Accuracy tests

Wouter J. Den Haan
London School of Economics

How to check for accuracy

1. Informal accuracy tests
2. Formal accuracy tests

Informal accuracy tests

These are possibly more important than formal ones

1. "Play" with your model/algorithm
   1. Understand properties of the model
   2. Change parameter values and understand how model properties change
   3. Open up the black box

Informal accuracy tests

2. Solve your model in a different way
    1. Linear instead of log-linear
    2. Use model equations to substitute out variables
    3. Approximate something else
        - c_t instead of k_{t+1}
        - $c^{-\gamma}$ instead of c_t

Formal accuracy tests

1. Euler-equation errors
   - require numerical integration (but this is not that difficult to do)
2. Dynamic Euler-equation errors
   - also requires numerical integration
3. Welfare measures (be careful)
4. DenHaan-Marcet (DHM) accuracy test
   - simple, but hard to interpret

Idea behind most accuracy tests

Model:

$$\mathsf{E}\left[f(x_{t-1}, x_t, y_t, y_{t+1}) | I_t\right] = 0$$

where $\mathsf{E}[f(\cdot)|I_t]$ is the Euler-equation errror

Accuracy tests:

- Euler-equation error: $\mathsf{E}[f(\cdot)|I_t]$ should be zero at *many* points in state space

Euler-eq. error & standard growth model

$$f_t = -c_t^{-\gamma} + \beta c_{t+1}^{-\gamma}(\alpha \exp(z_{t+1})k_t^{\alpha-1} + 1 - \delta)$$

with

$$z_{t+1} = \rho z_t + \sigma e_{t+1}$$

Euler-equation errors

- True solution satisfies

$$\mathsf{E}\left[f(x_{t-1}, x_t, y_t, y_{t+1}) | I_t\right] = 0$$

  for *all* points in the state space

- This can be checked for *any* numerical solution (including perturbation solutions) at *many* points in the state space

How to deal with integration?

- Easy if shocks have discrete support
- Numerical integration
  (this must be done accurately)

Growth model with discrete innovations

$$\max_{\{c_t,k_t\}_{t=1}^{\infty}} \mathsf{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \frac{c_t^{1-\gamma}-1}{1-\gamma}$$

$$\text{s.t.} \quad c_t + k_t = \exp(z_t) k_{t-1}^{\alpha} + (1-\delta) k_t \tag{1}$$

$$z_t = \rho z_{t-1} + \sigma e_t, \tag{2}$$

$$e_t = \begin{cases} +1 \text{ with probability } 1/2 \\ \text{-1 with probability } 1/2 \end{cases}$$

Basic idea

1. Construct *fine* grid with values for k_{-1} and z

2. Euler-equation error at (k_{-1}, z) equals

$$
\begin{aligned}
&-\boldsymbol{c}(k_{-1}, z)^{-\gamma} \\
&+0.5 * \beta \boldsymbol{c}(\boldsymbol{k}, \rho z + \sigma)^{-\gamma}(\alpha \exp(\rho z + \sigma)\boldsymbol{k}^{\alpha-1} + 1 - \delta) \\
&+0.5 * \beta \boldsymbol{c}(\boldsymbol{k}, \rho z - \sigma)^{-\gamma}(\alpha \exp(\rho z - \sigma)\boldsymbol{k}^{\alpha-1} + 1 - \delta)
\end{aligned}
$$

with $\boldsymbol{k} = \boldsymbol{k}(k_{-1}, z)$

When is a solution accurate

- When Euler eq. errors are small at many points
- Problem: magnitude of errors is hard to interpret

