Income profile of AFORE affiliates who contribute regularly

Tapen Sinha

Director, International Center for Pension Research, ITAM, Mexico

ING Chair Professor, ITAM, Mexico

Professor, School of Business, University of Nottingham, UK

tapen@itam.mx

tapen.sinha@nottingham.ac.uk

How did income increase during 1997-2005 among affiliates?

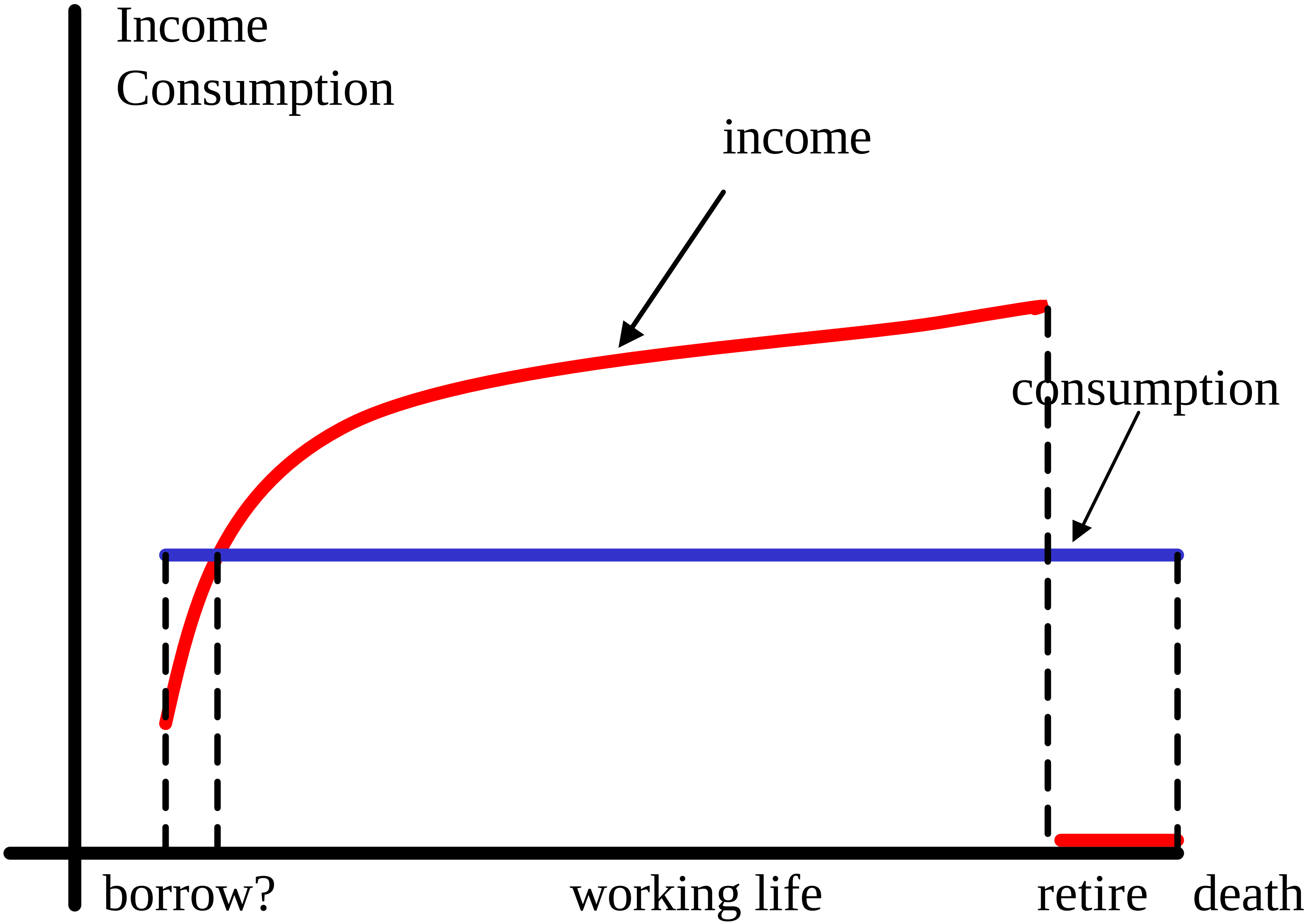
- Main question
- Why is it important?
- To know how much affiliates would have in their AFOREs during their lifetimes
- To have an idea of how income changes in a life cycle context in Mexico in the ***formal*** sector
- To disentangle the gender gap in income

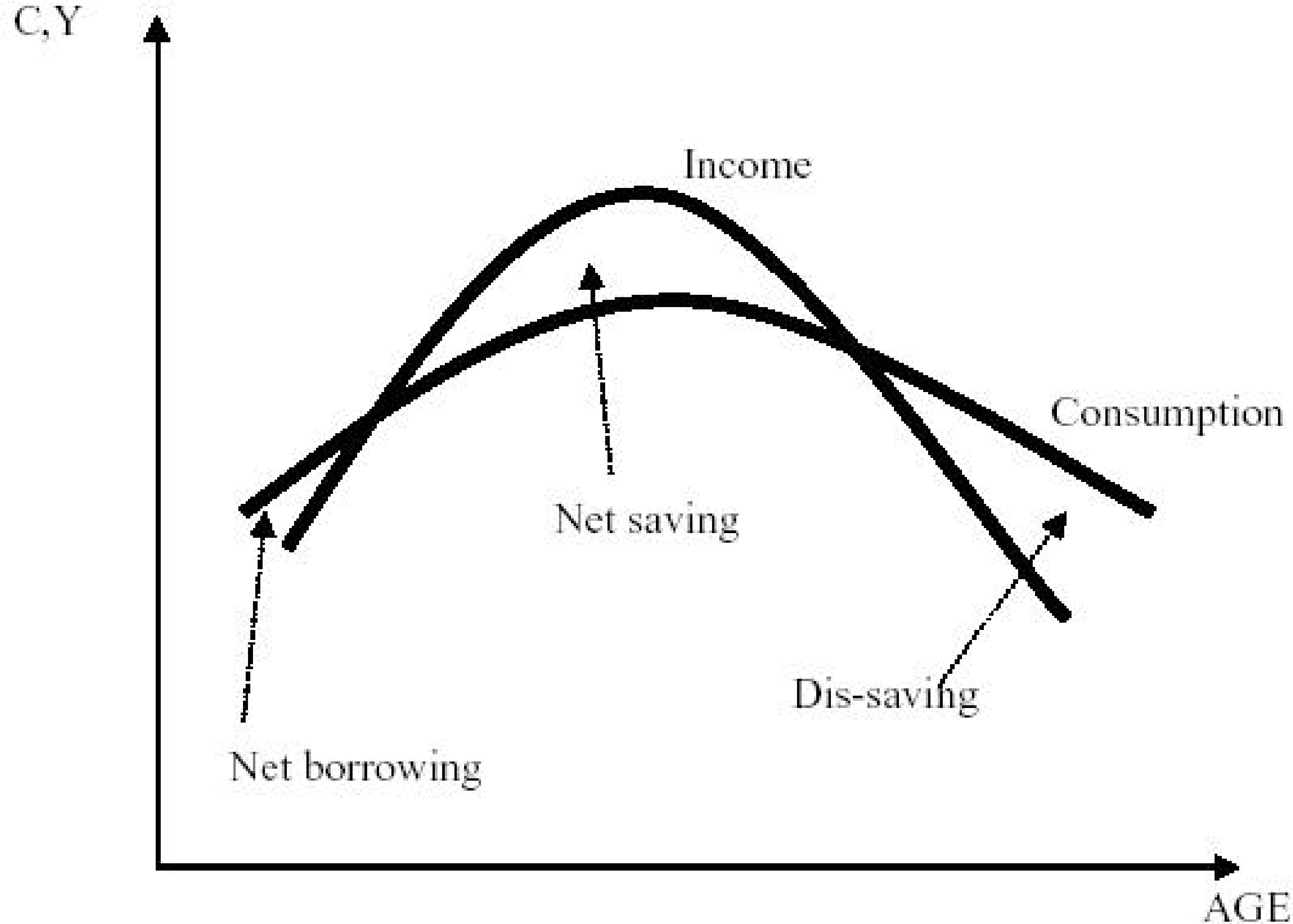
Limitations

- There are few observations in some of the cells – thus, they are not reliable (example: women in the highest quintile)
- We do not get a picture of ***all*** who are contributing – only the ones that are contributing regularly
- We only have data for 1997-2005 much less than ***lifetime*** data

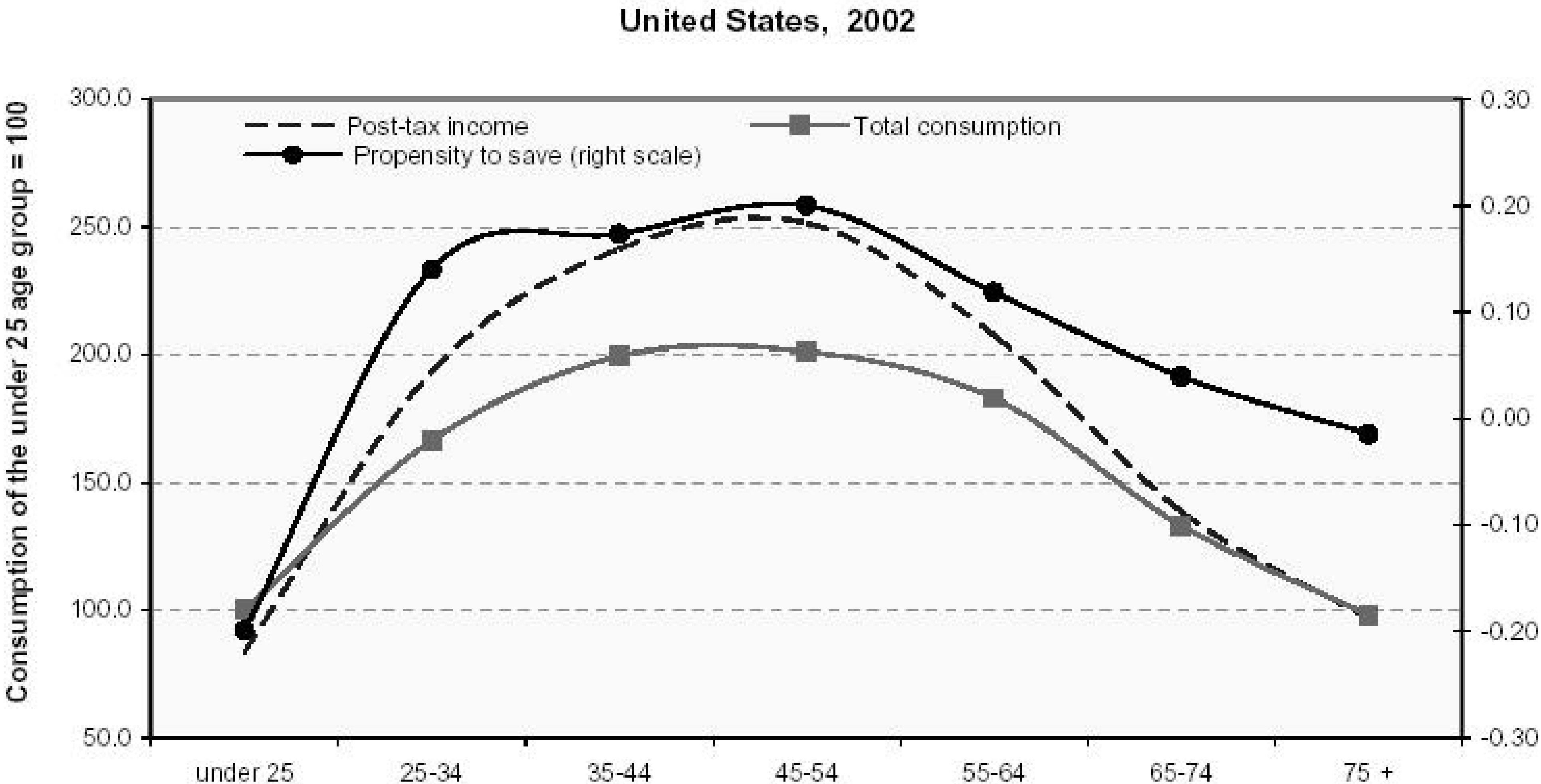
Original study

- With macroeconomic data
- Modigliani Bloomberg life cycle hypothesis
- Milton Friedman permanent income hypothesis
- Both posit that people smooth consumption over lifetime income
- Income has less volatility than consumption

