Causality in Economics: An empirical investigation

Asad Zaman and Zahid Asghar

- One of the major goals of empirical research in Economics: Determination of causal relationship.
- Economists ,econometricians and econometric statisticians have long been studying the issue of causality and causal laws, i.e., the issue of identifying a causal relation between an outcome and a set of factors that may have determined the outcome.

- Debate on Causality in economics dates backs to Hume (1752) when he explored the relationship between money and prices and David Hume was of the view that everything can be verified from what is observed
- Title of Adam Smith's book "An Enquiry into the Nature and Causes of the Wealth of Nations" provides evidence that causality concept is crucial in Economics.
- Ricardo and Stuart Mill were also explicitly involved in causality issues.

- There has been lot of controversy about causality and related concepts in Philosophy
- No consensus on definition of causality in Philosophy
- Why philosophers don’t need and economists need it

- Econometric modeling to find out the relation between different variables.
- Objectives of Econometrics Modeling

Description (Means,sd,correlation)

Forecasting (whole linear comb)

Causation and Policy analysis

- First two normally achieved easily
- Y=a+b1*x1+b2*x2+b3*x3+e
- Y=earnings,x1:Years of edu,x2:age,x3:age^2

- Experimental Studies
- Observational Studies
- Relatively simple to find in the former
- We are here interested in Observational Studies.

- Correlation symmetric, causation asymmetric
- Cholesterol level and Lung Cancer (Smoking)
- Shoe Size and Learning Skills
- Poor and Crime
- Export and Economic Growth (Money,FDI)etc.
- Love marriages and Divorce rate

- Main approaches
- Granger Causality (widely in use now)
- Structural causality
- Suppes' approach
- Graph theoretic approach.

Method for Detecting Structural Causality

- Given a bivariate series (X, Y) there are three possibilities for causality:
- (1) X &Y are jointly determined,
- (2) first X is determined and then Y is calculated from some equation like $Y = a + bX + u$, or
- (3) first Y is determined and then X is calculated from some equation like $X = c + d\,Y + v$.
- All three possibilities are observationally equivalent – data series generated by (1), (2) & (3) will be identical in all respects and hence it is impossible to detect causality by looking at the data *so long as there is no structural change.*
- Thus in a stable environment, it is impossible to tell whether Y causes X or whether X causes Y or whether there is mutual bi-directional causality. When there is some structural change, it will reveal the causal patterns provided that we look carefully.

- For example, suppose that the variance of X increases. If Y is caused by X, then there will be no change in the conditional distribution of Y given X. However, the conditional distribution of X given Y will change. Also the joint distribution of X and Y will change. So of the three possibilities listed above, only the causally correct one – number (2) – will stay the same after the structural change. From this we learn that causally correct relationships can survive certain types of structural change. This information can be used to differentiate between models which are causally correct and those which are not in period of structural change. In periods where we have stability and no structural changes, even models with incorrect causality will perform well.

- Let $X=\alpha+\epsilon$
- $Y=\beta+\delta*X+u$
- Where $\epsilon \sim niid(0,\sigma_{\epsilon}^2)$ and $u \sim niid(0,\sigma_{u}^2)$
- four probability distributions namely conditional of X given Y, Marginal of X , Conditional of Y given x and marginal of Y.

- $E(X) = \alpha, \quad V(X) = \sigma_{\epsilon}^2$
- $E(Y) = \beta + \delta\alpha, \; V(Y) = \delta^2\sigma_{\epsilon}^2 + \sigma_u^2$
- $\text{Covariance}(X,Y) = \delta\,\sigma_{\epsilon}^2$
- Now for conditional distribution of X we have
- $E(X/Y=y) = \alpha + (\delta\,\sigma_{\epsilon}^2/(\,\delta^2\sigma_{\epsilon}^2 + \sigma_u^2))*(Y - \beta - \delta\alpha)$

  $= \alpha\,\sigma_u^2 - \beta\delta\,\sigma_u^2 + (\delta\,\sigma_{\epsilon}^2/(\,\delta^2\sigma_{\epsilon}^2 + \sigma_u^2))*Y$

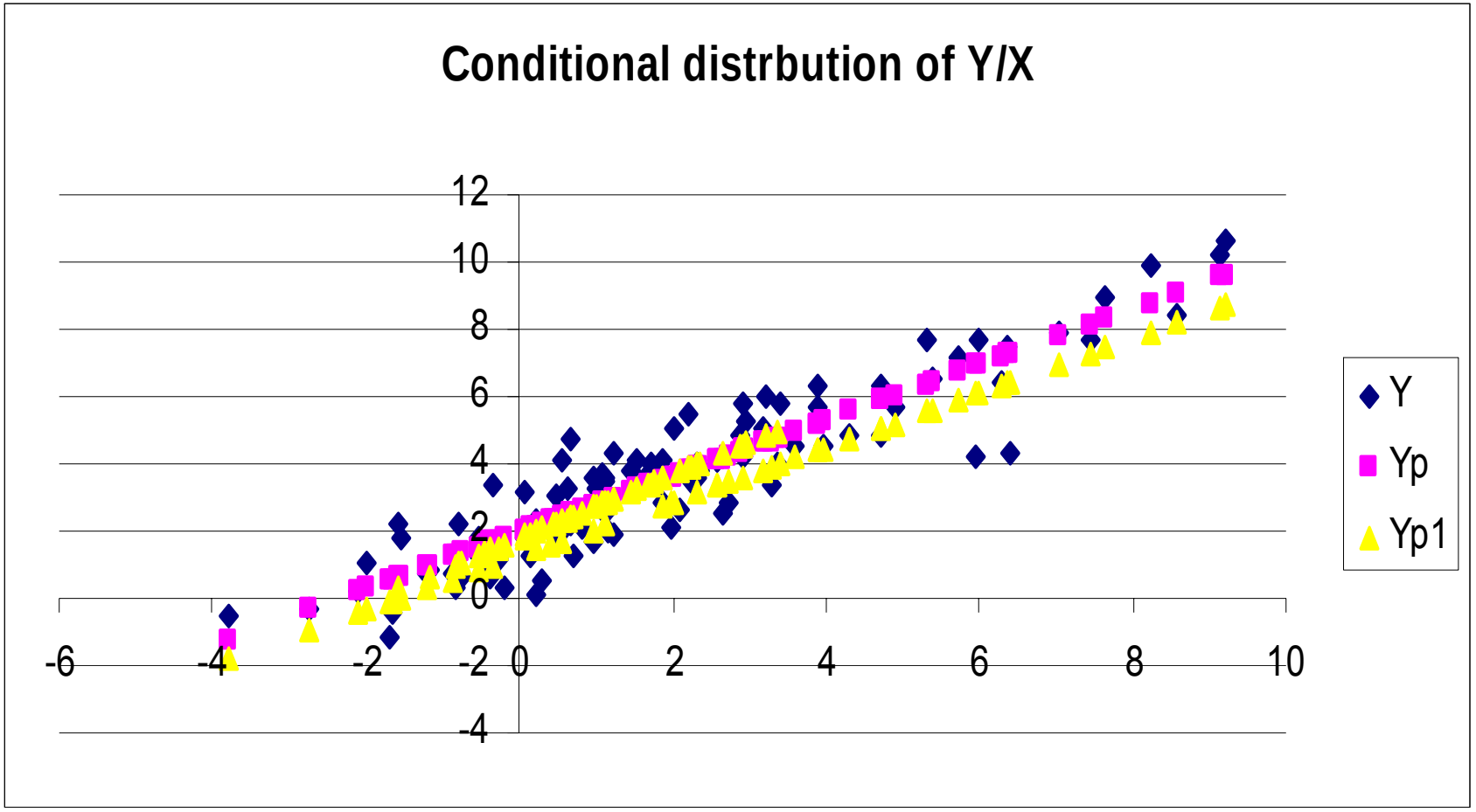
- Variannce(X/Y=y)= $\sigma_\epsilon^2 \sigma_u^2/(\delta^2\sigma_\epsilon^2+ \sigma_u^2)$
- f(X/Y)=N($\alpha \sigma_u^2 - \beta\delta \sigma_u^2 + (\delta \sigma_\epsilon^2/(\delta^2\sigma_\epsilon^2+ \sigma_u^2))$*Y, $\sigma_\epsilon^2 \sigma_u^2/(\delta^2\sigma_\epsilon^2+ \sigma_u^2)$)
- E(Y/X)=$\beta+\delta\alpha+(\delta \sigma_\epsilon^2/(\sigma_\epsilon^2))(X-\alpha)$

