Understanding Uncertainty Shocks and the Role of the Black Swan

Anna Orlik*[1] Laura Veldkamp**

*Federal Reserve, Board of Governors

**NYU Stern

[1]Disclaimer: The views expressed herein are those of the authors and do not necessarily reflect the position of the Board of Governors of the Federal Reserve or the Federal Reserve System

Introduction

- Many models explore effects of exogenous uncertainty shocks.
  But where do uncertainty shocks come from?

- Uncertainty: Stdev of a forecast (error) conditional on I_t.

$$U_t = Std[y_{t+1}|I_t] = \sqrt{E\left[\left(y_{t+1} - E\left(y_{t+1}|I_t\right)\right)^2|I_t\right]}$$

- Existing literature: Uncertainty measurement assumes I_t contains true y_t distribution and its parameters.

- Our paper: No one knows the true distribution of outcomes. They re-estimate it each period. $\rightarrow$ large, counter-cyclical uncertainty shocks.

Assuming agents know the true model parameters misses most fluctuations in uncertainty.

What We Do

Use only real-time GDP data (1968-2013, Philly Fed) to estimate the distribution of GDP growth and measure uncertainty, accounting for parameter uncertainty.

- Begin with prior beliefs over the state and parameters (estimated from 1947-68 data).
- Observe each quarter of data and apply Bayes' Law.
  - Metropolis-Hastings + change-of-measure technique
    $\rightarrow$ distributions of parameters, conditional on data history.
- Calculate $U_t = Std[y_{t+1}|I_t]$.
- How much of uncertainty changes come from not knowing parameters?

What Distribution to Estimate?

- Key feature: Agents estimate tail probabilities.

  A normal distribution fixes these $\rightarrow$ no U_t action.
  Need parameters that govern higher moments (skewness).
  - Can capture skewness in GDP data (-0.3)
  - Key for our forecasts to resemble SPF forecast data

- Solution: Take a linear hidden state model (Kalman filter system) and do an exponential twist.

- A form of g-and-h transformation used in statistics for Bayesian distribution fitting (Headrick '10).

Forecasting Model

- We estimate this:

$$\begin{aligned} y_t &= c + b\exp(-S_t - \sigma\varepsilon_t) \\ S_t &= \rho S_{t-1} + \sigma^S \xi_t \end{aligned}$$

  where ε_t and $\xi_t \sim iid\ N(0,1)$. $y_t =$ GDP growth.

- Why do statisticians use this forecasting model?
  1. small # parameters.
  2. parameters that regulate skewness
  3. small deviation from linear-normal with a one-to-one mapping (tractability)

Two Benchmarks

What comes from parameter learning, what comes from the skewness and what comes from the interaction?

We compare our forecasts to:

1. The same model with known parameters.
   Parameters from max likelihood on the full data sample.
   Called "volatility" (V_t).

2. A linear-normal model with parameter uncertainty

$$\begin{aligned} y_t &= C + S_t + \sigma\varepsilon_t \\ S_t &= \rho S_{t-1} + \sigma^S \xi_t \end{aligned}$$

Results

model:	unc/vol	normal	skewed
Mean	U_t	4.20%	4.53%
	V_t	3.45%	4.01%
Std deviation	U_t	0.48%	1.50%
	V_t	0%	0.05%
	Detrended uncertainty/volatility		
$\text{Corr}(\tilde{U}_t,\ E_t[y_{t+1}])$	U_t	0.04	-0.78
$\text{Corr}(\tilde{V}_t,\ E_t[y_{t+1}])$	V_t	0	-0.74
	Forecast properties		
	data	normal	skewed
Mean forecast	2.29%	2.82%	2.27%

Takeaways: Learning about parameters (esp. black swan risk) $\rightarrow$

1. Large uncertainty shocks (std U_t)
2. Counter-cyclical uncertainty shocks
3. Matches SPF forecast bias

Result 1: Large Uncertainty Shocks

	known parameters		*estimated params*	
	normal	skewed	normal	skewed
Std dev of U_t	0	0.05%	0.48%	1.50%