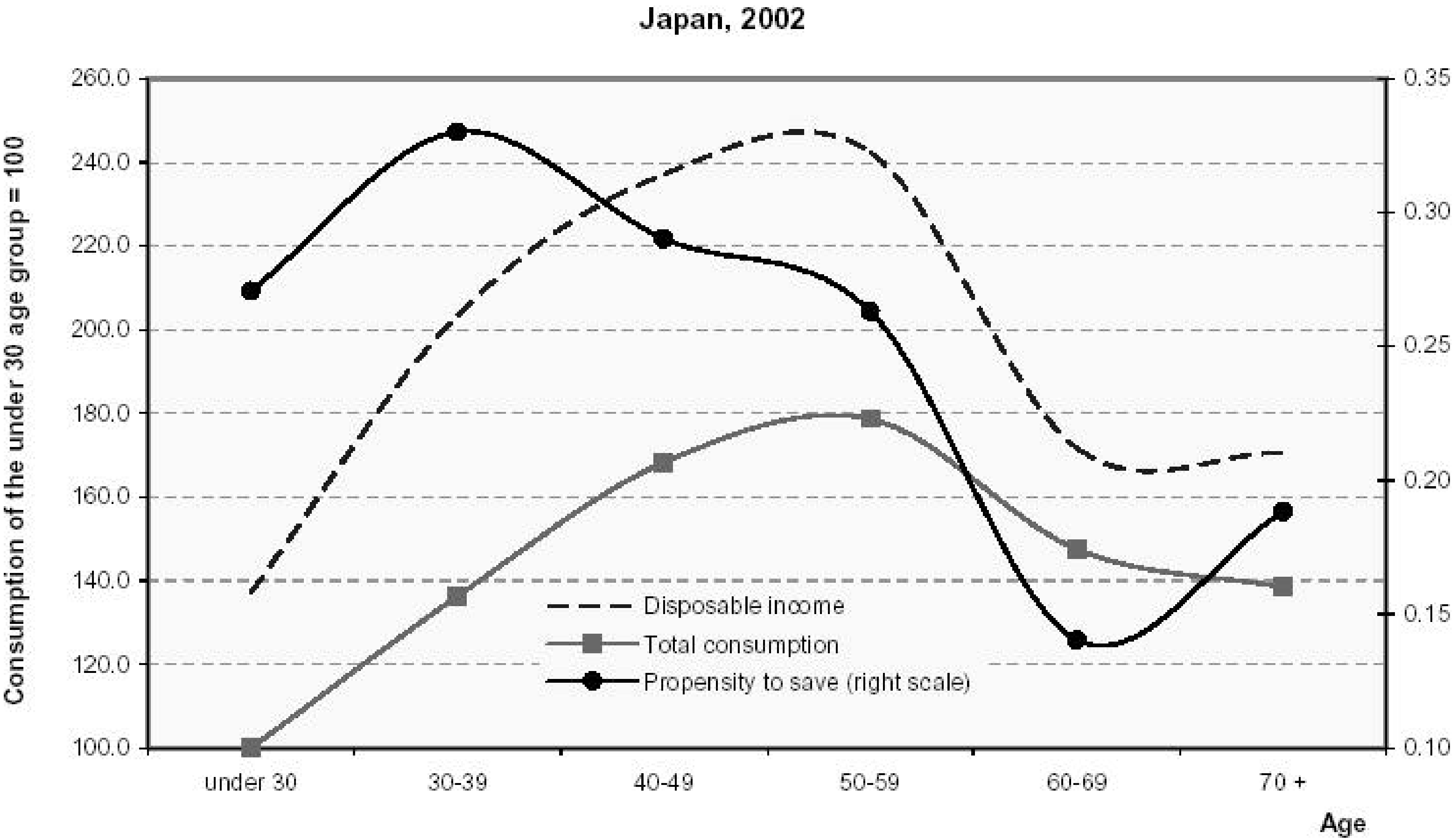

Income
Consumption
income
consumption
borrow?
working life
retire
death



“Typical” income consumption profile



Age consumption income profile in the US matches the typical one......



....but the Japanese data does not match the typical...

Data

- For men and women separate samples
- For each quintile, 1000 persons for each age group
- 20 or below, 21-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56-60, 61 and above
- Total ten categories
- Thus, (in theory) we have 50,000 observations for men and women – we have less numbers
- Each person has a salary figure for August 1997 to February 2005 every two months (46 obs)

Simple analysis

- We can collapse all the observations by examining the average salary over (almost) eight years and examine how ***average salary*** changes with all the persons put together taking into account sex of the person and the age of the person
- This will be similar to the analysis of the US and Japan (we saw in the OECD Report in the previous slides)

Estimating for the entire sample

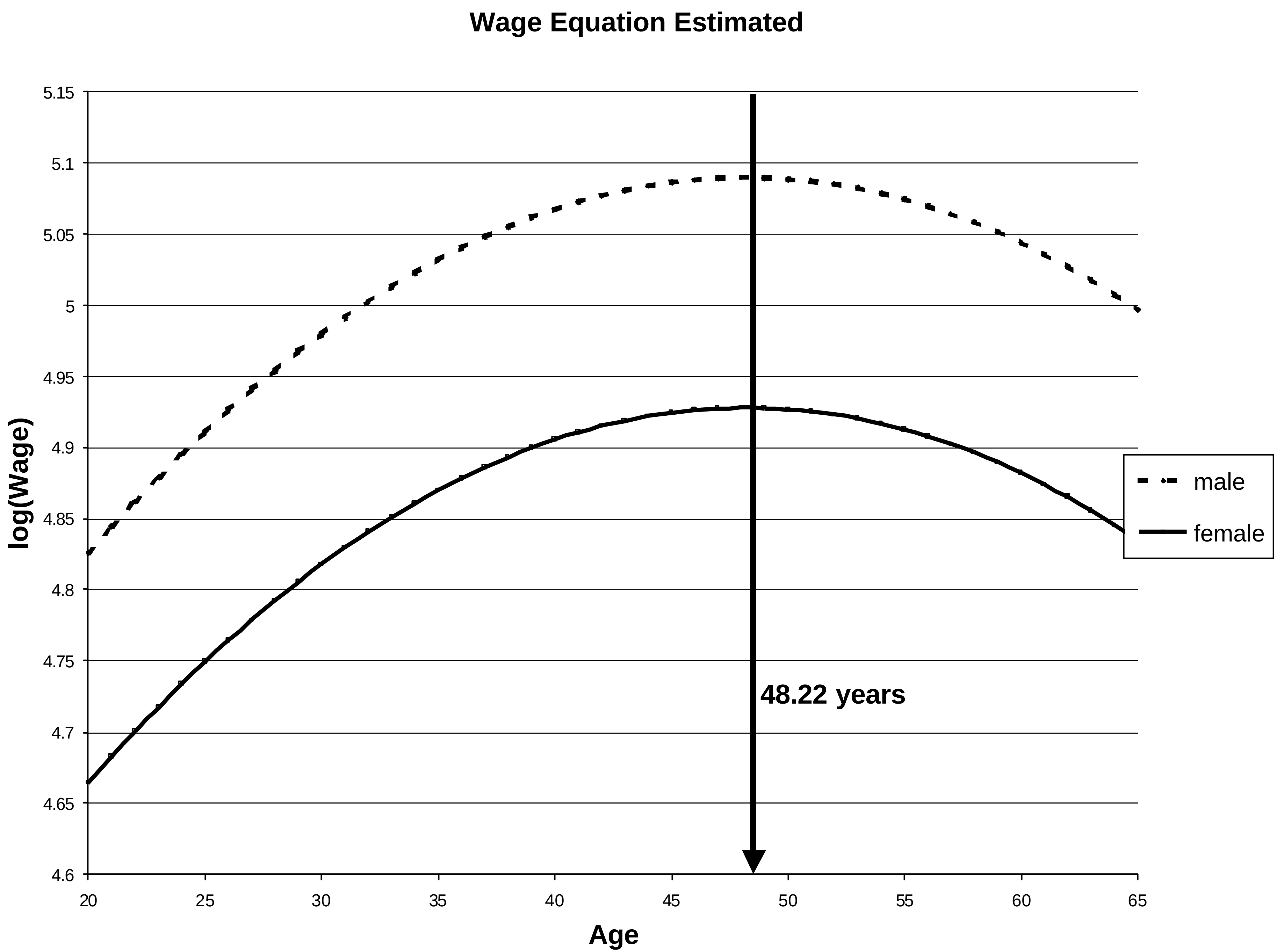
- Estimation Equation:
- $Log(SalProm) = c_0 + c_1.sexo + c_2 edad + c_3 edad^2$
- LOG(SalProm) = 4.482169548 - 0.1617908224*SEXO + 0.0318998435*EDAD - 0.0003307339208*EDAD*EDAD
- Highly significant coefficients

Separate estimates

- Males
- SalProm = -9.935 + 9.311*EDAD - 0.0887*EDAD*EDAD
- Females
- SalProm = 18.471 + 6.712*EDAD - 0.06534*EDAD*EDAD

When does the income starts to decline

- From separate estimates:
- Male 52.47
- Female 51.35
- From combined estimate:
- Highest income at 48.23



Wage Equation Estimated
log(Wage)
5.15
5.1
5.05
5
4.95
4.9
4.85
4.8
4.75
4.7
4.65
4.6
20
25
30
35
40
45
50
55
60
65
Age
48.22 years
male
female

What is wrong with the pictures?

- They are calculated from by combining all the people of different age groups
- They do not tell us anything about **cohort effect**
- Here, we can actually observe the real trajectory of persons over eight years
- What we need to know for constructing income profiles of actual workers is what they earn over their lifetimes – we have segments in our data