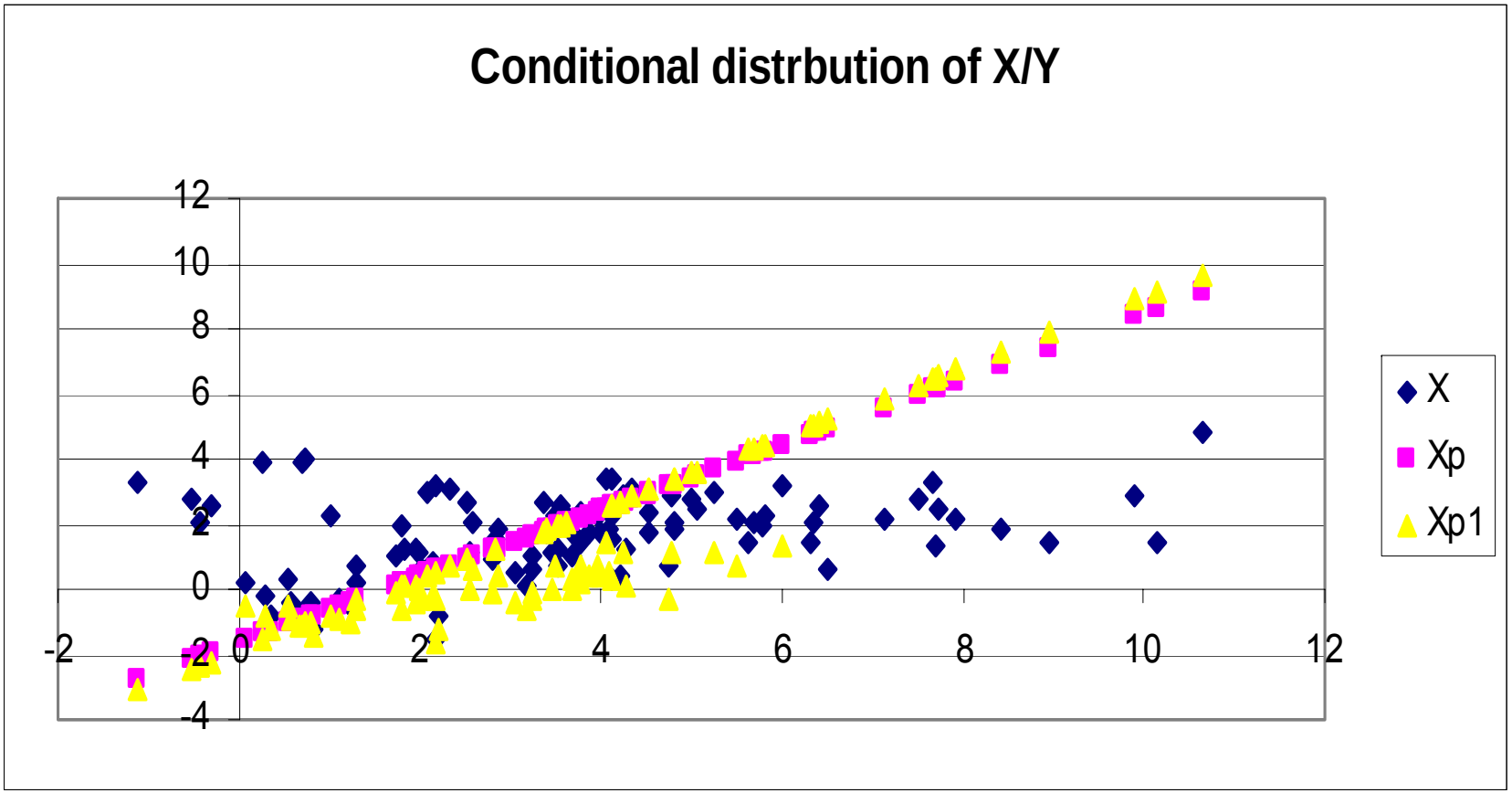
  $=\beta+\delta X$
- V(Y/X)=$(\delta^2\sigma_\epsilon^2+ \sigma_u^2)(1-(\delta \sigma_\epsilon^2)2/(\delta^2\sigma_\epsilon^2+ \sigma_u^2) \sigma_\epsilon^2)$
- $= \sigma_u^2$

