



*OECD SHORT-TERM ECONOMIC  
STATISTICS EXPERT GROUP (STESEG)  
MEETING, PARIS, 27-28 JUNE 2005*

THE OECD PROJECT ON REVISIONS ANALYSIS  
FIRST ELEMENTS FOR DISCUSSION

Tommaso Di Fonzo  
OECD and University of Padova



## Starting point of the project

OECD-ONS workshop on “Assessing and improving statistical quality. Revision analysis for the national accounts”, 7-8 October 2004

*“The OECD also agreed to explore the possibility to expand and maintain the revisions database that was set up for purposes discussed in the workshop.*

*An international dataset and the resulting summary measures of revisions should help to put countries’ individual revision analyses into an international context.”*



## Aims of the paper

- First elements of the ongoing project
- Remarks on strategy and tools for revisions analysis
- Building, maintaining and using revisions databases

## Statement of the problem: Revisions to economic data

- Part of wider consideration of statistical quality
- Data production process based on multiple, complex source data/procedures
- Usual statistical measures of sampling biases, variances and measurement errors properties cannot be used

# Statement of the problem: Revisions to economic data

- The main aim of revisions is to improve the quality of preliminary figures published
- 5 main sources of revision, loosely defined as:
  1. Incorporation of later and improved data
  2. replacement of judgemental estimates
  3. definitional changes
  4. methodological changes
  5. seasonal adjustment updating

# The accuracy and reliability dimension of data quality

- At least two needs for both producers and users of economic data:

*increase awareness on approaches to assess quality of statistical output coming from official economic datasets*

*stimulate thinking about additional approaches that might be used to assess accuracy and reliability*

# The reliability dimension of data quality

- Reliability is related to the closeness of the initial estimated value(s) to the subsequent estimates
- Assessing reliability involves comparing estimates over time
- Assessing reliability refers to revisions
- Timeliness: first preliminary estimates are involved (trade-off)

## The reliability dimension of data quality (IMF, Carson & Laliberté, 2002)

- 3.5 (Elements) Revisions studies – Revisions, as a gauge of reliability, are tracked and mined for the information they may provide
- 3.5.1 (Indicators) Studies and analysis of revisions are carried out routinely and used to inform statistical processes
- Best practices: ABS, ONS, BEA, Statistics Canada,...

# OECD Main Economic Indicators

- Published in paper format since the early 1960's
- Long time series for key short-term economic statistics for most OECD countries
  - Quarterly GDP and expenditure components
  - Index of industrial production & Composite Leading Indicators
  - Retail Trade, Consumer and Producer Prices
  - Wages, employment and unemployment
  - Interest rates, exchange rates, monetary aggregates
- Electronic monthly 'snapshot' databases available back at least until 1995

# MEI as a revisions database source

- MEI monthly snapshots in principle provide historical data on subsequent releases on a wide range of key variables
  - Provides a coherent set of variables representing the ‘information set’ available at successive monthly intervals

# OECD interest in a revisions database

- Analysis – are revisions in general becoming smaller? Are they random (i.e. centred around 0)?
- Support internationally coordinated research work to improve the quality of early estimation methods
- Encourage transparency in the statistical process
- Assess real-time performance of the OECD CLIs

# The OECD project

## Real time database & revisions analysis of GDP for OECD countries

- Initial OECD work (Ahmad, Bournot & Koechlin, 2004) was for G7 countries, period 1996 – 2000
- Current project
  - Expand coverage to all OECD countries
  - Expand time period covered
  - Devise a scheme which can support ongoing production of real time database & summary revision analysis indicators
  - Quarterly GDP and its components