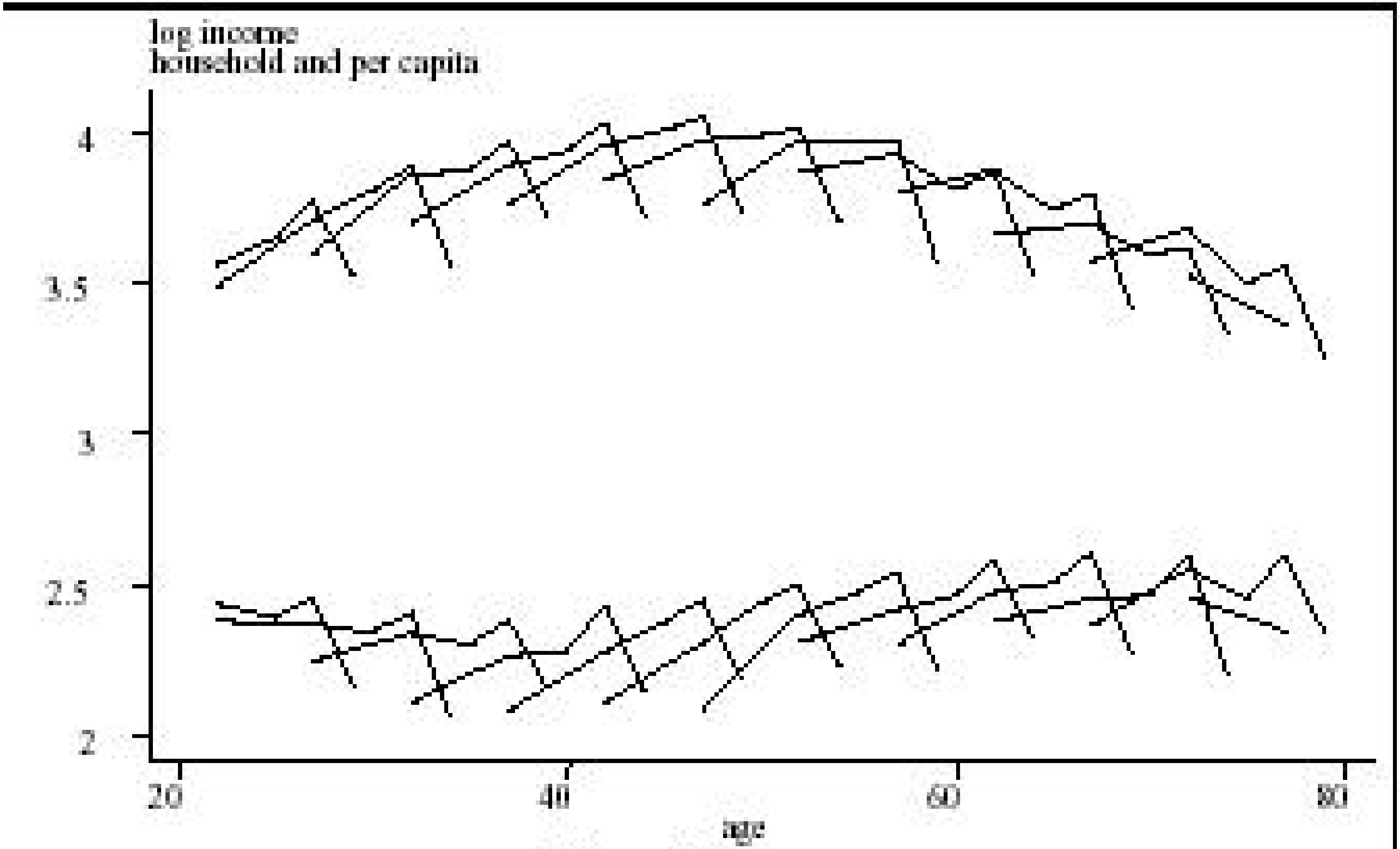


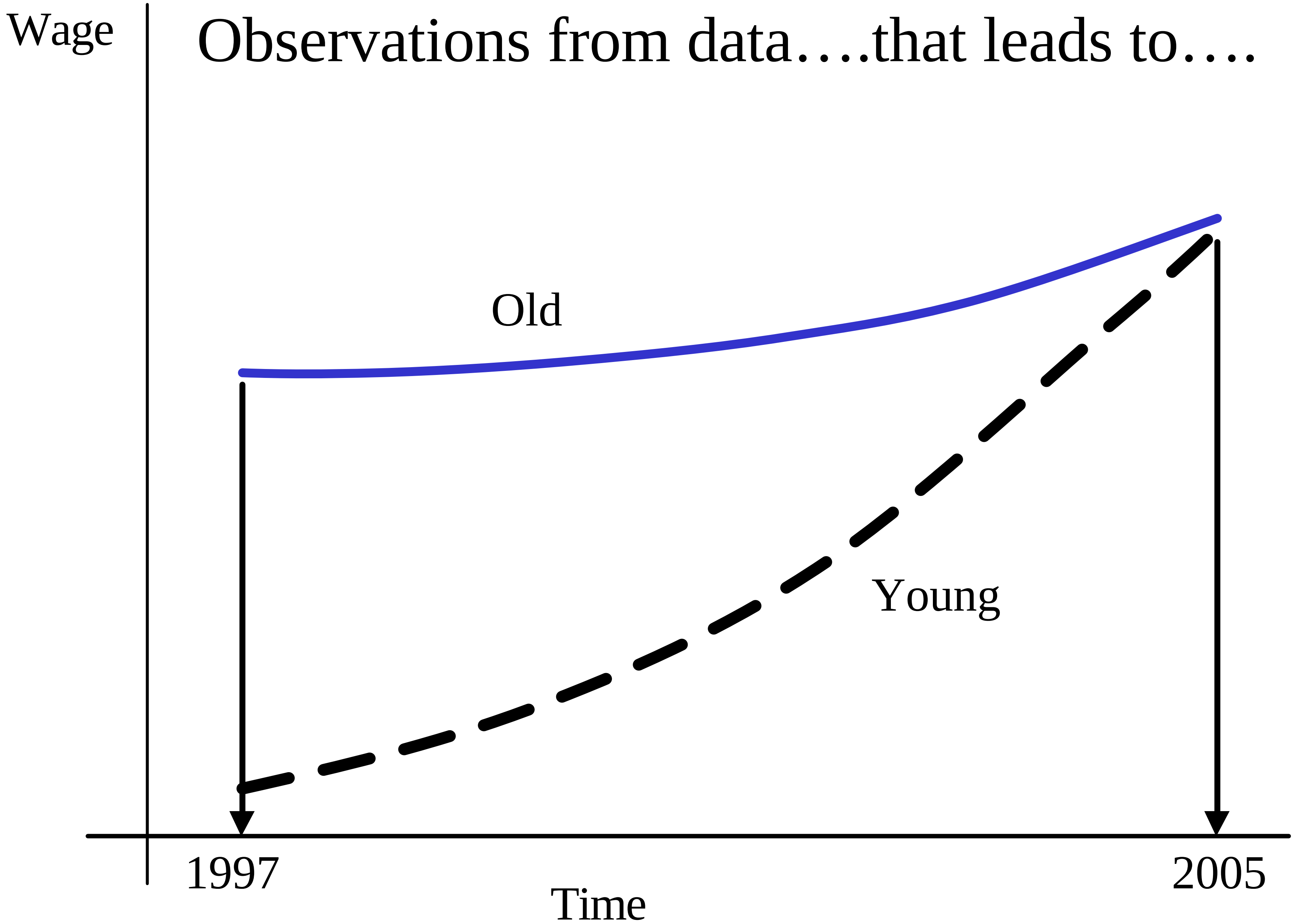
Figure 1b

Household Savings and Income Distribution in Mexico
Orazio P. Attanasio and Miguel Székely, 1998

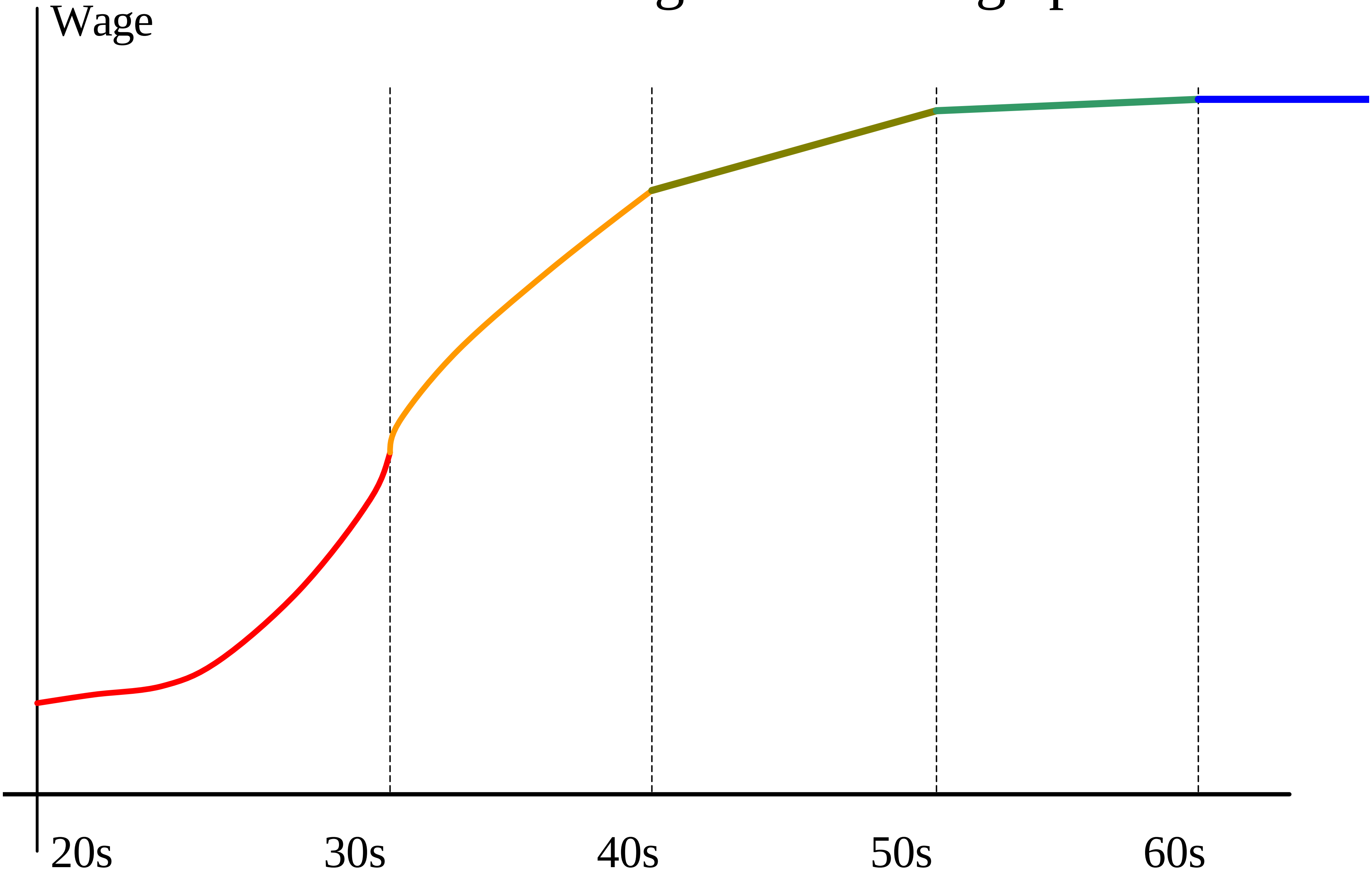
Analyzing data controlling for cohorts and gender

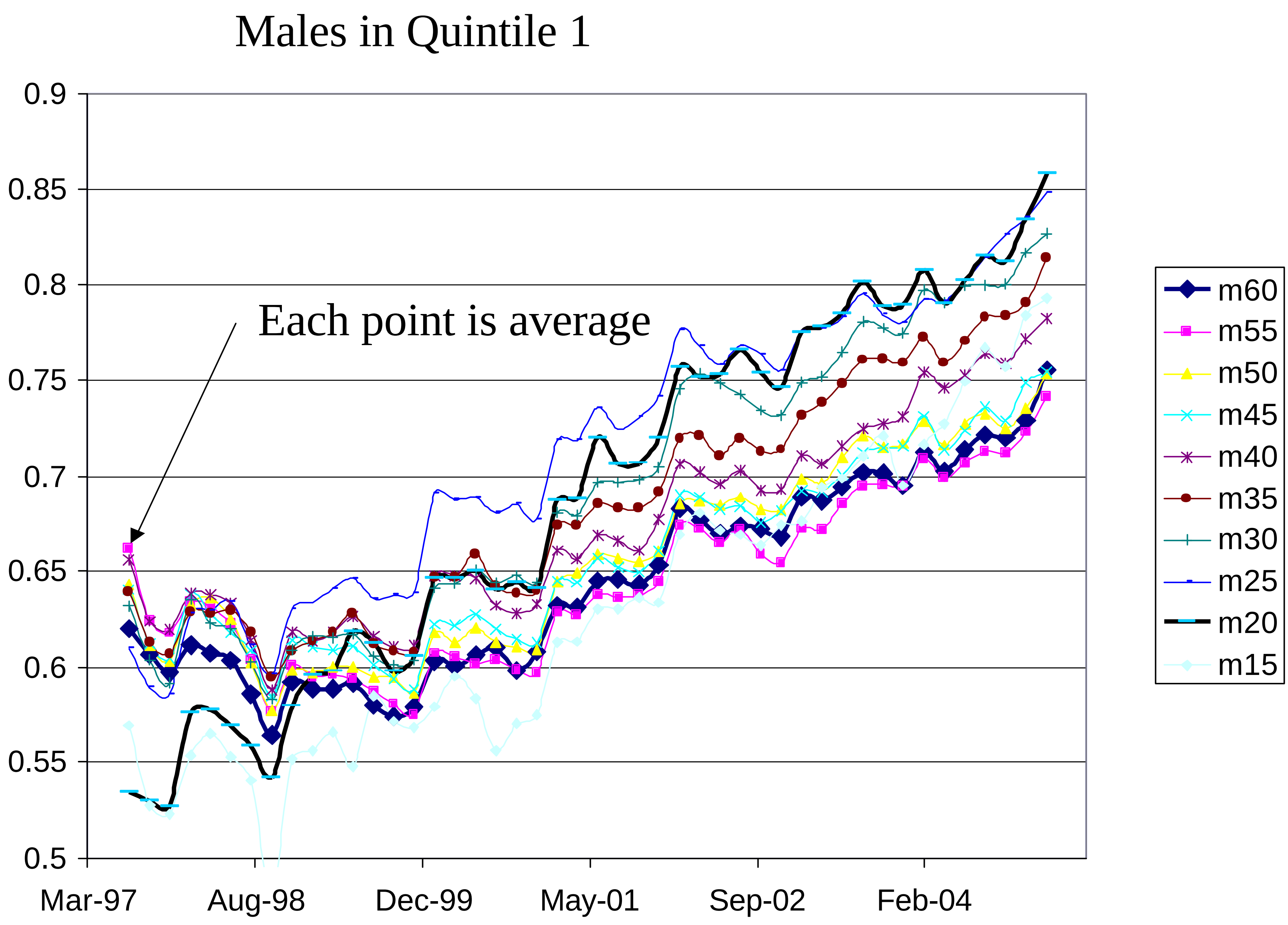
- Now examine the data, controlling for "starting income" (which is a proxy for cohort) and follow them for eight years
- How should it look like?
- Logically, it should depend on the economic performance (more about that at the end)
- It should also depend on cohort
- For young higher growth than old

Observations from data….that leads to….