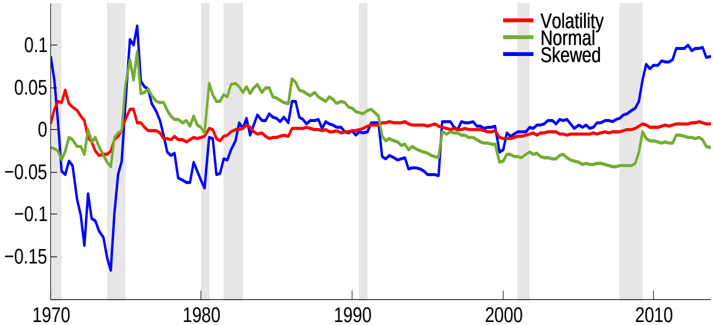
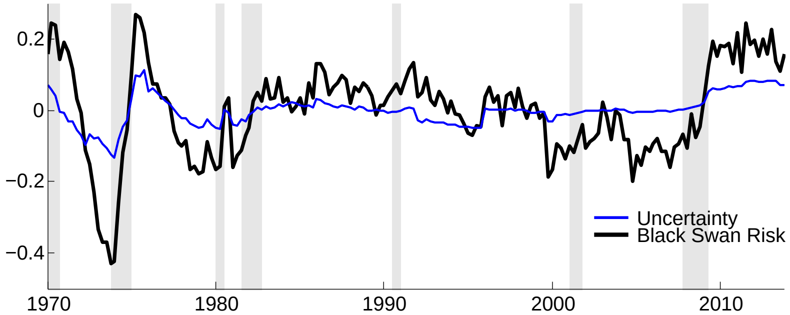


Figure: Uncertainty (U_t) in linear and skewed models, in mean-zero, log deviations from trend.

Parameter learning + Skewness = Large uncertainty shocks.

What Explains Large Shocks? Black Swans.

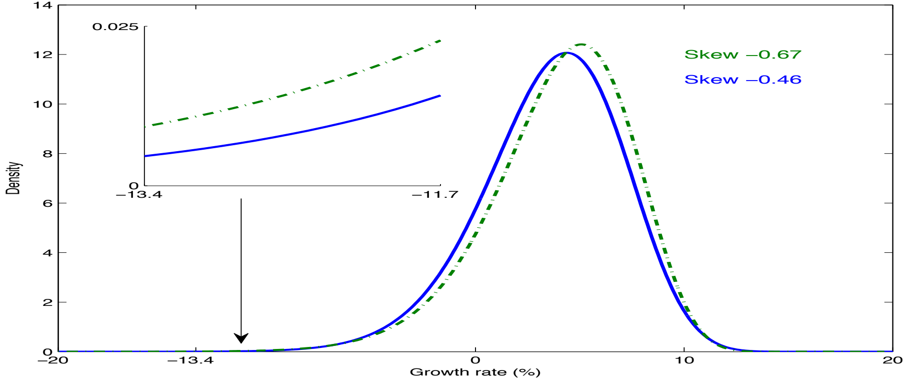
$Black\ Swan\ Risk_t \equiv Prob[y_{t+1} \leq -6.8\% | y^t]$ (1-in-100 year event)
Correlation(BSw, U_t) is 75% (both detrended).



Most changes in uncertainty come from re-estimating probability of unobserved tail events (black swans).

Why Are Black Swan Probabilities Volatile?

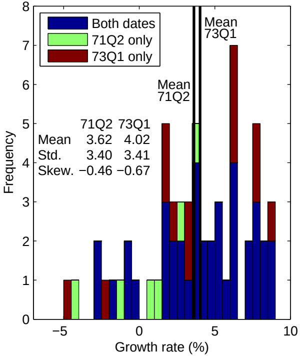
- Extreme event probabilities are very sensitive to small revisions in skewness.
- Skewness keeps fluctuating because it is hard to learn.

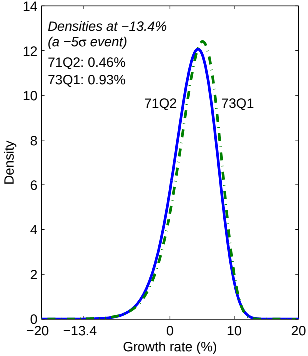


What events trigger black swan shocks?

Events that make skewness more negative:

1. Negative outliers
2. A string of mild positive realizations and a small negative revision. Makes left tail observations more extreme. (builds fragility)
   - Ex: 7 qtrs of mostly positive data in'70s. Black swan prob doubles.





