International Long-run Growth Dynamics

(work in progress)

Ulrich Müller, Jim Stock, Mark Watson

Central Bank of Chile, October 2018

Original motivation for work

Long-horizon predictive distributions for global GDP/Population as an input into determining the "Social Cost of Carbon" (SCC) from CO_2 emissions.

(SCC is used by regulators and others)

Reference: *NAS* (2017)

Damages are long-lived $\Rightarrow$ Predictive distributions over 100, 200, or more years.

Damages depend on location $\Rightarrow$ Joint predictive distributions for many countries.

Develop a statistical model for joint long-run dynamics for many countries

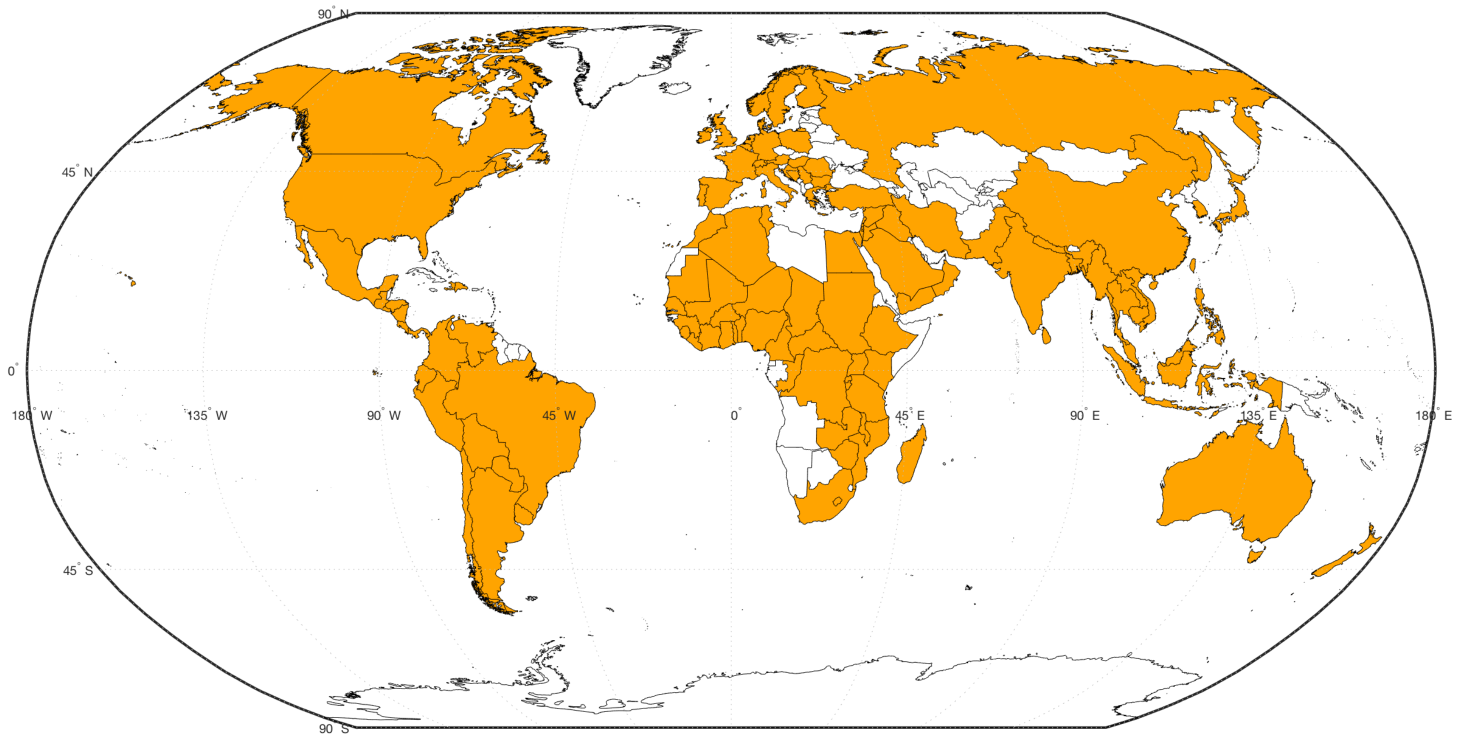
Useful for:

(1) Reduced form description of cross-country long-run growth dynamics (convergence, persistence of development gaps, etc.)