Interpretable Euler-equation errors

- At each grid point calculate *two* consumption values
  1. $\boldsymbol{c}(k_{-1}, z)$ using the numerical approximation
  2. implied value, $\boldsymbol{c}_{imp}(k_{-1}, z)$, using

$$\boldsymbol{c}_{imp}(k_{-1}, z) = g^{-1/\gamma}$$

with

$$g = \begin{array}{l} +0.5 * \beta \boldsymbol{c}(\boldsymbol{k}, z, +\sigma)^{-\gamma}(\alpha \exp(\rho z + \sigma)\boldsymbol{k}^{\alpha-1} + 1 - \delta) \\ +0.5 * \beta \boldsymbol{c}(\boldsymbol{k}, z, -\sigma)^{-\gamma}(\alpha \exp(\rho z - \sigma)\boldsymbol{k}^{\alpha-1} + 1 - \delta) \end{array}$$

that is, value implied by accurately calculated RHS of Euler equation

Interpretable Euler-equation errors

- Euler-equation error is equal to

$$\left| \frac{\boldsymbol{c}(k_{-1}, z) - \boldsymbol{c}_{imp}(k_{-1}, z)}{\boldsymbol{c}_{imp}(k_{-1}, z)} \right|$$

What to do with the errors?

- Calculate maximum and average of the errors
- Investigate
  - Pattern (e.g., are errors always of the same sign)
  - Are nodes with largest errors very likely?
  - What happens at nodes with largest errors?
    For example, if consumption is very small at those nodes, then small basically irrelevant errors may show up as large percentage errors

Growth model with continuous support

$$\max_{\{c_t,k_t\}_{t=1}^{\infty}} \mathsf{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \frac{c_t^{1-\gamma}-1}{1-\gamma}$$

$$\text{s.t.} \quad c_t + k_t = \exp(z_t) k_{t-1}^{\alpha} + (1-\delta) k_t$$

$$z_t = \rho z_{t-1} + \sigma e_t,$$

$$e_t \sim N(0,1)$$

Calculate conditional expectation

- Given are k_{-1}, z, and policy function $g(k_{-1},z)$
- $\delta = 1$ to simplify notation

Calculate conditional expectation

Use $k = g(k_{-1}, z)$ to get

$$
\begin{aligned}
& \mathrm{E}\left[\frac{\beta \exp(z_{+1})\alpha k^{\alpha-1}}{c_{t+1}}\right] \\
= \quad & \mathrm{E}\left[\frac{\beta \exp(z_{+1})\alpha k^{\alpha-1}}{\exp(z_{+1})k^{\alpha} - k_{+1}}\right] \\
= \quad & \mathrm{E}\left[\frac{\beta \exp(z_{+1})\alpha k^{\alpha-1}}{\exp(z_{+1})k^{\alpha} - g(k, z_{+1})}\right] \\
= \quad & \mathrm{E}\left[\frac{\beta \exp(\rho z + \sigma\varepsilon_{+1})\alpha k^{\alpha-1}}{\exp(\rho z + \sigma\varepsilon_{+1})k^{\alpha} - g(k, \rho z + \sigma\varepsilon_{+1})}\right]
\end{aligned}
$$

Conditional expectation

$$\mathsf{E}\left[\frac{\beta\exp(\rho z+\sigma\varepsilon_{+1})\alpha k^{\alpha-1}}{\exp(\rho z+\sigma\varepsilon_{+1})k^{\alpha}-g(k,\rho z+\sigma\varepsilon_{+1})}\right]$$

$$=\int_{-\infty}^{\infty}\frac{\beta\exp(\rho z+\sigma\varepsilon_{+1})\alpha k^{\alpha-1}}{\exp(\rho z+\sigma\varepsilon_{+1})k^{\alpha}-g(k,\rho z+\sigma\varepsilon_{+1})}\frac{\exp(-0.5\varepsilon_{+1}^{2})}{\sqrt{2\pi}}d\varepsilon_{+1}$$

$$=\int_{-\infty}^{\infty}\frac{\beta\exp(\rho z+\sigma\sqrt{2}\tilde{\varepsilon}_{+1})\alpha k^{\alpha-1}}{\exp(\rho z+\sigma\sqrt{2}\tilde{\varepsilon}_{+1})k^{\alpha}-g(k,\rho z+\sigma\sqrt{2}\tilde{\varepsilon}_{+1})}\frac{\exp(-\tilde{\varepsilon}_{+1}^{2})}{\sqrt{\pi}}d\tilde{\varepsilon}_{+1}$$

where $\varepsilon_{+1}=\tilde{\varepsilon}_{+1}\sqrt{2}$ and the Jacobian, $\sqrt{2}$, is used when implementing the change in variables