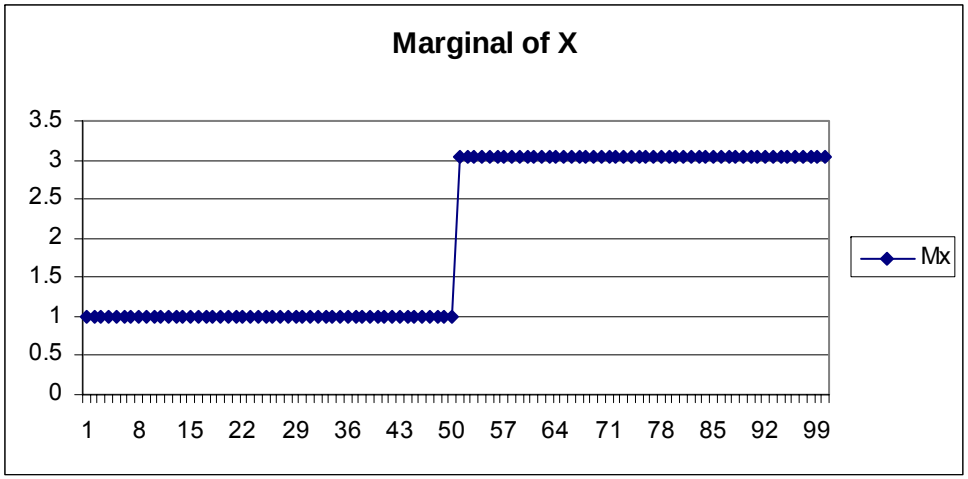
- $X=\alpha+\epsilon$
- $Y=\beta+\delta*X+u$
- $g(Y/X)=N(\beta+\delta X, \sigma_u^2)$
- $f(X/Y)=N(\alpha\ \sigma_u^2-\beta\delta\ \sigma_u^2+ (\delta\ \sigma_\epsilon^2/(\ \delta^2\sigma_\epsilon^2+ \sigma_u^2))*Y, \sigma_\epsilon^2\ \sigma_u^2/(\ \delta^2\sigma_\epsilon^2+\ \sigma_u^2))$
- $f(X)=N(\alpha, \sigma_\epsilon^2)$
- $h(Y)=N(\beta+\ \delta\alpha, \delta^2\sigma_\epsilon^2+\ \sigma_\epsilon^2)$

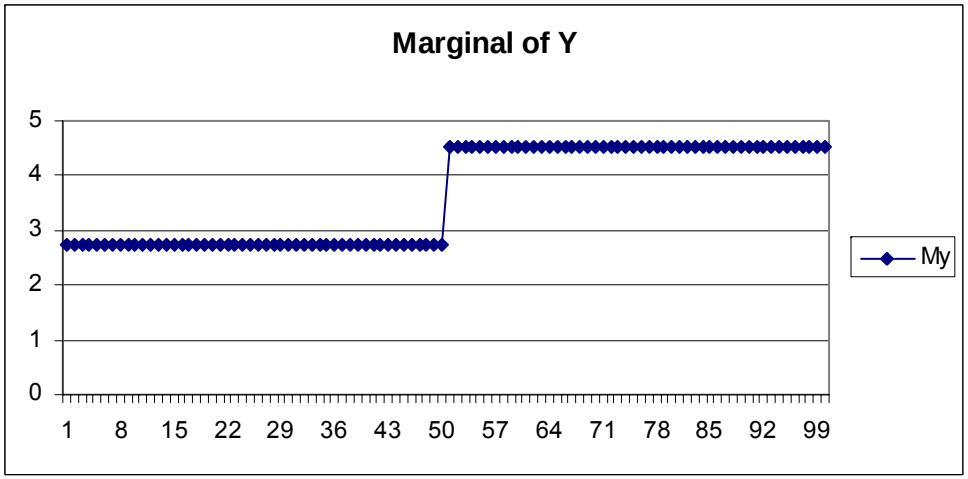
3 Evidence from Simulated Data

- X=1+ei and Y=2+0.8*X+ej, : both random errors are N(0,1) and covariance is zero between these two error terms. 100 observations on X and Y.
- We can not make a decision that whether its X which is causing Y or vice versa
- Then we generate first 50 values of X with ei~N(0,1) and next 50 with ei~N(2,3)
- Change is only in X


Conditional distrbution of Y/X
12
10
8
6
4
2
0
-2
-4
-6
-4
-2
0
2
4
6
8
10
Y
Yp
Yp1


Conditional distrbution of X/Y
12
10
8
6
4
2
0
-2
-4
-2
0
2
4
6
8
10
12
X
Xp
Xp1


Marginal of X
3.5
3
2.5
2
1.5
1
0.5
0
1 8 15 22 29 36 43 50 57 64 71 78 85 92 99
Mx


Marginal of Y
5
4
3
2
1
0
1 8 15 22 29 36 43 50 57 64 71 78 85 92 99
My

- This approach for testing causality requires lot of investigation in the underlying economic mechanism not only on theoretical grounds but also in historical prospective. How to find the period of intervention is a question of considerable importance.
- Extra statistical information as the change in Government policies, minutes of the Central Bank's monetary policy etc, signals an intervention in the investment policies, money-supply process. Purely statistical or econometric information is unlikely to be sufficient to identify an intervention.

- Hoover (2001) mentions that this intervention should be traced in historical prospective and then statistical tests should also be carried out to validate that whether intervention is there. Freedman(1991) also pointed out that determining causal direction requires an in depth knowledge of the problem at hand.

How to Proceed

- Have some knowledge from history on intervention in a series
- Apply some statistical test(e.g. Chow Test etc) to verify that intervention
- If chronological intervention is supported by statistical tests then apply regression on two data sets separately.
- The stable conditional distribution will be probably the true causal relation
- If such interventions exist for both of the variables at a particular time period then we can not find causal direction by such tests.

GDP and Macroceconomic Variables

Recently one of the most debated topic is how to determine the nature of relationship among different variables e.g between export and economic growth, money and economic growth, energy consumption and GDP, FDI and Pollution, etc.

