How much nominal rigidity is there in the US economy? Testing a New Keynesian DSGE Model using indirect inference

Vo Phuong Mai Le (Cardiff University) Patrick Minford (Cardiff University and CEPR) Michael Wickens (Cardiff University, University of York and CEPR)

Cardiff University

January 2010

- '**...if the conclusions (of a theory) have been falsified, then their falsification also falsifies the theory from which they were logically deduced.** It should be noticed that a positive decision can only temporarily support the theory, for subsequent negative decisions may always overthrow it.' Popper, The Logic of Scientific Discovery (p.10).

- ‘**...if the conclusions (of a theory) have been falsified, then their falsification also falsifies the theory from which they were logically deduced.** It should be noticed that a positive decision can only temporarily support the theory, for subsequent negative decisions may always overthrow it.’ Popper, The Logic of Scientific Discovery (p.10).
- ‘...my recollection is that Bob Lucas and Ed Prescott were initially very enthusiastic about rational expectations econometrics. After all, it simply involved imposing on ourselves the same high standards we had criticized the Keynesians for failing to live up to. **But after about five years of doing likelihood ratio tests on rational expectations models, I recall Bob Lucas and Ed Prescott both telling me that those tests were rejecting too many good models.**’ Tom Sargent, interviewed by Evans and Honkapohja (p.6)

Introduction

- Quotations reveal a) economists' Popperian desire to test by rejection b) problem faced by DSGE models with tests based on likelihood from direct inference.

Introduction

- Quotations reveal a) economists' Popperian desire to test by rejection b) problem faced by DSGE models with tests based on likelihood from direct inference.
- RBC/DSGE models fitted data poorly; so new testing procedure designed based on comparison of 'dynamic facts' with model's simulation of these. This is essentially the procedure used in Indirect Inference, though under Ind. Inf. the procedure is based on a formal test statistic with a distribution derived from the model and the data.

Introduction

- Quotations reveal a) economists' Popperian desire to test by rejection b) problem faced by DSGE models with tests based on likelihood from direct inference.
- RBC/DSGE models fitted data poorly; so new testing procedure designed based on comparison of 'dynamic facts' with model's simulation of these. This is essentially the procedure used in Indirect Inference, though under Ind. Inf. the procedure is based on a formal test statistic with a distribution derived from the model and the data.
- Here we test Smets and Wouters (SW, 2007) DSGE model of the US using Ind. Inf.. The model includes: habit persistence, variable capacity utilisation, q-investment theory, and Calvo wages/prices with lagged indexation. Close to Christiano, Eichenbaum and Evans (JPE, 2005) for US. New Neo- Keynesian Synthesis state of the art; estimated by Bayesian methods. Similarities in method and model to SW (2003) for EU.

Introduction

- Quotations reveal a) economists' Popperian desire to test by rejection b) problem faced by DSGE models with tests based on likelihood from direct inference.
- RBC/DSGE models fitted data poorly; so new testing procedure designed based on comparison of 'dynamic facts' with model's simulation of these. This is essentially the procedure used in Indirect Inference, though under Ind. Inf. the procedure is based on a formal test statistic with a distribution derived from the model and the data.
- Here we test Smets and Wouters (SW, 2007) DSGE model of the US using Ind. Inf.. The model includes: habit persistence, variable capacity utilisation, q-investment theory, and Calvo wages/prices with lagged indexation. Close to Christiano, Eichenbaum and Evans (JPE, 2005) for US. New Neo- Keynesian Synthesis state of the art; estimated by Bayesian methods. Similarities in method and model to SW (2003) for EU.
- We evaluate a version of this US model (using mean SW posterior coefficients) and also variants of it, using indirect inference.

Indirect v. direct inference

- estimating structural model directly (e.g. by FIML or Bayesian ML, 'direct inference') maximises fit of specified structure to current data, given lagged data and current exogenous variables — basically minimises sum of squared errors of current forecast. Likelihood tests essentially check model's forecasting ability.

Indirect v. direct inference

- estimating structural model directly (e.g. by FIML or Bayesian ML, 'direct inference') maximises fit of specified structure to current data, given lagged data and current exogenous variables — basically minimises sum of squared errors of current forecast. Likelihood tests essentially check model's forecasting ability.
- but we are interested in how well model replicates 'dynamic facts' Model may have small errors but poor dynamic behaviour. Also DSGE models typically (because highly restricted) have large errors but arguably represent dynamics well.

Indirect v. direct inference cont.

