

## Effective Calibration

by

A. Ronald Gallant  
Department of Economics  
University of North Carolina  
Chapel Hill NC 27599-3305 USA

WEB: [www.unc.edu/~arg](http://www.unc.edu/~arg)  
Working papers (paper, slides)  
Browse ftp site (code, user's guide)  
FTP: [ftp.econ.duke.edu](ftp://ftp.econ.duke.edu/pub/arg/papers)  
pub/arg/papers (paper, slides)  
pub/arg/edf (code, user's guide)

1

### The Context

We have a structural model with parameters  $\rho$  that can be simulated and that has been rejected by a statistical procedure.

### The Problem

What is the best way to determine the parameters of this model from data?

### Why Do This?

The structural model is in agreement with the notions of a scientific discipline and therefore arguably more reliable out-of-sample at large distances from the edge of the data cloud than the reduced form or nonparametric model that defeated it in-sample.

2

### Example: Endowment Economy

Endowment process:

$$\log(e_t/e_{t-1}) = \alpha_0 + \alpha_1 \log(e_{t-1}/e_{t-2}) + \sigma_e u_t$$

Utility function:

$$\mathcal{E}_0 \left( \sum_{t=0}^{\infty} \beta^t \frac{e_t^{1-\gamma} - 1}{1-\gamma} \right),$$

Asset pricing equation:

$$v_t = \beta \mathcal{E}_t [(1 + v_{t+1}) (e_{t+1}/e_t)^{1-\gamma}]$$

where

- $\mathcal{E}_t$  is conditional expectation with respect to the present and past of the endowment process
- $v_t = p_t/e_t$  is the price dividend ratio which is determined by the asset pricing equation
- $p_t$  is the price of the equity security.

3

### Implied Returns

Implied gross equity return:

$$r_t^e = [(1 + v_t)/v_{t-1}] (e_t/e_{t-1})$$

Implied gross risk-free return:

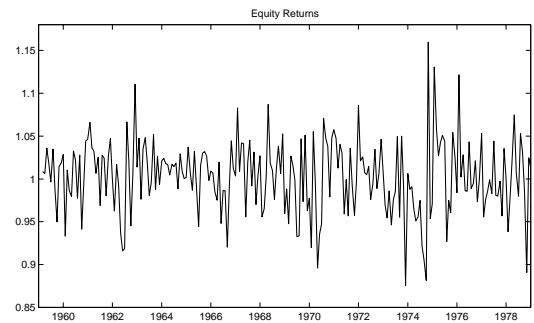
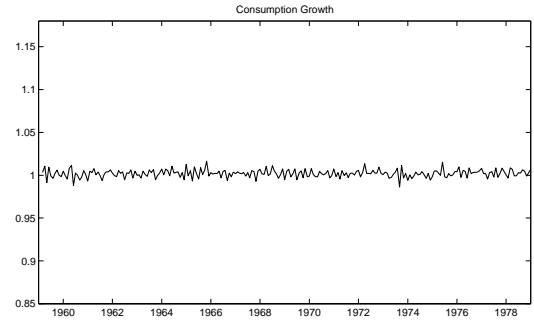
$$r_t^b = [\beta \mathcal{E}_t (e^{t+1}/e_t)^{-\gamma}]^{-1}$$

4

Table 1. Characteristics of the Data.

Quantiles	$\alpha_t / \alpha_{t-1}$	$r_t^e$
99%	1.013425	1.121526
95%	1.009645	1.057568
90%	1.007642	1.048265
75% Q3	1.004783	1.030098
50% Med	1.002235	1.005685
25% Q1	0.999164	0.978847
10%	0.996466	0.948152
5%	0.994561	0.932595
1%	0.991563	0.890846
IQR Q3-Q1	0.005619	0.051251
Mean	1.002060	1.002851
Std. Dev.	0.004573	0.042095
Annualized		
Med	2.715	7.039
IQR	1.946	17.754
Mean	2.500	3.475
Std. Dev.	1.584	14.582

Notes: The sampling frequency is monthly. Med is the median and IQR is the inter quartile range.



### Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

### Options Available

- Simulated Maximum Likelihood  
Durham and Gallant (2001)
- Asymptotic Equivalents of MLE
  - Indirect inference  
Gourieroux, Monfort and Renault (1993)
  - Efficient method of moments  
Gallant and Tauchen (2001)
- MCMC Bayes  
Eraker (2001)
- Simulated Method of Moments  
Duffie and Singleton (1993)

### Quasi Maximum Likelihood (1)

Density implied by structural model:  $f(y|\rho)$

True density of the data:  $f(y)$

Quasi MLE converges to the minimizer  $\rho^\#$  of Kullback-Liebler discrepancy

$$\int [\log f(y) - \log f(y|\rho)] f(y) dy.$$

No useful approximation properties flow from this measure of distance

### Quasi Maximum Likelihood (2)

Some standard norms are continuous with respect to Kullback-Liebler discrepancy, e.g.

$$\begin{aligned} & \int |f(y) - f(y|\rho)| dy \\ & \leq C \left[ \int [\log f(y) - \log f(y|\rho)] f(y) dy \right]^{1/2} \end{aligned}$$

Useful if Kullback-Liebler discrepancy tends to zero.

