NATIONAL BANK OF POLAND
WORKING PAPER
No. 101

Skew-normal shocks in the linear state space form DSGE model

Grzegorz Grabek, Bohdan Kłos, Grzegorz Koloch

Warsaw 2011

Grzegorz Grabek, Bohdan Kłos, Grzegorz Koloch – Skew-Normal Shocks in the Linear State Space Form DSGE Model, Bureau of Applied Research, Ecomomic Institute, National Bank of Poland, Warsaw, September, 2011

Address:

Bohdan.Klos@mail.nbp.pl	(0 48 22) 653 15 87
Grzegorz.Grabek@mail.nbp.pl	(0 48 22) 585 41 08
Grzegorz.Koloch@mail.nbp.pl	(0 48 22) 653 21 79

Design:

Oliwka s.c.

Layout and print:

NBP Printshop

Published by:

National Bank of Poland
Education and Publishing Department
00-919 Warszawa, 11/21 Świętokrzyska Street
phone: +48 22 653 23 35, fax +48 22 653 13 21

http://www.nbp.pl

Contents

Abstract

Observed macroeconomic data – notably GDP growth rate, inflation and interest rates – can be, and usually are skewed. Economists attempt to fit models to data by matching first and second moments or co-moments, but skewness is usually neglected. It is so probably because skewness cannot appear in linear (or linearized) models with Gaussian shocks, and shocks are usually assumed to be Gaussian. Skewness requires non-linearities or non-Gaussian shocks. In this paper we introduce skewness into the DSGE framework assuming skewed normal distribution for shocks while keeping the model linear (or linearized). We argue that such a skewness can be perceived as structural, since it concerns the nature of structural shocks. Importantly, the skewed normal distribution nests the normal one, so that skewness is not assumed, but only allowed for. We derive elementary facts about skewness propagation in the state space model and, using the well-known Lubik-Schorfheide model, we run simulations to investigate how skewness propagates from shocks to observables in a standard DSGE model. We also assess properties of an *ad hoc* two-steps estimator of models' parameters, shocks' skewness parameters among them.

JEL: C12, C13, C16, D58, E32

Introduction

Skewness is a statistical feature of observed economic data. For an arbitrary random variable, like output growth rate, inflation rate or an interest rate, skewness is typically manifested by the lack of symmetry of the probability density function which governs this variable[1]. Intuitively, if a random variable follows a skewed distribution, then its deviations above the mean value are (i.e. positive deviations from the mean) on average either larger or smaller in magnitude than the deviations below the mean (i.e. negative deviations from the mean). Also, either positive or negative deviations from the mean value tend to be more frequent (i.e. are more probable)[2]. Table 1 reports skewness coefficients[3] for four macro aggregates in wide range of countries calculated over 30 years using quarterly data. Application of the Bai and Ng (2005) test π_3 for skewness indicates that most of these coefficients are statistically significant. With the except of Canada, quarterly inflation rate is positively skewed, which means that, on one hand, positive deviations of inflation rate from the mean value tend to be bigger in magnitude than the negative ones, and, on the other, that we should expect more episodes of inflation rate below the mean than episodes of inflation rate above the mean. These two features of inflation imply that inflation risks are asymmetric, especially if the mean value turns out to be in line with the central bank inflation target. It is thus of no surprise, that nominal interest rates also reveal positive skewness pattern. This is a partial argument for the fact that interest rate inherit skewness pattern from inflation and not the other way round. Real output growth rate, in turn, tends to have more frequently values above the mean than below the mean, but negative deviations are on average greater in magnitude. In other words, GDP tends to grow at a moderate pace, but if a recession hits, it can be severe. Absolute changes[4] of domestic exchange rates versus the US dollar — with the exception of the Swiss franc and the Japanese yen, are positively skewed in the sample, which means that appreciation of the currencies is more frequent than depreciation and that appreciation tends to be moderate in magnitude, but there may be, although less frequently, episodes of substantial depreciations. This stands in line with the safe haven status of the US currency. Negative skewness of absolute changes of the Japanese yen and the Swiss franc vs. US dollar could be understood as them having a safe haven status with respect to the US dollar, which nowadays is true for the Swiss franc at least. Investigated skewness patterns hold not only for individual countries, but can also be seen in aggregate economies — in the European Union and in the OECD.

[1] Yet, it can be the case that probability distribution function of a non-skewed random variable is not symmetric.

[2] This gives rise to the positive and negative skewness respectively.

[3] We measure skewness using a skewness coefficient defined as the third central moment standardized by the second central moment to the power of 1.5, see eq. (1.6). For a review of other skewness measures see e.g. MacGillivray (1986).

[4] Positive changes denote depreciation and negative changes denote appreciation vs. the US currency.

Table 1. Skewness (measured by sample skewness coefficient) in macroeconomic data

Country	GDP growth	Inflation	Nom. interest rate	Exchange rate
Australia	-0.18	0.05	0.71$*$	1.33$*$
Canada	$-0.50*$	$-0.63*$	0.76$*$	0.74$*$
France	$-0.96*$	1.64$*$	0.49$*$	-
Japan	$-1.20*$	0.88$*$	0.47	$-0.46*$
Korea	$-1.48*$	0.49$*$	0.68$*$	2.92$*$
Switzerland	-0.12	0.97$*$	0.72$*$	-0.02
United Kingdom	$-1.34*$	1.04$*$	0.40$*$	0.85$*$
United States	$-0.97*$	1.82$*$	0.86$*$	-
European Union	$-1.96*$	1.57$*$	-	0.31
OECD	$-2.24*$	1.79$*$	-	-

Note: $*$ indicates significance at least at the 10% level.

Source: Own calculations based on OECD data

It is clear from this exemplary exposition that, at least for investigated samples, major macroeconomic time series reveal a meaningfully interpretable skewness pattern. This somehow stands in contrast with the fact that DSGE models, as far as their first order approximations are concerned[5], totally abstract from skewness of observed data, assuming that both structural innovations and measurement errors are normally distributed, hence symmetric.

Neglecting information provided by skewness of economic data distorts the balance of risks faced by the policy makers, which limits their ability to achieve assumed objectives. Additionally, unnoticed or neglected features of economic phenomena tend to limit the insight into them, especially if they can be perceived as structural ones.

In the DSGE domain, skewness in observed variables can appear as a result of three major factors. Firstly, skewness can appear as a result of models' non-linearities. A trivial example is when a normally distributed variable, e.g. a shock, is squared so that it obtains a χ^2_1 distribution. Shocks, for they influence states, would propagate skewness to observables. Such a mechanism would work if we allowed for higher order approximations of the economy. Secondly, skewness in observables can emerge as a result because of models' internal mechanisms, e.g. asymmetric preferences, see Christodoulakis and Peel (2009) or downward nominal or real rigidities, see (Fahr and Smets, 2008; Kim and Ruge-Murcia, 2009). In such a case skewness constitutes an endogenous feature of the model and there is a magnitude of degrees of freedom in which it can be introduced. However, skewness vanishes from states and observables if only the first order perturbation is used[6], which in practice is often the case, especially when the model is estimated, see (Amisano and Tristani, 2007). In a linearized state space form DSGE model skewness in observables can appear when shocks hitting the economy follow a skewed distribution[7]. In such a case skewness constitutes a structural feature of the modeled economy because it reflects a statistical feature of structural shocks hitting the economy.

[5]In this paper we focus only on the first order perturbation for reasons that are explained later in this section. Thus, excercises provided in this paper is more an econometric than an economic one.

[6]And shocks follow a symmetric distribution

[7]Outside the framework of DSGE models Ball and Mankiw (1995) showed how combination of non-linearity (firms adjust prices to shocks that are sufficiently large to justify paying menu costs) and skewed shocks (with zero mean) to desired price levels leads to skewed observed price changes.

In this paper we take the latter approach, i.e. we take a linear state space model — which is thought of as a first order perturbation of a DSGE economy — and assume that martingale difference shocks in the transition equation have a skewed distribution. Alternatively, we could assume that measurement errors are skewed. Both approaches result in skewed observables, but the latter one lacks any structural motivation, whereas the one we take seems to have a sound economic interpretation — shocks are skewed in a structural way. What is important for our motivation, is that in the class of linear state space models skewness in observed variables must be a reflection of skewness in stochastic disturbances, so the number of degrees of freedom through which skewness can be accounted for in the modeled economy is minimized to one — there is no other way for skewness to enter the model[8].

Working with skewed shocks gives rise to the question which family of probability distributions is appropriate for this purpose. Such a family, firstly, should nest a normal distribution, so that the typical (normal) specification is allowed for and skewness in shocks can be rejected if it does not find enough support in the data. Secondly, employed distribution should have properties which allow us to use the state space setting. Desired properties involve closure under most general linear transformations, under addition of independent variables, under taking joint and marginal distributions and under conditioning. Most of these features, but not all of them, are offered by the closed skewed normal distribution.

In the paper we do three things. First, we deliver elementary facts about propagation of skewness and of the closed skewed normal distribution in linear state space models. Second, we conduct simulation experiments designed to capture propagation of skewness from shocks to observed variables in a small open economy Lubik and Schorfheide (2007) DSGE model. Finally, we develop a simple, yet useful, two-step quasi-maximum likelihood estimation procedure, which is capable of handling skewness, but avoids computational difficulties which emerge in case of maximum likelihood estimation.

[8] We *a priori* reject the possibility that measurement errors are skewed since this would seem as an artificial, technical assumption.

Chapter 1

Skewness in linear models

This section presents the closed skewed normal distribution and provides elementary facts on propagation of skewness and of the closed skewed normal distribution in a linear state space model.

1.1 The closed skewed normal distribution

Let us denote a density function of a p-dimensional normal distribution with mean[1] $\tilde{\mu}$ and positive-definite covariance matrix $\tilde{\Sigma}$ by $\phi_p(z;\tilde{\mu},\tilde{\Sigma})$. Let us also denote a cumulative distribution function of a q-dimensional normal distribution with mean $\tilde{\mu}$ and nonnegative-definite covariance matrix $\tilde{\Sigma}$ by $\Phi_q(z;\tilde{\mu},\tilde{\Sigma})$. For $q>1$ function Φ_q does not have a closed form.

By $\mathrm{R}_{p\times q}$, $p,q\geq 1$, let us denote a space of linear operators from R^p to R^q. For every $M\in\mathrm{R}_{p\times q}$ let $|M|$ denote a determinant of M and let $r(M)$ denote a rank of M. We will define the closed skewed normal, possibly singular, distribution by means of the moment generating function (mgf). Then, under nonsingularity conditions, probability density function (pdf) will be provided.

Definition 1.1.1. (*csn distribution — mgf*) Let $\tilde{\mu}\in\mathrm{R}^p$ and $\vartheta\in\mathrm{R}^q$, $p,q\geq 1$. Let $\tilde{\Sigma}\in\mathrm{R}_{p\times p}$ and $\Delta\in\mathrm{R}_{q\times q}$, $|\tilde{\Sigma}|,|\Delta|\geq 0$, and let $D\in\mathrm{R}_{q\times p}$. We say that random variable z has a (p,q) dimensional closed skewed normal distribution with parameters $\tilde{\mu}$, $\tilde{\Sigma}$, D, ϑ and Δ if moment generating function of z, $M_z(t)$, is given by:

$$M_z(t)=\frac{\Phi_q(D\tilde{\Sigma}t;\vartheta,\Delta+D\Sigma D^T)}{\Phi_q(0;\vartheta,\Delta+D\Sigma D^T)}\mathrm{e}^{t^T\tilde{\mu}+\frac{1}{2}t^T\tilde{\Sigma}t}$$

which henceforth will be denoted by:

$$z\sim csn_{p,q}(\tilde{\mu},\tilde{\Sigma},D,\vartheta,\Delta)$$

Note that matrices $\tilde{\Sigma}$ and Δ are allowed to be singular. If $\tilde{\Sigma}$ is not positive definite, i.e. $|\tilde{\Sigma}|=0$, resulting distribution is called *singular*. If $\tilde{\Sigma}$ is positive definite, i.e. $|\tilde{\Sigma}|>0$, distribution is

[1] All vectors are column vectors throughout the paper.

called *nonsingular*. The *csn* distribution is "closed" in the sense, that it is closed under *full rank* linear transformations[2]. Isomorphic linear transformations transform nonsingular *csn* variables into nonsigular ones and singular variables into singular ones. Full row, but column rank deficient linear transformations (dimension shrinkage) transform nonsingular *csn* variables into nonsigular ones and singular variables into singular or nonsingular ones. Full column, but row rank deficient linear transformations (dimension expansion) transform nonsingular *csn* variables into singular ones, whereas singular variables remain singular. Both singular and nonsingular variables can be transformed into a non-*csn* distributed variable under a rank deficient transformation. The skewed normal distribution — when considered as consisting of both singular and nonsingular *csn* variables — is therefore not closed under arbitrary linear transformations, which entails computational difficulties for maximum likelihood estimation of state space models with *csn* shocks when the transition matrix in state space equations is singular, which typically is the case in DSGE modeling.

For $|\tilde{\Sigma}| > 0$, a *csn* random variable z has a probability density function:

Definition 1.1.2. (*csn distribution — pdf*) If a random variable z follows a (p,q)-dimensional, $p,q \geq 1$, closed skewed normal distribution with parameters $\tilde{\mu}$, $\tilde{\Sigma}$, D, ϑ and Δ, where $\tilde{\mu} \in \mathrm{R}^p$, $\vartheta \in \mathrm{R}^q$, $\tilde{\Sigma} \in \mathrm{R}_{p\times p}$, $|\tilde{\Sigma}| > 0$, $\Delta \in \mathrm{R}_{p\times p}$, $|\Delta| \geq 0$ and $D \in \mathrm{R}_{q\times p}$, than probability density function of z is given by:

$$p(z) = \phi_p(z;\tilde{\mu},\tilde{\Sigma})\frac{\Phi_q(D(z-\tilde{\mu});\vartheta,\Delta)}{\Phi_q(0;\vartheta,\Delta + D\Sigma D^T)} \tag{1.1}$$

Density function (1.1) defines a (p,q)-dimensional *nonsingular* closed skewed normal distribution in the sense that a random variable has (p,q)-dimensional nonsingular closed skewed normal distribution with parameters $\tilde{\mu}$, $\tilde{\Sigma}$, D, ϑ and Δ if and only if its density function for every $z \in \mathrm{R}^p$ equals $p(z)$. The probability density function (1.1) involves a probability distribution function of a q-dimensional normal distribution for, in principle, arbitrarily large q, which entails computational difficulties when working with a likelihood function based on $p(z)$. Closed skewed normal density function can be read as a a product of a normal density function (which is symmetric) and a skewing or weighting function given by a quotient of two normal probability distribution functions (in fact the distribution function in the denominator constitutes a constant of proportionality so that everything integrates to unity).

Parameters $\tilde{\mu}$, $\tilde{\Sigma}$ and D have interpretation of location, scale and skewness parameters respectively. Parameters ϑ and Δ are artificial, but inclusion of these additional dimensions allows for closure of the *csn* distribution under conditioning and marginalization respectively. The q-dimension is also artificial, but it allows for closure for sums and the joint distribution of independent (not necessarily *iid*) variables. When $\tilde{\Sigma}$, D and Δ are scalars, they will be denoted respectively by $\tilde{\sigma}$, d and δ.

[2]Under full rank transformation we mean full row or full column rank transformation and this definition embraces the case when matrix of the transformation is square and has a full rank. In the latter case the transformation is called an *isomorphism*. When both the row and the column ranks are not full, transformation is called rank deficient.

Let us note the following:

Remark 1.1.3. For $p = q = 1$, $\vartheta = 0$ and $\Delta = 1$ the *csn* distribution reduces to the Azzalini skewed normal distribution, see Azzalini and Valle (1996); Azzalini and Capitanio (1999).

Such a case will be denoted by:

$$z \sim sn(\tilde{\mu}, \tilde{\sigma}, d) \tag{1.2}$$

In the next sections we will find useful the following:

Corollary 1.1.4. *Let $z \sim csn_{1,1}(\tilde{\mu}, \tilde{\sigma}, d, \vartheta, \delta)$ for parameters as in definition (1.1.2), and assume that $\delta + d^2\tilde{\sigma} \neq 0$, then:*

$$\mathrm{E}(z) = \tilde{\mu} + \sqrt{\frac{2}{\pi}} \frac{d\tilde{\sigma}}{\sqrt{\delta + d^2\tilde{\sigma}}} \tag{1.3}$$

$$var(z) = \tilde{\sigma} - \frac{2}{\pi} \frac{d^2\tilde{\sigma}^2}{\delta + d^2\tilde{\sigma}} \tag{1.4}$$

$$\mathrm{E}(z - E(z))^3 = \left(2 - \frac{\pi}{2}\right)\left(\sqrt{\frac{2}{\pi}}\right)^3 \left(\frac{d\tilde{\sigma}}{(\delta + \tilde{\sigma}d^2)^{\frac{1}{2}}}\right)^3 \tag{1.5}$$

It follows that:

Remark 1.1.5. Let $z \sim csn_{1,1}(\tilde{\mu}, \tilde{\sigma}, d, \vartheta, \delta)$ for parameters as in definition (1.1.4), then $\mathrm{E}(z) = 0$ if and only if $\tilde{\mu} = -\sqrt{\frac{2}{\pi}} \frac{d\tilde{\sigma}}{\sqrt{\delta + d^2\tilde{\sigma}}}$.

We also need the following:

Corollary 1.1.6. *Let $z \sim csn_{p,q}(\tilde{\mu}, \tilde{\Sigma}, D, \vartheta, \Delta)$, $p, q \geq 1$, for parameters as in definition (1.1.1). Elements of z are independent if and only if matrices $\tilde{\Sigma}$ and D are diagonal.*

Since $\tilde{\Sigma}$ is $p \times p$ and D is $q \times p$, corollary (1.1.6) implies that it is impossible to have $q = 1$ while keeping elements of z independent for $p > 1$, because it has to be the case that $q = p > 1$ in order for D to be diagonal. This is relevant for state space models with *csn* distributed *iid* disturbances — e.g. shocks in the transition equation, because the state variable, say ξ_t, in every period consists of the *csn* distributed state from the previous period, say ξ_{t-1}, plus the *csn*-distributed disturbance[3], say u_t, and when we add two *csn* variables we have to add their q-dimensions, so that the q-dimension of ξ_t is the sum of q-dimensions of ξ_{t-1} and u_t, hence, according to corollary (1.1.6), contribution of u_t to q-dimension of ξ_t in every period cannot be squeezed to eg. 1, but must equal the number of elements of u_t, hence the q-dimension of ξ_t quickly expands which poses numerical difficulties for maximum likelihood estimation.

In further sections we will also need the following corollaries (1.1.7–1.1.10):

Corollary 1.1.7. *Let $z \sim csn_{p,q}(\tilde{\mu}, \tilde{\Sigma}, D, \vartheta, \Delta)$, $p, q \geq 1$, for parameters as in definition (1.1.1). Let also $x \sim N(\mu_x, \Sigma_x)$, $\Sigma_x > 0$, be independent of z, then:*

$$z + x \sim csn_{p,q}(\tilde{\mu} + \mu_x, \tilde{\Sigma} + \Sigma_x, D\tilde{\Sigma}(\tilde{\Sigma} + \Sigma_x)^{-1}, \vartheta, \Delta + (D(I - \tilde{\Sigma}(\tilde{\Sigma} + \Sigma_x)^{-1}))\tilde{\Sigma}D^T)$$

[3] Both state from the previous period and the disturbance are transformed by the linear transformation in state space models, but let us ignore this fact for the present argumentation (or assume this transformations are identities).