- For example recently one of the most debated topic is how to determine the nature of relationship between exports and national output. So far very little consensus has emerged. Central question in this debate is whether strong economic performance is export led or growth driven .
- proponents of ELG argue that growth of export has a stimulating influence across the economy as a whole in form of technological spillover and other externalities. Models by Grosman and Helpman (1991), Riveria-Batiz and Romer (1991), and Romer (1991) posit that expanded international trade increase the number of specialized imports, increasing growth rates as economies become open to international trade. Knowledge externalities, spillover affect etc

- There are very strong arguments which are put forward to support ELG hypothesis theoretically. From a demand side perspective, sustained demand growth in a small domestic economy can not maintained permanently since domestic demand exhausts very soon. On the contrary, export markets are limitless and hence there is no need for any restriction on output. Thus export can serve as a catalyst for income growth, as a component of aggregate demand.[1] Pakistan at the moment is facing this problem that it has sustained high growth rate for the last four five years through domestic demand by having more cars, investment in Telecom, Stock Market but without any reasonable increase in export and therefore, it can not continue this journey of high growth rate for very long.

- In addition to this direct demand side effect, export expansion may have an indirect affect by providing foreign exchange which allows for having more capital import.[2]This increase in capital goods in turn boost economic growth by raising the level of capital formation. As Chuang(1998) argues that capital goods import from technological advanced countries may increase productivity and thereby growth, since knowledge and technology is embodied in equipment and machinery and therefore transferred through international trade.
-
  [1] Agosin(1999)
- [2] Reizman, Summers, and Whiteman (1996)

- On theoretical grounds there are several possible channels through which exports can enhance productivity. A country can promote specialization in areas where it has comparative advantage through export expansion, and lead to reallocation of resources from the relatively inefficient non-trade sector to the more productive sector.

- Secondly, the growth of exports can increase productivity by offering larger economies of scale (Helpman and Krugman(1995).
- Thirdly,total factor productivity may increase through dynamic spillover effects on the rest of the economy when there is export growth. The possible sources of these knowledge externalities include productivity enhancements resulting from increased competitiveness, more efficient management styles, better forms of organizations, labor training , and knowledge about technology and intellectual markets (Chuang 1998). In short export growth has beneficial impact on output growth .[3]
- [3] For more detail Hertzer et al (2006)

- Pro GDP: have noted that an increase in GDP generally leads to a corresponding expansion of trade unless the pattern of growth induced supply and corresponding demand create an anti-trade bias. Neoclassical theory typically stresses the causality that runs from factor endowment and productivity to the supply of exports (e.g. Findlay 1984).
- A rise in total factor productivity should increase out put and through lower power per unit cost increase export; so there are very strong arguments on both sides.

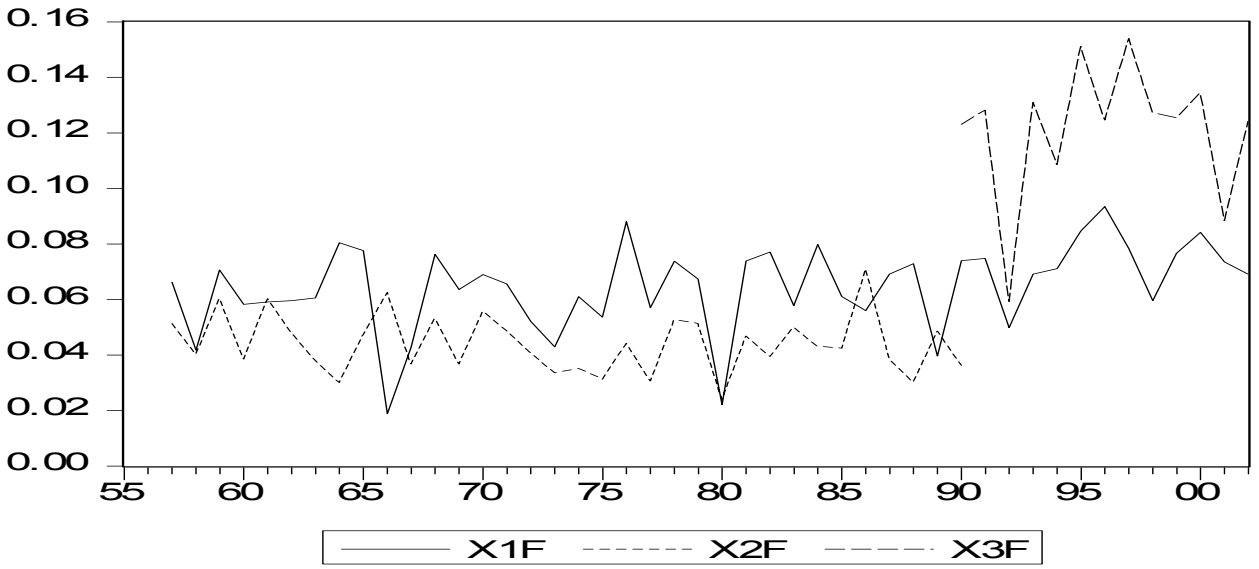
Causality test for India

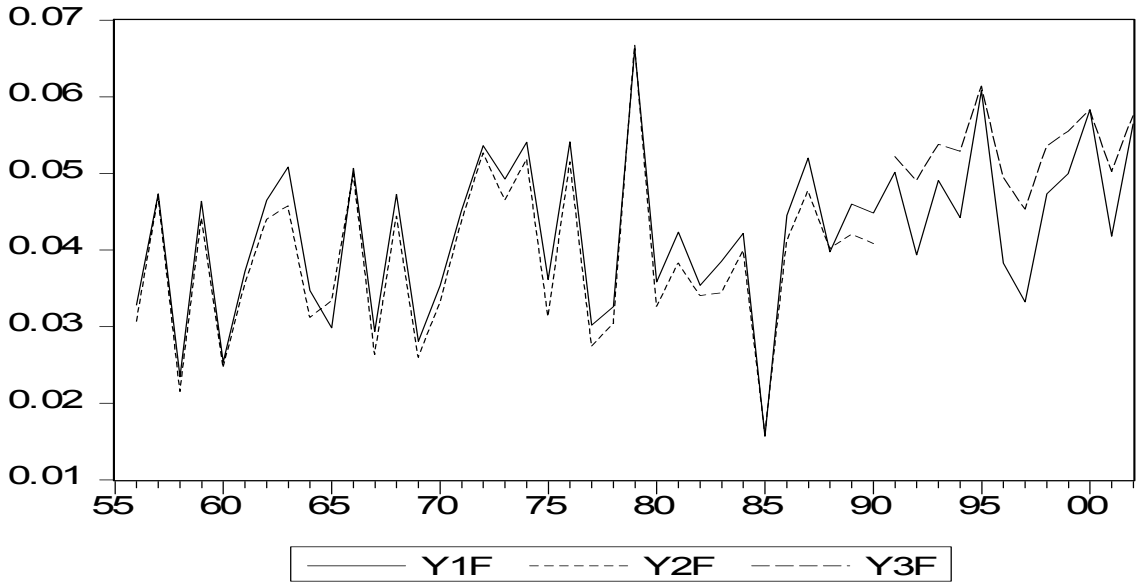
- History
- Three major events in Indian history from economic point of view
- 1965 war
- 1979 economic recession
- 1990s policy of open economy
- Both series export and GDP are tested at these three point for intervention

