

Robust official business statistics methodology during COVID-19-related and other economic downturns

Paul A. Smith^{a,c,*} and Boris Lorenc^{b,c}

^a*S3RI and Department of Social Statistics and Demography, University of Southampton, Southampton, UK*

^b*Bright Lynx Research, Tallinn, Estonia*

^c*European Network for Better Establishment Statistics*

Abstract. Official statistics has not properly researched and understood how its methods and models behave at times of downturns (and potentially in the corresponding situation of similarly paced (unpredictable and fast) growths). There is generally a wish to make methods robust to unusual changes, but these are often tackled situation by situation. Production of official statistics during COVID-19 has necessitated some radical changes in both data collection and statistical methods; these have been introduced with admirable speed and dedication, but this process would have been made easier with a body of research already in place to draw from. We discuss the issues with the robustness of statistical methods during downturns, and highlight the opportunity to gather data which can be analysed to give evidence for the most robust methods to use as protection against poor measurement during future downturns.

Keywords: Robust measurement, model stability, intrinsic relationships, statistical relations

1. Introduction

As we wrote this, in mid-2020, the world was fighting a global pandemic as well as enduring a recession not seen in the life-times of most of us living today. Official statistics was stepping up its efforts and introducing novel statistical products that provide a needed factual basis for informing citizens and politicians about the pandemic and its impact, as well as straining to apply existing methods to produce its ordinary statistics under the challenge of both the pandemic and the recession.

This has driven many changes: there has been a very rapid move away from in-person data collection, and the creative use of alternative data sources has increased dramatically, accompanied by changes in governance to support sharing of these data sources for statistical

purposes. There are also many methodological issues, and many of these are linked to the implicit or explicit use of models in statistical processing. Such models appear in many statistics production processes, including sampling, editing, imputation, estimation, and seasonal adjustment. These models are typically based on stable relationships and long-run behaviour, and have needed rapid review and adjustment in light of the unprecedented changes caused by restrictions to businesses' operations to deal with COVID-19. But this is not particular to the current situation – it happens when there is a recession, and there is a need for better study of the performance of methods during these atypical types of periods.

Here we focus on economic downturns as they have proved to be difficult to predict and thus have sudden effects. A corresponding situation in which parallel reasoning might apply would be of similarly paced (unpredictable and fast) growths – though in our view these occur very seldom (while downturns do appear relatively often but unpredictably). The discussion here

*Corresponding author: Paul A. Smith, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. E-mail: p.a.smith@soton.ac.uk.

therefore covers any unpredictable and fast change in the basic circumstances of an economy.

During the pandemic, producers of official statistics faced a number of issues, starting from the state of the sampling frame with respect to the population of interest (e.g. active businesses), unavailability of respondents, obstacles to direct data collection, unreliable validation rule limits for data editing, production and dissemination of relevant statistics for current user needs, and so on. A summary of how an official statistics producer addressed its current challenges is given in [1].

The COVID situation is also an opportunity for official statistics to improve its methods and make them more robust for the future. We mean that there is an opportunity now to reflect on how our methods are performing under the pressures of the recession and the pandemic, to collect empirical evidence about their performance, to conduct analysis of the evidence, and to publish, share and engage in discussions about why some methods, approaches and choices fared better during the crisis than others. These are rare opportunities – the previous recession occurred a decade ago after the 2008–2009 financial crisis, but perhaps was not utilised sufficiently – and it may only be possible to record some of the empirical evidence while a recession lasts. Also, the motivation to understand what causes the methods to underperform during recessions seems, unsurprisingly, to be strongest during recessions and to subside when the economy returns to ‘situation normal’.

Among the broad range of issues, we believe one is of particular interest. It permeates business statistics methodology in any situation but becomes more perceptible during challenges such as an economic downturn. It is the question of reliance on assumptions (explicit or implicit) of the relatedness of variables used in business statistics production; reliance on models (again, explicit or implicit); and reliance on assumptions about behaviour of variables over time (time series). The window of opportunity to collect relevant data to study this situation may already be closed in 2021 as economies recover.

A form of these assumptions often met in official statistics is the use of auxiliary variables, for sampling (like stratified sampling or sampling proportional to a size variable), for validation (where editing limits are commonly set on the basis of such a variable), and for estimation (with regression estimation or calibration). Auxiliary variables are expected or assumed to have *and maintain* certain relations to the variables of interest; the existence of such relations is needed for the

methods to work as expected. Use of auxiliary variables lies at the core of many of the methods developed and used in official business and economic statistics in the preceding 50 years. But such relationships may be disrupted in downturns.

