Some empirical observations on a sample of Icelandic employees

Helgi Tómasson
helgito@hi.is
Háskóli Íslands

January 26th 2009

Plan of talk

1.

Plan of talk

1. Background, IDA/LINDA Icelandic data

Plan of talk

1. Background, IDA/LINDA Icelandic data
2. Some comments on impact of population dynamics and within cohort income dynamics

Plan of talk

1. Background, IDA/LINDA Icelandic data
2. Some comments on impact of population dynamics and within cohort income dynamics
3. The development of an Icelandic data base

Plan of talk

1. Background, IDA/LINDA Icelandic data
2. Some comments on impact of population dynamics and within cohort income dynamics
3. The development of an Icelandic data base
4. A computer exercise with a sample from preliminary version of the data base

Plan of talk

1. Background, IDA/LINDA Icelandic data
2. Some comments on impact of population dynamics and within cohort income dynamics
3. The development of an Icelandic data base
4. A computer exercise with a sample from preliminary version of the data base
5. Summary, the goal of the conference, future research

Background, IDA/LINDA Icelandic data

1.

Background, IDA/LINDA Icelandic data

1. My personal story:

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee
2. The curse of data. Aggretated data can be misleading.

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee
2. The curse of data. Aggretated data can be misleading.
3. Important features of the underlying process have to captured by the model. Avoid the omitted variable bias.

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee
2. The curse of data. Aggretated data can be misleading.
3. Important features of the underlying process have to captured by the model. Avoid the omitted variable bias.
4. All statistical analysis are performed through a model

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee
2. The curse of data. Aggretated data can be misleading.
3. Important features of the underlying process have to captured by the model. Avoid the omitted variable bias.
4. All statistical analysis are performed through a model
5. We hope the individual data will make models available that give us an improved understanding of the labour market.

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee
2. The curse of data. Aggretated data can be misleading.
3. Important features of the underlying process have to captured by the model. Avoid the omitted variable bias.
4. All statistical analysis are performed through a model
5. We hope the individual data will make models available that give us an improved understanding of the labour market.
6. IDA and LINDA are huge projects that require a lot of labour

Background, IDA/LINDA Icelandic data

1. My personal story: Under the supervision of professor Anders Klevmarken I was introduced to the cocept of modelling income-distributions. The toolbox constisted of Markov-models and individual data of male employee
2. The curse of data. Aggretated data can be misleading.
3. Important features of the underlying process have to captured by the model. Avoid the omitted variable bias.
4. All statistical analysis are performed through a model
5. We hope the individual data will make models available that give us an improved understanding of the labour market.
6. IDA and LINDA are huge projects that require a lot of labour
7. In Iceland we are just beginning. After doctoral studies I worked with firm-individual based data (Kjararannsóknarnefnd). Organizing data of such type is a big project.

Aggregation can be misleading

	High wages	Low wages
Males	18	12
Females	7	3

Table: Firm A. 70% of females high pay, 60% of males high pay.

	High wages	Low wages
Males	2	8
Females	9	21

Table: Firm B. 30% of females high pay, 20% of males high pay.

	High wages	Low wages
Males	20	20
Females	16	24

Table: Firm A+B. 40% of females high pay, 50% of males high pay.

