A new approach for policymakers to testing models and policies

Money Macro and Finance Conference, Cardiff

Patrick Minford

Cardiff University & CEPR

September 2015

Introduction

- Policymakers about to introduce monetary policy reforms after crisis....

Introduction

- Policymakers about to introduce monetary policy reforms after crisis....
- How to design them?

Introduction

- Policymakers about to introduce monetary policy reforms after crisis....
- How to design them?
- Will they fail or succeed?

Introduction

- Policymakers about to introduce monetary policy reforms after crisis....
- How to design them?
- Will they fail or succeed?
- How to make sure they succeed to some extent at least?

Introduction

- Rival suggestions today

- Rival suggestions today
- a) Back to Economic History and build ad hoc models 'reflecting elasticities' (e.g.SVARs)? **Violates** Lucas' Critique: useless policy evaluation.

Introduction

- Rival suggestions today
- a) Back to Economic History and build ad hoc models 'reflecting elasticities' (e.g.SVARs)? **Violates** Lucas' Critique: useless policy evaluation.
- b) Build micro-founded models, calibrate and compare with stylised facts? **No** measure of statistical accuracy.

Introduction

- Rival suggestions today
- a) Back to Economic History and build ad hoc models 'reflecting elasticities' (e.g.SVARs)? **Violates** Lucas' Critique: useless policy evaluation.
- b) Build micro-founded models, calibrate and compare with stylised facts? **No** measure of statistical accuracy.
- c) Improve on above by using Bayesian methods, treating calibrated parameters as priors? **If priors wrong**, bias in probabilities. Also many possible models, with contradictory implications for policy!

Introduction

- Rival suggestions today
- a) Back to Economic History and build ad hoc models 'reflecting elasticities' (e.g.SVARs)? **Violates** Lucas' Critique: useless policy evaluation.
- b) Build micro-founded models, calibrate and compare with stylised facts? **No** measure of statistical accuracy.
- c) Improve on above by using Bayesian methods, treating calibrated parameters as priors? **If priors wrong**, bias in probabilities. Also many possible models, with contradictory implications for policy!
- d) Estimate by FIML (= Bayesian with flat priors)? **Bad small sample bias**/flat likelihood surfaces — poor statistical accuracy.

Introduction

- Rival suggestions today
- a) Back to Economic History and build ad hoc models 'reflecting elasticities' (e.g.SVARs)? **Violates** Lucas' Critique: useless policy evaluation.
- b) Build micro-founded models, calibrate and compare with stylised facts? **No** measure of statistical accuracy.
- c) Improve on above by using Bayesian methods, treating calibrated parameters as priors? **If priors wrong**, bias in probabilities. Also many possible models, with contradictory implications for policy!
- d) Estimate by FIML (= Bayesian with flat priors)? **Bad small sample bias**/flat likelihood surfaces — poor statistical accuracy.
- a)-d) all widely used today. For policymakers not useful: in practice they use 'suite of models' and rely on own beliefs.

This talk

- Introducing new practical powerful method for policymakers to test models and reforms: can reforms fail?

This talk

- Introducing new practical powerful method for policymakers to test models and reforms: can reforms fail?
- Method 'Indirect Inference': testing/estimating model on ability to replicate chosen relevant macro behaviour on Friedman's 'as if' principle.

This talk

- Introducing new practical powerful method for policymakers to test models and reforms: can reforms fail?
- Method 'Indirect Inference': testing/estimating model on ability to replicate chosen relevant macro behaviour on Friedman's 'as if' principle.
- Power lost in 'direct inference' (model estimated directly on data) such as FIML and Likelihood Ratio tests- will explain.

This talk

- Introducing new practical powerful method for policymakers to test models and reforms: can reforms fail?
- Method 'Indirect Inference': testing/estimating model on ability to replicate chosen relevant macro behaviour on Friedman's 'as if' principle.
- Power lost in 'direct inference' (model estimated directly on data) such as FIML and Likelihood Ratio tests- will explain.
- Current practices developed in Cardiff over past decade or more with coauthors Mike Wickens, David Meenagh, Mai Le and recently Yongdeng Xu; also many former PhD students (citations in paper).

This talk

- Introducing new practical powerful method for policymakers to test models and reforms: can reforms fail?
- Method ‘Indirect Inference’: testing/estimating model on ability to replicate chosen relevant macro behaviour on Friedman’s ‘as if’ principle.
- Power lost in ‘direct inference’ (model estimated directly on data) such as FIML and Likelihood Ratio tests- will explain.
- Current practices developed in Cardiff over past decade or more with coauthors Mike Wickens, David Meenagh, Mai Le and recently Yongdeng Xu; also many former PhD students (citations in paper).
- Matlab Programme INDIRECT now available for download, plus supporting manual, papers etc. — www.patrickminford.net/Indirect

Indirect v. Direct inference

- Direct Inference: fit DSGE model, DSGEM, to data (FIML) minimising forecast error. Likelihood Ratio tests model's relative forecasting errors.

Indirect v. Direct inference

- Direct Inference: fit DSGE model, DSGEM, to data (FIML) minimising forecast error. Likelihood Ratio tests model's relative forecasting errors.
- Indirect Inference: describe data in terms of relevant 'dynamic facts' ('Auxiliary model', AUXM). Could simulated DSGEM generate same behaviour?