In this discussion we will speak of a *relation* or *relatedness*, explicit or implicit, modelled or not, in order to be as general as possible.

Some of the relations may hold when the economic situation is normal but fail in an economic downturn, while others may hold in both normal times and a downturn. The ones holding during any part of the economic cycle we will term intrinsic and see them as robust. We believe that understanding which relations are robust to an economic downturn is important for production of better economic and business statistics in any future economic downturn (as well as in any ‘normal’ period in between) and that investing additional efforts during the present downturn to measure, collect, analyse, publish and discuss relevant data is an opportunity not to be missed.

In this article we give more background to the problem and its impact, provide some examples of possibly susceptible approaches, introduce the concept of robustness in this context, based on intrinsic relations between the variables, and review some institutional strategies we feel may be useful in improving the quality of official business and economic statistics in any conditions (downturn or not).

2. Background

Official statistics make use of models and assumptions as part of their production processes, and considerable research efforts are expended to make sure that the most appropriate methods are used. This research typically proceeds by gathering a selection of historical data, and then analysing it to examine the relationships between the variables of interest. These relationships are utilised to improve the estimates for the phenomena being measured. This kind of approach is used in many different types of official statistics, including those traditionally depending on sampling and surveys, through methods which make implicit and explicit use of models, and on to data science where techniques such as machine learning are used to discover and use relationships in data.

The historical data are frequently the most recent periods available, which makes sense if the primary interest is in up-to-dateness; it is logical to base decisions

about what to do in the future periods on the relationships found at times closest to those periods, since the relationships are most likely to be similar. However, to (mis)quote from investment advertising, “past performance is no guarantee of future results”, and so it is important to ensure that such investigations use robust methods that can deal with departures from the usual behaviour of series.

We may be confident that departures from usual behaviour will occur, be they downturns related to economic statistics or changes in subjective wellbeing due to changing conditions (of which the pandemic provides a clear example). However, the precise form of the departures is not known in advance. When they occur, there is heightened interest in official statistics as the factual basis for forecasting and policy formulation. But they are also the periods when those relationships and models that we rely on to support our statistical production processes are most likely to be affected, and this creates the paradox that official statistics are most scrutinised just at the time when they are least certain.

That there are risks for the quality of business and economic statistics produced at times of economic downturns has been noted before, not the least at the European Network for Better Establishment Statistics (ENBES)’s first workshop in 2009, which was held just at the end of the “Great Recession” of 2007–2009 [2].

Official statistics has not really researched and understood how its methods and models behave at these key times. There is a general wish to make methods robust to unusual changes, but these are often tackled situation by situation. More review, synthesis and targeted investigation seems to us to be in order. We would be delighted if this article exposes many examples of such studies and reviews, which would be a valuable resource for building these approaches into official statistics.

3. Not a disaster, not “special”

The downturns we are discussing are those which appear as an expected, if irregular, part of the economic cycle. For the purpose of this article, we specifically exclude dealing with catastrophic disasters, which generally require specific, targeted procedures to estimate the impact, not least because there is a comparison with an implicit counterfactual in which the disaster did not occur. See for example [3,4].

In the same vein there may be transient effects on series from severe weather, or particular events, and we are also not targeting these, which are likely not to

have a strong impact over more than one period. [5] discusses these types of “special events”.

4. Some examples of susceptible approaches

Generally, to process and produce results on variables of interest, official statistics relies on some auxiliary variables. In stratified sampling, a classical example of an auxiliary variable is the preceding period’s value of the same variable of interest (say, turnover the year before). At times of downturn, the relationship will be affected; inasmuch as every business is affected similarly, the procedures react appropriately to the changed values in data sources. But recessions often affect different businesses differently, which means that the relationship between the collected and auxiliary variables will be noisier (that is, not be as strong as in normal times). This will increase the variance of the estimate and – depending on the estimator used – will have a negligible or a not so negligible effect on the bias.

A part of the risk involved stems from the methodologists’ reliance on some explicit or implied assumptions, or implicit or explicit models. In editing, a score for selective editing may employ the preceding period’s value to identify for editing those observations with significant deviations from their historical values. Imputation models also rely on relationships, and may not work as intended when there are large changes in series. The ONS has recently examined the imputation for hours worked in its Labour Force Survey during the lockdown period to deal with the threat of COVID-19, saying:

“Imputation used for the LFS was not designed to deal with the changes experienced in the labour market in recent months. Experimental work with adjusted methodology suggests the use of the existing methodology has understated the reduction in the actual numbers of hours worked by approximately 5% to 6%.” [6]

Further examples of this kind can be given with respect to estimation, outlier detection, etc.

