ESSAYS IN HONOR OF JOON Y. PARK

ADVANCES IN ECONOMETRICS

Series editors: Thomas B. Fomby, R. Carter Hill, Ivan Jeliazkov, Juan Carlos Escanciano, Eric Hillebrand, Daniel L. Millimet, Rodney Strachan

Recent Volumes:

Volume 27A	Missing Data Methods: Cross-sectional Methods and Applications – Edited by David M. Drukker
Volume 27B	Missing Data Methods: Time-series Methods and Applications – Edited by David M. Drukker
Volume 28	DSGE Models in Macroeconomics: Estimation, Evaluation and New Developments – Edited by Nathan Balke, Fabio Canova, Fabio Milani and Mark Wynne
Volume 29	Essays in Honor of Jerry Hausman – Edited by Badi H. Baltagi, Whitney Newey, Hal White and R. Carter Hill
Volume 30	30th Anniversary Edition – Edited by Dek Terrell and Daniel Millmet
Volume 31	Structural Econometric Models – Edited by Eugene Choo and Matthew Shum
Volume 32	VAR Models in Macroeconomics – New Developments and Applications: Essays in Honor of Christopher A. Sims – Edited by Thomas B. Fomby, Lutz Kilian and Anthony Murphy
Volume 33	Essays in Honor of Peter C. B. Phillips – Edited by Thomas B. Fomby, Yoosoon Chang and Joon Y. Park
Volume 34	Bayesian Model Comparison – Edited by Ivan Jeliazkov and Dale J. Poirier
Volume 35	Dynamic Factor Models – Edited by Eric Hillebrand and Siem Jan Koopman
Volume 36	Essays in Honor of Aman Ullah – Edited by Gloria Gonzalez-Rivera, R. Carter Hill and Tae-Hwy Lee
Volume 37	Spatial Econometrics – Edited by Badi H. Baltagi, James P. LeSage and R. Kelley Pace
Volume 38	Regression Discontinuity Designs: Theory and Applications – Edited by Matias D. Cattaneo and Juan Carlos Escanciano
Volume 39	The Econometrics of Complex Survey Data: Theory and Applications – Edited by Kim P. Huynh, David T. Jacho-Chávez and Guatam Tripathi
Volume 40A	Topics in Identification, Limited Dependent Variables, Partial Observability, Experimentation, and Flexible Modelling Part A – Edited by Ivan Jeliazkov and Justin L. Tobias
Volume 40B	Topics in Identification, Limited Dependent Variables, Partial Observability, Experimentation, and Flexible Modelling Part B – Edited by Ivan Jeliazkov and Justin L. Tobias
Volume 41	Essays in Honor of Cheng Hsiao – Edited by Tong Li, M. Hashem Pesaran and Dek Terrell
Volume 42	The Econometrics of Networks – Edited by Áureo de Paula, Elie Tamer and Marcel-Cristian Voia
Volume 43A	Essays in Honor of M. Hashem Pesaran – Edited by Alexander Chudik, Cheng Hsiao and Allan Timmermann
Volume 43B	Essays in Honor of M. Hashem Pesaran – Edited by Alexander Chudik, Cheng Hsiao and Allan Timmermann
Volume 44A	Essays in Honour of Fabio Canova – Edited by Juan J. Dolado, Luca Gambetti and Christian Matthes
Volume 44B	Essays in Honour of Fabio Canova – Edited by Juan J. Dolado, Luca Gambetti and Christian Matthes
Volume 45A	Essays in Honor of Joon Y. Park – Edited by Yoosoon Chang, Sokbae Lee and J. Isaac Miller

ADVANCES IN ECONOMETRICS VOLUME 45B

ESSAYS IN HONOR OF JOON Y. PARK: ECONOMETRIC METHODOLOGY IN EMPIRICAL APPLICATIONS

EDITED BY

YOOSOON CHANG
Indiana University, USA

SOKBAE LEE
Columbia University, USA

And

J. ISAAC MILLER
University of Missouri, USA

emerald PUBLISHING

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2023

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-83753-213-1 (Print)
ISBN: 978-1-83753-212-4 (Online)
ISBN: 978-1-83753-214-8 (Epub)

ISSN: 0731-9053 (Series)