Indirect v. Direct inference

- Direct Inference: fit DSGE model, DSGEM, to data (FIML) minimising forecast error. Likelihood Ratio tests model's relative forecasting errors.
- Indirect Inference: describe data in terms of relevant 'dynamic facts' ('Auxiliary model', AUXM). Could simulated DSGEM generate same behaviour?
- AUXM can include: key Impulse Response Functions; moments and cross-moments. Here use VAR of key variables: approx to DSGEM VARMA reduced form (source of IRFs etc).

Indirect v. Direct inference

- Direct Inference: fit DSGE model, DSGEM, to data (FIML) minimising forecast error. Likelihood Ratio tests model's relative forecasting errors.
- Indirect Inference: describe data in terms of relevant 'dynamic facts' ('Auxiliary model', AUXM). Could simulated DSGEM generate same behaviour?
- AUXM can include: key Impulse Response Functions; moments and cross-moments. Here use VAR of key variables: approx to DSGEM VARMA reduced form (source of IRFs etc).
- Direct and Indirect Inf. both tests of DSGEM's 'specification': compare properties below.

Model evaluation by II

- Parameters, $\widehat{\theta}$, of DSGEM taken as given (by estimation/calibration): treated as null hypothesis.

Model evaluation by II

- Parameters, $\widehat{\theta}$, of DSGEM taken as given (by estimation/calibration): treated as null hypothesis.
- Compare AUXM, here VAR, estimated from data with the same VAR estimated from simulations of DSGEM.

Model evaluation by II

- Parameters, $\widehat{\theta}$, of DSGEM taken as given (by estimation/calibration): treated as null hypothesis.
- Compare AUXM, here VAR, estimated from data with the same VAR estimated from simulations of DSGEM.
- Use statistical distribution of VAR coeffs (α_S) implied by DSGEM: find by bootstrapping DSGEM (i.e. simulating with model's own implied shocks say 1000 times) and re-estimating the α on each of the 1000 bootstrap samples.

Model evaluation by II

- Parameters, $\widehat{\theta}$, of DSGEM taken as given (by estimation/calibration): treated as null hypothesis.
- Compare AUXM, here VAR, estimated from data with the same VAR estimated from simulations of DSGEM.
- Use statistical distribution of VAR coeffs (α_S) implied by DSGEM: find by bootstrapping DSGEM (i.e. simulating with model's own implied shocks say 1000 times) and re-estimating the α on each of the 1000 bootstrap samples.
- Variation of these α_S give their var-covar matrix, $\Omega(\alpha[\widehat{\theta}])$.

Joint distribution implied by DSGEM illustrated for 2 VAR coefficients only (constructed distribution, not SW):

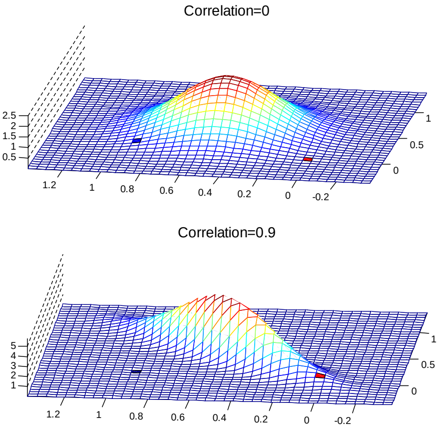


Figure: Bivariate Normal Distributions $(0.1, 0.9$ shaded blue and $0, 0$ shaded red) with correlation of 0 and 0.9.

Joint distribution implied by DSGEM illustrated for 2 VAR coefficients only (constructed distribution, not SW):

- Bottom is usual case: VAR coeffs covary. i.e. $\Omega(\alpha[\widehat{\theta}])$ has non-zero off-diagonal elements. e.g. AR coeffs of inflation and interest rates covary due to Fisher equation (restriction of DSGEM).

Joint distribution implied by DSGEM illustrated for 2 VAR coefficients only (constructed distribution, not SW):

- Bottom is usual case: VAR coeffs covary. i.e. $\Omega(\alpha[\widehat{\theta}])$ has non-zero off-diagonal elements. e.g. AR coeffs of inflation and interest rates covary due to Fisher equation (restriction of DSGEM).
- PDF of k simulated VAR coeffs, α_S, given by joint normal $\frac{1}{(2\pi)^k \left|\Omega(\alpha[\widehat{\theta}])\right|} e^{-0.5[a_s - \alpha_S(\widehat{\theta})]' \{\Omega(\alpha[\widehat{\theta}])\}^{-1} [a_s - \alpha_S(\widehat{\theta})]}$.

Joint distribution implied by DSGEM illustrated for 2 VAR coefficients only (constructed distribution, not SW):

- Bottom is usual case: VAR coeffs covary. i.e. $\Omega(\alpha[\widehat{\theta}])$ has non-zero off-diagonal elements. e.g. AR coeffs of inflation and interest rates covary due to Fisher equation (restriction of DSGEM).
- PDF of k simulated VAR coeffs, α_S, given by joint normal $\frac{1}{(2\pi)^k \left|\Omega(\alpha[\widehat{\theta}])\right|} e^{-0.5[a_s-\alpha_S(\widehat{\theta})]'\{\Omega(\alpha[\widehat{\theta}])\}^{-1}[a_s-\alpha_S(\widehat{\theta})]}$.
- Exponent is (-0.5 times) Wald statistic (IIW) based on the bootstrap distribution (implied by the assumed DSGEM coefficients $\widehat{\theta}$) of a_S around their bootstrap means, $\alpha_S(\widehat{\theta})$: $[a_s - \alpha_S(\widehat{\theta})]'\{\Omega(\alpha[\widehat{\theta}])\}^{-1}[a_s - \alpha_S(\widehat{\theta})]$. Has approx Chi-squared distribution (k).