Corollary 1.1.8. *Let $z \sim csn_{1,q}(\tilde{\mu}, \tilde{\sigma}, d, \vartheta, \delta)$, $q \geq 1$ and for parameters as in definition (1.1.1), let also $\rho \neq 0$ and $b \in \mathrm{R}$, then:*

$$\rho z + b \sim csn_{1,q}(\rho\tilde{\mu} + b, \rho^2\tilde{\sigma}, \frac{1}{\rho}d, \vartheta, \delta)$$

Corollary 1.1.9. *Let $z \sim csn_{p,q}(\tilde{\mu}, \tilde{\Sigma}, D, \vartheta, \Delta)$, $p, q \geq 1$, for parameters as in definition (1.1.1), let also $A \in \mathrm{R}_{p\times p}$, $|A| > 1$, and $b \in \mathrm{R}^p$, then:*

$$Az + b \sim csn_{p,q}(A\tilde{\mu} + b, A\tilde{\Sigma}A^T, D\tilde{\Sigma}A^{-1}, \vartheta, \Delta)$$

Corollary 1.1.10. *Let $z_i \sim csn_{p,q_i}(\tilde{\mu}_i, \tilde{\Sigma}_i, D_i, \vartheta_i, \Delta_i)$, $p, q_i \geq 1$, $i = 1, 2, ..., n$, for parameters as in definition (1.1.1), then $\sum_{i=1}^n z_i \sim csn_{p,\sum_{i=1}^n q_i}(\tilde{\mu}^\star, \tilde{\Sigma}^\star, D^\star, \vartheta^\star, \Delta^\star)$, where:*

$$\tilde{\mu}^\star = \sum_{i=1}^n \tilde{\mu}_i, \quad \Sigma^\star = \sum_{i=1}^n \tilde{\Sigma}_i, \quad D^\star = (\Sigma_1 D_1^T, ..., \Sigma_n D_n^T)^T (\Sigma^\star)^{-1},$$

$$\vartheta^\star = (\vartheta_1^T, \vartheta_2^T, ..., \vartheta_n^T)^T, \quad \Delta^\star = \Delta^\oplus + D^\oplus \tilde{\Sigma}^\oplus D^\oplus - [\bigoplus_{i=1}^n D_i \tilde{\Sigma}_i](\tilde{\Sigma}^\star)^{-1}[\bigoplus_{i=1}^n D_i \tilde{\Sigma}_i]^{-1}$$

for $\Delta^\oplus = \bigoplus_{i=1}^n \Delta_i$, $D^\oplus = \bigoplus_{i=1}^n D_i$ and $\Sigma^\oplus = \bigoplus \tilde{\Sigma}_i$, where operator $\oplus$, arbitrary matrices A and B, is defined as:

$$A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

We are now ready to investigate how skewness propagates from disturbances to states and observables in a linear state space setting. We will do this in a twofold manner. First we will show how skewness propagates through the state space form in general and then we will turn to the special case when shocks in the transition equation follow a closed skewed normal distribution.

1.2 Propagation of skewness

In this section we put forward elementary facts about skewness propagation in linear state space models. First we deal with state variables, and then with the observables. As a measure of skewness we employ the skewness coefficient[4], which, for an arbitrary random variable $z \in \mathrm{R}$, is defined as[5]:

[4]We choose a skewness coefficient for a handful of reasons. First, it is widely applied by many researchers, hence any results can be easily used by others. In addition, it satisfies properties stated by Arnold and Groneneveld (1995). Second, other skewness measures based on mode or quartiles may be not suitable for our purposes since the closed form formulas for mode and quartiles of skew-normal distribution have not been derived and we would find it difficult to provide results on propagation of skewness in general and in the case of the closed skewed distribution in particular. Third, as it will be clear from further analysis, for shocks, which we model as independent Azzalini-type variables, skewness coefficient represents an exhaustive measure of skewness. In case of observables, which are not independent, the univariate skewness coefficient is not exhaustive because it omits cross-skewness. However, from the economic point of view, we are only interested in skewness of observables perceived as single variables. The co-skewness between variables might enable us to gain a deeper insight into the issue and we think of it as of a promising direction of future research.

[5]We assumed that $z \in \mathrm{R}$ in (1.6), but in principle it can be the case that $z \in \mathrm{R}^p$ for $p > 1$ if only exponentiations and division in (1.6) are considered as elementwise operations.

1

$$\gamma(z) = \frac{\mathrm{E}(z - \mathrm{E}(z))^3}{(\mathrm{E}(z - \mathrm{E}(z))^2)^{\frac{3}{2}}} \tag{1.6}$$

provided that the second and the third central moment of z exist[6]. We will make use of the following:

Remark 1.2.1. For a random variable z with an n-times differentiable moment generating function $M_z(x)$ we have:

$$\mathrm{E}(z - \mathrm{E}(z))^n = \kappa_n(z) = \frac{\partial^n \ln M_z(x)}{\partial x^n}\,|_{x=0}$$

where $\kappa_n(z)$ denotes the n-th cumulant of z.

Remark 1.2.2. Let z be a random variable for which $\kappa_n(z)$ exists for $n = 2, 3$, then:

$$\gamma(z) = \frac{\kappa_3(z)}{(\kappa_2(z))^{\frac{3}{2}}}$$

Remark 1.2.3. Let z_i be independent random variables and let $\alpha_i \in \mathrm{R}$, $i = 1, 2, ..., m$, then:

$$\kappa_n\left(\sum_{i=1}^{m} \alpha_i z_i\right) = \sum_{i=1}^{m} \alpha_i^n \kappa_n(z_i)$$

State variables

We will start with a one-dimensional model and then move to the multidimensional case. Let us consider the following autoregressive model, which represents the state-space formulation without the measurement equation:

$$\xi_t = \rho\xi_{t-1} + u_t \tag{1.7}$$

$$u_t \sim p(\ldots) \tag{1.8}$$

$$\xi_0 \sim N(\mu_{\xi_0}, \sigma_0) \tag{1.9}$$

for $t = 1, 2, ..., T$, where $\xi_t, \xi_0, u_t \in \mathrm{R}$, $\rho \neq 0$, $\mu_{\xi_0} \in \mathrm{R}$, $\sigma_0 \geq 0$ and $p(\ldots)$ is any distribution[7] such that $\gamma(u_t)$ exists for every t and is constant, i.e. $\gamma(u_t) = \gamma(u)$ for every t.

Let us make use of the fact that ξ_t can be expressed as a weighted sum of innovations u_{t-k} for $k = 0, 1, ..., t-2$, and of ξ_1, i.e. we employ the moving average representation of ξ_t:

$$\xi_t = \rho^{t-1}\xi_1 + \sum_{k=0}^{t-2} \rho^k u_{t-k} = \sum_{k=0}^{t-1} \rho^k u_{t-k} \tag{1.10}$$

where the second equality comes from a simplifying assumption that $\mu_{\xi_0} = 0$ and $\sigma_0 = 0$, so that $\xi_1 = u_1$. This makes exposition simpler and does not change meaning of the results[8]. First

[6] Which is true in all cases considered in this paper.
[7] We write $p(\ldots)$ to denote that p can depend on some parameters.
[8] Because ξ_1 inherits skewness from u_1 and not from x_0, which is normally distributed.

we will investigate the effect on ξ_t exerted be innovation u_t at time $t = 1$ keeping $u_t = 0$ for $t > 1$, then we will see what happens if u_t is allowed to be nonzero also for $t > 1$[9].

When $u_t = 0$ for $t > 1$, (1.10) states that $\xi_t = \rho^{t-1}u_1$ and, employing remarks (1.2.1–1.2.3), we see that:

$$\gamma(\xi_t) = \frac{\rho^{3(t-1)}}{\left(\rho^{2(t-1)}\right)^{\frac{3}{2}}}\gamma(u) = sgn(\rho)^{t+1}\gamma(u) = \begin{cases} \gamma(u) & \rho > 0, \\ (-1)^{t+1}\gamma(u) & \rho < 0. \end{cases} \tag{1.11}$$

which means that univariate autoregressive models preserve skewness (as measured by the skewness coefficient) which originates from a one-time shock occurrence regardless of the value of the autoregressive coefficient $\rho > 0$ and preserve absolute skewness regardless of $\rho < 0$. This is true for all $t = 1, 2, ..., T$. It may come as a surprise, since the effect in magnitude of u_0 exerted on ξ_t evaporates totally (in the limit) as t increases.

Now, still being in the univariate case, let us drop the assumption that $u_t = 0$ for $t > 1$. Once again we use representation (1.10). Employing remark (1.2.2) we see that:

$$\kappa_n(\xi_t) = \begin{cases} \frac{1-\rho^{nt}}{1-\rho^n}\kappa_n(u) & |\rho^n| \neq 1, \\ t\kappa_n(u) & |\rho^n| = 1. \end{cases} \tag{1.12}$$

which, for $n = 2, 3$, means, that:

$$\kappa_n(\xi_t) = \begin{cases} \frac{1-\rho^{nt}}{1-\rho^n}\kappa_n(u) & |\rho| \neq 1, n = 2, 3 \\ t\kappa_n(u) & \rho = 1, n = 2, 3 \\ t\kappa_n(u) & \rho = -1, n = 2 \\ \frac{1-(-1)^t}{2}\kappa_n(u) & \rho = -1, n = 3. \end{cases} \tag{1.13}$$

Using remarks (1.2.1–1.2.3) we see that for $\rho = 1$ we have $\gamma(\xi_t) = t^{-\frac{1}{2}}\gamma(u) \to 0$ which is a well known property that skewness of a sum of *iid* random variables vanishes with time. Also for $\rho = -1$ we have that $\gamma(\xi_t) = \frac{1-(-1)^t}{2}t^{-\frac{3}{2}}\gamma(u) \to 0$. In the first case, i.e. for $\rho = 1$, convergence of $\gamma(\xi_t)$ is monotonic, whereas in the latter, i.e. for $\rho = -1$, $\gamma(\xi_t)$ oscillates with t. For $|\rho| \neq 1$ skewness coefficient of ξ_t equals:

$$\gamma(\xi_t) = \frac{\frac{1-\rho^{3t}}{1-\rho^3}}{\left(\frac{1-\rho^{2t}}{1-\rho^2}\right)^{\frac{3}{2}}}\gamma(u) = \theta(\rho, t)\gamma(u) \to \theta(\rho, \infty)\gamma(u) = \frac{(1-\rho^2)^{\frac{3}{2}}}{1-\rho^3}\gamma(u) \tag{1.14}$$

which, this time, depends both on ρ and t. We used a notation $\theta(\rho, t) = \frac{1-\rho^{3t}}{1-\rho^3}\left(\frac{1-\rho^{2t}}{1-\rho^2}\right)^{-\frac{3}{2}}$. Now we can form the following:

[9] The first case shows what is the response (in terms of skewness) of ξ_t, $t = 1, 2, ..., T$, to an impulse from $u_t \sim p(\ldots)$ which occurred in period $t = 1$. The second case predicts response of ξ_t, $t = 1, 2, ..., T$, generated according to (1.7).

Proposition 1.2.1. *Assume model (1.7) for ξ_t. Assume that $\mu_{\xi_0} = 0$ and $\sigma_0 = 0$ (so that $\xi_1 = u_1$). Let t be fixed. Then, $\gamma(\xi_t) = \theta(\rho, t)\gamma(u)$, and $\theta(\rho, t)$, as a function of $\rho \in (-1, 1)$, increases with $\rho \in (-1, 0)$, decreases with $\rho \in (0, 1)$ and reaches a maximum value of one for $\rho = 0$.*

This means, that for stationary models, i.e. for $|\rho| < 1$, if shocks u_t are positively skewed, i.e. if $\gamma(u) > 0$, the skewness coefficient of states is constant and maximal for $\rho = 0$ and it decreases as ρ departs from zero both to the left or to the right until it reaches 1 or -1. If shocks are negatively skewed, i.e. if $\gamma(u) < 0$, everything is the other way round, i.e. the skewness coefficient of states is constant and minimal for $\rho = 0$ and it increases as ρ departs from zero both to the left or to the right until it reaches 1 or -1. In both cases sign of $\gamma(\xi_t)$ equals the sign of γ_u. Proposition (1.2.1) states how $\gamma(\xi_t)$ behaves as a function of ρ. Behavior of $\gamma(\xi_t)$ as a function of t for fixed ρ is stated in the following:

Proposition 1.2.2. *Let assumptions be as in Proposition (1.2.1), but let $\rho \in (-1, 1)$ be fixed instead of t. Then $\gamma(\xi_t)$ is constant over time and equal to $\gamma(u)$ for $\rho = 0$ and decreases with time for $\rho \in (-1, 0) \times (0, 1)$ reaching the limit of $0 < \theta(\rho, \infty)\gamma(u) < \gamma(u)$ where $\theta(\rho, \infty) = \frac{(1-\rho^2)^{\frac{3}{2}}}{1-\rho^3}$ and $0 < \theta(\rho, \infty) < 1$. The limiting fraction of $\gamma(u)$, i.e. $\theta(\rho, \infty)$, is an increasing function of $\rho \in (-1, 0)$, decreasing function of $\rho \in (0, 1)$ and reaches a maximum value of one for $\rho = 0$.*

This means, that for stationary models, i.e. when $|\rho| < 1$, skewness of states $\gamma(\xi_t)$ evaporates with time, but it does not vanish totally, reaching in the limit some fraction $0 < \theta(\rho, \infty) < 1$ of skewness shocks $\gamma(u)$. The limiting fraction is an increasing function of $\rho \in (-1, 0)$ and a decreasing function of $\rho \in (0, 1)$. This is in contrast with the random walk specifications when skewness evaporates totally with the decay rate of $t^{-\frac{1}{2}}$ and $t^{-\frac{3}{2}}$ for $\rho = 1$ and $\rho = -1$ respectively.

Now we will discuss the multivariate case. More specifically, we will show how skewness propagates in a model of the form:

$$\xi_t = A\xi_{t-1} + u_t \tag{1.15}$$
$$u_t \sim p(\ldots) \tag{1.16}$$
$$\xi_0 \sim N(\mu_{\xi_0}, \Sigma_0) \tag{1.17}$$

for $t = 1, 2, ..., T$, where $A \neq 0$, $|\Sigma_0| \geq 0$, eigenvalues of A are less then one in modulus, so that model (1.15) is non-explosive[10] and $p(\ldots)$ is any distribution such that $\gamma(u_t)$ exists for every t and is constant, i.e. $\gamma(u_t) = \gamma(u)$ for every t[11].

In what follows, we assume[12] that $\vartheta_u = 0$ and that univariate elements of u_t are independent, see corollary (1.1.6). The difference between models (1.7) and (1.15) for $\xi_t \in \mathrm{R}^p$ is that in the latter case it is allowed that $p > 1$. As in the univariate case, in what follows we assume for simplicity that $\mu_{\xi_0} = 0$ and $\Sigma_0 = 0$, so that $\xi_1 = u_1$.

To determine effects of innovations u_t exerted on states ξ_t we resort to the moving average representation:

[10]This is true in case of DSGE models.
[11]Note that here $\gamma(u) \in \mathrm{R}^p$.
[12]This assumption simplifies the considerations and does not change meaning of obtained results.

$$\xi_t = A^{t-1}\xi_1 + \sum_{k=0}^{t-2} A^k u_{t-k} = \sum_{k=0}^{t-1} A^k u_{t-k} \tag{1.18}$$

where the last equality follows from the simplifying assumption about ξ_1.

It has to be made explicit that we are interested in skewness coefficients of elements of ξ_t, i.e. of one-dimensional variables $\xi_{t,i}$ for $i = 1,2,...,p$, $t = 1,2,...,T$, and not in synthetic multivariate skewness measures of ξ_t regarded as p-dimensional variables. Respective skewness coefficients will be denoted by $\gamma(\xi_{t,i}) = \frac{\kappa_3(\xi_{t,i})}{(\kappa_2(\xi_{t,i}))^{\frac{3}{2}}}$.

Since variables u_t are independent for $t = 1,2,...,T$, employing remark (1.2.3) we see that[13]:

$$\kappa_n(\xi_t) = \kappa_n\left(\sum_{k=0}^{t-1} A^k u_{t-k}\right) = \sum_{k=0}^{t-1}\left(A^k\right)^{\circ(n)}\kappa_n(u) \tag{1.19}$$

where $A^{\circ(n)}$ denotes the n-th Hadamard (or Schur) power of matrix A, i.e. $A^{\circ(n)} = \underbrace{A \circ A \circ ... \circ A}_{n}$, for $\circ$ denoting the Hadamard (or Schur) product, i.e. elementwise multiplication. From (1.19) we see, more explicitly, that:

$$\kappa_n(\xi_{t,i}) = \sum_{k=0}^{t-1}\sum_{j=1}^{p}(a_{ij}^k)^n \kappa_n(u_{.,j}) \tag{1.20}$$

where a_{ij}^k denotes the ij-th entry of A^k and time indexes for shocks u were suppressed so that $u_{.,j}$ denotes the j-th element of u for any t[14].

Let us now try to determine skewness of $\gamma(\xi_{t,i})$, $i = 1,2,...,p$, assuming that $u_1 \neq 0$ and $u_t = 0$ for $t > 1$. In this case, see eq. (1.20), $\kappa_n(\xi_t) = (A^k)^{\circ(n)}\kappa_n(u)$, which converges with t to a zero vector as long as A is nonexplosive. Unfortunately, not much can be said in general about skewness coefficients $\gamma(\xi_{t,i})$ as functions of elements of matrix A except for the fact, that using 1.20 we readily obtain closed-form formulae. Therefore, let us consider only a simple case in which exactly one of the shocks in u_1 has a nonzero value:

Proposition 1.2.3. *Assume model (1.15) for ξ_t. Assume also that $u_t = 0$ for $t > 1$ and that $u_{1,r} \neq 0$ for some $r \in \{1,2,...,p\}$, whereas $u_{1,j} = 0$ for all $j \in \{1,2,...,p\} \setminus r$. Then:*

$$\gamma(\xi_{t,i}) \propto \begin{cases} 1 & \text{if } a_{i,r}^{t-1} > 0, \\ -1 & \text{if } a_{i,r}^{t-1} < 0. \end{cases} \tag{1.21}$$

where $a_{i,j}^k$ denotes the (i,j)-th element of A^k. It follows, that the series of skewness coefficients $\gamma(\xi_{t,i})$, converges with t if and only if there exists $t' \geq 0$, such that $a_{i,r}^k > 0$ for all $t \geq t'$ or $a_{i,r}^k < 0$ for all $t \geq t'$. Moreover, if such t' exists, then $\gamma(\xi_{t,i})$ is constant for $t > t'$.

[13] We drop time indices for u_t since variables u_t are assumed to be *iid*. Also, for ξ_t being a p-dimensional variable, $\kappa_n(\xi_t)$ denotes a vector cumulant with entries $\kappa_n(\xi_{t,i})$ for $i = 1,2,...,p$.

[14] Less explicit, but considerably more parsimonious expressions for $\kappa_n(\xi_{t,i})$ for all n, t and i simultaneously can be obtained using notations of tensor calculus. The latter approach would also be advisable in case of nondegenerate dependency structure among entries of u_t. Since we do not pursuit higher-order cumulants than the third one and shocks are independent, we stay with the explicit notation (1.20).