# Initial OECD work

## Improvements in Timeliness

**Table 1: First published estimates in MEI of quarterly GDP data in G-7 countries**

	Q1	Q2	Q3	Q4	Year
<b>Canada</b>	<b>June</b> except in 1999 (July)	<b>September</b>	<b>December</b> except in 1997 (January)	<b>March t+1</b>	<b>March t+1</b>
<b>France</b>	<b>June</b> except in 1997 (July) and in 1999 (August)	<b>September</b> except in 2000 and 2001 (October)	<b>December</b> except in 1997 (January)	<b>March t+1</b>	<b>March t+1</b>
<b>Germany</b>	<b>June</b> from 2001 (July in 96, 97, 99 and 2000)	<b>October</b> until 1999 <b>September</b> since 2000	<b>December</b> since 1998 (January t+1 before)	<b>March t+1</b> except in 96 and 98 (April t+1)	<b>March t+1</b> except in 96 and 98 (April t+1)
<b>Italy</b>	<b>June</b> since 2001 (May, July or August for years prior to 2001)	From August to November but mainly <b>October</b>	<b>December</b> since 2000 (January or February t+1 before)	From March t+1 to June t+1 but <b>March t+1</b> from 2001	<b>March or April t+1</b>
<b>Japan</b>	<b>July</b> until 1999 <b>June</b> since 2000	<b>October</b> until 2001 <b>September</b> since 2002	January or December but mainly <b>December</b>	<b>April</b> until 2001 <b>March</b> since 2002	<b>April</b> until 2001 <b>March</b> since 2003
<b>United Kingdom</b>	<b>June</b> until 2000 <b>May</b> since 2001	<b>September</b> until 2000 <b>August</b> since 2001	From November to January t+1 but mainly <b>December</b>	<b>February or March t+1</b>	<b>February or March t+1</b>
<b>United States</b>	<b>May</b>	August or September but mainly <b>August</b>	<b>November</b> except in 1999 (December)	<b>February t+1</b> since 1998 (March and January before)	<b>February t+1</b> since 1997

## Initial OECD work

### Caveats for international comparisons

- $P$  are not necessarily flash estimates
- Timeliness  $P$  differs between countries
- No attempt is made to isolate the different types of revisions, (including errors)
- Differences may exist between estimates and national estimates
- No investigation into influence of economic cycle
- Period covers SNA93 introduction and other changes, fixed base, chain-linking



## Initial OECD work

### Main Indicators used

- Mean revision

$$\bar{R} = \frac{1}{n} \sum_{t=1}^n (L_t - P_t) = \frac{1}{n} \sum_{t=1}^n R_t$$