- how represent dynamic facts parsimoniously? A VAR is often used. This is 'auxiliary model' of indirect inference: it describes facts neutrally between competing structural models (the nulls), subject to not contradicting these nulls in specification. Can use VAR coefficients (as here) to summarise dynamics in the data. Or can use VAR to produce Impulse Response Functions (IRFs) of key variables to key shocks.

Indirect v. direct inference cont.

- how represent dynamic facts parsimoniously? A VAR is often used. This is 'auxiliary model' of indirect inference: it describes facts neutrally between competing structural models (the nulls), subject to not contradicting these nulls in specification. Can use VAR coefficients (as here) to summarise dynamics in the data. Or can use VAR to produce Impulse Response Functions (IRFs) of key variables to key shocks.
- then Indirect Inference asks: is this DSGE model the mechanism that generated the data, as described by this VAR/these IRFs?

Model estimation by indirect inference

- The idea is to move the parameters of the structural DSGE model around until the model implies a VAR representation as close as possible to the VAR (auxiliary model) estimated on the actual data. IE the DSGE model parameters are chosen not 'directly' to mimimise the model's forecast errors ; but 'indirectly' on the basis of the model's implied closeness to the VAR fitted to the data.

Model estimation by indirect inference

- The idea is to move the parameters of the structural DSGE model around until the model implies a VAR representation as close as possible to the VAR (auxiliary model) estimated on the actual data. IE the DSGE model parameters are chosen not 'directly' to mimimise the model's forecast errors ; but 'indirectly' on the basis of the model's implied closeness to the VAR fitted to the data.
- Closeness can be assessed on basis of various measures — e.g. the impulse response functions or the VAR coefficients themselves (we use these).

Model estimation by indirect inference

- The idea is to move the parameters of the structural DSGE model around until the model implies a VAR representation as close as possible to the VAR (auxiliary model) estimated on the actual data. IE the DSGE model parameters are chosen not 'directly' to mimimise the model's forecast errors ; but 'indirectly' on the basis of the model's implied closeness to the VAR fitted to the data.
- Closeness can be assessed on basis of various measures — e.g. the impulse response functions or the VAR coefficients themselves (we use these).
- References: Smith (1993), Gregory and Smith (1991, 1993), Gourieroux et al. (1993), Gourieroux and Montfort (1995) and Canova (2005)

Model evaluation by indirect inference

- Here the parameters of the macroeconomic model are taken as given; the model is treated as the null hypothesis. It is assumed either that it has already been estimated, or calibrated, as well as possible.

Model evaluation by indirect inference

- Here the parameters of the macroeconomic model are taken as given; the model is treated as the null hypothesis. It is assumed either that it has already been estimated, or calibrated, as well as possible.
- The aim is to compare the VAR estimated from the data with the VAR estimated from simulations of the macroeconomic model.

Model evaluation by indirect inference

- Here the parameters of the macroeconomic model are taken as given; the model is treated as the null hypothesis. It is assumed either that it has already been estimated, or calibrated, as well as possible.
- The aim is to compare the VAR estimated from the data with the VAR estimated from simulations of the macroeconomic model.
- So in effect the measure of fit of the DSGE model to the VAR is used now as a test of the model.

Model evaluation by indirect inference cont.

- The test is based on the statistical distribution of the VAR parameters (also the data variances) implied by the DSGE structural model, denoted by the vector a of criterion measures. This is found by bootstrapping the DSGE model (ie simulating with model's own implied shocks, solving by DYNARE, 1000 times) and re-estimating the VAR on the 1000 bootstrap pseudo-samples.

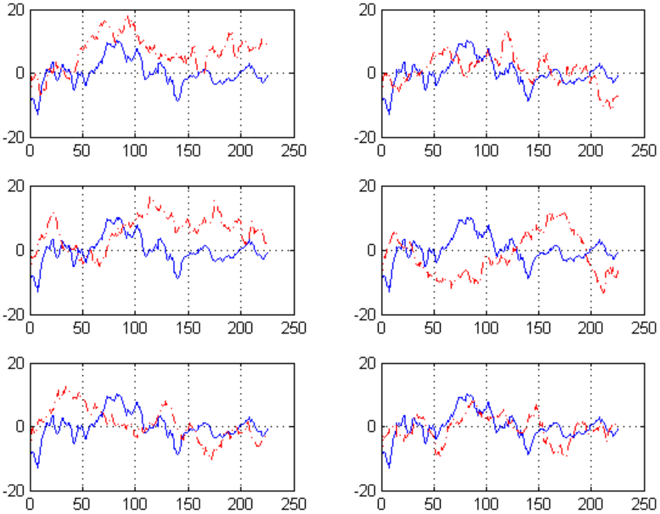
Model evaluation by indirect inference cont.

