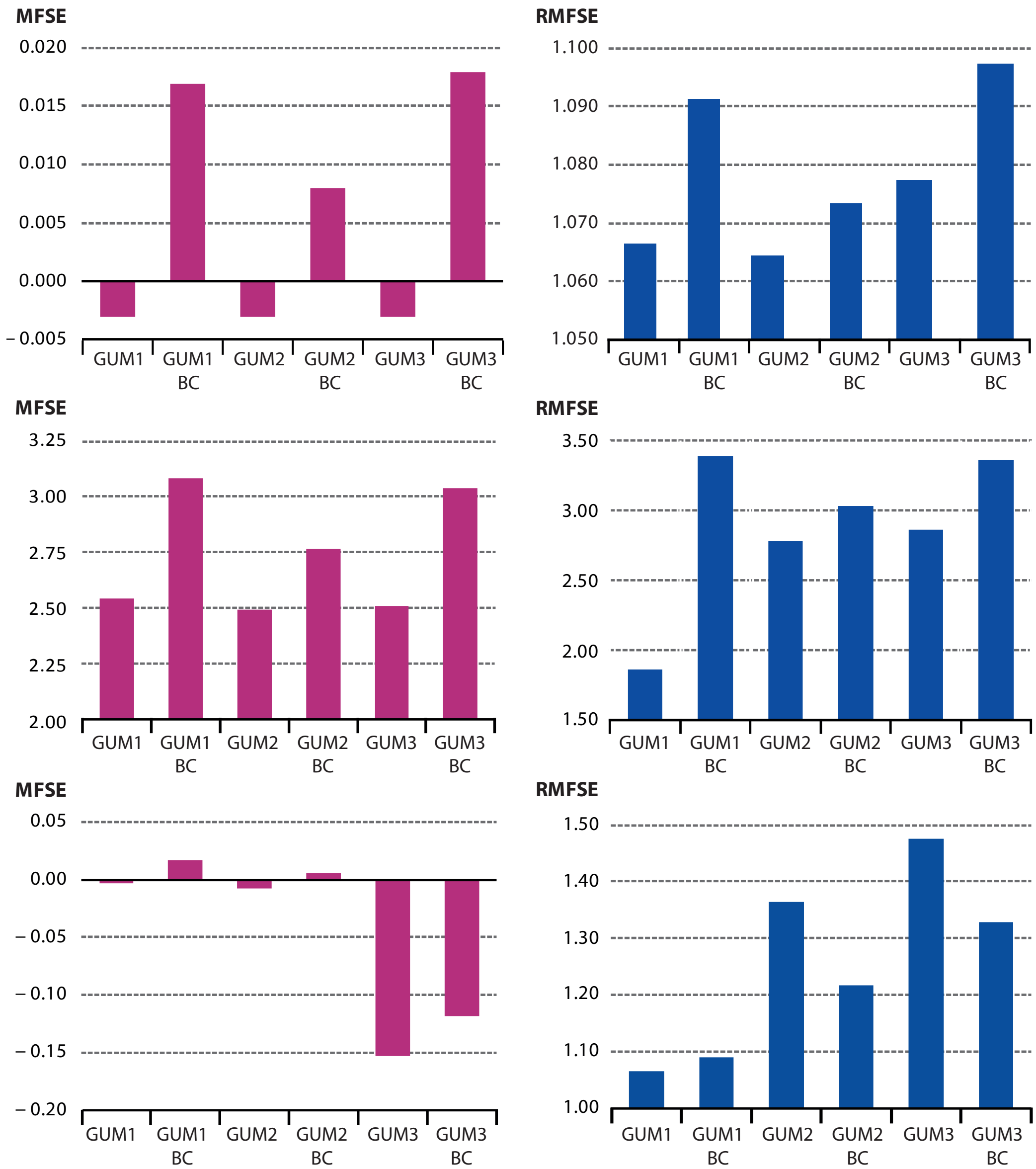
4.3.4. CONSIDERING BIAS CORRECTION

Hendry and Krolzig (2005) present a correction to reduce search bias in parameter estimates. The suggested procedure for bias correction takes into account that conditional on having been selected, the absolute values of parameter estates of irrelevant variables are upward biased. This procedure has been applied to parameter estimates in GUM 1 - GUM 3. Note that bias correction has only been applied to parameter estimates that are subject to search biases so any forced variables are not corrected (i.e., the intercept term and $x_1 - x_3$ in GUM 2).

Insight 7 states that correcting parameters for search is most beneficial if location shifts occur in genuinely irrelevant variables. Across the different breaking dynamics, we observe the potential benefit to be yielded from bias correction. Absent location shifts, correcting parameters for search uniformly worsens nowcast accuracy as displayed in plot (a) and (b) of Figure 4. At $\alpha =$ 0.05, and with only 4 irrelevant variables, bias correction mostly acts on the coefficients of the relevant variables. Bias correction of relevant parameters, however, has been found to increase their mean squared errors. Moreover, given the identical DGPs for the disaggregates, an irrelevant variable may in fact function as a close substitute for a relevant one. Together, these arguments can motivate why bias correcting, and hence setting some parameter estimates to 0, is not found to improve nowcast accuracy in this set-up. This also applies to simulations in which relevant and irrelevant disaggregates break by identical amounts, since the latter remain good proxies for relevant disaggregates and search costs are small, as can be seen in plots (c) and (d) of Figure 4. Of course, very high frequency data may reveal that shifts have occurred in some aspects of the economy and allow adjustments thereto, perhaps using a robust device such as intercept correction: see Castle *et al.* (2017).

With different breaking dynamics, and in line with insight 7, bias correction improves nowcast accuracy in terms of MFE and RMSFE as illustrated in plots (e) and (f) of Figure 4. In GUM 2, its impact on the MFE remains marginal, while the RMSFE is lowered by 10 %. With selection over all variables in GUM 3, the impact of bias correction becomes most pronounced. The MFE falls by over 20 % in magnitude, while the RMSFE declines by 10 %. With a higher ratio of irrelevant to relevant variables, and a looser selection criterion the usefulness of bias correction is likely to be higher.

Figure 4: MFE and RMSFE: Bias correction (BC)



Source: Authors' calculations.

5. Nowcasting consumption expenditure

Having demonstrated the relevance of the seven insights from the nowcast taxonomy in a simulation exercise, it is of interest to consider how they are applicable in practice. For this purpose, we present an ex-post nowcast of growth in final consumption expenditure by households and non-profit institutions serving households (NPISH) over two horizons, 2008Q1 -2010Q4 and 2015Q1-2016Q3. For the nowcast of consumption expenditure growth, a small-scale model with two disaggregated series is used. The Office for National Statistics (ONS) uses three approaches to compute GDP: the output, income and expenditure approach. The expenditure approach is obtained from the sum of final consumption expenditure by households, NPISH and government on goods and services, gross capital formation, and net exports of goods and services. This additivity property can be replicated based on current price data on the components of the expenditure approach as published by the ONS. Consequently, the nowcasts of consumption expenditure could be combined with data on the other components of the expenditure approach, or their nowcasts, to arrive at a nowcast of GDP growth. This section commences with a presentation of the data in Section 5.1, and a discussion of seasonals and trends in Section 5.2 to motivate the data transformations used in this empirical example. Subsequently, forecasts of the two disaggregated series are provided in Sections 5.3 and 5.4, and the nowcasts of consumption expenditure growth are presented in Section 5.5. Section 5.6 interprets the empirical findings in light of the nowcast taxonomy.

5.1. The data

For the nowcast of growth in final consumption expenditure, $\Delta c_t = (C_t - C_{t-1})/C_{t-1}$, we consider two disaggregates: the index for retail trade (IR) and the number of newly registered passenger cars (CAR). IR is one of the two primary measures in the computation of final consumption expenditure [9]. Approximately 40 % of average weekly household expenditure is on non-durable retail sales, and 4 % on car purchases as computed from ONS (2015), which may proxy for durable consumption expenditure. The index of retail sales and passenger car registrations become available with a lag of one month, and therefore need to be forecast to fill in ragged edges in the dataset. The forecasting models use four survey indicators on consumer confidence (CCI), retail trade confidence (RCI), service sector confidence (SCI) and economic sentiment (ESI). All data used in this empirical application are final estimates, so that we may abstract from measurement errors and revisions in the following analysis. The indicators are released at the end of the month that they refer to, and hence are treated as being available in real-time.

Table 7: Nowcasting horizons

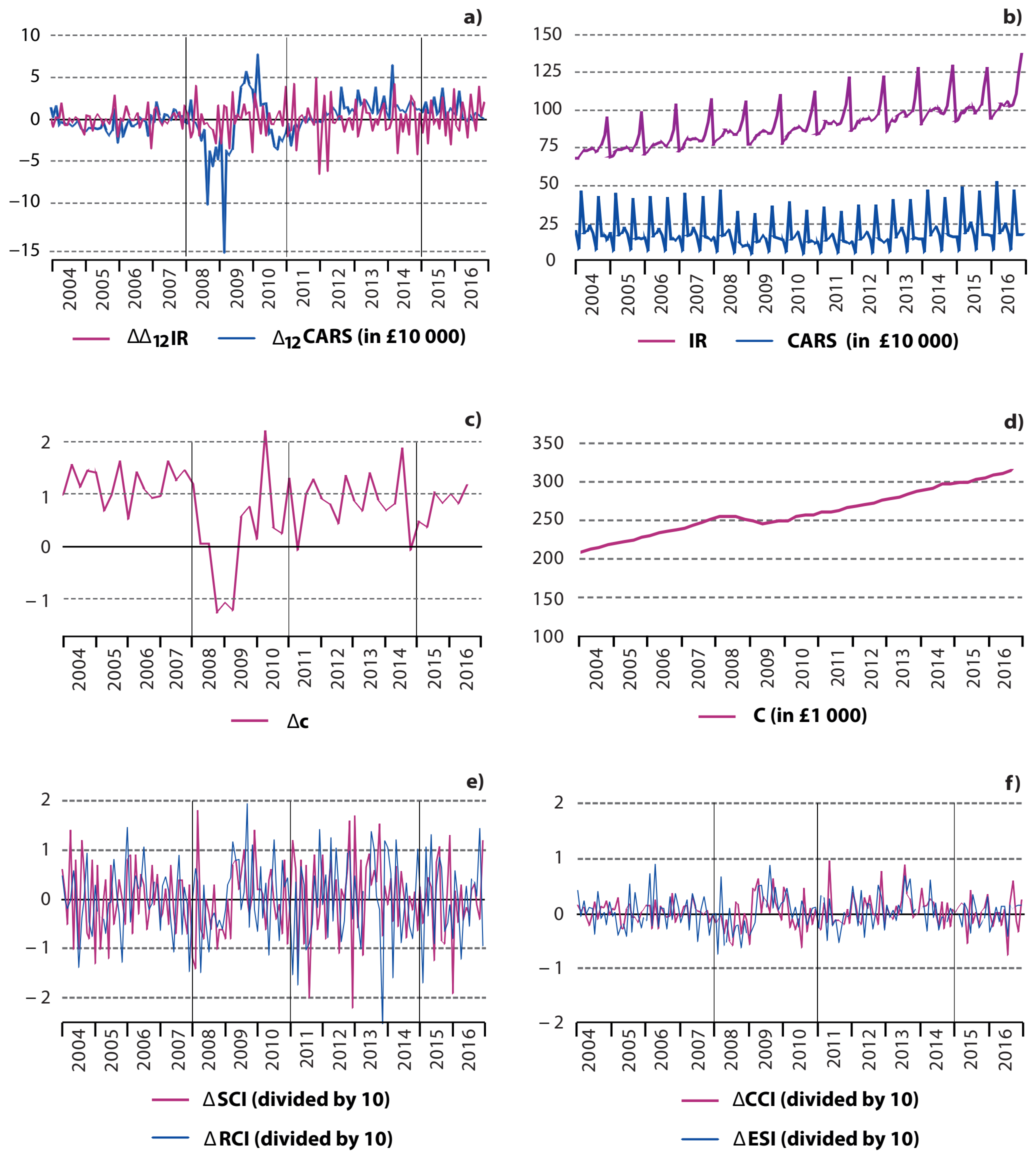
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Q1		$H1_{Q1}$	$H2_{Q1}$	$H3_{Q1}$								
Q2					$H1_{Q2}$	$H2_{Q2}$	$H3_{Q2}$					
Q3								$H1_{Q3}$	$H2_{Q3}$	$H3_{Q3}$		
Q4	$H3_{Q4}$										$H1_{Q4}$	$H2_{Q4}$

Subscripts refer to the respective quarter.

[9] See https://www.ons.gov.uk/economy/nationalaccounts/satelliteaccounts/qmis/consumertrendsqmi.

For the nowcasting period 2008Q1-2010Q4 we consider the in-sample period 2004Q1-2007Q4. In 2008-2009 the UK economy was subject to two large economic shocks, the deterioration in the functioning of financial markets and a fall in international trade, and went into its worst recession since the Great Depression, see Millard (2015). The recession is visible in the plot of consumption growth in panel (c) of Figure 5. The performance of the nowcasting framework during this turbulent period is then compared to nowcasts over the more stable quarters 2015Q1-2016Q3 with in-sample period 2004Q1-2014Q4. We consider three nowcasting origins,

Figure 5: **Data series over the period 01/2004–09/2016**



Notes: Δc: growth rate of C.
Source: Authors' calculations.

H1-H3, per quarter to be nowcast according to the structure in Table 7. The three nowcast horizons have been chosen to examine the impact of the accumulation of information on retail sales and passenger car registrations throughout the quarter to be nowcast. The first nowcast is estimated at *H1*, the second month of the reference quarter. At this point, data on retail sales and car registrations for the first month of the reference quarter are available, but ragged edges have to be filled in at the nowcast origin *H1*. At *H2*, data on the disaggregates for the first and second month of the reference quarter have been released, and forecast values of retail sales and car registrations for the third month are included. In the month following the end of the reference quarter, at *H3*, data for all three months is contemporaneously available. This coincides with the ONS release date of the initial estimate of GDP growth based on the production approach.

5.2. Trends and seasonals

It is a well established empirical fact that data on consumption is a unit-root process; the paper by Davidson *et al*. (1978) treats the statistical modelling of the aggregate consumption function, and laid the ground for subsequent work on co-integration. As co-integrated data introduces new statistical features in the face of structural breaks such as co-breaking, we choose to model the non-integrated growth rate of final consumption expenditure. For analogous reasons, we decide to model monthly changes of the retail sales index ([10]). Similar to working with first differences, taking growth rates implies that long-run information is dropped from the data. Importantly, structural breaks in the levels are turned into impulses in the first-differenced data. The decision to work with growth rates and first differences therefore ensures that standard asymptotic theory and regression techniques apply, however, it comes at the cost of information loss and makes the detection of structural breaks more challenging.

In addition to unit roots, economic time series exhibit seasonality. Seasonal effects in the data refer to systematic calendar-related fluctuations. Retail sales, for example, rise each year around Christmas. Other examples are effects due to weather, due to administrative measures such as the start of the school year, or variations in the length of months. The data may also include calendar effects which relate to factors that do not occur in the same month/quarter every year such as changing numbers of trading days, or moving holidays, e.g., Easter. Seasonal and calendar effects must be accounted for to make consecutive periods comparable. Seasonality may be modelled explicitly by including seasonal indicators. Alternatively, a simple method for removing seasonality is yearly differencing. Further, statistical offices provide seasonally adjusted data. The ONS uses the software X12-ARIMA to remove seasonality, see ONS (2007). The advantage of X12-ARIMA over seasonal indicators and yearly differencing is that calendar effects are accounted for before removing any systematic seasonality. However, it is also used to adjust for extreme values and outliers, exacerbating the loss of information on dynamics in the data. To preserve the additivity property, we model growth rates of final consumption expenditure by households and NPISHs in current prices using seasonally adjusted data. For retail sales and car registrations we work with non-seasonally adjusted data and take yearly differences. The data appendix summarises the relevant transformations undertaken to remove unit roots and seasonalities from the data series.

([10]) In order to avoid biased results on ADF tests due to the recession during 2008-2010, we perform ADF tests on the small sample 01/2004-12/2007. Given the low power of ADF tests in small samples, we extend the sample to include data up to 1997, however, there is no convincing evidence for the stationarity of the retail sales index.

5.3. Forecasting the retail sales index

In order to fill in missing data in the differenced retail sales index at the nowcast origins *H1* and *H2*, a forecasting model for retail sales has to be specified. *Autometrics* is used for variable selection in-sample based on the following initial GUM:

$$(12)\ \Delta\Delta_{12}\mathrm{IR}_{tm} = \beta_0 + \sum_{i=0}^{12} (\beta_{s,i}\Delta\mathrm{SCI}_{tm-i} + \beta_{R,i}\Delta\mathrm{RCI}_{tm-i} + \beta_{C,i}\Delta\mathrm{CCI}_{tm-i} + \beta_{E,i}\Delta\mathrm{ESI}_{tm-i})$$

$$+ \sum_{i=1}^{12}\beta_{I,i}\Delta\Delta_{12}\mathrm{IR}_{tm-i} + \sum_{k=2}^{Tm}\beta_{1,k}d_k + v_{t_m} \quad \text{for } t_m = 1, ..., T_m.$$

d_k is an indicator taking the value 1 at time $t_m = k$ to implement impulse-indicator saturation (IIS: see Hendry *et al.* (2008) and Johansen and Nielsen (2009)) to automatically detect unknown outliers and structural breaks at any point in the sample. We use the default significance level of 0.01 for diagnostic tests and the intercept is forced to be in the final model specification. Pre-search lag reduction was turned off. In order to account for potential serial correlation, 12 endogenous lags are included in the GUM. Since survey indicators may be leading movements in final consumption expenditure, 12 lags are incorporated into the GUM ([11]).

We allow for bias correction. Since correction for search costs has been shown to reinforce the Hurwicz bias, endogenous lags are excluded from bias correction. Parameter estimates of IIS are unbiased and are therefore not subject to bias correction. IIS and variable selection are applied in several steps. These steps are performed separately for the two in-sample periods, t_m = 01/2004, ..., 12/2007 and t_m = 01/2004, ..., 12/2014, and the respective nowcast horizons h_m = 01/2008, ..., 12/2010 and h_m = 01/2015, ..., 09/2016.

(i) Select over indicators at $\alpha = 0.001$, including an intercept.

(ii) Select over variables and any retained indicators from Step (i) at $\alpha = 0.05$.

(iii) For each new observation over the nowcasting horizon, $h_m = 1, ..., H_m$, re-estimate the parameters and test for the significance of an indicator for the last observation, $d_{T_m+h_m}$, at $\alpha = 0.05$. If significant, retain the indicator.

The selected model is kept the same across the nowcast horizon to reduce model uncertainty.

The retail sales index was differenced on an annual basis to remove systematic seasonality, with monthly differences of the annual change computed to remove the unit root. The remaining variation captures deviations from the long-run trend in retail sales. These irregular fluctuations may arise due to trend breaks such as the financial crisis. The financial crisis and recession is difficult to spot in Figure 5, panel (b). Indeed, Anagboso and McLaren (2009) confirm that retail sales remained strong over the economic downturn. Consequently, there are erratic fluctuations in the differenced retail sales index over 01/2008–12/2010 in panel (a) of Figure 5, with tentative evidence of location shifts having been turned into impulses over this period.

Equation (13) shows the preferred model specification of retail sales:

$$(13)\ \widehat{\Delta\Delta_{12}\mathrm{IR}}_{tm} = \underset{(0.11)}{0.05} - \underset{(0.10)}{0.25}\Delta\Delta_{12}\mathrm{IR}_{tm-1} - \underset{(0.11)}{0.29}\Delta\Delta_{12}\mathrm{IR}_{tm-5} + \underset{(0.18)}{0.42}\Delta\mathrm{SCI}_{tm-5} - \underset{(0.18)}{0.80}\Delta\mathrm{SCI}_{tm-8}$$

$$+ \underset{(0.17)}{0.71}\Delta\mathrm{SCI}_{tm-10} - \underset{(0.66)}{4.9}\Delta\mathrm{CCI}_{tm-7} - \underset{(0.77)}{3.1}\Delta\mathrm{CCI}_{tm-12}$$

([11]) It was attempted to augment the GUM with lags 1-4 of consumption growth to proxy for real personal disposable income, yet they were not retained and therefore omitted.

$$\hat{\sigma} = 0.63;\ R^2 = 0.75;\ \text{Obs.} = 35;\ \chi^2_N = 0.95;\ F(7, 27) = 0.00;$$
$$F_{AR}(3, 27) = 0.64;\ F_{ARCH}(3, 27) = 0.64;\ F_{Het}(7, 27) = 0.63$$

Numbers in parentheses refer to standard errors, $\hat{\sigma}$ is the standard error of the estimated equation, the p-value of the F-test on the joint significance of the included regressors, denoted F, as well as the R^2. p-values of mis-specification tests are also reported for the F-test of residual autocorrelation, F_{AR}, autoregressive conditional heteroskedasticity, F_{ARCH}, normality, χ^2_N, and heteroskedasticity including squares and cross-products, F_{Het}. Indicators are retained for the period 01/2008-12/2010.