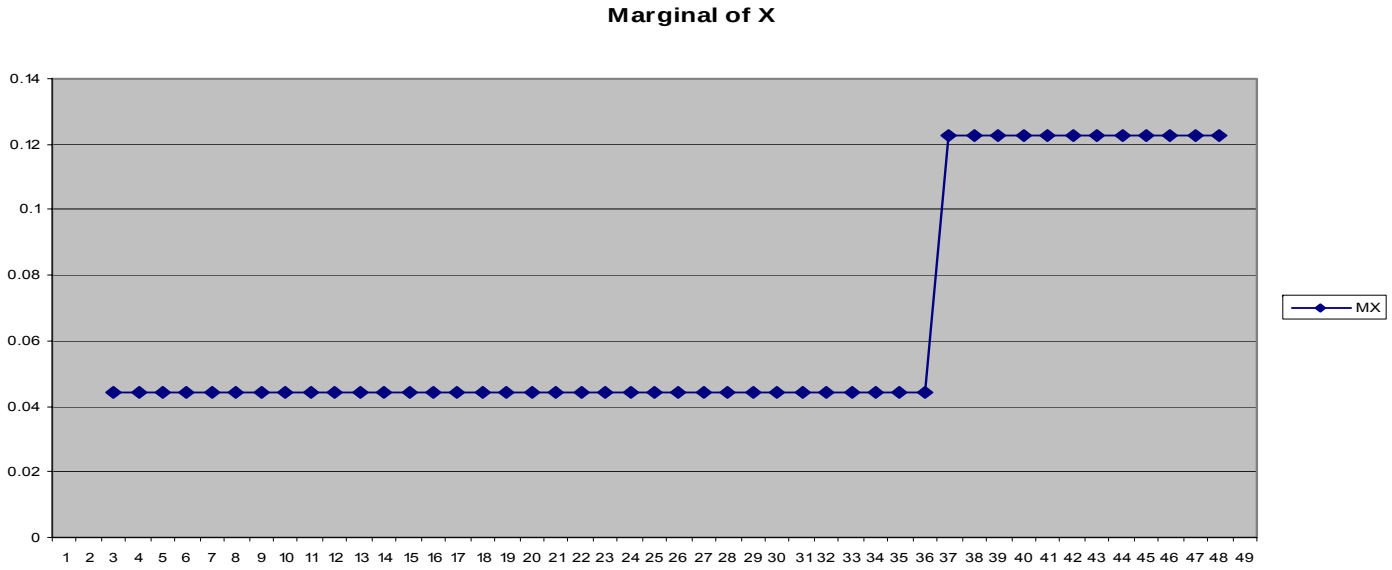
Year			
	1990	1979	1965
Export	0.018	0.274	0.3751
GDP	0.089	0.498	0.769

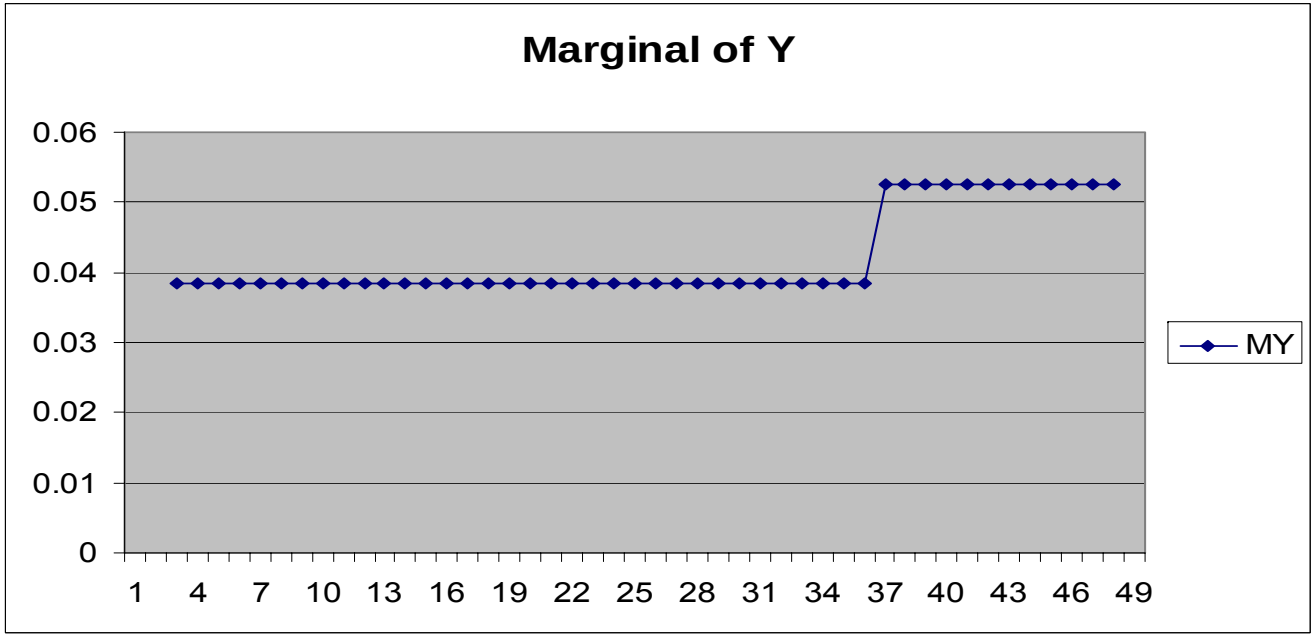
Distributions	Year	Results
GDP Conditional	1955-2002	DY=0.483DY(−1)+ 0.2039DX(−1)
		(0.115) (0.056)
	1955-1989	DY=0.410DY(−1)+0.2189DX(−1)
		(0.0778) (0.1426)
	1990-2002	DY=0.7083DY(−1)+0.0946 DX(−1)
		(0.2238) (0.0926)
Export Conditional	1955-2002	DX=1.0564 DY(−1)+0.0.1533 DX(−1)
		(0.2943) (0.1437)
	1955-1989	DX=0.700DY(−1) −0.0001DX(−1)
		(0.318657) (0.174)
	1990-2002	DX=2.4148DY(−1)−0.1126DX(−1)
		(0.7665) (0.3172)