- The test is based on the statistical distribution of the VAR parameters (also the data variances) implied by the DSGE structural model, denoted by the vector *a* of criterion measures. This is found by bootstrapping the DSGE model (ie simulating with model's own implied shocks, solving by DYNARE, 1000 times) and re-estimating the VAR on the 1000 bootstrap pseudo-samples.
- Bootstrapping is used to find small sample distributions implied by the actual errors implied by the model and data— as asymptotic small sample distributions based on an assumed (eg normal) distribution of the errors may be misleading (eg if error distributions are non-normal).

Model evaluation by indirect inference cont.

- The test is based on the statistical distribution of the VAR parameters (also the data variances) implied by the DSGE structural model, denoted by the vector a of criterion measures. This is found by bootstrapping the DSGE model (ie simulating with model's own implied shocks, solving by DYNARE, 1000 times) and re-estimating the VAR on the 1000 bootstrap pseudo-samples.
- Bootstrapping is used to find small sample distributions implied by the actual errors implied by the model and data— as asymptotic small sample distributions based on an assumed (eg normal) distribution of the errors may be misleading (eg if error distributions are non-normal).
- A Wald statistic is calculated based on the bootstrap distribution (implied by the DSGE model coefficients $\widehat{\theta}$) of a around their bootstrap means, as $[a_T - \alpha_S(\widehat{\theta})]' W(\widehat{\theta}) [a_T - \alpha_S(\widehat{\theta})]$,where $W(\widehat{\theta})$ is the inverse of the var-covar matrix of the a. a_T are the values from the data.

Method illustrated: 1) random bootstrap samples from DSGE model (output- weighted model-blue is actual):



2) Joint distribution implied by DSGE model illustrated for 2 coefficients only:

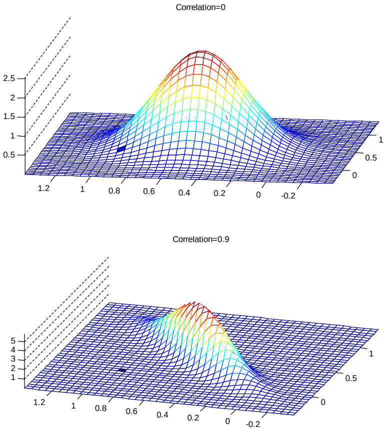


Figure: Bivariate Normal Distributions (0.1, 0.9 shaded), corr. of 0 & 0.9.

3) The Wald statistic implied by joint distribution for full 30 coefficients

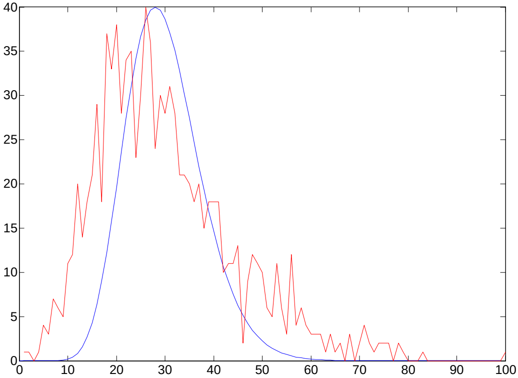


Figure: Histogram of Wald statistic for SW model of final section, with Chi-squared distribution (30 degrees of freedom)

4) The Wald statistic transformed to a normalised t-value (Mahalanobis Distance)

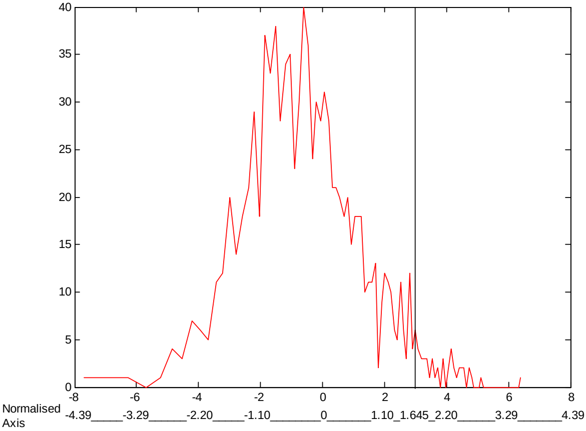


Figure: Normalised Mahalanobis Distance

Wald statistic discussed:

- with 2 coefficients Wald can be written as $\frac{t_1^2+t_2^2-2t_1t_2\rho_{12}}{1-\rho_{12}^2}$;so in Fig1a $\rho_{12}=0.9$, $t_1=-1.6$, $t_2=+1.6$; Wald=51.2 (if $\rho_{12}=0$, Wald=5.12) against 95% value for $\chi^2(2)$ of 5.99. Note how high correlation affects Wald and joint significance.