1

Proposition (1.2.3) states, that in this simple case skewness coefficients $\gamma(\xi_{k,i})$, $i = 1, 2, ..., p$, can converge with t or oscillate around 0 with a constant amplitude, and, if any of them converges, than it equals its limit starting from some t. This is somehow analogical to the univariate stationary case when skewness coefficient was constant for $\rho > 0$ and oscillated for $\rho < 0$ while being constant in magnitude.

Measurements

Transition equation (1.15) in a state space model is accompanied by a measurement equation of the form:

$$y_t = F\xi_t + He_t \tag{1.22}$$

$$e_t \sim N(0, \Sigma_e) \tag{1.23}$$

for $t = 1, 2, ..., T$, where y_t denote obserables, $F \in \mathrm{R}_{m\times n}$, $H \in \mathrm{R}_{m\times n}$ and $\Sigma_e \in \mathrm{R}_{m\times m}$, $|H| > 0$, $|\Sigma_e| > 0$. Employing remarks (1.2.1–1.2.3), we see that:

$$\gamma(y_t) = \frac{F^{\circ(3)}\kappa_3(\xi_t)}{\left(F^{\circ(2)}\kappa_2(\xi_t) + H^{\circ(2)}\kappa_2(u)\right)^{\frac{3}{2}}} < \gamma(Fy_t) \tag{1.24}$$

because $H^{\circ(2)}\kappa_2(u_t) > 0$ and $\gamma(Fy_t) = \frac{F^{\circ(3)}\kappa_3(\xi_t)}{\left(F^{\circ(2)}\kappa_2(\xi_t)\right)^{\frac{3}{2}}}$. Division and exponentiation in (1.24) is elementwise.

1.3 Propagation of the *csn* distribution

Having provided formulae for propagation of skewness in the state space setting, we now turn to the special case of *csn*-distributed disturbances. More specifically, we will track how distributions of states and measurements change over time. Knowing this is essential for a maximum likelihood estimation. As previously, we start with a univariate model and then extend results to the multivariate case, in which we show that the *csn* distribution does not in general propagate through the state space setting. The reason is that the autoregressive matrix *A* in the transition equation can be singular. Hence, analytical maximum likelihood estimation requires an appropriate regularization[15] for matrix *A*. In this paper, however, we alow that *A* be singular and consider a quasi-maximum likelihood alternative.

State variables

Let us consider model (1.7) with an additional assumption that:

$$p(\ldots) = csn(\tilde{\mu}_u, \tilde{\sigma}_u, d_u, \vartheta_u, \delta_u)$$

with $\tilde{\sigma}_u > 0$, $d_u \in \mathrm{R}$, $\mu_{\xi_0} \in \mathrm{R}$ where $\tilde{\mu}_u$ is set in such a way that[16] $\mathrm{E}(u_t) = \mathrm{E}(u) = 0$.

[15] Such a regularization is feasible by means of an appropriate model reformulation.
[16] Parameter $\tilde{\mu}_u$ is therefore not free, but equals $-\sqrt{\frac{2}{\pi}}\frac{d_u\tilde{\sigma}_u}{\sqrt{1+d_u^2\tilde{\sigma}_u}}$ which implies that $\mathrm{E}(u_t) = 0$.

First we will investigate the effect on ξ_t exerted by innovation u_t at time $t=1$ keeping $u_t=0$ for $t>1$. Since ξ_1 is a sum of a normally distributed variable $\rho\xi_0$ and a $csn_{1,1}$-distributed variable u_1, it is, according to corollary (1.1.7), a csn random variable with parameters $\tilde{\mu}_{\xi,1}=\tilde{\mu}_u$, $\tilde{\sigma}_{\xi,1}=\rho^2\sigma_0+\tilde{\sigma}_u$, $d_{\xi,1}=d_u\frac{\tilde{\sigma}_u}{\tilde{\sigma}_{\xi,1}}$, $\vartheta_{\xi,1}=\vartheta_u$ and $\delta_{\xi,1}=\delta_u+\tilde{\sigma}_u d_u^2(1-\frac{\tilde{\sigma}_u}{\tilde{\sigma}_{\xi,1}})$. To see how effects of u_1 propagate through ξ_t, let us notice that $\xi_t=\rho\xi_{t-1}=\rho^{t-1}\xi_1$ for $t>1$, hence, according to corollary (1.1.8), variable ξ_t has a $csn_{1,1}$ distribution with parameters:

$$\tilde{\mu}_{\xi,t}=\rho\tilde{\mu}_{\xi,t-1}=\rho^{t-1}\tilde{\mu}_{\xi,1},\quad \tilde{\sigma}_{\xi,t}=\rho^2\tilde{\sigma}_{\xi,t-1}=\rho^{2(t-1)}\tilde{\sigma}_{\xi,1}, \tag{1.25}$$

$$d_{\xi,t}=\frac{1}{\rho}d_{\xi,t-1}=\frac{1}{\rho^{t-1}}d_{\xi,1},\quad \vartheta_{\xi,t}=\vartheta_{\xi,t-1}=\vartheta_{\xi,1},\quad \delta_{\xi,t}=\delta_{\xi,t-1}=\delta_{\xi,1} \tag{1.26}$$

Hence, skewness parameter of ξ_t equals $d_{\xi,t}=\frac{1}{\rho^{t-1}}d_{\xi,1}=\frac{1}{\rho^{t-1}}d_u\frac{\tilde{\sigma}_u}{\tilde{\sigma}_{\xi,1}}$. If $|\rho|<1$, then $|d_{\xi,t}|$ increases with t without bound[17] regardless of the shocks' skewness parameter $d_u\neq 0$. If $|\rho|=1$, then $|d_{\xi,t}|$ equals $d_{\xi,1}=d_u\frac{\tilde{\sigma}_u}{\tilde{\sigma}_{\xi,1}}$ for all t and if $|\rho|>1$, then $|d_{\xi,t}|$ decreases with t reaching zero in the limit. If $\rho<0$, then sign of $d_{\xi,t}$ additionally oscillates.

This basic fact can easily be misunderstood, because, since the magnitude of $d_{\xi,t}$, as measured for example by $|d_{\xi,t}|$, implies in some sense the absolute (i.e. left or right) strength of skewness, one could conclude that absolute skewness intensifies with time in stationary models, is time invariant in case of random walk models and evaporates with time under explosive specifications. This is, however, not the case, because variance of ξ_t also changes with t. As a consequence, skewness of ξ_t is constant over time for $\rho>0$ and oscillates around zero with a constant amplitude for $\rho<0$, which is in line with results obtained in the previous section. In a more explicit way application of corollary (1.1.4) for $\rho>0$ states that:

$$\gamma(\xi_t)=\left(2-\frac{\pi}{2}\right)\left(\sqrt{\frac{2}{\pi}}\right)^3\frac{\left(\frac{\sigma_t d_t}{(1+\sigma_t d_t^2)^{\frac{1}{2}}}\right)^3}{\left(\sigma_t-\frac{2}{\pi}\frac{\sigma_t^2 d_t^2}{1+\sigma_t d_t^2}\right)^{\frac{3}{2}}}= \tag{1.27}$$

$$=\left(2-\frac{\pi}{2}\right)\left(\sqrt{\frac{2}{\pi}}\right)^3\frac{\left(\frac{\sigma_1 d_1}{(1+\sigma_1 d_1^2)^{\frac{1}{2}}}\right)^3}{\left(\sigma_1-\frac{2}{\pi}\frac{\sigma_1^2 d_1^2}{1+\sigma_1 d_1^2}\right)^{\frac{3}{2}}}=\gamma(\xi_1) \tag{1.28}$$

Time independency of skewness of the impulse response distribution in state space setting with csn disturbances appears also as a consequence of applying for $z_t=\xi_t$ the following more general:

Proposition 1.3.1. *Let $z_t\in\mathbb{R}$, for $t=1,2,...,T$, be distributed according to a $csn_{1,1}$ distribution with parameters $\tilde{\mu}_{z,t}$, $\tilde{\sigma}_{z,t}>0$, $d_{z,t}$, $\vartheta_{z,t}$ and $\delta_{z,t}>0$. Assume that $\vartheta_{z,t}=0$ and that $\delta_{z,t}=const$ for all t. If $\tilde{\sigma}_{z,t}d_{z,t}=const$ and $\tilde{\sigma}_{z,t}d_{z,t}^2=const$ for all t, then absolute value of the skewness coefficient of z_t, i.e. $|\gamma(z_t)|$, is constant over time and $sgn(\gamma(z_t))$ equals $sgn(d_t)$.*

[17] I.e. $\lim_{t\to\infty}|d_{\xi,t}|=\infty$.

Now we drop the assumption that $u_t = 0$ for $t > 1$. We make use of moving average representation (1.10)[18]. Distribution of ξ_1 is a $csn_{1,1}$ distribution with parameters[19]: $\tilde{\mu}_{\xi,1}$, $\tilde{\sigma}_{\xi,1}$, $d_{\xi,1}$, $\vartheta_{\xi,1}$ and $\delta_{\xi,1}$. Also, using corollary (1.1.8), random variables $v_k = \rho^k u_{t-k}$, for $k = 0, 1, ..., t-2$, have $csn_{1,1}$ distributions, but with parameters: $\tilde{\mu}_{v,k} = \rho^k \tilde{\mu}_u$, $\tilde{\sigma}_{v,k} = \rho^{2k}\tilde{\sigma}_u$, $d_{v,k} = \frac{1}{\rho^k} d_u$, $\vartheta_{v,k} = \vartheta_u$ and $\delta_{v,k} = \delta_u$. From corollary (1.1.10), ξ_t has therefore a $csn_{1,t}$ distribution with parameters:

$$\tilde{\mu}_{\xi,t} = \frac{1-\rho^{t-1}}{1-\rho}\tilde{\mu}_u + \rho^{t-1}\tilde{\mu}_{\xi,1}, \quad \tilde{\sigma}_{\xi,t} = \frac{1-\rho^{2(t-1)}}{1-\rho^2}\tilde{\sigma}_u + \rho^{2(t-1)}\tilde{\sigma}_{\xi,1} \tag{1.29}$$

$$D_{\xi,t} = r_t^T \frac{d_u \tilde{\sigma}_u}{\sigma_{\xi,t}}, \quad \vartheta_{\xi,t} = 1 \otimes \vartheta_{\xi,1} \quad \delta_{\xi,t} = \delta_{\xi,1} \tag{1.30}$$

where $r_t = (\rho^{t-1}, \rho^{t-2}, ..., \rho, 1)^T$ and $\otimes$ denotes the tensor (Kronecker) product. Above formulae are valid for $|\rho| \neq 1$. To derive them we only need to notice that that $d_{v,k}\sigma_{v,k} = \rho^k d_u \tilde{\sigma}_u$ and that $d_{\xi,1}\tilde{\sigma}_{\xi,1} = d_u \tilde{\sigma}_u$. For $\rho = 1$ the difference is only that $\tilde{\mu}_{\xi,t} = (t-1)\tilde{\mu}_u + \tilde{\mu}_{\xi,1}$, $\tilde{\sigma}_{\xi,t} = (t-1)\tilde{\sigma}_u + \tilde{\sigma}_{\xi,1}$ and $r_t = (1_t)^T$.

Now we will discuss the multivariate case. Let us consider model (1.15) with an additional assumption that:

$$p(\ldots) = csn(\tilde{\mu}_u, \tilde{\Sigma}_u, \tilde{D}_u, \vartheta_u, \Delta_u)$$

where $|\tilde{\Sigma}_u| > 0$, $|\Delta_u| > 0$ and $\tilde{\mu}_u$ is chosen in such a way that $\mathrm{E}(u_t) = 0$. We assume[20] that $\vartheta_u = 0$ and that univariate elements of u_t are independent. In what follows we assume for simplicity that $\mu_{\xi_0} = 0$ and $\Sigma_0 = 0$, so that $\xi_1 = u_1$.

The multivariate case differs from the univariate one in a fundamental way. The univariate model (1.7) assures that the state variable ξ_t is distributed according to a $csn_{1,t}$ distribution for all t, i.e. that the csn distribution is closed under transformations which model (1.7) applies to ξ_t. In the multivariate case this does not have to be the case. To check if ξ_t has a csn distribution, we give the following:

Proposition 1.3.2. *Let ξ_{t-1} be distributed according to a $csn_{p,q}$ for some $p, q \geq 1$ with parameters $\tilde{\mu}_{\xi,t-1}$, $\tilde{\Sigma}_{\xi,t-1} \geq 0$, $D_{\xi,t-1}$, $\vartheta_{\xi,t-1}$ and $\Delta_{\xi,t} > 0$. Let $z_t = A\xi_{t-1}$, $A \in \mathrm{R}_{p,p}$. Then, z_t has a csn (possibly singular) distribution if and only if $r(A^T) = r([A^T|w_i])$ for all $i = 1, 2, ..., q$, where $r(A)$ denotes rank of A and w_i denotes the i-th row $D_{\xi,t-1}$.*

Proposition (1.3.2) states, that for a csn variable ξ_{t-1}, variable $z_t = A\xi_{t-1}$ has a csn distribution if and only if rows of $D_{\xi,t-1}$ are linear combinations of rows of A. In other words, all rows of $D_{\xi,t-1}$ must belong to $span(A^T)$, i.e. to the image of A^T. This condition is always satisfied (regardless of $D_{\xi,t}$) if A has a full rank. However, for a rank deficient operator A this is a very restrictive condition, since $D_{\xi,t-1}$ can be in principle arbitrary. Although proposition (1.3.2) constitutes a negative result for ξ_t as a p-dimensional variable, it has to be stressed that it is assured that elements of ξ_t are csn distributed. In this paper we assume that A can be rank deficient, hence the distribution of states is in general not a csn distribution, hence we do not go for an analytical maximum likelihood estimation.

[18] Without the simplifying assumption that $\xi_1 = u_1$.

[19] Their values have already been provided in this section.

[20] This assumption simplifies the considerations.

1

Measurements

As stated in proposition (1.3.2), state variables ξ_t can fall out of the *csn* distribution family starting from some $t > 1$ if the autoregressive matrix A in model (1.15) is rank deficient. In case of DSGE models, especially larger ones, this is usually the case. In what follows, we notice that even if A is rank deficient, observed variables still follow a *csn* distribution for all t. The reason for this is that H has a full row rank. To see this notice that:

$$\begin{aligned} y_t &= F\xi_t + He_t = FA^{t-1}\xi_1 + F\sum_{k=0}^{t-2} A^k u_{t-k} + He_t = && (1.31) \\ &= F\sum_{k=0}^{t-1} A^k u_{t-k} + He_t = A_t\omega_t && (1.32) \end{aligned}$$

where $A_t = \left[FA^{t-1}|FA^{t-2}|...|FA|F|H\right]$ and $\omega_t = \left[u_1, u_2, ..., u_t, e_t\right]^T$ and we assumed for simplicity that $\xi_1 = u_1$. Matrix A_t has a full row rank since H is full rank and ω_t has a nonsingular *csn* distribution, hence y_t follows a nonsingular *csn* distribution, which may come as a surprise since ξ_t not only can have a singular *csn* distribution, but can have some other, i.e. not *csn*, distribution.

Chapter 2

2

A DSGE model with structural skewness

To investigate some issues related to skewness in DSGE models we employ the small open economy model of Lubik and Schorfheide (2007) (LS) which is a simplified version of Gali and Monacelli (2005) and extend it by allowing (some of the) structural shocks to follow a closed skewed normal distribution. LS model can be seen as a minimum set of equations for an open economy framework and its small size is an advantage because it reduces computational burden of simulations which we conduct. Below we present model's equation in already log-linearised form (denoted by hats over variables), more details can be found in Lubik and Schorfheide (2007) or Negro and Schorfheide (2008). The model is also implemented in YADA package (Warne, 2010).

2.1 The model

There are nine state variables in the model: a growth rate of a non-stationary world technology $\left(z_t \equiv \frac{A_t}{A_{t-1}}\right)$ where (A_t) denotes the non-stationary technology, foreign output $(\widehat{y}^\star_t)$ and inflation $(\widehat{\pi}^\star_t)$, terms of trade growth rate $(\Delta\widehat{q}_t)$, monetary policy shock $u(R)$, domestic output $(\widehat{y}_t)$ and inflation $(\widehat{\pi}_t)$, exchange rate growth rate $(\Delta\widehat{e}_t)$, and nominal interest rate $(\widehat{R}_t)$. The non-stationary technology process is assumed to be present in all real variables therefore, to ensure stationarity, all real variables are expressed as deviations from A_t.

Four state variables are approximated by autoregressions with normally distributed shocks:

$$\widehat{z}_t = \rho_z \widehat{z}_{t-1} + u(z)_t \tag{2.1}$$
$$\widehat{\pi}^\star_t = \rho_{\pi^\star} \widehat{\pi}^\star_{t-1} + u(\pi^\star)_t \tag{2.2}$$
$$\widehat{y}^\star_t = \rho_{y^\star} \widehat{y}^\star_{t-1} + u(y^\star)_t \tag{2.3}$$
$$\Delta\widehat{q}_t = \rho_q \Delta\widehat{q}_{t-1} + u(\Delta q)_t \tag{2.4}$$

Random processes $u(z)_t$, $u(\pi^\star)_t$, $u(y^\star)_t$, and $u(\Delta q)_t$ as well as the monetary policy shock $u(R)_t$ represent structural shocks or innovations. In the original formulation of Lubik and Schorfheide (2007) they are all normally, hence symmetrically, distributed. In our approach each structural shock follows a closed skewed normal $csn_{1,1}$ distribution:

$$u(z)_t \sim csn_{1,1}\left(\tilde{\mu}_z, \tilde{\sigma}_z, d_z, \vartheta_z, \delta_z\right) \tag{2.5}$$
$$u(\pi^\star)_t \sim csn_{1,1}\left(\tilde{\mu}_{\pi^\star}, \tilde{\sigma}_{\pi^\star}, d_{\pi^\star}, \vartheta_{\pi^\star}, \delta_{\pi^\star}\right) \tag{2.6}$$
$$u(y^\star)_t \sim csn_{1,1}\left(\tilde{\mu}_{y^\star}, \tilde{\sigma}_{y^\star}, d_{y^\star}, \vartheta_{y^\star}, \delta_{y^\star}\right) \tag{2.7}$$
$$u(q)_t \sim csn_{1,1}\left(\tilde{\mu}_{\Delta q}, \tilde{\sigma}_{\Delta q}, d_{\Delta q}, \vartheta_{\Delta q}, \delta_{\Delta q}\right) \tag{2.8}$$
$$u(R)_t \sim csn_{1,1}\left(\tilde{\mu}_R, \tilde{\sigma}_R, d_R, \vartheta_R, \delta_R\right) \tag{2.9}$$

which means that they can, but do not have to be, normally distributed. We demand that parametrization of shocks makes them martingale difference sequences.

2

Euler equation combined with perfect risk sharing and the market-clearing condition for the foreign good gives rise to an open economy dynamic IS curve:

$$\begin{aligned}\widehat{y}_t = & \operatorname*{E}_t \widehat{y}_{t+1} - [\tau + \alpha(2-\alpha)(1-\tau)]\left(\widehat{R}_t - \operatorname*{E}_t \widehat{\pi}_{t+1}\right) - \rho_z \widehat{z}_t \\ & - \alpha[\tau + \alpha(2-\alpha)(1-\tau)] \operatorname*{E}_t \Delta\widehat{q}_{t+1} + \alpha(2-\alpha)\frac{1-\tau}{\tau}\operatorname*{E}_t \Delta\widehat{y}^\star_{t+1}.\end{aligned} \tag{2.10}$$

Parameters τ and α denote the intertemporal substitution elasticity and the import share (hence $0 < \alpha < 1$ and for $\alpha = 0$ equation reduces to closed economy variant).