Joint distribution implied by DSGEM illustrated for 2 VAR coefficients only (constructed distribution, not SW):

- Bottom is usual case: VAR coeffs covary. i.e. $\Omega(\alpha[\widehat{\theta}])$ has non-zero off-diagonal elements. e.g. AR coeffs of inflation and interest rates covary due to Fisher equation (restriction of DSGEM).
- PDF of k simulated VAR coeffs, α_S, given by joint normal $\frac{1}{(2\pi)^k \left|\Omega(\alpha[\widehat{\theta}])\right|} e^{-0.5[a_s - \alpha_S(\widehat{\theta})]'\{\Omega(\alpha[\widehat{\theta}])\}^{-1}[a_s - \alpha_S(\widehat{\theta})]}$.
- Exponent is (-0.5 times) Wald statistic (IIW) based on the bootstrap distribution (implied by the assumed DSGEM coefficients $\widehat{\theta}$) of a_S around their bootstrap means, $\alpha_S(\widehat{\theta})$: $[a_s - \alpha_S(\widehat{\theta})]'\{\Omega(\alpha[\widehat{\theta}])\}^{-1}[a_s - \alpha_S(\widehat{\theta})]$. Has approx Chi-squared distribution (k).
- Test: reject if Wald with data-estimated α_T: $[a_T - \alpha_S(\widehat{\theta})]'\{\Omega(\alpha[\widehat{\theta}])\}^{-1}[a_T - \alpha_S(\widehat{\theta})]$ exceeds W_C (critical value from IIW distribution).

Wald statistic for joint distribution of k=30 coefficients (Smets Wouters DSGEM)

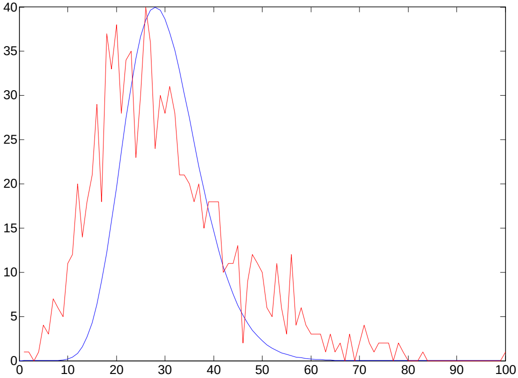


Figure: Histogram of Wald statistic for SW model, with asymptotic Chi-squared distribution for k=30

Model estimation by II

- Move $\widehat{\theta}$ of DSGEM around until IIW minimised: i.e. DSGEM simulated behaviour as close to that of AUXM as possible.

Model estimation by II

- Move $\widehat{\theta}$ of DSGEM around until IIW minimised: i.e. DSGEM simulated behaviour as close to that of AUXM as possible.
- Properties of II estimator of DSGEM parameters: consistent and asymptotically normal (regardless of data features included in AUXM).

Model estimation by II

- Move $\widehat{\theta}$ of DSGEM around until IIW minimised: i.e. DSGEM simulated behaviour as close to that of AUXM as possible.
- Properties of II estimator of DSGEM parameters: consistent and asymptotically normal (regardless of data features included in AUXM).
- Refs Estimation: Smith (1993), Gregory and Smith (1991, 1993), Gourieroux et al. (1993), Gourieroux and Montfort (1995), Canova (2005). Not testing or small sample properties.

Model estimation by II

- Move $\widehat{\theta}$ of DSGEM around until IIW minimised: i.e. DSGEM simulated behaviour as close to that of AUXM as possible.
- Properties of II estimator of DSGEM parameters: consistent and asymptotically normal (regardless of data features included in AUXM).
- Refs Estimation: Smith (1993), Gregory and Smith (1991, 1993), Gourieroux et al. (1993), Gourieroux and Montfort (1995), Canova (2005). Not testing or small sample properties.
- Various ways to get estimator: we use search algorithms (typically Simulated Annealing), find estimator halves FIML small sample bias.

The power of IIW and LR tests

- Generate 1000 Monte Carlo samples from Smets Wouters, SW, model as estimated for post-war US sample.

The power of IIW and LR tests

- Generate 1000 Monte Carlo samples from Smets Wouters, SW, model as estimated for post-war US sample.
- Stationary shocks/samples. Parallel set of non-stationary shocks/samples: here use VECM instead of VAR as AUXM.

The power of IIW and LR tests

- Generate 1000 Monte Carlo samples from Smets Wouters, SW, model as estimated for post-war US sample.
- Stationary shocks/samples. Parallel set of non-stationary shocks/samples: here use VECM instead of VAR as AUXM.
- Set up series of nulls ranked by degree of (general) 'falseness': perturb parameters by $+$ or $-$ $x\%$ alternately.