(2) Long-run international probabilistic forecasts (original motivation)

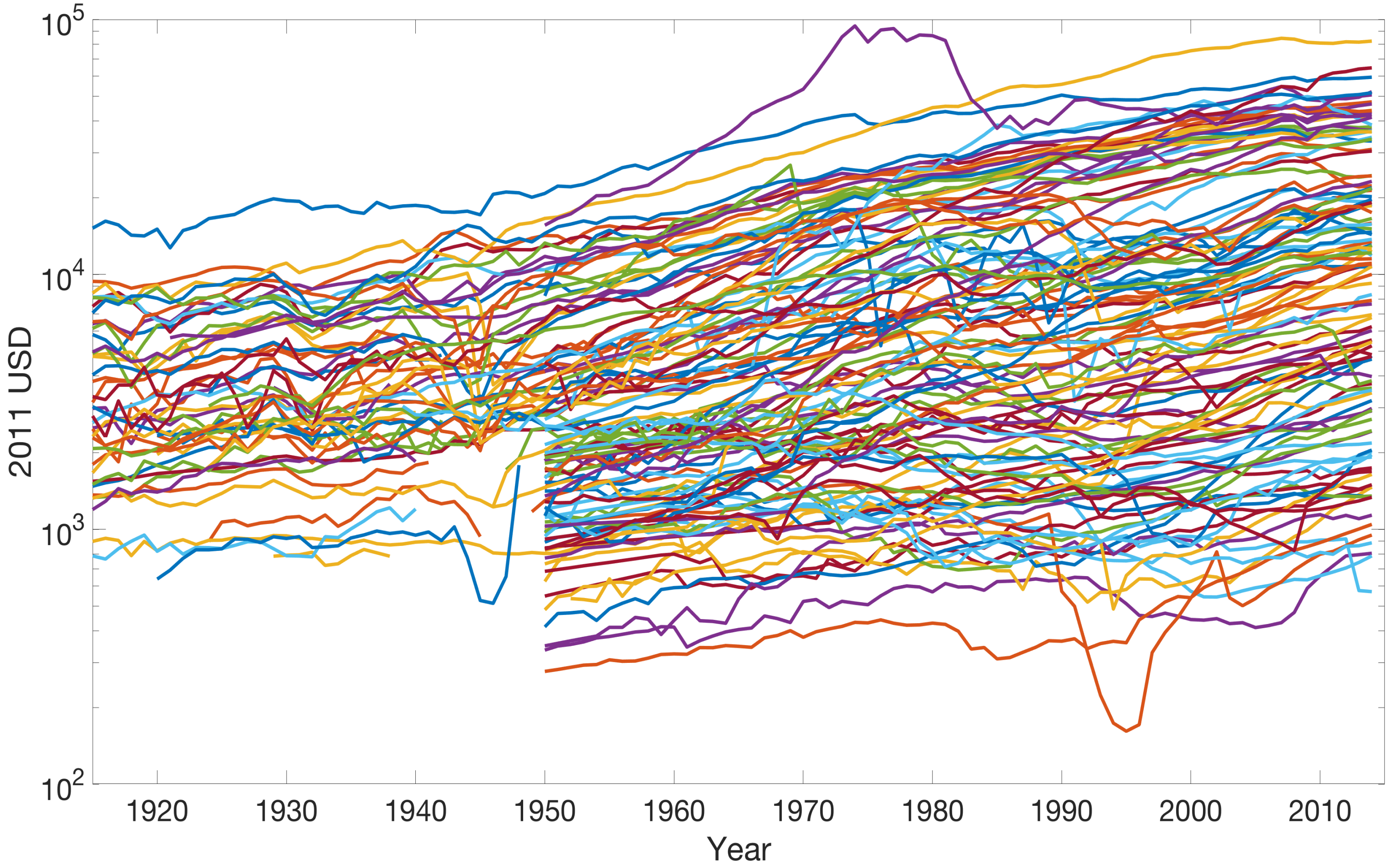
Data: Annual 1915-2014 for 112 countries

(Merged: PWT 1950-2014 and Maddison 1915-1949
countries with at least 50 years of post-1949 data and population > 3 million)