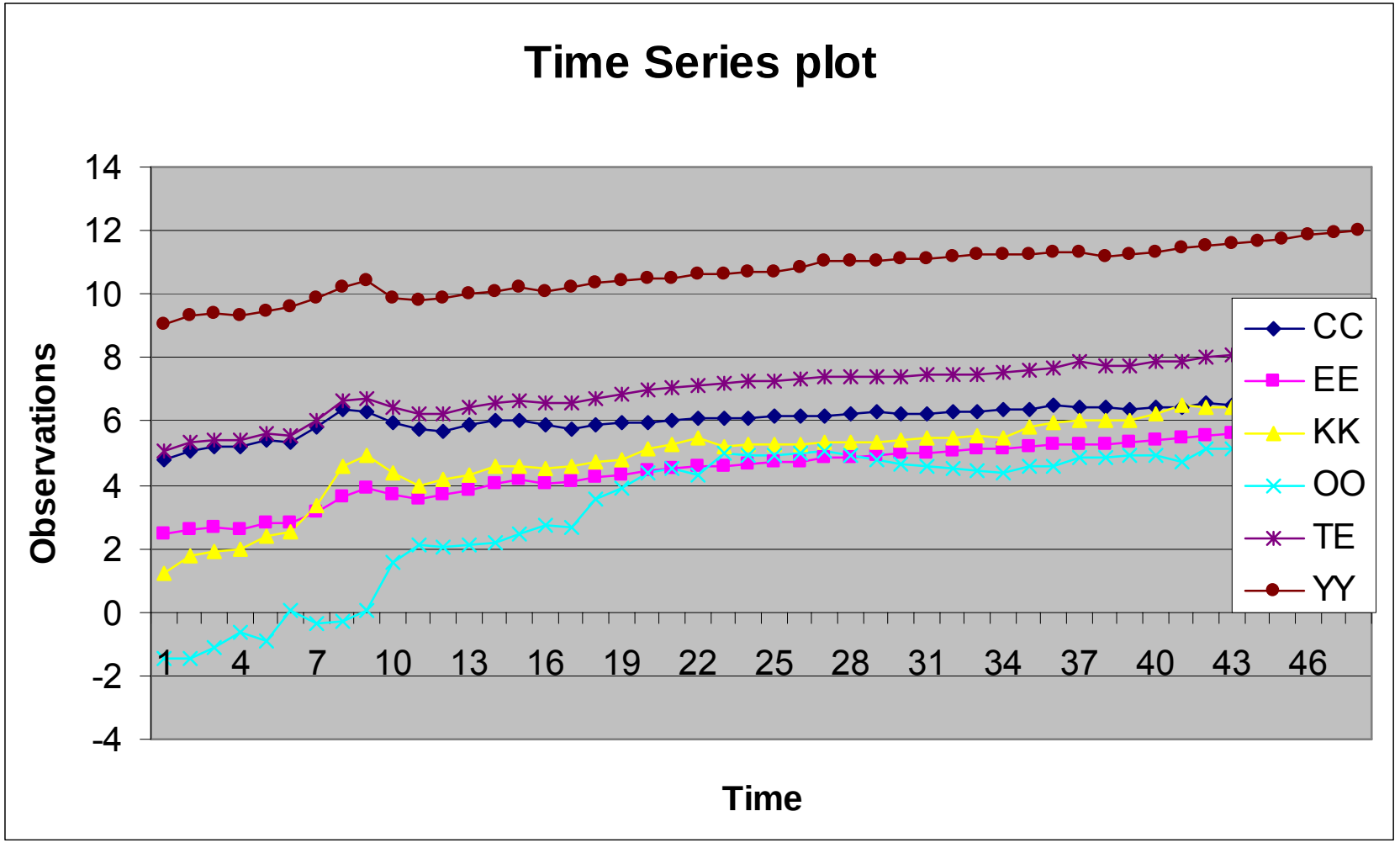
Export Marginal	1955-2002	DX=0.0611+0.058DX (−1)
		(0.016) (0.1533)
	1955-1989	DX=0.0480-0.151DX (−1)
		(0.016) (0.1786)
	1990-2002	DX=0.1281-0.0579DX (−1)
		(0.0426) (0.3147)
GDP Marginal	1955-2002	DY=0.0483−0.1476DY(−1)
		(0.0076) (0.1488)
	1955-1989	DY=0.0472-0.2481DY(-1)
		(0.0086) (0.174)
	1990-2002	DY=0.0492-0.0578DY(-1)
		(0.0.017) (0.3038)


0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
55
60
65
70
75
80
85
90
95
00
X1F
X2F
X3F


0.07
0.06
0.05
0.04
0.03
0.02
0.01
55
60
65
70
75
80
85
90
95
00
Y1F
Y2F
Y3F


Marginal of X
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49
MX

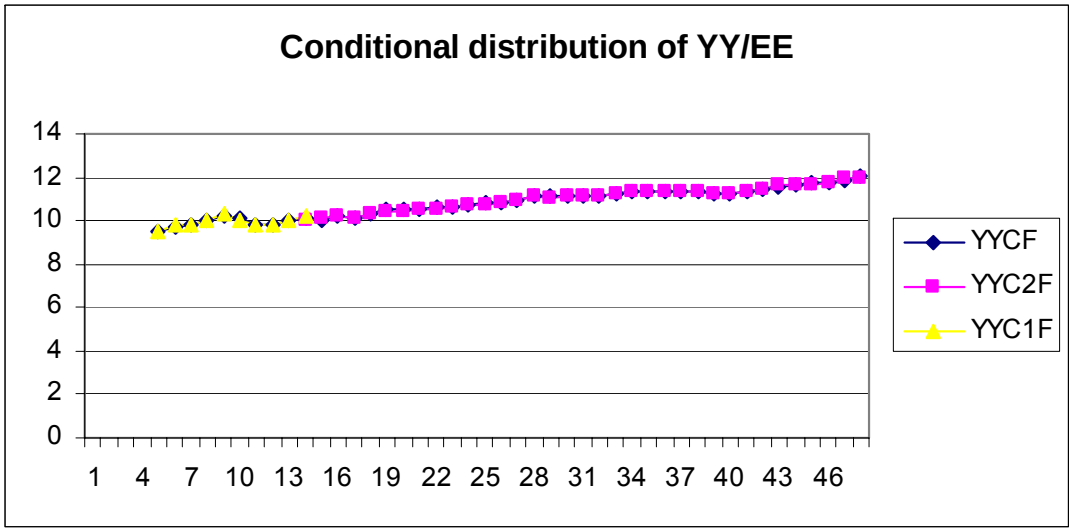

Marginal of Y
0.06
0.05
0.04
0.03
0.02
0.01
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49
MY