The forecasts and forecast errors are plotted in Figure 6, panels (a) and (b) respectively. The 95 % forecast intervals for conventional forecasts are plotted in panel (a), along with the forecasts resulting from the bias corrected coefficient estimates. Table 8 reports the RMSFEs. As the uncorrected and bias corrected forecasts are very similar, confidence intervals are not reported for the bias corrected forecasts. Bias correction marginally increases the RMSFE despite a large number of retained regressors, and location shifts in retained variables. While bias correction reduces the RMSFE by reducing the bias related to irrelevant variables, its overall impact on the squared nowcast error, or in this context forecast error, may be ambiguous due to its effect on relevant variables. This may serve as a possible explanation for this nevertheless counterintuitive finding. There is evidence of forecast failure over this volatile period.

Table 8: RMSFE in forecasting models

	IR		CAR	
	No BC	BC	No BC	BC
01/2008-12/2010	3.02	3.03	5.47	5.48
01/2015-09/2016	1.57	1.57	1.30	1.19

BC= bias correction

Source: Authors' calculations.

The results can be contrasted with the preferred model selected over 01/2004-12/2014, reported in (14):

$$(14)\quad \widehat{\Delta\Delta_{12}\text{IR}_{t_m}} = \underset{(0.12)}{0.07} - \underset{(0.08)}{0.50}\Delta\Delta_{12}\text{IR}_{t_m-1} - \underset{(0.08)}{0.25}\Delta\Delta_{12}\text{IR}_{t_m-2} + \underset{(0.07)}{0.25}\Delta\text{IR}_{t_m-11}$$
$$+ \underset{(0.17)}{0.51}\Delta\text{SCI}_{t_m-2} + 4 \text{ Impulse Indicators}$$

$$\hat{\sigma} = 1.31;\ R^2 = 0.57;\ \text{Obs.} = 119;\ \chi^2_N = 0.62;\ F(8, 110) = 0.00;$$
$$F_{AR}(7, 103) = 0.73;\ F_{ARCH}(7, 105) = 0.81;\ F_{Het}(8, 106) = 0.90$$

Figure 5 shows that changes in retail sales exhibit more regular movements over 01/2015-09/2016. Only one significant outlier in the period 01/2008-12/2010 is retained at the more conservative significance level of $\alpha = 0.001$, confirming the stability of retail sales over the financial crisis and recession. Significant indicators are found for 12/2011 and 01/2012, in line with more pronounced fluctuations during the end of 2011 and start of 2012. At the time, negative GDP growth indicated that the UK might be heading into a double-dip recession ([12]). In July 2012, the summer olympics started, providing a boost to retail sales. During the nowcast

([12]) See http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171766_263951.pdf.

horizon one positive outlier is retained at 12/2015 and 01/2016, capturing the increase in spending before Christmas.

The final model in Figure 7, panels (a) and (b), does a reasonably good job at predicting the sign of changes in retail sales. In line with the significant outliers during the Christmas period, the fit of forecasts deteriorates during this period though forecast failure can be avoided. Bias correction does not change the RMSFE

5.4. Forecasting passenger car registrations

To forecast yearly differences in passenger car registrations, we use the same steps (i)-(iii) for variable selection and outlier detection as described in the previous section. The initial GUM for CAR includes the following variables:

$$\Delta_{12}\mathrm{CAR}_{t_m} = \gamma_0 + \sum_{i=0}^{12} (\gamma_{s,i}\Delta\mathrm{SCI}_{t_m-i} + \gamma_{R,i}\Delta\mathrm{RCI}_{t_m-i} + \gamma_{C,i}\Delta\mathrm{CCI}_{t_m-i} + \gamma_{E,i}\Delta\mathrm{ESI}_{t_m-i})$$

$$+ \sum_{i=1}^{12} \gamma_{CA,i}\Delta_{12}\mathrm{CAR}_{t_m-i} + \sum_{k=2}^{T_m}\gamma_{1,k}d_k + v_{t_m} \qquad \text{for } t_m = 1, \ldots, T_m.$$

In Figure 5, a drop in the twelfth differences of passenger car registrations during 2008-2009 is evident, which is matched by the confidence indicators on the retail trade and service sector. Car registrations also indicates a pronounced outlier in 03/2009. In the second half of 2009, yearly differences in car registrations revert to being positive as a result of the car scrappage scheme implemented by the government. This scheme allowed for the scrappage of 400 000 old vehicles, and provided a £2 000 incentive to buy a new car, see Crossley, Leicester, and Levell (2010). This led to a significant short-term increase in car registrations. The statistical significance of outliers during this period is confirmed in the forecasting model selected at $\alpha = 0.05$. The clustered occurrence of outliers during the second half of 2008 and 2009 as well as during 2010 provides strong evidence for the relevance of structural breaks during this nowcast horizon.

The selected forecasting model is:

$$(15) \quad \widehat{\Delta_{12}\mathrm{CAR}}_{t_m} = \underset{(0.10)}{0.20} + \underset{(0.08)}{0.33}\,\Delta_{12}\mathrm{CAR}_{t_m-3} + \underset{(0.09)}{0.72}\Delta_{12}\mathrm{CAR}_{t_m-6} + \underset{(0.13)}{0.51}\,\Delta\mathrm{SCI}_{t_m-3}$$

$$+ \underset{(0.11)}{0.63}\,\Delta\mathrm{SCI}_{t_m-12} - \underset{(0.46)}{1.5}\,\Delta\mathrm{CCI}_{t_m-2} - \underset{(0.40)}{1.4}\,\Delta\mathrm{CCI}_{t_m-7} + \underset{(0.45)}{1.9}\,\Delta\mathrm{CCI}_{t_m-9}$$

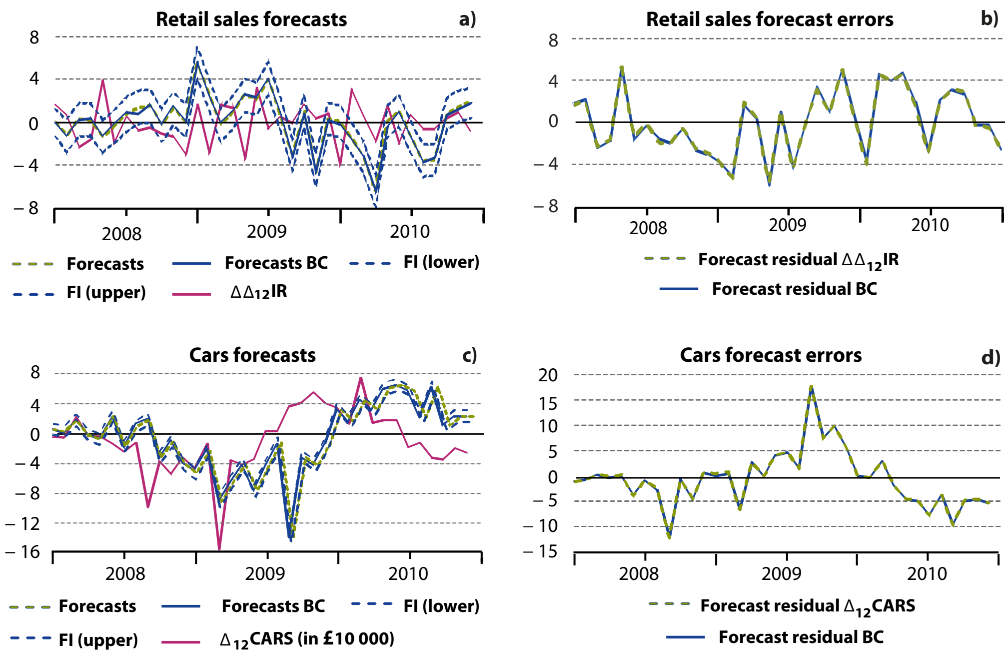
$$- \underset{(0.48)}{1.3}\,\Delta\mathrm{CCI}_{t_m-12} + \underset{(0.22)}{0.76}\,\Delta\mathrm{ESI}_{t_m} + \underset{(0.22)}{1.2}\,\Delta\mathrm{ESI}_{t_m-5} + \underset{(0.23)}{0.94}\,\Delta\mathrm{ESI}_{t_m-11}$$

$$\hat{\sigma} = 0.38;\ R^2 = 0.91;\ \text{Obs.} = 35;\ \chi^2_N = 0.49;\ \mathrm{F}(11, 24) = 0.00;$$

$$\mathrm{F}_{AR}(3, 21) = 0.08;\ \mathrm{F}_{ARCH}(3, 30) = 0.03;\ \mathrm{F}_{Het}(22, 13) = 0.46$$

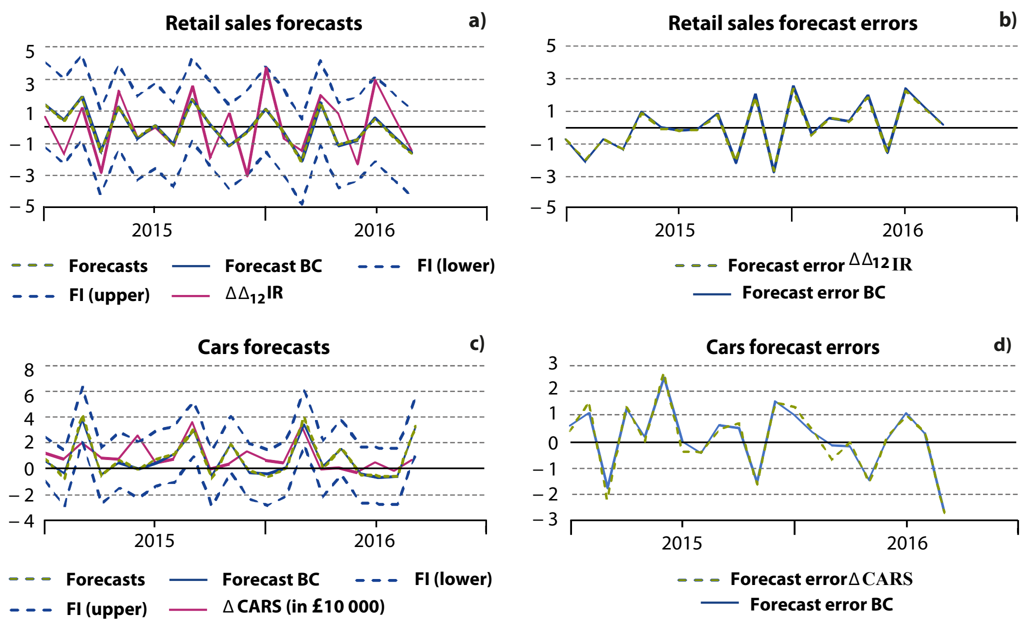
Compared to retail sales, more systematic, rather than purely irregular, fluctuations in the car registrations are retained. The performance of the preferred forecasting model is not enhanced by correcting parameters for search. Figure 6, panels (c) and (d) record uncorrected and bias corrected forecasts, along with the 95% confidence intervals for the forecasts from equation (15). The figure shows evidence of forecast failure over the forecast horizon.

Figure 6: Forecasts and forecast errors for retail sales and cars, 01=2008



Note: IIS is applied recursively as the forecast horizon advances for a xed model formulation, so estimated parameters and error variances do not change over the forecast horizon, hence the small forecast intervals for Cars.
Source: Authors' calculations.

Figure 7: Forecasts and forecast errors for retail sales and cars, 01/2015 - 09/2016



Source: Authors' calcuations.

Looking at the plot of car registrations, it is clear that the period of 01/2015 - 09/2016 is more stable. In order to account for the scrappage scheme, we retain an indicator variable covering the implementation period 05/2009 - 02/2010 in the forecasting model. IIS suggested systematically larger growth rates in car registrations in the months March and September from 2012 onwards. As stated by the ONS, car registrations are usually higher in Q1, Q3 than Q2, Q4, corresponding to the release of new number plates in the months March and September, see Grove (2012). We therefore specify, and retain, the indicator variable *Plate*, which takes the value 1 for March and September in the years 2012 - 2016, in the final forecasting model (16) over the full sample period:

$$(16)\quad \Delta_{12}\widehat{\mathrm{CAR}}_{t_m} = \underset{(0.09)}{-0.12} + \underset{(0.04)}{0.30}\,\Delta_{12}\mathrm{CAR}_{t_m-1} + \underset{(0.04)}{0.31}\,\Delta_{12}\mathrm{CAR}_{t_m-2} + \underset{(0.04)}{0.13}\,\Delta_{12}\mathrm{CAR}_{t_m-7}$$

$$+ \underset{(0.11)}{0.27}\,\Delta\mathrm{RCI}_{t_m-7} - \underset{(0.12)}{0.21}\,\Delta\mathrm{RCI}_{t_m-10} + \underset{(0.31)}{1.65}\,\Delta\mathrm{CCI}_{t_m-5} - \underset{(0.35)}{0.92}\,\Delta\mathrm{CCI}_{t_m-10}$$

$$+ \underset{(0.35)}{0.68}\,\mathrm{Scrappage}_{t_m} + \underset{(0.46)}{3.69}\,\mathrm{Plate}_{t_m} + 6\ \mathrm{Impulse\ Indicators}$$

$$\hat{\sigma} = 0.83;\quad R^2 = 0.92;\quad \mathrm{Obs.} = 120;\quad \chi^2_N = 0.00;\quad \mathrm{F}(17, 102) = 0.00;$$

$$\mathrm{F}_{AR}(7, 97) = 0.72;\quad \mathrm{F}_{ARCH}(7, 105) = 0.002;\quad \mathrm{F}_{Het}(13, 97) = 0.84$$

The large number of retained indicators over 2008 - 2010 underline the importance of structural breaks in yearly car registrations during that period. The bias corrected model provides more accurate forecasts of car registrations. Note the large fall in RMSFEs compared to the horizon 01/2008 - 12/2010 due to the absence of structural breaks with only three forecasts falling marginally outside the 95 % interval in Figure 7.

5.5. Nowcasting final consumption expenditure

To arrive at a nowcast of final consumption expenditure, an equation linking the differenced retail sales and car registrations to consumption expenditure needs to be specified. As described in Section 5.1, the monthly data are separated into three blocks r^1, r^2, r^3 to match the quarterly frequency of final consumption expenditure. These blocks refer to the first, second and third month of the quarter. Due to lack of knowledge of the DGP a more elaborate GUM compared to the Monte Carlo simulations is specified. Beyond the contemporaneous values, the GUM includes two lags of the disaggregates, and four lags of consumption expenditure growth, Δc, to capture potential serial correlation, as well as IIS to help achieve a congruent model specification. The steps (i)-(iii) of variable selection and outlier selection remain unchanged; outlier selection in-sample is performed at $\alpha = 0.001$, retaining an intercept. Variable selection including any retained indicators in-sample is performed at the significance level $\alpha = 0.05$. The same significance level holds for outlier detection over the nowcast horizon. The intercept and all contemporaneous values of the disaggregated components are retained in the final model specification in order to be able to evaluate the performance of bias correction for nowcasting.

At *H1*, the GUM takes the following form, with contemporaneous values of differences in retail sales and in car registrations of blocks r^1, r^2 being forced to be included in the final model. The nowcasting models for the three origins are selected in-sample, so all data may be treated as contemporaneously available.

$$\Delta c^{H1}_{tq} = \pi_0 + \pi_{r^2,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^2}_{tq}\\ \Delta_{12}CAR^{r^2}_{tq}\end{pmatrix} + \pi_{r^1,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^1}_{tq}\\ \Delta_{12}CAR^{r^1}_{tq}\end{pmatrix} + \sum_{k=2}^{Tq}\delta_{0,k}d_k + \sum_{k=1}^{4}\delta_{C,k}\Delta c_{t_q-k}$$

$$+\sum_{j=1}^{2}\left[\pi_{r^1,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^1}_{tq-j}\\ \Delta_{12}CAR^{r^1}_{tq-j}\end{pmatrix} + \pi_{r^2,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^2}_{tq-j}\\ \Delta_{12}CAR^{r^2}_{tq-j}\end{pmatrix} + \pi_{r^3,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^3}_{tq-j}\\ \Delta_{12}CAR^{r^3}_{tq-j}\end{pmatrix}\right] + \epsilon_{tq}$$

At the nowcasting origins there are ragged edges in $\Delta\Delta_{12}IR^{r^2}_{tq}$ and $\Delta_{12}CAR^{r^2}_{tq}$, which have to be forecast. At *H2*, contemporaneous values of retail sales and car registrations in blocks r^1, r^2, r^3, as well as the intercept, are always retained.

$$(17)\quad \Delta c^{H2}_{tq} = \pi_0 + \pi_{r^3,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^3}_{tq}\\ \Delta_{12}CAR^{r^3}_{tq}\end{pmatrix} + \pi_{r^1,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^1}_{tq}\\ \Delta_{12}CAR^{r^1}_{tq}\end{pmatrix} + \pi_{r^2,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^2}_{tq}\\ \Delta_{12}CAR^{r^2}_{tq}\end{pmatrix}$$

$$+\sum_{k=2}^{Tq}\delta_{0,k}d_k + \sum_{k=1}^{4}\delta_{C,k}\Delta c_{t_q-k} + \sum_{j=1}^{2}\sum_{i=1}^{3}\left[\pi_{r^i,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^i}_{tq-j}\\ \Delta_{12}CAR^{r^i}_{tq-j}\end{pmatrix}\right] + \epsilon_{tq}$$

At *H2*, it is $\Delta\Delta_{12}IR^{r^3}_{tq}$ and $\Delta CAR^{r^3}_{tq}$ that are missing and have to be filled in at the nowcast origin. The GUM for Δc^{H3}_{tq} is identical to that for Δc^{H2}_{tq} in equation (17). For nowcasts at *H3*, data for all three months of the reference quarter is available. Consequently, the final nowcasting model at *H3* is exclusively based on actual data. Again, the contemporaneous values of retail sales and car registrations of blocks r^1, r^2, r^3 are always retained.

$$(18)\quad \widehat{\Delta c^{H1}_{tq}} = \underset{(0.001)}{0.012} + \underset{(0.001)}{0.0016}\Delta\Delta_{12}IR^{r^1}_{tq} + \underset{(0.002)}{0.0016}\Delta\Delta_{12}IR^{r^2}_{tq} + \underset{(0.001)}{0.0018}\Delta_{12}CAR^{r^1}_{tq} - \underset{(0.002)}{0.004}\Delta_{12}CAR^{r^2}_{tq}$$

$$\hat{\sigma} = 0.03;\ R^2 = 0.57;\ Obs. = 12;\ X^2_N(2) = 0.60;\ F(4, 7) = 0.15;$$

$$F_{AR}(1, 6) = 0.83;\ F_{ARCH}(1, 10) = 0.17$$

$$(19)\quad \widehat{\Delta c^{H2}_{tq}} = \widehat{\Delta c^{H3}_{tq}} = \underset{(0.001)}{0.012} + \underset{(0.001)}{0.0017}\Delta\Delta_{12}IR^{r^1}_{tq} + \underset{(0.002)}{0.0005}\Delta\Delta_{12}IR^{r^2}_{tq} + \underset{(0.001)}{0.0004}\Delta\Delta_{12}IR^{r^3}_{tq}$$

$$+\underset{(0.001)}{0.0007}\Delta_{12}CAR^{r^1}_{tq} - \underset{(0.002)}{0.0033}\Delta_{12}CAR^{r^2}_{tq} + \underset{(0.001)}{0.0015}\Delta_{12}CAR^{r^3}_{tq}$$

$$\hat{\sigma} = 0.003;\ R^2 = 0.71;\ Obs. = 12;\ X^2_N(2) = 0.54;\ F(6, 5) = 0.23;$$

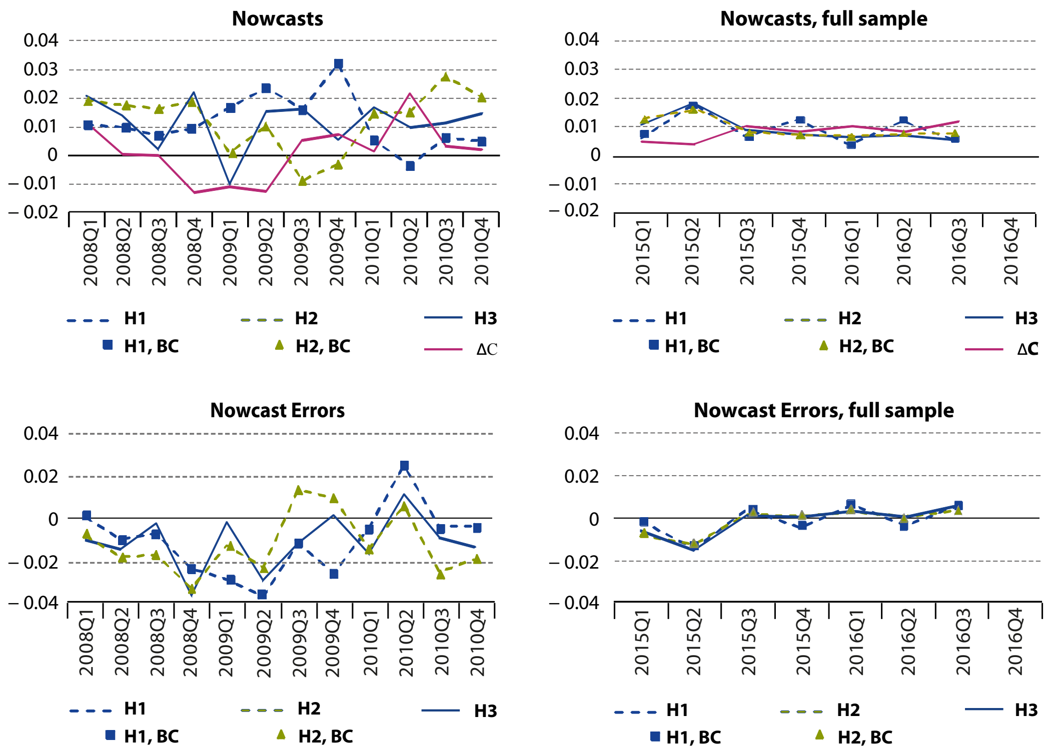
$$F_{AR}(1, 4) = 0.55;\ F_{ARCH}(1, 10) = 0.96$$

The above nowcasting models were selected over the in-sample period 2004Q1-2007Q4 ([13]). Only forced variables and indicators were retained. Indeed, as can be gathered from the size of parameter estimates relative to standard errors, without forcing the contemporaneous values of car registrations and retail sales into the final model, the preferred specification would have just included the intercept. At *H3*, it was also attempted to include the preliminary estimate of GDP that becomes available at this nowcast origin in the GUM. It was not retained in the final model specification. Recall that at *H3* the nowcasts are based entirely on actual data. Out of the

([13]) There are too few values to compute F_{HET}.

conventional nowcasts, *H3* minimises the RMSFE. Further, nowcasts at *H3* may be taken as a reference point for the explanatory power of the disaggregates used to model growth in final consumption expenditure.