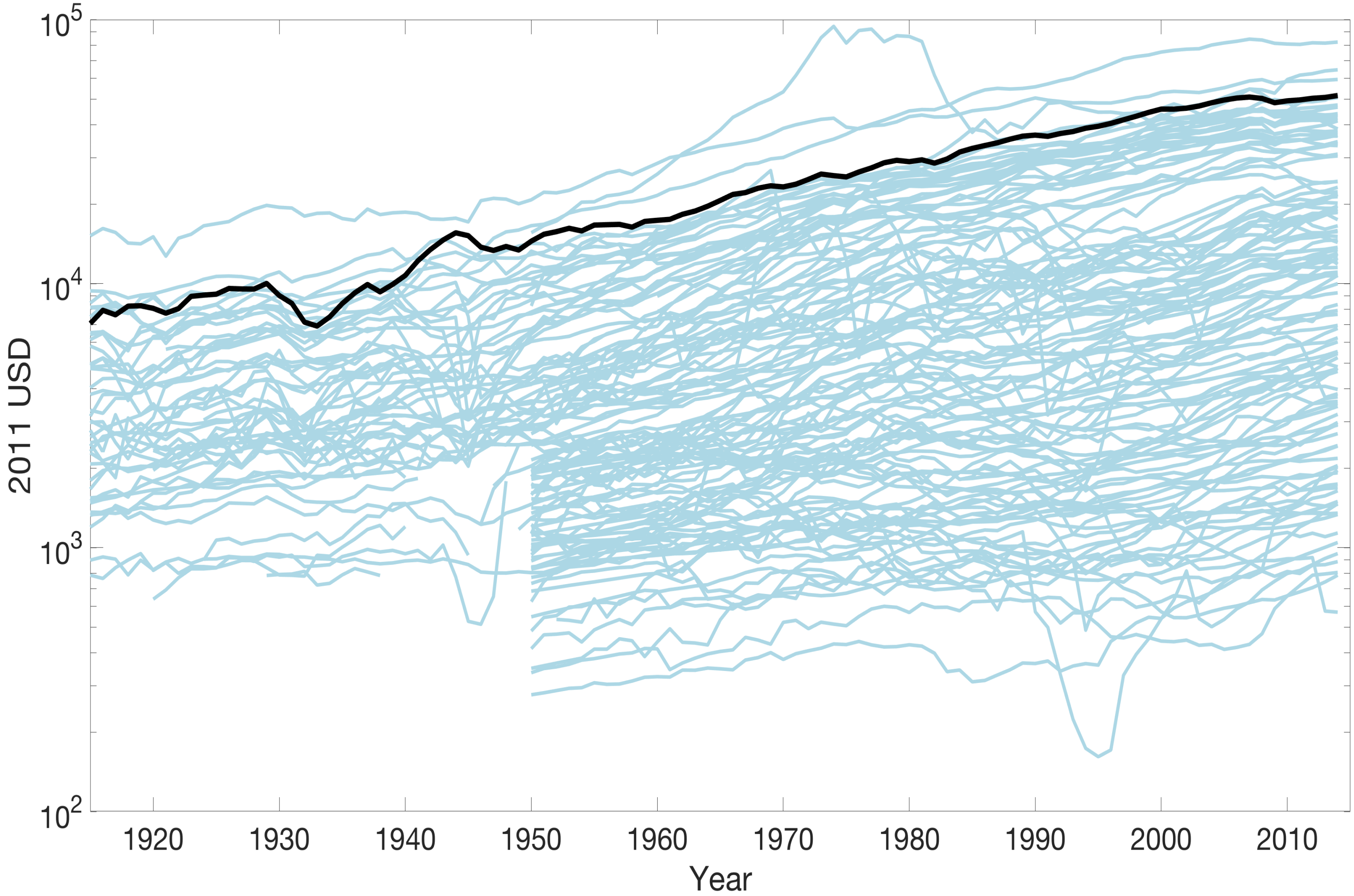


- 97% of World GDP in 2014 and 96% of World Population
- Unbalanced Panel (39-52 countries before 1950, 107 in 1950, 110 in 1952 and 112 in 1960)