Wald statistic discussed:

- with 2 coefficients Wald can be written as $\frac{t_1^2+t_2^2-2t_1t_2\rho_{12}}{1-\rho_{12}^2}$;so in Fig1a $\rho_{12}=0.9$, $t_1=-1.6$, $t_2=+1.6$; Wald=51.2 (if $\rho_{12}=0$, Wald=5.12) against 95% value for $\chi^2(2)$ of 5.99. Note how high correlation affects Wald and joint significance.
- Wald statistic distribution not an exact χ^2 because bootstrap distributions of a are not multivariate normal; skewed and with fat tails.

Wald statistic discussed:

- with 2 coefficients Wald can be written as $\frac{t_1^2+t_2^2-2t_1t_2\rho_{12}}{1-\rho_{12}^2}$;so in Fig1a $\rho_{12}=0.9$, $t_1=-1.6$, $t_2=+1.6$; Wald=51.2 (if $\rho_{12}=0$, Wald=5.12) against 95% value for $\chi^2(2)$ of 5.99. Note how high correlation affects Wald and joint significance.
- Wald statistic distribution not an exact χ^2 because bootstrap distributions of a are not multivariate normal; skewed and with fat tails.
- $\sqrt{\chi^2}$ is normal variable; $\sqrt{Wald}$ (Mahalanobis Distance, MD) here is skewed and fat-tailed. From it we can compute the distance from the 95% point; we express this a normalised t-value (where the 95% point is 1.645).

Wald statistic discussed:

- with 2 coefficients Wald can be written as $\frac{t_1^2+t_2^2-2t_1t_2\rho_{12}}{1-\rho_{12}^2}$;so in Fig1a $\rho_{12}=0.9$, $t_1=-1.6$, $t_2=+1.6$; Wald=51.2 (if $\rho_{12}=0$, Wald=5.12) against 95% value for $\chi^2(2)$ of 5.99. Note how high correlation affects Wald and joint significance.
- Wald statistic distribution not an exact χ^2 because bootstrap distributions of *a* are not multivariate normal; skewed and with fat tails.
- $\sqrt{\chi^2}$ is normal variable; $\sqrt{Wald}$ (Mahalanobis Distance, MD) here is skewed and fat-tailed. From it we can compute the distance from the 95% point; we express this a normalised t-value (where the 95% point is 1.645).
- we use Wald statistic percentile; when this=100 we also use normalised MD to show Distance from 95% point.

Practical issues

- we identify two types of Wald statistic:

Practical issues

- we identify two types of Wald statistic:
- a) Full Wald based on full joint distribution of VAR coefficients (with full covariance matrix as in Figure 1b above).

Practical issues

- we identify two types of Wald statistic:
- a) Full Wald based on full joint distribution of VAR coefficients (with full covariance matrix as in Figure 1b above).
- b) 'Directed Wald' statistics derived from one aspect alone of the model's performance — e.g. for groups of variables, we redo the VAR for these alone and compute Wald; or for a shock we compute joint distribution of its (average) IRFs.

Practical issues

- we identify two types of Wald statistic:
- a) Full Wald based on full joint distribution of VAR coefficients (with full covariance matrix as in Figure 1b above).
- b) 'Directed Wald' statistics derived from one aspect alone of the model's performance — e.g. for groups of variables, we redo the VAR for these alone and compute Wald; or for a shock we compute joint distribution of its (average) IRFs.
- relation to usual DSGE tests: eg moments/cross-moments, IRFs of DSGE and data? We find results are similar provided these comparisons are done on **joint** distribution of these measures, as in second illustration above of bivariate case; similar because same bootstrap samples produce distributions of these as of *a*.

A testing hierarchy — Wald tests increase in power the lower down the slide you go:

- begin with full universe of facts: **select key facts/variables and mode of summary description** — here 5 variables (output, inflation, interest rates, consumption, investment) :a consists of 25 VAR(1) coefficients and 5 data variances. NB the larger the selection, the higher the odds of rejection; models are not intended to explain everything!