The power of IIW and LR tests

- Generate 1000 Monte Carlo samples from Smets Wouters, SW, model as estimated for post-war US sample.
- Stationary shocks/samples. Parallel set of non-stationary shocks/samples: here use VECM instead of VAR as AUXM.
- Set up series of nulls ranked by degree of (general) 'falseness': perturb parameters by $+$ or $-$ $x\%$ alternately.
- IIW: perturb θ (incl. ρ) and 2nd/3rd/4th moments of innovations.

The power of IIW and LR tests

- Generate 1000 Monte Carlo samples from Smets Wouters, SW, model as estimated for post-war US sample.
- Stationary shocks/samples. Parallel set of non-stationary shocks/samples: here use VECM instead of VAR as AUXM.
- Set up series of nulls ranked by degree of (general) 'falseness': perturb parameters by $+$ or $-$ $x\%$ alternately.
- IIW: perturb θ (incl. ρ) and 2nd/3rd/4th moments of innovations.
- LR: perturb θ (excl. ρ). Shocks/ρ extracted/estimated for each sample; standard to avoid DSGEM going 'off track'.

The power of IIW and LR tests

- Generate 1000 Monte Carlo samples from Smets Wouters, SW, model as estimated for post-war US sample.
- Stationary shocks/samples. Parallel set of non-stationary shocks/samples: here use VECM instead of VAR as AUXM.
- Set up series of nulls ranked by degree of (general) 'falseness': perturb parameters by $+$ or $-$ $x\%$ alternately.
- IIW: perturb θ (incl. ρ) and 2nd/3rd/4th moments of innovations.
- LR: perturb θ (excl. ρ). Shocks/ρ extracted/estimated for each sample; standard to avoid DSGEM going 'off track'.
- Power of test: percent of 1000 MC samples which reject null on upper 5% tail.

Monte Carlo results — stationary data

Model Null	IIW	LR
True	5.0	5.0
False by- 1(%)	19.8	6.3
3	52.1	8.8
5	87.3	13.1
7	99.4	21.6
10	100.0	53.4
15	100.0	99.3
20	100.0	99.7

Table: Rejection Rates for Wald and Likelihood Ratio for 3 Variable VAR(1) on STATIONARY DATA

Monte Carlo results — non-stationary data

Model Null	IIW	LR
True	5.0	5.0
False by- 1(%)	7.9	5.2
3	49.2	5.8
5	97.8	6.2
7	100	7.4
10	100	9.6
15	100	15.6
20	100	26.5

Table: Rejection Rates for Wald and Likelihood Ratio for 3 Variable VAR(1) on NON-STATIONARY DATA

General Conclusion on comparison

- The power of IIW massively higher than of LR in small samples.

General Conclusion on comparison

- The power of IIW massively higher than of LR in small samples.
- Especially with non-stationary data.

Why the power differences above?

- Two reasons why IIW more powerful than LR.

Why the power differences above?

- Two reasons why IIW more powerful than LR.

1. With LR *reestimated* ρs. By bringing DSGEM 'back on track' weakens test power.

Why the power differences above?

- Two reasons why IIW more powerful than LR.

1. With LR *reestimated* ρs. By bringing DSGEM 'back on track' weakens test power.
2. If LR test done with same ρ as IIW ('like-for-like') this avoided: but *still much less power than IIW*.

Why the power differences above?

- Two reasons why IIW more powerful than LR.

1. With LR *reestimated* ρs. By bringing DSGEM 'back on track' weakens test power.
2. If LR test done with same ρ as IIW ('like-for-like') this avoided: but *still much less power than IIW*.

- LR test asymptotically equivalent to Standard Wald test (transformation of LR) which uses the *data-estimated* variance matrix (from data-based VAR), $\Omega(\alpha_T)$.

Why the power differences above?

- Two reasons why IIW more powerful than LR.

1. With LR *reestimated* ρs. By bringing DSGEM 'back on track' weakens test power.
2. If LR test done with same ρ as IIW ('like-for-like') this avoided: but *still much less power than IIW*.

- LR test asymptotically equivalent to Standard Wald test (transformation of LR) which uses the *data-estimated* variance matrix (from data-based VAR), $\Omega(\alpha_T)$.
- Also get $\Omega(\alpha_T)$ by bootstrapping data-based VAR = ' Wald UNRestricted': LR and Wald UNR asymptotically equivalent.

Why the power differences above?

- Two reasons why IIW more powerful than LR.

1. With LR *reestimated* ρs. By bringing DSGEM 'back on track' weakens test power.
2. If LR test done with same ρ as IIW ('like-for-like') this avoided: but *still much less power than IIW*.

- LR test asymptotically equivalent to Standard Wald test (transformation of LR) which uses the *data-estimated* variance matrix (from data-based VAR), $\Omega(\alpha_T)$.
- Also get $\Omega(\alpha_T)$ by bootstrapping data-based VAR = ' Wald UNRestricted': LR and Wald UNR asymptotically equivalent.
- **But** IIW uses *model-restricted* variance matrix, $\Omega(\alpha[\widehat{\theta}])$: this *changes* with changing Falseness, $\Omega(\alpha_T)$ does not.

Why the power differences above?