- Mean absolute revision

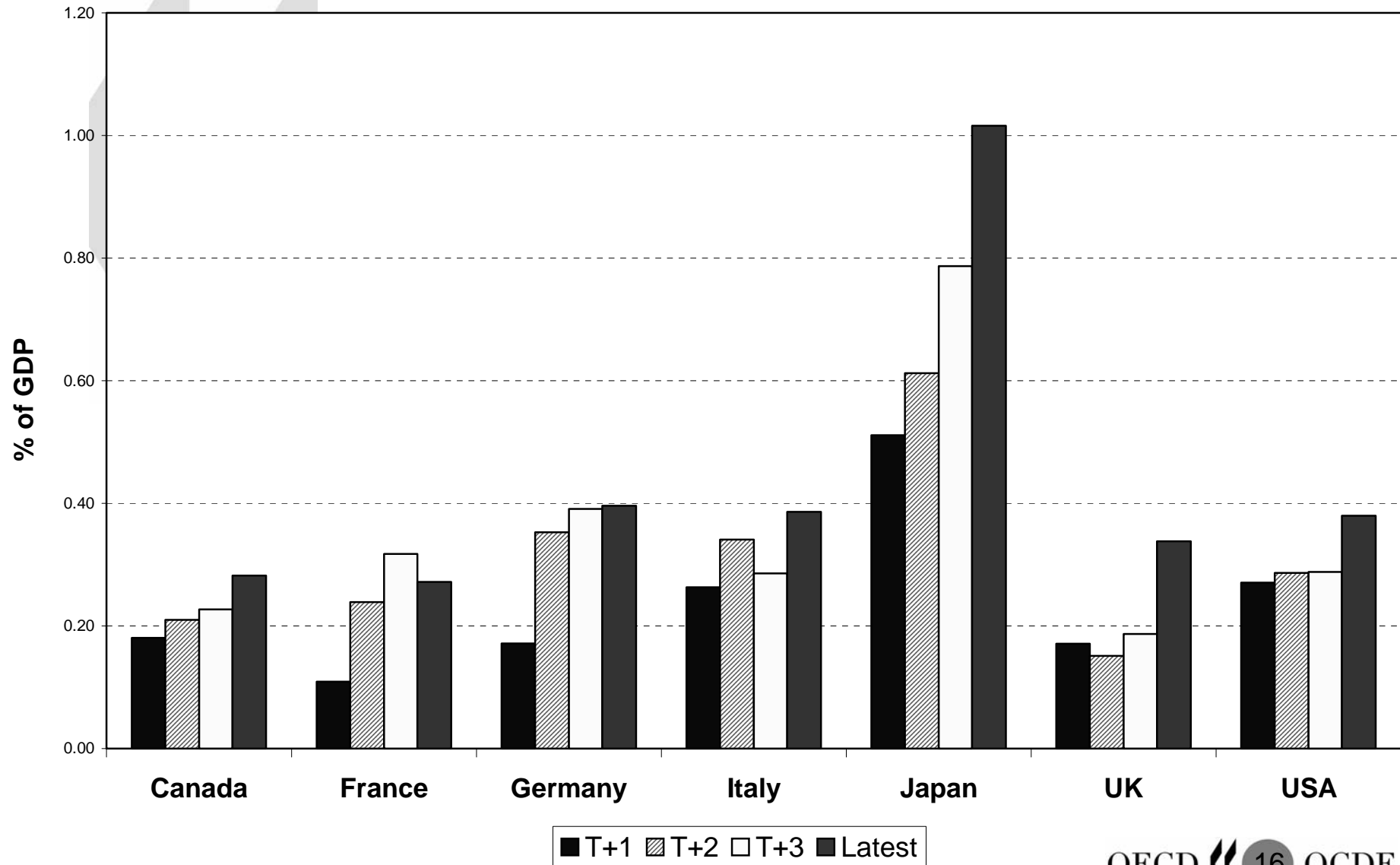
$$MAR = \frac{1}{n} \sum_{t=1}^n |L_t - P_t| = \frac{1}{n} \sum_{t=1}^n |R_t|$$

- Number of times  $L > P$

- Secondary analysis: mean absolute revisions between subsequent estimates

# Initial OECD work

## Mean Absolute Revision 1996-2000 - GDP



# The ongoing project

## Data availability for GDP (volume series, sa) in the OECD MEI real-time database (preliminary)

MEI editions	Vintages	Countries
First: May 95    Last: May 05	First: 95q1    Last: 05q1	1: USA
First: May 95    Last: May 05	First: 94q4    Last: 04q4	15: AUS, AUT, CAN, CHE, DEU, DNK, ESP, FIN, FRA, GBR, ITA, JPN, NLD, NOR, NZL
First: Jan-97    Last: May 05	First: 95q4    Last: 04q4	1: PRT
First: Mar-97    Last: May-05	First: 96q3    Last: 05q1	1: BEL
First: May-97    Last: May-05	First: 96q4    Last: 04q4	1: KOR
First: Jul-99    Last: May-05	First: 99q1    Last: 04q4	1: CZE
First: Apr-04    Last: May-05	First: 03q4    Last: 04q4	1: GRC
First: Nov-04    Last: May-05	First: 04q2    Last: 04q4	1: IRL
First: May 95    Last: Sep 00	First: 96q3    Last: 00q2	2: SWE, TUR
First: May 95    Last: Sep 00	First: 96q3    Last: 00q1	1: MEX
---	---	5: HUN, ISL, LUX, POL, SVK