CONTENTS

PART III
PANDEMIC, CLIMATE, AND DISASTER

PART IV
MICROECONOMETRICS AND PANEL DATA

PART V
RETROSPECTIVE

LIST OF CONTRIBUTORS

Martín Almuzara	Federal Reserve Bank of New York, USA
Ian J. Bateman	University of Exeter, UK
Randall C. Campbell	Mississippi State University, USA
Marine Carrasco	University of Montreal, Canada
Saraswata Chaudhuri	McGill University, Canada
Gabriele Fiorentini	Universitá di Firenze and RCEA, Italy
Alain Hecq	Maastricht University, Netherlands
Cheng Hsiao	University of Southern California, USA
Rustam Ibragimov	Imperial College London, UK; and St. Petersburg State University, Russia
J. Isaac Miller	University of Missouri, USA
Whayoung Jung	Korea Capital Market Institute, South Korea
Shakeeb Khan	Boston College, USA
Lutz Kilian	Federal Reserve Bank of Dallas, USA; and CEPR, UK
Namhyun Kim	University of Exeter, UK
Ji Hyung Lee	University of Illinois Urbana-Champaign, USA
Yoonseok Lee	Syracuse University, USA
Arnaud Maurel	Duke University and NBER, USA; and IZA, Germany
Lealand Morin	University of Central Florida, USA
Florens Odendahl	Banco de España, Madrid, Spain
Asli Ogunc	Texas A&M University-Commerce, USA
Eric Renault	University of Warwick, UK
Barbara Rossi	University Pompeu Fabra, Spain
Tatevik Sekhposyan	Texas A&M University, USA
Enrique Sentana	CEMFI, Spain; and CEPR, UK
Donggyu Sul	University of Texas at Dallas, USA
Idriss Tsafack	University of Montreal, Canada
Elisa Voisin	Maastricht University, Netherlands

Oscar Wahlstrom	Brown University, USA
Patrick Wongsa-art	Cardiff University, UK
Zeyu Xing	Imperial College London, UK; and UBP Investment Management, China
Yichong Zhang	Singapore Management University, Singapore
Qiankun Zhou	Louisiana State University, USA
Xiaoqing Zhou	Federal Reserve Bank of Dallas, USA

INTRODUCTION

Volume 45 of *Advances in Econometrics* honors Professor Joon Y. Park, who has made numerous and substantive contributions to the field of econometrics over a career spanning four decades since the 1980s and counting. Volume 45 consists of 28 chapters and is in fact split between two volumes with the first focusing on econometric theory and the second focusing on econometric applications. These papers have been contributed by Joon's friends, colleagues, coauthors, former students, and even his dissertation advisor, Professor Peter C. B. Phillips, and the volume is edited by his wife and most frequent collaborator, Professor Yoosoon Chang, and two of his former students.

In the typical fashion of *Advances in Econometrics*, the papers were to be submitted in early 2021 after a conference in Joon's honor in April 2020, which would have nearly coincided with his 65$^{\text{th}}$ birthday. Of course, the COVID-19 pandemic forced much of the world into lockdown in April 2020, so plans changed. Papers were still submitted in 2021, but the conference was delayed and, as of this writing, is scheduled for September 29–30, 2023, in Bloomington, Indiana, which Joon and Yoosoon have called home for nearly 15 years.

We introduce the 15 chapters of the second volume. The first 14 are grouped into 4 sections that are related – some closely and some very loosely – to Professor Park's work and especially to his more recent work. We conclude the volume with a retrospective article summarizing four decades of this series, *Advances in Econometrics*.

The first two decades of Joon's published record is dominated by contributions to theoretical time series that relate to empirical macroeconomics, broadly defined, more closely than to any other field outside of econometrics. However, much of his work then and up to the present has been motivated by a sincere interest in how the tools he pioneered could be used in empirical applications. These methodologies have influenced empirical work of all sorts.

In macroeconometrics, he has published influential work on regime switching (Chang et al., 2017) and common stochastic trends (Chang et al., 2010), for example. Intensively studying high-frequency time series in the early 2000s, his contributions to financial econometrics include Ait-Sahalia and Park (2012), Choi, Jeong, and Park (2014), and Kim and Park (2017). His recent record contains several well-cited papers with methodologies motivated by understanding and forecasting energy consumption for the Republic of Korea and the world more generally, including Chang et al. (2014) and Chang et al. (2021). Related to energy consumption is climate change, and Park continues to make contributions to the econometric analysis of climate change, starting with Chang et al. (2020).

Following the themes mentioned above, the chapters in this volume are grouped as follows: (I) macroeconometrics, (II) financial econometrics, (III) pandemic, climate, and disaster, and (IV) microeconometrics and panel data.

PART I: MACROECONOMETRICS

We open the volume with a contribution by Martín Almuzara, Gabriele Fiorentini, and Enrique Sentana, entitled “Aggregate Output Measurements: A Common Trend Approach,” which relates the study of common trends, an area in which Professor Park has contributed significantly, to macroeconomic aggregates. The authors analyze a model for different measurements of a single persistent latent time series with mean-reverting measurement errors, thereby allowing for a common trend among these measurements. They find that over-differencing drives potentially large biases in estimation and reduces the precision of smoothed estimates of the latent variable. They obtain an improved aggregate output measure using US quarterly data.