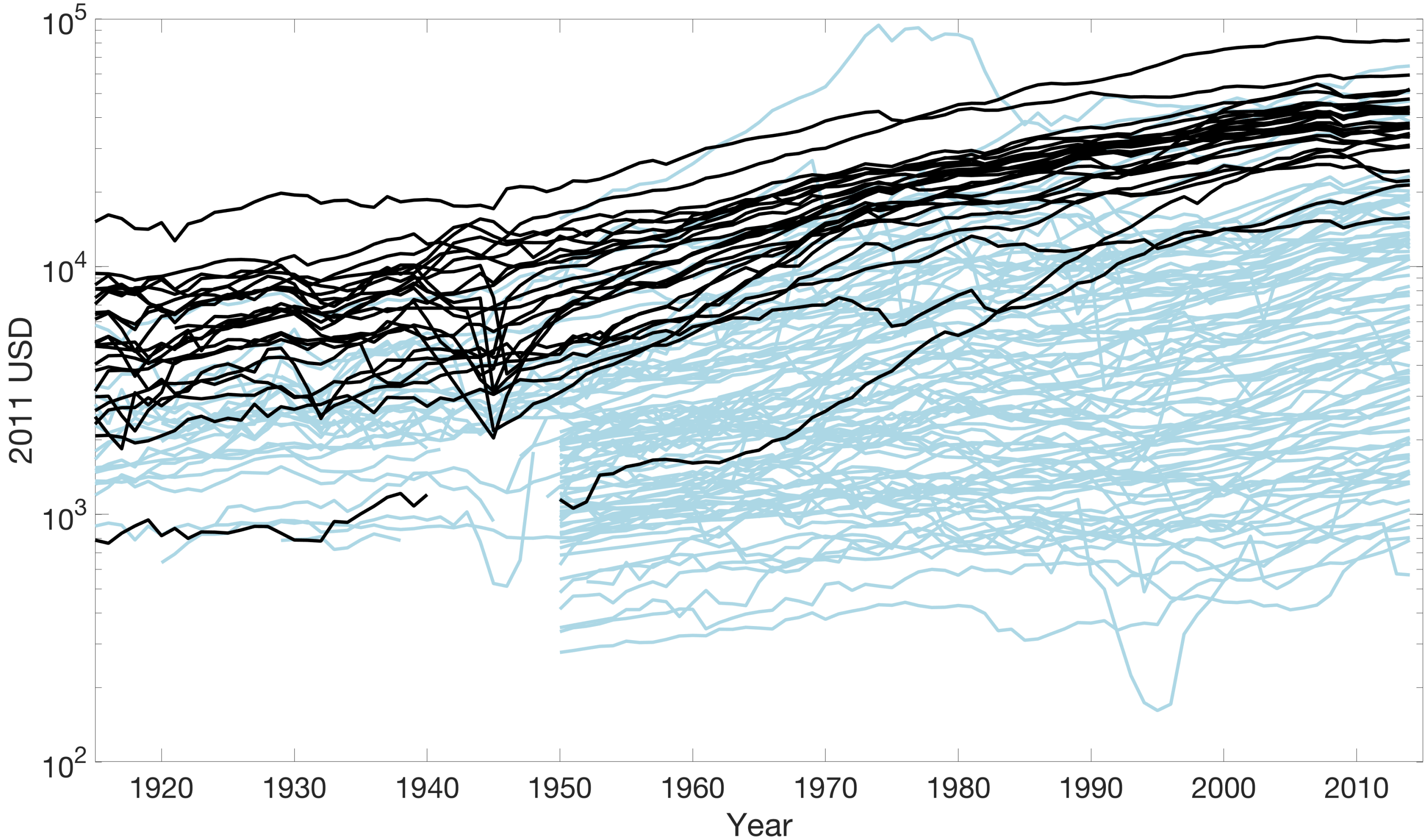
Data: GDP/Population for 112 countries



United States

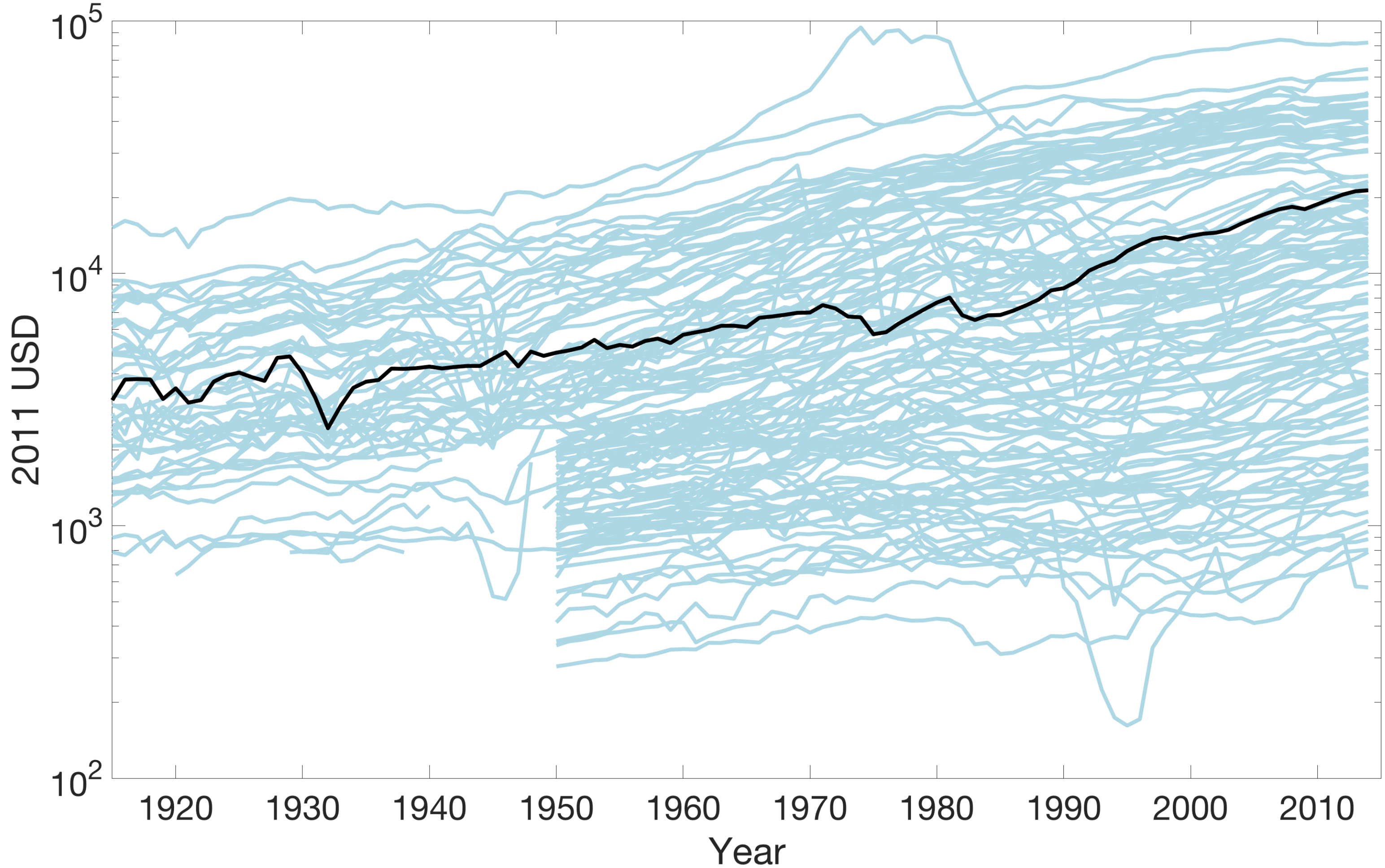