The omitted variable bias, avoid 2x2 tables

	job=1	job=2
Males	154.000	241.429
Females	126.667	200.000

Table: Average pay by job-category and gender

	age-group=1	age-group=2
Males	145.000	235.000
Females	122.500	190.000

Table: Average pay by age-category and gender

Some datasnooping with experimental version of an Icelandic database

-

Some datasnooping with experimental version of an Icelandic database

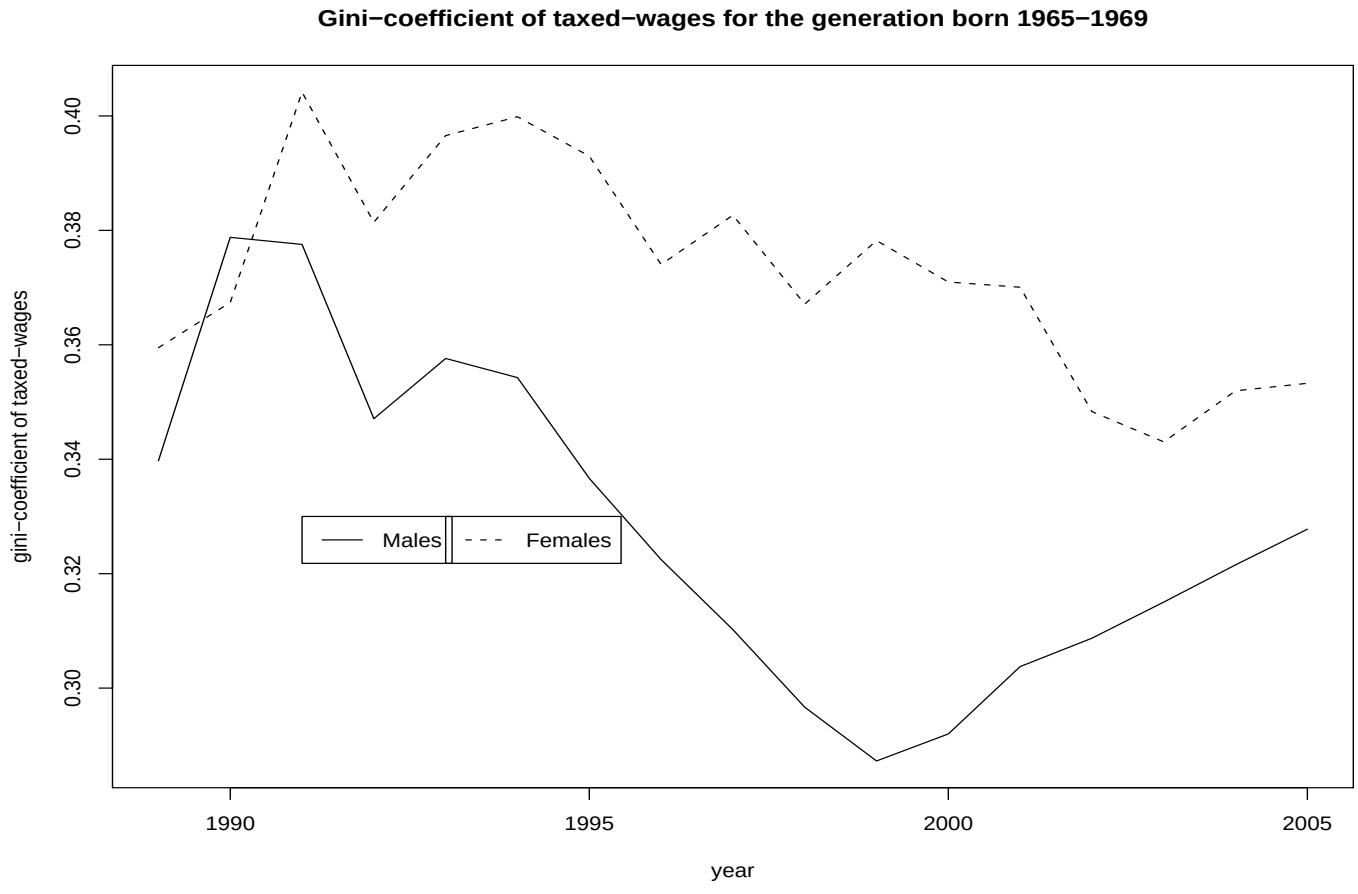
- A sample of individuals with taxed wages in the period 1989-2005.

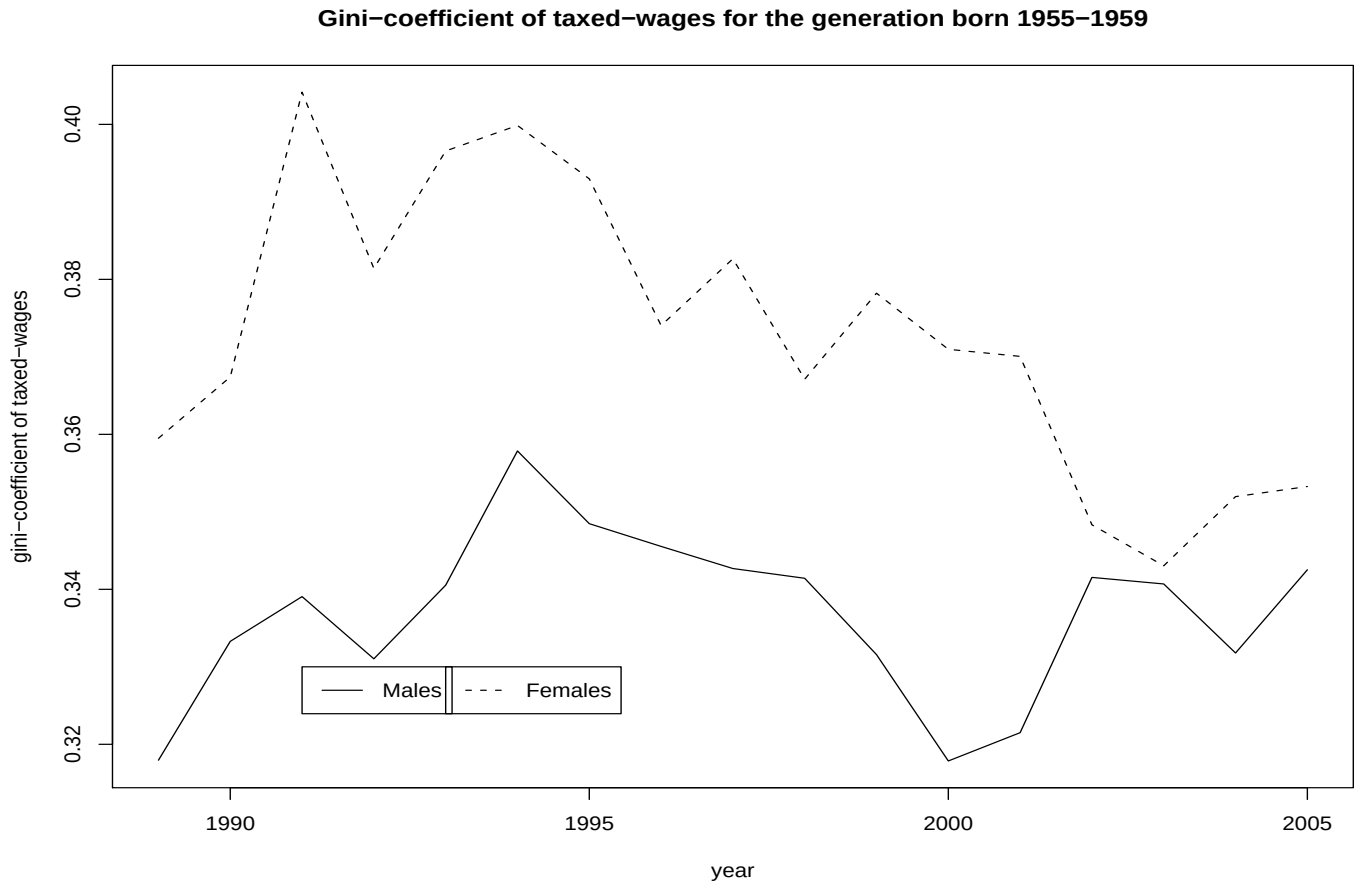
Some datasnooping with experimental version of an Icelandic database