Forecast rationality is a key principle of macroeconometrics. While existing forecast rationality tests either focus on constant deviations from forecast rationality over the full sample or are constructed to detect smooth deviations based on nonparametric techniques, in “Markov Switching Rationality,” Florens Odendahl, Barbara Rossi, and Tatevik Sekhposyan propose novel parametric tests for detecting Markov switching deviations from forecast rationality. They find that forecasters tend to systematically overpredict interest rates during periods of monetary easing, while the forecasts are unbiased otherwise. Their findings emphasize the special role played by monetary policy in shaping interest rate expectations above and beyond macroeconomic fundamentals.

Energy is a factor of production in the macroeconomic production function of any country, so few commodities are as important as oil in understanding macroeconomic fluctuations. Lutz Kilian and Xiaoqing Zhou survey the extensive literature on oil market VARs in their contribution, “The Econometrics of Oil Market VAR Models.” As this literature has expanded at a rapid pace, it has become increasingly difficult for most economists to track the differences between alternative oil market models and the basis for divergent conclusions reached in the literature. This survey provides a useful guide, with a particular focus on the econometric foundations of the analysis of oil market models.

PART II: FINANCIAL ECONOMETRICS

In their contribution entitled “Quantile Impulse Response Analysis with Applications in Macroeconomics and Finance,” Whayoung Jung and Ji Hyung Lee study the dynamic responses of the conditional quantiles and their applications in macroeconomics and finance. This chapter builds a multi-equation autoregressive conditional quantile model and proposes a new construction of quantile impulse response functions (QIRFs). The new QIRF toolset the authors provide adds nicely to the burgeoning research efforts that have been devoted to measuring distributional effects of economic shocks. Using the QIRFs, the authors find that the left tail of economic activity is most responsive to monetary and financial market shocks, and they use this result to evaluate the impact of economic shocks on the 5% quantile of economic activity, a measure of growth-at-risk, during the global financial crisis.

In “Risk Neutral Density Estimation with a Functional Linear Model,” Marine Carrasco and Idriss Tsafack propose a nonparametric estimator of the risk neutral density based on cross-sectional European option prices. They show that the risk neutral density can be viewed as the solution of an ill-posed integral equation and estimate it using an iterative method called Landweber-Fridman. They establish the consistency and asymptotic normality of their estimator and provide an application to S&P 500 options.

Lealand Morin, in his contribution entitled “Estimating Diffusion Models of Interest Rates at the Zero Lower Bound: From the Great Depression to the Great Recession and Beyond,” proposes a new way to properly estimate a class of parametric diffusion models that can be used to represent the interest rate over a long time span possibly including several episodes where the interest rate may stay near or at the zero lower bound. This approach makes it easier to learn about the interest rate dynamics from major historic zero lower bound episodes in the United States, most notably the Great Depression and Great Recession. This enhanced understanding may help us predict future responses of key macroeconomic variables to the interest rate that has recently gone through a new episode of zero lower bound, from the outset of the COVID-19 pandemic to monetary policy tightening implemented to moderate inflation in early 2022.

Rapid stock market growth without real economic back-up has led to the 2015 Chinese stock market crash. In “A Market Crash or Tail Risk? Heavy Tails and Asymmetry of Returns in the Chinese Stock Market,” Zeyu Xing and Rustam Ibragimov analyze structural breaks in heavy-tailedness and asymmetry properties of returns in Chinese A-share markets due to the crash using robust methods for inference on the tail index. Their empirical results show that the main determinants of heavy-tailedness in Chinese financial markets are liquidity and company size.

PART III: PANDEMIC, CLIMATE, AND DISASTER

Continuing with the theme of crashes, the next set of articles relates to disasters past, present, and future, with an emphasis on the COVID-19 pandemic and climate change. Alain Hecq and Elisa Voisin contribute “Predicting Crashes in Oil Prices During the COVID-19 Pandemic with Mixed Causal-Noncausal Models,” which sheds light on how data transformations can substantially impact predictions made by mixed causal–noncausal models that rely on specifications in which time series depend not only on their lags but also on their leads. The authors investigate oil prices and estimate probabilities of crashes before and during the first wave of the COVID-19 pandemic in 2020, comparing various mechanical detrending methods with a detrending performed using the level of strategic petroleum reserves.