OECD



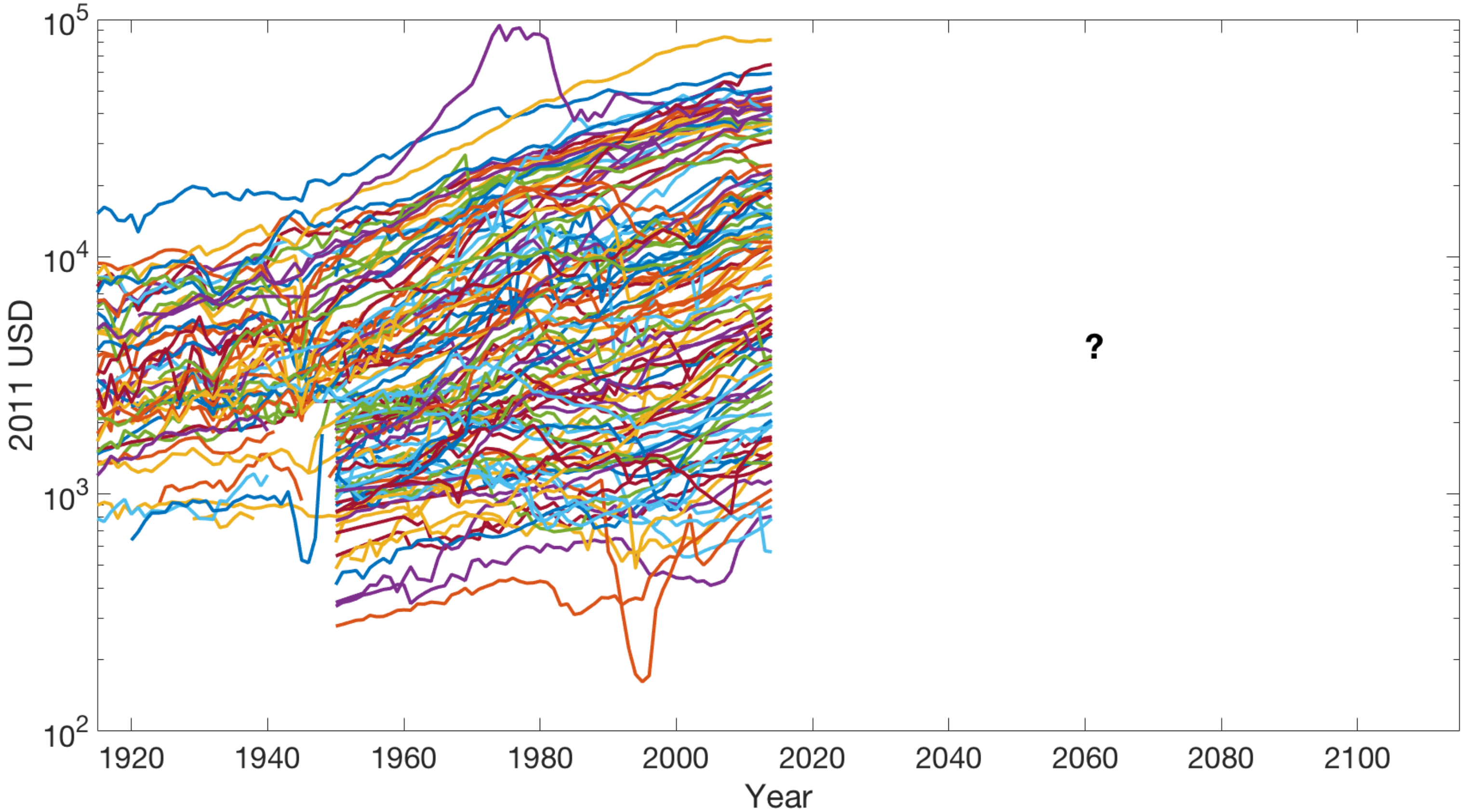
10^5
10^4
10^3
10^2
2011 USD
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
Year

