

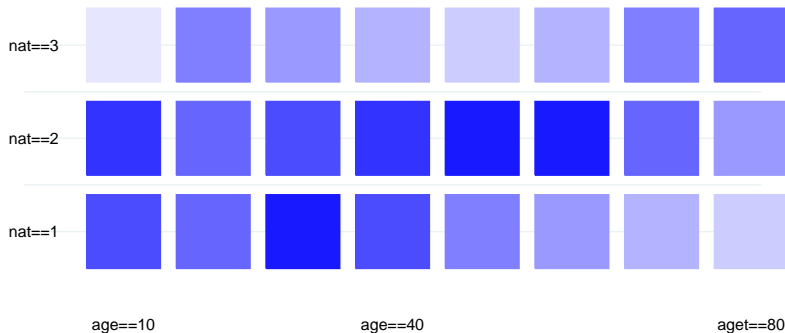
# **Stata graphics: Getting the graphs you want the hard way**

Philippe Van Kerm

CEPS/INSTEAD, Luxembourg  
`philippe.vankerm@ceps.lu`

Sixth Winter School on Inequality and Social Welfare Theory  
January 10–13 2011, Alba di Canazei, Italy

# Beware of visual distortions



# Beware of visual distortions

Cleveland and McGill: Graphical Perception

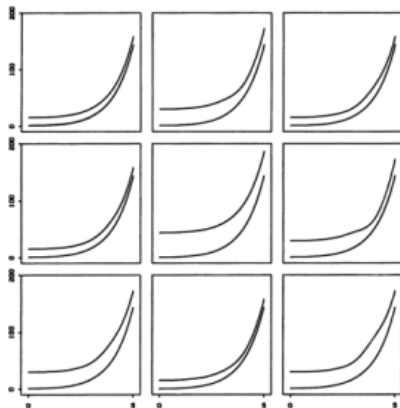


Figure 26. Curve-difference chart.

# Beware of visual distortions

Cleveland and McGill: Graphical Perception

549

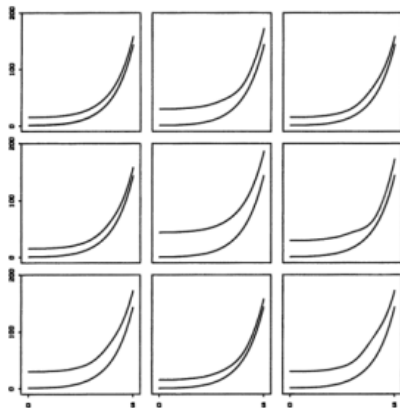


Figure 26. Curve-difference chart.

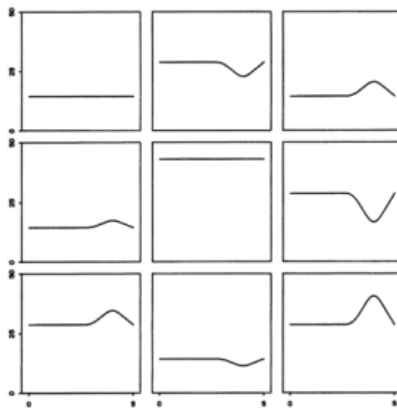


Figure 27. Curve differences.

# Outline

**Part I:**

A quick overview

**Part II:**

The anatomy of 'twoway' graphics

**Part III:**

Customizing graphs: Tips by examples

**Part IV:**

Q&A?

# Overview of Stata graphics

## Three main families

1. Estimation commands with graphical output (`lowess`, `kdensity`, some regression diagnostics, etc.)
2. Descriptive commands compute various summary statistics (means, tabulations, frequencies) and display results graphically (`graph pie`, `graph bar`, `graph box`, `hist`, ...)   
⇒ useful summaries over subsamples in particular
3. Simple (two-way) plotting commands do not calculate anything, and simply display data on two-way graphs (`graph twoway scatter`, `graph twoway line`, ... )

(My talk is about the third category of commands.)

# Overview of Stata graphics

## Three main families

1. Estimation commands with graphical output (`lowess`, `kdensity`, some regression diagnostics, etc.)
2. Descriptive commands compute various summary statistics (means, tabulations, frequencies) and display results graphically (`graph pie`, `graph bar`, `graph box`, `hist`, ...)   
⇒ useful summaries over subsamples in particular
3. Simple (two-way) plotting commands do not calculate anything, and simply display data on two-way graphs (`graph twoway scatter`, `graph twoway line`, ...)

(My talk is about the third category of commands.)

# Overview of Stata graphics

## Three main families

1. Estimation commands with graphical output (`lowess`, `kdensity`, some regression diagnostics, etc.)
2. Descriptive commands compute various summary statistics (means, tabulations, frequencies) and display results graphically (`graph pie`, `graph bar`, `graph box`, `hist`, ...)   
⇒ useful summaries over subsamples in particular
3. Simple (two-way) plotting commands do not calculate anything, and simply display data on two-way graphs (`graph twoway scatter`, `graph twoway line`, ...)

(My talk is about the third category of commands.)



# Overview of Stata graphics

## Three main families

1. Estimation commands with graphical output (`lowess`, `kdensity`, some regression diagnostics, etc.)
2. Descriptive commands compute various summary statistics (means, tabulations, frequencies) and display results graphically (`graph pie`, `graph bar`, `graph box`, `hist`, ...)   
⇒ useful summaries over subsamples in particular
3. Simple (two-way) plotting commands do not calculate anything, and simply display data on two-way graphs (`graph twoway scatter`, `graph twoway line`, ... )

(My talk is about the third category of commands.)

# Overview of Stata graphics

## Three main families

1. Estimation commands with graphical output (`lowess`, `kdensity`, some regression diagnostics, etc.)
2. Descriptive commands compute various summary statistics (means, tabulations, frequencies) and display results graphically (`graph pie`, `graph bar`, `graph box`, `hist`, ...)   
⇒ useful summaries over subsamples in particular
3. Simple (two-way) plotting commands do not calculate anything, and simply display data on two-way graphs (`graph twoway scatter`, `graph twoway line`, ... )

(My talk is about the third category of commands.)

# Overview of Stata graphics

## Documentation

- ▶ `help graph` is certainly the single most important command when creating almost any graph!
- ▶ Graphics manual has same information as online help (but with illustrations)
- ▶ Mitchell's book *A visual guide to Stata graphics* is a very useful companion with illustrations and code examples
- ▶ UCLA Academic Technology Services website (Google 'UCLA Stata library graph examples')

## The anatomy of `graph twoway`

```
[graph] twoway plot [if] [in] [, twoway options ]
```

where the syntax of `plot` is

```
[ ( ) plotype varlist ..., options [ ) ] [ ] ]
```

1. 'Overlaying' principle: plots can be overlaid on top of each other
2. plot-specific options vs. overall graph options

Plots are: scatter, line, spikes, paired-coordinates lines or arrows, functions, ... all of them can be overlaid!

Look of axis, titles, symbols, etc. is controlled by options

## Six examples

1. PDFs and CDFs
2. A decorated Lorenz curve
3. Dot plot by hand
4. Radar plot (DIY)
5. Thematic map
6. Heat/colour plot

(Now open Stata and run `go.do` – you will need to adjust the path)