A testing hierarchy — Wald tests increase in power the lower down the slide you go:

- begin with full universe of facts: **select key facts/variables and mode of summary description** — here 5 variables (output, inflation, interest rates, consumption, investment) :a consists of 25 VAR(1) coefficients and 5 data variances. NB the larger the selection, the higher the odds of rejection; models are not intended to explain everything!
- evaluation of individual features on their own (cf in figure above, a VAR coeff compared with its own distribution, not as in bottom panel jointly with other VAR coeffs.) Gives individual significance (**'t-values'**)

A testing hierarchy — Wald tests increase in power the lower down the slide you go:

- begin with full universe of facts: **select key facts/variables and mode of summary description** — here 5 variables (output, inflation, interest rates, consumption, investment) :*a* consists of 25 VAR(1) coefficients and 5 data variances. NB the larger the selection, the higher the odds of rejection; models are not intended to explain everything!
- evaluation of individual features on their own (cf in figure above, a VAR coeff compared with its own distribution, not as in bottom panel jointly with other VAR coeffs.) Gives individual significance **('t-values')**
- evaluate aspects of the model — i.e. groups of features evaluated jointly — **Directed Walds**

A testing hierarchy — Wald tests increase in power the lower down the slide you go:

- begin with full universe of facts: **select key facts/variables and mode of summary description** — here 5 variables (output, inflation, interest rates, consumption, investment) :a consists of 25 VAR(1) coefficients and 5 data variances. NB the larger the selection, the higher the odds of rejection; models are not intended to explain everything!
- evaluation of individual features on their own (cf in figure above, a VAR coeff compared with its own distribution, not as in bottom panel jointly with other VAR coeffs.) Gives individual significance (**'t-values'**)
- evaluate aspects of the model — i.e. groups of features evaluated jointly — **Directed Walds**
- evaluate all aspects of the model against a_T — **Full Wald**

Application of the method to US post-war data

- Quarterly post-war data (1947Q1–2004Q2); all detrended by simple constant + linear time trend, as in SW EU model; achieves stationarity on usual tests. Similar results were obtained with other filters- linear detrending chosen as removes least information from the data.

Application of the method to US post-war data

- Quarterly post-war data (1947Q1–2004Q2); all detrended by simple constant + linear time trend, as in SW EU model; achieves stationarity on usual tests. Similar results were obtained with other filters- linear detrending chosen as removes least information from the data.
- As a benchmark for the New Keynesian model (NK) we create a 'New Classical' version (NC), identical in all respects except that full price/wage flexibility, one-quarter information delay of households in labour supply, simpler Taylor Rule.

Application of the method to US post-war data cont.

- Also look at weighted combination of NK and NC, where firms/workers face both perfectly and imperfectly competitive product/labour markets in fixed proportions.

Application of the method to US post-war data cont.

- Also look at weighted combination of NK and NC, where firms/workers face both perfectly and imperfectly competitive product/labour markets in fixed proportions.
- Object is a) to allow relative testing of DSGE models; b) to revisit an old issue, extent of nominal rigidity, relating this to the 'weights' on imperfect competition. We interpret 'extent of nom. rig.' as reflecting relative degree of time dependence v state dependence of nominal contracts- evidence of micro studies for former (eg Bils and Klenow, 2004), for latter see Gertler and Leahy (2008).

Workings of the NK and NC model types

- NK workings: because capacity utilisation is flexible, demand shocks (consumption/investment/money) directly impact on output and (via Phillips Curve) inflation, then (via Taylor Rule) interest rates. Supply shocks (productivity, labour supply, wages/inflation mark-ups) play role as 'cost-push' inflation shocks; then (via Taylor Rule) affect interest rates and so output. Persistent demand shocks raise 'Q' persistently and produce an 'investment boom' which via demand effects reinforces itself. Thus the model acts as a 'multiplier/accelerator' of demand shocks.

Workings of the NK and NC model types

- NK workings: because capacity utilisation is flexible, demand shocks (consumption/investment/money) directly impact on output and (via Phillips Curve) inflation, then (via Taylor Rule) interest rates. Supply shocks (productivity, labour supply, wages/inflation mark-ups) play role as 'cost-push' inflation shocks; then (via Taylor Rule) affect interest rates and so output. Persistent demand shocks raise 'Q' persistently and produce an 'investment boom' which via demand effects reinforces itself. Thus the model acts as a 'multiplier/accelerator' of demand shocks.
- NC workings: inelastic labour supply means output variation dominated by supply shocks (productivity and labour supply) with investment/consumption reactions in 'pure' RBC manner. These reactions plus demand shocks create market-clearing movements in real interest rates and via the Taylor rule inflation. Thus supply shocks are prime movers of all variables, but demand shocks add to the variability of nominal ones.