Model Null		IIW	LR	LR-like-f-I	Wald UNR
	True	5.0	5.0	5.0	5.0
False by-	3(%)	52.1	8.8	21.8	7.5
	5	87.3	13.1	37.5	30.7
	7	99.4	21.6	58.9	75.0
	10	100.0	53.4	84.0	97.0
	15	100.0	99.3	99.0	100.0
	20	100.0	99.7	100.0	100.0

Table: Rejection Rates for Wald and Likelihood Ratio for 3 Variable VAR(1) using STATIONARY Data

Understanding power differences visually

- Wald UNR: could α_T distribution from unknown true model in data generate $\alpha_S(\widehat{\theta})$?

Understanding power differences visually

- Wald UNR: could α_T distribution from unknown true model in data generate $\alpha_S(\widehat{\theta})$?
- IIW: could α_S distribution from DSGEM generate α_T?

Understanding power differences visually

- Wald UNR: could α_T distribution from unknown true model in data generate $\alpha_S(\widehat{\theta})$?
- IIW: could α_S distribution from DSGEM generate α_T?
- Evaluating gap between α_T and $\alpha_S(\widehat{\theta})$ 'from different ends'.

Understanding power differences visually

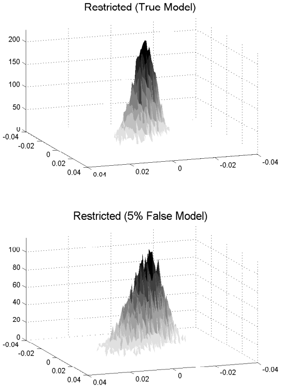


Figure: Restricted VAR and Unrestricted VAR Coefficient Distributions

Understanding power differences visually

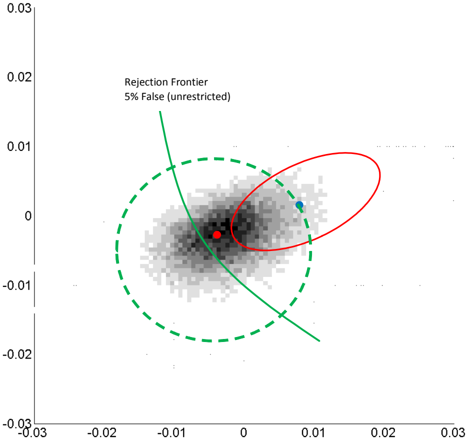
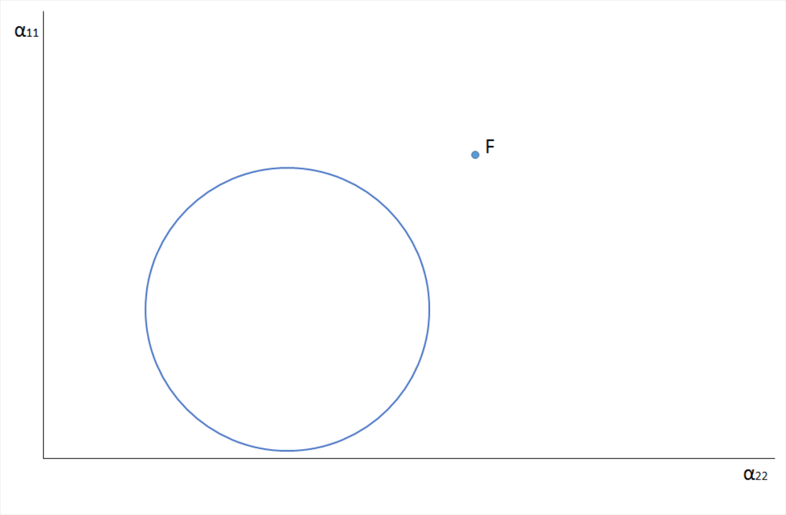
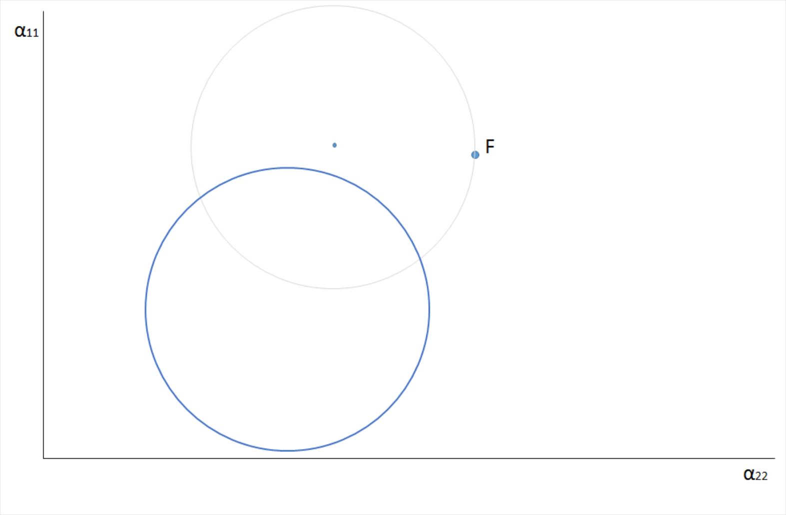


Figure: Two 95% contours for tests of 5% False Model — Green=Unrestricted; Red=Restricted.

