

Discussion of: Dynamic Hierarchical Factor Models

Emanuel Moench Serena Ng Simon Potter

by Christopher Otrok

Summary

- A useful new entrant into the large scale factor literature
- This class of models used for many applications
 - ▶ Documenting the extent and nature of comovement in large panels (countries, states, industries)
 - ▶ Forecasting (semi-parsimonious use of lots of data)
 - ▶ Measuring the state of the economy (in the spirit of Stock and Watson (1989))
- The application of this paper is to monitor the state of the economy
 - ▶ Structure allows for some interpretation of what is driving the economy (e.g. Housing)
 - ▶ Structure allows for updating of state of economy as data are released
- A very nicely done paper with state of the art econometric modelling

Summary

- A useful new entrant into the large scale factor literature
- This class of models used for many applications
 - ▶ Documenting the extent and nature of comovement in large panels (countries, states, industries)
 - ▶ Forecasting (semi-parsimonious use of lots of data)
 - ▶ Measuring the state of the economy (in the spirit of Stock and Watson (1989))
- The application of this paper is to monitor the state of the economy
 - ▶ Structure allows for some interpretation of what is driving the economy (e.g. Housing)
 - ▶ Structure allows for updating of state of economy as data are released
- A very nicely done paper with state of the art econometric modelling

Modelling Contribution

- In principal one could estimate the following factor model

$$Y_t = B_t(L)F_t + \epsilon_t \quad (1)$$

- where Y is $N \times 1$, B a $N \times M$ time varying matrix lag polynomial, F a $M \times 1$ vector of factors
- $E(\epsilon_t \epsilon_t') = \Omega$
- For large N and T the likelihood function is hard to deal with
- The literature has worked with various restricted version of this model
 - ▶ allow for limited correlation in errors, work out asymptotics
 - ▶ allow for some parameter instability, work out asymptotics
 - ▶ often these approaches allow variables to load on all factors
- This paper uses a parametric approach to estimation
 - ▶ impose various restrictions on the parameters (zero, symmetry)
 - ▶ try to find a clever blocking strategy to break the problem into smaller feasible ones
 - ▶ A new parameter reduction is used here: impose the hierarchical structure

Modelling Contribution

- In principal one could estimate the following factor model

$$Y_t = B_t(L)F_t + \epsilon_t \quad (1)$$

- where Y is $N \times 1$, B a $N \times M$ time varying matrix lag polynomial, F a $M \times 1$ vector of factors
- $E(\epsilon_t \epsilon_t') = \Omega$
- For large N and T the likelihood function is hard to deal with
- The literature has worked with various restricted version of this model
 - ▶ allow for limited correlation in errors, work out asymptotics
 - ▶ allow for some parameter instability, work out asymptotics
 - ▶ often these approaches allow variables to load on all factors
- This paper uses a parametric approach to estimation
 - ▶ impose various restrictions on the parameters (zero, symmetry)
 - ▶ try to find a clever blocking strategy to break the problem into smaller feasible ones
 - ▶ A new parameter reduction is used here: impose the hierarchical structure

Modelling Contribution

- In principal one could estimate the following factor model

$$Y_t = B_t(L)F_t + \epsilon_t \quad (1)$$

- where Y is $N \times 1$, B a $N \times M$ time varying matrix lag polynomial, F a $M \times 1$ vector of factors
- $E(\epsilon_t \epsilon_t') = \Omega$
- For large N and T the likelihood function is hard to deal with
- The literature has worked with various restricted version of this model
 - ▶ allow for limited correlation in errors, work out asymptotics
 - ▶ allow for some parameter instability, work out asymptotics
 - ▶ often these approaches allow variables to load on all factors
- This paper uses a parameteric approach to estimation
 - ▶ impose various restrictions on the parameters (zero, symmetry)
 - ▶ try to find a clever blocking strategy to break the problem into smaller feasible ones
 - ▶ A new parameter reduction is used here: impose the hierarchical structure

This paper

- one novelty: emphasis on separate blocks
 - ▶ Similar in some ways to Kose Otrok Whiteman (2003): zero restrictions to identify regional blocks
 - ▶ let some variables load on only some factors
 - ▶ factors then interpreted as 'labor market factor' etc.
- This paper does this in a parsimonious way that leads to an efficient algorithm
- The hierarchical structure means that we start with factor for sub-blocks
 - ▶ then higher level factors are estimated off of these factors
 - ▶ we don't have to estimate a factor on a large set of data
 - ▶ asymptotics not needed
- Advantage over KOW: measure of the state of entire block (perhaps the Euro area)
- Advantage of KOW: Is there a Euro business cycle?
- Economic analysis versus economic measurement