Optimal price setting by domestic firms leads to the neokeynesian Phillips curve:

$$\widehat{\pi}_t = \beta \operatorname*{E}_t \widehat{\pi}_{t+1} + \alpha\beta \operatorname*{E}_t \Delta\widehat{q}_{t+1} - \alpha\Delta\widehat{q}_t + \frac{\kappa}{\tau + \alpha(2-\alpha)(1-\tau)}\left(\widehat{y}_t - \widehat{\overline{y}}_t\right), \tag{2.11}$$

where $\widehat{\overline{y}}_t = -\alpha(2-\alpha)(1-\tau)/\tau\widehat{y}^\star_t$ is the potential output in the absence of nominal rigidities. The parameter β is the discount factor. The parameter κ is a function of underlying structural parameters (elasticities of labour supply and demand, price stickiness), and it is treated itself as structural.

Definition of consumer prices under the assumption of relative PPP allows to determine change in nominal exchange rate as:

$$\Delta\widehat{e}_t = \widehat{\pi}_t - \widehat{\pi}^\star_t - (1-\alpha)\Delta\widehat{q}_t. \tag{2.12}$$

The nominal interest rate is assumed to follow a policy rule:

$$\widehat{R}_t = \rho_R \widehat{R}_{t-1} + \left(1-\rho_R\right)\left[\Psi_\pi \widehat{\pi}_t + \Psi_y \widehat{y}_t + \Psi_e \Delta\widehat{e}_t\right] + u(R)_t, \tag{2.13}$$

where ρ_R is a smoothing parameter.

Following Lubik and Schorfheide, we use five observable variables to link the model with the data: real GDP growth, annualised inflation rate, annual nominal interest rate, change in exchange rate, and change in terms of trade. The measurement equations take the form:

$$\underline{\pi_t} = 4\left(\frac{\pi_A}{400} + 1\right)\widehat{\pi}_t + \pi_A, \tag{2.14}$$
$$\underline{\Delta e_t} = \Delta\widehat{e}_t, \tag{2.15}$$
$$\underline{y_t} = \Delta\widehat{y}_t + \widehat{z}_t + \gamma_Q, \tag{2.16}$$

$$\underline{\Delta q_t} = \Delta \widehat{q}_t, \tag{2.17}$$

$$\underline{R_t} = 4\left(\frac{\pi_A + r_A + 4\gamma_Q}{400} + 1\right)\widehat{R}_t, + \pi_A + r_A + 4\gamma_Q, \tag{2.18}$$

where underlined variables denote observable variables. Parameter π_A is annual rate of inflation, γ_Q is quarterly growth rate of non–stationary technology process (z_t in steady state), and r_A is an element of real interest rate $r = r_A + 4\gamma_Q$.

2

2.2 Simulation exercises on skewness propagation

In order to numerically assess how structural skewness propagates in the LS model, we simulated 10000 samples of observables, each consisting of 600 observations. Two cases were considered. In the first case shocks were assumed to be normal, whereas in the second one structural skewness was introduced by assuming that exactly one shock follows a closed skew normal distribution. Parameters of *csn* distributions were chosen in such a way, that the skewness coefficient of each shock was equal to 0.50, so that structural skewness is always positive, which means shocks draws from above the mean value are less probable than those from below the mean. Behavioral parameters and standard deviations of shocks as well as autocorrelation coefficients of states were motivated by LS's central values of priors (see Lubik and Schorfheide, 2007, p. 1077). Standard deviations of measurement errors are approximately 10% of observed variables' standard deviations. Table 2.1 shows the basic set of parameters of LS model.

In each of the considered cases skewness of states and observables was calculated. Results are reported in Table 2.2, columns of which contain skewness coefficients under normality (column 2), when exactly one shock is skewed (columns 3–7), and when all shocks are skewed (column 8), both in the block of state variables and observable ones. Let us first notice that the skewness of autoregressive variables[1] $\widehat{z}_t$, $\Delta\widehat{q}_t$, $\widehat{y}^\star_t$, $\widehat{\pi}^\star_t$, and $u_t(R)$ depends on their autoregressive coefficients (which are reported in Table 2.1), as predicted in Proposition 1.3.1, i.e. the higher the autoregressive coefficient, the smaller the skewness. Under reported parametrization skewness of states and observable variables when all shocks are assumed to be skewed is roughly equal to sum of skewnesses implied by each of the shocks, but in general this does not have to be the case. Furthermore, inflation does not have its own shock in the model, but factors which induce positive skewness of CPI inflation — foreign demand and foreign price dynamics, generate positive skewness of the nominal interest rate, which reveals the pattern of propagation through the monetary policy rule. And the other way round — positively skewed monetary policy shock is reflected by positive skewness of the interest rate and a negative contribution to skewness of inflation. Finally, skewness of output is driven mainly by skewness of growth rates of technology. Positive skewness of foreign inflation is also the main cause of skewness of changes in exchange rate.

The results raise the question whether it is possible to replicate with the LS model, the pattern of skewness observed in the real data as presented in the introduction — positive skewness for inflation, nominal interest rate and depreciation rate, and negative skewness of output growth rates. Although such combination of skewed shocks exists it would require highly positively

[1] The monetary policy shock $u_t(R)$ can also be perceived as an autoregressive process with autoregression coefficient equal to zero.

skewed foreign output shock which is not reasonable as we expect it to be negatively skewed. However, it is possible to replicate the pattern of skewness if we change specification of the exchange rate equation in the model, replacing relative PPP with some version of uncovered interest rate parity and introducing risk premium shock.

Table 2.1. The basic set parameters of Lubik-Schorfheide DSGE model.

Behavioral			Disturbances			Measurement errors		
Param.	Value	Note	Param.	Value	Note	Param.	Value	Note
ψ_π	1.500		$\rho_{\hat{z}}$	0.200		$\sigma_{\underline{y}}$	0.01	kept fixed
ψ_y	0.250		$\rho_{\Delta\hat{q}}$	0.400		$\sigma_{\underline{\pi}}$	0.09	kept fixed
$\psi_{\Delta e}$	0.100		$\rho_{\hat{y}^\star}$	0.900		$\sigma_{\underline{R}}$	0.09	kept fixed
ρ_R	0.600		$\rho_{\hat{\pi}^\star}$	0.800		$\sigma_{\Delta\underline{e}}$	0.16	kept fixed
α	0.150		$\rho_{\hat{\epsilon}^R}$	0.000	kept fixed	$\sigma_{\Delta\underline{q}}$	0.04	kept fixed
κ	0.500		$\sigma_{\hat{z}}$	1.000				
τ	0.500		$\sigma_{\Delta\hat{q}}$	1.900				
r_A	0.750		$\sigma_{\hat{y}^\star}$	1.890				
π_A	2.000	kept fixed	$\sigma_{\hat{\pi}^\star}$	3.000				
γ_Q	0.800		$\sigma_{\hat{R}}$	0.400				

σ_u — standard deviation of u; ρ_u —- autocorrelation coefficient of u

Table 2.2. Skewness (measured by sample skewness coefficient) in simulated data

Variable	Normal distribution	Skew normal distribution					
		$u(z)$	$u(\Delta q)$	$u(y^\star)$	$u(\pi^\star)$	$u(R)$	all
State variables							
$\widehat{y}$	0.00	0.00	0.00	−0.12	0.00	−0.01	−0.12
$\widehat{\pi}$	0.00	0.00	0.00	0.03	0.09	−0.03	0.08
$\widehat{r}$	0.00	0.00	−0.01	0.03	0.03	0.05	0.09
$\Delta\widehat{e}$	0.00	0.00	−0.02	0.00	−0.16	0.00	−0.19
$\widehat{z}$	0.00	0.47	0.00	0.00	0.00	0.00	0.47
$\Delta\widehat{q}$	0.00	0.00	0.41	0.00	0.00	0.00	0.41
$\widehat{y}^\star$	0.00	0.00	0.00	0.14	0.00	0.00	0.14
$\widehat{\pi}^\star$	0.00	0.00	0.00	0.00	0.21	0.00	0.22
$u(R)$	0.00	0.00	0.00	0.00	0.00	0.50	0.50
Observable variables							
GDP growth rate	0.00	0.25	0.00	−0.03	0.00	0.00	0.21
Inflation	0.00	0.00	0.00	0.03	0.08	−0.03	0.08
Interest rate	0.00	0.00	−0.01	0.03	0.03	0.04	0.09
Exchange rate	0.00	0.00	−0.02	0.00	−0.16	0.00	−0.19
Terms of trade	0.00	0.00	0.40	0.00	0.00	0.00	0.40

Chapter 3

Estimation of models' parameters

In order to work with (first order perturbations of) DSGE models with structural skewness we have to develop a parameter estimation technique for a (linear) state space model with skewed shocks. A state space model which represents a reduced form of a DSGE model under normal shocks is usually estimated by the Kalman filter (KF) maximum likelihood (ML) estimator (see e.g. Hamilton, 1994; Meinhold and Singpurwalla, 1983). Calculation of likelihood function value via KF typically constitutes also a step of Bayesian estimation (see Fernández-Villaverde (2009)). Popularity of KF estimation is motivated by the fact that for normally distributed shocks and measurement errors KF produces analytical filtration, i.e. it yields exact likelihood value — not an approximation, it is fast and easy to implement. Robustness of KF for non-Gaussian shocks in the transition equation is sometimes negated, e.g. (Meinhold and Singpurwalla, 1989). Nonetheless, we keep in mind that KF is an optimal[1] linear filter for arbitrary, hence also for closed skewed normal shocks[2].

Ideally, we would like to extend the KF formulation to the case of *csn* shocks, which would allow us to perform analytical filtration and obtain exact likelihood function in each step of the ML routine for inference of parameters. It is possible, but under assumptions which are not met in case of most DSGE models (see e.g. Naveau et al. (2005)). The problem is a reduced rank of the autoregressive matrix in the transition equation[3] and the fact that calculation of likelihood function value, which has to be a fast task for KF-type filters, requires calculation of a cumulative distribution function value of a highly dimensional normal distribution with arbitrary dependency structure. In practice, the latter task can be done only by Monte Carlo techniques. Also in comparison with numerical burden of methods like particle filtering (see for example Fernández-Villaverde and Rubio-Ramirez (2007); An (2005)) or fully Bayesian parameter estimation which simultaneously involves evaluation of the likelihood, numerical optimization, posterior sampling e.g. via Metropolis-Hastings type of methods and highly dimensional numerical integration via e.g. MCMC-type of methods, a simple two-step KF-based (limited information) approach can be desirable:

1. Kalman filter (quasi-) maximum likelihood (Q-ML) estimation[4] of models' parameters (deep parameters, second moments of shocks and measurement errors, etc.) neglecting skewness,

[1]Optimal in the sense that it produces minimal trace of one-step ahead prediction errors covariance matrix.
[2]This means that better filters are only nonlinear ones.
[3]The issue could be solved via proper transformation of the matrix, however.
[4]We use the therm *quasi maximum likelihood estimator* in a very broad sense, as a case when maximum likelihood principle is applied to a misspecified (statistical) model, see White (1982). We do not relay on properties of Q-ML estimators given by (e.g.) Wedderburn (1974), Gourieroux et al. (1984), Nadler and Lee (1992).

2. filtration (estimation) of shocks conditional upon parameter estimates from step 1; method of moments estimation of parameters of shocks' *csn* probability distribution functions (let us call them *shocks' parameters*), conditional upon filtration; testing for skewness of shocks.

Estimates of all the parameters obtained in step 1, but for shocks' parameters, become final estimates. Final estimates of shocks' parameters are in turn obtained in step 2, in which it is assumed that shocks have a *csn* distribution.

Above procedure omits direct optimization-based estimation of shocks' parameters. It may be an advantage, because the log-likelihood function for the *csn* distribution exhibits anomalies, e.g. improper shape, inflection points in profile likelihood (singularity of Fisher information matrix) at points where skewness vanishes, divergence of parameters of the distribution, see Azzalini and Capitanio (1999), Azzalini (2004)), see also Azzalini and Genton (2008). Only some of these anomalies may be removed via proper parametrization.[5].

3.1 Quasi-maximum likelihood estimation

A first order perturbation of a DSGE model with structural skewness obtains the following state space form:

$$\begin{cases} \xi_t = A\xi_{t-1} + Bu_t, & u_t \sim csn_{p_u,p_u}\left(\tilde{\mu}_u, \tilde{\Sigma}_u, D_u, 0, I\right), \\ & \mu_u = \mathrm{E}(u_t) = 0, \quad \mathrm{var}(u_t) = \Sigma_u; \\ & \xi_0 \sim \mathrm{N}\left(\mu_{\xi_0}, \Sigma_{\xi_0}\right), \\ y_t = F\xi_t + e_t, & e_t \sim \mathrm{N}\left(0, \Sigma_e\right). \end{cases} \tag{3.2}$$

for $t = 1, 2, ..., T$, where ξ_t denote states, y_t observables, e_t and u_t denote martingale difference measurement errors and structural shocks respectively. The covariance matrix Σ_{ξ_0} can be zero, i.e. states at $t = 0$ (ξ_0) can be non-stochastic. Matrices A, B, F are functions of models' deep parameters vector $\theta_{\mathcal{M}}$ and A can be singular (and generally is), but not explosive. If $D_u = 0$,

[5] Azzalini et al. (2010) considered a more general case of estimation of a skew-symmetric distribution's parameters. A simple version of the probability density function of a scalar skew-symmetric random variable may be written, up to a constant, in the form:

$$\mathrm{f}_{SS}(z) = \mathrm{f}_0(z;\theta_a)\,\pi(z;\theta_a,\theta_b), \qquad z \in \mathrm{R} \tag{3.1}$$

where: f_0 is a symmetric density, and π is a skewing function, such that $\pi(-z) = 1 - \pi(z) \geq 0$ for all $z \in \mathrm{R}$. The location and scale parameters are introduced via definition $Y = \nu + \omega Z$. The (closed) skew-normal distribution is a special case of skew-symmetric family, the multivariate extension is straightforward. „The class of distributions (3.1) can be obtained via a suitable censoring mechanism, regulated by $\pi(z)$, applied to samples generated by the base density f_0, [...]. Under this perspective, it is of interest to estimate the parameters of f_0 via a method which does not depend, or depends only to a limited extent, on the component $\pi(z)$, which in many cases is not known, or is not of interest to be estimated [...]", (see Azzalini et al., 2010, p. 2). A distributional invariance property of skew-symmetric distribution is a key concept in their method of estimating equations. The distributional invariance is defined in the following way: If X, Y are two random variables $X \sim \mathrm{f}_0$, $Y \sim \mathrm{f}_{SS}$, and $T(.)$ is an even function, then $T(X) \overset{d}{=} T(Y)$. This property ensures that for any choice of even function T_k the expected value

$$\mathrm{E}T_k\left(\frac{Y-\nu}{\omega}\right) = c_k, \ \ (k = 1, 2, ..)$$

depend only on f_0, provided they exist. The authors use that feature to build estimation equations. Ma et al. (2005) considered a semiparametric model, where the parameters of interest are mean and variance but the skewness parameter is a nuisance parameter. The authors tested properties of regular asymptotically linear (RAL) estimators of Newey (1990). Fletcher et al. (2008) tested a variant of method of moment estimators of *csn* parameters. Akdemir (2009) considered maximum product spacing estimation.

i.e. shocks' skewness parameter vanishes, then $\tilde{\mu}_u = \mu_u$, $\tilde{\Sigma}_u = \Sigma_u$ and shocks u_t are normally distributed. Structural shocks u_t are, by definition, independent, therefore matrices $\tilde{\Sigma}_u$, D_u and Σ_u are diagonal, see corollary (1.1.6), with: $\tilde{\Sigma}_u = \text{diag}(\tilde{\sigma}_{u_i}, i = 1, 2, ..., p_u)$, $D_u = \text{diag}(d_{u_i}, i = 1, 2, ..., p_u)$ and $\Sigma_u = \text{diag}(\sigma_{u_i}, i = 1, 2, ..., p_u)$. This, by remark (1.1.3), means, that each $u_{t,i}$, i.e. each component of u_t, has an Azzalini-type skewed normal distribution, see (Azzalini and Valle, 1996). Applying corollary (1.1.4), we see that first three central moments of $u_{t,i}$ are:

$$\text{E}(u_{t,i}) = \kappa_1(u_{t,i}) = \tilde{\mu}_{u_i} + \sqrt{\frac{2}{\pi}} \frac{d_{u_i} \tilde{\sigma}_{u_i}}{\sqrt{1 + d_{u_i}^2 \tilde{\sigma}_{u_i}}}, \tag{3.3}$$

$$\text{var}(u_{t,i}) = \kappa_2(u_{t,i}) = \tilde{\sigma}_{u_i} - \frac{2}{\pi} \frac{d_{u_i}^2 \tilde{\sigma}_{u_i}^2}{1 + d_{u_i}^2 \tilde{\sigma}_{u_i}} \tag{3.4}$$

$$\text{E}(u_{t,i} - E(u_{t,i}))^3 = \kappa_3(u_{t,i}) = \left(\sqrt{\frac{2}{\pi}}\right)^3 \left(2 - \frac{\pi}{2}\right) \left(\frac{d_{u_i} \tilde{\sigma}_{u_i}}{\sqrt{1 + d_{u_i}^2 \tilde{\sigma}_{u_i}}}\right)^3 \tag{3.5}$$

which means, that skewness coefficients of $u_{t,i}$ are equal to:

$$\gamma(u_{t,i}) = \gamma(u_i) = \left(\frac{4 - \pi}{2}\right) \frac{\left(\sqrt{\frac{2}{\pi}} \frac{d_{u_i} \tilde{\sigma}_{u_i}}{\sqrt{1 + d_{u_i}^2 \tilde{\sigma}_{u_i}}}\right)^3}{\left(\tilde{\sigma}_{u_i} - \frac{2}{\pi} \frac{d_{u_i}^2 \tilde{\sigma}_{u_i}^2}{1 + d_{u_i}^2 \tilde{\sigma}_{u_i}}\right)^{\frac{3}{2}}}. \tag{3.6}$$

and satisfies the condition: $|\gamma(u_{t,i})| < \gamma_{max} \approx 0.995$[6].

Note that shocks u_t, as structural ones, are required to be martingale differences for every $\tilde{\Sigma}_u$ and D_u, which is obtained by forcing $\tilde{\mu}_u$ to adjust so that $\text{E}(u_{t,i}) = 0$, see the first eq. in (3.3).

If $\theta_{\mathcal{M}}$ is fixed, skewness of shocks has therefore no impact on the steady state of the model. Variances of shocks Σ_u are functions of shocks' distribution parameters $\tilde{\Sigma}_u$ and D_u, however, given $\theta_{\mathcal{M}}$ and Σ_u, these parameters imply structure of shocks' variance (shape and scale) and not its magnitude, see the second eq. in (3.3). This is a heuristic motivation of our estimation procedure which first step neglects skewness of the distribution, i.e. the skewing function, and approximates likelihood using the normal distribution.