Figure 8: Nowcasts and nowcast errors



Source: Authors' calculations.

Note: BC: with bias correction (excluding endogenous lags); H1, H2, H3 refer to the nowcasting horizon. Since nowcasts at H3 are based on actual data, it is only included once.

Nowcasts in 2015Q1-2016Q3 are more accurate as a result of the smaller fluctuations, and more regular behaviour in consumption expenditure growth, as is evident from Figure 8. Differences in RMSFE become negligible (the figures in Table 9 have been multiplied by 100). Searching over indicators in-sample at $\alpha = 0.001$ in step (i), d_{tq}=*2008Q4*, d_{tq}=*2009Q1*, d_{tq}=*2009Q2* are retained at *H1 – H3*. These are dropped once the disaggregates are included.

$$(20) \quad \widehat{\Delta c}^{H1}_{tq} = \underset{(0.001)}{0.003} + \underset{(0.13)}{0.54}\Delta c_{tq-3} - \underset{(0.001)}{0.001}\Delta\Delta_{12}IR^{r^{1}}_{tq} - \underset{(0.001)}{0.0003}\Delta\Delta_{12}IR^{r^{2}}_{tq} + \underset{(0.001)}{0.0005}\Delta\Delta_{12}CAR^{r^{1}}_{tq} + \underset{(0.001)}{0.002}\Delta_{12}CAR^{r^{2}}_{tq}$$

$$\hat{\sigma} = 0.006;\ R^2 = 0.52;\ \text{Obs.} = 40;\ X^2_N(2) = 0.57;\ F(5, 34) = 0.00;$$
$$F_{AR}(3, 31) = 0.13;\ F_{ARCH}(3, 34) = 0.22;\ F_{Het}(10, 29) = 0.99$$

$$(21)\quad \widehat{\Delta c^{H2}_{tq}} = \widehat{\Delta c^{H3}_{tq}} = \underset{(0.001)}{0.003} + \underset{(0.13)}{0.58}\Delta c_{t_q-3} - \underset{(0.0006)}{0.0005}\Delta\Delta_{12}IR^{r^1}_{tq} - \underset{(0.0008)}{0.0001}\Delta\Delta_{12}IR^{r^2}_{tq} + \underset{(0.0006)}{0.0005}\Delta\Delta_{12}IR^{r^3}_{tq}$$
$$- \underset{(0.001)}{0.002}\Delta_{12}CAR^{r^1}_{tq} + \underset{(0.001)}{0.003}\Delta_{12}CAR^{r^2}_{tq} + \underset{(0.0004)}{0.001}\Delta_{12}CAR^{r^3}_{tq}$$

$$\hat{\sigma} = 0.005;\ R^2 = 0.58;\ Obs. = 40;\ X^2_N\ (2) = 0.53;\ F(5, 34) = 0.00;$$
$$F_{AR}(3, 29) = 0.24;\ F_{ARCH}\ (3, 34) = 0.28;\ F_{Het}(14, 25) = 0.99$$

In the longer in-sample period, the third lag of consumption growth becomes significant and is retained at all nowcast origins. The size of the parameter estimates on the endogenous lag relative to the disaggregates highlights that this lag dominates the nowcasts. At all three nowcasting origins, the nowcasts miss the lower growth at the start of 2015 in Figure 8. Nowcasts based on bias-corrected forecasts at *H1* marginally improve the fit of the nowcast, while they have adverse effects on nowcasting accuracy at *H2*. Further, the model at *H3* is found to be less accurate in terms of RMSFE than *H2*. In general, findings in the full sample confirm that the disaggregates seem to have little explanatory power for the direction of change in consumption growth.

Table 9: Percentage RMSFE in nowcasting models of final consumption expenditure

	H1		H2		H3
	No BC	BC	No BC	BC	
01/08-12/10	1.845	1.845	1.773	1.773	1.581
01/15-0/16	0.684	0.669	0.589	0.590	0.651

BC= bias correction. Reported figures have been multiplied by 100.

Source: Authors' calculations.

5.6. Interpretation using the 7 insights

This section interprets the empirical example in light of the seven insights from the nowcast error taxonomy, though such an interpretation remains inconclusive without knowledge of the underlying DGP, and hence lack of information on relevant variables, and breaking dynamics. With this caveat in mind, the empirical example, and in particular the comparison of the two nowcasting periods, remain useful in considering how the intuitions of the theoretical insights can be applied in practice.

Over the volatile period 2008Q1-2010Q4, there is nowcast failure. As discussed, retail sales remained strong over the recent recession, while final consumption expenditure was subject to a substantial downward shift. Though taking the first differences reduces the mean of the retail sales index to a small value, the relationship between differenced retail sales and consumption expenditure arguably shifted over the recession. In consideration of this argument, nowcast failure is consistent with the first insight, which states that a shift in dynamics causes nowcasts to deteriorate if the exogenous variable has a non-zero mean.

Taking into account the transformation of the retail sales index, only car registrations include relevant information on the persistent drop in consumption growth rates over 2008-2010. This does not seem to be sufficient to reliably predict consumption expenditure over 2008Q1-2010Q4. However, there is no systematic nowcast failure in the more stable period, 2015Q1-2016Q3, based on the same information set. This comparison is in line with the second insight, which states that incorrect omission of relevant variables matters for the accuracy of nowcasts if the omitted variables are subject to location shifts, while it does not lead to nowcast failure in stable times. The omission of relevant information in this small-scale nowcasting exercise therefore provides an additional explanation for the deterioration of nowcasts over the great recession.

Equally, the discussion on the forecasts of car registrations in the preceding section made clear that there was systematic forecast failure over 01/2008-12/2010, whilst forecasts were more accurate over the later nowcasting period. We know from the third and fourth insight that failing to accurately forecast the post-shift mean may cause systematic nowcast failure, while there is no mean accuracy gain from accurate forecasts absent location shifts in the exogenous variable. Both insights are consistent with the findings in Tables 8 and 9, and forecast failure likely exacerbated the deterioration in nowcasting accuracy over 2008Q1- 2010Q4.

Insights five and six relate to the retention of irrelevant information, and show that absent mean shifts the inclusion of irrelevant variables is not costly, while it is detrimental to the nowcasting performance if there is a shift in irrelevant variables that is not mirrored in the dependent variable. Given the low explanatory power of the included explanatory variables for movements in consumption expenditure growth, it is questionable whether the variables under consideration may be thought of as 'relevant'. If they are deemed to be irrelevant, then the accuracy of nowcasts was reduced not only due to omission of relevant information, but also due to retention of irrelevant variables that were subject to shifts, providing an additional source of nowcast failure over 2008Q1-2010Q4. As alluded to previously, a limitation to this interpretation is that the distinction between relevant versus irrelevant variables becomes unclear when taking into account that in a real-world settings most variables are correlated with each other, so that the retail sales index, and car registrations may proxy for relevant variables.

Insight seven states that bias correction on irrelevant variables is likely to be more beneficial in turbulent periods. In the empirical example, bias correction was applied to the forecasting models, and was found to improve forecasting performance over the stable period, though its overall effect remained negligible. While the confidence indicators in the forecasting models cannot be thought of as causing movements in retail sales or car registrations, and hence are not part of the respective DGPs, insight seven, in line with simulation evidence in Section 4.3.4, would suggest that the included variables in forecasting models are good proxies for relevant information. In particular with respect to car registrations, it seems equally plausible that any gains in forecasting accuracy from bias correcting parameter estimates of irrelevant variables were overshadowed by forecast failure.

While the empirical example offers a rich set-up for applying the nowcast taxonomy, it should be acknowledged that the performance of the nowcasting framework has been unreliable. From insight four, a robust method of forecasting during the great recession would be required. Further, the current model assumes that the shift in consumption growth may be explained by the disaggregates. This assumption is too restrictive, since accounting for shifts at the aggregate level could alleviate the impact of forecast failure from disaggregates, of the impact of dynamic shifts from insight one, or of shifts at the aggregate level that exceed those in disaggregates.

In addition, the empirical exercise underlines the exigencies towards the data that may be used to successfully nowcast with the suggested framework. Taking first differences results in semi-robust data, from which any regular trends have been removed, and location shifts have been turned into impulses. The lack of information contained in semi-robust data increases the difficulty of nowcasting accurately. A natural step is to consider the nowcasting framework in a co-integrated framework to be able to model variables with unit roots. Moreover, the two disaggregates were modelled using survey or 'soft' data only. While surveys are a timely source, the nowcasting literature is undecided about the explanatory content of soft data, see, inter alia, Antolin-Diaz *et al.* (2017) or Mitchell (2009). Based on similar Eurostat surveys to those used here, Mitchell (2009) finds that only in addition to hard data did soft data improve nowcasts at the onset of the recession. Further he found that, due to their projective nature, surveys became less successful further into the quarter to be nowcast. It seems advisable to include hard data beyond endogenous lags, e.g., appropriately proxying for real personal disposable income. In the evaluation of the empirical application, however, it should be taken into account that consumer expenditure is difficult to model empirically. Under certain assumptions on the economic model, an inter-temporal utility maximisation problem with rational expectations predicts that consumption follows a random walk, see Muellbauer (1994).

6. Conclusion

In providing preliminary estimates of economic aggregates, statistical agencies have to select a nowcasting model from a large number of disaggregates, and deal with missing data of mixed frequency. The nowcast error taxonomy presented in this paper offers a framework for thinking about sources of nowcast errors that may be faced by statistical agencies. The nowcast error taxonomy makes clear that the main sources for systematic nowcast failures are unmodelled location shifts, and underlines the importance of accurate infilling of any missing data that is needed to construct the macroeconomic aggregate. Further, the nowcast error taxonomy highlights the costs associated with omitting relevant disaggregates or retaining irrelevant ones, and showed that these costs are most pronounced if the regressors are subject to location shifts. Mis-specification and estimation uncertainty have comparatively small impacts on nowcast accuracy. The theoretical insights have been confirmed by evidence based on a simulation exercise, and have been applied in an empirical nowcasting exercise of final consumption expenditure by households and NPISH using a retail sales index and passenger car registrations to interpret sources of nowcast failure.

Beyond the seven insights, the nowcast error taxonomy helps to interpret trade-offs involved in variable selection. In practice, the distinction between relevant and irrelevant variables is not known. The trade-off in variable selection thus consists of reducing costs of search by specifying a tight significance level for variable selection to reduce the probability of falsely retaining irrelevant variables, whilst simultaneously increasing costs of inference by increasing the likelihood of falsely rejecting relevant variables. The simulations made clear that depending on out-of-sample breaking dynamics, costs of inference or costs of search may dominate. In simulations with location shifts occurring to all relevant disaggregates, and non-zero correlations between variables, costs of inference dominated costs of search with irrelevant variables proxying for relevant ones. Once asymmetric breaking dynamics were specified in the simulations, the adverse impact of costs of search became apparent. The empirical example highlighted additional challenges, such as non-stationarity of real-world data requiring data transformations that lead to information loss, which further complicate the distinction between relevant and irrelevant information in practice. Equally, relevant variables may have low non-centralities in practice due to noise in the data, impacting on the choice of the optimal significance level in variable selection. In line with the difficulty of isolating relevant information in variable selection, results on bias correction after model selection were mixed, both in simulations and in the empirical example. Beyond insights into nowcast failure, the taxonomy may thus inform model selection in nowcasting.

Data Appendix

Table 10: Data Sources

Label	Description	Source	Period	Transfor-mation	Release lag
C_t	Household & NPISH final consumption expenditure in million (CPSA)	[1]	1997Q1-2016Q3	5	4
IR_t	Index for value of retail sales, all retailing incl. automotive fuel	[1]	1997M1-2016M12	1; 2	1
CAR_t	New passenger car registrations UK	[2]	1997M1-2016M12	2; 3	1
SCI_t	Service sector confidence indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0
CCI_t	Consumer confidence indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0
RCI_t	Retail trade confidence indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0
ESI_t	Economic sentiment indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0

SA - seasonally adjusted, CP - current price.
Transformations: (1) $= \Delta X_t = X_t - X_{t-1}$; (2) $= \Delta_{12} X_t$; (3) $= X_t/10\,000$; (4) $= \Delta X_t/10$; (5) $= \Delta X_t/X_{t-1}$.
The transformed time series were tested for unit roots using Dickey-Fuller tests including available data up to 12/2007. 12 lags were included for the monthly data, and 4 lags for quarterly data..

Sources: [1] ONS (ons.gov.uk), [2] Eurostat database (http://ec.europa.eu/eurostat/data/database).

Acknowledgements

Financial support from the Economic and Social Research Council (award ES/J500112/1), the Robertson Foundation (award 9907422), Institute for New Economic Thinking (grant 20029822) and Statistics Norway through Research Council of Norway Grant 236935 are all gratefully acknowledged.

References

Anagboso, M. and C. McLaren (2009), 'The impact of the recession on retail sales volumes', *Economic and Labour Market Review*, Volume 3. Office for National Statistics.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017), 'Hard vs. Soft Data in Nowcasting GDP: A Comment', https://drive.google.com/file/d/0B7xdKeAL7xDbWUVhMlVJUEtoeG8/view.

Castle, J. L., D. F. Hendry, and O. Kitov (2017), 'Forecasting and nowcasting macroeconomic variables: A methodological overview', *Handbook on Rapid Estimates*, UN and Eurostat.

Castle, J. L. and N. Shephard (Eds.) (2009), *The Methodology and Practice of Econometrics*, Oxford: Oxford University Press.

Clements, M. P. and D. F. Hendry (1998), 'Forecasting economic processes', *International Journal of Forecasting*, 14, pp. 111–131.

Clements, M. P. and D. F. Hendry (2006), 'Forecasting with breaks', in G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Econometrics on Forecasting*, pp. 605– 657, Amsterdam: Elsevier.

Davidson, J. E. H., D. F. Hendry, F. Srba, and S. Yeo (1978), 'Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom', *The Economic Journal*, pp. 661–692.

Doornik, J. A. (2009), 'Autometrics', See Castle and Shephard (2009), pp. 88–121.

Ericsson, N. R. (2017), 'Economic forecasting in theory and practice: An interview with David F. Hendry', *International Journal of Forecasting,* 33, pp. 523–542.

Crossley, F. T., A. Leicester, and P. Levell (2010), 'Fiscal stimulus and the consumer', in R. Chote, C. Emmerson, and J. Shaw (Eds.), IFS Green Budget. IFS.

Giannone, D., L. Reichlin, and S. Simonelli (2009), 'Nowcasting Euro Area economic activity in Real-Time: The role of confidence indicators', *National Institute Economic Review,* 210, pp. 90–97.

Giannone, D., L. Reichlin, and D. Small (2008), 'Nowcasting GDP and Inflation: The real-time informational content of macroeconomic data', *Journal of Monetary Economics,* 55, pp. 665–676.

Grove, J. (2012), 'Vehicle licensing statistics, Great Britain: Quarter 3 2012', Technical report, Department for Transport.

Hendry, D. F. and J. A. Doornik (2014), *Empirical Model Discovery and Theory Evaluation*, Cambridge, Mass., MIT Press.

Hendry, D. F. and K. Hubrich (2011), 'Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate', *Journal of Business and Economic Statistics*, 29, pp. 216–227.

Hendry, D. F., S. Johansen, and C. Santos (2008), 'Automatic selection of indicators in a fully saturated regression', *Computational Statistics,* 33, pp. 317–335, Erratum, pp. 337–339.

Hendry, D. F. and H.-M. Krolzig (2005), 'The properties of automatic GETS modelling', *The Economic Journal*, 115, pp. C32–C61.

Hendry, D. F. and G. E. Mizon (2012), 'Open-model forecast-error taxonomies', in X. Chen and N. R. Swanson (Eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*, pp. 219–240. New York: Springer.

Johansen, S. and B. Nielsen (2009), 'An analysis of the indicator saturation estimator as a robust regression estimator', in Castle and Shephard (2009), pp. 1–36.

Millard, S. (2015), 'The Great Recession and the UK labour market', *Working Paper,* 566, Bank of England.

Mitchell, J. (2009), 'Where are we now? The UK recession and nowcasting GDP growth using statistical models', *National Institute Economic Review,* 209(1), 60–69.

Muellbauer, J. (1994), 'The assessment: consumer expenditure', *Oxford Review of Economic Policy* 10(2), pp. 1–41.

ONS (2007), 'Guide to seasonal adjustment with X-12-ARIMA', Technical report, Office for National Statistics.

ONS (2015), 'Family spending: 2015', Technical report, Office for National Statistics.

3

Advancements in measuring intangibles for European economies

CAROL CORRADO [1], JONATHAN HASKEL [2], MASSIMILIANO IOMMI [3], CECILIA JONA-LASINIO [4], MATILDE MAS [5] AND MARY O'MAHONY [6]

Abstract: This paper provides an overview of the methods developed so far to measure intangible assets and to evaluate their contribution to economic growth. We review the work carried out under the INTAN-Invest research collaboration and the SPINTAN EU Funded project to generate cross-country estimates of intangible investments coherent with the Corrado, Hulten and Sichel (2005) model. We also provide a summary of recent empirical evidence on the role of intangible capital as a driver of growth across industries and sectors in European countries and the US as an illustration of uses of the intangibles assets data.

JEL codes: O30, O47, E22

Keywords: Intangible investments, productivity growth, innovation

[1] The Conference Board and Georgetown Center on Business and Public Policy.
[2] Imperial College Business School, Centre for Economic Policy Research and IZA Institute of Labor Economics.
[3] International Monetary Fund.
[4] ISTAT (Italian National Institute for Statistics) and LUISS Lab of European Economics, Rome.
[5] University of Valencia and Valencian Institute of Economic Research.
[6] King's College London and National Institute of Economic and Social Research, corresponding author.

1. Introduction

Investments in intangible assets are widely recognized as major determinants of innovation, growth and employment in the 'knowledge economy.' Endogenous growth models emphasize that knowledge and skills are important determinants of growth and stress that knowledge spillovers generate persistent growth (e.g. Romer, 1986; Lucas, 1988). The importance of R&D and innovation was also explicitly recognized in the 'Lisbon process', and in its successor the 'Europe 2020' agenda, aimed at improving the growth and employment performance of the EU.

Corrado, Hulten and Sichel (2005, 2009), hereafter CHS, addressed the conceptual problem of defining intangible assets using an inter-temporal framework e.g. Weitzmann (1976, 2003). The CHS analysis leads to the conclusion that 'any use of resources that reduces current consumption in order to increase it in the future qualifies as investment' and that all types of capital should be treated symmetrically. Therefore, 'investment in knowledge capital should be placed on the same footing as that of investment in plant and equipment' (Corrado *et al*. (2005), p. 19 and Corrado *et al.* (2009), p. 666). A convenient consequence of the CHS approach and its emphasis on the symmetric treatment of all assets is also that one does not have to worry too much about defining 'intangibles' by way of specific characteristics. Rather it is more important to determine whether a spending type meets the test of being a current outlay that enhances the future capacity of producers (and thereby future consumption). Additionally, the CHS approach does not require explicit econometric techniques and rather offers a practical approach to monitoring intangible capital as part of the measurement program carried out by a statistical office (which, after all, already counts investment in some intangibles such as software and R&D).

Building on Lev (2001) and Nakamura (1999, 2001), CHS developed expenditure measures for intangible investment in the United States, classifying intangible capital into three broad categories: computerized information, innovative property, and economic competencies. At that time only software and artistic and entertainment originals were recognized as assets in official guidelines for national accounts. Since then, the national accounts fixed asset boundary has been expanded to include R&D (SNA 2008/ESA 2010).