Chile

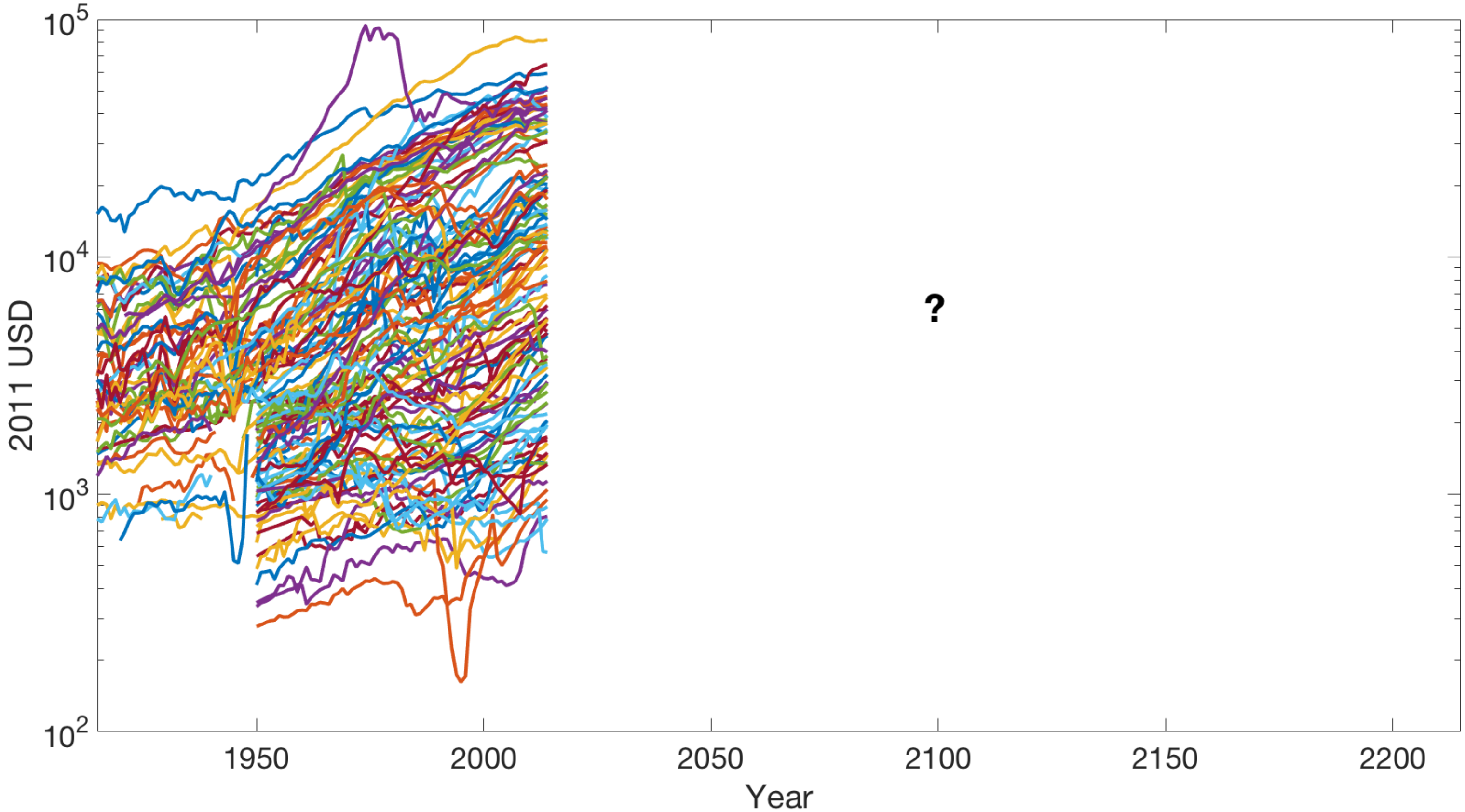


10^5
10^4
10^3
10^2
2011 USD
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
Year