This paper

- one novelty: emphasis on separate blocks
 - ▶ Similar in some ways to Kose Otrok Whiteman (2003): zero restrictions to identify regional blocks
 - ▶ let some variables load on only some factors
 - ▶ factors then interpreted as 'labor market factor' etc.
- This paper does this in a parsimonious way that leads to an efficient algorithm
- The hierarchical structure means that we start with factor for sub-blocks
 - ▶ then higher level factors are estimated off of these factors
 - ▶ we don't have to estimate a factor on a large set of data
 - ▶ asymptotics not needed
- Advantage over KOW: measure of the state of entire block (perhaps the Euro area)
- Advantage of KOW: Is there a Euro business cycle?
- Economic analysis versus economic measurement

This paper

- one novelty: emphasis on separate blocks
 - ▶ Similar in some ways to Kose Otrok Whiteman (2003): zero restrictions to identify regional blocks
 - ▶ let some variables load on only some factors
 - ▶ factors then interpreted as 'labor market factor' etc.
- This paper does this in a parsimonious way that leads to an efficient algorithm
- The hierarchical structure means that we start with factor for sub-blocks
 - ▶ then higher level factors are estimated off of these factors
 - ▶ we don't have to estimate a factor on a large set of data
 - ▶ asymptotics not needed
- Advantage over KOW: measure of the state of entire block (perhaps the Euro area)
- Advantage of KOW: Is there a Euro business cycle?
- Economic analysis versus economic measurement

Some comments

- In principle model has lots of flexibility
 - ▶ Lagged factors
 - ▶ multiple factors per block
 - ▶ all of this is possible in the KOW framework, but leads to a proliferation of parameters
 - ▶ in practice don't find many second factors for common factors, use few lags

Some comments

- In principle model has lots of flexibility
 - ▶ Lagged factors
 - ▶ multiple factors per block
 - ▶ all of this is possible in the KOW framework, but leads to a proliferation of parameters
 - ▶ in practice don't find many second factors for common factors, use few lags

Some comments

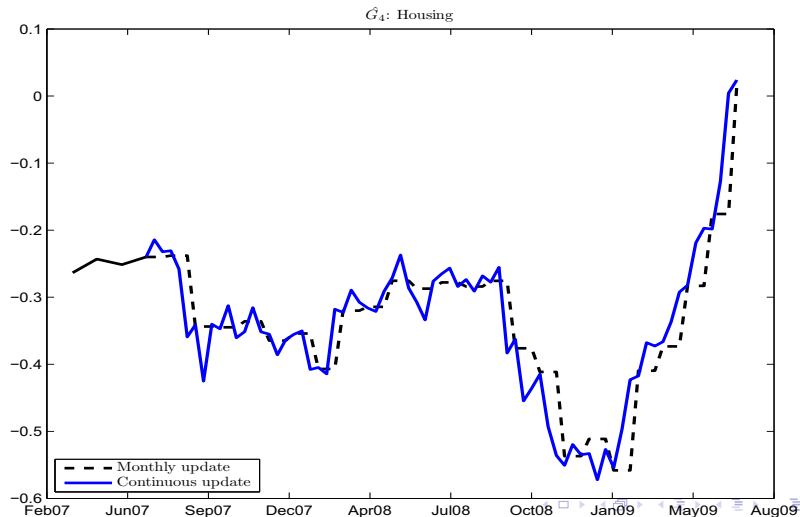
- Parsimony is achieved with restrictions on parameter values
 - ▶ Response to common factor restricted to be identical within a block
 - ▶ Why do labor market variables have the same response to the common factor?
 - ▶ The estimate factor is essentially a weighted average of the data with weights given by factor loadings
 - ▶ What if a bunch of variables should have 0 loadings?
 - ▶ Why not test this on smaller scale versions of the model?
- Some variables within a block are given a factor loading of 1
 - ▶ If two people use the same model will we get the same result?
 - ▶ Which variable is given the big weight probably matters