Testing the SW NK model

- SW's error properties estimated from Bayesian posteriors somewhat, but not massively, different from those implied by the DSGE models and the data (the 'actual' errors under the null of the DSGE model):

	Govt Spend	Pref	Inv	Mon	Prod	Price Mark -up	Wage Mark -up
SW stdev	0.53	0.23	0.45	0.24	0.45	0.14	0.24
Data stdev	0.67	0.37	0.70	0.34	0.55	0.24	0.31
SW $AR(1)$	0.97	0.22	0.71	0.15	0.95	0.89	0.96
SW $MA(1)$	0.52					-0.69	-0.84
Est. $AR(1)$	0.94	-0.06	0.53	-0.06	0.97	0.93	0.92
Est. $MA(1)$	0.55					-0.71	-0.85

Table: Standard deviations of innovations and coefficients of shocks (actual vs. assumed)

- SW NK model based on ‘own errors’ is strongly rejected; normalised t-value of Mahalanobis Distance is 3.4. Nominal variable variances far too small relative to data; eg interest rate variance in data 0.65 (% per quarter) v. model 95% bounds 0.19-0.57.

- SW NK model based on ‘own errors’ is strongly rejected; normalised t-value of Mahalanobis Distance is 3.4. Nominal variable variances far too small relative to data; eg interest rate variance in data 0.65 (% per quarter) v. model 95% bounds 0.19-0.57.
- Using errors derived from actual data and SW model, MD t-value is similar at 3.6. Nominal variances even smaller relative to data.

- SW NK model based on 'own errors' is strongly rejected; normalised t-value of Mahalanobis Distance is 3.4. Nominal variable variances far too small relative to data; eg interest rate variance in data 0.65 (% per quarter) v. model 95% bounds 0.19-0.57.
- Using errors derived from actual data and SW model, MD t-value is similar at 3.6. Nominal variances even smaller relative to data.
- NC model is worse: MD t-value worsens to 4.7. Also variance of nominal variables now far too large; eg inflation variance in data 0.44 (% per quarter) v model 95% bounds 2.34-3.61.

- SW NK model based on 'own errors' is strongly rejected; normalised t-value of Mahalanobis Distance is 3.4. Nominal variable variances far too small relative to data; eg interest rate variance in data 0.65 (% per quarter) v. model 95% bounds 0.19-0.57.
- Using errors derived from actual data and SW model, MD t-value is similar at 3.6. Nominal variances even smaller relative to data.
- NC model is worse: MD t-value worsens to 4.7. Also variance of nominal variables now far too large; eg inflation variance in data 0.44 (% per quarter) v model 95% bounds 2.34-3.61.
- Weighted model is best: MD t-value=2.8; all data variances lie within model bounds. Assumes best-fitting weights for imperfectly competitive share: 0.2 (product market), 0.1 (labour market); best fit found by Indirect Inference. This model behaves like NC- so supply shocks dominant-, except that nominal variables are heavily dampened by NK weights (disproportionate effect because dampen expected inflation as well as inflation directly; hence 'dampening multiplier' effect).

- SW NK model based on 'own errors' is strongly rejected; normalised t-value of Mahalanobis Distance is 3.4. Nominal variable variances far too small relative to data; eg interest rate variance in data 0.65 (% per quarter) v. model 95% bounds 0.19-0.57.
- Using errors derived from actual data and SW model, MD t-value is similar at 3.6. Nominal variances even smaller relative to data.
- NC model is worse: MD t-value worsens to 4.7. Also variance of nominal variables now far too large; eg inflation variance in data 0.44 (% per quarter) v model 95% bounds 2.34-3.61.
- Weighted model is best: MD t-value=2.8; all data variances lie within model bounds. Assumes best-fitting weights for imperfectly competitive share: 0.2 (product market), 0.1 (labour market); best fit found by Indirect Inference. This model behaves like NC- so supply shocks dominant-, except that nominal variables are heavily dampened by NK weights (disproportionate effect because dampen expected inflation as well as inflation directly; hence 'dampening multiplier' effect).
- Note however all models rejected overall.