For instance, the headline estimates from the Netherlands’ Labour Force Survey are derived from a modelling procedure which uses the autocorrelation pattern in the series to improve the quality of the estimates [7]. But at times of quick turns in the economic cycle the unsupervised model may not react quickly enough, and therefore needs to be reviewed [8].

Then there are impacts affecting seasonal adjustment of time series, which rely on repeating patterns of corre-

lation over time, which may be substantially disrupted during a downturn. These methods then require expert intervention to make a suitable judgement on the balance between the signal and noise in the series. In the same way as before, this makes the seasonal adjustment less certain just at the moment that the series are most under scrutiny. Although the methodology for assessing the mean squared error accounting for seasonal adjustment has been developed [9], it does not seem to have been applied to consider the changes in quality of seasonal adjustment during recessions.

As an example of an assumption, [10] discuss the way that the balance of births and deaths of businesses is incorporated in business survey estimation in the ONS, and that it may be less valid during a downturn when business formation is generally at its lowest. National accounts processes require some assumptions about relationships or even direct proportionality of different components. It is challenging even to keep track of all of these so that the effects of a downturn can be considered for each, but such assumptions are typically the result of a lack of data sources, so even in the case where the assumptions are not felt to be reasonable, there may be no alternative evidence with which to make an adjustment.

Data science approaches which are predicated on finding relationships among variables and using them as the basis for statistical outputs, are clearly susceptible to changes in those relationships. They also need robustness in training and regular review, the same kinds of procedures as in statistical approaches, but often with different terminology.

5. Robustness and intrinsic relationships

While use of a preceding period's values may work well when the economy is changing smoothly, there are clearly risks associated with their use at times of economic turmoil. Then the preceding period is not such a good predictor of the performance in the current period. The stratification will be less than optimal, the ratio estimator's bias might become noticeable, the editing efforts might be misdirected and far more resource demanding than actually intended.

Preceding period's values of interest at the micro-data level have a relation to the current period's values through the knowledge that they refer to the same unit (same business). However, the circumstances in which a business is operating may change, for economic cycle reasons or, as recently, a pandemic. The relation for

which the preceding period's values are used in producing the statistics of interest is based on the values' "good behaviour" through time, which is in turn based on the smooth performance of the economy. We refer to this as the "statistical relation" of an auxiliary variable and a variable of interest.

To this can be contrasted an auxiliary variable from the same time period, for instance (just as an example, which would need empirical validation) turnover as the variable of interest and employment as an auxiliary variable. As the current economic downturn clearly demonstrates, if the turnover of a business is impacted, in a subset of industrial sectors its employment will also be affected. Refer to this as the "intrinsic relation" of an auxiliary variable and a variable of interest. Businesses in these sectors simply function so that if there is a loss of turnover, it is highly likely there will also be a loss of employees.

We believe knowledge of which variables (indicators of business's performance) are intrinsic and which only statistical can increase business statistics methodologies' robustness against variations in the economic cycle and safeguard against the difficult to predict downturns.

To distinguish between intrinsic and only statistical relations is not simple. During the COVID period there is an opportunity – not existing in 'normal' times – to gather empirical evidence and learn. Survey methodologists should collaborate with business economists and experts in business functions to further their knowledge of business operations and identify intrinsic relations.

6. A minority of periods ... but which ones?

When we are undertaking research using past data, recessions (and other similar changes) affect a minority of periods, if they appear in the data at all. So the training data are typically unbalanced, heavily favouring times of growth, which predominate. If we are using mean behaviour to make a judgement about the best approach, which is typical, there is no guarantee that the selected approach will be suitable for downturn periods. Yet it is during those periods that we rely on having methods which are as objective as possible, because we need them to continue to work well when things are least certain.

This suggests some avenues for research and collaboration. Each country individually has relatively small amounts of information about downturns, but if this could be brought together, there would be potential for

learning much more about the performance of methods in this situation and possible alternatives. We could also consider weighting the training data so that more emphasis is placed on the downturn periods; but we still need to be sure that the approaches work well in normal times too.

At first sight it may seem that different methods or models should be used in downturns to cope with the changes in relationships. However, this would require a crystal ball to know when a downturn was happening (or about to happen), to know when to change methods. Even determining which are the right periods to upweight in a training data set is not obvious... what is the reach of a downturn?