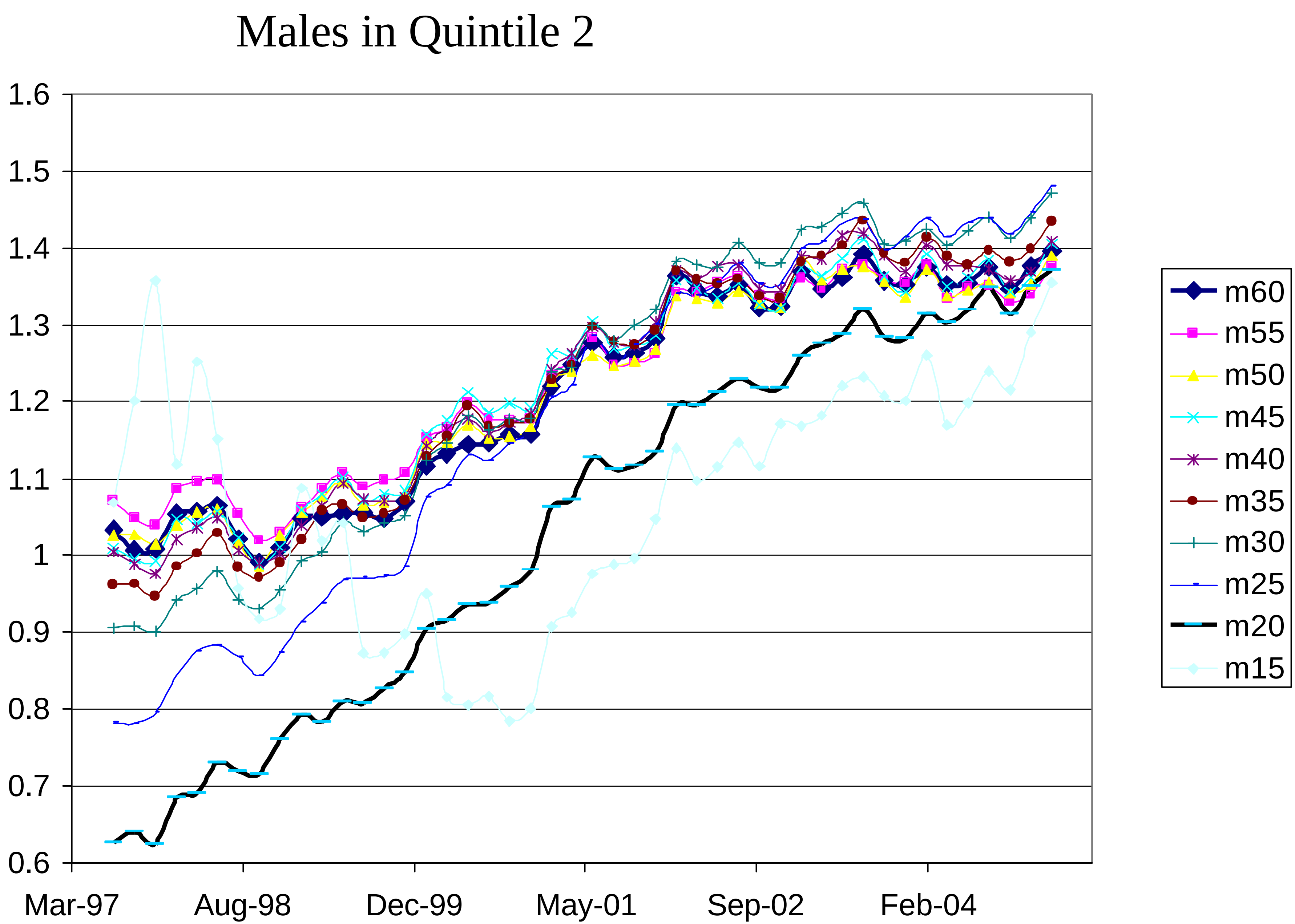


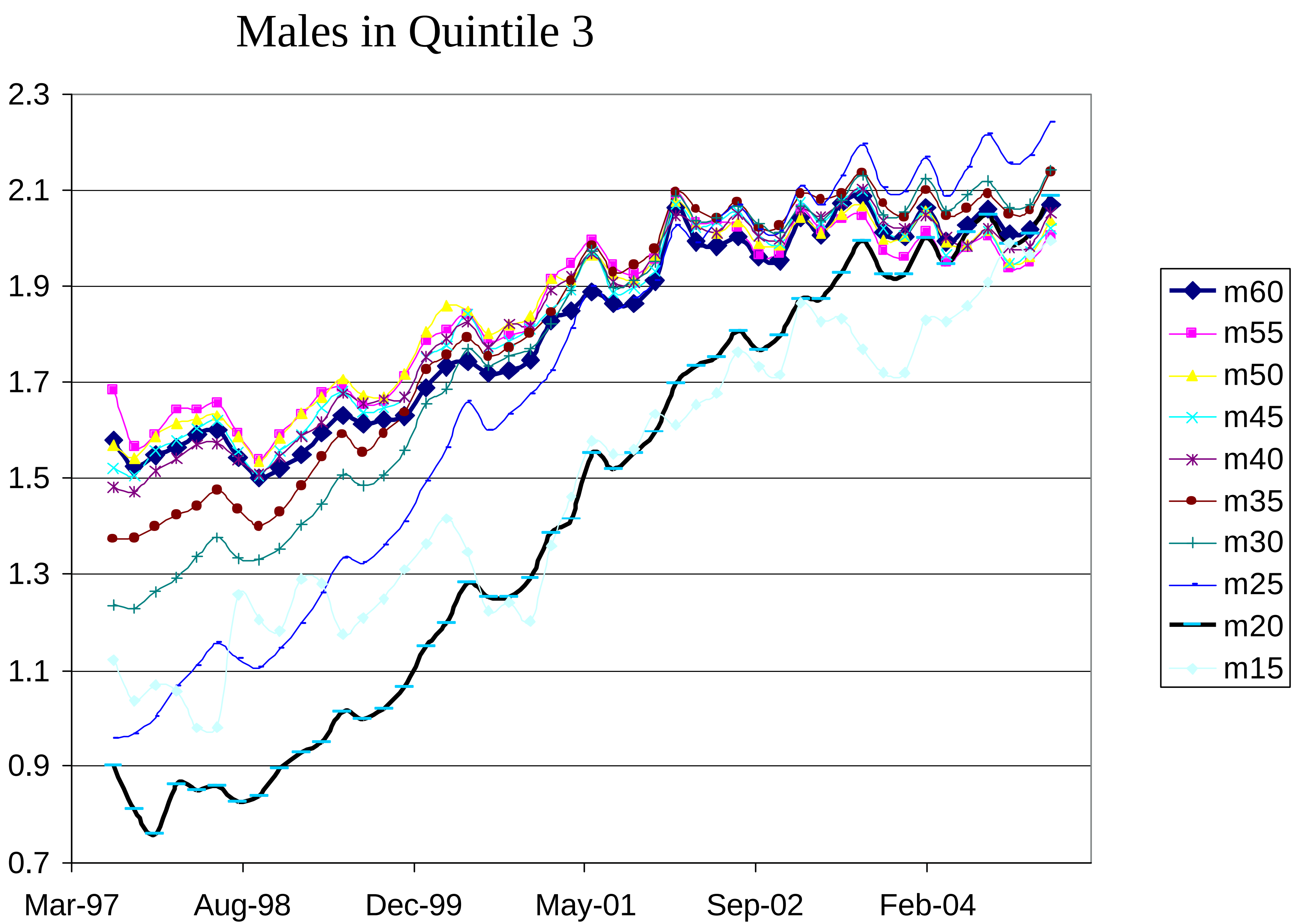
…the following lifetime wage profile

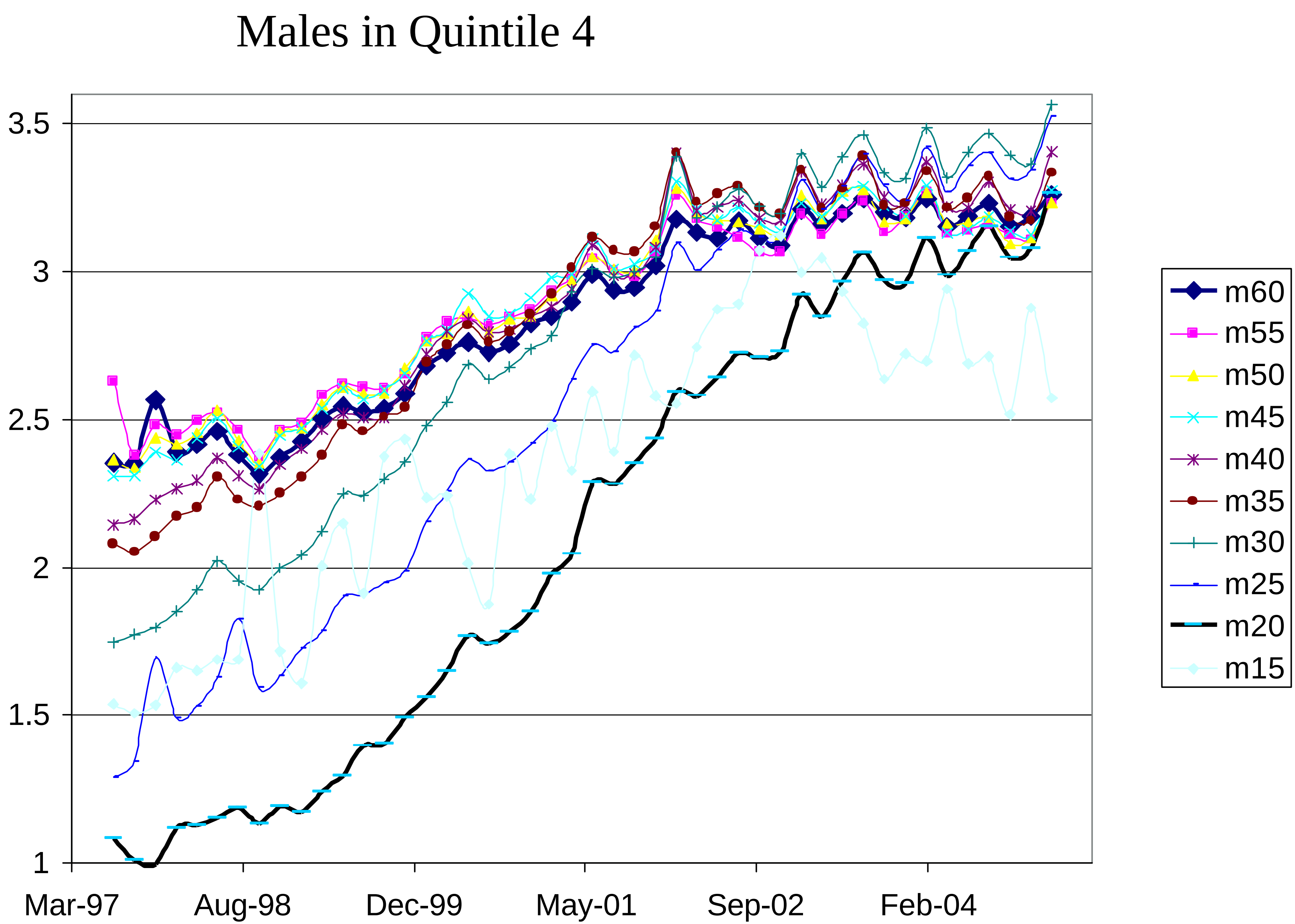




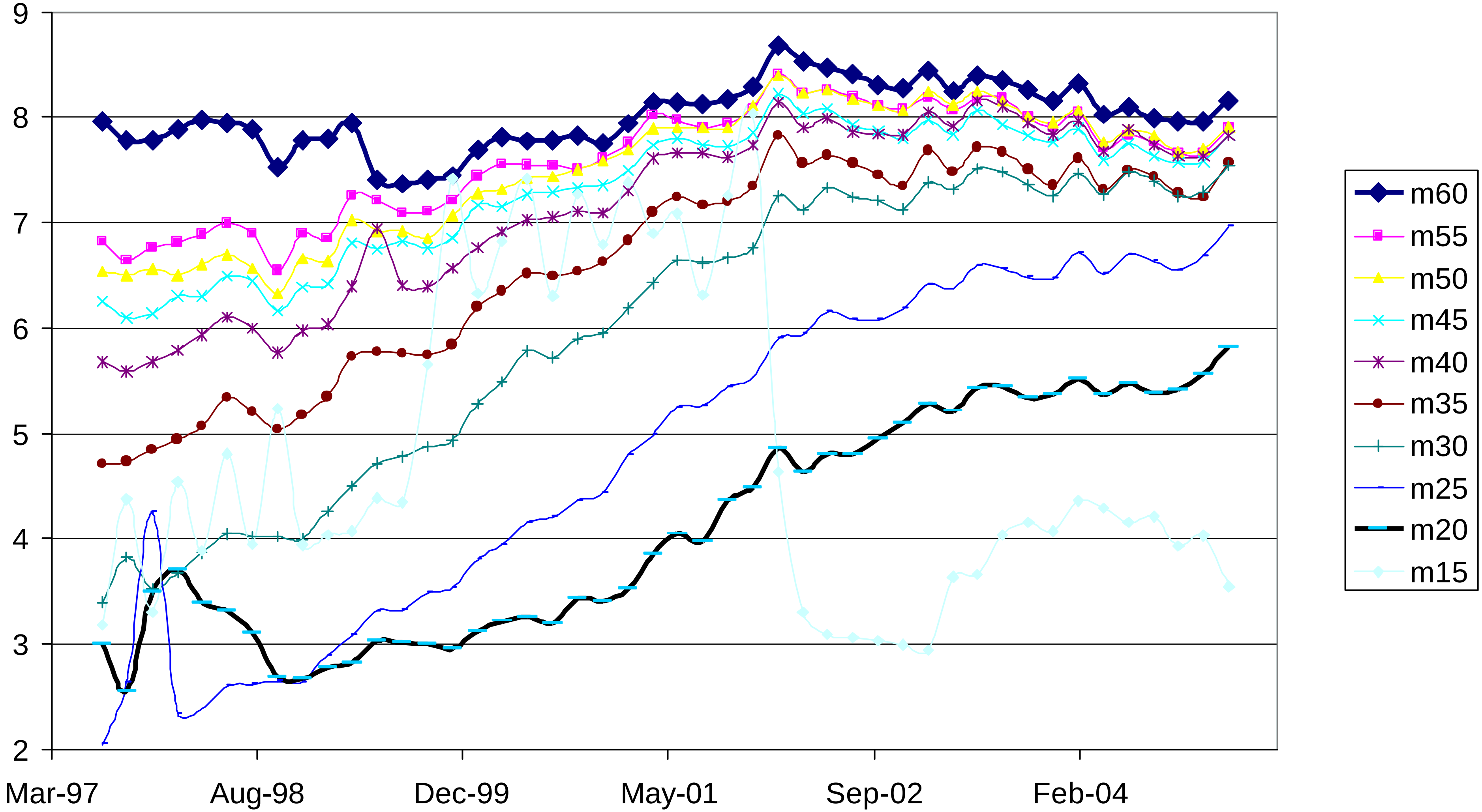
Males in Quintile 1
Each point is average
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
m60
m55
m50
m45
m40
m35
m30
m25
m20
m15