Variable combinations	Direct Wald
Y, C, INV	98.3
Y, C, INV, R	99.0
Y, C, INV, π	100
Y, R, π	99.4
Y, π	97.6
R, π	96.2
$\sigma_Y^2, \sigma_R^2, \sigma_\pi^2, \sigma_C^2, \sigma_I^2$	97.0

Table: Directed Wald statistics BY VARIABLE COMBINATIONS

Single Variables	Direct Wald
$Y\,(AR\,(3))$	96.2
$R\,(ARMA\,(1,1))$	98.4
$\pi\,(AR\,(3))$	90.3
$C\,(AR\,(3))$	98.8
$INV\,(AR\,(2))$	95.2

Table: Directed Wald statistics BY VARIABLE COMBINATIONS

- Taking the best model we now examine why it fails and how it succeeds, using Directed Walds.

- Taking the best model we now examine why it fails and how it succeeds, using Directed Walds.
- First, ignoring variances, only one variable singly is accepted at 95% : inflation, π (90.3); investment (INV) and output (Y) nearly (95.2, 96.2); interest rates (R, 98.4) and consumption (C, 98.8) rejected clearly.

- Taking the best model we now examine why it fails and how it succeeds, using Directed Walds.
- First, ignoring variances, only one variable singly is accepted at 95% : inflation, π (90.3); investment (INV) and output (Y) nearly (95.2, 96.2); interest rates (R, 98.4) and consumption (C, 98.8) rejected clearly.
- Second, combinations of real variables (Y, C, INV) are rejected (98.3) while combinations of nominal (π, R) nearly accepted (96.2).

- Taking the best model we now examine why it fails and how it succeeds, using Directed Walds.
- First, ignoring variances, only one variable singly is accepted at 95% : inflation, π (90.3); investment (INV) and output (Y) nearly (95.2, 96.2); interest rates (R, 98.4) and consumption (C, 98.8) rejected clearly.
- Second, combinations of real variables (Y, C, INV) are rejected (98.3) while combinations of nominal (π, R) nearly accepted (96.2).
- Third, the most limited combinations of real and nominal variables (Y, π and Y, π, R) are rejected plainly (97.6, 99.4).

- Taking the best model we now examine why it fails and how it succeeds, using Directed Walds.
- First, ignoring variances, only one variable singly is accepted at 95% : inflation, π (90.3); investment (INV) and output (Y) nearly (95.2, 96.2); interest rates (R, 98.4) and consumption (C, 98.8) rejected clearly.
- Second, combinations of real variables (Y, C, INV) are rejected (98.3) while combinations of nominal (π, R) nearly accepted (96.2).
- Third, the most limited combinations of real and nominal variables (Y, π and Y, π, R) are rejected plainly (97.6, 99.4).
- Fourth, variances of real and nominal variables are captured individually but as a group rejected jointly (97.0).

- Taking the best model we now examine why it fails and how it succeeds, using Directed Walds.
- First, ignoring variances, only one variable singly is accepted at 95% : inflation, π (90.3); investment (INV) and output (Y) nearly (95.2, 96.2); interest rates (R, 98.4) and consumption (C, 98.8) rejected clearly.
- Second, combinations of real variables (Y, C, INV) are rejected (98.3) while combinations of nominal (π, R) nearly accepted (96.2).
- Third, the most limited combinations of real and nominal variables (Y, π and Y, π, R) are rejected plainly (97.6, 99.4).
- Fourth, variances of real and nominal variables are captured individually but as a group rejected jointly (97.0).
- Conclusion: the model can capture real and nominal relationships separately, but fails to capture real-nominal relationships and scale of joint movement.

Interpreting the failure as regime change

- The failure identified could lie anywhere in the model. But one plausible avenue lies in monetary regime change (eg from fixed exchange rates to money supply rules, to Taylor Rules), causing change in price/wage rigidity as price/wage variability changed.

Interpreting the failure as regime change

- The failure identified could lie anywhere in the model. But one plausible avenue lies in monetary regime change (eg from fixed exchange rates to money supply rules, to Taylor Rules), causing change in price/wage rigidity as price/wage variability changed.
- We find evidence of regime change in the sample using the Perron-Wu test on our VAR:

The estimated breaks are:	1965.02	1984.02
The 95% C.I. for the 1st break is	(1964.04;1965.04)	
The 95% C.I. for the 2nd break is	(1983.02-1985.02)	

Table: Perron-Qu Multivariate Structural Break Test

Interpreting the failure as regime change

- The failure identified could lie anywhere in the model. But one plausible avenue lies in monetary regime change (eg from fixed exchange rates to money supply rules, to Taylor Rules), causing change in price/wage rigidity as price/wage variability changed.
- We find evidence of regime change in the sample using the Perron-Wu test on our VAR:

The estimated breaks are:	1965.02	1984.02
The 95% C.I. for the 1st break is	(1964.04;1965.04)	
The 95% C.I. for the 2nd break is	(1983.02-1985.02)	

Table: Perron-Qu Multivariate Structural Break Test

- Break in 1965 could be related to rising inflation and early monetary conflict with Germany in Bretton Woods; break in 1985 could be related to shift towards interest-rate setting with implicit inflation targeting.