Hermite Gaussian Quadrature

$$
\begin{aligned}
x &\sim N(\mu, \sigma^2) \\
\mathsf{E}\left[H(x)\right] &\approx \sum_{j=1}^{J} \left(\frac{H\left(\mu + \sqrt{2}\sigma\zeta_j\right)\omega_j}{\sqrt{\pi}} \right)
\end{aligned}
$$

Hermite Gaussian Quadrature

$$\int_{-\infty}^{\infty} \frac{\beta \exp(\rho z + \sigma\sqrt{2}\tilde{\varepsilon}_{+1})\alpha k^{\alpha-1}}{\exp(\rho z + \sigma\sqrt{2}\tilde{\varepsilon}_{+1})k^{\alpha} - g(k, \rho z + \sigma\sqrt{2}\tilde{\varepsilon}_{+1})} \frac{\exp(-\tilde{\varepsilon}_{+1}^2)}{\sqrt{\pi}} d\tilde{\varepsilon}_{+1}$$

$$\approx$$

$$\sum_{j=1}^{J} \frac{\beta \exp(\rho z + \sigma\sqrt{2}\zeta_j)\alpha k^{\alpha-1}}{\exp(\rho z + \sigma\sqrt{2}\zeta_j)k^{\alpha} - g(k, \rho z + \sigma\sqrt{2}\zeta_j)} \frac{1}{\sqrt{\pi}} \omega_j$$

Euler-equation errors - Pros & Cons

1. Pro: if checked at *fine* grid then close to definition solution
2. Con: only tests for *one-period* ahead forecast errors; ignores possibility of accumulation of small errors
    - Dynamic Euler-equation error could pick those up
    - DHM statistic could pick those up

Dynamic Euler-equation errors

- Generate time series for z_t and choose k_0
- Generate two time paths for endogenous variables c_t and k_t
  1. generate time series for c_t & k_t with numerical approximation
  2. generate *alternative* series doing the following in each period
     - use numerical approx. to calculate cond. expect. accurately
     - use this conditional expectation to calculate implied consumption value
     - get capital from this implied consumption value & budget constraint
     - (numerical approximation only used to calculate cond. expect.)

Details of step 2

1. Generate time series for z_t and set $k_{imp,0} = k_0$
2. Given values of z_t and k_t: calculate conditional expectation $(= g_t)$ exactly as with regular Euler-eq. errors.
   (Thus use your numerical solution to evaluate choices inside the integral)
3. Calculate $c_{imp,t} = g^{-1/\gamma}$
4. Calculate $k_{imp,t} = z_t k^{\alpha}_{imp,t-1} + (1-\delta)k_{imp,t-1} - c_{imp,t}$

Welfare-based accuracy tests

- Be careful
- Welfare loss of using $k_t = k_{ss}\ \forall t$, instead of the optimal policy function is relatively small
  $\Longrightarrow$ different approximations can be similar in terms of welfare

DHM Accuracy test

$$\mathsf{E}\left[f(x_{t-1}, x_t, y_t, y_{t+1}) | I_t\right] = 0$$

$$\Longrightarrow$$

$$\mathsf{E}\left[f(x_{t-1}, x_t, y_t, y_{t+1}) h(s_t) | I_t\right] = 0$$

$$\Longrightarrow$$

$$\mathsf{E}\left[f(x_{t-1}, x_t, y_t, y_{t+1}) h(s_t)'\right] = 0$$

for any $s_t \in I_t$ and any measurable function $h(\cdot)$

Use simulated data to test

$$\frac{\sum_{t=1}^{T} f(x_{t-1}, x_t, y_t, y_{t+1}) h(s_t)'}{T} \approx 0$$

Simple DHM Accuracy test

1. Calculate $\bar{u}$, the average of

$$u_t = c_t^{-\gamma} - \beta c_{t+1}^{-\gamma}(\alpha \exp(z_{t+1}) k_t^{\alpha-1} + 1 - \delta)$$