Yoonseok Lee and Donggyu Sul also investigate the recent pandemic in their contribution, “Depth-Weighted Forecast Combination: Application to COVID-19 Cases.” They develop a novel forecast combination approach based on the order statistics of individual predictability from panel data forecasts. Defining the notion of forecast depth based on normalized forecast errors during the training period, they derive the limiting distribution of the depth-weighted forecast combination. Using this novel forecast combination, they predict the national level of new COVID-19 cases in the United States and find that the proposed method yields more accurate and robust predictions compared with other popular forecast combinations, including the ensemble forecast from the Centers for Disease Control and Prevention.

While the recent COVID-19 pandemic provides a tangible example of an economic disaster, economic disasters may have disparate causes. Saraswata Chaudhuri, Eric Renault, and Oscar Wahlstrom examine economic disasters more broadly in their contribution, entitled "Identification of Beliefs in the Presence of Disaster Risk and Misspecification." They reconsider the equity premium puzzle and related asset market puzzles in light of the effect of rare disasters on asset prices. Low-probability economic disasters can restore the validity of model-implied moment conditions only if the amplitude of disasters may be arbitrarily large in due proportion. Yet they prove that there is no such thing as a population empirical likelihood-based model-implied probability distribution in the presence of unbounded disasters.

The next two chapters do not consider disasters explicitly, but there is a broad consensus within the scientific community on the potential for climate change to induce economic disasters. Land use is an issue that is inextricably tied to both the effects of climate change, by way of changes in arable land, for example, and to the causes of climate change, by way of changes in albedo, or reflection of solar energy. In their contribution "A New Model for Agricultural Land-Use Modeling and Prediction in England Using Spatially High-Resolution Data," Namhyun Kim, Patrick Wongsa-art, and Ian J. Bateman contribute to a better understanding of farmers' responses to behavioral drivers of land-use decisions by establishing an alternative analytical procedure that overcomes various drawbacks suffered by methods currently used in existing studies. Specifically, high-resolution spatial data ameliorates the idiosyncratic effects of the physical environment, and their model is equipped to deal with censoring, spatial dependence, and heterogeneity in the data and errors.

Also on the topic of spatial heterogeneity, J. Isaac Miller contributes "Local Climate Sensitivity: What Can Time Series of Distributions Reveal About Spatial Heterogeneity of Climate Change?" He introduces an easily implemented semiparametric statistical approach based on a physical energy balance climate model to estimate net heat transport and allow for spatial heterogeneity in the response of temperature to climate forcings. He finds that areas dominated by ocean tend to import energy and are relatively more sensitive to climate forcings, but that these areas warm more slowly than areas dominated by land.

PART IV: MICROECONOMETRICS AND PANEL DATA

In "Maximum Likelihood Estimation of Dynamic Panel Data Models with Interactive Effects: Quasi-Differencing Over Time or Across Individuals?" Cheng Hsiao and Qiankun Zhou consider the quasi-maximum likelihood estimation (MLE) of dynamic panel models with quasi-differencing to remove interactive effects. They show that the quasi-difference MLE over time is inconsistent when T is large, whether N is fixed or large, and is consistent and asymptotically unbiased when the difference is across individuals when N is large, whether T is fixed or large.

Factor structures have been employed in a variety of settings in cross sectional and panel data models. In "Informational Content of Factor Structures in Simultaneous Binary Response Models," Shakeeb Khan, Arnaud Maurel, and Yichong Zhang investigate the informational content of factor structures in discrete triangular systems. Their main finding is that imposing a factor structure yields point identification of parameters of interest, such as the coefficient associated with the endogenous regressor in the outcome equation, under weaker assumptions than are usually required in these models.

PART V: RETROSPECTIVE

Advances in Econometrics is a series of research volumes first published in 1982. Professor Park's contribution of the variable addition test for cointegration to *Advances in Econometrics* in Park (1990) is both one of his most highly cited works and one of the most highly cited in the series. So, it seems appropriate to conclude this volume in his honor with a retrospective piece on the series. Asli Ogunc and Randall C. Campbell, with "Forty Years of *Advances in Econometrics*," present an update to the history of the *Advances in Econometrics* series published in 2012. They describe key events in the history of the series and provide information about key authors, contributors, and other historical data on the series.

One of the joys of compiling this volume, "*Essays in Honor of Joon Y. Park: Econometric Methodology in Empirical Applications*," has been to see the many in ways in which Joon's research has influenced the methodologies used in empirical applications throughout the discipline. We hope you enjoy reading these chapters as much as we have.

ADDITIONAL REFERENCES

Ait-Sahalia, Y., & Park, J. Y. (2012). Stationarity-based specification tests for diffusions when the process is nonstationary. *Journal of Econometrics*, *169*, 279–292.