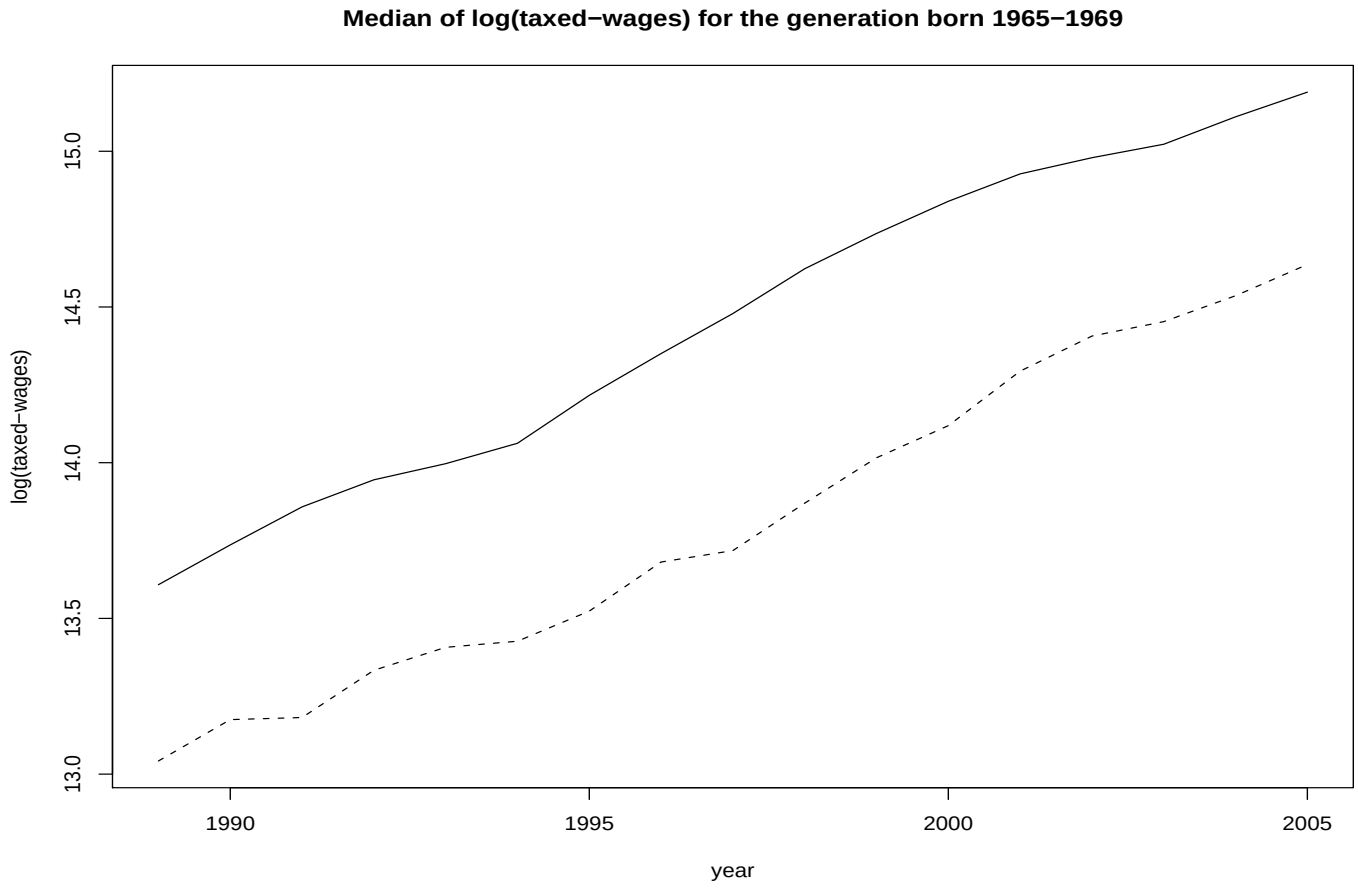
- A sample of individuals with taxed wages in the period 1989-2005.
- Everyone who married into to sample was also included