Result 2: Uncertainty is Counter-Cyclical

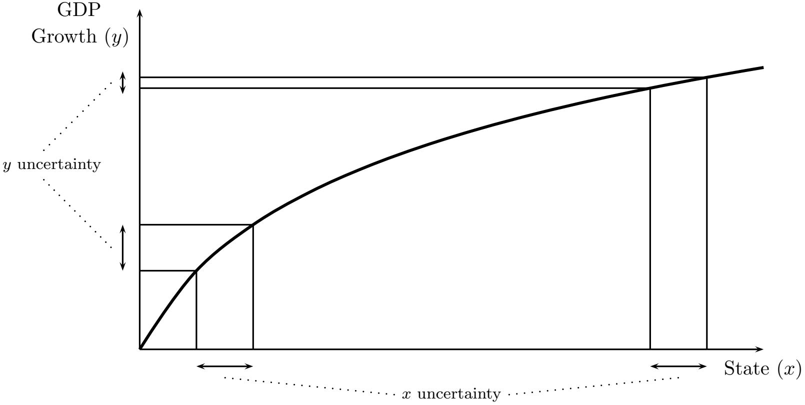
	known parameters		estimate params	
	normal	skewed	normal	skewed
$corr(U_t, \Delta GDP)$	0	-0.74	0.04	-0.78

- Why? Think of probability distribution as a having a normal pdf with a change-of-measure function g^{-1}. $y = g(x)$ where x is normal.
- Then, by the Radon-Nikodym theorem,

$$Var[y_{t+1}|y^t] = E\left[\frac{dg}{dx}\middle| x^t\right] Var[x_{t+1}|x^t] + cov(\frac{dg}{dx}, (x_{t+1} - E[x_{t+1}])^2)$$

- Lemma: If g is concave, y is negatively skewed.
  $\rightarrow$ dg/dx is a decreasing function of x.
  $\rightarrow$ $Var[y_{t+1}|y^t]$ is decreasing in $E[x_{t+1}|x^t]$.

Counter-Cyclical Uncertainty and Changing Measure



Concavity of the change-of-measure function is key to counter-cyclical uncertainty. It arises endogenously because GDP growth is negatively skewed.

Result 3: Forecast Bias

$E[y_{t+1}|y^t, \theta]$ is mean GDP growth $= 2.68\%$.

$E[y_{t+1}|y^t]$ is average growth forecast $= 2.29\%$ in data, $= 2.27\%$ in model.
Bias declines over time in the model and data.

- Forecast bias is known in the forecasting literature.
  Explanations focus on forecasters' objectives.
- Our forecasters just use Bayes law. They are just as biased as professional forecasters!
- We prove: skewness + param uncertainty = forecast bias.

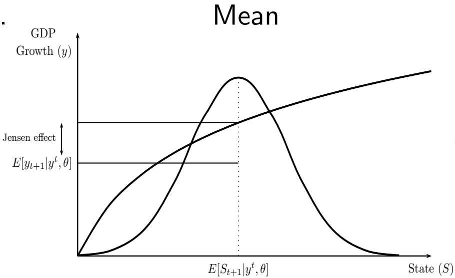
Lemma

Suppose $y = g(x)$, with g concave and $x \sim N(\mu, \sigma)$. If param distributions $h(\mu'|\sigma')$ and $k(\sigma')$ have means μ and σ, then mean > forecast: $\int y f(y|\mu,\sigma)dy > \int \int \int y\, f(y|\mu',\sigma')\, h(\mu'|\sigma')\, k(\sigma')\, dy\, d\mu'\, d\sigma'$.

Result 3: Why are forecasts low?

- Mean is an expectation, conditional on true parameters.
- Forecast is conditional on distributions of params. More uncertainty.
- GDP growth is a concave fn of a normal variable. Expectation has a Jensen inequality term. (mean $<$ median) More uncertainty makes Jensen term bigger, forecast lower.

.



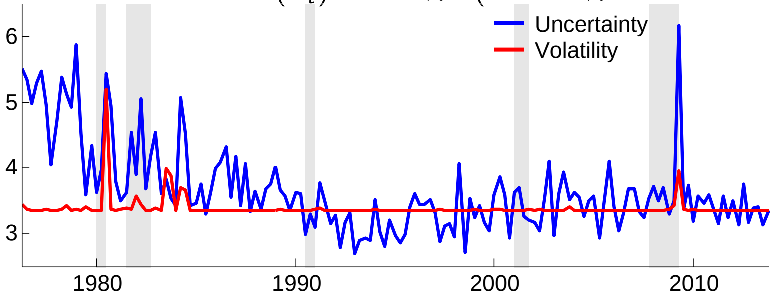
Perhaps this captures how forecasters form beliefs?