Chang, Y., Choi, Y., Kim, C. S., Miller, J. I., & Park, J. Y. (2021). Forecasting regional long-run energy demand: a functional coefficient panel approach. *Energy Economics*, *96*, 105117.

Chang, Y., Choi, Y., & Park, J. Y. (2017). A new approach to model regime switching., *Journal of Econometrics*, *196*, 127–143.

Chang, Y., Kaufmann, R. K., Kim, C. S., Miller, J. I., Park, J. Y., & Park, S. (2020). Evaluating trends in time series of distributions: A spatial fingerprint of human effects on climate. *Journal of Econometrics*, *214*, 274–294.

Chang, Y., Kim, C. S., Miller, J. I., Park, J. Y., & Park, S. (2014). Time-varying long-run income and output elasticities of electricity demand with an application to Korea. *Energy Economics*, *46*, 334–347.

Chang, Y., Miller, J. I., & Park, J. Y. (2009). Extracting a common stochastic trend: Theory with some applications. *Journal of Econometrics*, *150*, 231–247.

Choi, H., Jeong, M., & Park, J. Y. (2014). An asymptotic analysis of likelihood-based diffusion model selection using high frequency data. *Journal of Econometrics*, *178*, 539–557.

Kim, J., & Park, J. Y. (2017). Asymptotics for recurrent diffusions with application to high frequency regression. *Journal of Econometrics*, *196*, 37–54.

Park, J. Y. (1990). Testing for unit roots and cointegration by variable addition. In G. F. Rhodes & T. B. Fomby (Eds.), *Advances in econometrics* (pp. 107–133). JAI Press.

PART I

MACROECONOMETRICS

This page intentionally left blank

CHAPTER 1

AGGREGATE OUTPUT MEASUREMENTS: A COMMON TREND APPROACH

Martín Almuzara[a], Gabriele Fiorentini[b] and Enrique Sentana[c]

[a]*Federal Reserve Bank of New York, New York, United States*
[b]*Universitá di Firenze and RCEA, Florence, Italy*
[c]*CEMFI and CEPR, Madrid, Spain*

ABSTRACT

The authors analyze a model for N different measurements of a persistent latent time series when measurement errors are mean-reverting, which implies a common trend among measurements. The authors study the consequences of overdifferencing, finding potentially large biases in maximum likelihood estimators (MLE) of the dynamics parameters and reductions in the precision of smoothed estimates of the latent variable, especially for multiperiod objects such as quinquennial growth rates. The authors also develop an R^2 measure of common trend observability that determines the severity of misspecification. Finally, the authors apply their framework to US quarterly data on GDE and GDI, obtaining an improved aggregate output measure.

Keywords: Cointegration; GDE; GDI; overdifferencing; signal extraction; statistical discrepancy

JEL Classification: C32; E01

Essays in Honor of Joon Y. Park: Econometric Methodology in Empirical Applications
Advances in Econometrics, Volume 45B, 3–33

ISSN: 0731-9053/doi:10.1108/S0731-90532023000045B001

1. INTRODUCTION

Aggregate measurements, particularly those of output, are a key input to research economists and policy makers. Assessing the state of business cycles, making predictions of future economic activity, and detecting long-run trends in national income are some of their most popular uses. These measurements are typically regarded as noisy estimates of the quantities of interest, but accounting for the role of measurement error in applications is a difficult task. An important exception arises when more than one measurement of the same quantity is available. This makes it possible to combine the different measurements to produce a better estimate, ideally assigning higher weights to more precise ones.

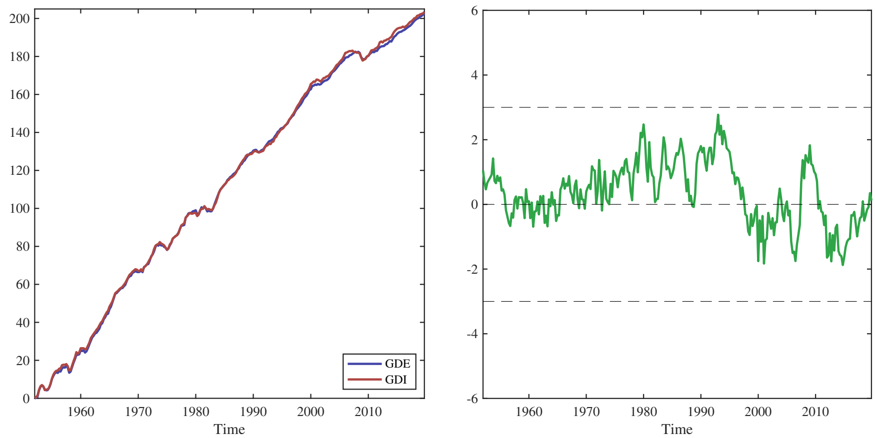
In the United States, the Bureau of Economic Analysis (BEA) reports both the expenditure-based Gross Domestic Expenditure (GDE) measure of output and its income-based Gross Domestic Income (GDI) counterpart. If the sources and methods of the statistical office were perfect, then the two would be identical. In practice, however, they differ (see Landefeld et al., 2008, for a review). The frequent, and at times noticeable, discrepancy between them (officially known as *statistical discrepancy*) has been recently the subject of active debate in academic and policy circles,[1] and various proposals for improved measures of economic activity have been discussed (see, e.g., Aruoba et al., 2016; Greenaway-McGrevy, 2011; Nalewaik, 2010, 2011).[2] The *GDPplus* measure of Aruoba et al. (2016), for example, is currently released on a monthly schedule by the Federal Reserve Bank of Philadelphia.