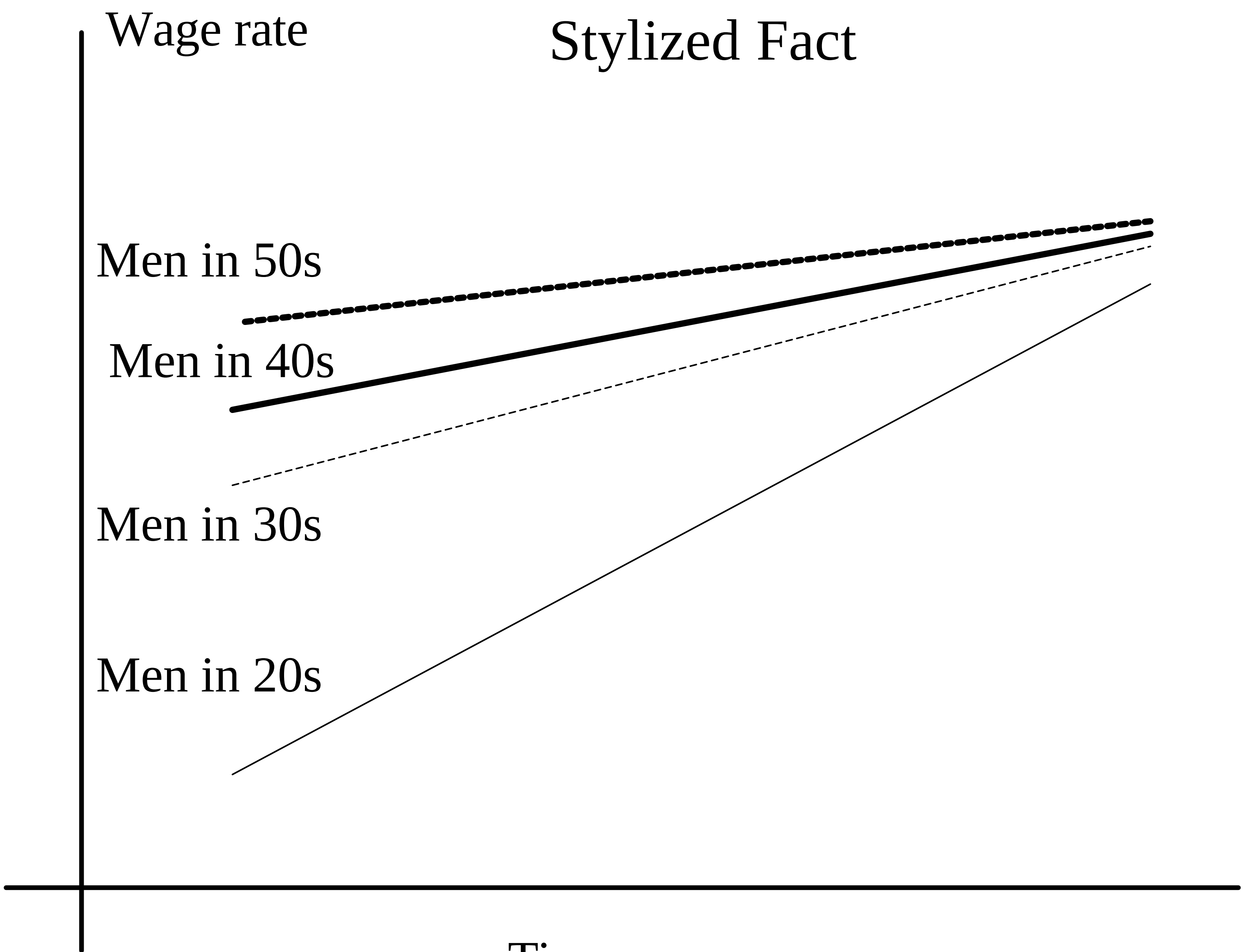

Males in Quintile 2
1.6
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
m60
m55
m50
m45
m40
m35
m30
m25
m20
m15

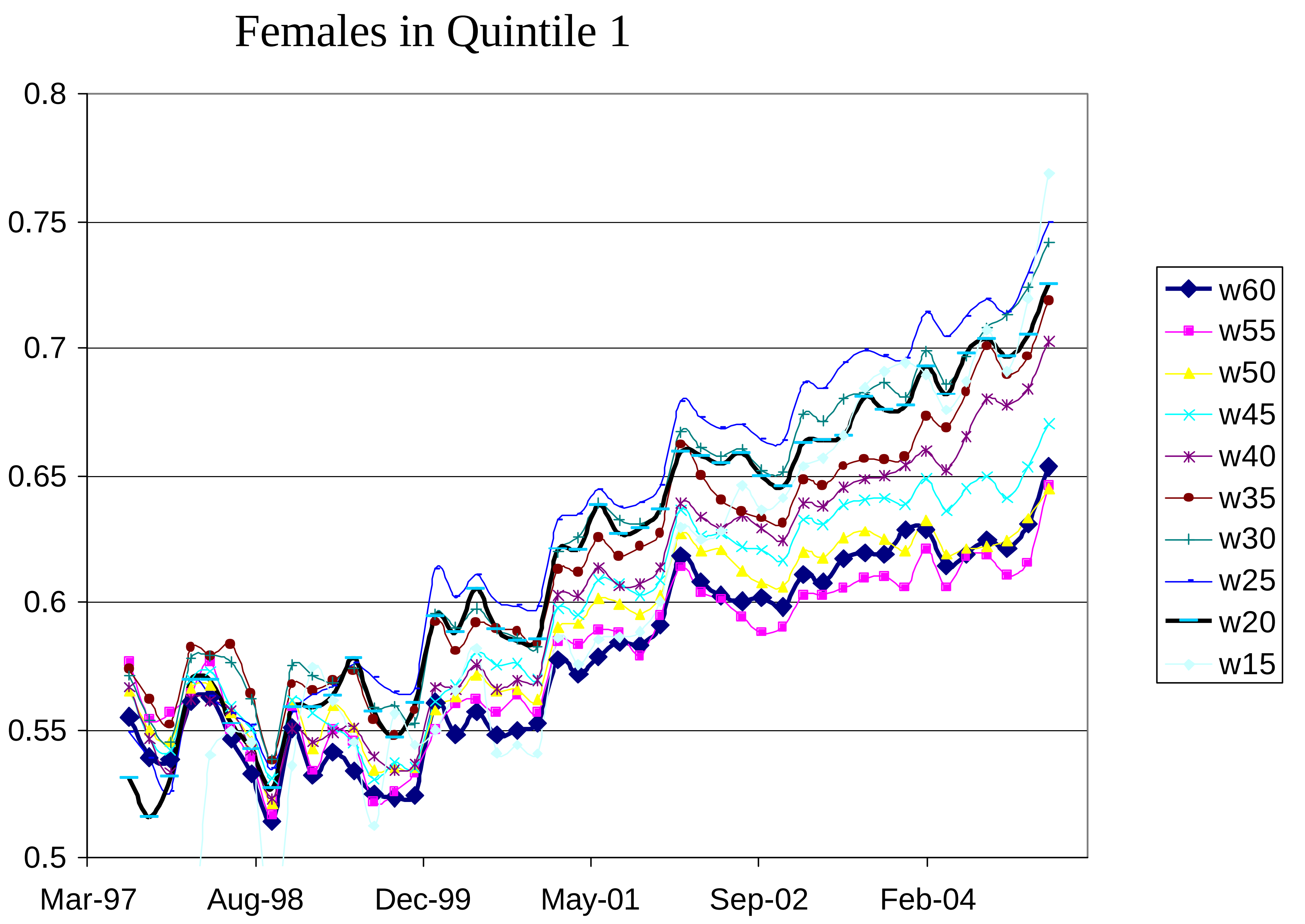

Males in Quintile 3
2.3
2.1
1.9
1.7
1.5
1.3
1.1
0.9
0.7
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
m60
m55
m50
m45
m40
m35
m30
m25
m20
m15