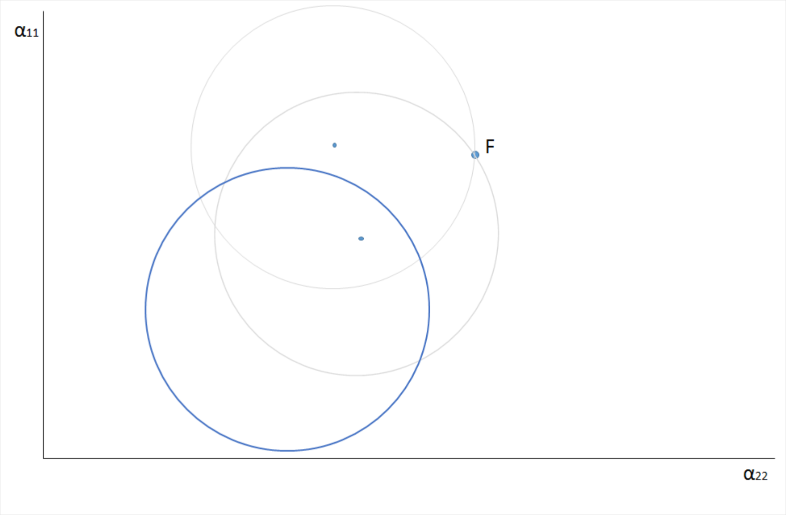
Understanding power differences visually



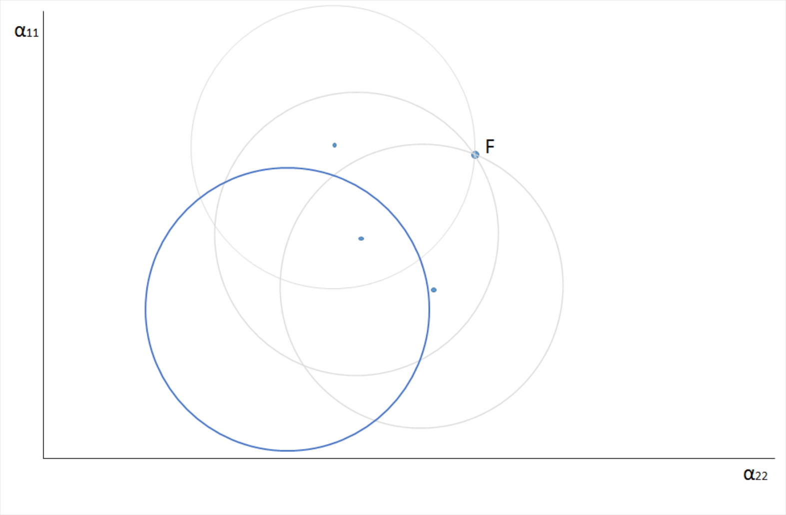
Understanding power differences visually



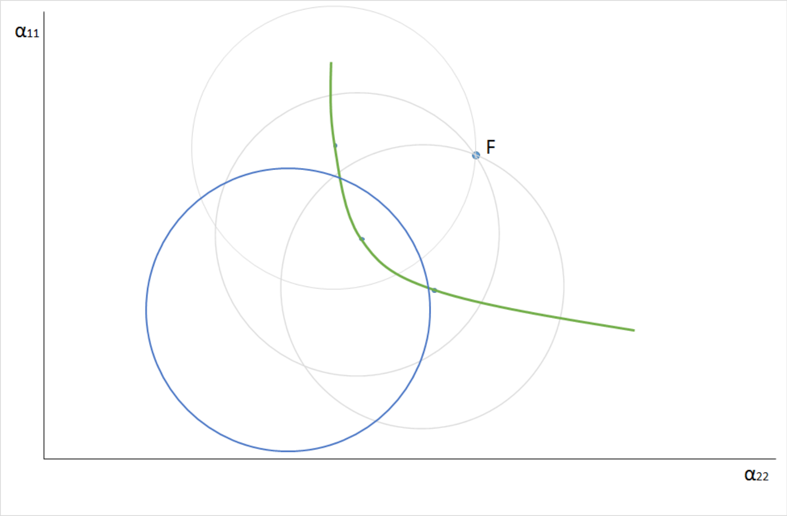
Understanding power differences visually



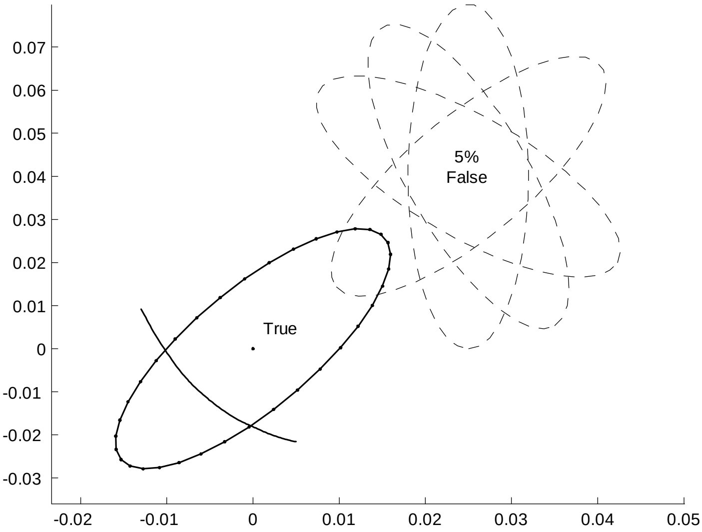
Understanding power differences visually



Understanding power differences visually



Understanding power differences visually



Why the power difference?