Therefore we need methods which are robust to downturns and work effectively both at these times and in normal situations. Such a happy situation may be easier to describe than to achieve! But we still need to make an assessment of the methods' properties so that we can assess how well that robustness is achieved.

In situations where a robust one-size-fits-all approach is not possible (or has not been discovered yet), what is the best approach? Should we have adaptive methods which change in an automated (and objective) way? Or is human intervention the only satisfactory solution?

7. Institutional research strategies

Interest in recessions peaks when we are in one, as then a National Statistical Institute (NSI)'s outputs are of particular and wide interest. In our experience this is translated within an NSI into questions about whether the methods in statistical production are the most appropriate for the current situation, whether they are affected by the recession itself, and whether some variation in procedures might be more appropriate to deal with the downturn. But, it has been thought that this is not the right time to do the required research, which needs time and resources when they are already most stretched. Once the downturn passes, there is no longer a pressing urgency for these answers, so the research is not followed through. Therefore at the next recession the same questions are asked, still without answers.

[2,10] set out a series of investigations ("tasks") following the 2007–9 recession, and some of these were undertaken in the ONS. Some, such as reviewing seasonal adjustments, require only a deeper look with existing techniques. But other tasks, involving evaluating estimation approaches and checking the assumptions underlying a range of methods, have not been written up and disseminated, and it is not clear how many of these tasks were actually completed.

8. Conclusions

The COVID-related economic downturn is an opportunity for business survey methodologists to gather empirical evidence on which methods and choices performed well during the economic downturn and which didn't. Sharing analyses of empirical evidence and bringing country experiences together is essential. A systematic review of this evidence will improve knowledge of the robustness of the methods and improve the choice and tuning of methods so that they perform better in future downturns

Towards that goal, we recommend considering which variables – among those that are in use or can be accessed – are only statistically related, and thus prone to model breakdowns in times of economic swings, and which are intrinsically related and thus less risky in models and similar parts of statistical processing. That theoretical understanding, supported by empirical evidence collected during economically challenging times like these, may permanently bring the practice of business survey methodology to safer ground.

References

- [1] HCSO. HCSO Covid-19 Activities. European Network for Better Establishment Statistics (ENBES) website, 2020. [updated 2020 Sep 18; cited 2021 Sep 4]. Available from: <https://bit.ly/2EzN6au>.
- [2] Brand M. Methodological issues arising for the Office for National Statistics from the recession. European Establishment Statistics Workshop, Stockholm, 2009. Available from: https://statswiki.unece.org/download/attachments/122323607/2009_S1_2_Brand.pdf.
- [3] Session IPS073: Response to natural disasters. World Statistics Congress, Hong Kong, 2013. Available from: <https://2013.isiprceedings.org/index.php?pn=1&r1=L&r4=All&r3=NR&r5=All&r6=1&r7=IPS073&rows=242>.
- [4] Lam NS, Arenas H, Pace K, LeSage J, Campanella R. Predictors of business return in New Orleans after Hurricane Katrina. *PLoS One*. 2012; 7(10): e47935. doi: 10.1371/journal.pone.0047935.
- [5] Smith P. Developing a policy on special events in ONS. Newport: ONS; 2012. Available from <https://webarchive.nationalarchives.gov.uk/20121103011233/http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/special-events-group/developing-a-policy-on-special-events.pdf>.
- [6] ONS. Employment in the UK: July 2020. Newport: ONS; 2020. Available from: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/employmentintheuk/july2020>.
- [7] Van den Brakel JA, Krieg S. Estimation of the monthly unemployment rate through structural time series modelling in a rotating panel design. *Surv Meth*. 2009; 35: 177–190.
- [8] Van den Brakel J, Souren M, Krieg S. Estimating monthly labour force figures during the COVID-19 pandemic in the Netherlands. Discussion Paper. The Hague: Statistics Nether-

- lands. Available from: <https://www.cbs.nl/en-gb/background/2021/08/monthly-labour-force-figures-during-covid-19>.
- [9] Pfeffermann D, Sverchkov M. Estimation of mean squared error of X-11-ARIMA and other estimators of time series components. *J Off Stat.* 2014; 30: 811–838.
- [10] Simkins A, Smith P, Brand M. Financial crisis and recession: How ONS has addressed the statistical and analytical challenges. *Econ Lab Market Rev.* 2010; 4: 30–37.