European Economic Review 100 (2017) 293–317

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/euroecorev

A tale of fat tails[☆]

Chetan Dave, Samreen Malik[*]

New York University (AD), PO Box 903, New York, NY 10276-0903, United States

ARTICLE INFO

Article history:
Received 26 September 2016
Accepted 28 August 2017
Available online 8 September 2017

Keywords:
Adaptive learning
DSGE models
Fat tails
Power law

ABSTRACT

We document the extent to which major macroeconomic series, used to inform linear DSGE models, can be characterized by power laws whose indices we estimate via maximum likelihood. Assuming data follow a linear recursion with multiplicative noise, low estimated indices suggest fat tails. We then ask whether standard DSGE models under constant gain learning can replicate those fat tails by an appropriate increase in the estimated gain and without much change in the transmission mechanism of shocks. We find that is largely the case via implementation of a minimum distance estimation method that eschews any allegiance to distributional assumptions.

1. Introduction

The ability of linear dynamic stochastic general equilibrium (DSGE) models to adequately account for macroeconomic fluctuations has come under scrutiny in light of the Great Recession. Such large but rare events manifest themselves in the form of fat tails for data that are usually employed in standard Gaussian empirical DSGE modeling under rational expectations (RE). However, in the absence of additional assumptions on the stochastic nature of innovations, standard DSGE models are unable to replicate observed large fluctuations.[1] We show that a DSGE model under adaptive learning (AL) endogenously delivers model dynamics that better replicate observed fat tails.

Existing DSGE analyses have explored at least three avenues to model large macroeconomic fluctuations. The first avenue replaces the assumption of Normally distributed innovations with a fat-tailed specification (e.g. a Student's-*t* or Laplace). Fat tails in the distribution of innovations allow for a higher probability that a large shock occurs and works its way through standard transmission mechanisms. The second avenue replaces the assumption of a constant variance for structural innovations with exogenous stochastic volatility specifications. The idea is that if one introduces exogenous volatility into a DSGE shock specification then macroeconomic variables will also exhibit the sort of volatility associated with rare but large fluctuations. The third avenue introduces time variation in the structural parameters of a model which in turn generates time or state dependent responses of economies to an otherwise constant variance shock process. All three avenues modify models so that exogenous sources of volatility are introduced in order to match observed volatility. Our analysis is closest

☆ We especially thank editor Eric M. Leeper, associate editor and anonymous referees for many comments that helped us significantly improve the paper. We extend special thanks to Jess Benhabib for his guidance, comments and suggestions throughout the process. We also thank Christian Haefke, Jean Imbs, John Leahy, Max Mihm, Romain Ranciere, Michael Reiter, Gilles Saint-Paul and seminar participants at the Midwest Macroeconomics meetings, EEA-ESEM congress, University of Winnipeg, Indian Statistical Institute and NYU Shanghai for helpful comments and suggestions. The usual disclaimer applies.

* Corresponding author.
E-mail addresses: chetdav@gmail.com (C. Dave), samreen.malik@nyu.edu (S. Malik).

[1] We use the term large fluctuations to mean that the frequency of realizations that deviate substantially from the trend are higher than the frequencies of the same realizations under a Normal distribution.

http://dx.doi.org/10.1016/j.euroecorev.2017.08.010
0014-2921/

to the third avenue and presents an endogenous channel via stochastic gradient constant gain (SGCG) AL that delivers fat tails for endogenous variables in an otherwise standard model.

AL is increasingly used by macroeconomists to bridge data-model gaps. Seminal work on statistical learning (Sargent, 1993; Evans and Honkapohja, 2001), additional insights from a similar literature (Deak et al., 2015; Gaspar et al., 2006; Massaro, 2013; Milani, 2007; Orphanides and Williams, 2005; De Grauwe, 2012), and experimental evidence on the importance of the learning process in accounting for business cycle fluctuations and volatility (Adam and Woodford, 2012; Duffy, 2012; Heemeijer et al., 2012; Jaimovich and Rebelo, 2007) provide strong support that AL is a reasonable alternative to the standard RE DSGE setting. In contrast to the standard RE DSGE model, in which agents know the true stochastic process of an economy, under AL agents revise their forecast rules in response to incoming data so as to ascertain that stochastic process over time.[2] This difference in how expectations are formed influences model dynamics. In particular, under RE, model dynamics are described by a fixed coefficient vector autoregression (VAR). Under SGCG learning however, model dynamics are described by a linear recursion with multiplicative and additive noise (LRMN) written as