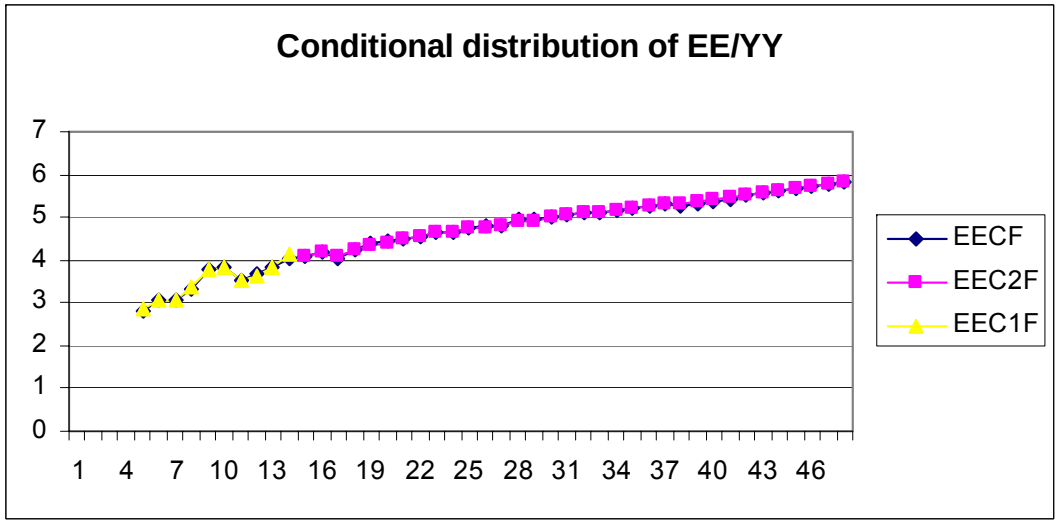

Time Series plot
Observations
14
12
10
8
6
4
2
0
-2
-4
1
4
7
10
13
16
19
22
25
28
31
34
37
40
43
46
Time
CC
EE
KK
OO
TE
YY

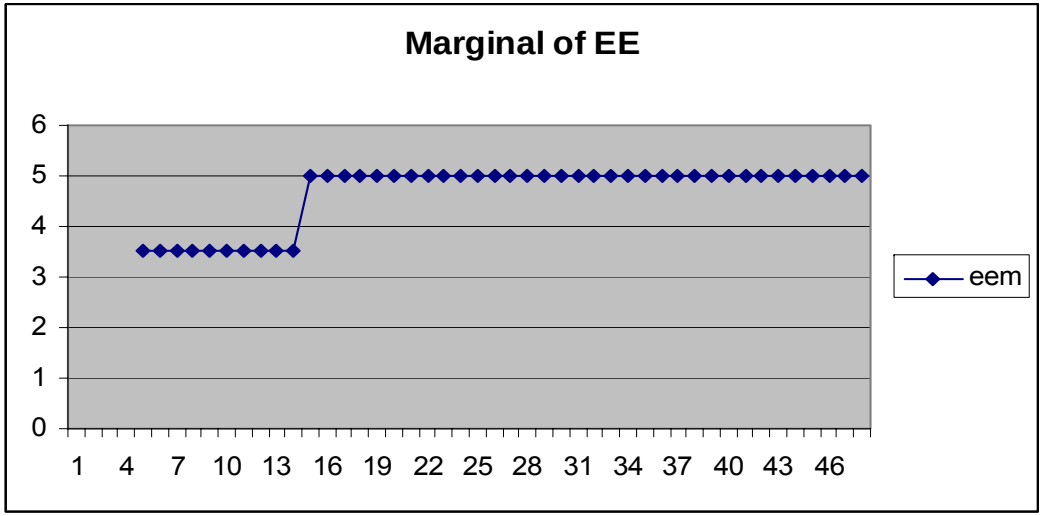
Results of Causality from 1952-1999(at lag length 4)		
Source	1952-1999(lag4)	1965-1999(Lag4)
CC ⇒ YY	0.000016	0.305372
YY⇒CC	0.222	0.324841
KK⇒YY	0.0019	0.730152
YY⇒KK	0.3525	0.957640
TE⇒YY	0.0032	0.919491
YY⇒TE	0.2461	0.244372
OO⇒YY	0.8968	0.861737
YY⇒OO	0.6975	0.153969
EE⇒YY	0.0018	0.899251
YY⇒EE	0.1050	0.541047

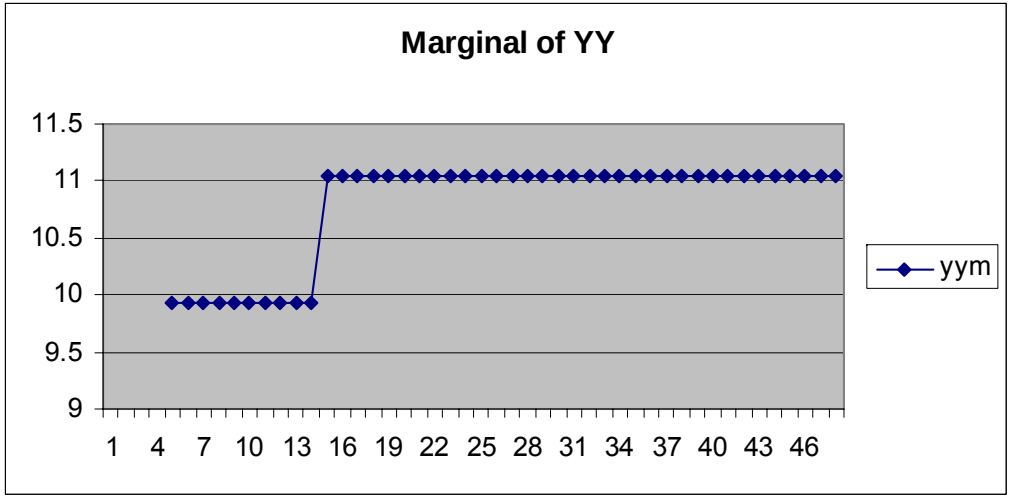
Pair wise Granger Causality Tests				
	Sample Range	1952 1999	1965-1999	1952-1990
Source	Lags:	Probability	Probability	Probability
CC⇒YY	4	0.000	0.018	0.000
CC⇒YY	3	0.00061	0.50	0.0031
CC⇒YY	2	0.009	0.53	0.018
YY⇒CC	4	0.009	0.068	0.056
YY⇒CC	3	0.013	0.007	0.071
YY⇒CC	2	0.01	0.002	0.06
EE⇒YY	4	0.002	0.27	0.002
EE⇒YY	3	0.07	0.33	0.14
EE⇒YY	2	0.46	0.33	0.61
YY⇒EE	4	0.182	.56	0.306
YY⇒EE	3	0.212	0.72	0.34
YY⇒EE	2	0.14	0.89	0.21