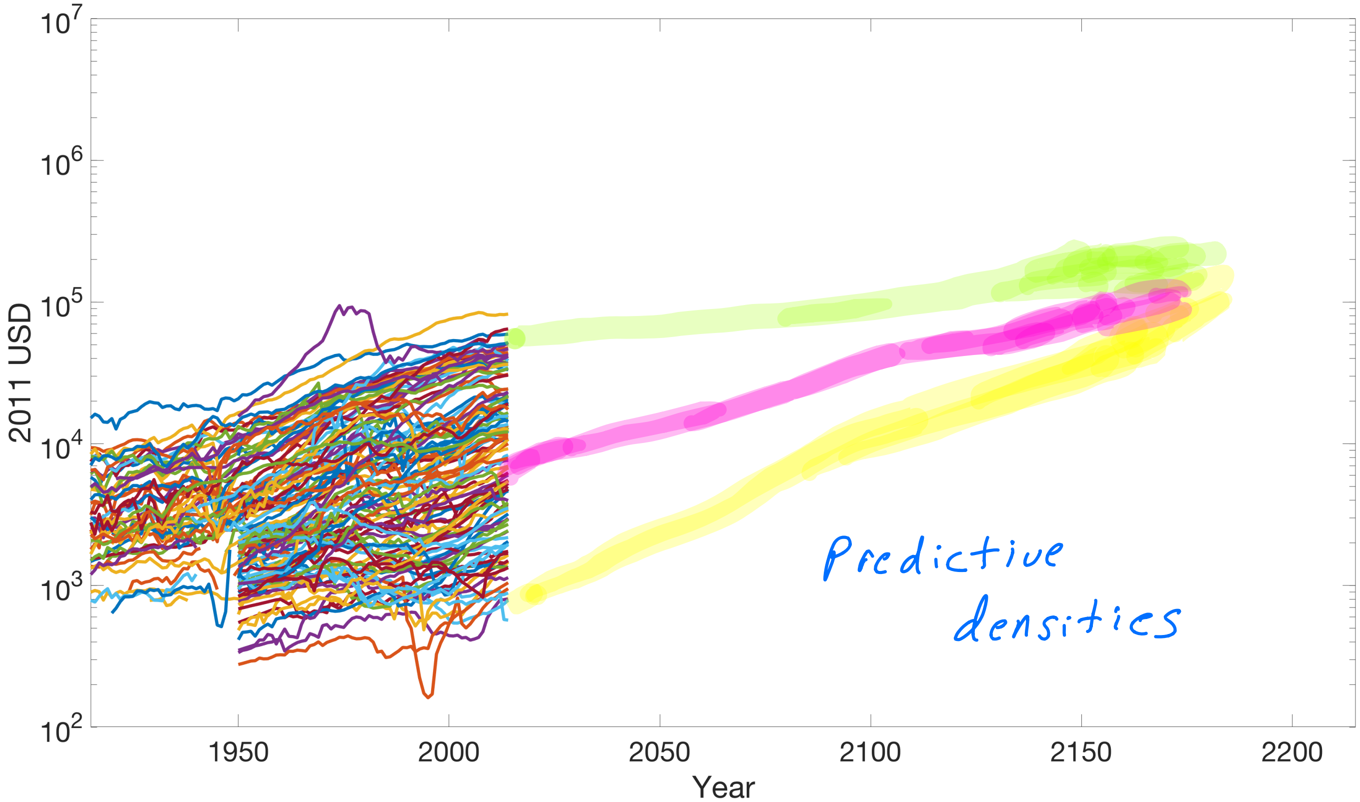
Long-Run Forecasting Problem