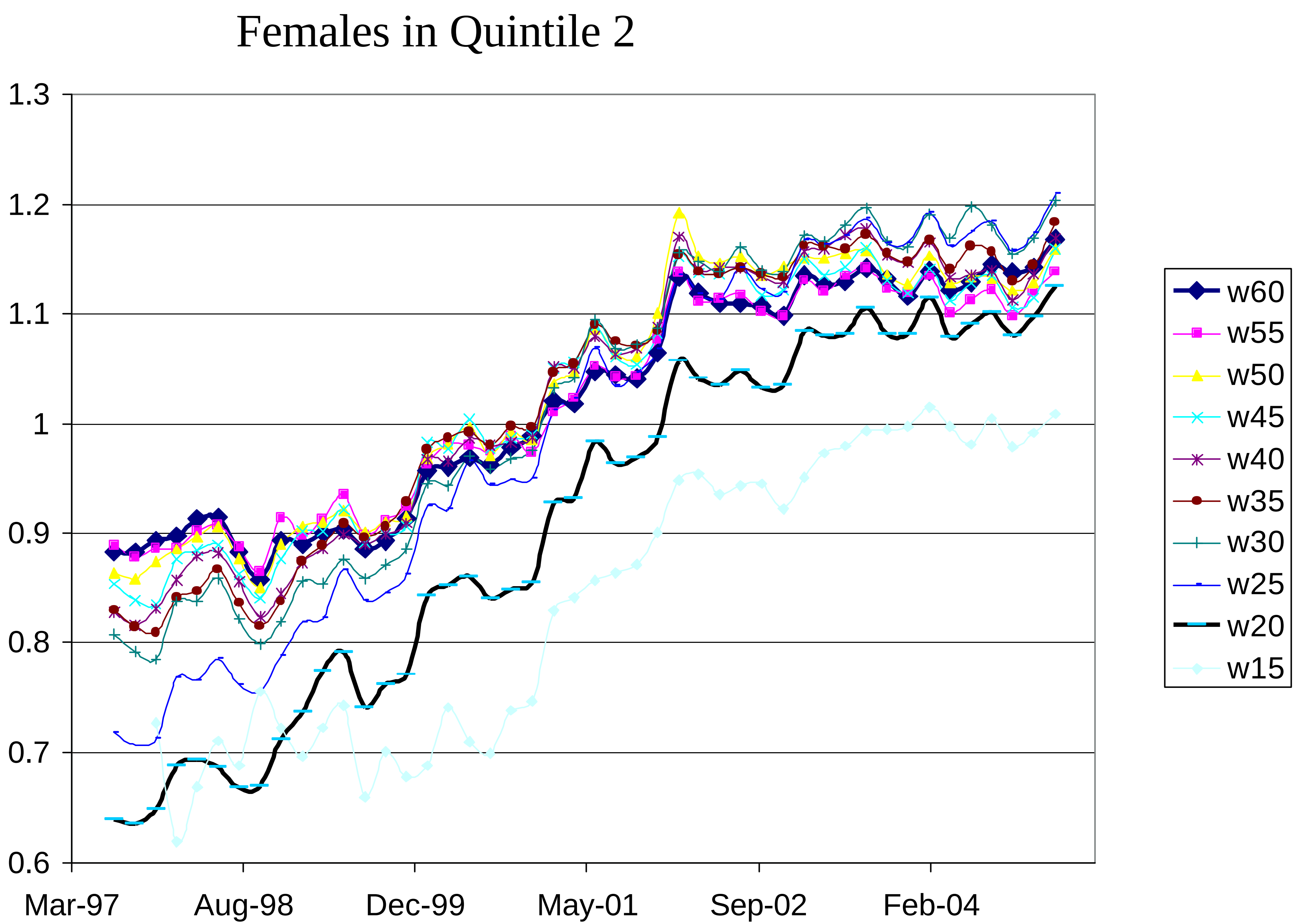

Males in Quintile 4
3.5
3
2.5
2
1.5
1
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
m60
m55
m50
m45
m40
m35
m30
m25
m20
m15

Males in Quintile 5

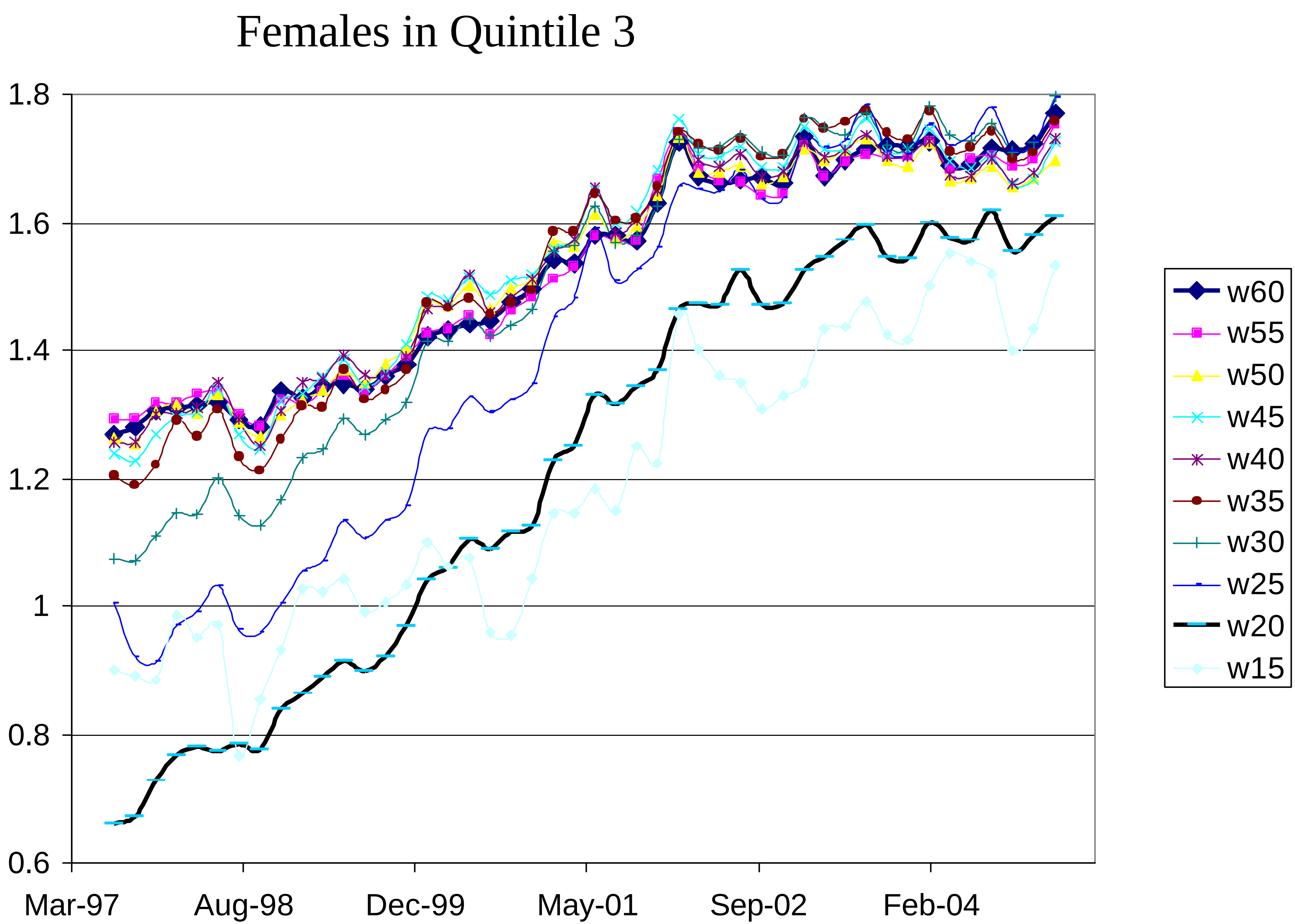



Wage rate
Stylized Fact
Men in 50s
Men in 40s
Men in 30s
Men in 20s
Time


Females in Quintile 1
0.8
0.75
0.7
0.65
0.6
0.55
0.5
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
w60
w55
w50
w45
w40
w35
w30
w25
w20
w15

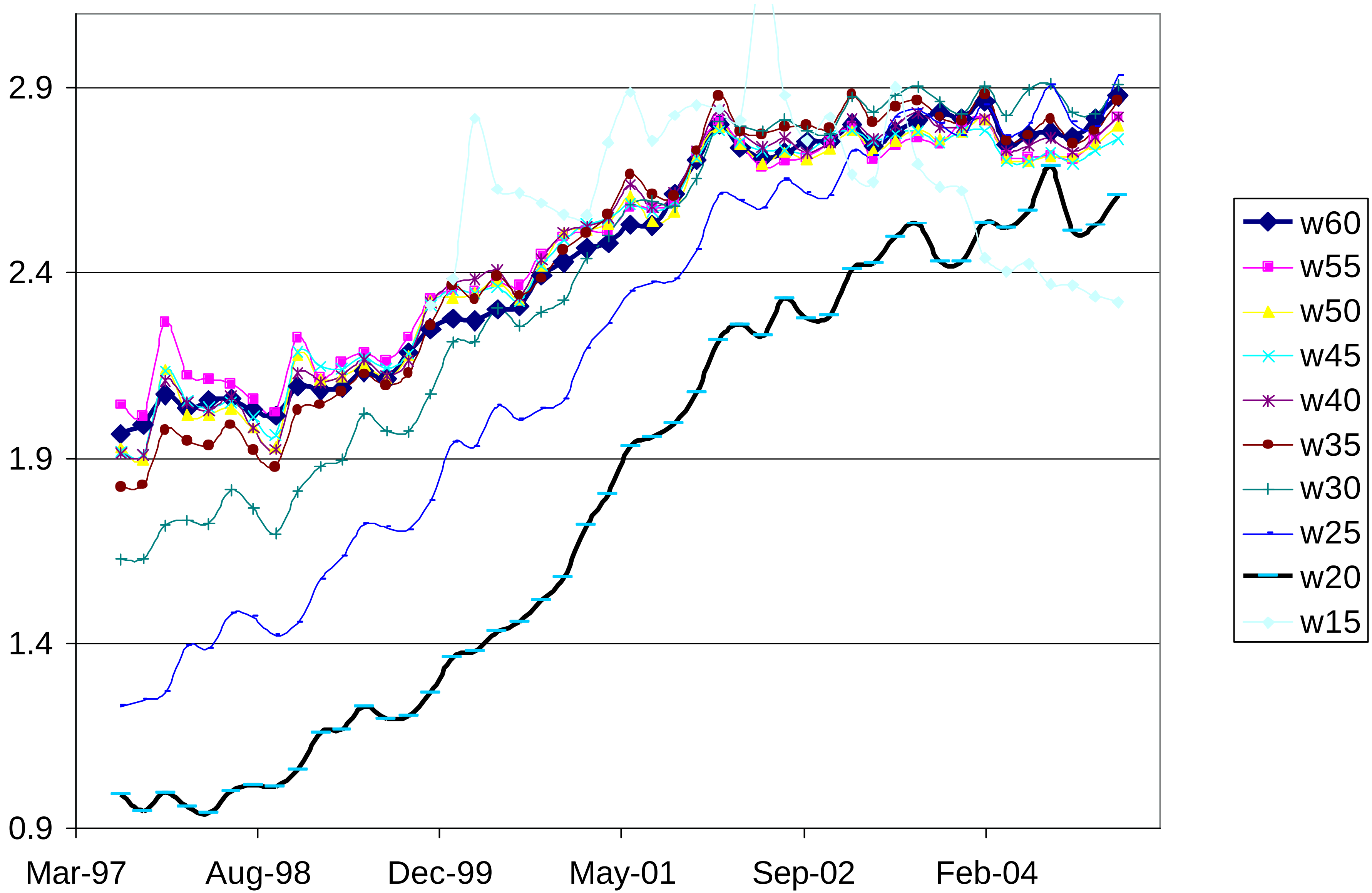


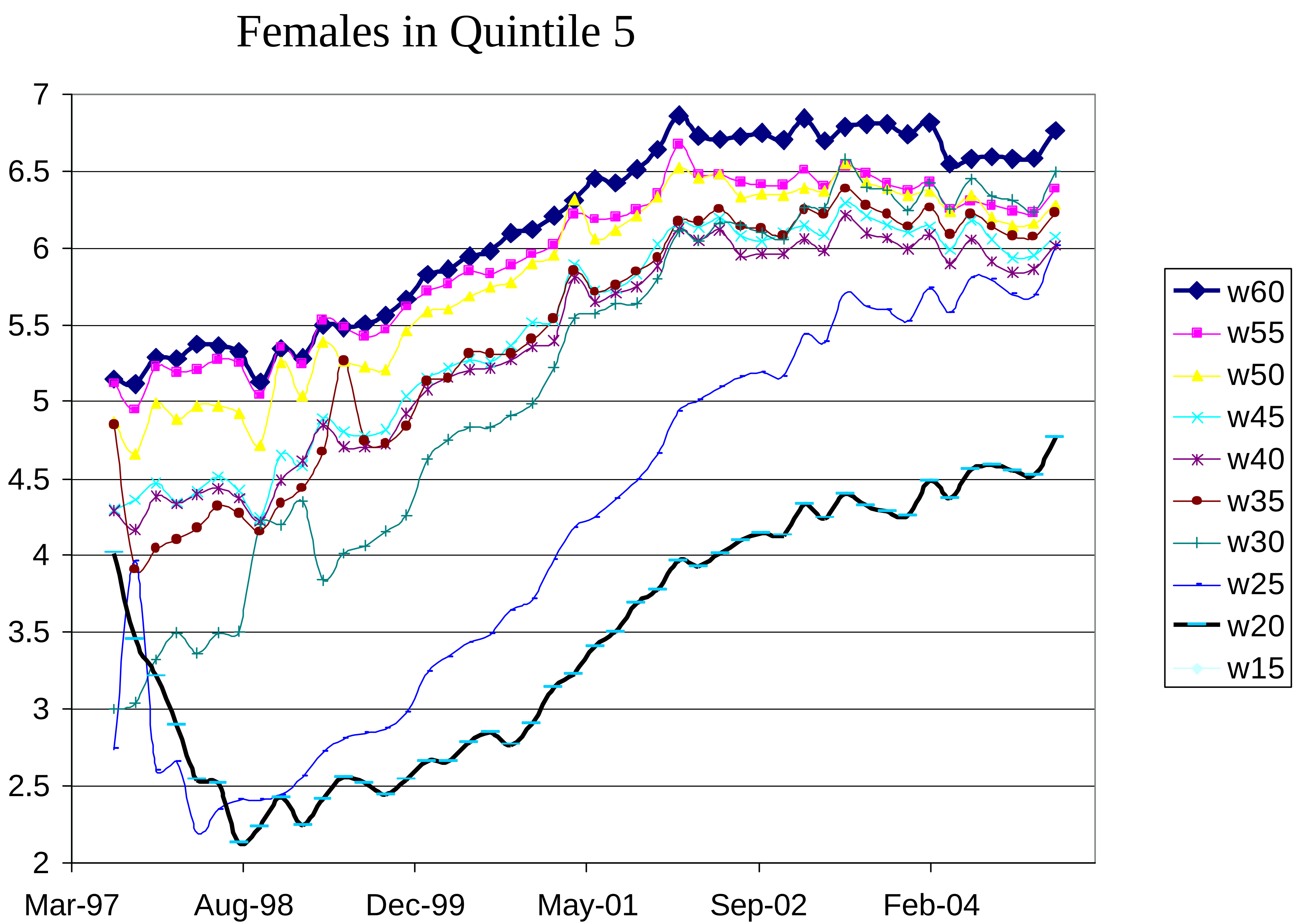
Females in Quintile 2
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
w60
w55
w50
w45
w40
w35
w30
w25
w20
w15