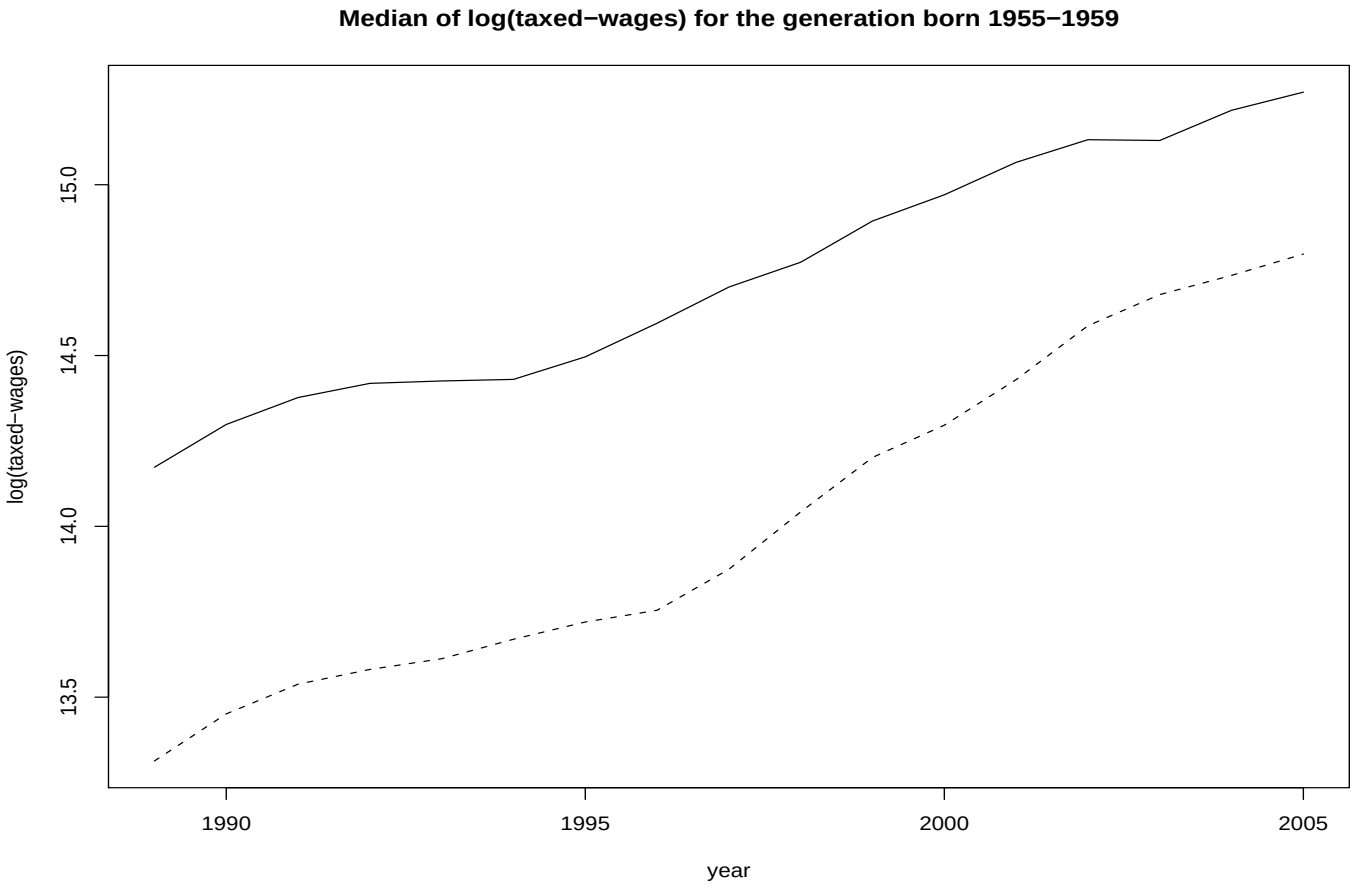
Some datasnooping with experimental version of an Icelandic database