Of little value in if Kullback-Liebler discrepancy tends to a positive constant.

### Quasi Maximum Likelihood (3)

Structural models tend to be better behaved than actual data:  $f(y|\rho)$  will have thinner tails than  $f(y)$ .

When this happens, quasi maximum likelihood calibrations will be sensitive to observations at the edges of the data cloud.

### Quasi Maximum Likelihood (4)

It would seem from these observations that quasi maximum likelihood and its asymptotic equivalents should be avoided for the purpose of calibration.

## Bayes

The calibration problem as stated here is formally outside the Bayesian system of inference. The issues are as follows:

To bring the problem within the Bayesian system of inference, one must envisage an exhaustive collection of models of which the structural model under consideration is a member in order to make comparative probability statements and give probabilistic estimates of the reliability of an estimate.

Geweke (1999) suggests a reduced form model to augment the class of models so that the requisite encompassing is plausible. But probability statements relative to a reduced form model seem hardly in the spirit of calibration as envisaged here.

Without reduced form augmentation, one must envisage all possible candidate structural models that one would be willing to entertain, which is an impossible task.

The alternative to a collection of models that exhaust all possibilities within the Bayesian paradigm is to condition inference on a model believed to be false. If one actually does so, then calibration has the structural problems of lack of objectivity or instability, depending on how the prior is chosen, and the task of designing an MCMC algorithm.

13

## Method of Moments (1)

Method of moments converges to the minimizer  $\rho^\#$  of

$$\left\| \int \psi(y)f(y) dy - \int \psi(y)f(y|\rho) dy \right\|$$

where  $\|\cdot\|$  is some measure of distance and  $\psi$  some vector of moment functions, both selected subjectively by the calibrator.

The integral

$$\int |\psi(y)|f(y)dy$$

on which the method relies may not be finite.

14

## Method of Moments (2)

Method of moments calibration is not invariant to change in scale of measurement.

For example, suppose the structural model implies log gross returns are normally distributed with location parameter  $\mu$  and scale parameter  $\sigma$  whereas observed log returns data actually follow the Laplace distribution with location parameter  $\nu$  and scale parameter  $\tau$ .

If a match to first and second moments is done in log returns then, as sample size increases, the calibration converges to

$$\begin{aligned}\mu &= \nu \\ \sigma^2 &= 2\tau^2.\end{aligned}$$

15

## Method of Moments (3)

If the match is done in returns and  $\tau < \frac{1}{2}$  then the calibration converges to

$$\begin{aligned}\mu &= \nu - 2 \log(1 - \tau^2) + \frac{1}{2} \log(1 - 4\tau^2) \\ \sigma^2 &= 2 \log(1 - \tau^2) - \log(1 - 4\tau^2).\end{aligned}$$

If the match is done in returns and  $\tau \geq \frac{1}{2}$  then the second moment does not exist and the calibration fails.

16

### Method of Moments (4)

This criticism of method of moments would collapse immediately if one matched to a collection of bounded moments that had an objective rationale in a way that was invariant to change in scale of measurement. This observation motivates our approach.

17

### Rationale (1)

The step functions  $r_u(y) = I_{(-\infty, u]}(y)$ , where  $y \in \mathbb{R}^d$  and  $(-\infty, u]$  is a rectangle in  $\mathbb{R}^d$ , are a basis for all bounded continuous functions of the data.

Therefore, a natural calibration procedure would be to choose model parameters to match the moments of step functions for some values of the index  $u$ .

This is the same as matching the empirical distribution of the data  $\hat{F}_n(u)$  to the distribution implied by the model  $F(u|\rho)$ , where the latter may be computed by simulation.

18

### Rationale (2)

Matching to the entire collection  $\{r_u\}_{u \in \mathbb{R}^d}$  simultaneously can be accomplished by minimizing  $\|\hat{F}_n(u) - F(u|\rho)\|$  with respect to some norm  $\|\cdot\|$  or other distance function that is analogous to one of the standard norms.

The sup norm is not a good choice because simulation methods often introduce granularities that destroy its effectiveness. Some minimal smoothing is necessary. The two natural candidates that provide some smoothing are the  $L_1$  and  $L_2$  norms. Of them, the  $L_2$  is analytically more tractable and is also more traditional.

It must be a weighted  $L_2$  norm because distribution functions are not  $L_2$ -Lebesgue. The natural weighting function is the true density of the data.

19

### Rationale (3)

Because the true density of the data is unknown, one integrates with respect to the empirical distribution of the data instead. Integrating with respect to the empirical distribution is equivalent to averaging over the sample  $\{y_t\}_{t=1}^n$ .