- Reason b) summarised: IIW a non-standard Wald test of unrestricted AUXM against DSGEM-restricted AUXM (as the null).

Why the power difference?

- Reason b) summarised: IIW a non-standard Wald test of unrestricted AUXM against DSGEM-restricted AUXM (as the null).
- Wald UNR and IIW asymptotically equivalent when model is true (stationary, no exog variables) but not when model is false (hence power difference).

Why the power difference?

- Reason b) summarised: IIW a non-standard Wald test of unrestricted AUXM against DSGEM-restricted AUXM (as the null).
- Wald UNR and IIW asymptotically equivalent when model is true (stationary, no exog variables) but not when model is false (hence power difference).
- IIW requires $\Omega(\alpha[\widehat{\theta}])$ which only calculable numerically as in INDIRECT.

Over-identification of DSGEM allows power to rise as VAR expanded

- IIW asks: could coeffs of data-based (approx) reduced form come from DSGEM? Is DSGEM 'footprint' the one found in data?

Over-identification of DSGEM allows power to rise as VAR expanded

- IIW asks: could coeffs of data-based (approx) reduced form come from DSGEM? Is DSGEM 'footprint' the one found in data?
- DSGEMs dealt with here highly over-identified (tested by Le et al, 2013, using IIW test; generate from DSGEM large samples and large number of replications; can we find NO other DSGEM also rejected by these at 5%? If so DSGEM identified).

Over-identification of DSGEM allows power to rise as VAR expanded

- IIW asks: could coeffs of data-based (approx) reduced form come from DSGEM? Is DSGEM 'footprint' the one found in data?
- DSGEMs dealt with here highly over-identified (tested by Le et al, 2013, using IIW test; generate from DSGEM large samples and large number of replications; can we find NO other DSGEM also rejected by these at 5%? If so DSGEM identified).
- Hence each VAR coeff is different nonlinear combination of DSGEM θ: more VAR coeffs mean more combinations for DSGEM to match.

Over-identification of DSGEM allows power to rise as VAR expanded

- IIW asks: could coeffs of data-based (approx) reduced form come from DSGEM? Is DSGEM 'footprint' the one found in data?
- DSGEMs dealt with here highly over-identified (tested by Le et al, 2013, using IIW test; generate from DSGEM large samples and large number of replications; can we find NO other DSGEM also rejected by these at 5%? If so DSGEM identified).
- Hence each VAR coeff is different nonlinear combination of DSGEM θ: more VAR coeffs mean more combinations for DSGEM to match.
- Can exploit this to raise power of IIW test — eg by raising order of VAR, by adding variances of VAR errors, by adding more variables to VAR.

Over-identification of DSGEM allows power to rise as VAR expanded

- IIW asks: could coeffs of data-based (approx) reduced form come from DSGEM? Is DSGEM 'footprint' the one found in data?
- DSGEMs dealt with here highly over-identified (tested by Le et al, 2013, using IIW test; generate from DSGEM large samples and large number of replications; can we find NO other DSGEM also rejected by these at 5%? If so DSGEM identified).
- Hence each VAR coeff is different nonlinear combination of DSGEM θ: more VAR coeffs mean more combinations for DSGEM to match.
- Can exploit this to raise power of IIW test — eg by raising order of VAR, by adding variances of VAR errors, by adding more variables to VAR.
- 'Full Wald' puts in all variables, full order of reduced form VAR, and all VAR error variances: huge power. In practice choose 'Directed Wald'.

Over-identification of DSGEM allows power to rise as VAR expanded

- IIW asks: could coeffs of data-based (approx) reduced form come from DSGEM? Is DSGEM 'footprint' the one found in data?
- DSGEMs dealt with here highly over-identified (tested by Le et al, 2013, using IIW test; generate from DSGEM large samples and large number of replications; can we find NO other DSGEM also rejected by these at 5%? If so DSGEM identified).
- Hence each VAR coeff is different nonlinear combination of DSGEM θ: more VAR coeffs mean more combinations for DSGEM to match.
- Can exploit this to raise power of IIW test — eg by raising order of VAR, by adding variances of VAR errors, by adding more variables to VAR.
- 'Full Wald' puts in all variables, full order of reduced form VAR, and all VAR error variances: huge power. In practice choose 'Directed Wald'.
- Gives rise to trade-off between power and tractability.