2. Calculate how much this error would change steady state consumption

$$\begin{aligned} c^{-\gamma} &= \bar{u} + c_{ss}^{-\gamma} \\ c &= \left(\bar{u} + c_{ss}^{-\gamma}\right)^{-1/\gamma} \end{aligned}$$

3. Express error as fraction of steady state value

$$\frac{c - c_{ss}}{c_{ss}}$$

Formal DHM Accuracy test

1. Simulate sample of T obs. (Say $T = 3,500$ & discard 500)
2. Calculate

$$J_T = TM_T' W_T^{-1} M_T$$

$$M_T = \frac{\sum_{t=1}^{T} h(s_t) f(x_{t-1}, x_t, y_t, y_{t+1})}{T}$$

$$W_T = \frac{\sum_{t=1}^{T} f(x_{t-1}, x_t, y_t, y_{t+1}) h(s_t)' h(s_t) f(x_{t-1}, x_t, y_t, y_{t+1})}{T}$$

Formal DHM Accuracy test

- J_T has a χ^2 distribution with n_h degrees of freedom
- If $h(s_t)$ is a scalar, then

$$J_T = \left(\frac{M_T}{\sqrt{W_T/T}} \right)^2$$

Implementation of DHM statistic

1. Do the DHM statistic N times
2. Check the fraction of times the statistic is in the lower and upper 5% range; inaccurate solutions are typically blown away (because of having too many realizations in the upper critical region)
3. Personally, I prefer to do the test multiple times for scalar $h(s_t)$ because this provides more information. In fact, using $h(s_t) = 1$ can already be quite informative

Limits of DHM statistic

1. Even accurate solutions are rejected more often than 5% for high enough T; thus the higher the value of T for which you get good results the better
2. Results are random so inaccurate solutions could get through by sheer chance
3. The opposite of #2 turns out to be a bigger problem in practice: DHM is often difficult to pass in the sense that solutions that in many aspects are close to the true or an extremely accurate solution can fail the DHM statistic miserably

Example - Matching model

Household side

$$\max_{\{C_t,K_t\}_{t=1}^{\infty}} \mathsf{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \frac{C_t^{1-\gamma}-1}{1-\gamma}$$

$$\text{s.t. } C_t + K_t = W_t N_{t-1} + R_t K_{t-1} + (1-\delta)K_{t-1} + P_t \qquad (3)$$

$$N_t = (1-\rho^x)N_{t-1} + M_t \qquad (4)$$

Household takes the number of "matches", M_t, the wage rate, W_t, the rental rate R_t, and profits, P_t, as given.
FOC

$$C_t^{-\gamma} = \mathsf{E}_t\left[\beta C_{t+1}^{-\gamma}(R_{t+1}+1-\delta)\right]$$

Matching model example

Problem for firm matched with worker

$$\max_{k_t} z_t k_t^{\alpha} - W_t - R_t k_t$$

FOC:

$$R_t = \alpha z_t k_t^{\alpha - 1}$$

Firm-level profits are (at optimal k) equal to

$$p_t = (1 - \alpha) z_t k_t^{\alpha} - W_t$$

Wages are given by the following rule

$$W_t = (1 - \omega_0) \times [\omega_1 * p_t + (1 - \omega_1)\bar{p}]$$

where $\bar{p}$ are steady state level profits. Wages are completely sticky if ω_1 is equal to 0.