These considerations suggest that the appropriate calibration  $\hat{\rho}_n$  is the minimizer of the sample objective function

$$s_n(\rho) = \frac{1}{n} \sum_{t=1}^n [\hat{F}_n(y_t) - F(y_t|\rho)]^2.$$

That is the proposal.

To give it a name, call it edf-calibration.

In the statistical literature it would be described as a minimum chi squared method where the objective function is a variant of the Cramer von Mises test statistic.

20

### Main Theoretical Results

Under weak regularity conditions that accommodate both time series and cross sectional data,  $\hat{\rho}_n$  tends to the parameter value  $\rho^o$  that minimizes

$$s^o(\rho) = \int [F(y) - F(y|\rho)]^2 f(y) dy,$$

where  $F(y)$  and  $f(y)$  are the distribution and density functions, respectively, that actually govern the data.

Under stronger conditions  $\sqrt{n}(\hat{\rho}_n - \rho^o)$  is asymptotically normal.

The stationary bootstrap is used to assess variability. Two sources of variability: sampling variability and computational instability.

21

### Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

22

### Notation (1)

To facilitate computations and proofs, an empirical distribution function is given a stylized representation.

$I(u)$  is the univariate distribution function that puts all its mass at zero; that is,

$$I(u) = I_{[0,\infty)}(u)$$

for  $u \in \mathbb{R}^1$ .

In terms of  $I(\cdot)$ , the empirical distribution function of a sequence of vectors in  $\mathbb{R}^d$

$$y_t = (y_{1,t}, \dots, y_{d,t}) \quad t = 1, \dots, n$$

is

$$\hat{F}_n(v) = \frac{1}{n} \sum_{t=1}^n \prod_{i=1}^d I(v_i - y_{it})$$

for  $v \in \mathbb{R}^d$ .

23

### Notation (2)

$F(v|\rho)$  is the distribution function implied by the structural model which is approximated by

$$F_N(v|\rho) = \frac{1}{N} \sum_{t=1}^N \prod_{i=1}^d I(v_i - \hat{y}_{it})$$

where

$$\hat{y}_t = (\hat{y}_{1,t}, \dots, \hat{y}_{d,t}) \quad t = 1, \dots, N$$

is a simulation from the structural model with parameters set to  $\rho$ .

24

### EDF-Calibration

$$s_n(\rho) = \frac{1}{n} \sum_{t=1}^n [\hat{F}_n(y_t) - F(y_t|\rho)]^2$$

$$\hat{\rho}_n = \operatorname{argmin}_{\rho \in R} s_n(\rho)$$

If the distribution function implied by the structural model is continuous, then edf-calibration is invariant to invertible, increasing, differentiable transformations that are applied elementwise.

Proof: Application of the change of variable formula for integration and summation coupled with the observation that such transformation do not change the value of  $I(y - u)$ .

25

### Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

26

### Computations (1)

Minimizing

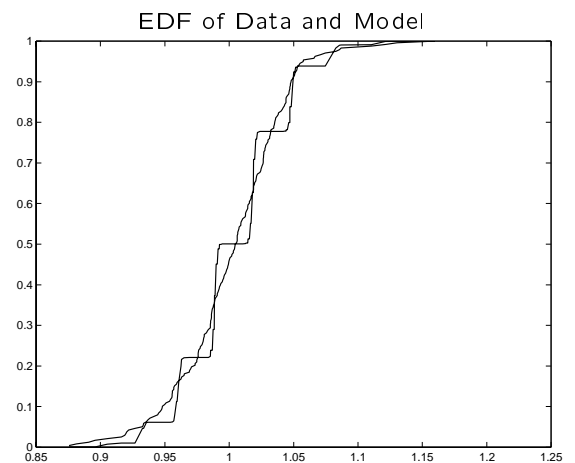
$$s_n(\rho) = \frac{1}{n} \sum_{t=1}^n [\hat{F}_n(y_t) - F_N(y_t|\rho)]^2$$

is not an easy problem because  $s_n(\rho)$  is not a smooth function of  $\rho$  due to the indicator functions that appear in the definition of  $\hat{F}_n(y)$  and  $F_N(y|\rho)$ .

27

### Computations (2)

The problem is harder still if the simulation strategy granularizes the data.



28

### Computations (3)

Smoothing the distributions  $\hat{F}_n(y)$  and  $F_N(y|\rho)$  prior to matching stabilizes the computations considerably. The smoothed problem is not difficult to solve numerically.

The solution to the smoothed problem can either be accepted as the solution to the original problem or used as a start value for solving the more difficult original problem.