## The ongoing project

Assessing the significance of the mean revisions

- Standard t-test
- Modified t-test using AR(1)-based corrections (ONS)
- Modified t-statistic using HAC (Heteroskedasticity, Autocorrelation Consistent) estimate of the variance of the mean revision (technical details in the paper)
- HAC procedure has been implemented in the OECD prototype spreadsheet for revisions analysis

## The ongoing project

### Mean squared revision and its decomposition

- Summary measure based on a quadratic loss function

$$MSR = \frac{1}{n} \sum_{t=1}^n (L_t - P_t)^2 = \frac{1}{n} \sum_{t=1}^n R_t^2$$

$$MSR = \bar{R}^2 + (S_P - \rho S_L)^2 + (1 - \rho^2) S_L^2$$

## The ongoing project

Mean squared revision and a useful decomposition

$$1 = UM + UR + UD$$

- *UM*: proportion of *MSR* due to systematic differences between the preliminary and the latest estimates
- *UR*: proportion of *MSR* due to the slope coefficient in the regression  $L_t = \alpha + \beta P_t + u_t$
- *UD*: disturbance proportion of *MSR*, part of the revision which is not explained by the mean or slope error
- ‘Good’ preliminary estimates will have low values of *UM* and *UR* and a high value of *UD*

# Revisions analysis spreadsheet

## From levels ...

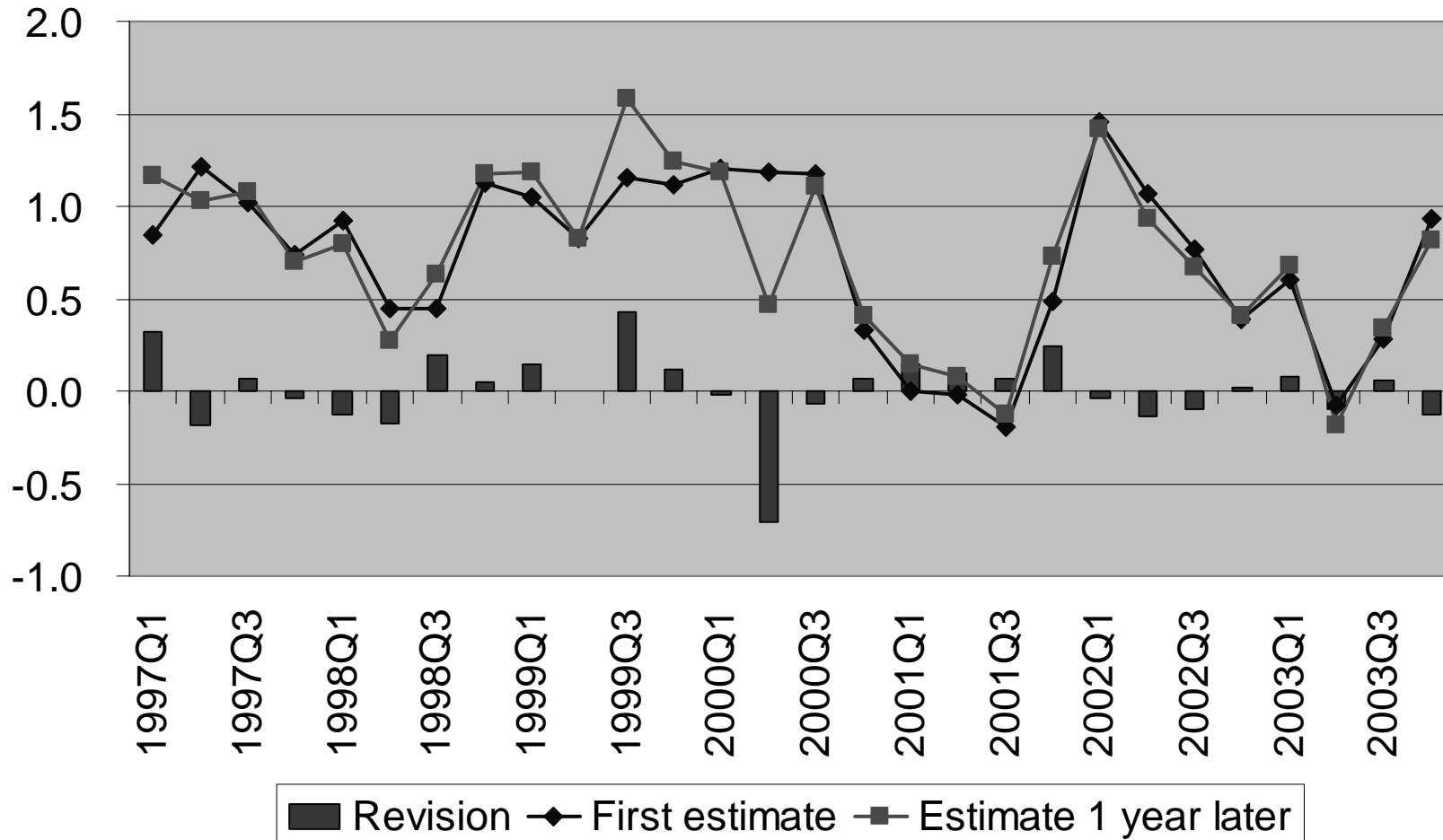
	1996Q4	1997Q1	1997Q2	1997Q3	1997Q4	1998Q1	1998Q2	1998Q3	1998Q4
<b>First estimate</b>	624.2	629.5	637.6	803.8	809.9	816.0	819.0	830.7	840.4