The empirical understanding of the contribution of intangibles assets to economic performance improved substantially over recent years. A significant research effort generated measures of intangible investment for business sectors for twenty-eight European member states plus the US (INTAN-Invest ([7]), drawing on the COINVEST and INNODRIVE projects ([8])). In addition, industry level estimates of intangible investment were developed as part of the INDICSER project ([9]) (Niebel *et al.* (2016)) and INTAN-Invest has recently incorporated industry estimates into their database (see below). At the same time, researchers in other countries have looked at intangible investment in, for example, Japan (Fukao *et al.* (2009)).

More recently, Corrado *et al.* (2017b), under the SPINTAN ([10]) project, extended and modified the CHS framework for application to the public sector. They proposed the construction of a satellite national account to capture public investments in intangibles at the level and detail needed for modeling the creation and use of knowledge-based capital in a society. Merging the INTAN-Invest and SPINTAN measures of intangibles allows completing the coverage of intangible investment by industry sector, making possible the generation of total economy

([7]) www.intaninvest.net.
([8]) www.coinvest.org.uk; www.innodrive.org.
([9]) www.indicser.net.
([10]) www.spintan.net.

growth accounts with intangibles as productive assets. This is a crucial advancement because policy analysis of an economy's growth and productivity performance requires complete data on both private and public intangible investments.

The aim of this paper is to provide an overview of recent developments in measuring intangible investment in the EU countries and the US. The paper is structured as follows. The theoretical framework is set out in the next section. This is followed in Section 3 by a description of the measurement methods and data. Section 4 presents some summary descriptive measures for the total economy whereas Section 5 quantifies the empirical implications of capitalizing intangibles for growth. The paper concludes with a brief summary and discussion of future measurement challenges.

2. The theoretical framework

CHS (2005, 2009) advanced a simple three-sector model that specifies production functions for consumer goods, conventional investment goods, and intangibles. The model was used to show how an economy's input and output growth changed when business investment in intangibles was capitalised. The model was also adopted to identify the prices and quantities that needed to be measured in order to capitalise intangibles and study their contribution to growth.

The approach outlined below follows Corrado *et al.* (2011), integrating the various approaches to innovation (this section), and implementation into a national accounts measurement framework (Section 3) — see also Corrado *et al*. (2013). The main assumptions of the model are the following. Knowledge (ideas) is an input needed to produce consumption and tangible investment goods, together with labour and tangible capital. There exist two types of knowledge. One is knowledge that is generated without using factors of production and that is freely available to firms (free knowledge). The other is knowledge that is produced using inputs and that firms must pay for to use in their production process (commercialised knowledge). Commercialised knowledge is accumulated over time, generating the stock of commercial knowledge via the standard perpetual inventory relation and with its own user cost.

To be more precise, the model considers a simplified economy with just two industries/sectors. The innovation ('upstream') sector produces new finished ideas, i.e. it commercializes knowledge (e.g. a way of organizing production, or a software programme adapted to the needs of the organisation that implements pay and pension calculations), while the 'production' ('downstream') sector uses the knowledge to produce consumption and tangible investment goods. The innovation sector can, at least for some period, appropriate returns to its knowledge, and so this model is identical to Romer (1990) (where patent-protected knowledge is sold at a monopoly price to the final output sector during the period of appropriability), while the production sector is a price taker for commercialised knowledge. Both sectors are price takers for labour and tangible capital.

The first implication of the model is a broad definition of investment, which includes expenditure to purchase both tangible goods and commercialised knowledge, and a broad definition of aggregate output, which includes not only consumption goods and tangible investment goods but also commercialised knowledge.

(1) $P^Q Q = P^Y Y + P^N N = P^C C + P^I I + P^N N$

Where Q is real value added in the whole economy, Y is the output of the downstream sector N is commercialised knowledge, C is consumption, I is tangible investment, and P with the appropriate superscript are the corresponding prices.

The idea of including intangible investment as part of GDP can be thought of by analogy to tangible investment. Suppose an aircraft factory buys in aluminium and produces both final output and its own machines. Then its value added should be properly treated as both the final aeroplanes and the machines, i.e. one might think of the factory as consisting of both an aircraft factory and also a machine factory. Its investment should be treated as equal to the output of the new machines. Now, suppose the factory also writes its own long-lived software to run the machines. Then we should think of it as both an aircraft factory and machine factory and also a software factory and its investment should include not only the machines but also the new software that is produced. The second implication is that the expression for the sources of growth in value added is,

$$(2)\quad dlnQ = s_Q^L dlnL + s_Q^K dlnK + s_Q^R dlnR + dlnTFP$$

where *dlnTFP* is defined as the growth in Q (extended output including commercialised knowledge) over and above the growth contributions of labour (L), the accumulated stock of tangible capital (K) and the accumulated stock of commercialised knowledge (R) and where s^X_Q is the share of nominal value added accounted for by payments to factor X.

The final implication is that the model provides a measure of innovation. Equation (2) says that value added growth is due in part to growth in L and K. This formalises the idea of Jorgenson (2007), that growth can be achieved by duplication i.e. adding more labour and tangible capital, but also by innovation, that is, adding more ideas. It further says that growth can be due to the increased use of paid-for ideas, *dlnR*, weighted by their rental price (the licence fee to use a patent in an industrial process, for example); hence their contribution to *dlnQ* is $s_Q^R dlnR$. The final term, *dlnTFP* is the growth impact of everything else, which in this model can only be free ideas used in both sectors. Thus in this model, innovation in the sense of use of ideas is also growth net of K and L usage, i.e.

$$(3)\quad Innovation = s_Q^R dlnR + dlnTFP = dlnQ - (s_Q^L dlnL + s_Q^K dlnK)$$

Many innovation studies have attempted to distinguish between innovation and diffusion, the latter being the spread of new ideas. If the ideas come for free, they are, in this framework, counted in *TFP* growth. So the part of innovation measured by $s_Q^R dlnR$ is investment in commercialised new ideas and that part measured by *dlnTFP* might be regarded as the diffusion of free ideas.

3. Measuring intangibles for the total economy: concepts and methods

The empirical counterpart of the model outlined above requires measures of intangible investment. Corrado *et al.* (2013) set a general expression for estimating nominal intangible investment for a country or a region as follows:

$$(4)\ \ P^N N_t = \sum_{i=1\ i=1}^{N} \sum_{s=1}^{S} (\gamma_{i,s,t}^{own\text{-}account} \lambda_{i,s,t}^{own\text{-}account} OwnCost_{i,s,t}^{Indicator} + \gamma_{i,s,t}^{Purchased} \lambda_{i,,s,t}^{Purchased} Purchased_{i,s,t}^{Indicator})$$

where $P^N N$ is nominal expenditure, *i* is a subscript for industries, and *s* is a subscript for sectors. *OwnCost* and *Purchased* are time-series indicators of the own-account and purchased components of intangible investment, respectively. The other symbols, which though fully subscripted (i.e., by industry, sector, and time), are parameters: λ and γ are sector — and asset-specific capitalization factors that adjust the own cost and purchased indicators to benchmarks for each asset and sector. More specifically, λ is a time series indicator that is needed to transform the intermediate expenditure on intangibles into a sector-industry gross output (own-account) or gross spending measure varying over time ([11]); γ is the capitalization factor ([12]), namely, a parameter that adjusts a spending measure to a measure of investment — a fraction of revenues or employee time, say, devoted to long-lived activities (see Corrado *et al*. (2005)).

Intangible assets can thus be distinguished between assets that are already classified as investment in national accounts (software, R&D, mineral explorations and entertainment and artistic originals) and those assets that are not considered as investment (brands, organizational capital, design, training). Each intangible asset can be assumed to be composed of two different parts: purchased and own-account. In what follows we will take a closer look at the distinction between purchased and own-account intangibles distinguishing between intangibles already classified as investments in national accounts (NA) and assets that are not included in the NA asset boundary (non-NA).

Purchased intangible investment

With regard to the purchased component of non-national accounts CHS intangibles the time series for *Purchased* indicators are obtained from use tables in current prices (NACE Rev. 2 basis), available from most national statistical offices (NSOs) from 2002 onwards; for earlier years, it is possible to resort to the input-output tables generated from the WIOD project (see Bacchini *et al.* (2016)), for a detailed description of sources and methods). The use tables provide intermediate purchases by industry (columns) and by product (rows) according to the Classification of Products by Activity (CPA) codes. For the four CHS purchased assets, design, brands, organizational capital, and training, the following codes are selected: Architectural and engineering services, technical testing and analysis services (CPA M71); Advertising and market research services (CPA M73); Legal and accounting services, services of head offices and management consulting services (CPA M69 and M70); and Education services (CPA P85).

([11]) If annual time series of the use tables are available the λ parameters for the purchased component are implicit in the time series or can be estimated based on the relationship between those series and an indicator of intangible expenditure.

([12]) The capitalization factors are percentage values applied to the total expenditure on intangibles classified as intermediate costs to determine what part of it can be included in the asset boundary. For the non-national account assets these are set as follows (Corrado *et al.* (2016b)): design, 0.5; advertising and market research, 0.6; purchased organizational capital, 0.8; own-account organizational capital, 0.2; and training, 1.

Then once intangible expenditure by market and non-market industries is identified, the CHS methodology is adopted to capitalize each series (Corrado *et al.* (2005)).

As for national accounts intangibles, estimates rely on R&D and software data released by NSOs but then it is necessary to make further elaborations to generate intangible investment measures cross-classified by industry and institutional sector. Data availability depends on two different scenarios and these can vary also depending on the asset: one when NSIs provide gross fixed capital formation (GFCF) by industry and by sector but not the cross-classification between them and the other when GFCF data are available only by industry for software and/or R&D and there is no information classified by institutional sector. To deal with both situations two calculation methods have been identified within the SPINTAN project and they are exhaustively described in (Bacchini *et al.* (2016)).

Own-account intangible investment

The standard approach to measure gross fixed capital formation for own final use is based on the costs of production, i.e. the sum of compensation of employees, intermediate consumption and the cost of capital (consumption of fixed capital and, only for market producers, net operating surplus). The key variable in the calculation is the labour cost component frequently measured on the basis of compensation of specific occupational groups directly involved in the production of the asset for internal use (thus for example own-account software spending might be inferred from the wages of software occupations outside the software-producing industry).

Estimate of own-account training is a bit problematic since there is no information available about the labour costs of specific occupational groups directly involved in internally produced training activity; thus the lion's share of such costs are the opportunity costs of workers undergoing firm-specific training.

As for the remaining assets it is assumed that the own-account production of design, advertising and market research in the non-market sector is negligible and might be omitted while for the market sector this remains an open issue (see Bacchini *et al.* (2016), and Corrado *et al.* (2016b)). INTAN-Invest (2017) generates measures of brand and design only for the purchased component because there is not detailed occupational information available to estimate the own-account portion of these assets.

Measures of own-account organizational capital are produced by estimating total compensation of managers and then applying the corresponding capitalization factor (that takes into account also the other components of the cost of production, besides labour cost).

Prices and volume measures of intangibles

Generating measures of intangible investment in real terms is a big challenge because units of knowledge cannot be readily determined. Most intangible assets are unique products (with the exception of copies, e.g. in the case of pre-packaged software) and a large amount is produced on own account. Thus to get volume measures of intangibles it is necessary to make some assumptions, taking into account the current practice in NSOs. Purchased intangible assets, independently of the sector performing the investment, are generally deflated using average price indices because sector specific price information is not available. Own-account intangibles in real terms instead are obtained with an input based approach built on cost indices varying across sectors.

Specific recommendations about price measures for intangibles are provided by the *Handbook* *on Deriving Capital Measures for Intellectual Property Products* (OECD (2010)), that identifies three broad categories of intangibles and the corresponding prices: copies for sale, originals for sale, and originals for own-use. Hedonic methods are suggested to deflate copies; Producer prices (see the Producer Price Index Manual) are considered the best price measures for originals for sale, and originals for own-use has to be evaluated by means of productivity-quality adjusted price measures, and when these are not available it is recommended to adopt input-based methods.

The IPP suggestions can be easily followed to deflate purchased organizational capital, design, advertising and market research, because producer price indices for the corresponding industries are generally available, even if the statistical practice varies across countries. Service Producer Price Indexes (SPPIs) are generated taking into account quality adjustments and they are rather heterogeneous across countries and industries. Further, they are asset specific. Thus we assume, that currently, they are the best available price measure to deflate purchased intangible assets not included in the SNA asset boundary.

National accounts intangibles (software, R&D, mineral explorations and entertainment, literary and artistic originals) are deflated following two different approaches. Software is deflated adopting the harmonized price deflator developed by Corrado *et al.* (2012) and based on the OECD method. The harmonized price is obtained using a country-specific input cost index, the US pre-packaged software price index, and adjusting it for the relative inflation differential between the country of interest and the US.

Volume measures of R&D, mineral explorations and entertainment, literary and artistic originals, are obtained resorting to official national accounts deflators. The guidelines from Eurostat suggest using an input-based deflation method for R&D. The input-cost approach is currently the only viable option to deflate R&D because it guarantees a satisfactory degree of international comparability ([13]).

Summing up, the volume measures of purchased non-NA intangible investments are obtained applying national accounts value added prices of the industry corresponding to the main producer of each asset. Real measures of national accounts intangibles, besides software, are built applying investment deflators by branch of economic activity, and when these are not available, the asset price for the total economy.

Sources and data

In this paper we provide empirical evidence for the total economy drawing on INTAN-Invest estimates of business sector intangibles and on the SPINTAN measures of public intangible investment ([14]). The two sets of intangible estimates, although generated from two different and independent projects, share the same measurement approach and refer to two non-overlapping cross-classifications of sectors and industries. INTAN-Invest and SPINTAN estimates, taken together, provide harmonized measures of investment in intangible assets for the total economy cross classified by 21 industries (corresponding to the sections of the NACE rev. 2 classification) and two institutional sectors (market and non-market) for 15 European countries and the US.

The main pillar of the SPINTAN and INTAN-Invest estimation strategy is the adoption of the

([13]) A contrasting approach is in a paper by Corrado *et al.* (2011), which casts the calculation of a price deflator for R&D in terms of estimating its contribution to productivity. Applying their method to the United Kingdom yielded a price deflator for R&D that fell at an average rate of 7.5 percent per year from 1995 to 2005 and thus implied that real UK R&D rose 12 percent annually over the same period.

([14]) INTAN-Invest measure of intangible investment are available at www.intaninvest.net and SPINTAN estimates are downloadable from www.SPINTAN.net.

expenditure-based approach to measure the value of investment in intangible assets (i.e., expenditure data are used to develop direct measures of intangible investment). Moreover, both projects have the goal of generating measures of harmonized intangible investment satisfying (as much as possible) the following criteria: exhaustiveness, reproducibility, comparability across countries and over time, and consistency with official national accounts data. The above characteristics are assured by the adoption of official data sources homogeneous across countries. The main data sources are national accounts, whose availability dictates starting estimates of intangibles in 1995.

SPINTAN provides estimates of intangible investment performed by the non-market sector in a set of industries of interest. More precisely, the SPINTAN non-market sector consists of the non-market producers classified in the following industries: (1) Scientific research and development (NACE divison M72); (2) Public administration and defence; compulsory social security (NACE section O); (3) Education (NACE section P); (4) Human health and social work activities (NACE section Q), and (5) Arts, entertainment and recreation (NACE divisons R90-92) – see Corrado *et al.* (2017b). Non-market producers are defined consistently with national accounts definitions (i.e. establishments that supply goods or services free, or at prices that are not economically significant and that are classified in the Government sector (S.13) or in the Non Profit Institutions Serving Households (NPISH) sector (S.15)).

In the system of national accounts, units are classified by industry according to the activity they carry out, being market or non-market producers. Therefore, each industry can (potentially) consist of a mix of market and non-market producers. In particular, this is true for all the industries covered by the SPINTAN estimates, with the exception of the industry 'Public administration and defence; compulsory social security' (NACE section O), that includes only units belonging to sector S.13. We refer to these industries as SPINTAN mixed industries. Note that the SPINTAN non-market sector differs from the total of sectors S.13 and S.15 from national accounts as it does not cover non-market producers that are not classified in the industries of interest listed above.

INTAN-Invest covers investment by asset in industries from NACE sections from A to M (excluding M72) and Section S plus the market sector component of NACE M72, P, Q and R. In other words, it is the business sector complement to SPINTAN necessary to cover the total economy. For the sake of simplicity, we refer to the INTAN-Invest estimates as covering the market sector ([15]). Details of the calculations and assumptions required to calculate investments in intangible assets for the business sector can be found in Corrado *et al*. (2016b) and for the public sector in Corrado *et al.* (2016a).

3.1. Market and non-market intangible investment

What then are intangible assets? Table 1 summarizes the CHS list of business intangible assets (on the left) and maps them to the public or non-market sector (on the right). The correspondence between the two columns is not one-to-one. As may be seen, the asset boundary is slightly different depending on the market-non-market nature of the sector ([16]). But before we discuss differences across the two columns, let us make a few points about the similarities. First, while the character of some assets are rather different when produced by public institutions, e.g.,

([15]) In fact they also include the non-market sector component not covered by SPINTAN. The industry and sector coverage in INTAN-Invest 2017 has changed with respect to the previous INTAN-Invest estimates that did not cover industries P and Q and covered all industry R.

([16]) For a detailed discussion about the different nature of intangible investment in the market and non-market sector see Corrado *et al.* (2017).

Table 1: Knowledge capital for a total economy

Market sector		Non-market sector	
Computerized information		**Information, scientific, and cultural assets**	
1	Software	1	Software
2	Databases	2	Databases, including open data
Innovative property			
3	R&D, broadly defined to include all new product development costs *	3	Basic and applied science research, industrial and defense R&D
4	Entertainment & artistic originals	4	Cultural and heritage, including design
5	Design		
6	Mineral exploration	5	Mineral exploration
Economic competencies		**Societal competencies/social infrastructure**	
7	Brands	6	Brands
8	Organizational capital	7	Organizational capital
(8a)	Managerial capital	(7a)	Professional/managerial capital
(8b)	Purchased organizational services	(7b)	Purchased organizational services
9	Firm-specific human capital (employer-provided training)	8	Function-Specific human capital (employer-provided training)
		9	Schooling-produced human capital

* New product development costs include expenditures for testing and development of new products (including financial products and other services products) not included in conventional science-based R&D, software, databases, and design.

R&D, organizational, and mineral exploration, one may still draw a correspondence between these assets across sectors. For example, Jarboe (2009) defines public investments in brand as expenditures for export promotion, tourism promotion, and consumer product and food and drug safety (i.e, investments in product reputation). The correspondence for computer software, purchased investments in organizational capital, and function-specific worker capital (employer-provided training) is even closer.

The circled items are rather different in a public sector context. Open data refers to information assets in the form of publicly collected data issued and curated for public use. This runs the gamut from patent records to demographic statistics and national accounts to geographic information and local birth/death records. Indeed, after asking the question, 'What are public sector intangible assets in the United Kingdom?' Blaug and Lekhi (2009, p.53) concluded that 'perhaps the most important . . . is information assets.' Jarboe (2009) includes government information creation as a high-level category in his estimates of U.S. federal government intangible investments. The category includes spending on statistical agencies, the weather service, federal libraries, nonpartisan reporting and accounting offices, and the patent office, which suggests information assets loom large in the United States as well.

Cultural assets are public intangible assets whose services are used in production in cultural domains dominated or influenced by the public and non-market sectors; cultural domains are defined by the UNESCO Framework for Cultural Statistics. The capital used in many domains is included in existing estimates of private capital (tangible and intangible), but public investments (or funding) for new asset creation needs to be identified and newly capitalized. Note that cultural assets are notionally grouped with public architectural and engineering design, on the grounds that the British Museum's tessellated glass ceiling or the Louvre Pyramid are as valuable (and as incalculable) as the museums' contents although of course their correspondence to private counterparts is apparent.

3.2. National account intangibles: software and R&D

R&D, software and other intangible assets are already incorporated in the national accounts under the ESA 2010 revision.

Computer software together with large databases were recognized as intangible fixed assets under ESA 95 (par. 3.110). Both assets are a subcomponent of intellectual property products, together with research and development, mineral exploration, and artistic originals.

Computer software and databases and other originals of intellectual property products are valued at the acquisition price when traded on markets in the national accounts. The initial value is estimated by summing their costs of production, appropriately revalued to the prices of the current period. If it is not possible to establish the value by this method, the present value of expected future receipts arising from using the asset is estimated. GFCF, net capital stocks, GFCF at previous year's prices and the corresponding implicit deflators for price-volume decomposition for software and databases by industry based on ESA 2010 are widely available from Eurostat for most European countries (especially EU-15 countries) from 1995 onwards. Research by the OECD suggests that the valuation of databases varies across country statistical agencies (OECD (2010)).