Let $\theta_u = \left(\tilde{\Sigma}_u, D_u\right)$ and $\theta_e = (\Sigma_e)$. We know (see section 1.3) that observables y_t are distributed according to a $csn\left(\tilde{\mu}_{y,t}, \tilde{\Sigma}_{y,t}, D_{y,t}, \vartheta_{y,t}, \Delta_{y,t}\right)$ distribution and, given observables, likelihood function of the models' parameters $\theta = \left(\theta_{\mathcal{M}}, \theta_u, \theta_e\right)$ is denoted by $\mathcal{L}(\theta)$. We are interested in finding θ which maximizes $\mathcal{L}(\theta)$. Maximizer of θ, denoted by $\hat{\theta}$, will be approximated in two steps. Let $\bar{\theta}_u = (\tilde{\Sigma}_u, 0) = (\Sigma_u, 0)$, $\bar{\theta} = (\theta_{\mathcal{M}}, \bar{\theta}_u, \theta_e)$ and $\bar{\mathcal{L}}(\bar{\theta}) = \mathcal{L}(\bar{\theta})$. $\bar{\mathcal{L}}(\bar{\theta})$ is the quasi-likelihood function in the sense that it represents the original likelihood function conditioned upon $D = 0$, which means that it neglects shocks' skewness[7]. In the first step a maximizer:

$$\hat{\bar{\theta}} = (\hat{\theta}_{\mathcal{M}}, \hat{\bar{\theta}}_u, \hat{\theta}_e) = \underset{\bar{\theta} \in \bar{\Theta}}{\arg\max} \left\{\bar{\mathcal{L}}\left(\bar{\theta}\right)\right\} \tag{3.7}$$

[6] In a general case γ_{max} is an increasing function of ϑ_i and a decreasing function of $\tilde{\delta}_i$, $(i = 1, ..., q)$, see the definition 1.1.2 of csn probability function.

[7] Notice, that in the case under consideration $\tilde{\Sigma}_u = \Sigma_u$, hence the variances of shocks are estimated.

is found. With $D = 0$, this is a standard maximum likelihood estimation of a state space model with normally distributed shocks. Then, shocks u_t are filtered using model (3.2) with parameters $\hat{\bar{\theta}}$ plugged in it[8], and sample estimates of shocks' skewness coefficients (method of moment estimators) $\hat{\gamma}_{u_i}$ for $i = 1, 2, ..., p_u$ are established. If only skewness coefficients are of interest, then the procedure ends yielding $\hat{\gamma}_u = (\hat{\gamma}_{u_i},\ i = 1, 2, ..., p_u)$. Otherwise, original shocks parameters $\tilde{\Sigma}_u$ and D_u are recovered from $\hat{\gamma}_u$ and $\hat{\Sigma}_u$ according to equations (3.3–3.6), which results in estimates of $\tilde{\Sigma}_u$ and D_u respectively, and final estimate of θ becomes $\hat{\hat{\theta}} = (\hat{\theta}_{\mathscr{M}}, \hat{\theta}_u, \hat{\theta}_e)$ where $\hat{\theta}_u = \left(\hat{\tilde{\Sigma}}_u(\hat{\gamma}_u, \hat{\Sigma}_u),\ \hat{D}_u(\hat{\gamma}_u, \hat{\Sigma}_u)\right)$.

3.2 The procedure of stochastic simulations

To asses properties of our two-steps quasi-maximum likelihood/method of moments estimator we run several stochastic simulation experiments. A single iteration of our stochastic simulation procedure looks as follows:

1. A sample of shocks u_t and measurement errors e_t are simulated. States ξ_t and observables y_t are computed according to (3.2).
2. Given observables from step 1, a Newton-type optimization routine is applied to find $\bar{\theta}$, i.e. the maximizer of the quasi likelihood function $\mathscr{L}(\bar{\theta})$ (the first step of the estimation procedure described earlier). If optimization fails to converge, steps 3–4 are skipped and estimation results are discarded[9]. In this situation a new iteration is initiated.
3. Given $\hat{\bar{\theta}}$, i.e. parameters obtained in step 2, states, observables and shocks $\hat{u}_t$ are filtered using the Kalman smoother.
4. Smoothed shocks $\hat{u}_t$ are used to investigate some characteristic of skewness estimators and tests for skewness. In particular we analyze properties of adjusted sample skewness coefficients estimator as well as and size and power of tests for skewness of shocks.

All parameters, except for shocks' parameters, i.e. $\theta_{\mathscr{M}}$, θ_e as well as matrix of second moments of shocks Σ_u (a part of θ_u), are common for all simulation trials, see Table2.1. Table 3.1 reports shocks' skewness parameters D_u (a component of θ_u) and equivalent skewness coefficients γ_{u_i} given d_{u_i} and σ_{u_i}. We use these parameters to generate three variants of data in this experimental setup. These variants are: normal shocks variant, which is our benchmark, moderate skewness of all shocks (CSN-1), and strong skewness of all shocks (CSN-2).

Random number generator for the skewed normal distribution follows Gupta et al. (2004, Prop. 2.5, p. 184), see also Roch and Valdez (2009); Dunajeva et al. (2003); González-Farías et al. (2004); Iversen (2010). The length of samples varies from 75 („small sample") up to 600 („large sample"). For each case of given length over 2000 replications were generated.

[8] In fact, filtration of shocks is a byproduct of estimation of $\bar{\theta}$ using the Kalman filter. From now on we will denote filtered (smoothed) shocks by $\hat{u}$.

[9] The number of rejected trials varied with sample size. It was up to 40% for samples of small sample, and just a few for large samples.

Table 3.1. Simulations specific skewness parameters of shocks

Variant	$d_{u(z)}$	$\gamma_{u(z)}$	$d_{u(\Delta q)}$	$\gamma_{u(\Delta q)}$	$d_{u(y^\star)}$	$\gamma_{u(y^\star)}$	$d_{u(\pi^\star)}$	$\gamma_{u(\pi^\star)}$	$d_{u(R)}$	$\gamma_{u(R)}$
Normal	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
CSN-1	1.498	0.50	0.788	0.50	0.756	0.50	0.499	0.50	3.744	0.50
CSN-2	5.688	0.95	2.994	0.95	2.873	0.95	1.896	0.95	14.20	0.95

$\gamma_{u(i)}$ — skewness coefficient of $u(i)$; $d_{u(i)}$ — *csn* distribution parameter of shock $u(i)$.

3.3 Quasi-maximum likelihood estimator of models' parameters

The selected results of stochastic simulations are presented in Table 3.2[10]. Table 3.2 reports relative percentage biases[11] and standard deviations of models parameter estimates $\widehat{\theta}_{\mathcal{M}}$ obtained in the second step of the simulation procedure (the first step of the estimation procedure). The general point is that results obtained for the normal case (the first row, the ML estimator) and for variants CSN1–CSN2 (rows 2–3, the Q-ML estimator) do not differ substantially, although shocks skewness is neglected during the estimation in *csn* variants. Bias of the Q-ML procedure in short sample is considerable, but this is also the case for the ML estimator. There is likely an identification problem for interest rate rule parameters (ψ_π, ψ_y $\psi_{\Delta e}$) as well as for r_A and $\sigma_{y^\star}$[12]. The magnitude of Q-ML estimators' bias and ML estimators' bias is similar. However, ML estimators are often slightly more precise (taking into account their standard deviations). The biases as well as the standard deviations of estimators are (approximatively) declining functions of sample size. This means that our *ad hoc* Q-ML estimators have properties of consistent estimators, at least in the problem at hand.

[10] Detailed results are available from the authors upon request.

[11] The relative bias is defined as: $100\frac{\widehat{\theta}-\theta}{\theta}$.

[12] It might be seen as a support for Cochrane (2007) thesis, who noticed that parameters of the Taylor rule in a simple new-Keynesian model of economy are unidentified (the model specification issue), but if J.H. Cochrane is right, sample size should not matter. The results of our exercise indicate however, that we likely faced a data related issue.

Table 3.2. Summary of stochastic simulation exercises. Properties of (Q)-LM estimators of the basic set of parameters. 2000 replications.

Simul. Variant	Estim. Feature	$\widehat{\psi}_\pi$	$\widehat{\psi}_y$	$\widehat{\psi}_{\Delta e}$	$\widehat{\rho}_R$	$\widehat{\alpha}$	$\widehat{\kappa}$	$\widehat{\tau}$	$\widehat{r}_A$	$\widehat{\gamma}_Q$	$\widehat{\rho}_{\hat{z}}$	$\widehat{\rho}_{\hat{d}q}$	$\widehat{\rho}_{y^\star}$	$\widehat{\rho}_{\pi^\star}$	$\widehat{\sigma}_{\hat{z}}$	$\widehat{\sigma}_{\hat{d}q}$	$\widehat{\sigma}_{y^\star}$	$\widehat{\sigma}_{\pi^\star}$	$\widehat{\sigma}_{\epsilon_R}$
		Estimator of parameter																	
Sample size = 75																			
Normal	%Bias	32.09	87.50	56.44	2.462	1.407	6.891	-1.912	23.49	-4.276	3.530	-2.593	-2.188	-0.786	-4.970	-0.657	61.11	-1.017	3.816
Normal	Std	1.215	0.532	0.136	0.120	0.029	0.174	0.202	0.562	0.127	0.047	0.102	0.059	0.053	0.113	0.161	3.376	0.253	0.096
CSN-1	%Bias	36.14	93.96	62.88	2.797	1.708	7.383	-2.613	24.57	-4.351	3.944	-1.842	-1.965	-0.834	-5.669	-1.098	60.16	-0.418	4.659
CSN-1	Std	1.290	0.551	0.148	0.124	0.029	0.172	0.200	0.563	0.123	0.048	0.103	0.059	0.052	0.119	0.172	3.446	0.269	0.121
CSN-2	%Bias	32.33	86.84	57.45	2.567	1.577	6.856	-2.928	22.75	-4.351	3.825	-2.707	-2.031	-0.974	-5.916	-1.138	53.64	-0.949	3.536
CSR-2	Std	1.223	0.509	0.137	0.121	0.028	0.169	0.195	0.565	0.122	0.047	0.101	0.060	0.053	0.123	0.185	3.265	0.289	0.075
Sample size = 150																			
Normal	%Bias	15.72	39.72	27.51	1.463	0.769	2.522	0.443	9.338	-1.723	3.750	-0.637	-0.976	-0.580	-2.546	-0.470	40.06	-0.217	1.561
Normal	Std	0.746	0.290	0.080	0.093	0.021	0.112	0.155	0.438	0.094	0.035	0.071	0.039	0.036	0.080	0.111	2.423	0.180	0.057
CSN-1	%Bias	15.51	39.80	27.21	0.969	0.539	2.885	0.767	8.598	-1.769	3.304	-1.394	-0.865	-0.554	-3.187	-0.177	45.29	-0.466	1.224
CSN-1	Std	0.831	0.344	0.091	0.092	0.022	0.113	0.159	0.427	0.091	0.034	0.072	0.038	0.037	0.082	0.120	2.598	0.193	0.050
CSN-2	%Bias	13.72	37.51	24.14	0.346	0.007	3.752	0.907	9.116	-1.893	2.570	-1.028	-1.190	-0.549	-3.247	-0.555	44.73	-0.503	1.264
CSN-2	Std	0.745	0.298	0.081	0.094	0.022	0.116	0.159	0.412	0.090	0.034	0.071	0.038	0.037	0.083	0.129	2.555	0.214	0.054
Sample size = 600																			
Normal	%Bias	2.990	7.635	5.516	0.145	-0.043	0.546	0.276	1.511	-0.190	1.305	-0.073	-0.265	-0.237	-0.675	-0.097	11.96	-0.136	0.227
Normal	Std	0.264	0.103	0.029	0.049	0.011	0.055	0.088	0.239	0.051	0.020	0.035	0.017	0.018	0.041	0.055	0.999	0.088	0.024
CSN-1	%Bias	2.859	6.533	4.820	0.056	-0.019	0.635	1.269	-0.026	0.054	1.765	-0.073	-0.196	-0.084	-0.754	-0.061	13.89	-0.157	0.003
CSN-1	Std	0.291	0.113	0.032	0.050	0.012	0.055	0.085	0.239	0.051	0.019	0.037	0.017	0.018	0.042	0.059	0.988	0.096	0.026
CSN-2	%Bias	2.680	6.525	4.537	-0.141	0.047	1.080	0.684	0.515	-0.171	1.303	-0.549	-0.254	-0.180	-0.808	-0.189	11.96	-0.127	0.332
CSN-2	Std	0.313	0.114	0.032	0.048	0.011	0.058	0.084	0.240	0.051	0.019	0.035	0.018	0.019	0.044	0.065	0.978	0.103	0.054

$\gamma_{u(i)}$ — skewness coefficient of $u(i)$; $d_{u(i)}$ — csn distribution parameter of shock $u(i)$.

3

3.4 Method of moments estimator of shocks' skewness

To estimate skewness one may use method of a moment estimator — the sample skewness coefficient $\widehat{\gamma}$ defined as:

$$\widehat{\gamma}(Z) = \frac{\frac{1}{T}\sum_{t=1}^{T}\left(Z_t - \overline{Z}\right)^3}{\left(\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Z_t - \overline{Z}\right)^2}\right)^3}, \tag{3.8}$$

where $\overline{Z}$ is the sample mean. Bai and Ng (2005, p. 55) noticed that skewness measured by the sample skewness coefficient is usually underestimated. That observation agrees with our findings. Taking it into account data presented in Table A.6 (see Appendix), one can conclude that, when shocks are skew-normal, a variant of sample skewness coefficient ($\widehat{\gamma}$) is biased (skewness is underestimated) in a limited sample[13]. The range of skewness coefficient is the second problem worth noting. The skewness coefficient of closed skew-normal shocks is limited, it must satisfy the condition: $|\gamma_i| < \gamma_{max} \approx 0.995$. The sample skewness coefficient could be arbitrary large. Hence the sample coefficient of skewness should be simultaneously rescaled into proper range and scaled up to minimize bias, but these two transformations are contradictory. We checked several propositions of such adjustments and chose one that remains asymptotic properties of the sample estimator of skewness coefficient. We find that it is reasonable to treat separately irregular cases where the coefficient is outside the admissible range. Our adjusted sample skewness coefficient estimator is defined as follows:

$$\widehat{\widehat{\gamma}}_{T,i} = \begin{cases} \tilde{\gamma}_i & \text{if} \quad |\tilde{\gamma}_i| \leq \gamma_{max} \\ \operatorname{sig}(\tilde{\gamma}_i)\left[\gamma_0(T) + \left(\gamma_{max} - \gamma_0(T)\right)\operatorname{erf}\left(|\tilde{\gamma}_i|\right)\right] & \text{otherwise} \end{cases} \tag{3.9}$$

where: T — is the sample size, $\gamma_{max} = 0.995$ and the functions $\tilde{\gamma}_i$ and $\gamma_0(T)$ are defined as:

$$\tilde{\gamma}_i \equiv \tilde{\gamma}_i(Z) = \frac{T}{T-3}\frac{\sqrt{T(T-1)}}{T-2}\widehat{\gamma}_i(Z), \qquad \text{for} \quad T > 10 \tag{3.10}$$

$$\gamma_0(T) = \max\left\{0.45,\ \frac{2}{\pi}\left[0.5 + \operatorname{erf}\left(1 - \sqrt{3}\frac{\pi}{\sqrt{T}}\right)\right]\right\} \tag{3.11}$$

The results of stochastic simulations presented in Table A.6 (see Appendix) allow to assess its main properties when the estimator is applied to simulated shocks u_t. The adjusted estimator of skewness coefficient is not very precise. It still underestimates skewness if the sample size T is small or even moderate. The bias is, however, a decreasing function of sample, the variance of the estimator is a decreasing function of sample size as well. Even when the „true" coefficient of skewness γ is close to the bound $|0.995|$, the bias declines but very slowly, so it behaves as a consistent estimator.

Table 3.3 shows properties of the estimator applied to smoothed shocks $\widehat{u}_t$. In this case the bias seems to be a declining function of sample so the estimator behaves as asymptotically

[13] Dunajeva et al. (2003) derived an approximate formula for the bias.

unbiased as well. Nevertheless the fall of bias is slow and even in the large sample the estimates of skewness for the „difficult" shocks, $u(y^{\star})$ and $u(z)$ are very imprecise. However, this is an outcome of imprecise estimation of variances of these shocks (the first step of the estimation procedure). We suspect a model and/or data related issues, e.g. an identification problem, because that phenomenon occurs also for normal shocks and ML estimation. Imprecise estimates of shocks' variances distort filtering and the estimated (smoothed) shocks $\widehat{u}_t$ are imprecise approximation of "true" shocks u_t (see Table A.4 in Appendix).