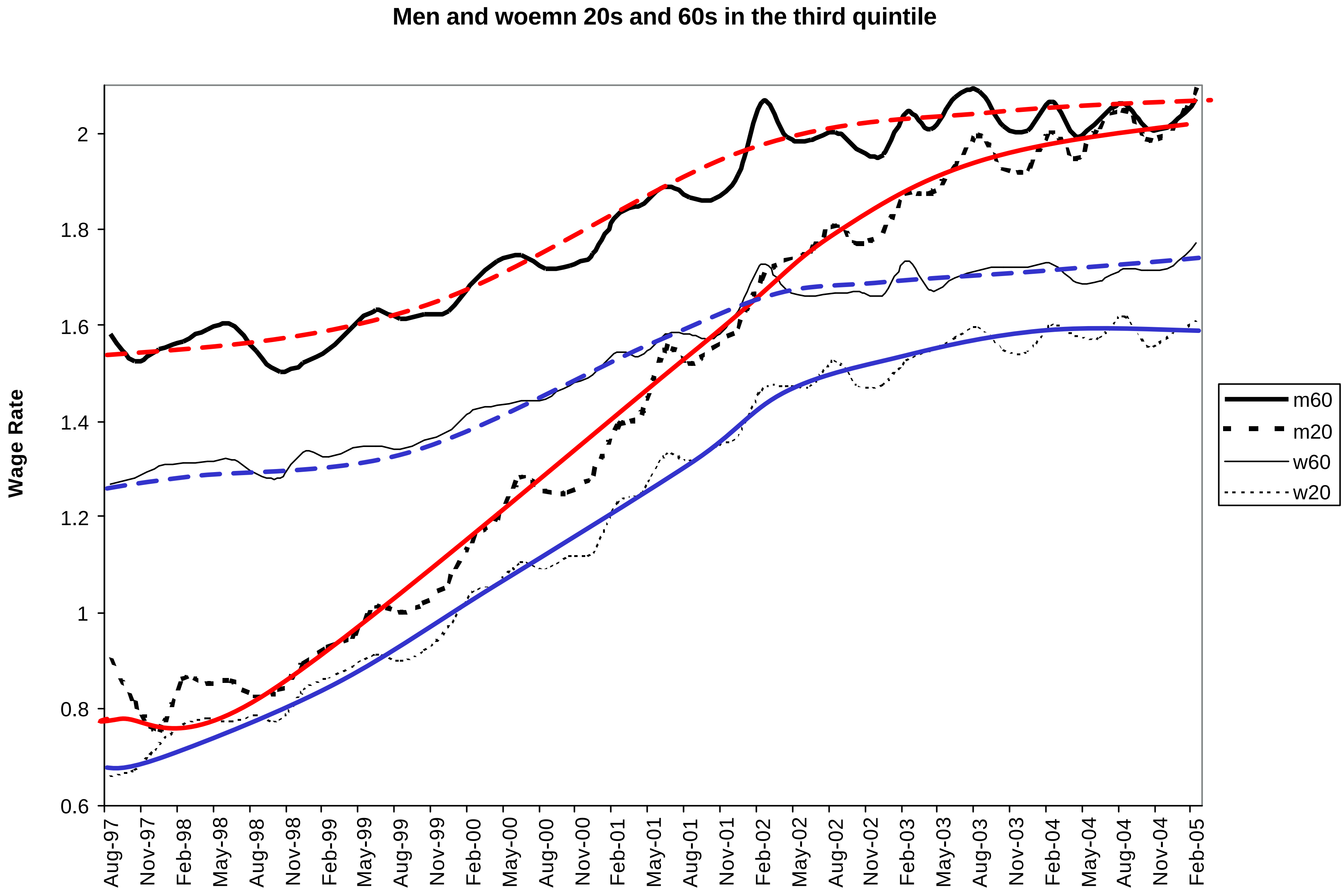


Females in Quintile 3
1.8
1.6
1.4
1.2
1
0.8
0.6
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
w60
w55
w50
w45
w40
w35
w30
w25
w20
w15

Females in Quintile 4



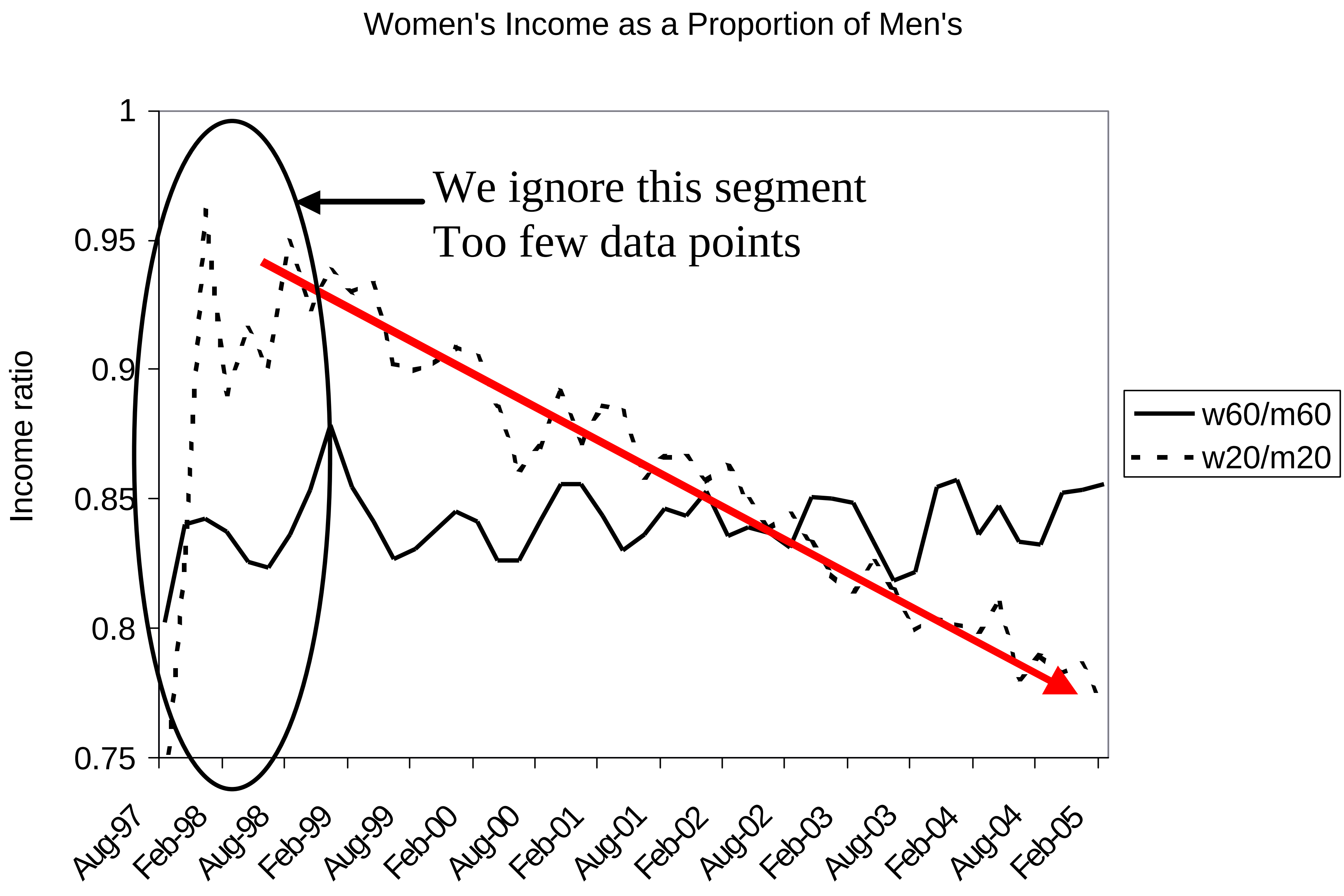

Females in Quintile 5
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
Mar-97
Aug-98
Dec-99
May-01
Sep-02
Feb-04
w60
w55
w50
w45
w40
w35
w30
w25
w20
w15



Men and woemn 20s and 60s in the third quintile
Wage Rate
2
1.8
1.6
1.4
1.2
1
0.8
0.6
Aug-97
Nov-97
Feb-98
May-98
Aug-98
Nov-98
Feb-99
May-99
Aug-99
Nov-99
Feb-00
May-00
Aug-00
Nov-00
Feb-01
May-01
Aug-01
Nov-01
Feb-02
May-02
Aug-02
Nov-02
Feb-03
May-03
Aug-03
Nov-03
Feb-04
May-04
Aug-04
Nov-04
Feb-05
m60
m20
w60
w20

Stylized facts

- For the three middle quintiles, all income of all ages grow
- Men's income always stays above that of women's
- For men and women in their twenties, the income grows faster than for men in their forties, fifties and sixties
- Wage gap between men and women is expanding for women in their twenties but **not** for them in their sixties (see the following slide)


Women's Income as a Proportion of Men's
Income ratio
1
0.95
0.9
0.85
0.8
0.75
Aug-97
Feb-98
Aug-98
Feb-99
Aug-99
Feb-00
Aug-00
Feb-01
Aug-01
Feb-02
Aug-02
Feb-03
Aug-03
Feb-04
Aug-04
Feb-05
We ignore this segment
Too few data points
w60/m60
w20/m20

For the men in their twenties in the third quintile

Dependent Variable: M20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DATE	0.036276	0.003567	10.16877	0
DATE*DATE	-7.20E-05	7.36E-05	-0.978116	0.3335
C	0.639283	0.036348	17.58766	0

R-squared	0.970604	Mean dependent var	1.43933
Durbin-Watson s	0.439952	Prob(F-statistic)	0