Research and development (R&D) is defined in ESA 2010 as: 'Research and [experimental] development consists of the value of expenditures on creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and use of this stock of knowledge to devise new applications.' The value of research and development should be determined in terms of the economic benefits it is expected to provide in the future. This includes the provision of public services in the case of R&D acquired by government. In principle, R&D that does not provide an economic benefit to its owner does not constitute a fixed asset and should be treated as intermediate consumption. Unless the market value of the R&D is observed directly, it may, by convention, be valued at the sum of costs, including the cost of unsuccessful R&D, as described in chapter 6 of SNA 2008 par. 10.103. Thus R&D can be classified as a fixed asset if some criteria are satisfied, such as the provision of an economic benefit ([17]) to its owner. The INTAN-Invest and SPINTAN measures for market and non-market R&D and software are based on national accounts data ([18]).

3.3. Non-national accounts intangibles: brand, organizational capital and training

Given the complex nature of intangible assets, there is no definition of, or single method to, measure intangibles not included in national accounts asset boundaries that is accepted worldwide (Corrado *et al.* (2005)). Most of the literature simply identifies three critical attributes of intangibles: i) they are viewed as sources of probable future economic profits, ii) they lack physical substance, and iii) to some extent, they can be retained and traded by a firm (OECD (2008)). Yet, characteristics (i) and (iii) are also largely reflected in the more general definition

([17]) Economic benefit refers to the repeated and continuous use in the production process over a long period of time (more than one year). The SNA clarifies that the concept of economic benefit includes also the provision of public services in the case of R&D acquired by government.
([18]) See Corrado *et al.* (2016b) and for the public sector in Corrado *et al.* (2016a).

of economic assets provided by the System of National Accounts (SNA) that classifies them as those entities: over which ownership rights are enforced by institutional units, individually or collectively; and from which economic benefits may be derived by their owners by holding them or using them over a period of time (Harrison (2006)).

On the other hand, Corrado *et al.* (2005) proposed the widest definition of intangibles, referring to a standard intertemporal framework that leads to the conclusion that any use of resources that reduces current consumption in order to increase it in the future qualifies as an investment. This definition implies that all types of capital should be treated symmetrically, thus leading to a very broad definition of capital, including, for example, intellectual and human capital as well as organisational assets (Schreyer (2007)).

Non-national accounts intangibles, as seen above, include Innovative Property, other than R&D, designed to capture a range of assets that may have intellectual property protection associated with them, e.g. design rights. Economic competencies, instead aim at capturing a range of knowledge assets that firms invest in to run their businesses, but that might have no intellectual property. These include the costs of marketing and launching new products, including ongoing investments to maintain the value of a brand, and firm provided human capital in the form of training (Corrado *et al.* (2005, 2009)). These assets are conceptually straightforward although require assumptions to implement as detailed in Corrado *et al.* (2016b).

Economic competencies also include organisational capital which is conceptually more difficult and has a different characterization according to if we refer to the market or non-market sector. In the literature there is a broad consensus that organisational capital can have a significant impact on the outcome and performance of a firm (see for example Aral and Weill (2007) and Kapoor and Adner (2011)). Organisational capital is the cumulated knowledge that is built up in firms through investment in organising and changing the production process. These investments can be purchased externally by the firm, through expenditures on management consultancy and similar services, or can be own-account, produced within the firm through the actions of employees. Corrado *et al.* (2005, 2009) see own-account organisational capital as knowledge produced by persons in authority in a firm (managers), which yields a firm specific capital good jointly produced with output, and embodied in the organisation itself. This begs the question if managers, as defined in standard codes of occupations, are the only persons within the firm who have such authority. In particular in the public services there may be other high level employees who also possess authority. In SPINTAN the definition of own-account organisational capital was broadened to include some professionals such as senior doctors, who have the specific knowledge to set goals and the authority to ensure they are implemented.

4. Empirical evidence

The overall picture

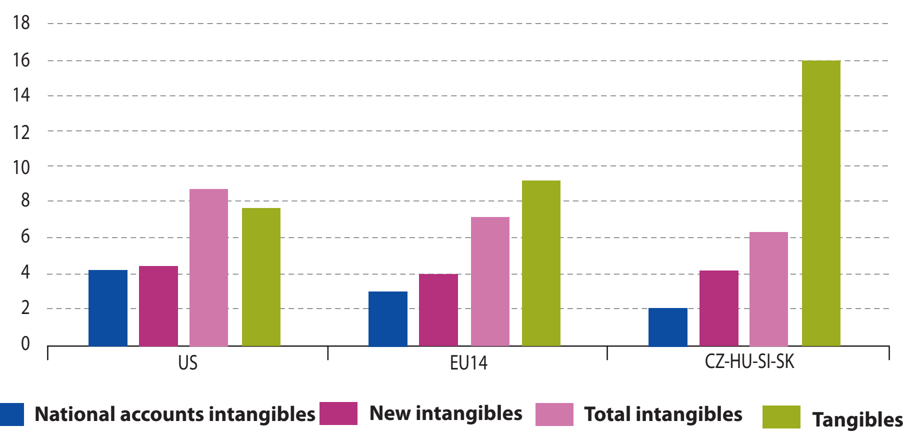
In what follows we look first at the relevance of intangible investment over GDP distinguishing between what is already capitalized in national accounts and what is not (non-NA intangibles) to get a sense of the weight of the assets left outside the asset boundary (Figure 1). Then we move to a framework where we assume that all intangibles are capitalized and so we analyse the shares of intangible investment over adjusted (19) value added.

In 2000-2013, as shown in Figure 1, the average share of intangible investment in GDP as measured in the national accounts was higher in the US (4.2 %) than in the EU14 (20) (3.1 %) as well as in the four new EU Member States (NMS) (21) included in the analysis (2.2 %). Moreover, national accounts data suggest that the GDP share of tangible investment in the three regions (7.7 %, 9.2 % and 16.0 % respectively) is higher than the intangible share.

However, when new intangible assets are included in the analysis, the intangible investment gap between the European economies and the US broadens. New intangibles account for 4.6 % of GDP, adjusted to include the new intangibles, in the US, and 4.1 % and 4.2 % in the EU14 and NMS respectively. Adding new intangibles to national account assets makes the GDP share of total intangible investment increase to 8.8 % in the US, 7.2 % in the EU14 and 6.4 % in the NMS. Hence in the US intangibles share of GDP was greater than the share of tangible investment while in the European economies the opposite was the case (22).

Within the EU14 economies, total intangible shares of GDP vary considerably, revealing an interesting geographical pattern. Northern Europe (Denmark, Finland, Ireland, Sweden and

Figure 1: Intangible and tangible investment
(% GDP, average 2000-2013)



Source: INTAN-Invest and authors' elaborations on national accounts.

(19) Value added is adjusted to account for the capitalization of non-NA intangible assets (Corrado *et al*. (2012)).
(20) The sample countries are EU15 member economies excluding Luxemburg.
(21) New member states are: Czech Republic, Hungary, Slovenia and Slovakia.
(22) Although intangible intensity in the four NMS was slightly lower than in the EU14 region, the ratio of tangible investment to GDP (16 %) was almost 50 % higher than in the US and almost 60 % higher than in the EU14 region.

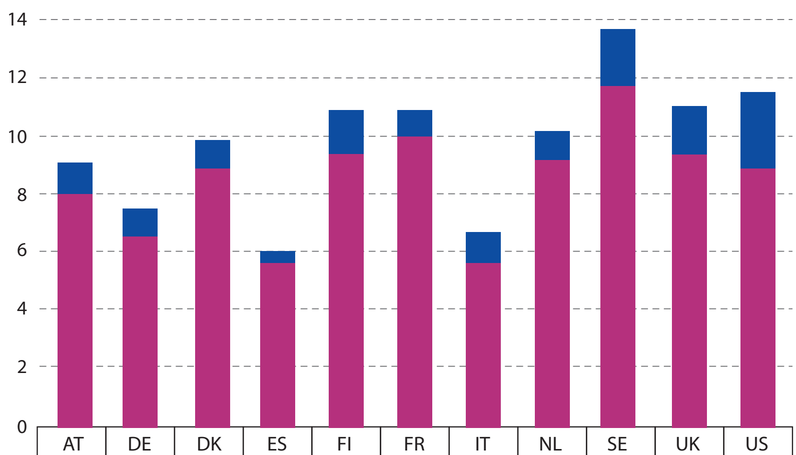
the UK) and non-German-speaking continental European countries (France, Netherlands and Belgium) are highly intangible intensive and characterised by higher intangible than tangible shares of GDP over the years 2000-2013. Sweden is the leading country with an intangible GDP share of 10.4 %, followed by the UK (9.0 %), Finland (8.8 %), France (8.7 %), the Netherlands and Ireland (both at 8.5 %) and Belgium (8.1 %) and Denmark (7.8 %) lagging slightly behind.

The Mediterranean and German-speaking countries are relatively less intangible intensive economies. In Austria, the intangible investment rate (6.7 %) is lower compared to the more intangible-oriented economies but still close to the average of the EU14. Portugal (6.0 %) and Germany (5.9 %) are below the EU14 average intangible share of GDP whereas Italy (5.3 %) and Spain (4.6 %) are far behind. Greece shows the lowest average share over the period (3.7 %), being an outlier also in terms of the tangible GDP share of investment.

Looking at sectoral value added shares of public and private intangibles reveals that overall (market and non-market) intangible investments account for from nearly 14 % (Sweden) to just under 6 % (Spain) of value added (Figure 2). The market sectors accounts for the main component of intangible investment as a share of value added - averaged across countries, the market sector shares of value added are 8 % compared to 1.5 % for the non-market sectors.

In those countries with the highest shares of intangibles (Sweden, the US and the UK), intangible investments now account for a larger value added share than tangible capital investments (Figure 3). Countries such as Spain and Italy have a much lower share of intangibles than tangibles.

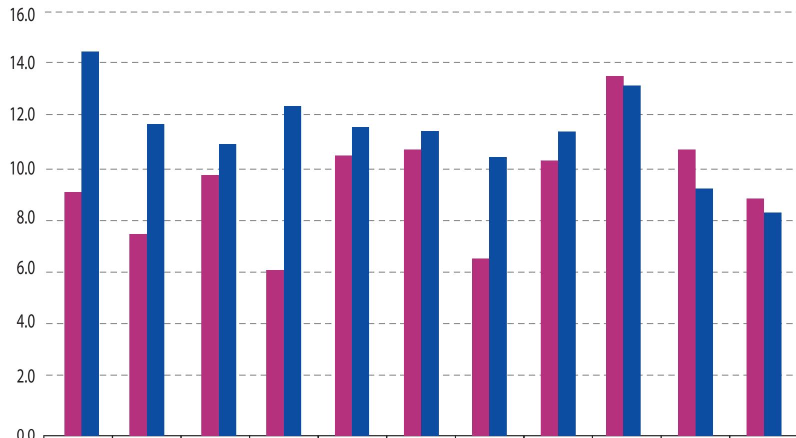
Figure 2: Market and non-market intangibles (2013)
(Adjusted value added shares of intangible investment)



	AT	DE	DK	ES	FI	FR	IT	NL	SE	UK	US
Non-market	1.1	1.0	1.0	0.5	1.2	0.7	1.0	1.1	1.9	1.6	2.7
Market	8.0	6.5	8.9	5.5	9.4	9.9	5.6	9.2	11.6	9.3	8.9

Source: authors' elaboration on SPINTAN and INTAN-Invest data.

Figure 3: Tangible and intangible investment (2013)
(Adjusted value added shares)



	AT	DE	DK	ES	FI	FR	IT	NL	SE	UK	US
Intangibles	9.1	7.5	9.8	6.0	10.6	10.7	6.6	10.3	13.5	10.8	8.9
Tangibles	14.6	11.8	11.1	12.3	11.7	11.6	10.5	11.5	13.3	9.3	8.3

Source: authors' elaboration on SPINTAN and INTAN-Invest data.

The differences in intangible intensity across countries also mirror the industrial structure of the economies, with countries such as Germany that are heavily concentrated in manufacturing having a higher share of tangible than intangible investments.

5. Intangibles and growth

Arguably the main interest in constructing measures of capital services is their use in explaining international growth trends. Therefore we conclude this paper with a brief summary on efforts to link intangibles to growth. Note that in the papers referred to in this section, intangible inputs are based on the capital services that arise from their use. Interested readers should consult the papers for details on moving from investment to capital stocks and capital services. A common finding is that the spillovers referred to in Section 2 are evident in the data.

Using data for the market economy for 13 countries, Roth and Thum (2013) suggest that once accounting for business intangibles i) capital deepening becomes the dominant source for explaining labour productivity growth and ii) the explanatory power of TFP growth is diminished from 25 percentage points to 10 percentage points. In econometric production function estimates, these authors report a coefficient of intangible investment of about one quarter - this turns out to be much higher than the coefficient identified by this asset's factor share in growth accounting.

Using the INTAN-Invest data, Corrado *et al.* (2017a) find a coefficient of similar, in some specifications even larger magnitude. They formally investigate the presence of spillovers that are suspected if the estimated marginal product of a factor exceeds the marginal product implied by the factor remuneration under competitive markets. Their results strongly support the possibility of spillovers. Moreover, they find evidence of a complementarity between intangible assets at the aggregate level and ICT capital at the sectoral level.

A first attempt to produce internationally comparable estimates of intangible investments at the industry level was undertaken by Niebel *et al.* (2013) for the INDICSER project. The growth accounting estimates by industry suggest that the importance of intangible capital assets by type varies across sectors, with R&D the most important asset in manufacturing whereas organisational capital dominates in many service sectors. In terms of contributions to labour productivity growth, however, there appear to be common sectoral patterns across countries, with high investment in all sectors in some countries (the UK and the Netherlands) and low investment in others (Italy and Spain). The paper performed an econometric estimation of the relationship between indicators of intangible capital and labour productivity growth at a sectoral level. This confirms the positive impact of intangible capital on economic performance as found by previous authors. However, the paper estimates coefficients on intangibles, ranging from 10 % to 17 %, which is much lower than the coefficients using aggregate data. The paper suggests that unexplained heterogeneity at the macro level is likely to explain this difference and such biases are partially addressed using industry data. Nevertheless these estimates remain higher than average growth accounting impacts, consistent with some spillovers from this asset type.

Recent empirical evidence (Corrado *et al.* (2016b)) confirm that intangible investment is a key policy variable. A relevant characteristic of intangible capital is that it is growth-promoting (Corrado *et al.* (2014)), thus potentially contributing to reducing the growth gap between the EU and the US. Therefore policies designed to foster innovation and to make the economic environment more conducive to investment in intangible assets should adopt a view of innovation that is broader than R&D. Corrado *et al.* (2016b) show that the investment gap between the EU14 and the US is more related to the lower contributions of computer software and databases, artistic originals, mineral exploration, brand and training than to the contribution of R&D.

In a recent paper Corrado *et al.* (2016b) use the recently constructed INTAN-Invest cross-country cross-industry dataset on investment in tangible and intangible assets for 18 European countries and the US, in a growth accounting framework, to analyse the impact of capital before and after the Great Recession in 2008-2009. The major findings are the following. First, tangible investment fell massively during the Great Recession and has hardly recovered, whereas intangible investment has been relatively resilient; it recovered fast in the US but lagged behind in the EU. Second, the sources of growth analysis including only national account intangibles (software, R&D, mineral exploration and artistic originals), suggest that, over the period 2000-2013, capital deepening is the main driver of growth, with tangibles and intangibles accounting for 80 % and 20 % in the EU, while both account for 50 % in the US. Extending the asset boundary to the intangible assets not included in the national accounts increases the contribution of capital deepening. The contribution of tangibles is reduced both in the EU and the US (60 % and 40 % respectively) while intangibles account for a larger share (40 % in EU and 60 % in the US). Their analysis shows that since the Great Recession, the slowdown in labour productivity growth has been driven by a decline in TFP growth with relatively minor roles for tangible and intangible capital. Finally, they document a significant correlation between stricter employment protection rules and less government investment in R&D, and a lower ratio of intangible to tangible investment.

6. Conclusion

The main purpose of this paper was to set out the existing research on measuring intangible capital and its impact on economic performance. It illustrated the theoretical framework for understanding the impact of intangibles on output and then discussed which types of activities might be deemed to be intangible assets. The paper then uses the newly developed data from INTAN-Invest and SPINTAN to investigate differences across countries and suggests that the EU lags the US in investing in this type of asset. Recent empirical evidence using these data show that intangible investment is important for understanding the pattern of economic growth, both across time and countries. The econometric analysis using these data yield estimates that are consistent with spillovers of intangibles to growth.

This paper has reviewed the research on intangibles from a macro or industry perspective. It should be noted that there is also a considerable body of evidence emerging on the importance of intangibles at the firm level. Examples include Görzig *et al.* (2010), Piekkola (2016), Riley and Robinson (2011) and Riley and Rosazza Bondibene (2017), which highlight a positive relationship between firm performance and the use of intangibles. However, as with some of the macro estimates, intangible assets are indirectly measured using occupation data.

Further analysis could also consider the regional dimension, to link in with the extensive literature on the concentration of knowledge assets by geographical location, especially the role of cities. In this respect, some progress has been made by Mas *et al.* (2017) who construct intangible data for the 17 Spanish regions and 24 industries covering the period 1995-2014.

In all these approaches, macro/industry, firm and region, the analysis would benefit from more and better data, e.g. direct surveys of the intangible investment behaviour of firms or use of administrative databases.

Acknowledgements

We are grateful to the European Commission FP-7 grant agreement 612774 (SPINTAN) for financial support. Errors are our own.

References

Aral, S. and P. Weill, (2007), 'IT assets, organizational capabilities, and firm performance: How resource allocations and organizational differences explain performance variation', *Organization Science,* 18 (5), pp. 763-780.

Bacchini, F., C. Corrado, J. Hao, J. Haskel, R. Iannaconne, M., Iommi, K., Jager and C. Jona-Lasinio, (2016), 'Estimates of Intangible Investment in the Public Sector: The EU, US, China and Brazil', SPINTAN working paper, No. 11.

Blaug, R. and R. Lekhi (2009), Accounting for Intangibles: Financial Reporting and Value Creation in the Knowledge Economy, London: The Work Foundation.

Corrado, C., P. Goodridge and J. Haskel (2011), 'Constructing a Price Deflator for R&D: Estimating the Price of Knowledge as a Residual', The Conference Board Economics Program Working Paper, EPWP #11-03.

Corrado, C., J. Haskel and C. Jona-Lasinio (2017a). 'Knowledge spillovers, ICT and productivity growth', *Oxford Bulletin of Economics and Statistics*, vol. 79, pp. 592-618, ISSN: 1468-0084, doi: 10.1111/obes.12171.

Corrado, C., J. Haskel and C. Jona-Lasinio (2017b), 'Public Intangibles: the Public Sector and Economic Growth in the SNA', forthcoming in *Review of Income and Wealth*, DOI: 10.1111/roiw.12325.

Corrado, C., J. Haskel, C. Jona-Lasinio and M. Iommi (2012), 'Intangible Capital and Growth in Advanced Economies: Measurement Methods and Comparative Results', Working Paper, June, available at http://www.intaninvest.net.

Corrado, C., J. Haskel, C. Jona-Lasinio and M. Iommi (2013), 'Innovation and intangible investment in Europe, Japan, and the United States', *Oxford Review of Economic Policy*, vol. 29, pp. 261-286, ISSN: 0266-903X, doi: 10.1093/oxrep/grt017.

Corrado C., J. Haskel, C. Jona-Lasinio, M. Iommi (2016b), 'Intangible investment in the EU and US before and since the Great Recession and its contribution to productivity growth,' in 'Investment and Investment Finance in Europe', Chapter 2, pp. 73-102, European Investment Bank Report, November 2016.

Corrado, C., C. Hulten and D. Sichel (2005), 'Measuring Capital and Technology: An Expanded Framework', In: Corrado, C., J. Haltiwanger, and D. Sichel (eds), *Measuring Capital in the New Economy*, Chicago, IL: University of Chicago Press, pp. 11–46.

Corrado, C., C. Hulten and D. Sichel (2009), 'Intangible Capital and U.S. Economic Growth', *Review of Income and Wealth* 55: pp. 661–685.

Corrado, C., K. Jaeger and C. Jona-Lasinio (Eds.) (2016a), 'Measuring Intangible Capital in the Public Sector: A Manual', available at www.spintan.net.

Fukao, K., T. Miyagawa, K. Mukai, Y. Shinoda, and K. Tonogi (2009), 'Intangible Investment in Japan: Measurement and Contribution to Economic Growth', *Review of Income and Wealth* 55 (3), pp. 717–736.

Görzig, B., H. Piekkola and R. Riley (2010), 'Production of Own Account Intangible Investment: Methodology in Innodrive Project', INNODRIVE Working Paper 1.

Jarboe, K. (2009), 'US Policies for Fostering Intangibles,' Presentation, 6th Summit on Monetizing & Maximizing IP.

Jorgenson, D. (2007), Public Comment to the Measuring Innovation in the 21st Century Advisory Committee.

Kapoor, R. and R. Adner (2012), 'What firms make vs. what they know: how firms' production and knowledge boundaries affect competitive advantage in the face of technological change', *Organization Science*, 23 (5), pp. 1227-1248.

Lev, B. (2001), 'Intangibles: Management:, Measurement, and Reporting', Washington, D.C., Brookings Institution Press.

Lucas, R. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics* 22 (1).