Matching model example

Free entry

posting cost = prob of success × value if success

$$\psi = \frac{M_t}{V_t} g_t$$

$$g_t = \mathsf{E}_t \left[\beta \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} p_{t+1} + (1 - \rho^x) g_{t+1} \right]$$

Matching model example

Matching technology

$$M_t = \frac{U_t V_t}{\left(U_t^{\xi} + V_t^{\xi}\right)^{1/\xi}}$$

with

$$U_t = 1 - N_{t-1}$$

Matching model example

Equilibrium

Equilibrium in the rental market

$$K_{t-1} = N_{t-1} k_t$$

profits transferred to households

$$P_t = N_{t-1} p_t - \psi V_t$$

Equations: Household

$$C_t^{-\gamma} = \mathsf{E}_t \left[\beta C_{t+1}^{-\gamma} (R_{t+1} + 1 - \delta) \right]$$

```
exp(-nu*c)=dfactor*exp(-nu*c(+1))*(exp(r(+1))+1-delta)
```

$$C_t + K_t + \psi V_t = z_t K_{t-1}^{\alpha} N_{t-1}^{1-\alpha} + (1-\delta) K_{t-1} \text{ or}$$

$$C_t + I_t + \psi V_t = Y_t,\ Y_t = z_t K_{t-1}^{\alpha} N_{t-1}^{1-\alpha},\ I_t = K_t - (1-\delta) K_{t-1}$$

```
exp(c)+exp(i)+pcost*exp(v)=exp(y)
exp(k)=(1-delta)*exp(k(-1))+exp(i)
y=varz+alpha*k(-1)+(1-alpha)*n(-1)
```

Equations: Matching

$$N_t = (1-\rho^x)N_{t-1} + \frac{U_t V_t}{\left(U_t^\xi + V_t^\xi\right)^{1/\xi}}$$

```
exp(n)
=(1-rox)*exp(n(-1))+exp(u+v)
/((exp(u*etam)+exp(v*etam))^(1/etam))
```

$$U_t = 1 - N_{t-1}$$

exp(u)=1-exp(n(-1))

Equations: rental rate & productivity

$$R_t = \alpha z_t k_t^{\alpha-1}$$

```
r=log(alpha)+varz+(alpha-1)*(k(-1)-n(-1))
```

$$\ln(z_t) = \rho \ln(z_{t-1}) + \varepsilon_t$$

```
varz=rho*varz(-1)+e
```

Equations: free entry

$$\psi = \frac{M_t}{V_t} g_t$$

```
pcost=
exp(eta)*exp(u)/((exp(u*etam)+exp(v*etam))^(1/etam))
```

$$g_t = \mathsf{E}_t \left[\beta \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} p_{t+1} + (1 - \rho^x) g_{t+1} \right]$$

```
exp(eta)=
dfactor*(exp(c(+1))/exp(c))^(-nu)
*(exp(prof(+1))+(1-rox)*exp(eta(+1)))
```

$$p_t = (1-\alpha)z_t k_t^{\alpha} - W_t$$

```
prof
=
log(
(1-omega1*omega0)*(1-alpha)*exp(varz+alpha*(k(-1)-n(-1))
-(1-omega1)*omega0*profitss
)
```

System

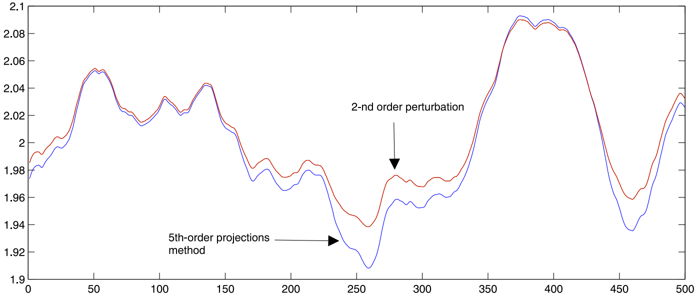
11 equations in 11 unknowns:

- $N_t, g_t, V_t, C_t, K_t, R_t, U_t, p_t, \ln(z_t), Y_t, I_t$
- n, eta, v, c, k, r, u, prof, y, varz, i

Accuracy errors

	2-nd order perturbation	5-th order projections
Capital Euler equation		
average	0.034%	0.026%
max	0.34%	0.33%
Employment Euler equation		
average	0.89%	0.004%
max	2.31%	0.086%

Log capital level



Log capital level