$$X_t = \Phi_t X_{t-1} + \varepsilon_t, \tag{1}$$

and with a stationary distribution for X_t different than that of a VAR due to the interplay between the *stochastic* multiplicative term (Φ_t) and the stochastic additive term (ε_t).[3] As a function of the nature of the interplay, the applied mathematics literature shows that the tail of the stationary distribution of X_t can be fatter than that of a Normal distribution (e.g. Kesten, 1973). This implies that X_t can take on extreme values with a higher probability than under a Normal distribution and thus this equation forms an alternative lens with which to view data and an associated model economy. In a sense the intuition of a LRMN system is as follows: as grains of sand pile up into a dune, at some point a grain falls that shifts the dune dramatically, and this dramatic movement occurs with some regularity.[4]

Since under SGCG learning a DSGE model is written as a LRMN, one needs to change the underlying assumption on the data generating process (DGP) from a fixed coefficient VAR to a LRMN. Under this new assumption, the tail of the stationary distribution of data (Y_t) can be fat. We measure the thickness of the tail by estimating the tail index p under the assumption that $Y_t \sim Y^{-p}$ (a power law), since under SGCG learning model variables are distributed similarly. We also conduct a formal test to establish whether the hypothesized power-law is a plausible fit to the data and our test statistics show that data that enter a DSGE empirical exercise are usually not Normal.

We find that data exhibit characteristics consistent with a LRMN assumption on the DGP. In particular, the data employed in DSGE models have fatter tails than would be warranted under a Normality assumption. We implement a minimum distance estimation exercise that allows us to jointly estimate model parameters including the constant gain (g). Were our estimates of g small or near zero then model dynamics would approximate those of a RE DSGE model. However, we find that estimates of g are non-zero and in fact higher than in the current literature.[5] A finding of a higher constant gain does not violate any theoretical or empirical requirement that g be near zero (so an AL model is in a small vicinity of its RE solution). Further we show that a model under RE with Normally distributed innovations, or a model with fat-tailed distributions for innovations, is not able to come as close as a DSGE model under SGCG AL in terms of being able to replicate fat tails observed in data.

These empirical and simulation results allow us to establish a central intuition, and therefore our key contribution: given that a larger g reflects a shorter memory (learning horizon), as g rises, macroeconomic variables are more likely to visit extreme values (deep recessions and booms), simply because agents do not remember as much of history as they could and therefore are bound to repeat it. Overall, we are able to show that without departing fundamentally from the standard DSGE model, SGCG AL is enough to account for observed fat tails without altering the model's fit to other distributional dimensions of the data.

Having established the model and described related literature in Sections 2 and 3, we set out the stylized facts for all major macroeconomic time series in Section 4. In Section 5 we reconcile data with model with a minimum distance exercise. Section 6 follows with simulation exercises and we conclude in Section 7 with a re-statement of our central contribution: that given the recurrent manner in which macroeconomic time series exhibit large fluctuations, LRMN model representations may be more suitable to empirical analyses of linear DSGE models.

2. Rational expectations vs. adaptive learning

Linear DSGE models begin by specifying a familiar form

$$X_t = A(\theta)E_t(X_{t+1}) + B(\theta)U_t, \tag{2}$$

$$U_t = P(\theta)U_{t-1} + \varepsilon_t, \; E(\varepsilon_t\varepsilon_t') = \Sigma(\theta), \tag{3}$$

[2] In the limit as the constant gain (g) of a SGCG learning process tends to zero, a model economy approaches the RE solution (see Evans and Honkapohja, 2001).

[3] Our definitions for the process $X_t = \Phi_t X_{t-1} + \varepsilon_t$ are as follows. If Φ_t is a constant matrix Φ then we call it a fixed coefficient VAR. If Φ_t varies over time in a deterministic manner we call it a variable coefficient VAR. If Φ_t is itself a stochastic process then we refer to the equation as a LRMN.

[4] For similar intuition relating to the notion of self-organized criticality that a LRMN represents, see Blume et al. (2010).

[5] Malmendier and Nagel (2016) find evidence in favor of a constant gain (approximately 0.02) using survey data on expectations across generations.