KK⇒YY	4	0.003	0.53	0.003
KK⇒YY	3	0.0015	0.33	0.003
KK⇒YY	2	0.39	0.49	0.46
YY⇒KK	4	0.196	0.74	0.30
YY⇒KK	3	0.3	0.77	0.3
YY⇒KK	2	0.24	0.75	0.34
TE⇒YY	4	0.004	0.017	0.003
TE⇒YY	3	0.024	0.183	0.03
TE⇒YY	2	0.62	0.29	0.56
YY⇒TE	4	0.321	0.25	0.34
YY⇒TE	3	0.22	0.14	0.31
YY⇒TE	2	0.19	0.31	0.25
OO⇒YY	4	0.95	0.93	0.624
OO⇒YY	3	0.634	0.96	0.18
OO⇒YY	2	0.33	0.72	0.08
YY⇒OO	4	0.80	0.24	0.703
YY⇒OO	3	0.685	0.103	0.47
YY⇒OO	2	0.82	0.27	0.52


Conditional distribution of YY/EE
14
12
10
8
6
4
2
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46
YYCF
YYC2F
YYC1F


Conditional distribution of EE/YY
7
6
5
4
3
2
1
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46
EECF
EEC2F
EEC1F


Marginal of EE
6
5
4
3
2
1
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46
eem


Marginal of YY
11.5
11
10.5
10
9.5
9
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46
yym

Results for Pakistan 1970-2003

Causal Direction	Lag Length	P-Value	AIC	SC
X⇒Y	(2,2)	0.0174	-5.1070	-4.878
X⇒Y	(2,3)	0.4955	-5.1359	-4.8584
X⇒Y	(3,3)	0.5607	-5.077	-4.7537

Results for Pakistan 1974-2003

Causal Direction	Lag Length	P-Value	AIC	SC
X⇒Y	(2,2)	0.7524	-5.096	-4.8584
X⇒Y	(2,3)	0.7720	-5.079	-4.7920
X⇒Y	(3,3)	0.850	-5.011	-4.6752
X⇒Y	(3,4)	0.574	-5.2142	-4.8271
X⇒Y	(3,3)	0.850	-5.011	-4.6752

Table 6
Results for Pakistan in Multivariate Framework

Lag2	GDP	Export	TOT
GDP	_____	0.0266	0.8197
Export	0.3586	_____	0.1996
TOT	0.7071	0.0458	_____
Lag3			
GDP	_____	0.7226	0.6455
Export	0.6390	_____	0.3592
TOT	0.4532	0.0039	_____

Results for India 1955-2002

Causal Direction	Lag Length	P-Value	AIC	SC
X⇒Y	(2,2)	0.0547	-4.23	-4.03
X⇒Y	(2,3)	0.1346	-4.24	-3.99
X⇒Y	(2,4)	0.3361	-4.17	-3.88
X⇒Y	(3,3)	0.097	-4.2662	-3.98
X⇒Y	(1,3)	0.1049	-4.2661	-4.06
X⇒Y	(1,2)	0.038	-4.2389	-4.079

Table 8

Results for India 1955-1998

Causal Direction	Lag Length	PValue	AIC	SC
X⇒Y	(2,2)	0.0317	-4.187	-3.9806
X⇒Y	(2,3)	0.0653	-4.201	-3.9504
X⇒Y	(3,3)	0.0437	-4.2390	-3.9465

India in Multivariate Framework

Results for the data range 1955-2002

Source	Lag	GDP	X	TOT
GDP	2	_____	0.201	0.621
	3	_____	0.118	0.868
	4	_____	0.306	0.935
X	2	0.293	_____	0.049
	3	0.084	_____	0.114
	4	0.367	_____	0.200
TOT	2	0.346	0.360	_____
	3	0.402	0.306	_____
	4	0.254	0.218	_____

Results for Sri Lanka 1965-1997

Causal Direction	Lag Length	P-Value	AIC	SC
X⇒Y	(2,2)	0.6152	-2.2268	-1.996
X⇒Y	(2,3)	0.6308	-2.169	-1.889
X⇒Y	(2,4)	0.8112	-2.07	-1.741
X⇒Y	(3,3)	0.6278	-2.1059	-1.779
X⇒Y	(1,3)	0.5419	-2.2033	-1.97
X⇒Y	(1,2)	0.307	-2.2695	-2.0845

From Y to X 1965-1997

Causal Direction	Lag Length	P-Value	AIC	SC
Y⇒X	(2,2)	0.889	-0.9224	-0.6911
X⇒Y	(1,2)	0.687	-0.9820	-0.7970
X⇒Y	(3,3)	0.992	-0.7508	-0.4239
X⇒Y	(2,3)	0.9959	-0.8169	-0.5367

Table 11

Results for Sri Lanka 1970-1997

Causal Direction	Lag Length	P-Value	AIC	SC
X⇒Y	(2,2)	0.6432	-2.183	-1.945
X⇒Y	(2,3)	0.6426	-2.1344	-1.85
X⇒Y	(2,4)	0.8347	-2.063	-1.73
X⇒Y	(3,3)	0.6278	-2.063	-1.73
X⇒Y	(1,3)	0.5513	-2.1632	-1.925
X⇒Y	(1,2)	0.2889	-2.220	-2.030

- We conclude as follows
- Granger Causality does not lead us to any single direction of causation and most of the times the basic axiom proposed by Granger that direction of causation should remain same is violated.
- The proposed method of structural causality is superior to that of the former in the sense it is not only based on statistical grounds rather careful functioning of the underlying mechanism is also required. This idea is in close agreement to that of Hoover and Hendry. Nevertheless, there is still a long way to go on this issue of causality in Economics.

- Thank You