In this chapter, we propose improved output measures under the assumption that alternative measurements in levels do not systematically diverge from each other over the long run. While economic activity, like several other macro aggregates, arguably displays a strong stochastic trend, one would expect statistical discrepancies to mean-revert. In that case, measurements in levels would share a common trend. Somewhat surprisingly, though, the standard practice is to rely on models that do not impose this common trend, working instead with the growth rates of measurements. To cite a few references, Smith et al. (1998), Nalewaik (2010, 2011), Greenaway-McGrevy (2011), and Aruoba et al. (2016) all apply signal extraction techniques to a model of the first differences of log GDP. Similarly, the literature on GDP data revisions also works with growth rates, e.g., Aruoba (2008) and Jacobs and van Norden (2011) and Jacobs et al (2020).

In this respect, our main goal is to explore the implications of neglecting a common trend in levels for both parameter estimators and smoothed estimates of latent variables. Specifically, we follow Smith et al. (1998) in analyzing a model in which N different measurements y_t of an unobserved quantity x_t are available, so that

$$y_t = x_t 1_{N\times 1} + v_t,$$

with v_t denoting measurement errors in levels and $1_{N\times 1}$ a vector of N ones. In contrast to the literature, though, we model x_t as $I(1)$ – i.e., Δx_t is stationary and strictly invertible – but v_t as $I(0)$. The discrepancies between measurements $y_{it} - y_{jt} = v_{it} - v_{jt}$ are thus cointegrating relationships, reflecting that mean



(a) Output measurements in levels

(b) Statistical discrepancy

Fig. 1. GDE and GDI. *Notes*: We use November 2020's release of BEA national accounts estimates spanning 1952Q1–2019Q4; (a) 100×logs of GDE and GDI subtracting their 1952Q1 values, i.e., percentage (log) growth in GDE and GDI accumulated since 1952Q1; (b) Differences between 100×log of GDE and 100×log of GDI.

reversion keeps alternative measurements from diverging. As a result, the measurement errors in first differences, Δv_t, are overdifferenced.

Fig. 1 shows US data counterparts to y_t and $y_{it} - y_{jt}$:

The parameters that describe the dynamics of x_t are typically of interest in themselves, as they inform important dimensions of business cycles and enter signal-extraction calculations. For that reason, we begin by studying the effects of ignoring cointegration among the elements of y_t on estimation procedures. We focus on Gaussian MLE in a simple parametric setup in which the model for x_t is correctly specified but that of v_t is not because of the neglected common trend. Our main finding is that even if x_t and v_t are stochastically independent, estimators of the autocorrelation parameters of x_t will be affected by misspecification in the dynamics of v_t, displaying potentially large biases and increased asymptotic variances. At the same time, we show that if the statistical model assumes Gaussian autoregressive dynamics for both Δx_t and Δv_t, then the estimators of their unconditional means and variances will be asymptotically unaffected. Consequently, the impact of misspecification will be confined to the autocorrelation structure of Δx_t.

Moreover, we prove that the extent to which inferences will be impaired is governed by (i) the severity of overdifferencing in measurement errors, and (ii) the overall signal-to-noise ratio. The more severely overdifferenced the elements of Δv_t are (i.e., the further away from unit root processes those measurement errors are), the stronger the dynamic misspecification resulting from the omitted common trend will be. In addition, a low degree of signal observability, which we quantify by means of an R^2 measure of the relative contribution of x_t and v_t to the variation in observables, amplifies the role of incorrect modeling assumptions on v_t. In the limiting case of $R^2 = 1$, x_t is observable and misspecification in v_t inconsequential.[3] Our results therefore complement those in Chang et al. (2009), who derive the asymptotic distribution of the Gaussian MLE in a dynamic factor model with a single common trend. While Chang et al. (2009) study the case of unknown loadings under correct specification, we focus on the case of known loadings (equal to $1_{N\times 1}$) but subject to the dynamic misspecification induced by overdifferencing.[4]