Finding a data-acceptable model for the Great Moderation period

	Direct Walds
$Y\,(AR\,(3))$	54.7
$R(AR(1))$	97.2
$R\,(AR\,(2))$	100
$\pi\,(AR\,(1))$	69.6
Y, π	88.5
Y, R, π	96.6
$var(Y), var(R), var(\pi)$	60.5

Table: Direct Walds for different combinations of output, inflation and interest rate

- Using data from 1984Q3-2004Q2, H-P filtered (as linear detrending was no longer sufficient for stationarity), best model had weights on imperfect competition of 0.8 for both product and labour markets- ie far more nominal rigidity than sample overall. Overall model is still rejected strongly (Wald=100); MD=4.2 (about the same as weighted model full sample on HP Data, 3.9).

- Using data from 1984Q3-2004Q2, H-P filtered (as linear detrending was no longer sufficient for stationarity), best model had weights on imperfect competition of 0.8 for both product and labour markets- ie far more nominal rigidity than sample overall. Overall model is still rejected strongly (Wald=100); MD=4.2 (about the same as weighted model full sample on HP Data, 3.9).
- But model can now fit combination of real and nominal variables (Y, π, R; 96.6; Y, π, 88.5). It fits the joint real and nominal variances easily (60.5); adding variances to Walds leaves them virtually unchanged.

- Using data from 1984Q3-2004Q2, H-P filtered (as linear detrending was no longer sufficient for stationarity), best model had weights on imperfect competition of 0.8 for both product and labour markets- ie far more nominal rigidity than sample overall. Overall model is still rejected strongly (Wald=100); MD=4.2 (about the same as weighted model full sample on HP Data, 3.9).
- But model can now fit combination of real and nominal variables (Y, π, R; 96.6; Y, π, 88.5). It fits the joint real and nominal variances easily (60.5); adding variances to Walds leaves them virtually unchanged.
- Conclusion: a model with a Taylor Rule and high nominal rigidity fits the Great Moderation in respect of main real/nominal variables (but not, as usual with such DSGE models, overall with consumption and investment included).

Conclusions

- In this paper we have used indirect inference to test a DSGE model the US estimated by Smets and Wouters (2007) following Christiano et al. (2005) The original New Keynesian (NK) version has been compared with a New Classical (NC) version and with weighted versions, where the weights are the assumed shares of imperfect competition in product and labour markets.

Conclusions

- In this paper we have used indirect inference to test a DSGE model the US estimated by Smets and Wouters (2007) following Christiano et al. (2005) The original New Keynesian (NK) version has been compared with a New Classical (NC) version and with weighted versions, where the weights are the assumed shares of imperfect competition in product and labour markets.
- The most successful model for the whole post-war period is a weighted model with small weights on imperfect competition in both markets, indicating low nominal rigidity. But even this is rejected by the data overall and is incapable of capturing joint behaviour of nominal and real variables (even just output and inflation together).

Conclusions

- In this paper we have used indirect inference to test a DSGE model the US estimated by Smets and Wouters (2007) following Christiano et al. (2005) The original New Keynesian (NK) version has been compared with a New Classical (NC) version and with weighted versions, where the weights are the assumed shares of imperfect competition in product and labour markets.
- The most successful model for the whole post-war period is a weighted model with small weights on imperfect competition in both markets, indicating low nominal rigidity. But even this is rejected by the data overall and is incapable of capturing joint behaviour of nominal and real variables (even just output and inflation together).
- A plausible reason for this failure is monetary regime change and the data suggest two structural breaks, in mid-60s and mid-80s.

Conclusions cont.

- For period since 1984 best version has high weights (80%) on imperfect competition; this model (uniquely among those examined) accepted in respect of the joint behaviour of output, inflation and interest rates.

Conclusions cont.

- For period since 1984 best version has high weights (80%) on imperfect competition; this model (uniquely among those examined) accepted in respect of the joint behaviour of output, inflation and interest rates.
- Suggests that whereas for most of post-war period state dependence may have been dominant force in pricing, implying little rigidity, with the Great Moderation it was dominated by time dependence and nominal rigidity increased sharply.