Mas, M., J. Quesada, E. Uriel, J. Fernández de Guevara, C. Albert, E. Benages, L. Hernández, C. Mínguez and J. C. Robledo (2017), *La economía intangible en España: Evolución y distribución por territorios y sectores (1995-2014)*, Fundación COTEC, Madrid; 77.

Nakamura, L. (1999), 'Intangibles: What put the new in the new economy?', *Federal Reserve Bank of Philadelphia Review* (July/August), pp. 3.

Nakamura, L. (2001), 'What is the US gross investment in intangibles? (at least) one trillion dollars a year!' Working Paper No. 01-15, Federal Reserve Bank of Philadelphia, Philadelphia, PA.

Niebel, T., M. O'Mahony and M. Saam (2016), 'The contribution of Intangible assets to sectoral productivity growth', *Review of Income and Wealth*, 63, S49-S67

OECD (2010), 'Handbook on deriving capital measures of intellectual property products', Technical report, OECD.

Piekkola, H. (2016), 'Intangible Investment and Market Evaluation', *Review of Income and Wealth*, 62, pp. 28-51.

Riley, R. and C. Robinson (2011), 'Skills and Economic Performance: The Impact of Intangible Assets on UK Productivity Growth', Evidence Report No. 39, UK Commission for Employment and Skills.

Riley, R. and C. Rosazza Bondibene (2017), 'Driving Productivity Growth: The Role of Firm-Specific Knowledge Assets', LLAKES Research Paper (forthcoming).

Romer, P. M. (1986), 'Increasing returns and long run growth', *Journal of Political Economy,* 94,1002.

Roth, F. and A. Thum (2013), 'Intangible capital and labor productivity growth—Panel evidence for the EU from 1998-2005', *Review of Income and Wealth,* 59, pp. 486-508.

Weitzman, M. L (1976), 'On the Welfare Significance of National Product in a Dynamic Economy,' *Quarterly Journal of Economics*, 90, pp.156–162.

Weitzmann, M. L. (2003), *Income, Wealth, and the Maximum Principle*, Cambridge, MA: Harvard University Press.

4 A focused strategy for the estimation of Italian quarterly external trade by geographical area, open issues and solutions

GIANCARLO LUTERO [1] AND PAOLO FORESTIERI [2]

Abstract: The European System of Accounts (ESA 2010) introduced significant changes in national accounts for external trade estimates, which impacted also on data broken down by geographical area. This work will present the new estimation by geographical macro-domains in both its general aspects and most significant theoretical issues. It will specifically focus on quarterly estimates, released with ESA 2010 for the very first time, explaining all the operational steps needed to achieve consistent estimates of goods and services components. An empirical exercise has been carried out to assess the forecasting performance of a set of dynamic models and their combination, and to examine the impact of parameters calibration on quarterly discrepancies in the final reconciliation step of the external trade accounts system as a whole.

JEL codes: C53, C61, F17

Keywords: external trade, geographical area, quarterly national accounts, point forecasts combination, reconciliation of quarterly time series

[1] ISTAT, National Accounts Directorate.
[2] ISTAT, National Accounts Directorate.

1. Introduction

In 2014, the Italian National Institute of Statistics (ISTAT) reviewed the whole system of national accounts to implement the changes introduced by the new European System of Accounts (ESA 2010) and the ESA transmission programme (see Eurostat (2013a)). ESA 2010 introduces an important conceptual innovation for international flows estimation, giving priority to the criterion of changing ownership of goods rather than the physical movement across borders. This new concept affects the registration of economic phenomena like processing and merchanting. The definitions of ESA 2010 influence the whole system of external trade estimates, including data broken down by geographical area and the estimates of imports and exports of goods by country (3). This paper briefly describes the revisions of the annual estimates by geographical area and the introduction of quarterly estimations. Specifically, the latter will be analysed and all the steps needed to achieve consistent quarterly estimates of goods and services components will be described.

Section 2 examines in depth the main methodological changes mentioned above and their impact on geographical data estimation. Section 3 is divided into two subsections. The first one describes the annual estimation by area, whilst the other gives insight in the annual and quarterly estimation of goods and processing services by country, which is an important component for the compilation of the Balance of Payments (BoP). Section 4 focuses on the issues related to quarterly estimates by geographical area (forecasts by proxy indicators, seasonal and calendar adjustments, temporal disaggregation and final reconciliation). Finally, Section 5 shows the computational exercise used to verify the forecast performance of a wide collection of dynamic models and several of their combinations. It also illustrates the impact of parameters calibration on quarterly discrepancies in the final step of reconciliation of the whole external trade accounts system. Section 6 presents a short summary of the results.

2. Changes introduced in external trade by ESA 2010

The transition process to new estimates of national accounts stems from the introduction of the 2008 version of the System of National Accounts (see United Nations *et al* (2009)). The new SNA has been adopted by the European Union and it inspired the new version of the European System of Accounts (ESA 2010), which replaces the previous ESA95. At the same time, coherently with the SNA, the international manual of balance of payments (BPM6, see IMF(2010)) was modified. ESA 2010 and BPM6 affect significantly the compilation rules of international flows in national accounts and consequently, the estimates of imports and exports of goods and services by geographical breakdown. The ESA 2010 introduces an important conceptual innovation regarding the registration of international flows. It gives priority to the criterion of changing ownership of goods rather than to physical movements across borders. This change impacts the recording of two phenomena: goods sent abroad for processing (in short, processing) and international trade of goods that do not cross the border of the seller's country of residence (merchanting).

(3) A correct definition of goods will be specified in the next sections.

Processing affects the estimates of imports and exports broken down by geographical areas. The value of processing services on goods either sent to or received from abroad for processing — without change of ownership — has to be registered as services. Goods undergoing processing should not be included in imports or exports aggregates. In ESA95, the flows described above were registered as imports and exports of goods, and the services were incorporated in the value of the goods. The new system should not affect, in principle, the net balance of flows with foreign countries. However, it should impact the levels of import and export flows; more specifically a large reduction in goods flows and a smaller increase in services.

According to the new definitions of ESA 2010, it is necessary to estimate the value of processing services abroad on goods owned by residents, and the processing services in Italy on goods owned by non-resident. Likewise, goods moving across the border for or after processing need to be deducted from the gross flows of imports and exports. International Merchandise Trade Statistics (IMTS) produced by ISTAT could theoretically provide enough information to estimate service flows and to adjust the gross flows of goods. Nevertheless, it presents some statistical inconsistencies (time delay problems, inconsistency between the declared value of goods before and after processing and the related processing service) and coverage problems (when processed goods do not cross the border). Hence, IMTS data are supplemented by the Intra-Community Trade in Services (ICTS) data collected by the tax authority. This enables to directly estimate the processing services trade of Italian companies within the EU. Through records linked to IMTS data, it allows to adjust the trade flows in goods to the ESA 2010 and BPM6 definitions. Therefore, final estimates are based on IMTS micro data containing information on the product country, operator and nature of the transaction, integrated with ICTS. The changes introduced by the ESA 2010 created a sharp disjunction between IMTS data and national accounts data for external trade of goods (mostly) and services.

The introduction of the ESA 2010 involved a review of imports and exports data broken down by area. In the next sections, we will present the yearly and quarterly estimates, the latter being introduced for the first time. This will emphasise the importance of consistency between the two domains.

3. Estimation of imports and exports by area and by country

Before describing the quarterly estimation of imports and exports by area (see Section 4), we focus on the annual estimation. Both are built at the same level of geographical disaggregation. The annual data provide the constraints for the quarterly raw data at current and chained prices. It should be underlined here that the annual estimates by geographical area are calculated by product, although only total data are transmitted to Eurostat. Section 3.1 describes the estimates by geographical area.

Section 3.2 deals with the estimation of goods and of processing services by country, which is needed for the BoP compilation. Both annual and quarterly data are considered by country and are calculated raw and at current prices, coherently with annual and quarterly data estimated by product and transmitted aggregated. Indeed, our goal is to briefly describe all the estimations by country and by area as they have to be considered as a whole. This draws attention to the efforts spent to achieve coherence among estimates.

3.1. Geographical area

The ESA 2010 Transmission programme requests that import and export of goods and services to be broken down into the following geographic area:

- Member States whose currency is the euro, the European Central Bank and other institutions and bodies of the euro area (EMU);
- Member States whose currency is not the euro, and institutions of the European Union (except the institutions mentioned above) (EU);
- Non-Member States whose currency is not the euro.

EMU and EU breakdowns show the actual composition at the end of each reference period (evolving composition). The estimation of external trade data by geographical area has been reviewed according to the principles of the ESA 2010. As mentioned above, changes such as the treatment of processing services or merchanting affect geographical breakdowns. The review has also taken into account the introduction of BPM6. This has an impact on the imports and exports of services estimates provided by the Bank of Italy, which is one of the sources used for the estimation by geographical area of imports and exports. Furthermore, quarterly data estimations of geographical data were produced for the first time in order to be fully compliant with the transmission programme. The main sources used for the annual estimate are:

- International Merchandise Trade Statistics (IMTS) produced by ISTAT;
- Intra-Community Trade in Services (ICTS) produced by ISTAT;
- Bank of Italy imports and exports of services data.

The methodology to estimate imports and exports of goods is based on IMTS and ICTS data. IMTS data are adjusted to remove the flows of goods crossing the border for or after processing. The processing adjustment procedure preserves the country information and the six-digit level of the Classification of Products by Activity (CPA). It also enables to aggregate data by area and product. Then, to ensure the exhaustiveness of estimation of imports and exports of goods by area, the following items are added:

- Imports and exports of illegal goods
- Merchanting (exports);
- Goods procured in ports (imports) ([4])

The main source of estimates of imports and exports of services is Bank of Italy's data except for processing services, which are estimated using ICTS and IMTS. Bank of Italy data are used to estimate the CIF/FOB ([5]) adjustments broken down by area. Since IMTS provides imports at CIF value, this enables to adjust imports of goods at FOB value, as required by the ESA 2010. Bank of Italy data broken down by geographical area are elaborated at fixed composition ([6]). Therefore this source cannot be used directly in estimates and data broken down by country are needed.

The final step of annual estimation is the sum of goods and services broken down by area. The estimates obtained are fully consistent with the total imports and exports of goods and services ([7])

([4]) Merchanting and goods procured in ports (imports) are estimate by the Bank of Italy.
([5]) The acronym FOB stands for 'free on board', while CIF stands for 'cost, insurance and freight'.
([6]) EMU and EU breakdowns are the last composition for all reference periods.
([7]) Data broken down by area are estimated at previous year price and chain-linked volumes.

3.2. Countries

Imports and exports of goods and processing services by country are estimated net of imports and exports of illegal goods, merchanting and imports of goods procured in ports. Imports and exports of goods by country are an important input to compile the BoP. At the same time, they provide an interesting tool for analysing international flows from the national accounts point of view because the estimates embody the definition of the ESA 2010 and BPM6. As reported above, they set a clear distinction between national accounts and IMTS data, which is one of the sources of the estimation process. The comparison between national accounts and IMTS data, together with the processing services data, is important to grasp the features of external trade.

Although the quarterly estimation process will be broadly discussed in the next sections, the aim of this paragraph is to specifically analyse the estimation by country as a whole. Imports and exports by country are estimated only in their raw version and at current prices. Therefore, discussions about calendar effects, seasonal adjustments and prices do not apply.

The quarterly estimates of imports and exports of goods are carried out for a selection of countries at t+65 days from the reference quarter ([8]) on the basis of IMTS data.

It is possible to split the procedure into two steps: the first one is the temporal disaggregation of data and the second one is a reconciliation of data. In the first step the source does not provide all the information to estimate imports and exports according to ESA 2010. Hence, only IMTS micro-data registered as 'not for processing' are used as reference indicators in an indirect estimation procedure ([9]): the quarterly disaggregation is obtained using annual data and reference indicators. In the second step, data are reconciled with quarterly estimates by product considering the annual constraint by country in order for the estimations to remain coherent with quarterly data by product and annual data by country. The method of reconciliation used here is a two-constraint one, either contemporaneously or temporally. This aspect will be discussed more broadly in Section 4.4. The contemporaneous constraint ensures the consistency with quarterly national accounts data, while the temporal aggregation constraint ensures the consistency between annual and quarterly data by country.

Table 1 summarizes the process of reconciliation, where the marginal values represent the constraints for the estimates.

Table 1: Framework of quarterly geographical area accounts

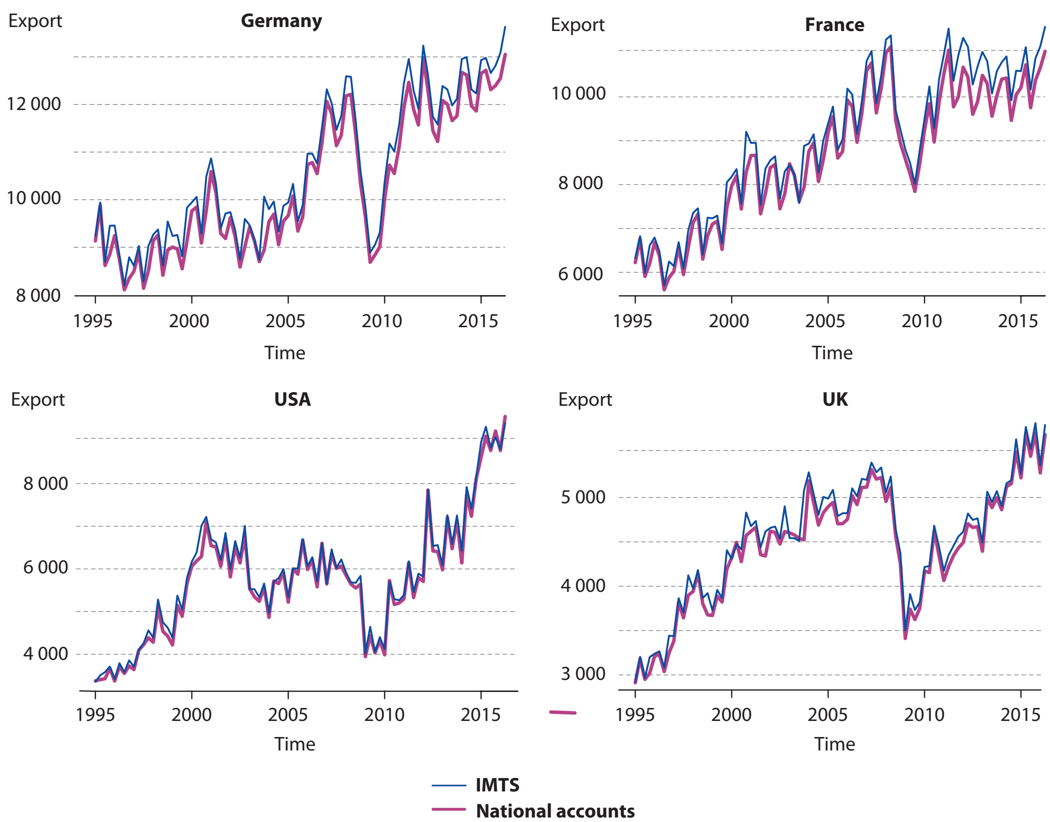
Year	Quarter	C1	C2	C3	.	Cn	tot
T	q1	Y_{q1}	Y_{q1}	Y_{q1}	.	Y_{q1}	$Z_{tot,q1}$
	q2	Y_{q2}	Y_{q2}	Y_{q2}	.	Y_{q2}	$Z_{tot,q2}$
	q3	Y_{q3}	Y_{q3}	Y_{q3}	.	Y_{q3}	$Z_{tot,q3}$
	q4	Y_{q4}	Y_{q4}	Y_{q4}	.	Y_{q4}	$Z_{tot,q4}$
	tot	Y_{c1}	Y_{c2}	Y_{c3}	.	Y_{cn}	Z_{tot}

In Figure 1, IMTS data and national accounts data for selected countries are plotted together in order to underline the strict relationship between them but also their differences. The graphs are showing exports of Italy to the main external trade partners. For all the partner countries considered, it is possible to grasp the fall of exports during the 2008 crisis and the

([8]) The estimates are used by Bank of Italy to compile the balance of payments.
([9]) The estimation procedure will be broadly discussed in the next paragraph.

subsequentrebounds; e.g. exports to the USA increased more than exports to EU countries. It appears that countries with higher differences between IMTS and national accounts exports present a higher level of imports of processing services. In fact, it can be said that higher levels of imports of processing services are associated to higher amounts of goods exported for processing (recorded in the IMTS data but not in the national accounts data).

Figure 1: IMTS and national accounts exports for countries with higher value, raw series at current prices
(million EUR)



Source: Authors' calculations.

4. Quarterly estimates by geographical area in four steps

As reported in Section 2, according to the new European System of Accounts (ESA 2010) and the ESA 2010 transmission programme, ISTAT has to release quarterly estimates of imports and exports broken down by geographical area. The geographical areas have been identified as follows:

- European Union countries + EMU Institutions;
- European Monetary Union countries;
- European Union countries not included in EMU;
- EMU Institutions;
- Extra EU countries.

Quarterly series have to be compiled at current prices, previous year's prices and chain-linked volumes, raw, calendar and seasonally adjusted, at t+60 days from the reference quarter, and for five geographical areas. Data are not published but are released, at planned deadlines, both to the Bank of Italy for the purpose of BoP compilation and to Eurostat.

In the system of Italian national accounts, a top-down approach is applied, in which annual accounts are the binding reference system. The other sub-systems — namely the quarterly accounts, institutional accounts, territorial accounts and environmental accounts — are deduced from it. The annual aggregates estimation is carried out using a wide range of diversified statistical tools. As for quarterly accounts data, they are evaluated through the so-called indirect methods or data interpolation procedures (temporal disaggregation of annual series with the use of reference indicators ([10])). This enables to overcome the limitations due to the substantial information gaps at quarterly level. The most important statistical sources to be taken into account in geographical area estimates are the same identified and used in quarterly national accounts:

- monthly external trade statistics of goods (internal ISTAT source);
- quarterly BoP data (Bank of Italy) relating to imports and exports of services;
- producer price indices for imports and exports of goods (ISTAT);
- consumer and services price indices (Eurostat, OECD), for imports of services and a single output price for exports.

The steps to develop a consistent system of quarterly geographical area accounts are the following:

(1) a forecasting procedure, needed to cope with the gap of information concerning the reference indicators for the most recent period (typically the last quarter or month), or a direct forecast on aggregated quarterly variables;

(2) seasonal and calendar (working days correction) adjustments;

(3) temporal disaggregation of annual data;

([10]) For an introduction see Bee-Dagum and Cholette (2006).

(4) final time series reconciliation to achieve at the same time internal consistency between the longitudinal dimension (geographical-institutional) and the temporal dimension.

In the next subsections, the theoretical aspects of each phase will be outlined.

4.1. Forecasting procedure with reference variables

One of the most significant problems faced in the quarterly national accounts domain is the possibility that some proxy-indicator were not available for the most recent quarter or month, which is generally the most interesting one for users. In the event one or two months of the current quarter were not available at the time of the publication of data, it would be necessary to use forecasting methods to fill in the missing information and proceed with the subsequent steps of the estimation process. A key question in forecasting is whether the point-forecast of an aggregate (direct method) improves upon those derived from a combination of forecasts with a specific — parametric or not — aggregation rule (indirect approach). Given the theoretical considerations and the empirical results available in the very large econometric literature, it is not possible to determine whether the direct approach is more appropriate than the indirect. The forecasts performance depends on many elements such as the type of statistical model used, the kind of time series, the forecasting horizon and many other factors.

For instance, David Hendry states in several papers ([11]) that sources of forecast or prediction errors may arise for many reasons. Examples of this are: model specifications or variables identifications included in the models, stochastic model functional forms (essentially linear model versus nonlinear), model selection criteria for dynamic models, general estimation uncertainty, data measurement errors and the presence of structural breaks over the forecast horizon. Therefore, in the context of macroeconomic forecasting, experimentation is necessary to exclude the impact of events that considerably influence results from the analysis of the outcome. Such events include aggregation bias, the presence of structural breaks or of a clear trending, the use of too many predictors, or the impact of seasonal adjustment procedure and data revision.

A *point forecast* $\hat{y}_{t+h|t}$, with forecast horizon *h*, is usually defined as the value that minimizes the mean squared error given the available information set Ω at time *t*:

$$\hat{y}_{t+h|t} = E(y_{t+h|t}|\Omega_t)$$

The forecast or *Prediction error* is defined as follows:

$$e_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$$

Among several available forecast accuracy measures, we will examine the most popular like: the *Mean Absolute Error*,

(1) $MAE = E(|e_{t+h}|)$

the *Mean Squared Error*,

(2) $MSE = E(e^2_{t+h})$

or the *Mean Error*, very useful to study possible persistent asymmetries in the empirical distribution of forecast error,

(3) $ME = E(e_{t+h})$

([11]) For example see Hendry (2004) and the classic reference Hendry (1995).

and finally, the *Root Mean Squared Error.*

(4) $RMSE = E[(e^2_{t+h})]^{(1/2)}$

It is always advisable to remain mindful of the distribution of the forecast error to select the most efficient stochastic model, to examine alternative measures like the median values, or even to set empirical distributions in which a percentage of tails values is discarded. A recommended solution, especially in the context of official statistics and univariate estimates domain, is to implement forecast combination techniques ([12]). Several arguments support the adoption of combination forecasts methods. Timmermann (2006) lists several of them:

- portfolio diversification because of incomplete information about target variable;
- presence of structural breaks in sample;
- misspecification bias of unknown source and form in individual fore-casting model;
- different approach of forecasters;
- a pure technical reason like computational capability.