For the men in their fifties in the third quintile

Dependent Variable: M50

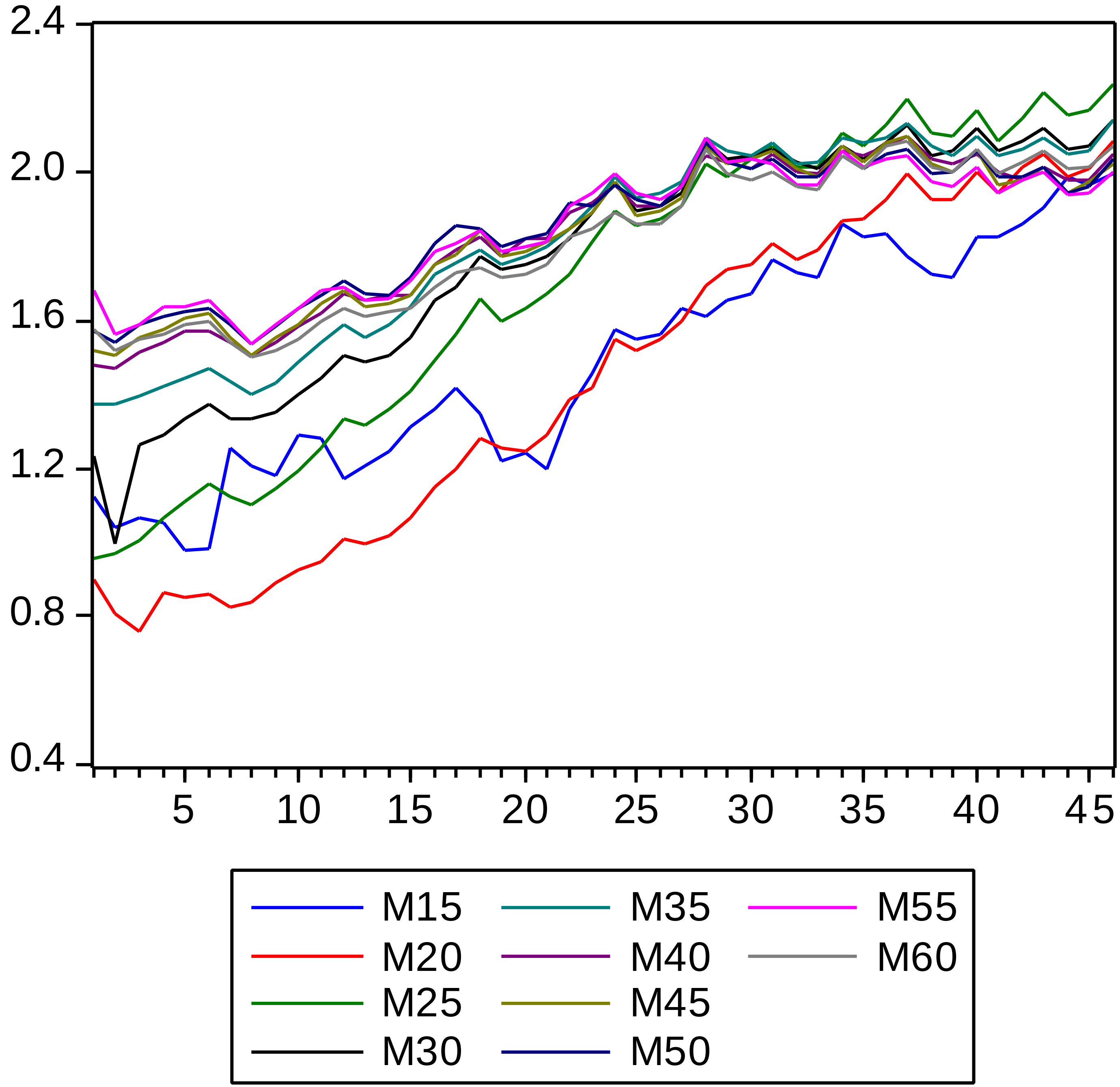
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DATE	0.026763	0.002397	11.16667	0
DATE*DATE	-0.000313	4.94E-05	-6.324	0
C	1.451821	0.02442	59.45233	0

R-squared	0.915035	Mean dependent var	1.85297
Durbin-Watson s	0.754211	Prob(F-statistic)	0

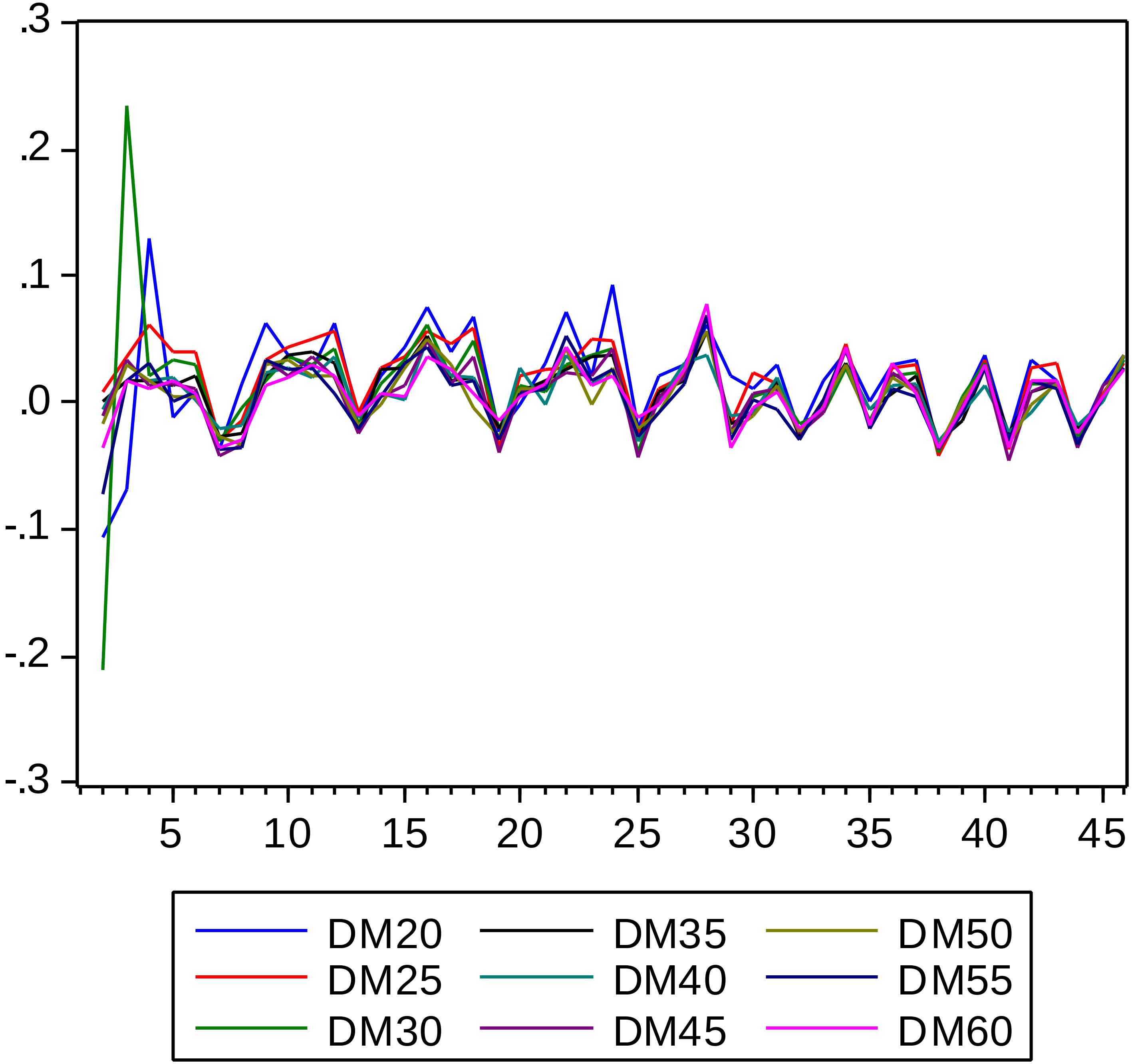
Applying the naïve model

Econometric issues

- Wage rate series for almost all age groups for almost all income quintiles have unit roots
- Thus, simple correlations among level series give spurious correlations
- How do we deal with the problem?
- One solution: Instead of the original series, manage the growth rate of the series



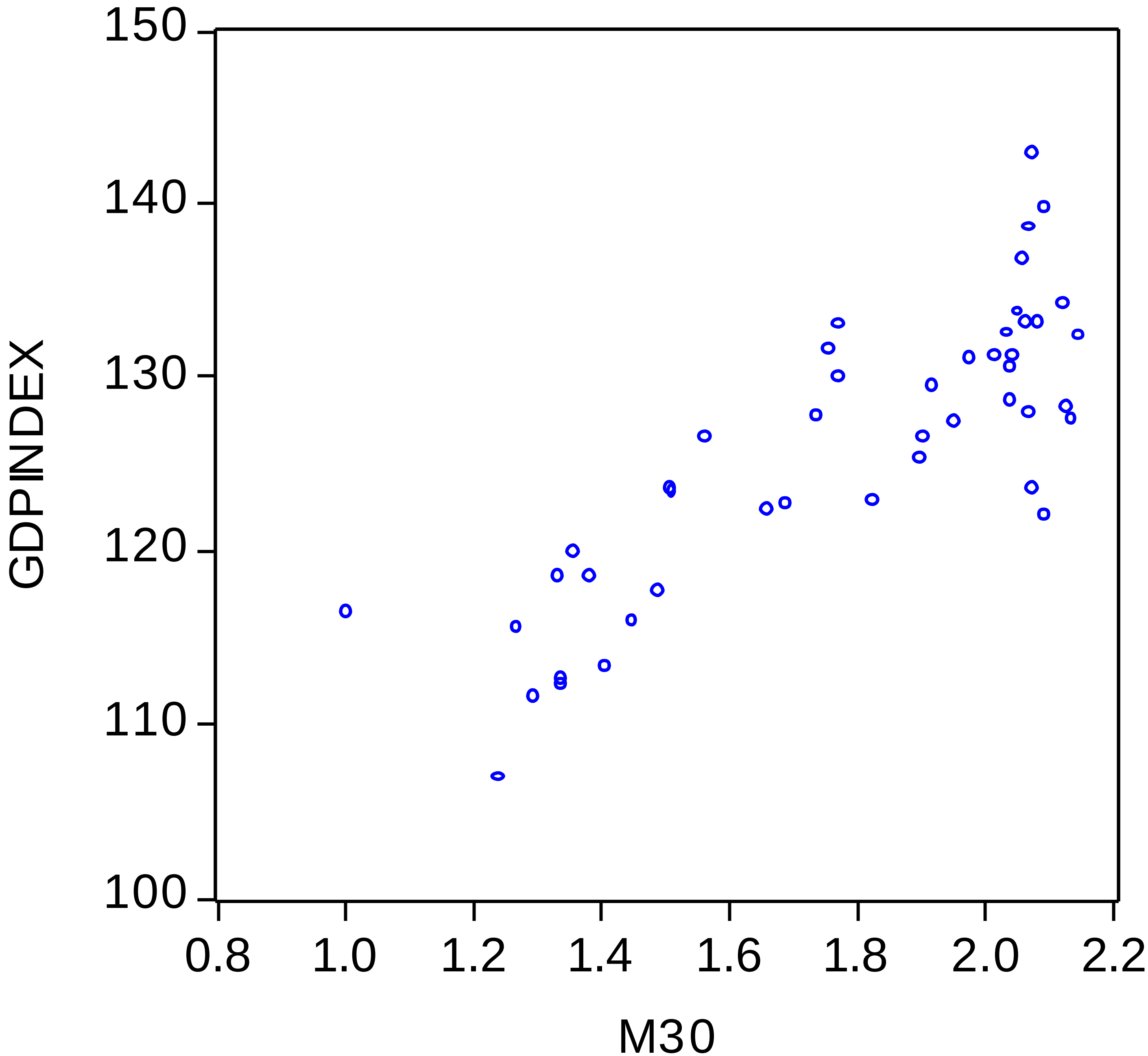
2.4
2.0
1.6
1.2
0.8
0.4
5
10
15
20
25
30
35
40
45
M15
M20
M25
M30
M35
M40
M45
M50
M55
M60



.3
.2
.1
.0
-.1
-.2
-.3
5
10
15
20
25
30
35
40
45
DM20
DM25
DM30
DM35
DM40
DM45
DM50
DM55
DM60

Result

- Even though the original series look quite dispersed, the growth rate (log difference) are almost perfectly synchronized
- What would they be related to?
- One possible culprit: the monthly GDP Index produced by INEGI (actual GDP figures are only available every three months, our series is available every two)



150
140
130
120
110
100
GDPINDEX
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
M30

A simple model

Dependent Variable: M30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDPINDEX	0.032815	0.003322	9.878397	0
C	-2.3587	0.419274	-5.62567	0
R-squared	0.689228	Mean dependent va		1.774597
Adjusted R-squ	0.682165	S.D. dependent var		0.322315
S.E. of regress	0.181711	Akaike info criterio		-0.5303
Sum squared re	1.452829	Schwarz criterion		-0.45079
Log likelihood	14.19678	F-statistic		97.58273
Durbin-Watson	1.370915	Prob(F-statistic)		0

But the model is wrong

- Why?
- Using Augmented Dickey Fuller test we cannot reject the hypothesis that M30 and GDP Index have unit roots
- Next logical pass: cointegrated model

Cointegration test results

Unrestricted Cointegration Rank Test (Trace)				
Hypothesized		Trace	0.05	
No. of CE(s)	Eigenvalue	Statistic	Critical Val	Prob.**
None *	0.34862	28.62281	15.49471	0.0003

We reject the null hypothesis that there is no unit root: one unit root

What comes first? Change in GDP or Wage rise?

Pairwise Granger Causality Tests
Sample: 1 46
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Probability
GDPINDEX does not Granger Cause M30	42	4.54527	0.00492
M30 does not Granger Cause GDPINDEX		5.01425	0.00286

Acknowledgement

- Without CONSAR’s data this research would have been impossible
- In particular, thanks to David Madero and Oscar Roldan
- Thanks to the Institute for Social and Economic Research, Osaka University (computational support: software, hardware)