Prediction, filtering and smoothing of x_t given data on y_t – signal extraction, for short – constitute the other main focus of our chapter. Given that the uncertainty of signal extraction calculations does not vanish in large T samples, unlike that of parameter estimators, we study their behavior at the pseudo-true parameter values, i.e., at the probability limits of ML estimators. Thus, we leverage on our estimation results to establish the suboptimality as a signal extraction technique of the Kalman-filter-based methods that neglect the common trend.

We find that the effect of ignoring the common trend is substantially different when signal extraction targets a short-run object and a long-run one. In particular, confidence sets for a long-run object such as an average of Δx_t over a relatively large time span are highly sensitive to even modest amounts of overdifferencing in Δv_t. This result is important because long-run objects are relevant to empirical questions about slowly evolving trends in macro variables. One example originates in the recurrent debate about growth deceleration in industrialized economies (e.g., Gordon, 2016). Another instance is the secular stagnation hypothesis, which implies a downward trend in interest rates (e.g., Hansen, 1939; Summers, 2015). Similarly, the apparent secular decline in labor shares (e.g., Blanchard, 1997; Kaldor, 1957; Karabarbounis & Neiman, 2014) provides another case in point.

On the empirical side, we fit our proposed common trend model to US data on GDE and GDI. Through standard Kalman smoothing calculations, we obtain an improved measure of economic activity, which we compare to other existing measures in the literature. We then use our improved measure to assess the robustness of a variety of empirical facts on economic activity, involving both short- and long-run objects. Our main findings are the following: (1) point estimates of the serial correlation structure of economic activity appear robust to common trend assumptions, (2) the same seems to be true of point estimates of the quarterly average rate of growth in GDP, but (3) our common trend model gives rise to lower signal extraction uncertainty about economic activity than its competitors. Our third finding is conceptually important because point estimates of latent variables cannot be justified by an appeal to consistency – uncertainty about latent variables remains high regardless of the sample size, implying that such estimates must be accompanied by a measure of their precision. This is particularly important from an empirical point of view because the "putative" precision of estimates of economic activity which do not impose a common trend is so low that no sharp conclusion can be drawn about trends in growth from them. In contrast, our common-trend model provides noticeably more precise inference about such long-run objects.[5]

Of course, whether or not there is a common trend is an empirical question in its own right. The evidence that the statistical discrepancy between US GDE and GDI, although persistent, is mean-reverting is suggestive but not conclusive.[6] Yet, the fact that, absent a common trend, the probability of observing large deviations between different measurements tends to one, lends strong support to our framework in the context of aggregate measurement problems.

The rest of the chapter is organized as follows. In Section 2, we present the basic setup. Section 3 discusses the properties of MLE while Section 4 is devoted to filtering. We report the results of our empirical analysis in Section 5. Finally, Section 6 concludes. Additional results are relegated to the Appendix and the supplemental material.

Notation. We use $\omega_{t_0:t_1}$ to denote the sequence $\{\omega_t\}_{t=t_0}^{t_1}$. If ω_t is a $d_1 \times d_2$ array for all t, and if it raises no confusion, we also use $\omega_{t_0:t_1}$ to denote the $d_1(t_1 - t_0 + 1)\times d_2$ array obtained by vertical concatenation of the terms of $\{\omega_t\}_{t=t_0}^{t_1}$. Analogously, $\psi_{1:N}$ denotes the column vector $(\psi_1,\ldots,\psi_N)'$. We write $\mathbb{E}_T[\omega_t] = T^{-1}\sum_{t=1}^{T}\omega_t$ for the sample average of $\omega_{1:T}$, $\mathbb{E}[\omega_t]$ for its population counterpart, "$\xrightarrow{p}$" for convergence in probability and "$\Rightarrow$" for weak convergence.

2. MODEL

In our setup, the statistical office collects N (log) measurements y_t of an unobserved scalar (log) quantity x_t. Let v_t be the vector of (multiplicative) measurement errors so that, in first differences,

$$\Delta y_t = \Delta x_t 1_{N\times 1} + \Delta v_t,\ t = 1,\ldots,T. \tag{1}$$

For a sample $\Delta y_{1:T}$, the data generating process is given by the probability distribution $\mathbb{P}$.