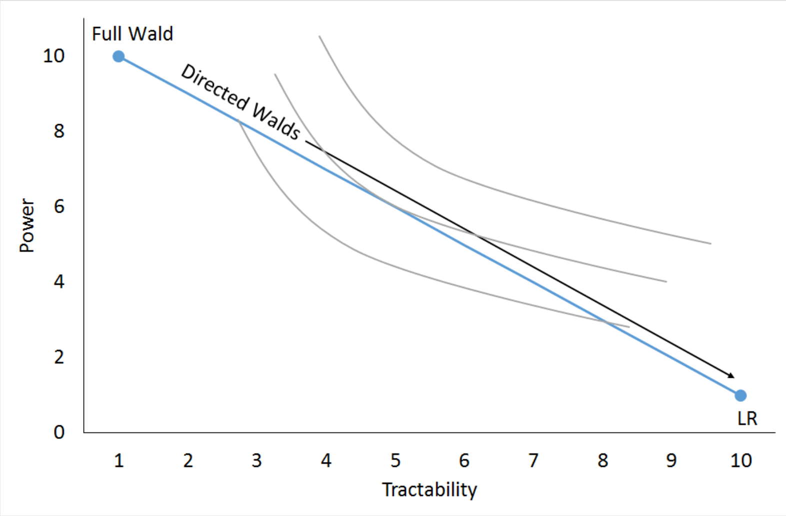


Figure: Maximising Friedman utility

- **Application to monetary policy reform.** Le et al (2014) reworking of SW model — US non-stationary data from 1980s, hybrid, banking+money, zero bound: passes IIW test 3 variable VAR1 (incl 3 error variances).

Application to DSGEM about monetary reform

- **Application to monetary policy reform.** Le et al (2014) reworking of SW model — US non-stationary data from 1980s, hybrid, banking+money, zero bound: passes IIW test 3 variable VAR1 (incl 3 error variances).
- Policy reforms (monetary base rule, Taylor Rules targeting Price Level or Nominal GDP) evaluated by numbers of 'crises' (crisis= GDP fall, 3 years before returns to start level).

Application to DSGEM about monetary reform

- **Application to monetary policy reform.** Le et al (2014) reworking of SW model — US non-stationary data from 1980s, hybrid, banking+money, zero bound: passes IIW test 3 variable VAR1 (incl 3 error variances).
- Policy reforms (monetary base rule, Taylor Rules targeting Price Level or Nominal GDP) evaluated by numbers of 'crises' (crisis= GDP fall, 3 years before returns to start level).
- **Use tests to establish worst case scenarios for policy reforms. If reforms still viable, policymakers happy!**

Application to DSGEM about monetary reform

Crises per 1000 yrs	Base case	Monetary Reform	PLT	NGDPT	PLT+ M. Refm	NGDPT+ M. Refm
MODEL						
True	20.8	6.62	2.15	1.83	1.41	1.31
IIW: 7% False	57.4	18.6	10.3	8.7	11.8	10.3
LR: 50% F	70.4	*Explosive*	33.3	33.4	34.4	34.2

Notes:
Base Case: monetary policies as estimated over the sample period;
Monetary Reform: Monetary Base rule (responds to credit premium) + Taylor Rule;
PLT:substituting Price Level Target for Inflation Target in Taylor Rule;
NGDPT: substituting Nominal GDP target for inflation and output targets in Taylor Rule.

Table: Policy analysis when model have varying falseness

Application to DSGEM about monetary reform

- Power function IIW gives 100% rejection rate if 7% False: worst case scenario, crisis every 100 years. No need for distortionary macro-pru.

Application to DSGEM about monetary reform

- Power function IIW gives 100% rejection rate if 7% False: worst case scenario, crisis every 100 years. No need for distortionary macro-pru.
- Power function LR gives 100% rejection rate if 50% False: worst case scenario. Gives crisis every 30 years. Need macro-pru to be safe.

Application to DSGEM about monetary reform

- Power function IIW gives 100% rejection rate if 7% False: worst case scenario, crisis every 100 years. No need for distortionary macro-pru.
- Power function LR gives 100% rejection rate if 50% False: worst case scenario. Gives crisis every 30 years. Need macro-pru to be safe.
- **Illustration only!** Policymakers should redo Monte Carlo analysis for particular model/data set. Power functions could alter.

Application to DSGEM about monetary reform

- Power function IIW gives 100% rejection rate if 7% False: worst case scenario, crisis every 100 years. No need for distortionary macro-pru.
- Power function LR gives 100% rejection rate if 50% False: worst case scenario. Gives crisis every 30 years. Need macro-pru to be safe.
- **Illustration only!** Policymakers should redo Monte Carlo analysis for particular model/data set. Power functions could alter.
- Explore possible 'evil agent' combinations for power/robustness.

Application to DSGEM about monetary reform

- Power function IIW gives 100% rejection rate if 7% False: worst case scenario, crisis every 100 years. No need for distortionary macro-pru.
- Power function LR gives 100% rejection rate if 50% False: worst case scenario. Gives crisis every 30 years. Need macro-pru to be safe.
- **Illustration only!** Policymakers should redo Monte Carlo analysis for particular model/data set. Power functions could alter.
- Explore possible 'evil agent' combinations for power/robustness.
- Procedure here for getting realistic idea of risks for policy reforms so can adapt as necessary.

Conclusions

- IIW method has good power when relevant features of interest chosen by policymakers when contemplating a policy reform.

Conclusions

- IIW method has good power when relevant features of interest chosen by policymakers when contemplating a policy reform.
- Able to explore relevant weaknesses of model for policy and establish *with sureness* whether these will upset policy reform.

Conclusions

- IIW method has good power when relevant features of interest chosen by policymakers when contemplating a policy reform.
- Able to explore relevant weaknesses of model for policy and establish *with sureness* whether these will upset policy reform.
- Programmes available for download to implement in user-friendly way. Happy to help early users and get feedback!