Solving Asset Pricing Models: A Statistical Approach

Walt Pohl

University of Zurich
Quantitative Business Administration

July 31, 2013

Agenda

The purpose of this talk is two-fold:

Agenda

The purpose of this talk is two-fold:

- Introduce a new general-purpose estimation method for structural models, the *empirical projection method*.

Agenda

The purpose of this talk is two-fold:

- Introduce a new general-purpose estimation method for structural models, the *empirical projection method*.
- Demonstrate the method on several asset-pricing models.

Structural Models

A *structural model* is a model given in terms of Euler equations.

Structural Models

A *structural model* is a model given in terms of Euler equations. These are conditional expectations,

$$E_t(F(x_t, x_{t+1})) = 0.$$

that depend on some parameters.

Structural Models

A *structural model* is a model given in terms of Euler equations. These are conditional expectations,

$$E_t(F(x_t, x_{t+1})) = 0.$$

that depend on some parameters.

Stock example:

Structural Models

A *structural model* is a model given in terms of Euler equations. These are conditional expectations,

$$E_t(F(x_t, x_{t+1})) = 0.$$

that depend on some parameters.

Stock example: The price P_t of a risky asset that pays a random dividend D_t.

Structural Models

A *structural model* is a model given in terms of Euler equations. These are conditional expectations,

$$E_t(F(x_t, x_{t+1})) = 0.$$

that depend on some parameters.

Stock example: The price P_t of a risky asset that pays a random dividend D_t.

Asset prices P_t must satisfy an equation of the form

Structural Models

A *structural model* is a model given in terms of Euler equations. These are conditional expectations,

$$E_t(F(x_t, x_{t+1})) = 0.$$

that depend on some parameters.

Stock example: The price P_t of a risky asset that pays a random dividend D_t.

Asset prices P_t must satisfy an equation of the form

$$E_t(P_t - m_{t+1}(P_{t+1} + D_{t+1})) = 0,$$

Structural Models

A *structural model* is a model given in terms of Euler equations. These are conditional expectations,

$$E_t(F(x_t, x_{t+1})) = 0.$$

that depend on some parameters.

Stock example: The price P_t of a risky asset that pays a random dividend D_t.

Asset prices P_t must satisfy an equation of the form

$$E_t(P_t - m_{t+1}(P_{t+1} + D_{t+1})) = 0,$$

where m_{t+1} is a function of some other variables, such as aggregate consumption, C_t.

Comparing Structural Models to the Data

We want to compare structural models to the data.

Comparing Structural Models to the Data

We want to compare structural models to the data.

Unfortunately, we can't directly measure conditional expectations, only unconditional ones.

Comparing Structural Models to the Data

We want to compare structural models to the data.

Unfortunately, we can't directly measure conditional expectations, only unconditional ones.

There are two standard solutions to this problem:

Comparing Structural Models to the Data

We want to compare structural models to the data.

Unfortunately, we can't directly measure conditional expectations, only unconditional ones.

There are two standard solutions to this problem:

- Parametric approaches.

Comparing Structural Models to the Data

We want to compare structural models to the data.

Unfortunately, we can't directly measure conditional expectations, only unconditional ones.

There are two standard solutions to this problem:

- Parametric approaches.
- GMM.

Comparing Structural Models to the Data

We want to compare structural models to the data.

Unfortunately, we can't directly measure conditional expectations, only unconditional ones.

There are two standard solutions to this problem:

- Parametric approaches.
- GMM.

I propose a third.

Parametric Approaches

Parametric approaches involve:

Parametric Approaches

Parametric approaches involve:

- Completely specifying the underlying distribution of the variables in the model.

Parametric Approaches

Parametric approaches involve:

- Completely specifying the underlying distribution of the variables in the model.
- This permits solving the conditional moment equation in terms of the other variables.

Parametric Approaches

Parametric approaches involve:

- Completely specifying the underlying distribution of the variables in the model.
- This permits solving the conditional moment equation in terms of the other variables.