The quadratic squasher is an effective smoother

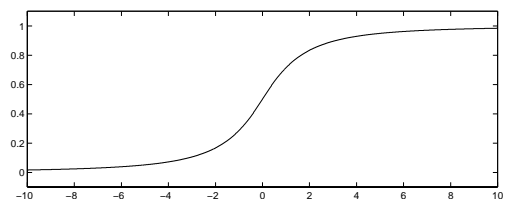
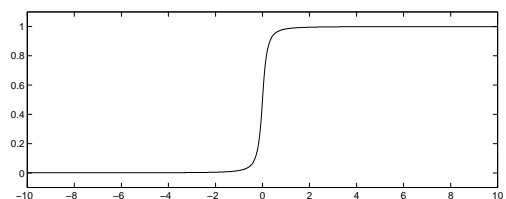
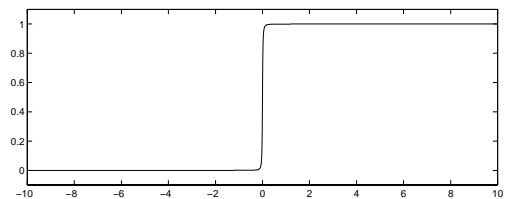
$$S(u) = \frac{u^2 + u|u| + 2u + 2|u| + 4}{2u^2 + 4|u| + 8}.$$

The scaled version is

$$S_\sigma(u) = S\left(\frac{u}{\sigma}\right).$$

29

### Quadratic Squasher, $\sigma = 0.01, 0.1, 1.0$



30

### Computations (4)

A smooth approximation to the empirical distribution  $\hat{F}_n(y)$  is

$$\hat{F}_{\sigma,n}(v) = \frac{1}{n} \sum_{t=1}^n \prod_{i=1}^d S_\sigma(v_i - y_{it}).$$

For a simulation  $\{\hat{y}_{it}\}_{t=1}^N$  from the structural model with parameters set to  $\rho$ , a smooth approximation to  $F_N(y|\rho)$  is

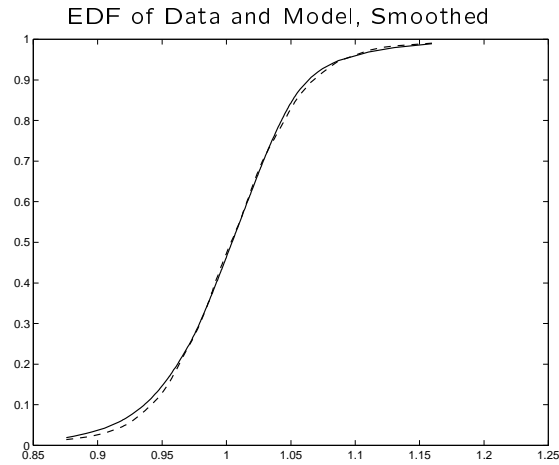
$$F_{\sigma,N}(v|\rho) = \frac{1}{N} \sum_{t=1}^N \prod_{i=1}^d S_\sigma(v_i - \hat{y}_{it}).$$

31

This slide intentionally left blank

32

### Computations (5)



33

### Computations (7)

The smoothed problem can be solved by standard nonlinear optimizers.

A method for solving the original problem is to evaluate  $s_n(\rho)$  over thousands of randomly selected points in a ball that has the smoothed solution as its center, accepting the minimum of these as the answer.

The unsmoothed objective function seems to be locally smooth because it pays to let the optimizer polish each of these randomly selected points with a few iterations.

34

### Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

35

### Regularity Conditions (1)

Data generating process is

- $\alpha$ -mixing of size  $-r/(r-2)$  for some  $r > 2$
- stationary with distribution  $F(y)$  and density  $f(y)$

Alpha mixing:

$$\alpha(\mathcal{F}, \mathcal{G}) = \sup_{F \in \mathcal{F}, G \in \mathcal{G}} |P(FG) - P(F)P(G)|$$

$$\alpha_m = \sup_t \alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+m}^\infty)$$

$$\alpha_m = \mathcal{O}\left(m^{-r/(r-2)-\delta}\right) \quad \text{for some } \delta > 0$$

36

### Glivenko-Cantelli Result

$$\lim_{n \rightarrow \infty} \sup_y |\hat{F}_n(y) - F(y)| = 0 \quad \text{a.s.}$$

Proof: Mixing implies a slln whence

$$\frac{1}{n} \sum_{t=1}^n r(y_t) \rightarrow \int r(u) dF(u)$$

for each  $r$  in

$$\mathcal{R} = \{r : r(y) = I_{(-\infty, u]}(y), u \in \mathbb{R}^d\}$$

An argument of Nobel and Olshen (1996, Appendix III) can be used to construct an  $\epsilon$ -bracketing for  $\mathcal{R}$ . Apply Theorem 2 of Pollard (1984, p. 8) to conclude that

$$\sup_{r \in \mathcal{R}} \left| \frac{1}{n} \sum_{t=1}^n r(y_t) - \int r(u) dF(u) \right| \rightarrow 0$$

37

### Regularity Conditions (2)

Structural model:

- stationary with distribution  $F(y|\rho)$  and density  $f(y|\rho)$
- $\rho$  restricted to a compact set  $R$
- smoothness conditions on  $F(y|\rho)$
- $\alpha$ -mixing of size  $-r/(r-2)$  for some  $r > 2$  for each  $\rho \in R$

38

### Calibration in Large Samples

$$\lim_{n \rightarrow \infty} \sup_{\rho \in R} |s_n(\rho) - s^o(\rho)| = 0$$

where

$$s^o(\rho) = \int [F(y) - F(y|\rho)]^2 f(y) dy.$$

and

$$\lim_{n \rightarrow \infty} \hat{\rho}_n = \rho^o$$

where

$$\rho^o = \operatorname{argmin}_{\rho \in R} s^o(\rho)$$

provided the minimum of  $s^o(\rho)$  is unique

Proof:

$$\begin{aligned} \sup_{\rho \in R} \left| s_n(\rho) - \frac{1}{n} \sum_{t=1}^n [F(y_t) - F(y_t|\rho)]^2 \right| \\ \leq \sup_y [\hat{F}_n(y) - F(y)]^2 + 4 \sup_y |\hat{F}_n(y) - F(y)| \end{aligned}$$

39

### Regularity Conditions (3)

Data generating process is

- $\beta$ -mixing of size  $-r/(r-2)$  for some  $r > 2$
- $k(u, v) = \left[ \prod_{i=1}^d I(v_i - u_i) - F(v|\rho^o) \right] \frac{\partial}{\partial \rho^i} F(v|\rho^o)$  is smooth and  $k^r(u, v)$  is dominated
- $\rho^o$  is an interior point of  $R$

Beta mixing:

$$\alpha_m \leq \beta_m = \sup_t \mathcal{E} \left\{ \sup_{F \in \mathcal{F}_{\infty}^*} |P(F|\mathcal{F}_{-m}^t) - P(F)| \right\},$$

40

### Asymptotic Normality

$\sqrt{n}(\hat{\rho}_n - \rho^o)$  is asymptotically normal

Proof: Sample and population first order conditions are

$$0 = \frac{\partial}{\partial \rho} s_n(\hat{\rho}_n) = \frac{-2}{n} \sum_{t=1}^n [\hat{F}_n(y_t) - F(y_t|\hat{\rho}_n)] \frac{\partial}{\partial \rho} F(y_t|\hat{\rho}_n)$$

$$0 = \frac{\partial}{\partial \rho} s^o(\rho^o) = -2 \int [F(y) - F(y|\rho^o)] \frac{\partial}{\partial \rho} F(y|\rho^o) dF(y)$$

After algebra, sample first order equation is

$$\mathcal{J}_n(\hat{\rho}_n) \sqrt{n}(\hat{\rho}_n - \rho^o) = \frac{1}{\sqrt{n}} A_n - \frac{2(n-1)}{n} \sqrt{n} U_n$$

where

$$U_n = \frac{1}{n(n-1)} \sum_{s \neq t} \frac{1}{2} [k(y_s, y_t) + k(y_t, y_s)]$$

A result of Yoshihara (1976) shows  $U_n$  is asymptotically normal with a mean that can be shown to be zero using the population first order condition. The rest is standard arguments from dynamic nonlinear models.

41

### Standard Errors

- The choice of simulation method may affect the accuracy of the asymptotics.
- There may be additional variation not reflected in the asymptotics due to numerical instability in computations.
- Use Politis and Romano (1994) bootstrap.

42

### Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

43

### Endowment Economy

Endowment process:

$$\log(e_t/e_{t-1}) = \alpha_0 + \alpha_1 \log(e_{t-1}/e_{t-2}) + \sigma_e u_t$$

Utility function:

$$\mathcal{E}_0 \left( \sum_{t=0}^{\infty} \beta^t \frac{e_t^{1-\gamma} - 1}{1-\gamma} \right),$$

Asset pricing equation:

$$v_t = \beta \mathcal{E}_t [(1 + v_{t+1}) (e_{t+1}/e_t)^{1-\gamma}]$$

- $v_t = p_t/e_t$  is the price dividend ratio
- $p_t$  is the price of the equity security.
- $r_t^e = [(1 + v_t)/v_{t-1}]/(e_t/e_{t-1})$  equity return.
- $r_t^b = [\beta \mathcal{E}_t (e^{t+1}/e_t)]^{-1}$  is the bond return

44

### Consumption Process

$$\log(c_t/c_{t-1}) = a_0 + a_1 \log(c_{t-1}/c_{t-2}) + \sigma_c z_t$$

The calibration to  $(c_t/c_{t-1}, c_{t-1}/c_{t-2}) \in \mathfrak{R}^2$  is

	$a_0$	$a_1$	$\sigma_c$
No Smoothing:	0.002578	-0.25929	0.004160
IQR:	0.000412	0.10255	0.000277
Smoothing:	0.002600	-0.25911	0.004277
IQR:	0.000372	0.11503	0.000268
Least Squares:	0.002547	-0.25908	0.004355
StdErr:	0.000309	0.06236	0.000414