<b>1 year later</b>	777.4	786.5	794.6	812.2	817.9	824.4	833.7	839.0	848.9
<b>2 years later</b>	783.5	793.4	807.8	819.3	825.8	834.1	836.4	842.4	855.1
<b>3 years later</b>	790.8	800.0	810.0	821.5	825.8	834.1	911.9	922.2	936.7
Jun-97	<b>624.2</b>	<b>629.5</b>							
Jul-97	624.2	629.5							
Aug-97	624.2	629.5							
Sep-97	624.2	<b>629.9</b>	<b>637.6</b>						
Oct-97	624.2	629.9	637.6						
Nov-97	624.2	629.9	637.6						
Dec-97	624.2	629.9	637.6						
Jan-98	<b>777.4</b>	<b>785.3</b>	<b>795.7</b>	<b>803.8</b>					
Feb-98	777.4	785.3	795.7	803.8					
Mar-98	777.4	<b>786.4</b>	<b>796.3</b>	<b>804.0</b>	<b>809.9</b>				
Apr-98	777.4	786.4	796.3	804.0	809.9				
May-98	777.4	786.4	796.3	804.0	809.9				
Jun-98	777.4	<b>786.5</b>	<b>794.6</b>	<b>803.0</b>	<b>808.6</b>	<b>816.0</b>			
Jul-98	777.4	786.5	794.6	803.0	808.6	816.0			
Aug-98	777.4	786.5	794.6	803.0	808.6	816.0			
Sep-98	777.4	786.5	794.6	803.0	808.6	<b>815.4</b>	<b>819.0</b>		
Oct-98	777.4	786.5	794.6	803.0	808.6	815.4	819.0		
Nov-98	777.4	786.5	794.6	803.0	808.6	815.4	819.0		
Dec-98	<b>783.5</b>	<b>793.4</b>	<b>803.5</b>	<b>812.2</b>	<b>817.9</b>	<b>824.0</b>	<b>827.0</b>	<b>830.7</b>	
Jan-99	783.5	793.4	803.5	812.2	817.9	824.0	827.0	830.7	
Feb-99	783.5	793.4	803.5	812.2	817.9	824.0	827.0	830.7	
Mar-99	783.5	793.4	803.5	812.2	817.9	<b>824.4</b>	<b>827.5</b>	<b>831.1</b>	<b>840.4</b>