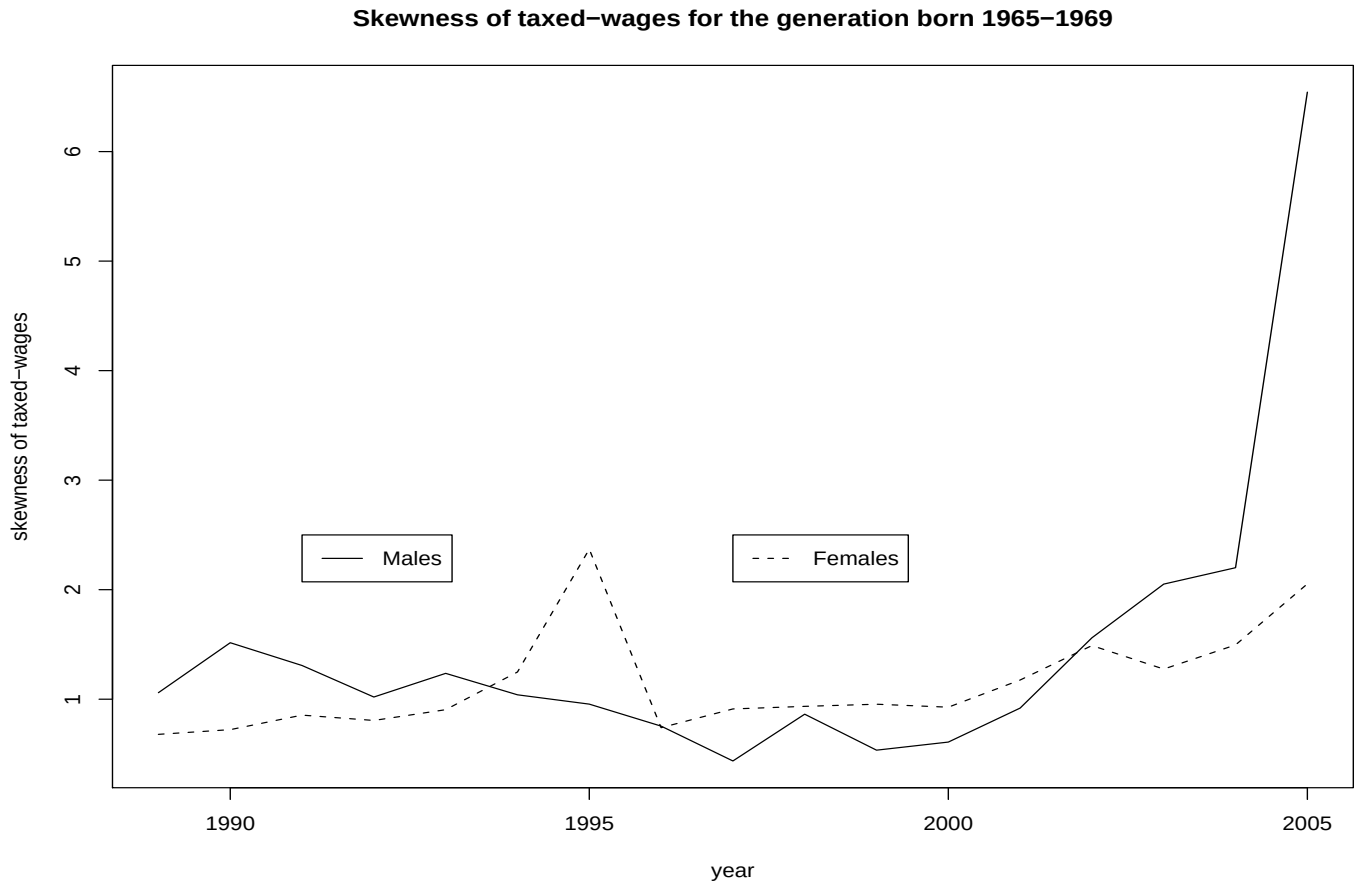
- A sample of individuals with taxed wages in the period 1989-2005.
- Everyone who married into to sample was also included
- Data consists of 13.671 individuals on a yearly basis.


Gini−coefficient of taxed−wages for the generation born 1965−1969
gini−coefficient of taxed−wages
0.30
0.32
0.34
0.36
0.38
0.40
1990
1995
2000
2005
year
Males
Females


Gini−coefficient of taxed−wages for the generation born 1955−1959
gini−coefficient of taxed−wages
0.32
0.34
0.36
0.38
0.40
1990
1995
2000
2005
year
Males
Females


Median of log(taxed−wages) for the generation born 1965−1969
log(taxed−wages)
13.0
13.5
14.0
14.5
15.0
1990
1995
2000
2005
year


Median of log(taxed−wages) for the generation born 1955−1959
log(taxed−wages)
13.5
14.0
14.5
15.0
1990
1995
2000
2005
year


Skewness of taxed−wages for the generation born 1965−1969
skewness of taxed−wages
1
2
3
4
5
6
1990
1995
2000
2005
year
Males
Females

Building block of stochastic dynamics

-

Building block of stochastic dynamics

- Markov-processes, discrete or continuous time.