Some comments

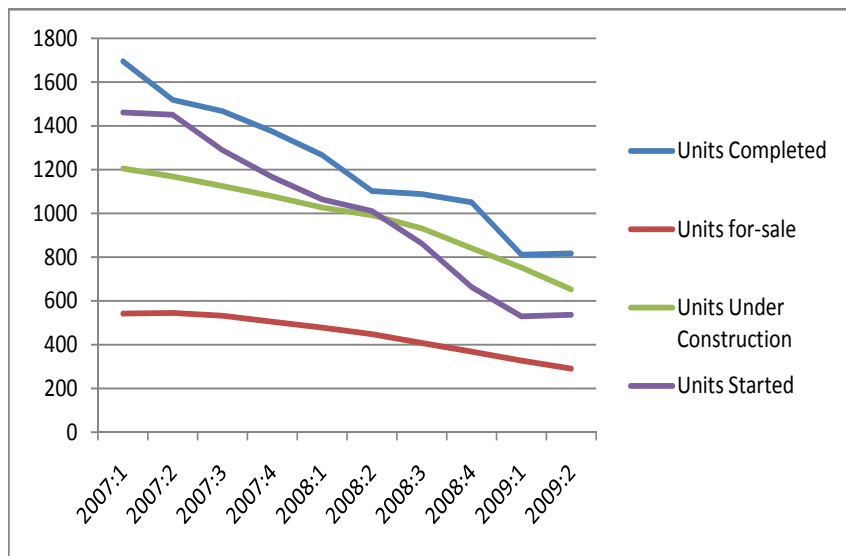
- Parsimony is achieved with restrictions on parameter values
 - ▶ Response to common factor restricted to be identical within a block
 - ▶ Why do labor market variables have the same response to the common factor?
 - ▶ The estimate factor is essentially a weighted average of the data with weights given by factor loadings
 - ▶ What if a bunch of variables should have 0 loadings?
 - ▶ Why not test this on smaller scale versions of the model?
- Some variables within a block are given a factor loading of 1
 - ▶ If two people use the same model will we get the same result?
 - ▶ Which variable is given the big weight probably matters

Housing Activity

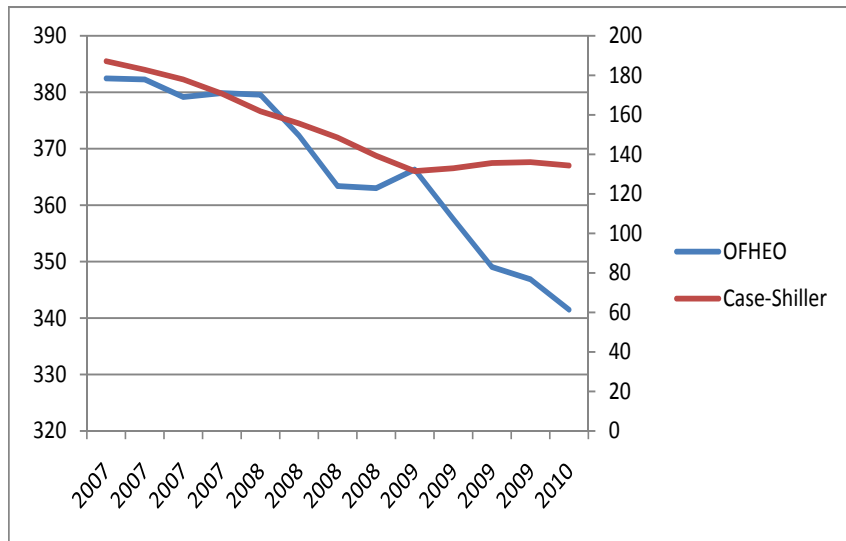
Figure 6: Monitoring Housing Activity in a Four Level Model with 5 Blocks and 14 Subblocks



Housing Data



Housing Data



Model Specification

- Specification: Assumption is block structure is correct
 - ▶ “standard model ignores block structure”
 - ▶ “Instead of imposing a possibly invalid structure ...” (on weak correlation in errors)
 - ▶ A priori no reason to think one assumption is better than the other
- Correct specification may be block with subblocks of 'demand' and 'output'
- Problem: large number of possible permutations

Model Specification

- Specification: Assumption is block structure is correct
 - ▶ “standard model ignores block structure”
 - ▶ “Instead of imposing a possibly invalid structure ...” (on weak correlation in errors)
 - ▶ A priori no reason to think one assumption is better than the other
- Correct specification may be block with subblocks of 'demand' and 'output'
- Problem: large number of possible permutations

More Comments

- Uncertainty in recession? why just point estimates?
- How is a measure of things sold a measure of demand?
- Forecasting performance (is this really better than other large N and T methods?)
- Compare real time updates with ex-post estimates of economic state
- Would permutations of the model lead to different estimates of the state of the economy?

Conclusion

- Valuable contribution to large scale factor literature
- A new approach to parameters reduction
- Valuable real time updates of state of economy
- Seems to have promise in forecasting literature