## ... to QoQ rates of change ...




	1997Q1	1997Q2	1997Q3	1997Q4	1998Q1	1998Q2	1998Q3	1998Q4
<b>First estimate</b>	0.8	1.2	1.0	0.7	0.9	0.4	0.4	1.1
<b>1 year later</b>	1.2	1.0	1.1	0.7	0.8	0.3	0.6	1.2
<b>2 years later</b>	1.3	1.1	1.4	0.8	0.7	0.3	0.7	1.5
<b>3 years later</b>	1.2	1.2	1.4	0.9	0.7	0.4	1.1	1.6
Jun-97	<b>0.8</b>							
Jul-97	0.8							
Aug-97	0.8							
Sep-97	<b>0.9</b>	<b>1.2</b>						
Oct-97	0.9	1.2						
Nov-97	0.9	1.2						
Dec-97	0.9	1.2						
Jan-98	<b>1.0</b>	<b>1.3</b>	<b>1.0</b>					
Feb-98	1.0	1.3	1.0					
Mar-98	<b>1.1</b>	<b>1.3</b>	<b>1.0</b>	<b>0.7</b>				
Apr-98	1.1	1.3	1.0	0.7				
May-98	1.1	1.3	1.0	0.7				
Jun-98	<b>1.2</b>	<b>1.0</b>	<b>1.1</b>	<b>0.7</b>	<b>0.9</b>			
Jul-98	1.2	1.0	1.1	0.7	0.9			
Aug-98	1.2	1.0	1.1	0.7	0.9			
Sep-98	1.2	1.0	1.1	0.7	<b>0.8</b>	<b>0.4</b>		
Oct-98	1.2	1.0	1.1	0.7	0.8	0.4		
Nov-98	1.2	1.0	1.1	0.7	0.8	0.4		
Dec-98	<b>1.3</b>	<b>1.3</b>	<b>1.1</b>	<b>0.7</b>	<b>0.8</b>	<b>0.4</b>	<b>0.4</b>	
Jan-99	1.3	1.3	1.1	0.7	0.8	0.4	0.4	
Feb-99	1.3	1.3	1.1	0.7	0.8	0.4	0.4	
Mar-99	1.3	1.3	1.1	0.7	<b>0.8</b>	<b>0.4</b>	<b>0.4</b>	<b>1.1</b>

... to charts ...

Chart to compare first to 1 year later estimate



## ... to revisions analysis tables

Basic indicators	Whole sample	Last 20 quarters
number of observations (n)	28	20
mean absolute revision	0.1396	0.1382
mean revision	0.0099	0.0086
std. dev. of mean revision	0.0396	0.0522
t-statistic	0.2502	0.1647
t-crit (n-1 degrees of freedom)	2.0518	2.0930
significant mean revision or not?	NO	NO
Other indicators		
Correlation	0.901	0.908
Min revision	-0.7	-0.7
Max revision	0.4	0.4
Range	1.1	1.1
% Later > Early	57.1	60.0
% Sign(Later) = Sign(Early)	92.9	  

## Extension of the work to other variables?

- Need to establish operational process for GDP and its components first
- Most likely next variable would be the Index of Industrial Production
  - But we seek this groups views on which variables should have priority

## Obstacles to extending the work

- Possible gaps in our electronic sources, which are only available to mid 1995
  - How far back should the revisions database go?
- Operational problems (e.g. series code changes)
- Methodological issues
  - Constancy of national series over time (e.g. various methodological changes causing breaks & revisions)
  - OECD procedures (linking, seasonal adjustment)
- Data management (produces big databases)

## Obstacles to extending the work

- Hence it's a very resource intensive process to establish a historical real time database for each variable
  - We are always looking for partners to denote resources in kind



## Elements for discussion

- Presentation of revisions, other (simple) descriptive measures?
- Possible use of an OECD revisions analysis database by member countries - what role for STESEG?
- Priority for adding other short-term economic statistics to the revisions database and how to use the results?
  - e.g. as an input to internationally coordinated methodological work?
  - To reconsider policies for communicating information on revisions to users?