Forecast combinations for data prediction, especially in official macroeconomic estimation, are suitable in several contexts:

- when there is an uncertainty over the 'best' forecasting model;
- to achieve a most efficient methodology;
- if there is an aversion with respect to large size forecast errors;
- when there is a great aversion versus the persistence of sign in prediction error.

In very general terms, combining forecasts include the following essential steps:

(1) selection of models to be implemented in the process;

(2) single model estimate (specification, diagnostics, computational problem);

(3) selection of loss function;

(4) forecasting summary (aggregation law).

In Section 5, a forecast exercise on real monthly data by geographical area will be proposed. This will be done considering several stochastic models and evaluating some forecast combination technique.

4.2. Seasonal and calendar adjustments

Removing the seasonal and calendar effects from raw figures is important to carry out the economic analysis of official macroeconomic data. In econometrics, two approaches enable to identify and decompose infra-annual series in several parts: the filter-based approach and the model-based approach. Eurostat recommends the use of the model-based approach, in particular the *Tramo/Seats* ([13]) method, and relative softwares like *Demetra+* ([14]) or *JDemetra* developed by Eurostat. The Demetra software was intended to provide a convenient and

([12]) See Timmermann (2006) for a very good introduction.
([13]) See Maravall (1995) for an introduction to Tramo/Seats
([14]) The National Bank of Belgium developed Demetra+ at the request of Eurostat. The Eurostat-European Central Bank (ECB) high level group of experts on seasonal adjustment steered the development work. In 2009, the same group produced the European Statistical System (ESS) guidelines on seasonal adjustment.

flexible tool for seasonal adjustment using Tramo/Seats and X-12-Arima methods. Both methods are divided in two main parts: the first one works on a pre-adjustment and removes the deterministic effects from the series with a regression model and Arima noises, whilst the second decomposes original series.

Tramo/Seats is a seasonal adjustment method developed by Spanish scholars Victor Gómez and Agustin Maravall (Bank of Spain). It consists of two linked programs: Tramo and Seats. Tramo (Time Series Regression with Arima Noise, Missing Observations, and Outliers) performs the estimation, forecasting and interpolation of regression models with missing observations and Arima errors. It also addresses several outliers types. Seats (Signal Extraction in Arima Time Series) performs Arima-based decomposition of observed time series into unobserved components, in particular the seasonal, irregular and the trend-cycle components.

Regarding the seasonal adjustment procedure in external trade by geographical area, it is necessary to consider the following choices:

- the direct or the indirect approach, i.e. whether to apply a transformation on the reference indicators disaggregated by product, or to operate directly on the aggregated final output;
- the frequency for seasonal and working days adjustments (monthly or quarterly).

Corrections have to be made on reference indicators (indirect approach) to produce geographical quarterly data. Then, monthly data are adjusted both for calendar and seasonal signals. Hence, transformed monthly indicators are aggregated at quarterly frequency, before proceeding to the temporal disaggregation.

4.3. Temporal disaggregation

Temporal disaggregation is a mathematical-econometric method that enables to interpolate or to generate unavailable economic time series at selected high-frequency level (monthly or quarterly data). These are temporally consistent with given low-frequency data (in our case annual frequency), the starting point of the procedure. The use of temporal disaggregation techniques is supported by a strong economic argument, as it provides a cheaper alternative to the high costs involved in the direct collection of infra-annual statistical data. Several methods are available to compute quarterly data:

- use of pure mathematical function (linear interpolation, Lisman-Sandee);
- least squares methods (Boot-Feibes-Lisman, Stram-Wei, Jacobs et al. and Denton);
- model based approach (Chow-Lin, Fernandez, Litterman, Guerrero).

The first two techniques are useful when no short-term data are available, and when it is necessary to generate unobserved series using mathematical criteria. Quarterly figures are obtained applying a smoothing method to available data (typically at annual frequency). For example, in the Boot-Feibes-Lisman method, an optimization mathematical rule is applied, which minimizes the square of a differenced sub annual measure, in accordance to the temporal aggregation constraint.

The regression model approach uses information on infra-annual dynamics of one or more proxy variables at the same frequency adjusted with current annual discrepancies. A significant proportion of Italian quarterly national accounts are estimated using these methods. Reference indicators are mainly selected due to their economic and statistical accordance with known annual variables. In a very general way, the quarterly estimate will be:

(5) $\hat{y} = X\hat{\beta} + VC'(CVC')^{-1}(y_* - X_*\hat{\beta})$,

where X is the matrix of reference indicators at quarterly frequency (with intercept), $L = VC'(CVC')^{-1}$ is a smoothing matrix applied to the quantity $(y_* - X_*\hat{\beta})$; the latter represents the discrepancies between annual data y_* and annualized indicators matrix X_* and where the parameter $\hat{\beta}$ is the GLS (Generalized Least Squares) estimator of annualized data:

(6) $\hat{\beta} = (X'_*(CVC')^{-1}X_*)^{-1}X'_*(CVC')^{-1}y_*$,

where V is the quarterly variance-covariance matrix and C is an aggregation matrix ([15]). The Chow-Lin and Fernández, together with the Denton pure benchmarking technique are at this moment, the most commonly used methods within the Italian quarterly national accounts because of their flexibility, easiness of application once it needs to compile quality data in a very limited time. These methods are characterized by the assumptions on the dynamic behavior of annual residuals $(y_* - X_*\hat{\beta})$, which are essential to calculate the a priori unknown covariance matrix V.

4.4. Final geographical area data reconciliation

In this subsection, methods of time series reconciliation in presence of both contemporaneous and temporal aggregation constraints will be briefly discussed. Reconciliation also involves, at the same time, the application of balancing and benchmarking methodologies. In the quarterly accounts estimates, outputs may present discrepancies regarding exogenous or endogenous quantities which are binding constraints: for example, by looking at the annual estimates of imports and exports by geographical area versus the quarterly accounts for the total economy. Hence, a discrepancies treatment has to be applied (see Table 2). In addition, since complete yearly data is not available and once quarters are extrapolated, only the contemporaneous constraint applies.

Table 2: Quarterly geographical area accounts constraints framework

Year	Quarter	EMU	EU-noEMU	Extra-EU	tot
T	q1	Y_{q1}	Y_{q1}	Y_{q1}	$Z_{tot,q1}$
	q2	Y_{q2}	Y_{q2}	Y_{q2}	$Z_{tot,q2}$
	q3	Y_{q3}	Y_{q3}	Y_{q3}	$Z_{tot,q3}$
	q4	Y_{q4}	Y_{q4}	Y_{q4}	$Z_{tot,q4}$
	tot	Y_{c1}	Y_{c2}	Y_{c3}	Z_{tot}

The account system constraints (the marginal values in Table 2) may be of two types:

- contemporaneous constraints, in which, at any given time, total value must be equal to the sum of all areas (last column in Table 2)
- temporal aggregation constraints, in which there must be consistency between high frequency time series and the same data at lower temporal frequency (last row in Table 2).

Discrepancies may be produced by several factors:

([15]) See Chow and Lin (1971), Fernàndez (1981), Denton (1971), or Bee-Dagum and Cholette (2006) for a general introduction to these methodologies.

- different collection data methodologies;
- use of several sample techniques;
- bad specification of regressors in a certain level of disaggregation;
- mathematical transformation on series (i.e. seasonal adjustment);
- use of stochastic models;
- application of forecasting methods.

As reported in di Fonzo and Marini (2011), in the context of reconciliation of temporal data, it is necessary to consider three basic issues:

(1) the dimension of the statistical problem, which is proportional to the number of longitudinal units-variables, and to the length and the frequency of the time span;

(2) the number of constraints imposed on the series;

(3) the preservation of temporal profiles of the preliminary series.

A reconciliation problem may be represented with the following constrained objective function:

$$\text{argmin } (R-P)'\Omega(R-P)$$

$$\text{sub } HR = Y_a$$

where R is the vectorized series of reconciliated final data, P the vectorized series of preliminary data to be balanced, H a partitioned aggregation matrix containing all the restrictions, Y_a the vectorized series of annual data and finally Ω is the variance-covariance matrix which is expression of a proportional first difference between the reconciliated series and preliminary values (Proportional First Differences, PFD). ([16])

$$\Omega = \hat{P}^{-1}(I_m \otimes \Delta'_n \Delta_n)\hat{P}^{-1}$$

Where $\hat{P}^{-1}$ is the inverse of diag(P) and Δ_n is the first difference matrix, in particular:

$$\Delta_n = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

The first order condition is:

$$\begin{bmatrix} \Omega & H' \\ H & 0 \end{bmatrix} \begin{bmatrix} R \\ \lambda \end{bmatrix} = \begin{bmatrix} \Omega P \\ Y_a \end{bmatrix}$$

The simultaneous solution of the system is the following expression:

$$(7)\quad R = P + \Omega^{-1}H'(H\Omega^{-1}H')^{-}(Y_a - HP)$$

where $(H\Omega^{-1}H')^{-}$ is the Moore Penrose generalized inverse matrix. This simultaneous solution could be too laborious from a computational point of view. Fortunately, an easy solution to this problem is available: it is a two-step procedure that uses Denton benchmarking for the

([16]) For details see Bee-Dagum and Cholette (2006) first step and least squares balancing for the second step ([17]). More specifically, the procedure is structured into the following two steps:

(1) for the first step, a benchmarking with Denton method computed via modified proportional first differences (PFD) is applied to restore the temporal additivity of every time series (area);

(2) for the second step, a least squares balancing procedure (de facto a Stone-Byron balancing approach) respecting the contemporaneous and the annual constraints is applied one year at a time. The elements of diagonal variance-covariance matrix are placed equal to the square of the preliminary series ([18]).

Briefly Denton's methodology ([19]) has this generalized form with Lagrangian function

(8) $L = (y - p)' M (y - p) + 2\lambda (z - Hp)$

where y are final benchmarked series, p are preliminary series, H is the usual aggregation matrix, z is the vector of constraints, λ is the usual vector of Lagrange multipliers, and M is a non-singular symmetric linear operator. The matrix M can assume different configurations, depending on the characterization of loss function defined in the objective function and the assumptions about the first observation of the sample period. For example, using a naive benchmarking can lead to a jump in the series due to a significant difference between adjacent annual values. This happens when the matrix M is defined as an identity matrix. The introduction of a difference operator of several orders minimizes or resolves the problem of jump in the series. For instance, the use of a first difference, with the lag operator such that $\Delta y_t = y_t - y_{t-1}$, implies that the modified PFD ([20]) loss function is as follows:

(9) $$K(p, y) = \sum_{t=2}^{T} \left(\frac{y_t - p_t}{p_t} - \frac{y_{t-1} - p_{t-1}}{p_{t-1}} \right)^2 = \sum_{t=2}^{T} \left(\frac{y_t}{p_t} - \frac{y_{t-1}}{p_{t-1}} \right)^2$$

Denton's method, in particular its PFD variant — which has an effect on growth rates — is characterized by preserving dynamics. In other words, the final benchmarked series y must reproduce in the best possible way the movement and the sign of the original series and their growth rate. The movement preservation principle is one of the most important features of the Denton's procedure and is a fundamental numerical feature in data reconciliation techniques.

The Stone-Byron procedure is a consolidated data balancing methodology using regression models ([21]). The problem can be formalized as follows: given a matrix of vectorized preliminary data p, a variance-covariance matrix of discrepancies V, an assumed diagonal, an unknown vector of balanced data y, an aggregation matrix H, a z-constraint vector, and a vector λ of Lagrange multipliers, the Lagrange function to be minimized is:

(10) $L = (y - p)' V (y - p) + 2\lambda(Hy - z)$

The BLUE Stone estimator — that is, the unbiased one with a minimum variance among linear estimators — takes the following form:

(11) $y = p - V H'(HV H')^{-1}(Hp - z)$

([17]) See Quenneville and Rancourt (2005), di Fonzo and Marini (2009) and di Fonzo and Marini (2011) for a technical deepening about two steps approach in reconciliation.)

([18]) The whole of two step procedures are different according the way in which the quadratic additive terms of objective function are normalized.

([19]) See Denton (1971).

([20]) See Cholette (1984).

([21]) See Stone *et al.* (1942), Stone (1961), and Byron (1978) for more information.

so the balanced final data are equal to original series corrected by a linear combination of the discrepancies $(Hp - z)$, modified by a smoothing matrix. This estimator can be calculated directly when $(HVH')^{-1}$ matrix dimensions are reduced, or when it is a dense matrix [22], otherwise, it is necessary to resort to iterative techniques, such as the one adopted for balancing the Italian annual accounts, based on the conjugated gradient algorithm. The variance-covariance V matrix is important to determine the size and direction of the adjustments. The elements of this matrix [23] are not known, so Stone assumed that they are equal to their respective preliminary value p_i, modified by a subjective correction factor θ_i. Let us denote:

$$\sigma_i^2 = \theta_i p_i \quad \forall i$$

Every element of the variance-covariance matrix V is generally weighted with a so-called confidence or alterability parameter [24], depending on the degree of confidence on the reliability of preliminary series. In the exercise introduced in Section 5, it is assumed that the diagonal elements of matrix V are equal to the squared preliminary series in the second step, so:

$$\sigma_i^2 = (\theta_i p_i)^2 \quad \forall i$$

An alternative approach might use a parametric technique to estimate them, or fix them according to a statistical indicator such as the coefficient of variability (as proposed by several scholars). The parameter θ_i has generally a value between zero and one. In the case of preliminary series, whose estimate is considered reliable or constrained for some reason, a value close but not equal to zero is assigned. Conversely, a value close to one could be given to variables that are not well specified or that are considered unconstrained. Currently, the distribution of discrepancies is affected both by the relative values of the alterability parameters and by the relative values of whole variables [25].

5. Application on real data

The application proposed here deals with the computation of quarterly data on imports and exports of goods and services separately in the five geographical areas (European Union countries + EMU Institutions, European Monetary Union countries, European Union countries not included in EMU, EMU Institutions, and Extra-EU countries), using seasonally adjusted and calendar effects corrected data, evaluated at current prices. The temporal horizon considered here is 1999:1-2015:3. As already mentioned, the geographical quarterly series is evaluated in the whole sample, considering the composition of every area at the end of each reference period (evolving composition). This means that countries may move from one area to the other, i.e. from Extra-EU to EU area when a country joins the European Union. This explains the presence of shifts and the variability visible in several time series.

The exercise will focus on two aspects: the forecasting practice on quarterly output of external trade and the calibration experiment on the final reconciliation of the whole system of data by geographical area. All calculations relative to temporal disaggregation, models identification and reconciliation have been developed by using ModelEasy+, Demetra+ for Windows

(22) See di Fonzo and Marini (2011) for computation facilities with sparse matrices.
(23) In most cases only the diagonal elements are of interest.
(24) These parameters could be interpreted as confidence levels respect the individual preliminary series reliability, so they could be parametrized as a subjective 'probability measure'. In our application the alterability parameters are chosen on the basis of a-priori subjective information.
(25) See Section 5 and Table 7 for an exercise with real data.

software. The seasonal adjustments were made with Tramo/Seats (release 5.0) in ModelEasy+ under Linux. The forecasting exercise, the residuals computation and the design of the graphs have been carried out with the open source platform and language R (release 3.0.2 ([26])).

IIn the first part of the exercise, an out-of-sample forecast exercise has been developed directly on aggregated output, considering (i) a logarithmic first difference transformation of the imports and exports of goods variables (as an approximation of growth rate), (ii) an historical dataset from 1999:1 to 2009:4 and (iii) an out-of-sample test dataset from 2010:1 to 2015:3 (representing 23 point-forecasts plus the connected forecast errors for four geographical areas, since the EU-Institutions area is null).

The purpose is to verify the predictive performance of eight different models (six linear and two nonlinear, see Table 3 for an explanatory legend) and two forecast combination types. The first one considers only the simple mean of the predictions of the linear models and the second one computes the arithmetic mean of all available models.

The obtained results are interesting; moreover, they are diversified by geographical area.

First, regarding exports data (Table 4), the combination of forecasts does not always achieve the best outcome, even if there is an improvement in terms of forecast errors minimization. Looking at the Mean Error, there is not a clear indication of persistence in sign of forecast errors. In this context, the Mean Absolute Error and Root Mean Squared Error best performance of SETAR(1) and LSTAR(1) model, for respectively Extra-EU and NO-EMU areas, prove to be very interesting. It shows how nonlinear parametrization may be useful in representation of real dynamic data beyond the scepticism of many users. It also performs very well in forecasting ([27]), even better than linear models.

Table 3: List of parametric models estimated in the out-of-sample forecast performance (see tables 4 and 5)

Model	Description
AR(1)	AutoRegressive of order p = 1
AR(p)1	AutoRegressive of order p (info criterion sel. AICC)
AR(p)2	AutoRegressive of order p (info criterion sel. BIC)
ES	Exponential Smoothing via State-Space
RWD	Random Walk with Drift
ARIMA(p,d,q)(P,D,Q)	AutoRegressive Integrated Moving Average of order (p,d,q)(P,D,Q) (info criterion sel. BIC)
SETAR(1)	Self-exciting Threshold AutoRegressive of order p = 1
LSTAR(1)	Logistic Smooth Transition AutoRegressive of order p = 1
CLM	Combination of Linear Model
CAM	Combination of All available Models

([26]) URL http://www.R-project.org.
([27]) See Granger and Teräsvirta (1993) and Tong (1990) for a general introduction to nonlinear time series theory.

Second, regarding imports data (Table 5), nonlinear models behave very well for EU and EMU areas, in particular the SETAR(1) model. Moreover, when the nonlinear model performs poorly, the combination of forecasts achieves acceptable results. It is important to underline that it would be possible to use a very large number of forecasts combination techniques: we have considered only the easiest and fastest ones for the sake of computation performance, a key criterion once looking at the compilation of quarterly accounts procedures. Another unexpected result is the global good performance of elementary models like AR(1) in both variables and also in all geographical areas.

Table 4: Exports forecast performance classified by areas, statistical models and indices

	AR(1)	AR(p)1	AR(p)2	ES	RWD
			Mean Absolute Error		
EU	3.68	3.943	3.831	3.527	5.161
EMU	4.194	3.922	3.88	4.045	5.776
EU-noEMU	3.949	4.01	4.01	3.968	6.959
Extra-EU	4.35	4.511	5.101	7.862	15.639
			Mean Error		
EU	0.103	1.165	1.05	0.037	-0.524
EMU	0.178	1.284	1.327	-0.35	-0.552
EU-noEMU	-0.096	1.314	1.314	-0.09	-0.435
Extra-EU	0.441	0.74	2.454	-0.08	-0.347
			Root Mean Squared Error		
EU	1.925	2.039	1.933	1.875	2.829
EMU	2.192	2.07	2.041	2.16	3.076
EU-noEMU	1.974	2.054	2.054	1.986	3.579
Extra-EU	2.453	2.611	2.727	3.837	7.244
	ARIMA	**SETAR(1)**	**LSTAR(1)**	**CLM**	**CAM**
			Mean Absolute Error		
EU	3.515	5.195	5.409	3.853	3.715
EMU	3.803	4.132	5.031	3.948	3.773
EU-noEMU	4.01	3.332	2.98	4.365	4.006
Extra-EU	5.101	3.618	3.053	6.298	4.937
			Mean Error		
EU	0.996	-1.735	-1.48	0.471	-0.04
EMU	1.072	- 0.811	-1.513	0.492	0.079
EU-noEMU	1.314	-0.848	-1.279	0.553	0.149
Extra-EU	2.454	1.585	1.516	0.942	1.094
			Root Mean Squared Error		
EU	1.782	2.704	2.724	1.946	1.965
EMU	2.07	2.193	2.745	2.038	2.016
EU-noEMU	2.054	1.749	1.689	2.216	2.032
Extra-EU	2.727	2.46	1.71	3.229	2.616

EU=European Union countries, EMU=European Monetary Union countries, EU-noEMU = European Union countries not included in EMU, Extra-EU=Extra EU countries

Source: Authors' calculations.