Building block of stochastic dynamics

- Markov-processes, discrete or continuous time.
- E.g. stochastic differential equations (SDE):

$$dX = \mu(X(t), t)dt + \sigma(X(t), t)dW$$

Building block of stochastic dynamics

- Markov-processes, discrete or continuous time.
- E.g. stochastic differential equations (SDE):

$$dX = \mu(X(t), t)dt + \sigma(X(t), t)dW$$

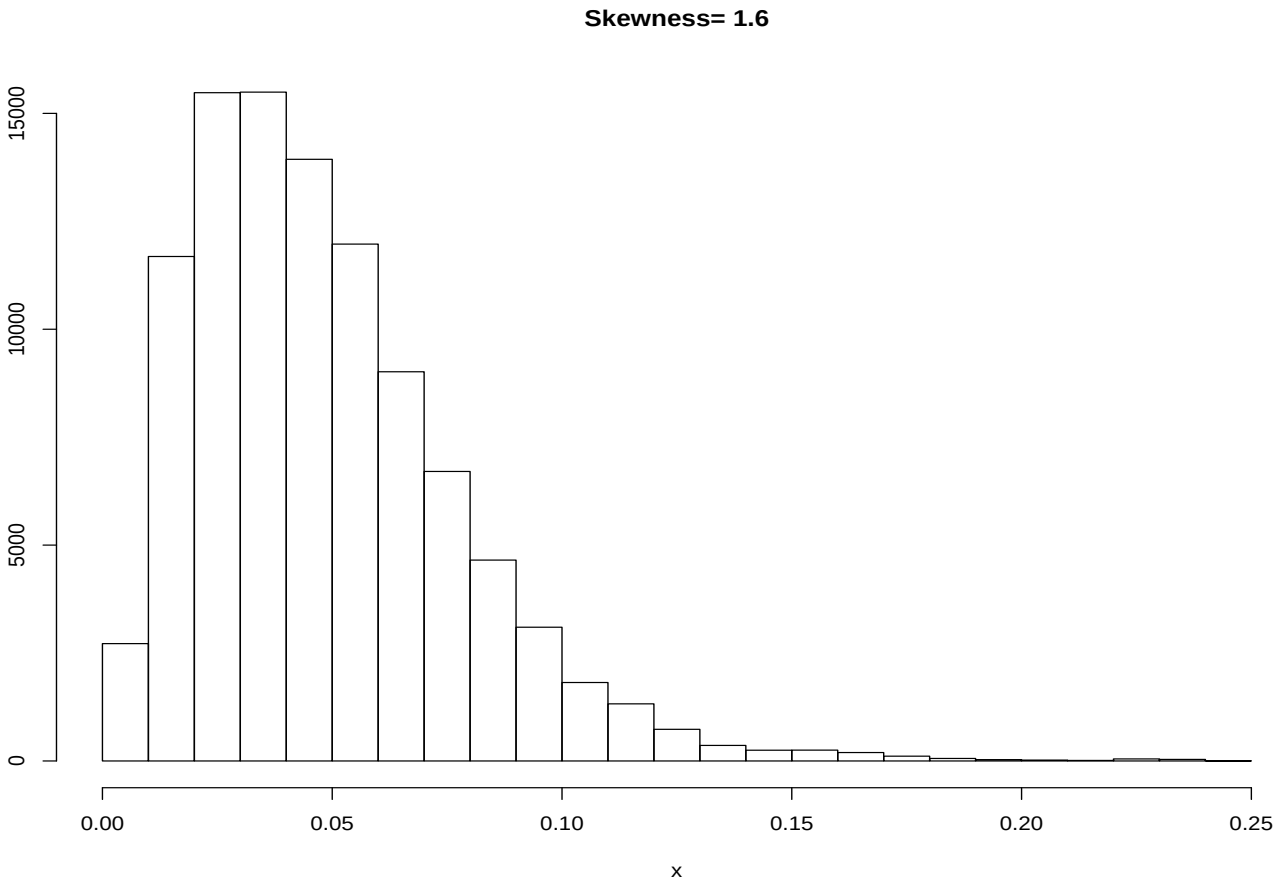
- Some very simple SDE can generate quite wild patterns.

Building block of stochastic dynamics

- Markov-processes, discrete or continuous time.
- E.g. stochastic differential equations (SDE):

$$dX = \mu(X(t), t)dt + \sigma(X(t), t)dW$$

- Some very simple SDE can generate quite wild patterns.
- Sometimes an equilibrium distribution can be derived.