So for example if we specified the distributions of D_t and C_t, we could compute P_t as a function of C_t and D_t.

Parametric Approaches

Parametric approaches involve:

- Completely specifying the underlying distribution of the variables in the model.
- This permits solving the conditional moment equation in terms of the other variables.

So for example if we specified the distributions of D_t and C_t, we could compute P_t as a function of C_t and D_t.

Parameters are then chosen to either match specific moments, or maximize a likelihood.

Parametric Approaches: Pros and Cons

Parametric approaches has one advantage:

Parametric Approaches: Pros and Cons

Parametric approaches has one advantage:

The model makes *definite predictions*. For example, the stock model predicts what the actual stock price, which we can compare with the data. It predicts various moments, etc.

Parametric Approaches: Pros and Cons

Parametric approaches has one advantage:

The model makes *definite predictions*. For example, the stock model predicts what the actual stock price, which we can compare with the data. It predicts various moments, etc.

And one big disadvantage:

Parametric Approaches: Pros and Cons

Parametric approaches has one advantage:

The model makes *definite predictions*. For example, the stock model predicts what the actual stock price, which we can compare with the data. It predicts various moments, etc.

And one big disadvantage:

The answer depends on the choice of distributions.

GMM

GMM bypasses the need to specify distributional assumptions by using instrumental variables estimation.

GMM

GMM bypasses the need to specify distributional assumptions by using instrumental variables estimation.

Choose several variables, I_t^i, known as instruments. Since the instruments are known at time t, we have

GMM

GMM bypasses the need to specify distributional assumptions by using instrumental variables estimation.

Choose several variables, I_t^i, known as instruments. Since the instruments are known at time t, we have

$$E_t(F) = 0,$$

GMM

GMM bypasses the need to specify distributional assumptions by using instrumental variables estimation.

Choose several variables, I_t^i, known as instruments. Since the instruments are known at time t, we have

$$E_t(F) = 0,$$

which implies

$$E_t(F) I_t^i = 0,$$

GMM

GMM bypasses the need to specify distributional assumptions by using instrumental variables estimation.

Choose several variables, I_t^i, known as instruments. Since the instruments are known at time t, we have

$$E_t(F) = 0,$$

which implies

$$E_t(F)I_t^i = 0,$$

so

$$0 = E(E_t(F)I_t^i) = E(FI_t^i).$$

GMM, cont'd

In GMM, parameters are chosen to minimize

$$\sum_i E(FI_t^i)^2$$

or something like it.

GMM: Pros and Cons

GMM has one big advantage:

GMM: Pros and Cons

GMM has one big advantage:

We can let the data tell us how variables are distributed.

GMM: Pros and Cons

GMM has one big advantage:

We can let the data tell us how variables are distributed. For example, we don't need to assume anything about how D_t and C_t are distributed.

GMM: Pros and Cons

GMM has one big advantage:

We can let the data tell us how variables are distributed. For example, we don't need to assume anything about how D_t and C_t are distributed.

GMM has one big disadvantage:

GMM: Pros and Cons

GMM has one big advantage:

We can let the data tell us how variables are distributed. For example, we don't need to assume anything about how D_t and C_t are distributed.

GMM has one big disadvantage:

It doesn't make any specific *predictions*. For example, we can't say what the model predicts for P_t.

GMM: Pros and Cons

GMM has one big advantage:

We can let the data tell us how variables are distributed. For example, we don't need to assume anything about how D_t and C_t are distributed.

GMM has one big disadvantage:

It doesn't make any specific *predictions*. For example, we can't say what the model predicts for P_t.

Instead, with GMM we must evaluate models indirectly. You pick *too many* instruments. The model is then *rejected* if

GMM: Pros and Cons

GMM has one big advantage:

We can let the data tell us how variables are distributed. For example, we don't need to assume anything about how D_t and C_t are distributed.