Result 4: With Stochastic Volatility, U_t Is Stationary

$$
\begin{aligned}
y_t &= c + b\exp(-S_t - \sigma\varepsilon_t) \\
S_t &= \rho S_{t-1} + \sigma_t^S \xi_t \quad \sigma_t^S \in \{\sigma_L^S, \sigma_H^S\} \ (\textit{both estimated}).
\end{aligned}
$$

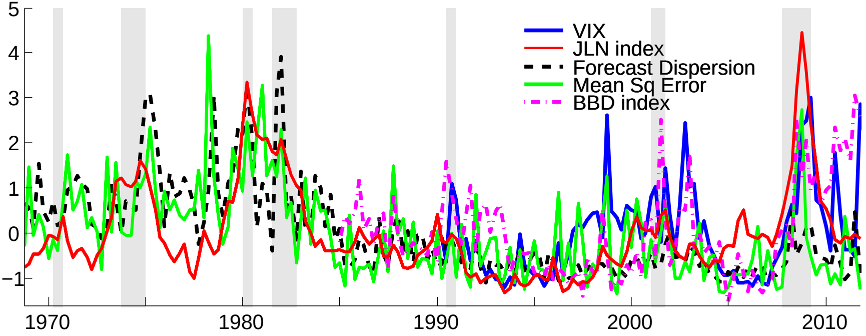
where $\rho = \sigma = 0$. ρ is small We set $= 0$ to facilitate estimation. Then σ is redundant.

Std(U_t) = 1.61% (vs. 1.50% with fixed vol).



When parameters change, beliefs don't converge. No U_t trend.

Comparing U_t to Other Uncertainty Measures



Highest correlation is U_t with BBD = 60% and VIX = 40%.
For stochastic vol, highest correlation is with forecast dispersion = 26%.

Evaluating Model Fit

Forecast properties	data	normal	skewed
Mean forecast	2.29%	2.82%	2.27%
Mean $\lvert F\ Err \rvert$	1.87%	2.25%	2.51%
Std forecast	2.25%	1.17%	0.64%
Std $\lvert F\ Err \rvert$	1.46%	2.17%	2.39%

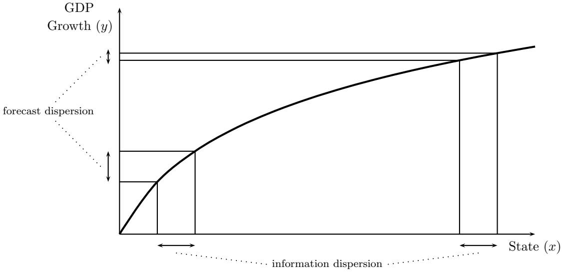
Does this forecasting model perform well?

- Not bad for using only GDP data.
  Adding signals reduces $\lvert F\ Err \rvert$, Std $\lvert F\ Err \rvert$ and increases std forecast. Can explain forecast dispersion as well (Kozeniauskas, O-V, 2014).
- Mean forecast looks like SPF data.
- Forecast error is high, but once you remove forecast bias (desirable), it falls to 2.1%.

Ultimately: We know skewed model fits better bc normal is almost-nested and not selected by the data.

Why Do Shades of Uncertainty Covary?

- High correlations are puzzling: Higher-order and micro uncertainty are distinct phenomena, arising from info or firm shock dispersion.
- Explanation: Black swan risk.
  Same skewness mechanism amplifies uncertainty *and disagreement.*
  "Black Swans and the Many Shades of Uncertainty" Kozeniauskas, O&V (2014)



Bad news → neg skewness → ↑ disaster risk → ↑ disagreement (H-O uncert) → ↑ production dispersion (micro uncert). Works quantitatively.

A unified explanation for many belief shocks.

Conclusions

- Macro theories typically assume agents know the true model and its parameters. The only source of uncertainty is which draw from a known distribution. This limited view rules out important sources of uncertainty. Learning about skewness is one example.
- When we allow agents to learn about model skewness, they experience large uncertainty shocks.
  - Skewness is tough to learn in small samples, so new data causes revisions.
  - Small revisions cause large changes in the probability of extreme events (black swans).
  - Changes in black swan risk affect conditional variance $\rightarrow$ uncertainty shocks.
  - In progress: Use this mechanism to explain financial crises and equity risk premia.