Skewness= 1.6
15000
10000
5000
0
0.00
0.05
0.10
0.15
0.20
0.25
x

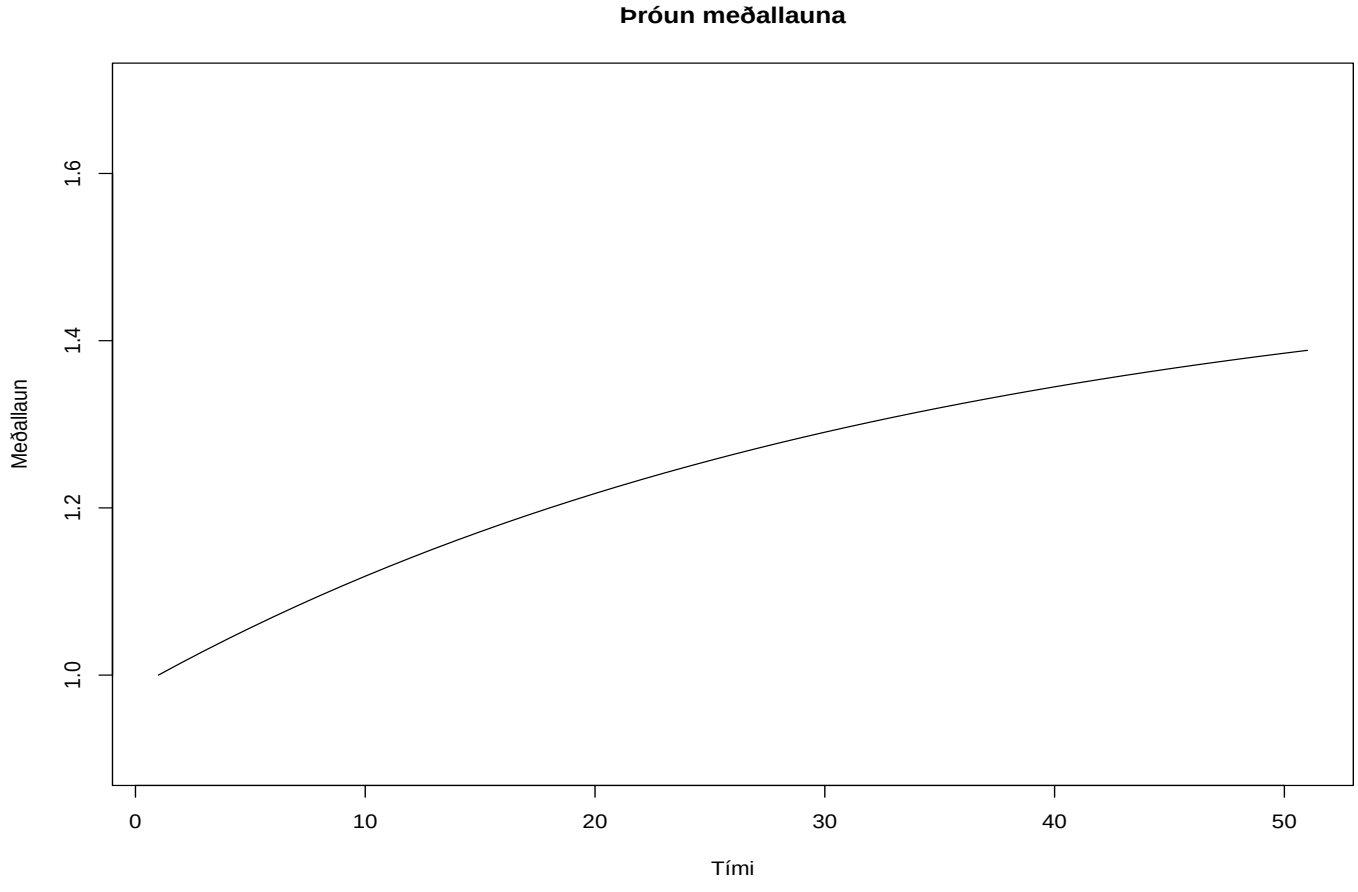


Figure: The expected average of $L(t) = 1 + X(t)$ where $dX = \kappa(\alpha - X)dt + \sigma X dW$.
$\kappa = 0.03$, $\alpha = 0.5$, $\sigma = 0.3$