Assumption 1. $\mathbb{P}$ *satisfies the following:*

(A) The time series $\Delta x_{0:T}, v_{1,0:T},\ldots,v_{N,0:T}$ *are cross-sectionally independent;*
(B) Δx_t *is a Gaussian AR(1) process: For some values* $\mu_0, \rho_0 \in (-1,1), \sigma_0 > 0$,

$$\Delta x_0 \sim N(\mu_0, \sigma_0^2),$$
$$\Delta x_t \mid \Delta x_{0:(t-1)} \sim N\left(\mu_0 + \rho_0(\Delta x_{t-1} - \mu_0), (1-\rho_0^2)\sigma_0^2\right),\ t = 1,\ldots,T;$$

(C) v_{it} *is a Gaussian AR(1) process: For some values* $\rho_i \in (-1,1], \sigma_i > 0$,

$$v_{i0} \sim N\left(0, \frac{(1+\rho_i)}{2}\sigma_i^2\right),$$
$$v_{it} \mid v_{i,0:(t-1)} \sim N\left(\rho_i v_{i,t-1}, \frac{(1+\rho_i)}{2}\sigma_i^2\right),\ t = 1,\ldots,T,\ i = 1,\ldots,N.$$

Assumptions 1(A) and 1(B) are made in essentially every paper in the literature (e.g., Almuzara et al., 2019; Aruoba et al., 2016; Greenaway-McGrevy, 2011; Smith et al., 1998). Independence between Δx_t and measurement errors rules out cyclical patterns in the statistical discrepancy. Although potentially of substantive interest, introducing dependence between Δx_t and v_t or across the v_{it}'s complicates identification of the spectra of latent variables. Similarly, AR(1) dynamics for Δx_t is generally agreed to be a reasonable benchmark for economic activity data. Normality is unnecessary for most of our analysis, but since our focus is on the modeling of measurement errors and the role of dynamic misspecification, we adopt it for ease of exposition.

According to Assumption 1(B), we can regard $\Delta x_{0:T}$ as a segment from a strictly stationary process $\Delta x_{-\infty:\infty}$,

$$\Delta x_t = (1-\rho_0)\mu_0 + \rho_0 \Delta x_{t-1} + \sqrt{1-\rho_0^2}\,\sigma_0 \varepsilon_{0t},$$

with $\varepsilon_{0t} \overset{iid}{\sim} N(0,1)$. Our parameterization of the process for the signal ensures that $\mathbb{E}[\Delta x_t] = \mu_0$ and $\mathrm{Var}(\Delta x_t) = \sigma_0^2$, which do not depend on ρ_0, thereby separating these unconditional moments from the parameters governing the dynamics of Δx_t. Thus, we can summarize the serial dependence structure of the growth rate by its spectral density

$$f_0(\lambda) = \sigma_0^2 \frac{(1-\rho_0^2)}{(1-\rho_0 e^{i\lambda})(1-\rho_0 e^{-i\lambda})} = \sigma_0^2 \left(\sum_{\ell=-\infty}^{\infty} \rho_0^{|\ell|} e^{i\ell\lambda} \right).$$

Assumption 1(C) implies Δv_{it} is overdifferenced, the severity of overdifferencing increasing as ρ_i moves away from unity. In fact, Δv_{it} is a strictly noninvertible ARMA(1,1) process, except in the limiting case $\rho_i = 1$, when Δv_{it} becomes white noise. We can view $\Delta v_{i,0:T}$ as a segment from a strictly stationary process $\Delta v_{i,-\infty:\infty}$,

$$\Delta v_{it} = \rho_i \Delta v_{i,t-1} + \sqrt{\frac{(1+\rho_i)}{2}}\sigma_i \Delta\varepsilon_{it},$$

with $\varepsilon_{it} \overset{iid}{\sim} N(0,1)$. We have $\mathbb{E}[\Delta v_{it}] = 0$ and $\mathrm{Var}(\Delta v_{it}) = \sigma_i^2$, and the spectral density of Δv_{it} is

$$f_i(\lambda) = \sigma_i^2 \frac{(1+\rho_i)(1-e^{i\lambda})(1-e^{-i\lambda})}{2(1-\rho_i e^{i\lambda})(1-\rho_i e^{-i\lambda})},$$

which vanishes at frequency $\lambda = 0$ if $\rho_i \neq 1$ – an unequivocal symptom of overdifferencing.

When $\rho_i \neq 1$ for all i, the spectral density matrix of Δy_t at $\lambda = 0$ is $f_0(0)1_{N\times N}$. Therefore, it is singular with finite positive diagonal, implying the cointegration (of rank $N-1$) of y_t. Thus, y_t is driven by a single common trend, x_t, while the statistical discrepancies $d_{ij,t} = y_{it} - y_{jt}$ are cointegrating relationships.[7]