GMM has one big disadvantage:

It doesn't make any specific *predictions*. For example, we can't say what the model predicts for P_t.

Instead, with GMM we must evaluate models indirectly. You pick *too many* instruments. The model is then *rejected* if

$$E(FI_t^i) = 0,$$

A New Method

I introduce a new method, which shares the advantages of both methods:

A New Method

I introduce a new method, which shares the advantages of both methods:

- It's *non-parametric*. No distributional assumptions necessary.

A New Method

I introduce a new method, which shares the advantages of both methods:

- It's *non-parametric*. No distributional assumptions necessary.
- It makes *definite predictions* about the variables in the model.

A New Method

I introduce a new method, which shares the advantages of both methods:

- It's *non-parametric*. No distributional assumptions necessary.
- It makes *definite predictions* about the variables in the model.

The method is an empirically-based version of projection methods.

Projection Methods

In the stock example, the price is some unknown function of the other variables,

Projection Methods

In the stock example, the price is some unknown function of the other variables,

$$P_t = P(C_t, D_t).$$

Projection Methods

In the stock example, the price is some unknown function of the other variables,

$$P_t = P(C_t, D_t).$$

We can't find this function exactly, but we can approximate it,

$$P(C_t, D_t) = \sum_{i=1}^{N} a_i P^i(C_t, D_t)$$

Projection Methods, cont'd

We can't choose a_i so that the conditional moment equation holds exactly,

Projection Methods, cont'd

We can't choose a_i so that the conditional moment equation holds exactly,

$$E_t(F) = 0,$$

Projection Methods, cont'd

We can't choose a_i so that the conditional moment equation holds exactly,

$$E_t(F) = 0,$$

so instead what we do is choose N linear *projection operators*, L_i, such that

Projection Methods, cont'd

We can't choose a_i so that the conditional moment equation holds exactly,

$$E_t(F) = 0,$$

so instead what we do is choose N linear *projection operators*, L_i, such that

$$L_i E_t(F) = 0$$

holds instead.

Projection Methods, cont'd

We can't choose a_i so that the conditional moment equation holds exactly,

$$E_t(F) = 0,$$

so instead what we do is choose N linear *projection operators*, L_i, such that

$$L_i E_t(F) = 0$$

holds instead.

This gives N equations in N unknowns.

Galerkin method

In the Galerkin method, we pick some weighing function, W, and let

Galerkin method

In the Galerkin method, we pick some weighing function, W, and let

$$L_i g = \int g(x) P^i(x) W(x) dx.$$

Galerkin method

In the Galerkin method, we pick some weighing function, W, and let

$$L_i g = \int g(x) P^i(x) W(x) dx.$$

The choice of W is arbitrary.

Galerkin method

In the Galerkin method, we pick some weighing function, W, and let

$$L_i g = \int g(x) P^i(x) W(x) dx.$$

The choice of W is arbitrary. I exploit this.

Empirical Galerkin method

What if W is the pdf of the time 0 distribution of the random variables?

Empirical Galerkin method

What if W is the pdf of the time 0 distribution of the random variables?

Then

$$\begin{aligned} L_i E_t(F) &= E(E_t(F)P^i) \\ &= E(FP^i). \end{aligned}$$

Empirical Galerkin method

What if W is the pdf of the time 0 distribution of the random variables?

Then

$$\begin{aligned} L_i E_t(F) &= E(E_t(F)P^i) \\ &= E(FP^i). \end{aligned}$$

But I can compute this without knowing the distribution of x!

Empirical Galerkin method, cont'd

Just compute

$$\sum_{t=1}^{N} F(x_t, x_{t+1}) P^i(x_t)$$

from the data, for each i.

Empirical Galerkin method, cont'd

Just compute

$$\sum_{t=1}^{N} F(x_t, x_{t+1}) P^i(x_t)$$

from the data, for each i.

Then choose the a_i so that each of the above equations are zero.

What Can I Do Now?