Table 3.3. Properties of the adjusted sample skewness estimator $\widehat{\gamma}(\widehat{u})$. 2000 replications

Shocks	Bias	Bias %	Mean	Mode	Median	St.Dev	Skewn.	Kurtosis	5%	95%
CSN-1	Sample size = 75+7									
$u(z)$	-0.178	-35.641	0.322	0.319	0.319	0.291	0.070	2.741	-0.151	0.846
$u(\Delta q)$	-0.025	-5.058	0.475	0.471	0.471	0.278	-0.086	2.447	0.009	0.947
$u(y^{\star})$	-0.270	-53.956	0.230	0.228	0.228	0.290	-0.031	3.058	-0.238	0.715
$u(\pi^{\star})$	-0.051	-10.234	0.449	0.449	0.449	0.268	-0.002	2.450	0.011	0.933
$u(\epsilon^R)$	-0.077	-15.498	0.422	0.426	0.426	0.284	-0.122	2.718	-0.048	0.931
CSN-1	Sample size = 150+7									
$u(z)$	-0.165	-32.916	0.336	0.332	0.277	0.210	0.120	3.027	-0.006	0.691
$u(\Delta q)$	-0.023	-4.567	0.478	0.472	0.395	0.207	0.062	2.772	0.149	0.837
$u(y^{\star})$	-0.241	-48.149	0.259	0.254	0.136	0.215	0.203	3.210	-0.084	0.615
$u(\pi^{\star})$	-0.037	-7.393	0.463	0.449	0.405	0.210	0.143	2.932	0.129	0.840
$u(\epsilon^R)$	-0.063	-12.565	0.437	0.431	0.431	0.215	0.137	2.919	0.103	0.820
CSN-1	Sample size = 600+7									
$u(z)$	-0.140	-27.901	0.361	0.361	0.294	0.110	0.069	3.075	0.182	0.548
$u(\Delta q)$	-0.010	-2.080	0.490	0.487	0.458	0.111	0.165	3.299	0.313	0.673
$u(y^{\star})$	-0.219	-43.796	0.281	0.280	0.244	0.106	0.028	3.090	0.110	0.464
$u(\pi^{\star})$	-0.027	-5.366	0.473	0.466	0.466	0.110	0.316	3.178	0.308	0.666
$u(\epsilon^R)$	-0.046	-9.288	0.453	0.448	0.361	0.111	0.199	3.365	0.271	0.641
CSN-2	Sample size = 75+7									
$u(z)$	-0.350	-36.797	0.600	0.623	0.935	0.266	-0.496	2.776	0.126	0.966
$u(\Delta q)$	-0.137	-14.377	0.813	0.882	0.933	0.174	-1.102	3.497	0.470	0.984
$u(y^{\star})$	-0.524	-55.209	0.426	0.422	0.311	0.288	-0.120	2.690	-0.042	0.935
$u(\pi^{\star})$	-0.142	-14.980	0.808	0.878	0.945	0.179	-1.029	3.303	0.465	0.982
$u(\epsilon^R)$	-0.201	-21.144	0.749	0.799	0.939	0.212	-0.840	3.030	0.356	0.978
CSN-2	Sample size = 150+7									
$u(z)$	-0.317	-33.384	0.633	0.636	0.636	0.197	-0.092	2.393	0.309	0.957
$u(\Delta q)$	-0.097	-10.258	0.853	0.906	0.947	0.132	-1.074	3.396	0.594	0.980
$u(y^{\star})$	-0.462	-48.672	0.488	0.481	0.392	0.215	0.076	2.810	0.140	0.881
$u(\pi^{\star})$	-0.113	-11.901	0.837	0.876	0.957	0.139	-0.923	3.132	0.568	0.980
$u(\epsilon^R)$	-0.152	-16.049	0.798	0.828	0.950	0.161	-0.744	2.832	0.495	0.978
CSN-2	Sample size = 600+7									
$u(z)$	-0.268	-28.210	0.682	0.678	0.597	0.111	0.140	3.127	0.505	0.874
$u(\Delta q)$	-0.049	-5.106	0.901	0.922	0.961	0.075	-0.978	3.438	0.761	0.979
$u(y^{\star})$	-0.412	-43.383	0.538	0.534	0.479	0.115	0.186	3.072	0.356	0.735
$u(\pi^{\star})$	-0.066	-6.926	0.884	0.902	0.963	0.085	-0.771	2.861	0.725	0.978
$u(\epsilon^R)$	-0.103	-10.811	0.847	0.857	0.962	0.096	-0.506	2.619	0.679	0.972

$\widehat{u}$ — estimated (smoothed) shocks

3.5 Tests for skewness of smoothed shocks

The final step of the estimation procedure involves testing for skewness of filtered (smoothed) shocks. We employ significance test of shocks' skewness coefficients (one-tailed)[14], the test based on adjusted sample skewness coefficient $\widehat{\gamma}$ as well as two parametric tests developed

[14]In general we prefer one-tailed tests, because of theirs higher power. The direction of skewness may be investigated having estimates of skewness coefficient.

by Bai and Ng (2005). We verify properties of these tests, since to our best knowledge their sampling distributions have not been established for the smoothed variables[15]. Given asymptotic distribution of $\widehat{\gamma}$ (under null hypothesis of normality[16] /symmetry/) $\sqrt{T}\,\widehat{\gamma} \xrightarrow{d} N(0,6)$, it is easy to notice, that (under null) $\sqrt{T}\,\tilde{\gamma} \xrightarrow{d} N(0,6)$ as well. However, since the number of irregular cases (under null) declines when the sample size grows, therefore one concludes that $\sqrt{T}\,\widehat{\widehat{\gamma}} \xrightarrow{d} N(0,6)$.

Table 3.4 reports the rejections ratios for some skewness tests computed for 10% critical values[17]. We verify size and power of the tests for simulated ("true") shocks and estimated (smoothed) shocks. The collected data indicate, that the rejection ratio of true hypothesis (symmetric shocks) is approximately 8–11%. The distortion created by LM estimator and two-sided Kalman filter (smoother) is quite moderate in the case of normal shocks — the size of tests is similar for simulated and estimated (smoothed) shocks. Only, the Bai-Ng joint test (χ^2 test) reject slightly less frequently. This is a feature of this test however. In the case of skew-normal shocks, the rejection ratio is sensitive to sample type. The power of tests is lower for estimated shocks. The loss of test power is differentiated. There are two „difficult" shocks, $u(y^\star)$ and $u(z)$, where the decline is considerable and two shocks where the decline is small. From the other hand, the power of test in small sample is rather low, especially when skewness is moderate (variant CSN-1). Hence, it may be difficult to identify skewness if the sample is small and/or skewness is small/moderate. Nevertheless, even in the worst case (small sample, moderate skewness and „difficult" shocks) one correctly detects skewness in 34% cases (and 60–67% for regular shocks) using out test based on $\widehat{\widehat{\gamma}}$. Bai-Ng π_3 skewness test gives very similar results.

The data shown the Table suggest that one-tailed Bai-Ng π_3 as well as one-tailed sample skewness coefficients $\widehat{\gamma}, \widehat{\widehat{\gamma}}$ perform similar. Our test based on adjusted sample skewness coefficient $\widehat{\widehat{\gamma}}$ has slightly hight power (and more precise size) but it could be a result of sampling noise[18].

[15] Bai and Ng (2005) skewness tests are valid also for likely serially correlated disturbances of the linear regression model — they proved asymptotic equivalence of test based on disturbances and estimated regression residuals.

[16] Compare (Bai and Ng, 2005). In general, under null of symmetry, asymptotic variance of sample skewness coefficient $\widehat{\gamma}$ is given by the following formula:

$$\mathrm{var}\left(\widehat{\gamma}\right) = \frac{\widehat{\mu}_6 - 6\,\widehat{\mu}_2\,\widehat{\mu}_4 + 9\,\widehat{\mu}_2^3}{N\,\widehat{\mu}_2^3}$$

provided that population central moments μ_i $(t = 1, ..., 6)$ exist. The formula is valid for any symmetric distribution and iid samples, see Gupta (1967) for details.

[17] Note, that the alternative hypothesis of these tests are not the same. The alternative for the test based on $\widehat{\widehat{\gamma}}$ indicates constrained value of the skewness coefficient.

[18] In addition, we run several stochastic simulation exercises using the triples test (a non-parametric test for asymmetry designed by Randles et al. (1980)) and normal as well as skew-normal shocks. The results suggest that the power of the triples test is slightly lower that the power of our adjusted sample skewness coefficient test applied to simulated data.

Table 3.4. Size and power /rejection ratio/ of skewness rests, 2000 replications.

Shock	Var.	1-tailed $\widehat{\gamma}$					1-tailed $\widehat{\widehat{\gamma}}$					1-tailed Bai-Ng π_3					Bai-Ng χ^2				
Type	Type	$u(z)$	$u(\Delta q)$	$u(y^*)$	$u(\pi^*)$	$u(\epsilon^R)$	$u(z)$	$u(\Delta q)$	$u(y^*)$	$u(\pi^*)$	$u(\epsilon^R)$	$u(z)$	$u(\Delta q)$	$u(y^*)$	$u(\pi^*)$	$u(\epsilon^R)$	$u(z)$	$u(\Delta q)$	$u(y^*)$	$u(\pi^*)$	$u(\epsilon^R)$
Sample size = 75+7																					
Norm.	sim	0.08	0.09	0.09	0.08	0.08	0.09	0.10	0.10	0.09	0.09	0.10	0.11	0.11	0.10	0.11	0.06	0.07	0.07	0.06	0.06
	est	0.08	0.10	0.09	0.08	0.09	0.09	0.11	0.10	0.09	0.10	0.10	0.10	0.11	0.10	0.10	0.06	0.07	0.06	0.06	0.06
CSN-1	sim	0.65	0.65	0.62	0.64	0.65	0.67	0.68	0.65	0.67	0.68	0.66	0.68	0.65	0.68	0.68	0.30	0.30	0.30	0.30	0.29
	est	0.43	0.64	0.32	0.62	0.57	0.46	0.67	0.34	0.63	0.60	0.44	0.67	0.33	0.63	0.59	0.14	0.28	0.09	0.26	0.22
CSN-2	sim	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.93	0.92	0.90	0.92	0.90
	est	0.80	0.98	0.58	0.98	0.94	0.83	0.98	0.60	0.98	0.95	0.82	0.99	0.60	0.98	0.96	0.47	0.89	0.24	0.88	0.76
Sample size = 150+7																					
Norm.	sim	0.09	0.10	0.09	0.09	0.09	0.10	0.10	0.10	0.09	0.10	0.10	0.11	0.11	0.10	0.11	0.07	0.06	0.06	0.07	0.07
	est	0.09	0.10	0.11	0.09	0.10	0.10	0.10	0.11	0.09	0.11	0.10	0.10	0.12	0.10	0.11	0.07	0.07	0.07	0.06	0.07
CSN-1	sim	0.88	0.88	0.87	0.87	0.86	0.88	0.88	0.88	0.88	0.87	0.89	0.88	0.89	0.89	0.88	0.60	0.61	0.59	0.59	0.59
	est	0.64	0.85	0.49	0.84	0.79	0.65	0.86	0.50	0.85	0.80	0.65	0.87	0.49	0.85	0.81	0.30	0.58	0.21	0.54	0.48
CSN-2	sim	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00
	est	0.98	1.00	0.86	1.00	1.00	0.98	1.00	0.87	1.00	1.00	0.98	1.00	0.87	1.00	1.00	0.86	1.00	0.59	1.00	0.99
Sample size = 600+7																					
Norm.	sim	0.10	0.10	0.09	0.10	0.09	0.10	0.10	0.09	0.10	0.10	0.11	0.10	0.10	0.11	0.10	0.07	0.09	0.08	0.07	0.08
	est	0.10	0.10	0.09	0.09	0.10	0.10	0.10	0.09	0.09	0.10	0.11	0.10	0.10	0.10	0.10	0.08	0.09	0.07	0.08	0.08
CSN-1	sim	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	est	0.98	1.00	0.93	1.00	1.00	0.98	1.00	0.93	1.00	1.00	0.98	1.00	0.92	1.00	1.00	0.92	1.00	0.73	0.99	0.99
CSN-2	sim	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	est	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

The 10% critical values are 1.28 (one-tailed), and 4.61; Bai-Ng tests: Newley-West kernel, no prewithening.

Source: Prepared by the authors

3

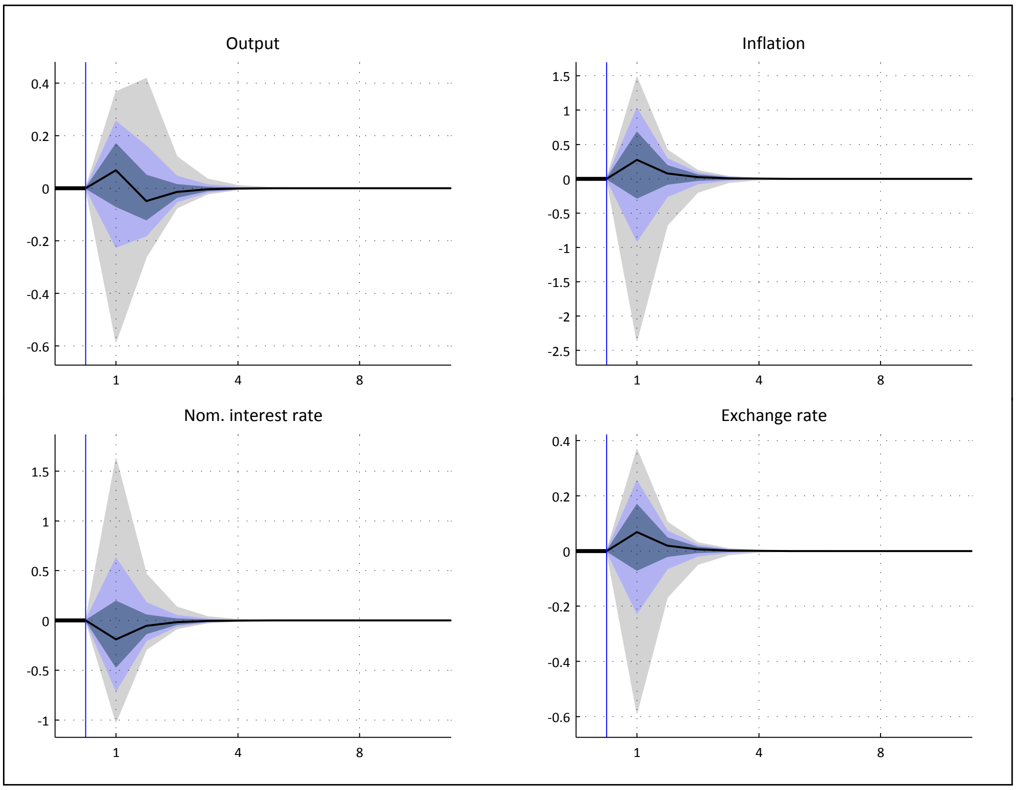
Chapter 4

Impulse response functions

4

For shocks which follow skewed distribution instead of typically assumed symmetric distribution, it may be more meaningful not to depict impulse response function to a single impulse (e.g. one standard deviation) but to create confidence intervals based on uncertainty of a shock. Figure 4.1 shows 30, 60, and 90% confidence intervals, as well as median, of response to a monetary policy shock. Confidence intervals include only shock uncertainty with parameters kept fixed. The exact procedure looks as follows. Based on a sample of smoothed shocks obtained after applying Kalman smoother, we calculate desired empirical percentiles of shocks distribution. Then, we obtain impulse responses to shocks equal to each of them. Monetary policy shock in this exercise is assumed to be positively skewed hence positive skewness of nominal interest rate. It means that nominal interest rate in response to a shock is more often below zero than above, but the probability of very high interest rate is higher than the probability of very low interest rate. If so, inflation, output and exchange rate will be negatively skewed.

Figure 4.1. Monetary policy shock



Chapter 5

Concluding remarks

Skewness of observables is not accounted for in the domain of DSGE models, at least when first order perturbations are employed, which is quite often the case both in the literature as well as in practice. On the other hand, most important macroeconomic time series – notably output growth, inflation and interest rates – reveal skewness. This paper attempts to fill this gap.

In the paper we stressed the fact, that skewness in observed variables can be a result of skewness in structural shocks. In fact, in a linear (or a linearized) DSGE model there is no other way to get skewed observables. Propagation of skewness in liner state-space models undergoes certain laws, e.g. skewness of states of univariate autoregressions decreases with time reaching a zero or non-zero limit for random walks and stationary specifications respectively.

Simulation exercises indicate that a simple two-step quasi-maximum likelihood/method of moments parameters' estimation procedure, which neglects shocks' skewness in the first step, does not distort estimates of models' parameter, at least for the problem at hand. This allows us to filter shocks, given parameters, and then estimate shocks' skewness parameters. Properties of skewness tests of filtered shocks are far less satisfactory. Then, quality of estimates of shocks skewness parameters seems to be shock dependent.

Appendix A

Tests for skewness and results of stochastic simulations

A.1 Tests for skewness

Distribution of sample skewness coefficient

For a series $\{X_t\}_{t=1}^{T}$ with mean μ and variance σ we defined its r-th central moment by $\mu_r = \mathrm{E}\left[(x-\mu)^r\right]$, the coefficient of skewness is defined as (see also equation (1.6)):

$$\gamma = \frac{\mu_3}{\sigma^3} = \frac{\mathrm{E}\left[(x-\mu)^3\right]}{\mathrm{E}\left[(x-\mu)^2\right]^{\frac{3}{2}}}. \tag{A.1}$$

The sample estimate of skewness coefficient is given by (compare equation (3.8));

$$\widehat{\gamma} = \frac{\frac{1}{T}\sum_{t=1}^{T}\left(X_t - \overline{X}\right)^3}{\left(\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(X_t - \overline{X}\right)^2}\right)^3}. \tag{A.2}$$

where $\overline{X}$ is the sample mean. If X_t is *iid* and normally distributed then $\sqrt{T}\,\widehat{\gamma} \xrightarrow{d} N(0,6)$.

Bai-Ng tests for skewness

Below we present tests for skewness proposed by Bai and Ng (2005). If X_t is weakly dependent and stationary up to sixth order, under the null hypothesis that $\gamma = 0$

$$\sqrt{T}\widehat{\gamma} = \frac{\boldsymbol{\alpha}}{\widehat{\sigma}^3}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{Z}_t + o_p(1) \tag{A.3}$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 & -3\sigma^2 \end{bmatrix}, \qquad \mathbf{Z}_t = \begin{bmatrix} (X_t-\mu)^3 \\ (X_t-\mu) \end{bmatrix}, \tag{A.4}$$

and

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{Z}_t \xrightarrow{d} N\left(\mathbf{0},\boldsymbol{\Gamma}\right), \qquad \boldsymbol{\Gamma} = \lim_{T\to\infty} TE\left(\overline{\mathbf{Z}\mathbf{Z}}^T\right) \tag{A.5}$$

with $\overline{\mathbf{Z}}$ being the sample mean of $\mathbf{Z}_t$ and Γ is the spectral density matrix at frequency 0 of $\mathbf{Z}_t$. Additionally

$$\sqrt{T}\widehat{\gamma} \xrightarrow{d} N\left(0, \frac{\boldsymbol{\alpha}\boldsymbol{\Gamma}\boldsymbol{\alpha}^T}{\sigma^6}\right) \qquad \text{or} \qquad \sqrt{T}\widehat{\mu}_3 \xrightarrow{d} N\left(0, \boldsymbol{\alpha}\boldsymbol{\Gamma}\boldsymbol{\alpha}^T\right) \tag{A.6}$$

Let $\widehat{\sigma}^2$ and $\widehat{\boldsymbol{\Gamma}}$ be consistent estimates of σ^2 and $\boldsymbol{\Gamma}$. Let $\widehat{\boldsymbol{\alpha}} = \left[1, -3\widehat{\sigma}^2\right]$, $s\left(\widehat{\mu}_3\right) = \left(\widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\Gamma}}\widehat{\boldsymbol{\alpha}}^T\right)^{1/2}$ and $s\left(\widehat{\gamma}\right) = \left(\widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\Gamma}}\widehat{\boldsymbol{\alpha}}^T/\widehat{\sigma}^6\right)^{1/2}$. Under the null hypothesis that $\gamma = 0$

$$\widehat{\pi}_3 = \frac{\sqrt{T}\widehat{\mu}_3}{s\left(\widehat{\mu}_3\right)} = \frac{\sqrt{T}\widehat{\gamma}}{s\left(\widehat{\theta}\right)} \xrightarrow{d} N(0,1). \tag{A.7}$$

Long-run variance matrix can be obtained nonparametrically by kernel estimation, e.g. the Bartlett kernel (see Newey and West (1987)).

Possible low power of the test can be increased by applying either a one-tailed test (direction of skewness is usually suspected) or a joint test of two odd moments, r_1 and r_2. Let

$$\mathbf{Y}_t = \begin{bmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(X_t-\mu\right)^{r_1} \\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(X_t-\mu\right)^{r_2} \end{bmatrix} \tag{A.8}$$

It can be shown that

$$\mathbf{Y}_t = \boldsymbol{\alpha}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{Z}_t + o_p(1) \tag{A.9}$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 & 0 & -r_1\mu_{r_1-1} \\ 0 & 1 & -r_2\mu_{r_2-1} \end{bmatrix}, \qquad \mathbf{Z}_t = \begin{bmatrix} \left(X_t-\mu\right)^{r_1} \\ \left(X_t-\mu\right)^{r_2} \\ \left(X_t-\mu\right) \end{bmatrix}. \tag{A.10}$$

Under the null hypothesis of symmetry, if

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{Z}_t \xrightarrow{d} N\left(\mathbf{0},\boldsymbol{\Gamma}\right), \qquad \boldsymbol{\Gamma} = \lim_{T\to\infty} TE\left(\overline{\mathbf{Z}}\overline{\mathbf{Z}}^T\right) \tag{A.11}$$

then $\mathbf{Y}_T \xrightarrow{d} N\left(\mathbf{0}, \boldsymbol{\alpha}\boldsymbol{\Gamma}\boldsymbol{\alpha}^T\right)$. Let $\widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\Gamma}}\widehat{\boldsymbol{\alpha}}^T$ be a consistent estimate of $\boldsymbol{\alpha}\boldsymbol{\Gamma}\boldsymbol{\alpha}^T$. Then

$$\widehat{\mu}_{r_1,r_2} = \mathbf{Y}_T^T\left(\widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\Gamma}}\widehat{\boldsymbol{\alpha}}^T\right)^{-1}\mathbf{Y}_T \xrightarrow{d} \chi_2^2 \tag{A.12}$$

We refer to the above joint test as Bai-Ng χ^2 test. The Gauss code of the test was downloaded from Serena Ng's webpage (http://www.columbia.edu/~sn2294/).