This is actually a very wild process

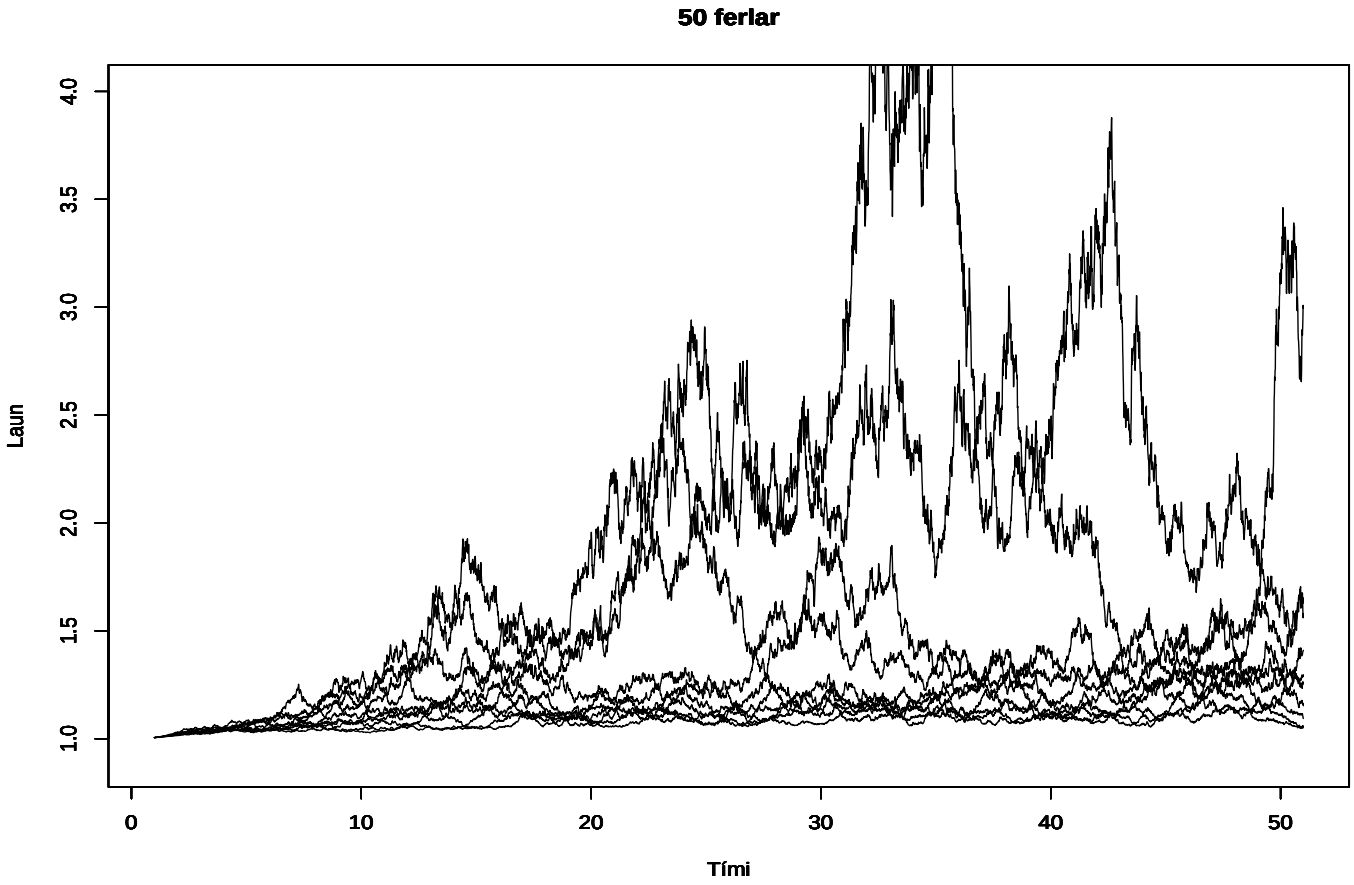


Figure: 50 simulated paths of $L(t) = 1 + X(t)$ where $dX = \kappa(\alpha - X)dt + \sigma X dW$.
$\kappa = 0.03$, $\alpha = 0.5$, $\sigma = 0.3$

Summary, the goal of the conference, future research

-

Summary, the goal of the conference, future research

- Many open problems both academic and technical

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed.

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed. Which categories of stochastic models to use? Developing sensible models is a difficult task. I shall introduce one in a seminar March 5th.

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed. Which categories of stochastic models to use? Developing sensible models is a difficult task. I shall introduce one in a seminar March 5th. !! probably a very stupid one !!

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed. Which categories of stochastic models to use? Developing sensible models is a difficult task. I shall introduce one in a seminar March 5th. !! probably a very stupid one !!
- Issues: Modeling unobservable components, adding extra observed variables

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed. Which categories of stochastic models to use? Developing sensible models is a difficult task. I shall introduce one in a seminar March 5th. !! probably a very stupid one !!
- Issues: Modeling unobservable components, adding extra observed variables
- Organizing data is a huge task

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed. Which categories of stochastic models to use? Developing sensible models is a difficult task. I shall introduce one in a seminar March 5th. !! probably a very stupid one !!
- Issues: Modeling unobservable components, adding extra observed variables
- Organizing data is a huge task
- We thank Henning Bunzel and Lennart Flood for giving us an insight in to the modeling of individual data in labour market research.

Summary, the goal of the conference, future research

- Many open problems both academic and technical
- What income/wage definition to use? Cohorts are of different size, A demographic component of a model is often needed. Which categories of stochastic models to use? Developing sensible models is a difficult task. I shall introduce one in a seminar March 5th. !! probably a very stupid one !!
- Issues: Modeling unobservable components, adding extra observed variables
- Organizing data is a huge task
- We thank Henning Bunzel and Lennart Flood for giving us an insight in to the modeling of individual data in labour market research.
- We hope to see you soon again in Iceland.