Given parameters, you can compute the a_i.

What Can I Do Now?

Given parameters, you can compute the a_i.
Given the a_i you have an approximation,

$$P = \sum a_i P^i.$$

What Can I Do Now?

Given parameters, you can compute the a_i.
Given the a_i you have an approximation,

$$P = \sum a_i P^i.$$

You can use this to compute a time series of P predicted by the model. Then you can compute moments, test goodness-of-fit, etc.

Application to Asset Pricing

I apply this to several standard asset pricing models:

Application to Asset Pricing

I apply this to several standard asset pricing models:

- CRRA – people are impatient, and hate risk.

Application to Asset Pricing

I apply this to several standard asset pricing models:

- CRRA – people are impatient, and hate risk.
- Internal habit – people also get used to a certain level of income, and hate it when it goes down.

Application to Asset Pricing

I apply this to several standard asset pricing models:

- CRRA – people are impatient, and hate risk.
- Internal habit – people also get used to a certain level of income, and hate it when it goes down.
- External habit – people compare their level of income to the average level of income.

Application to Asset Pricing

I apply this to several standard asset pricing models:

- CRRA – people are impatient, and hate risk.
- Internal habit – people also get used to a certain level of income, and hate it when it goes down.
- External habit – people compare their level of income to the average level of income.

I take several estimates from the literature, and see how they do in replicating the time series of risk-free rates and S&P 500 price-dividend ratios.

Application to Asset Pricing, cont'd

Use polynomials in degree 3 of the following function as a basis:

Application to Asset Pricing, cont'd

Use polynomials in degree 3 of the following function as a basis:

- For CRRA: log consumption growth rate and dividend/consumption ratio.

Application to Asset Pricing, cont'd

Use polynomials in degree 3 of the following function as a basis:

- For CRRA: log consumption growth rate and dividend/consumption ratio.
- For internal habit: current and lagged log consumption growth rate, and log dividend/consumption ratio.

Application to Asset Pricing, cont'd

Use polynomials in degree 3 of the following function as a basis:

- For CRRA: log consumption growth rate and dividend/consumption ratio.
- For internal habit: current and lagged log consumption growth rate, and log dividend/consumption ratio.
- For external habit: log consumption growth rate and dividend/consumption ratio, and log S_t – the surplus consumption ratio of Campbell-Cochrane.

Application to Asset Pricing, cont'd

Use polynomials in degree 3 of the following function as a basis:

- For CRRA: log consumption growth rate and dividend/consumption ratio.
- For internal habit: current and lagged log consumption growth rate, and log dividend/consumption ratio.
- For external habit: log consumption growth rate and dividend/consumption ratio, and log S_t – the surplus consumption ratio of Campbell-Cochrane.

All data is Shiller annual dataset.

Sources of Estimates

I take model parameter estimates from the literature:

- Hansen-Singleton (1982) – CRRA
- Ferson-Constanides (1991) – internal habit
- Campbell-Cochrane (1999) – external habit (calibrated)
- Tallarini-Zhang (2005) – external habit (estimated)

Table: CRRA Estimates

	Data	Hansen and Singleton (1982)
β	NA	0.9751
γ	NA	0.9001
$E(R)$	0.0752	0.0472
$\sigma(R)$	0.1660	0.0542
$E(\log(P/D))$	3.2105	3.5572
$\sigma(\log(P/D))$	0.3573	0.1496
$E(R^f)$	0.0450	0.0439
$\sigma(R^f)$	0.0243	0.0114