A.2 Results of stochastic simulations

A.2.1 Estimation of the basic set of parameters

Table A.1. Normal shocks — Kalman filter ML estimators, 2000 replications=2000

Parameter	Bias	Bias %	Mean	Mode	Median	St.Dev	Skewn.	Kurtosis	5%	95%
Sample size = 75+7										
ψ_π	0.481	32.09	1.981	1.604	1.208	1.215	2.895	13.47	1.022	4.552
ψ_y	0.219	87.50	0.469	0.295	0.214	0.532	3.375	18.19	0.075	1.518
$\psi_{\Delta e}$	0.056	56.44	0.156	0.115	0.073	0.136	3.030	14.89	0.048	0.411
ρ_R	0.015	2.462	0.615	0.613	0.615	0.120	0.016	2.884	0.420	0.819
α	0.002	1.407	0.152	0.150	0.151	0.029	0.256	3.027	0.107	0.201
κ	0.034	6.891	0.534	0.511	0.432	0.174	1.202	6.239	0.299	0.842
τ	-0.010	-1.912	0.490	0.486	0.382	0.202	-0.065	2.316	0.146	0.825
r_A	0.176	23.49	0.926	0.861	0.614	0.562	0.637	3.131	0.135	1.955
γ_Q	-0.034	-4.276	0.766	0.774	0.789	0.127	-0.383	3.048	0.535	0.959
$\rho_{\hat{z}}$	0.007	3.530	0.207	0.206	0.207	0.047	0.751	7.872	0.135	0.285
$\rho_{\Delta\hat{q}}$	-0.010	-2.593	0.390	0.397	0.434	0.102	-0.358	3.284	0.214	0.552
$\rho_{\hat{y}^\star}$	-0.020	-2.188	0.880	0.888	0.926	0.059	-0.918	4.279	0.770	0.961
$\rho_{\hat{\pi}^\star}$	-0.006	-0.786	0.794	0.797	0.793	0.053	-0.445	3.210	0.698	0.875
$\sigma_{\hat{z}}$	-0.050	-4.970	0.950	0.951	0.951	0.113	-0.216	3.897	0.771	1.135
$\sigma_{\Delta\hat{q}}$	-0.012	-0.657	1.888	1.886	1.815	0.161	0.126	3.074	1.623	2.149
$\sigma_{\hat{y}^\star}$	1.155	61.11	3.045	1.830	0.818	3.376	2.188	8.164	0.233	10.847
$\sigma_{\hat{\pi}^\star}$	-0.031	-1.017	2.969	2.964	2.961	0.253	0.058	3.035	2.562	3.388
σ_{ϵ^R}	0.015	3.816	0.415	0.405	0.390	0.096	13.21	379.4	0.315	0.549
Sample size = 150+7										
ψ_π	0.236	15.72	1.736	1.538	1.285	0.746	3.298	23.58	1.033	3.024
ψ_y	0.099	39.72	0.349	0.263	0.201	0.290	2.948	17.80	0.083	0.902
$\psi_{\Delta e}$	0.028	27.51	0.128	0.106	0.079	0.080	2.858	17.00	0.051	0.273
ρ_R	0.009	1.463	0.609	0.603	0.597	0.093	0.192	2.846	0.466	0.767
α	0.001	0.769	0.151	0.150	0.149	0.021	0.227	2.916	0.118	0.187
κ	0.013	2.522	0.513	0.503	0.438	0.112	0.594	3.666	0.349	0.713
τ	0.002	0.443	0.502	0.505	0.568	0.155	-0.086	2.777	0.243	0.763
r_A	0.070	9.338	0.820	0.777	0.699	0.438	0.425	2.720	0.159	1.608
γ_Q	-0.014	-1.723	0.786	0.792	0.819	0.094	-0.287	2.834	0.624	0.931
$\rho_{\hat{z}}$	0.008	3.750	0.208	0.205	0.187	0.035	0.428	3.301	0.155	0.270
$\rho_{\Delta\hat{q}}$	-0.003	-0.637	0.397	0.398	0.437	0.071	-0.113	3.078	0.280	0.509
$\rho_{\hat{y}^\star}$	-0.009	-0.976	0.891	0.896	0.896	0.039	-0.752	3.683	0.823	0.945
$\rho_{\hat{\pi}^\star}$	-0.005	-0.580	0.795	0.799	0.799	0.036	-0.340	3.077	0.731	0.851
$\sigma_{\hat{z}}$	-0.025	-2.546	0.975	0.977	0.950	0.080	-0.049	3.222	0.843	1.102
$\sigma_{\Delta\hat{q}}$	-0.009	-0.470	1.891	1.891	1.926	0.111	0.083	3.165	1.711	2.076
$\sigma_{\hat{y}^\star}$	0.757	40.06	2.647	1.950	1.570	2.423	2.541	11.52	0.480	7.618
$\sigma_{hat\pi^\star}$	-0.007	-0.217	2.993	2.988	2.959	0.180	0.042	2.953	2.697	3.294
σ_{ϵ^R}	0.006	1.561	0.406	0.400	0.398	0.057	6.028	126.6	0.335	0.498
Sample size = 600+7										
ψ_π	0.045	2.990	1.545	1.504	1.395	0.264	0.907	4.343	1.195	2.027
ψ_y	0.019	7.635	0.269	0.252	0.219	0.103	1.034	4.696	0.132	0.463
$\psi_{\Delta e}$	0.006	5.516	0.106	0.101	0.104	0.029	0.923	4.409	0.067	0.157
ρ_R	0.001	0.145	0.601	0.601	0.605	0.049	0.010	2.900	0.522	0.681
α	-0.000	-0.043	0.150	0.150	0.146	0.011	0.047	2.930	0.131	0.169
κ	0.003	0.546	0.503	0.499	0.468	0.055	0.341	3.175	0.417	0.599
τ	0.001	0.276	0.501	0.503	0.503	0.088	-0.053	3.001	0.358	0.642
r_A	0.011	1.511	0.761	0.751	0.629	0.239	0.141	2.873	0.381	1.163
γ_Q	-0.002	-0.190	0.798	0.799	0.778	0.051	-0.057	2.871	0.713	0.882
$\rho_{\hat{z}}$	0.003	1.305	0.203	0.201	0.199	0.020	0.502	3.475	0.173	0.238
$\rho_{\Delta\hat{q}}$	-0.000	-0.073	0.400	0.400	0.387	0.035	-0.042	2.836	0.342	0.457
$\rho_{\hat{y}^\star}$	-0.002	-0.265	0.898	0.899	0.898	0.017	-0.382	3.173	0.868	0.924
$\rho_{\hat{\pi}^\star}$	-0.002	-0.237	0.798	0.799	0.792	0.018	-0.323	3.095	0.766	0.827
$\sigma_{\hat{z}}$	-0.007	-0.675	0.993	0.995	0.997	0.041	-0.147	3.080	0.925	1.057
$\sigma_{\Delta\hat{q}}$	-0.002	-0.097	1.898	1.897	1.876	0.055	0.075	2.846	1.807	1.989
$\sigma_{\hat{y}^\star}$	0.226	11.96	2.116	1.915	1.894	0.999	1.577	7.571	0.906	3.931
$\sigma_{\hat{\pi}^\star}$	-0.004	-0.136	2.996	2.995	2.999	0.088	-0.021	2.987	2.849	3.141
σ_{ϵ^R}	0.001	0.227	0.401	0.400	0.406	0.024	0.292	3.326	0.363	0.443

Table A.2. CSN-1 shocks — Kalman filter Q-ML estimator, 2000 replications

Param.	Bias	Bias %	Mean	Mode	Median	St.Dev	Skewn.	Kurtosis	5%	95%
Sample size = 75+7										
ψ_π	0.542	36.14	2.042	1.608	1.212	1.290	2.634	11.44	1.032	4.713
ψ_y	0.235	93.96	0.485	0.307	0.174	0.551	3.314	17.92	0.077	1.425
$\psi_{\Delta e}$	0.063	62.88	0.163	0.114	0.074	0.148	3.036	15.54	0.048	0.452
ρ_R	0.017	2.797	0.617	0.610	0.580	0.124	0.039	2.785	0.417	0.832
α	0.003	1.708	0.153	0.151	0.147	0.029	0.309	3.073	0.107	0.203
κ	0.037	7.383	0.537	0.517	0.457	0.172	1.061	5.169	0.299	0.864
τ	-0.013	-2.613	0.487	0.486	0.497	0.200	-0.004	2.316	0.160	0.826
r_A	0.184	24.57	0.934	0.877	0.769	0.563	0.510	2.810	0.115	1.917
γ_Q	-0.035	-4.351	0.765	0.772	0.812	0.123	-0.244	2.788	0.555	0.951
$\rho_{\hat{z}}$	0.008	3.944	0.208	0.204	0.190	0.048	0.562	4.488	0.137	0.292
$\rho_{\Delta\hat{q}}$	-0.007	-1.842	0.393	0.401	0.411	0.103	-0.335	3.128	0.208	0.549
$\rho_{\hat{y}^\star}$	-0.018	-1.965	0.882	0.893	0.911	0.059	-0.896	3.823	0.770	0.960
$\rho_{\hat{\pi}^\star}$	-0.007	-0.834	0.793	0.796	0.789	0.052	-0.545	3.571	0.700	0.869
$\sigma_{\hat{z}}$	-0.057	-5.669	0.943	0.945	0.879	0.119	-0.134	3.378	0.749	1.136
$\sigma_{\Delta\hat{q}}$	-0.021	-1.098	1.879	1.876	1.934	0.172	0.115	3.077	1.601	2.177
$\sigma_{\hat{y}^\star}$	1.137	60.16	3.027	1.783	0.322	3.446	2.217	8.208	0.255	10.86
$\sigma_{\hat{\pi}^\star}$	-0.013	-0.418	2.987	2.988	2.907	0.269	0.026	3.013	2.548	3.423
$\sigma_{\hat{\epsilon}^R}$	0.019	4.659	0.419	0.406	0.400	0.121	18.16	551.8	0.312	0.551
Sample size = 150+7										
ψ_π	0.233	15.51	1.733	1.524	1.320	0.831	4.076	29.27	1.048	3.053
ψ_y	0.100	39.80	0.350	0.257	0.219	0.344	4.317	32.21	0.084	0.937
$\psi_{\Delta e}$	0.027	27.21	0.127	0.105	0.089	0.091	4.165	30.71	0.051	0.268
ρ_R	0.006	0.969	0.606	0.601	0.597	0.092	0.276	3.127	0.463	0.763
α	0.001	0.539	0.151	0.150	0.147	0.022	0.310	3.072	0.117	0.188
κ	0.014	2.885	0.514	0.502	0.495	0.113	0.711	3.727	0.356	0.723
τ	0.004	0.767	0.504	0.504	0.437	0.159	-0.048	2.705	0.243	0.767
r_A	0.064	8.598	0.814	0.786	0.946	0.427	0.362	2.727	0.158	1.568
γ_Q	-0.014	-1.769	0.786	0.791	0.817	0.091	-0.217	2.852	0.630	0.926
$\rho_{\hat{z}}$	0.007	3.304	0.207	0.205	0.194	0.034	0.274	3.038	0.152	0.268
$\rho_{\Delta\hat{q}}$	-0.006	-1.394	0.394	0.397	0.408	0.072	-0.205	3.024	0.273	0.508
$\rho_{\hat{y}^\star}$	-0.008	-0.865	0.892	0.897	0.899	0.038	-0.777	3.753	0.823	0.945
$\rho_{\hat{\pi}^\star}$	-0.004	-0.554	0.796	0.797	0.780	0.037	-0.293	3.071	0.732	0.852
$\sigma_{\hat{z}}$	-0.032	-3.187	0.968	0.969	0.948	0.082	-0.096	3.158	0.834	1.102
$\sigma_{\Delta\hat{q}}$	-0.003	-0.177	1.897	1.891	1.815	0.120	0.148	2.997	1.704	2.100
$\sigma_{\hat{y}^\star}$	0.856	45.29	2.746	1.961	0.940	2.598	2.580	11.82	0.480	7.758
$\sigma_{\hat{\pi}^\star}$	-0.014	-0.466	2.986	2.978	2.940	0.193	0.154	2.936	2.686	3.317
$\sigma_{\hat{\epsilon}^R}$	0.005	1.224	0.405	0.398	0.379	0.050	0.821	4.490	0.334	0.499
Sample size = 600+7										
ψ_π	0.043	2.859	1.543	1.502	1.432	0.291	1.661	10.72	1.169	2.070
ψ_y	0.016	6.533	0.266	0.247	0.216	0.113	1.932	12.04	0.128	0.466
$\psi_{\Delta e}$	0.005	4.820	0.105	0.100	0.092	0.032	1.706	11.04	0.064	0.162
ρ_R	0.000	0.056	0.600	0.599	0.600	0.050	0.178	3.263	0.519	0.684
α	-0.000	-0.019	0.150	0.150	0.152	0.012	-0.024	2.971	0.130	0.169
κ	0.003	0.635	0.503	0.499	0.487	0.055	0.332	3.126	0.418	0.599
τ	0.006	1.269	0.506	0.505	0.509	0.085	0.114	3.029	0.369	0.649
r_A	-0.000	-0.026	0.750	0.753	0.837	0.239	-0.001	2.920	0.354	1.150
γ_Q	0.000	0.054	0.800	0.800	0.793	0.051	0.002	2.849	0.717	0.885
$\rho_{\hat{z}}$	0.004	1.765	0.204	0.202	0.197	0.019	0.684	3.824	0.176	0.239
$\rho_{\Delta\hat{q}}$	-0.000	-0.073	0.400	0.400	0.415	0.037	-0.133	3.021	0.335	0.460
$\rho_{\hat{y}^\star}$	-0.002	-0.196	0.898	0.900	0.900	0.017	-0.403	3.185	0.868	0.924
$\rho_{\hat{\pi}^\star}$	-0.001	-0.084	0.799	0.800	0.802	0.018	-0.211	3.051	0.767	0.829
$\sigma_{\hat{z}}$	-0.008	-0.754	0.992	0.993	0.991	0.042	-0.125	3.055	0.923	1.059
$\sigma_{\Delta\hat{q}}$	-0.001	-0.061	1.899	1.900	1.881	0.059	-0.013	3.036	1.800	1.995
$\sigma_{\hat{y}^\star}$	0.263	13.89	2.153	1.959	1.496	0.988	1.716	8.943	0.975	3.984
$\sigma_{\hat{\pi}^\star}$	-0.005	-0.157	2.995	2.991	2.967	0.096	0.085	3.114	2.840	3.153
$\sigma_{\hat{\epsilon}^R}$	0.000	0.003	0.400	0.399	0.400	0.026	0.400	3.874	0.360	0.445

Table A.3. CSN-2 shocks — Kalman filter Q-ML estimator, 2000 replications.

Param.	Bias	Bias %	Mean	Mode	Median	St.Dev	Skewn.	Kurtosis	5%	95%
Sample size = 75+7										
ψ_π	0.485	32.33	1.985	1.575	1.227	1.223	2.672	11.63	1.026	4.488
ψ_y	0.217	86.84	0.467	0.303	0.125	0.509	2.885	13.96	0.073	1.506
$\psi_{\Delta e}$	0.057	57.45	0.157	0.114	0.064	0.137	2.879	14.38	0.048	0.422
ρ_R	0.015	2.567	0.615	0.609	0.599	0.121	0.076	2.804	0.427	0.830
α	0.002	1.577	0.152	0.151	0.153	0.028	0.262	3.032	0.108	0.202
κ	0.034	6.856	0.534	0.515	0.444	0.169	1.028	5.020	0.309	0.837
τ	-0.015	-2.928	0.485	0.485	0.521	0.195	-0.031	2.365	0.156	0.812
r_A	0.171	22.75	0.921	0.872	0.448	0.565	0.596	3.038	0.121	1.942
γ_Q	-0.035	-4.351	0.765	0.770	0.769	0.122	-0.316	2.984	0.551	0.955
$\rho_{\hat{z}}$	0.008	3.825	0.208	0.205	0.203	0.047	0.449	3.526	0.137	0.290
$\rho_{\Delta\hat{q}}$	-0.011	-2.707	0.389	0.392	0.413	0.101	-0.177	3.024	0.220	0.549
$\rho_{y^\star}$	-0.018	-2.031	0.882	0.892	0.896	0.060	-1.082	4.957	0.768	0.961
$\rho_{\pi^\star}$	-0.008	-0.974	0.792	0.795	0.811	0.053	-0.308	2.944	0.699	0.874
$\sigma_{\hat{z}}$	-0.059	-5.916	0.941	0.943	0.978	0.123	-0.033	3.503	0.740	1.138
$\sigma_{\Delta\hat{q}}$	-0.022	-1.138	1.878	1.874	2.029	0.185	0.217	3.003	1.583	2.193
$\sigma_{\hat{y}^\star}$	1.014	53.64	2.904	1.791	0.630	3.265	2.368	9.373	0.253	9.801
$\sigma_{\hat{\pi}^\star}$	-0.028	-0.949	2.972	2.960	2.907	0.289	0.288	3.132	2.524	3.457
$\sigma_{\hat{\epsilon}^R}$	0.014	3.536	0.414	0.404	0.428	0.075	0.887	4.172	0.313	0.553
Sample size = 150+7										
ψ_π	0.206	13.72	1.706	1.509	1.414	0.745	3.025	17.49	1.038	3.030
ψ_y	0.094	37.51	0.344	0.261	0.147	0.298	2.840	14.27	0.084	0.890
$\psi_{\Delta e}$	0.024	24.14	0.124	0.103	0.073	0.081	2.940	17.09	0.049	0.274
ρ_R	0.002	0.346	0.602	0.597	0.563	0.094	0.245	3.173	0.458	0.767
α	0.000	0.007	0.150	0.149	0.145	0.022	0.216	3.136	0.115	0.186
κ	0.019	3.752	0.519	0.505	0.466	0.116	0.768	4.228	0.355	0.726
τ	0.005	0.907	0.505	0.503	0.505	0.159	-0.031	2.591	0.250	0.770
r_A	0.068	9.116	0.818	0.799	0.744	0.412	0.331	2.830	0.179	1.529
γ_Q	-0.015	-1.893	0.785	0.791	0.806	0.090	-0.265	2.908	0.626	0.923
$\rho_{\hat{z}}$	0.005	2.570	0.205	0.203	0.193	0.034	0.414	3.558	0.152	0.264
$\rho_{\Delta\hat{q}}$	-0.004	-1.028	0.396	0.397	0.363	0.071	-0.157	3.100	0.279	0.511
$\rho_{\hat{y}^\star}$	-0.011	-1.190	0.889	0.894	0.899	0.038	-0.786	3.987	0.819	0.944
$\rho_{\hat{\pi}^\star}$	-0.004	-0.549	0.796	0.799	0.801	0.037	-0.359	3.196	0.730	0.852
$\sigma_{\hat{z}}$	-0.032	-3.247	0.968	0.968	0.971	0.083	-0.024	3.082	0.832	1.104
$\sigma_{\Delta\hat{q}}$	-0.011	-0.555	1.889	1.887	1.837	0.129	0.234	2.973	1.686	2.112
$\sigma_{\hat{y}^\star}$	0.845	44.73	2.735	1.940	1.165	2.555	2.462	11.19	0.460	7.847
$\sigma_{\hat{\pi}^\star}$	-0.015	-0.503	2.985	2.985	2.984	0.214	0.205	3.275	2.640	3.335
$\sigma_{\hat{\epsilon}^R}$	0.005	1.264	0.405	0.396	0.381	0.054	0.791	4.076	0.331	0.502
Sample size = 600+7										
ψ_π	0.040	2.680	1.540	1.489	1.469	0.313	6.702	134.6	1.190	2.022
ψ_y	0.016	6.525	0.266	0.249	0.206	0.114	4.565	73.65	0.135	0.452
$\psi_{\Delta e}$	0.005	4.537	0.105	0.100	0.095	0.032	4.866	81.35	0.066	0.156
ρ_R	-0.001	-0.141	0.599	0.598	0.597	0.048	0.122	2.969	0.522	0.682
α	0.000	0.047	0.150	0.150	0.155	0.011	0.038	2.986	0.132	0.169
κ	0.005	1.080	0.505	0.503	0.476	0.058	1.799	27.49	0.417	0.598
τ	0.003	0.684	0.503	0.502	0.483	0.084	0.067	3.179	0.369	0.643
r_A	0.004	0.515	0.754	0.756	0.693	0.240	0.040	2.790	0.349	1.149
γ_Q	-0.001	-0.171	0.799	0.799	0.804	0.051	-0.005	2.712	0.714	0.885
$\rho_{\hat{z}}$	0.003	1.303	0.203	0.200	0.194	0.019	0.593	3.895	0.175	0.239
$\rho_{\Delta\hat{q}}$	-0.002	-0.549	0.398	0.399	0.381	0.035	-0.076	2.956	0.338	0.455
$\rho_{y^\star}$	-0.002	-0.254	0.898	0.899	0.902	0.018	-0.511	3.425	0.866	0.925
$\rho_{\pi^\star}$	-0.001	-0.180	0.799	0.799	0.797	0.019	-0.127	3.228	0.767	0.829
$\sigma_{\hat{z}}$	-0.008	-0.808	0.992	0.991	1.007	0.044	0.008	2.998	0.919	1.066
$\sigma_{\Delta\hat{q}}$	-0.004	-0.189	1.896	1.895	1.880	0.065	0.124	3.040	1.789	2.005
$\sigma_{\hat{y}^\star}$	0.226	11.96	2.116	1.913	1.590	0.978	1.891	10.32	0.951	3.873
$\sigma_{\hat{\pi}^\star}$	-0.004	-0.127	2.996	2.996	2.999	0.103	0.102	2.996	2.829	3.168
$\sigma_{\hat{\epsilon}^R}$	0.001	0.332	0.401	0.400	0.392	0.054	30.473	1198.	0.361	0.445