Forecast and Uncertainty Moments

model:	unc/vol	normal	skewed	stoch.vol
Mean	U_t	4.20%	4.53%	4.29%
	V_t	3.45%	4.01%	3.43%
Std deviation	U_t	0.48%	1.50%	2.00%
	V_t	0%	0.05%	0.34%
Autocorrelation	U_t	0.99	0.97	0.83
	V_t	0	0.93	0.22
	Detrended uncertainty/volatility			
$\text{Corr}(\tilde{U}_t,\ E_t[y_{t+1}])$		0.04	-0.78	-0.60
$\text{Corr}(\tilde{V}_t,\ E_t[y_{t+1}])$		0	-0.74	-0.99
	Forecast properties			
	data	normal	skewed	
Mean forecast	2.29%	2.82%	2.27%	2.05%
Mean $\|F\ Err\|$	1.87%	2.25%	2.51%	2.61%
Std forecast	2.25%	1.17%	0.64%	0.54%
Std $\|F\ Err\|$	1.46%	2.17%	2.39%	2.37%

Priors for Estimating Normal and Skewed Models

.

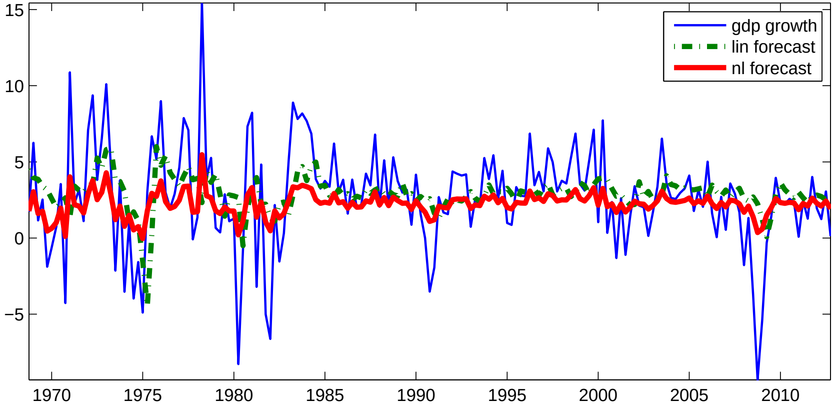
	Normal Model		Skewed Model		
Parameter	Mean	Stdev	Mean	Stdev	Distribution
c	2.35	0.68	41.27	6.97	inverse Γ
ρ	0.47	0.12	0.05	0.07	β
σ^2	4.89	3.45	0.02	0.01	inverse Γ
σ_s^2	15.92	4.47	0.005	0.007	inverse Γ

Table: Priors are estimated from 1947-68Q3 data.

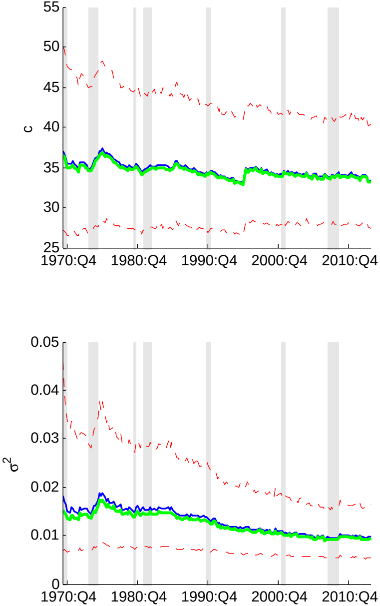
Estimation Procedure

- For skewed model, first transform to normal: $\ln(y_t - c)$. Then, for a given c, estimate other params with Metropolis-Hastings.
- Metropolis-Hastings steps are random walks (prportional r.w.s for σ and σ_s.
- For given parameter vector, compute the probability of the data history y^t analytically.
- Based on the probability, accept or reject the parameter vector.
- After many draws, the set of accepted has a frequency that approximates the conditional distribution of parameters.
- Also draw c's, transform data back to y's and accept/reject c proposals using the same criterion.