Figure: CRRA Predictions: Risky Asset

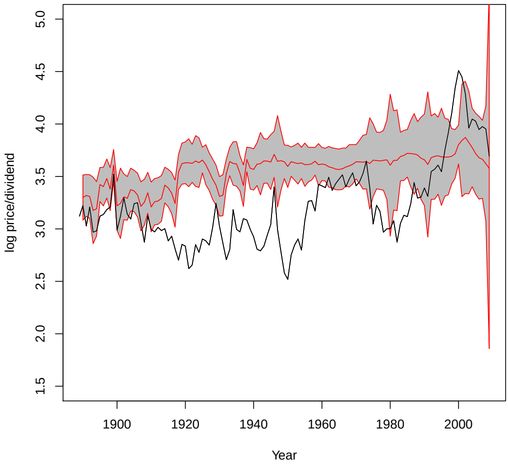


Table: One-lag Habit Estimates

	Data	Ferson-Constanides (1991)
β	NA	1.0250
γ	NA	2.5800
h	NA	0.1200
$E(R)$	0.0752	0.0308
$\sigma(R)$	0.1660	0.0821
$E(\log(P/D))$	3.2105	4.6178
$\sigma(\log(P/D))$	0.3573	0.2072
$E(R^f)$	0.0450	0.0296
$\sigma(R^f)$	0.0243	0.0515

Figure: Comparison of one-lag habit estimates

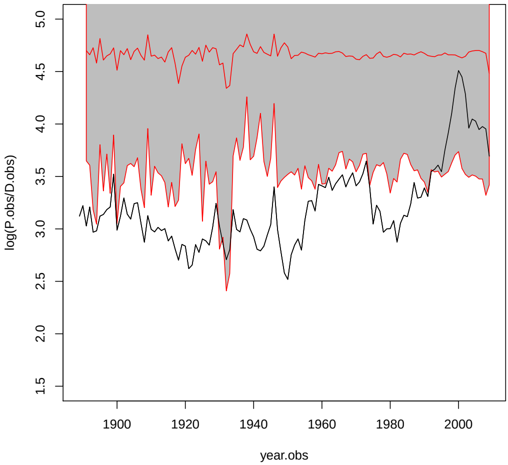
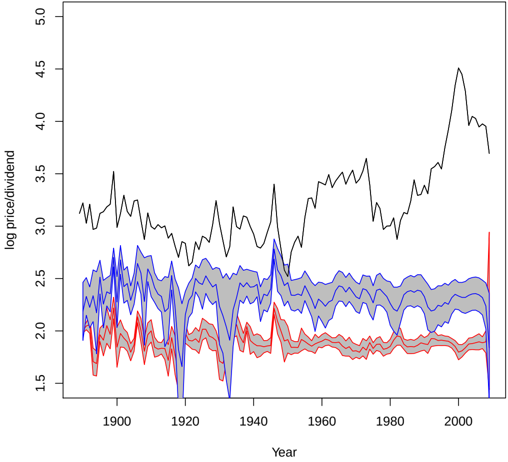


Table: External Habit Estimates

	Data	Campbell-Cochrane (1999)	Tallarini-Zhang (2005)
β	NA	0.8900	1.0100
γ	NA	2.0000	6.0000
ϕ	NA	0.8700	0.9000
$E(R)$	0.0752	0.1719	0.1223
$\sigma(R)$	0.1660	0.0994	0.1725
$E(\log(P/D))$	3.2105	1.8948	2.3441
$\sigma(\log(P/D))$	0.3573	0.0748	0.1066
$E(R^f)$	0.0450	0.1660	0.1017
$\sigma(R^f)$	0.0243	0.0403	0.0998

Figure: Time Series of External Habit Price Predictions



Conclusion

- I introduce a statistical method to solve models non-parametrically.

Conclusion

- I introduce a statistical method to solve models non-parametrically.
- It is a GMM estimator, but turned to a different purpose. Ordinary GMM takes data and gives you parameter estimates. Here we take parameter estimates, and produce predictions of the prices.

Conclusion

- I introduce a statistical method to solve models non-parametrically.
- It is a GMM estimator, but turned to a different purpose. Ordinary GMM takes data and gives you parameter estimates. Here we take parameter estimates, and produce predictions of the prices.
- This allows us to evaluate the failures of models along different dimensions, rather than simply accepting or rejecting.