45

### Model 1

$$\beta = 1\%, 2\%, 3\%, 4\%, 5\%$$

$$\gamma = \text{calibrated from } r_t^e$$

$$\alpha_0 = a_0 = 0.002600$$

$$\alpha_1 = a_1 = -0.25911$$

$$\sigma_e = \sigma_c = 0.004277$$

46

### Calibrated Parameters for Model 1

Time Preference Annual Rate	Parameter Calibration Data			
	Equity Returns $\gamma$	Equity Returns $\delta_n(\beta/n)$	Consumption Growth $a_0$	Consumption Growth $a_1$
1%	0.99917	3.16105 (2.7462)	3.385334	-0.2591061 (0.115025)
2%	0.99835	2.64983	3.462604	-0.2591061 (0.115025)
3%	0.99754	2.12672	3.532321	-0.2591061 (0.115025)
4%	0.99674	1.59152	3.593626	-0.2591061 (0.115025)
5%	0.99594	1.03804 (3.2180)	3.646292	-0.2591061 (0.115025)

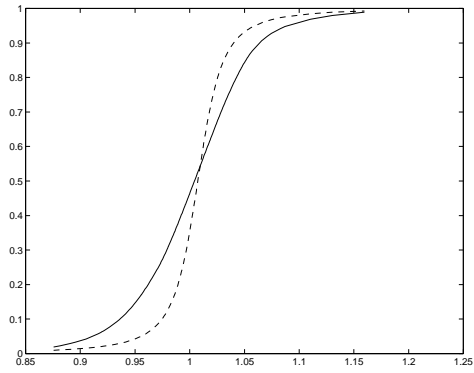
47

### Implied Returns for Model 1

Time Preference Annual Rate	Monthly Rate	Risk Aversion $\gamma$	Annualized Returns	
			Bond Std.Dev.	Equity Std.Dev.
1%	0.99917	3.16105	9.120	1.264
2%	0.99835	2.64983	8.846	1.060
3%	0.99754	2.12672	8.524	0.850
4%	0.99674	1.59152	8.152	0.636
5%	0.99594	1.03804	7.727	0.415

48

Model 1. EDF of Data and Model



Model 2

$$\beta = 1\%, 2\%, 3\%, 4\%, 5\%$$

$$\gamma = \text{calibrated from } r_t^e$$

$$\alpha_0 = a_0 = 0.002600$$

$$\alpha_1 = a_1 = -0.25911$$

$$\sigma_e = \text{calibrated from } r_t^e$$

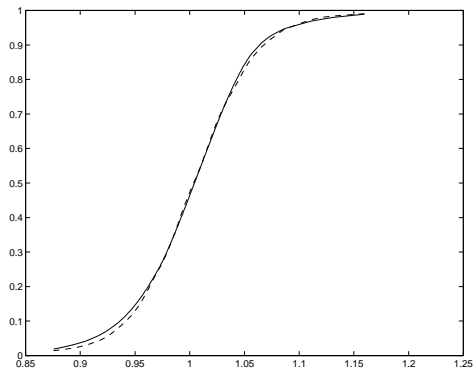
Calibrated Parameters for Model 2

Time Preference Annual Rate	Monthly $\beta$	Parameter Calibration Data			
		$\gamma$	Equity Returns $\sigma_e$	$s_h(\hat{\beta}_h)$	Consumption Growth $\alpha_1$
1%	0.99917	1.24340 (1.8785)	0.0355053 (0.018405)	0.019484	0.0026004 (0.000372)
2%	0.99835	0.79303	0.0399196	0.019750	0.0026004
3%	0.99754	0.46847	0.0435702	0.018653	0.0026004
4%	0.99674	0.21924	0.0465804	0.017875	0.0026004
5%	0.99594	0.00625 (0.9175)	0.0492922 (0.014159)	0.017543	0.0026004 (0.000372)

Implied Returns for Model 2

Time Preference Annual Rate	Monthly $\beta$	Risk Aversion $\gamma$	Annualized Returns		
			Bond	Equity	Std.Dev.
1%	0.99917	1.24340	3.035	4.109	5.102
2%	0.99835	0.79303	3.445	2.948	4.969
3%	0.99754	0.46847	3.960	1.901	4.963
4%	0.99674	0.21924	4.503	0.956	5.018
5%	0.99594	0.00625	5.019	0.029	5.044

Model 2. EDF of Data and Model



Model 3

$$\beta = 1\%, 2\%, 3\%, 4\%, 5\%$$

$$\gamma = 1.1$$

$$\alpha_0 = \text{calibrated from } r_t^e$$

$$\alpha_1 = a_1 = -0.25911$$

$$\sigma_e = \text{calibrated from } r_t^e$$

The bond return is 3.102 with a standard deviation of 3.772 and the equity return is 5.015 with a standard deviation of 13.606.