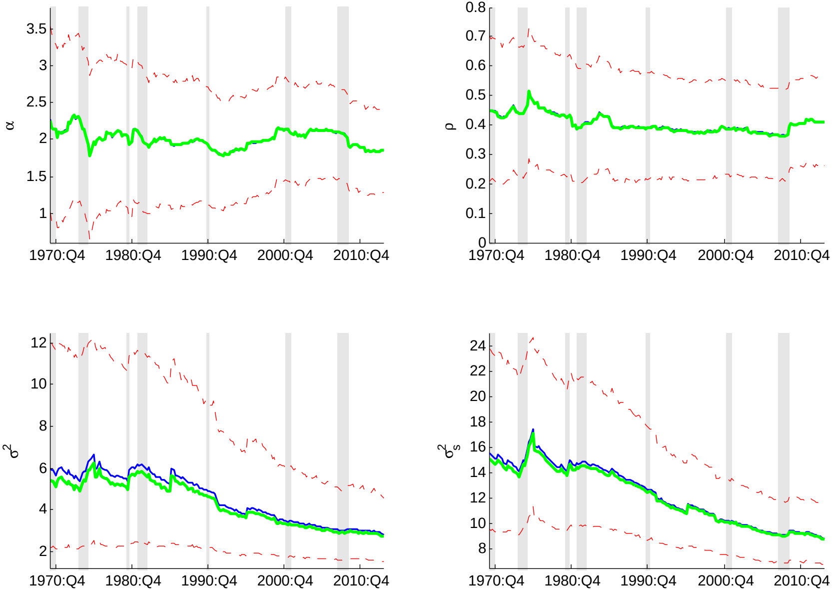
RGDP growth and forecasted growth



Skewed Model Parameter Estimates



Normal Model Parameter Estimates



Isn't Forecast Dispersion a "Model-free" Uncertainty Measure?

A general orthogonal decomposition:

$$y_{t+1} = E\left(y_{t+1}|I_t\right) + \eta_t + \epsilon_{it}$$

Then, uncertainty and forecast dispersion are

$$\begin{aligned} U_{it}^2 &= E\left[\left(\eta_t + \epsilon_{it}\right)^2 |I_t\right] = Var\left(\eta_t|I_{it}\right) + Var\left(\epsilon_{it}|I_t\right) \\ D_t^2 &= \frac{1}{N}\sum_i \left(E\left(y_{t+1}|I_t\right) - \overline{E}_t\right)^2 = \frac{1}{N}\sum_i Var\left(\epsilon_{it}|I_{it}\right) \end{aligned}$$

Dispersion measures uncertainty with the following model assumptions:

1. $Var\left(\eta_t|I_t\right) = 0$
2. $Var\left(\epsilon_{it}|I_{it}\right) = Var\left(\epsilon_{jt}|I_t\right)$ for all i, j, t.

Linear Forecasting with Forecast Dispersion

- Is there any relationship between forecast dispersion and model uncertainty?
- Same hidden state model with $I_{it} = \{M, y^t, z_i^t\}$.

$$z_{it} = y_{t+1} + \sigma_\xi \xi_{it} + \sigma_\varepsilon \varepsilon_t$$

- Calibrate σ_ξ and σ_ε to match forecast dispersion and average forecast error in the SPF.
- Findings
  - Generates forecast dispersion and avg forecast error (by construction). **But despite changes in U_t, no changes in dispersion!**
  - Gets close to corr(forecast, GDP): 71% in data 77% in model (30% baseline).
  - Lowers uncertainty (2.85%) and dampens the uncertainty shocks (0.12% std).

Forecast Bias Falls Over Time (Model and Data)

	Forecast Bias	
	Survey Data	Model Forecasts*
1968-1990	-0.67	-1.09
1991 - 2012	-0.22	-0.22

- Why does forecast bias shrink?
- Because forecasters learn and parameter uncertainty shrinks. When forecaster learns the true distribution, the forecast mean and true mean converge.

* Model results are for the skewed model with stochastic volatility.
Results for homoskedastic, skewed model: $-1.05 \rightarrow 0$ (faster learning).

The g-and-h Transformation

- Let z be a standard normal random variable.
- Then q has a g-and-h distribution if

$$q = \frac{1}{g}(\exp\{gz\} - 1)\exp\left\{\frac{1}{2}hz^2\right\}$$

where $g \neq 0$ and $h > 0$.

- g controls the skew.
- h is positively related to the kurtosis.