A.2.2 Estimation of state variables

Table A.4. Simulated and estimated /smoothed/ shocks, 2000 replications.

Simul. Variant	Shock u(i)	Simulated shocks /average in sample/					Estimated shocks /average in sample /				
		Mean	Median	St.Dev	Skewn.	Kurt.	Mean	Median	St.Dev	Skewn.	Kurt.
		Sample size = 75+7									
Normal	z	-0.025	-0.023	0.995	-0.009	2.988	0.001	0.004	0.859	-0.008	2.988
	Δq	0.001	0.001	1.898	-0.001	3.007	0.002	0.002	1.877	-0.004	3.004
	$y^\star$	0.001	0.000	1.880	-0.003	3.011	0.007	0.012	2.627	-0.004	2.985
	$\pi^\star$	-0.015	-0.006	2.983	-0.012	3.013	-0.012	-0.004	2.929	-0.013	3.012
	ϵ^R	0.004	0.003	0.398	-0.002	2.955	-0.001	-0.001	0.405	-0.005	2.964
CSN-1	z	-0.025	-0.113	0.989	0.483	3.352	0.002	-0.047	0.853	0.315	3.204
	Δq	0.003	-0.164	1.890	0.480	3.323	0.002	-0.158	1.871	0.470	3.315
	$y^\star$	-0.001	-0.168	1.877	0.459	3.274	-0.017	-0.128	2.608	0.223	3.101
	$\pi^\star$	-0.013	-0.294	2.997	0.469	3.267	-0.009	-0.269	2.947	0.440	3.251
	ϵ^R	0.004	-0.032	0.399	0.475	3.304	-0.000	-0.032	0.408	0.412	3.253
CSN-2	z	-0.025	-0.219	0.986	0.926	3.724	0.003	-0.101	0.850	0.605	3.423
	Δq	0.002	-0.359	1.891	0.914	3.724	0.003	-0.340	1.870	0.879	3.689
	$y^\star$	-0.002	-0.355	1.872	0.895	3.666	0.007	-0.210	2.503	0.420	3.262
	$\pi^\star$	-0.014	-0.580	2.985	0.918	3.732	-0.007	-0.541	2.932	0.870	3.682
	ϵ^R	0.002	-0.074	0.399	0.905	3.685	0.000	-0.065	0.404	0.782	3.571
		Sample size = 150+7									
Normal	z	-0.010	-0.012	0.999	-0.003	3.005	0.001	0.000	0.878	-0.001	2.994
	Δq	0.004	0.004	1.898	0.007	3.000	0.005	0.008	1.880	0.006	3.000
	$y^\star$	-0.009	-0.010	1.885	0.005	3.004	-0.014	-0.015	2.264	0.008	3.000
	$\pi^\star$	-0.004	0.005	2.998	-0.005	2.991	-0.003	0.001	2.951	-0.007	2.992
	ϵ^R	0.002	0.002	0.399	-0.006	3.014	0.000	0.000	0.395	-0.002	3.011
CSN-1	z	-0.011	-0.101	0.995	0.483	3.316	0.000	-0.052	0.871	0.330	3.197
	Δq	0.003	-0.169	1.902	0.484	3.309	0.004	-0.166	1.886	0.470	3.297
	$y^\star$	-0.001	-0.174	1.888	0.484	3.314	0.003	-0.099	2.349	0.255	3.145
	$\pi^\star$	-0.007	-0.286	2.994	0.482	3.310	-0.005	-0.258	2.944	0.456	3.292
	ϵ^R	0.000	-0.036	0.399	0.484	3.321	0.000	-0.031	0.394	0.430	3.278
CSN-2	z	-0.011	-0.205	0.992	0.929	3.737	0.001	-0.111	0.872	0.630	3.436
	Δq	0.001	-0.366	1.895	0.928	3.743	0.001	-0.353	1.878	0.904	3.717
	$y^\star$	-0.002	-0.364	1.884	0.925	3.734	0.000	-0.216	2.335	0.481	3.304
	$\pi^\star$	0.003	-0.567	2.990	0.927	3.765	0.005	-0.528	2.942	0.881	3.717
	ϵ^R	0.001	-0.076	0.399	0.932	3.772	0.000	-0.067	0.394	0.827	3.665
		Sample size = 600+7									
Normal	z	-0.001	-0.001	1.001	0.001	3.008	0.000	0.000	0.894	0.002	3.004
	Δq	0.001	0.002	1.899	-0.001	3.001	0.001	0.001	1.887	-0.001	3.001
	$y^\star$	0.001	-0.001	1.891	-0.000	3.002	-0.001	0.000	1.784	-0.001	2.998
	$\pi^\star$	-0.003	-0.006	2.998	-0.001	2.990	-0.004	-0.006	2.953	-0.001	2.988
	ϵ^R	0.001	0.001	0.400	-0.001	2.992	0.000	0.000	0.389	-0.000	2.995
CSN-1	z	0.000	-0.092	1.000	0.501	3.351	0.000	-0.058	0.893	0.359	3.230
	Δq	0.002	-0.172	1.900	0.498	3.347	0.003	-0.166	1.888	0.488	3.337
	$y^\star$	-0.001	-0.175	1.890	0.496	3.345	0.001	-0.088	1.814	0.280	3.166
	$\pi^\star$	-0.003	-0.280	2.997	0.494	3.337	-0.003	-0.260	2.952	0.471	3.317
	ϵ^R	0.000	-0.037	0.400	0.496	3.338	0.000	-0.033	0.388	0.451	3.303
CSN-2	z	-0.001	-0.196	0.999	0.950	3.824	-0.000	-0.119	0.893	0.679	3.533
	Δq	0.001	-0.369	1.898	0.944	3.799	0.001	-0.359	1.886	0.926	3.779
	$y^\star$	0.004	-0.364	1.891	0.950	3.830	0.005	-0.173	1.782	0.535	3.392
	$\pi^\star$	-0.000	-0.584	2.999	0.942	3.786	-0.000	-0.548	2.953	0.901	3.741
	ϵ^R	-0.000	-0.078	0.400	0.943	3.787	-0.000	-0.068	0.390	0.853	3.685

A.2.3 The test for skewness

Table A.5. The modified sample skewness estimator $\tilde{\gamma}$. Percent of cases outside the admissible range. 100000 replications, $u \sim csn$, $\text{var}(u) = 1$, $\text{E}(u) = 0$.

Shock's skewness coeff.	Sample size 50	75	150	300	600	1200	4800	9600
0.000	0.97	0.16	0.00	0.00	0.00	0.00	0.00	0.00
0.250	2.92	1.14	0.11	0.00	0.00	0.00	0.00	0.00
0.500	9.12	6.04	1.87	0.26	0.10	0.00	0.00	0.00
0.750	22.73	19.27	12.94	6.73	2.17	0.26	0.00	0.00
0.900	35.76	33.88	30.57	25.59	19.36	12.31	1.28	0.09
0.950	40.35	39.58	38.21	36.45	32.94	27.85	13.46	6.39
0.995	45.00	45.35	46.45	46.88	47.87	48.31	48.74	49.13

Table A.6. Properties of the adjusted sample skewness estimator $\widehat{\widehat{\gamma}}(u)$. 100000 replications, $u \sim csn$, $\text{var}(u) = 1, \text{E}(u) = 0$.

Skewn. coeff.	Estim. feature	Sample size 50	75	150	300	600	1200	4800	9600
0.000	Bias	-0.002	0.000	-0.001	0.000	0.000	-0.000	0.000	0.000
	St.dev.	0.356	0.288	0.201	0.143	0.101	0.071	0.035	0.025
	Skewness	0.007	-0.008	0.007	-0.007	-0.014	-0.003	0.001	0.016
	Kurtosis	2.979	3.172	3.168	3.149	3.066	3.050	2.978	2.978
-0.250	Bias	0.016	0.008	0.002	0.001	0.001	0.000	0.000	-0.000
	St.dev.	0.356	0.296	0.213	0.150	0.107	0.075	0.038	0.027
	Skewness	0.058	-0.068	-0.158	-0.151	-0.106	-0.087	-0.032	-0.014
	Kurtosis	2.800	2.950	3.166	3.165	3.082	3.114	2.981	2.988
0.500	Bias	-0.042	-0.022	-0.008	-0.003	-0.001	-0.001	-0.001	0.000
	St.dev.	0.331	0.283	0.212	0.156	0.112	0.078	0.039	0.028
	Skewness	-0.281	-0.137	0.107	0.213	0.189	0.126	0.088	0.043
	Kurtosis	2.520	2.512	2.741	3.047	3.110	3.098	3.034	3.019
-0.750	Bias	0.086	0.060	0.027	0.010	0.003	0.002	0.000	0.000
	St.dev.	0.273	0.235	0.183	0.144	0.110	0.080	0.041	0.029
	Skewness	0.700	0.621	0.393	0.179	-0.025	-0.157	-0.091	-0.066
	Kurtosis	2.751	2.673	2.403	2.406	2.615	2.941	3.019	3.019
0.900	Bias	-0.131	-0.103	-0.064	-0.040	-0.021	-0.009	-0.001	-0.000
	St.dev.	0.223	0.188	0.143	0.112	0.087	0.068	0.040	0.029
	Skewness	-1.085	-1.040	-0.991	-0.861	-0.733	-0.520	-0.074	0.070
	Kurtosis	3.475	3.337	3.245	2.977	2.809	2.559	2.696	2.963
0.950	Bias	-0.152	-0.123	-0.085	-0.057	-0.038	-0.023	-0.006	-0.002
	St.dev.	0.205	0.171	0.128	0.095	0.072	0.054	0.032	0.026
	Skewness	-1.257	-1.242	-1.252	-1.248	-1.167	-1.083	-0.708	-0.426
	Kurtosis	3.950	3.894	3.963	3.958	3.731	3.581	2.964	2.690
-0.995	Bias	0.172	0.144	0.106	0.079	0.059	0.045	0.028	0.023
	St.dev.	0.189	0.156	0.112	0.080	0.056	0.039	0.018	0.012
	Skewness	1.396	1.451	1.520	1.638	1.722	1.887	1.886	1.290
	Kurtosis	4.357	4.570	4.773	5.325	5.655	6.629	7.885	7.177

Bibliography

Akdemir, D. 2009. A class of mutlivariate skew distributions: Properies and inferential issues. Doctor of philosophy disseratation, Graduate College of Bowling Green State Univeristy.

Amisano, G. and Tristani, O. 2007. Euro area inflation persistence in an estimated nonlinear DSGE model. Working Paper Series 754, European Central Bank.

An, S. 2005. Bayesian estimation of DSGE model: Lessons from second-order aproximations. Univeristy of Pennsylvania, Manuscript, mimeo.

Arnold, B. C. and Groneneveld, R. A. 1995. Measuring skewness with respect to the mode. *The American Statistician* 49:34–38.

Azzalini, A. 2004. The skew-normal distribution and related skew normal distribution multivariate families. Paper, presented at the 20th Nordic Conference on Mathematical Statistics, Jyvaskyla, (NORDSTAT 2004).

Azzalini, A. and Capitanio, A. 1999. Applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 61:579–602.

Azzalini, A. and Genton, M. G. 2008. Robust likelihood methods based on the skew-t and related distributions. *International Statistical Review* 76:106–129.

Azzalini, A., Genton, M. G., and Scarpa, B. 2010. Invariance-based estimating equations for skew-symmetric distributions. Manuscript, mimeo.

Azzalini, A. and Valle, A. D. 1996. The multivariate skew-normal distribution. *Biometrika* 83:715–726.

Bai, J. and Ng, S. 2005. Test for skewness, kurtosis, and normality for time series data. *Journal of Business and Economic Statstics* 23:43–60.

Ball, L. and Mankiw, N. G. 1995. Relative-price changes as aggregate supply shocks. *The Quarterly Journal of Economics* 110:161–93.

Christodoulakis, G. and Peel, D. 2009. The central bank inflation bias and the presence of asymmetric preferences and non-normal shocks. *Economic Bulletion* 29:1608–1620.

Cochrane, J. H. 2007. Determinacy and identification with taylor rules. Working Paper 13409, NBER.

Dunajeva, O., Kollo, T., and Traat, I. 2003. Bias correction for the shape parameter of the skew normal distributions. *Tatra Mountains Mathematical Publications* 26:281–289.

Fahr, S. and Smets, F. 2008. Downward wage rigidities and optimal monetary policy in a monetary union. Manuscript.

Fernández-Villaverde, J. 2009. The econometrics of DSGE models. Working Paper 14677, NBER.

Fernández-Villaverde, J. and Rubio-Ramirez, J. F. 2007. Estimationg macroeconomic models: A likelihood approach. *Review of Economic Studies* pp. 1059–1087.

Fletcher, C., Noveau, P., and Allard, D. 2008. Estimating the closed skew-normal distributions parameters using weighted moments. Research report 40, Institue National de la Recherche Agronomique. *Review of Economic Studies* 72.

González-Farías, G., Domínguez-Molina, J. A., and Gupta, A. K. 2004. Additive properties of skew normal random vectors. *Journal of Statistical Planning and Inference* 126:521–534.

Gourieroux, C., Monfort, A., and Trognon, A. 1984. Pseudo maximim likelihood method: Theory. *Econometrica* 52:681–700.

Gupta, A. K., González-Farías, G., and Domínguez-Molina, J. A. 2004. A multivariate skew normal distribution. *Journal of Multivariate Analysis* 89:181–190.

Gupta, M. K. 1967. An asymptotically nonparametric test of symmetry. *The Annals of Mathematical Statistics,* 38:849–866.

HAMILTON, J. D. 1994. Time Series Analysis. Princeton University Press.

IVERSEN, D. H. 2010. Closed-skew distribution. Simulation, inversion and parameter estimation. Unpublished manuscript, Norvegian University of Science and Technology.

KIM, J. AND RUGE-MURCIA, F. J. 2009. How much inflation is necessary to grease the wheels? *Journal of Monetary Economics* 56:365–377.

LUBIK, T. A. AND SCHORFHEIDE, F. 2007. Do central banks respond to exchange rate movements? A structural investigation. *Journal of Monetary Economics* 54:1069–1087.

MA, Y., GENTON, M., AND TSIATSIS, A. A. 2005. Locally efficient semiparametric estimators for generalized skew-elliptical distributions. *Journal of the American Statistical Association* 100:980–989.

MACGILLIVRAY, H. L. 1986. Skewness and asymmetry: measures and orderings. *The Annals of Statistics* 14:994–1011.

MEINHOLD, R. J. AND SINGPURWALLA, N. D. 1983. Understanding the Kalman filter. *The American Statistician* 37:123–127.

MEINHOLD, R. J. AND SINGPURWALLA, N. D. 1989. Robustification of Kalman filter models. *Journal of the American Statistical Association* 84:479–486.

NADLER, J. A. AND LEE, Y. 1992. Likelihood, quasi-likelihood and pseudolikelihood: Some comparisons. *Journal of Royal Statistical Society. Series B (Methodological)* 54:273–284.

NAVEAU, P., GENTON, M. G., AND SHEN, X. 2005. A skewed Kalman filter. *Journal of Multivariate Analysis* 94:382–400.

NEGRO, M. D. AND SCHORFHEIDE, F. 2008. Inflation dynamics in a small open economy model under inflation targeting: Some evidence from Chile. Working Papers Central Bank of Chile 486, Central Bank of Chile.

NEWEY, W. K. 1990. Semmiparametric efficiency bounds. *Journal of Applied Econometrics* 5:99–135.

NEWEY, W. K. AND WEST, K. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–708.

RANDLES, R. H., FLIGNER, M. A., POLICELLO II, G. E., AND WOLFE, D. A. 1980. An asymptotically distribution-free test for symmetry versus asymmetry. *Journal of the American Statistical Association* 75:168–172.

ROCH, O. AND VALDEZ, E. A. 2009. Lower convex order bound approximations for sums of log-skew normal random variables. Manuscript, mimeo, Universitat de Barcelona.

WARNE, A. 2010. YADA manual — Computational details. EBC, Manuscript. URL http://www.texlips.net/yada/index.html.

WEDDERBURN, R. 1974. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrica* 61:439–447.

WHITE, H. 1982. Maximum likelihood estimation of misspecified models. *Econometrica* 50:1–25.