Calibrated Parameters for Model 3

Time Preference Annual Rate	Preference Monthly Rate	Risk Aversion	$\gamma$	Parameter Calibration Data		
				Equity Returns $\alpha_0$	Returns $\sigma_e$	Consump. $a_1$
1%	0.99917	1.10000		0.0028380 (0.004248)	0.036833 (0.00281)	0.019771 (0.115025)
2%	0.99835	1.10000		0.0018963	0.036834	0.019771
3%	0.99754	1.10000		0.0009692	0.036834	0.019771
4%	0.99674	1.10000		0.0000508	0.036835	0.019771
5%	0.99594	1.10000		-0.0008682 (0.004248)	0.036835 (0.00281)	0.019771 (0.115025)

Summary

Consumption, represented as a first order autoregression in log differences, does a poor job of serving as a single factor for equity returns at reasonable values of the agent's preference parameters for two reasons: the scale parameter is too small and the location parameter (growth rate) is too large.

Allowing the scale parameter to adjust freely allows for some plausible values for preference parameters and implied returns.

Allowing both to adjust allows preference parameters to be chosen at will.

## Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

57

## Alternative Computational Strategy

Endowment process:

$$\log(e_t/e_{t-1}) = \alpha_0 + \alpha_1 \log(e_{t-1}/e_{t-2}) + \sigma_e u_t$$

Asset pricing equation:

$$v_t = \beta \mathcal{E}_t [(1 + v_{t+1}) (e_{t+1}/e_t)^{1-\gamma}]$$

To implement edf one is obliged to generate a long simulation of the endowment process

$$\{e_t/e_{t-1}\}_{t=1}^N \quad N \sim 50,000$$

Use the simulation in connection with reasonable basis functions

$$\mathcal{P}(c) = \alpha_0 + \alpha_1 \phi_1(c) + \alpha_2 \phi_2(c) + \alpha_3 \phi_2(c)$$

to solve the asset pricing equation by GMM:

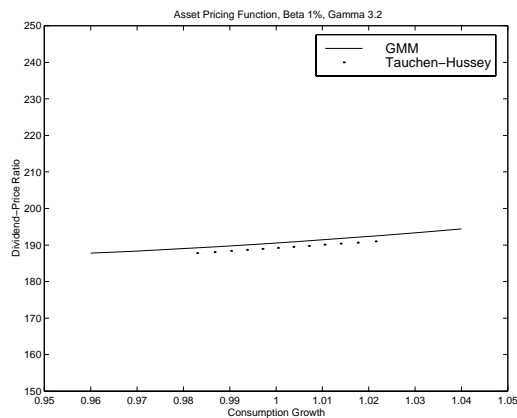
Unconditional Moment Condition:

$$0 = \mathcal{E}(\mathcal{P}(e_t/e_{t-1}) - \beta \{[1 + \mathcal{P}(e_{t+1}/e_t)] (e_{t+1}/e_t)^{1-\gamma}\})$$

Instruments:

$$Z = [1, \phi_1(e_t/e_{t-1}), \phi_2(e_t/e_{t-1}), \phi_2(e_t/e_{t-1})]$$

58



59

## Performance, Tauchen-Hussey vs GMM:

- Tauchen-Hussey simulations are cheaper to compute.
- But optimizations are less stable.
- So that many more start values required.
- Therefore, not much difference in total computational time.

60

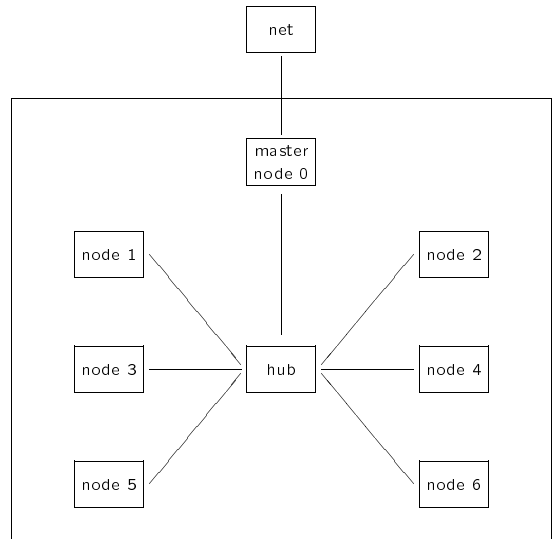
Calibrated Parameters for Model 1, No S-moothing.