In the second part of the exercise, the impact of transformation (seasonal and calendar effects, temporal disaggregation) on quarterly series has been studied in terms of size of discrepancies with respect to the double system of constraints in two macro-sectors (goods and services) and geographical areas. A two-step reconciliation approach has been adopted to solve possible computational impasses of the simultaneous solution illustrated in Section 4.4. The reconciliation procedure concerns only the three basic geographical areas: the European Monetary Union countries + EMU Institutions, the European Union countries not included in EMU and the Extra-EU countries. It covers the EU countries and the aggregation of the first two areas. The EMU Institutions area is excluded from the balancing. It is structured, as already specified in Section 4.4, as follows:

Table 5: imports forecast performance classified by areas, statistical models and indices

	AR(1)	AR(p)1	AR(p)2	ES	RWD
			Mean Absolute Error		
EU	5.147	5.351	5.351	5.54	10.571
EMU	5.223	5.275	5.275	5.781	10.243
EU-noEMU	5.509	5.928	5.928	6.332	12.063
Extra-EU	5.151	5.206	5.361	5.174	6.02
			Mean Error		
EU	-0.272	1.033	1.033	-0.56	-0.999
EMU	-0.266	0.901	0.901	-1.13	-0.985
EU-noEMU	-0.48	1.397	1.397	-0.68	-1.054
Extra-EU	-0.799	0.38	0.662	-0.85	-0.419
			Root Mean Squared Error		
EU	2.82	2.833	2.833	3.144	5.519
EMU	2.821	2.825	2.825	3.171	5.418
EU-noEMU	2.986	3.026	3.026	3.418	6.079
Extra-EU	2.682	2.702	2.768	2.726	3.334
	ARIMA	**SETAR(1)**	**LSTAR(1)**	**CLM**	**CAM**
			Mean Absolute Error		
EU	4.744	3.384	3.505	5.234	4.216
EMU	5.027	3.37	4.667	5.27	4.547
EU-noEMU	4.899	5.595	6.793	5.757	5.396
Extra-EU	5.361	5.552	5.224	5.182	5.19
			Mean Error		
EU	1.415	2.512	2.633	0.275	0.849
EMU	0.961	2.094	3.335	0.063	0.726
EU-noEMU	1.672	-1.322	0.16	0.375	0.136
Extra-EU	0.662	-0.578	0.07	-0.06	-0.11
			Root Mean Squared Error		
EU	2.694	1.788	1.838	2.902	2.358
EMU	2.796	1.861	3.77	2.926	2.523
EU-noEMU	2.726	2.935	3.678	3.183	3.028
Extra-EU	2.768	3.179	2.979	2.684	2.707

EU=European Union countries, EMU=European Monetary Union countries, EU-noEMU = European Union countries not included in EMU, Extra-EU=Extra EU countries

Source: Authors' calculations.

(1) benchmarking using Denton method and computed via modified PFD applied to each series (area);

(2) selection of confidence parameters for longitudinal units and balancing according to contemporaneous and annual constraints.; a least squares balancing procedure is applied one year at a time and elements of variance-covariance matrix are placed equal to the square of preliminary series.

To assess a measure of discrepancies, two indices have to be defined: the *Mean Squared Percentage Adjustment* index (MSPA) and the *Mean Squared Adjustment* (MSA) index.

$$(12) \quad MSPA\,(R_j, P_j) = 100 * \left[\frac{1}{n} \sum_{t=1}^{n} \left(\frac{R_{jt} - P_{jt}}{P_{jt}} \right)^2 \right]^{1/2}$$

$$(13) \quad MSA\,(r_j, p_j) = 100 * \left[\frac{1}{n-1} \sum_{t=2}^{n} (r_{jt} - p_{jt})^2 \right]^{1/2}$$

where the measures *r* and *p* are growth rates, defined as follows

$$r_{jt} = \frac{R_{jt} - R_{jt-1}}{R_{jt-1}} \qquad\qquad p_{jt} = \frac{P_{jt} - P_{jt-1}}{P_{jt-1}}$$

Data reported in Table 6 show the last available in sample observations (third quarter of year 2015). Considering a basic hypothesis of confidence parameters equal to one for all

Table 6: Values of external trade by macro-sector and geographical area, seasonally adjusted data, reference period 2015:3: preliminary, final data and differences
(million EUR)

	EU	EMU	EU-noEMU	Inst.	Extra-EU
			Preliminary series		
Export of goods	53 491.37	39 274.38	14 039.34	0.00	4 4228.44
Export of services	15 511.67	10 908.54	4 603.13	141.44	11 836.07
Import of goods	48 481.68	38 600.90	9 889.19	0.00	35 772.15
Import of services	14 417.06	11 101.68	3 315.38	54.82	10 531.00
			Reconciliated series		
Export of goods	53 467.88	39 411.07	14 056.81	0.00	44 401.79
Export of services	15 595.70	10 979.87	4 615.83	141.44	11 920.03
Import of goods	48 675.02	38 774.45	9 900.58	0.00	35 921.19
Import of services	14 362.35	11 051.45	3 310.90	54.82	10 485.80
			Differences		
Export of goods	23.49	-136.69	-17.47	0.00	-173.35
Export of services	-84.03	-71.33	-12.70	0.00	-83.96
Import of goods	-193.34	-173.55	-11.39	0.00	-149.04
Import of services	54.71	50.23	4.48	0.00	45.20

EU=European Union countries, EMU=European Monetary Union countries, EU-noEMU=European Union countries not included in EMU, Extra-EU=Extra EU countries.

Source: Authors' calculations.

geographical areas, and looking at the boxplot in Figure 2 and at the histograms in Figure 3, it is clear that imports and exports of services present the most relative dispersion in percentage discrepancies. On one hand, this is likely due to the underlying stochastic models and to the not-so-good reference indicators. On the other hand, the extra-EU countries area needs greater adjustment in absolute values. Figure 4 is a graphical comparison between preliminary and final series, with various confidence parameters assigned to this geographical area. Moreover, exports of goods reveal a clear strong positive asymmetry in discrepancies distribution.

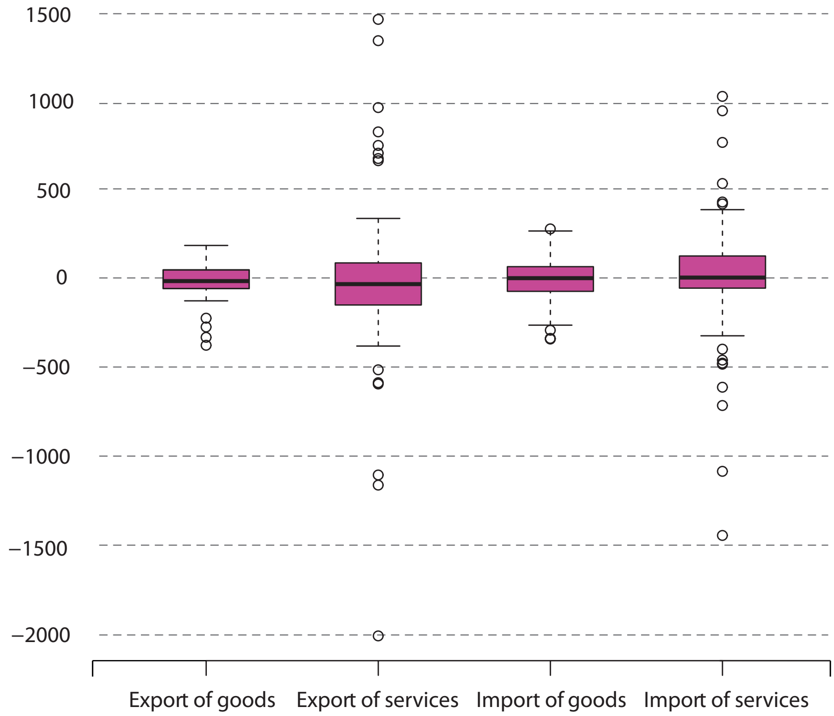
Table 7: Values of adjustment indices by macro-sector and geographical area, several alterability parameters

	EU	EMU	EU-noEMU	Inst.	Extra-EU
Alterability Parameters (1,1,1)					
MSPA					
Export of goods	0.08	0.12	0.04	NA	0.13
Export of services	2.52	3.02	1.09	0.00	3.16
Import of goods	0.15	0.17	0.04	NA	0.16
Import of services	1.78	2.12	0.64	0.00	2.44
MSA					
Export of goods	0.11	0.14	0.04	NA	0.15
Export of services	3.88	4.65	1.68	0.00	4.80
Import of goods	0.21	0.24	0.05	NA	0.23
Import of services	3.23	3.84	1.14	0.00	4.47
Alterability Parameters (0.1,0.0001,1)					
MSPA					
Export of goods	0.07	0.00	0.00	NA	0.26
Export of services	0.05	0.06	0.00	0.00	6.45
Import of goods	0.03	0.00	0.00	NA	0.34
Import of services	0.03	0.04	0.00	0.00	4.39
MSA					
Export of goods	0.04	0.00	0.00	NA	0.30
Export of services	0.07	0.10	0.00	0.00	9.80
Import of goods	0.04	0.01	0.00	NA	0.50
Import of services	0.05	0.07	0.00	0.00	7.98
Alterability Parameters (1,0.0001,0.5)					
MSPA					
Export of goods	0.13	0.21	0.00	NA	0.06
Export of services	3.89	5.36	0.00	0.00	1.40
Import of goods	0.22	0.27	0.00	NA	0.06
Import of services	2.95	3.83	0.00	0.00	1.13
MSA					
Export of goods	0.18	0.24	0.00	NA	0.06
Export of services	5.92	8.07	0.00	0.00	2.12
Import of goods	0.31	0.38	0.00	NA	0.09
Import of services	5.39	6.98	0.00	0.00	2.07

EU=European Union countries, EMU=European Monetary Union countries, EU-noEMU = European Union countries not included in EMU, Extra-EU=Extra EU countries

Source: Authors' calculations.

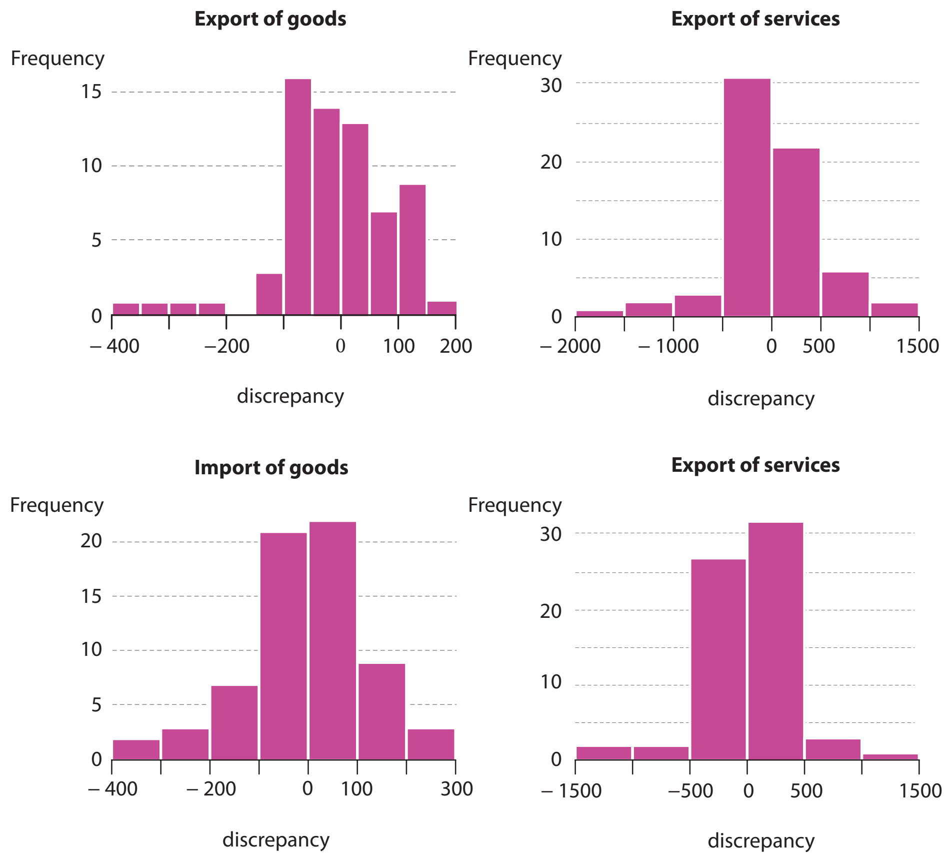
Figure 2: Total discrepancies by macro-sector
(% discrepancies)



Source: Authors' calculations.

Table 7 shows the results of a calibration exercise on confidence parameters used in reconciliation comparing the impact of adjustments in relation to several alterability parameter values. It must be underlined that a confidence parameter $\theta = 1$ (considering the maximum value) involves greater adjustment in balancing. Otherwise, $\theta = 0.0001$ keeps time series unchanged, shifting all the differences to the other units (areas). To reallocate the imbalances, the relative size of parameters among areas is more relevant than their absolute values. Given the context, the relative parameter values has been arbitrarily identified. In the future it will be advisable to identify a parametric procedure and a relative loss function, in order to better compute the relative values of these parameters. In this example, we have put $\theta = 1$ for the Extra-EU area and $\theta = 0.0001$ for EU-nonEMU area to clarify the impact of parameters value. It is clear that the gaps greatly impact Extra-EU countries. The allocation of discrepancies is considerably modified establishing the values $\theta =$ 0.5 for Extra-EU, $\theta = 1$ for EMU countries and $\theta =$ 0.0001 for EU-noEMU countries (clearly, this choice involves MSPA and MSA indices equal to zero).

Figure 3: Histograms of total discrepancies grouped by macro-sector



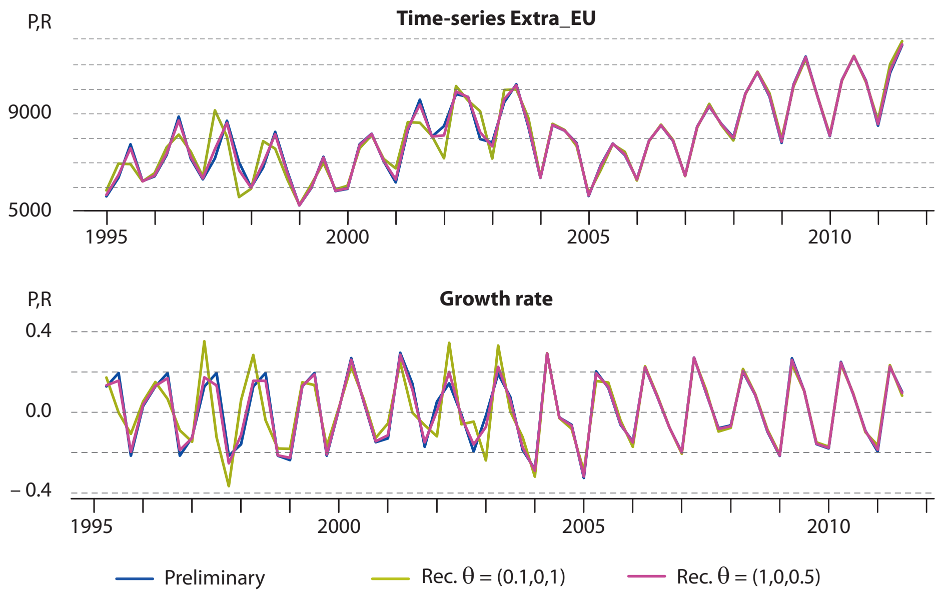
Source: Authors' calculations.

6. Concluding remarks

The ESA 2010 introduced significant changes in national accounts regarding external trade data. Together with the review of imports and exports data, a new estimation procedure for quarterly data broken down respectively by geographical area (proposed for the very the first time) and by country has been produced. These series are evaluated considering the composition of every area at the end of each reference period (evolving composition approach), which clearly affects the features of time series. In this paper, after a brief introduction to external trade innovations, the quarterly estimation procedure by geographical area has been described at length, illustrating all the steps necessary to produce the final quarterly series.

An empirical application has been carried out to determine the best strategy for future improvements of geographical quarterly estimates. An out-of-sample forecast exercise on aggregated outputs (imports and exports of goods) has been developed to verify the performance of a wide collection of dynamic models (six linear and two nonlinear), and of some of their combination (the first one considering only a linear combination of linear models,

Figure 4: Preliminary and final reconciliated series for several confidence parameters, export of services, Extra-EU area



Source: Authors' calculations.

and the other a combination of all considered models). Forecast combinations produce good results in all areas and it is very promising to verify that the two nonlinear models considered (SETAR and Logistic STAR of lag=1) prove to be more performing than the linear ones and also with respect to the naive combinations that has been considered. In terms of discrepancies allocation, the impact of a calibration in alterability parameters has also been examined in the final reconciliation of the accounts system. The results confirm that the value of confidence parameters may heavily influence the destination of discrepancies among areas and therefore, the final quarterly evolution of geographical reconciliated series. This applies strongly when the quarterly discrepancies are very large in some quarter as it occurs for example in the services component data. Hence, for these parameters, it could be more convenient to identify a parametric procedure, through the specification of a loss function, to calculate their relative values.

All in all, a substantial effort has already been made ensuring the consistency of the different estimates and using a reliable statistical basis. However, further improvements could be envisaged in order to make the whole system of external trade accounts fully coherent and methodologically more robust.

Acknowledgements

The authors would like to thank Miss Carla Sciullo for the help provided in the revision of the text and for her critical remarks. Usual disclaimers apply.

References

Bates, J. M. and C. W. J. Granger (1969), 'The combination of forecasts', *Operations Research Quarterly*, 20, pp. 451–468.

Bee-Dagum, E. and P. Cholette (2006), *Benchmarking, temporal distribution and reconciliation methods for time series*, Springer.

Byron, R. P. (1978), 'The estimation of large social account matrices', *Journal of the Royal Statistical Society*, Series A (General), Vol. 141, No. 3, pp. 359–367.

Cholette, P. (1984), 'Adjusting sub-annual series to yearly benchmarks', *Survey Methodology*, 10, pp. 35–49.

Chow, H. and A. Lin (1971), 'Best linear unbiased interpolation, distribution and extrapolation of time series by related series', *The Review of Economics and Statistics*, 53, pp. 372–375.

Denton, F. T. (1971), 'Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization', *Journal of the American Statistical Association*, 66, pp. 99–102.

di Fonzo, T. and M. Marini (2009), 'Simultaneous and two-step reconciliation of systems of time series'. Working Paper Series 9, Università di Padova.

di Fonzo, T. and M. Marini (2011), 'Simultaneous and two-step reconciliation of systems of time series: Methodological and practical issues', *Journal of the Royal Statistical Society C*, 60, pp. 143–164.

Diebold, F. X. and J. A. Lopez (1996), 'Forecast evaluation and combination', in *Handbook of Statistics (Statistical Methods in Finance)* (edited by Maddala, G. S. and Rao, C. R.), pp. 241–268. North- Holland.

Eurostat (2013a), *The European System of Accounts (ESA) 2010*, Luxembourg.

Eurostat (2013b), *Handbook on quarterly national accounts*, Luxembourg.

Fernàndez, R. (1981), 'A methodological note on the estimation of time series', *The Review of Economics and Statistics*, 63, pp. 471–478.

Granger, C. and T. Teräsvirta (1993), *Modelling Non-Linear Economic Relationship,* Oxford University Press.

Granger, C. W. J. and M. J. Machina (2006), 'Forecasting and decision theory', in *Handbook of economic forecasting* (edited by Elliot, G., Granger, C. W. J. and Timmermann, A.), chap. 2, pp. 81–98, Amsterdam, Elsevier.

Hendry, D. (1995), *Dynamic econometrics*, Oxford University Press.

Hendry, D. (2004), 'Unpredictability and the foundations of economic forecasting', Working paper 15, Economics Department, Oxford University.

IMF (2009), *Balance of Payments and International Investment Position Manual, sixth edition,* Washington D.C.

Maravall, A. (1995), 'Unobserved components in economic time series', in *Handbook of Applied Econometrics*, (edited by M. Hashem Pesaran and M.R. Wickens), chap. 1, pp. 12-72, Blackwell.

Quenneville, B. and E. Rancourt (2005), 'Simple methods to restore the additivity of a system of time series', in *Frontiers in benchmarking techniques and their application to official statistics*, Luxembourg.

R Development Core Team (2008), A Language and Environment for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Stock, J. H. and M. W. Watson (2004), 'Combination forecasts of output growth in a seven-country data set', *Journal of Forecasting*, 23, pp. 405–430.

Stone, R. (1961), *Input-output and national accounts*, OECD, Paris.

Stone, R., D. Champernowne and J. Meade (1942), 'The precision of national income estimates', *Review of Economic Studies*, 9, pp. 111–125.

Teräsvirta, T. (2006), 'Forecasting economic variables with nonlinear models', in *Handbook of economic forecasting* (edited by Elliot, G., Granger, C. W. J. and Timmermann, A.), chap. 8, pp. 413–457. Amsterdam, Elsevier.

Timmermann, A. (2006), 'Forecast combinations', in *Handbook of economic forecasting* (edited by Elliot, G., Granger, C. W. J. and Timmermann, A.), chap. 4, pp. 135–196, Amsterdam, Elsevier.

Tong, H. (1990), *Nonlinear time series, a dynamical system approach*, Oxford University Press.

United Nations *et al.* (2009), *System of National Accounts 2008*, New York.

West, K. D. (2006), 'Forecast evaluations', in *Handbook of economic forecasting* (edited by Elliot, G., Granger, C. W. J. and Timmermann, A.), chap. 3, pp. 99–134. Amsterdam, Elsevier.

Getting in touch with the EU

In person
All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europa.eu/contact

On the phone or by email
Europe Direct is a service that answers your questions about the European Union. You can contact this service:
- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by email via: http://europa.eu/contact

Finding information about the EU

Online
Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU publications
You can download or order free and priced EU publications from EU Bookshop at: http://publications.europa.eu/eubookshop. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).

EU law and related documents
For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU
The EU Open Data Portal (http://data.europa.eu/euodp) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.