Time Preference Annual Rate	Monthly $\beta$	Parameter Calibration Data				
		Equity Returns $\gamma$	Returns $s_n(\hat{p}_n)$	Consumption Growth $a_0$	$a_1$	$\sigma_c$
1%	0.99917	7.40528	9.576138	0.0026004	-0.2591061	0.0042770
2%	0.99835	7.12381	9.987357	0.0026004	-0.2591061	0.0042770
3%	0.99754	6.78966	10.403083	0.0026004	-0.2591061	0.0042770
4%	0.99674	6.52332	10.825775	0.0026004	-0.2591061	0.0042770
5%	0.99594	6.25214	11.253684	0.0026004	-0.2591061	0.0042770
6%	0.99516	5.98160	11.678955	0.0026004	-0.2591061	0.0042770
7%	0.99438	5.69407	12.110277	0.0026004	-0.2591061	0.0042770
8%	0.99361	5.50729	12.546775	0.0026004	-0.2591061	0.0042770
9%	0.99284	5.19477	12.990102	0.0026004	-0.2591061	0.0042770
10%	0.99209	4.93249	13.428356	0.0026004	-0.2591061	0.0042770

Calibrated Parameters for Model 1, S-moothing,  $\sigma = 0.01$ .

Time Preference Annual Rate	Monthly $\beta$	Parameter Calibration Data				
		Equity Returns $\gamma$	Returns $s_n(\hat{p}_n)$	Consumption Growth $a_0$	$a_1$	$\sigma_c$
1%	0.99917	3.16105	3.385334	0.0026004	-0.2591061	0.0042770
2%	0.99835	2.64983	3.462604	0.0026004	-0.2591061	0.0042770
3%	0.99754	2.12672	3.532321	0.0026004	-0.2591061	0.0042770
4%	0.99674	1.59152	3.593626	0.0026004	-0.2591061	0.0042770
5%	0.99594	1.03804	3.646292	0.0026004	-0.2591061	0.0042770
6%	0.99516	0.48247	3.688274	0.0026004	-0.2591061	0.0042770
7%	0.99438	-0.08586	3.720079	0.0026004	-0.2591061	0.0042770
8%	0.99361	-0.65554	3.740774	0.0026004	-0.2591061	0.0042770
9%	0.99284	-1.22972	3.750362	0.0026004	-0.2591061	0.0042770
10%	0.99209	-1.78924	3.748831	0.0026004	-0.2591061	0.0042770

Linux Cluster



Common configuration is 2 CPU's, 1MB memory, 80GB disk per node.

Calibrated Parameters for Model 1, S-moothing,  $\sigma = 0.001$ .

Time Preference Annual Rate	Monthly $\beta$	Parameter Calibration Data				
		Equity Returns $\gamma$	Returns $s_n(\hat{p}_n)$	Consumption Growth $a_0$	$a_1$	$\sigma_c$
1%	0.99917	7.21274	9.101063	0.0026004	-0.2591061	0.0042770
2%	0.99835	6.91029	9.489556	0.0026004	-0.2591061	0.0042770
3%	0.99754	6.61139	9.881308	0.0026004	-0.2591061	0.0042770
4%	0.99674	6.31336	10.275778	0.0026004	-0.2591061	0.0042770
5%	0.99594	6.01606	10.677502	0.0026004	-0.2591061	0.0042770
6%	0.99516	5.72516	11.075821	0.0026004	-0.2591061	0.0042770
7%	0.99438	5.43437	11.480594	0.0026004	-0.2591061	0.0042770
8%	0.99361	5.14717	11.886475	0.0026004	-0.2591061	0.0042770
9%	0.99284	4.85869	12.298401	0.0026004	-0.2591061	0.0042770
10%	0.99209	4.57345	12.705093	0.0026004	-0.2591061	0.0042770

### Coding Strategies

- *Shell Scripts*. The interconnected nodes are secure so that rcp, rsh, etc. do not need to be disabled. Some programs, such as nonlinear optimizers that use multiple, random starts, are so embarrassingly parallelizable, that parallelization can be done with shell scripts alone.
- *Message Passing Interface (MPI)*. The industry-standard protocol for implementing parallel processing. Allows communication among processes running on different processors. Architecture independent: Code written for a Linux cluster will run on multiple-processor, shared-memory machines. The recommended strategy.
  - <http://www.mpi-forum.org> MPI reference
  - <http://www.mpi.nd.edu/lam/downloads> software
  - <http://www.hku.hk/cc/sp2/ftp/mpiguide.ps> Fortran
  - <ftp://math.usfca.edu/pub/MPI/mpi.guide.ps> C & C++
- *Parallel Virtual Machine (PVM)*. Syntax similar to MPI. Runs on heterogeneous architectures. Can be installed and run on a user's account without requiring the user to have administrative privileges.

65

### Coding Strategies (Continued)

- *Parallelized Libraries*. Allows sequential code to have some of the benefits of parallelism. Works best on shared-memory, multiple-processor machines. Can actually impede performance if coupled with MPI or PVM.
- *High Performance Fortran*. A sort of hybrid of the strategies above, allows both threads and message passing. Worked poorly for us. The leading vendor is Portland Group whose compilers are of such low quality that they cannot compile LAPACK properly.

66

### Outline

- Rationale
- EDF-Calibration
- Computations
- Asymptotics
- Application
- More on Computations

67