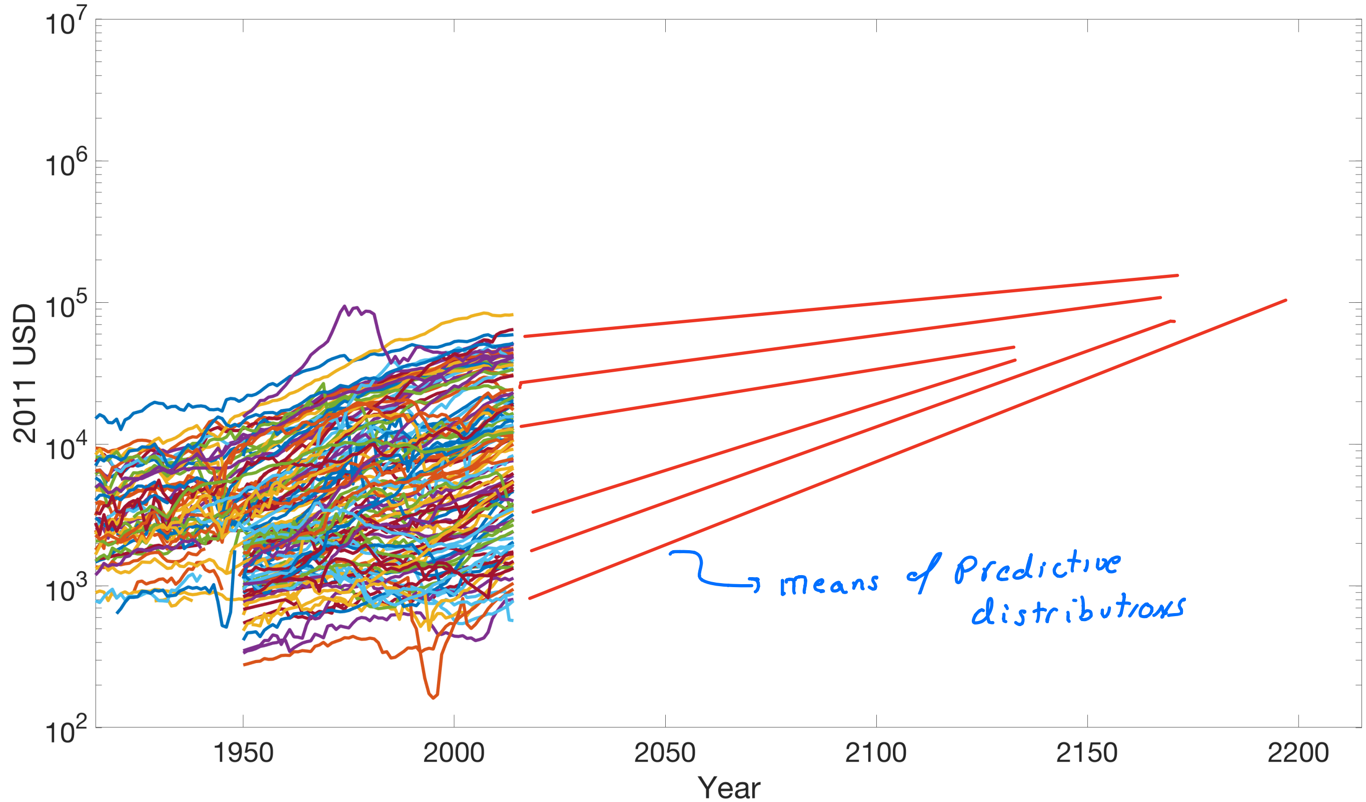


or

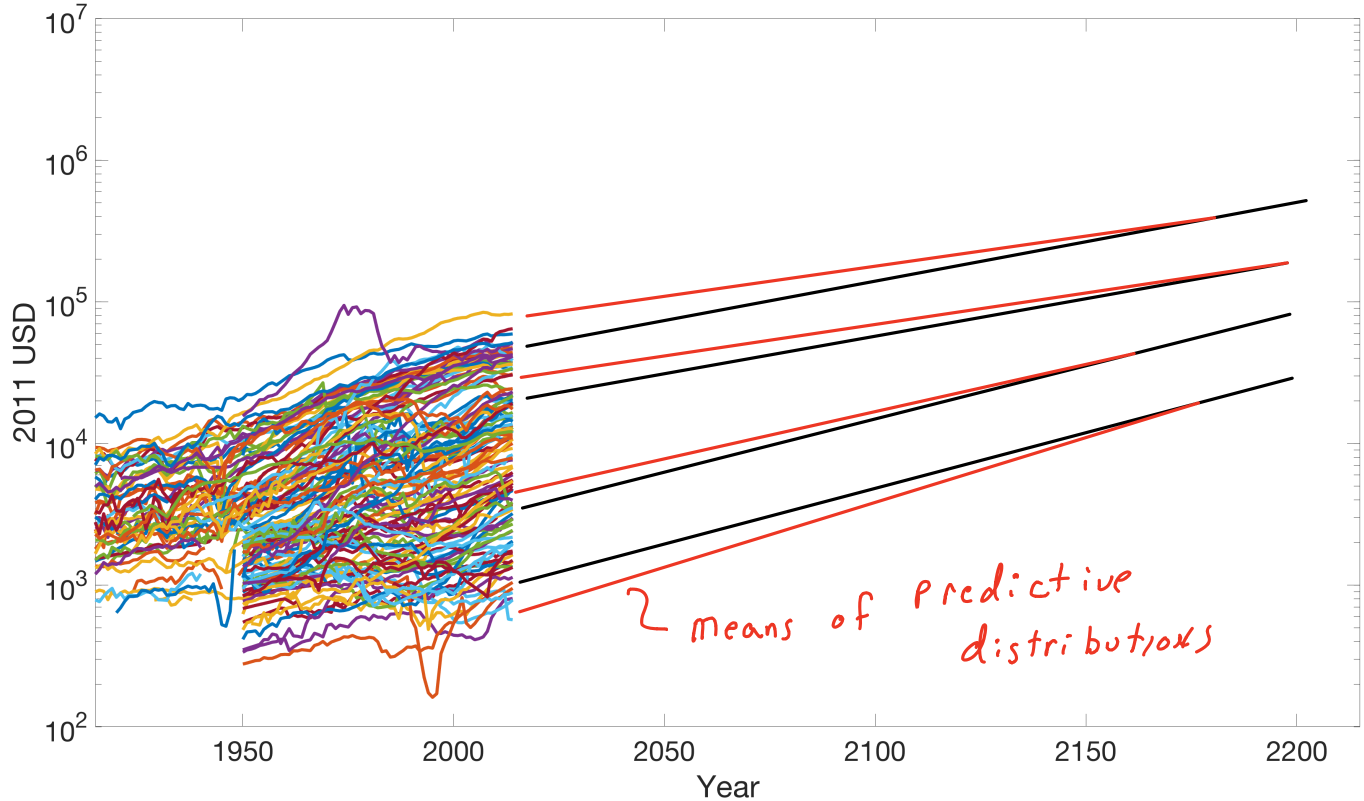


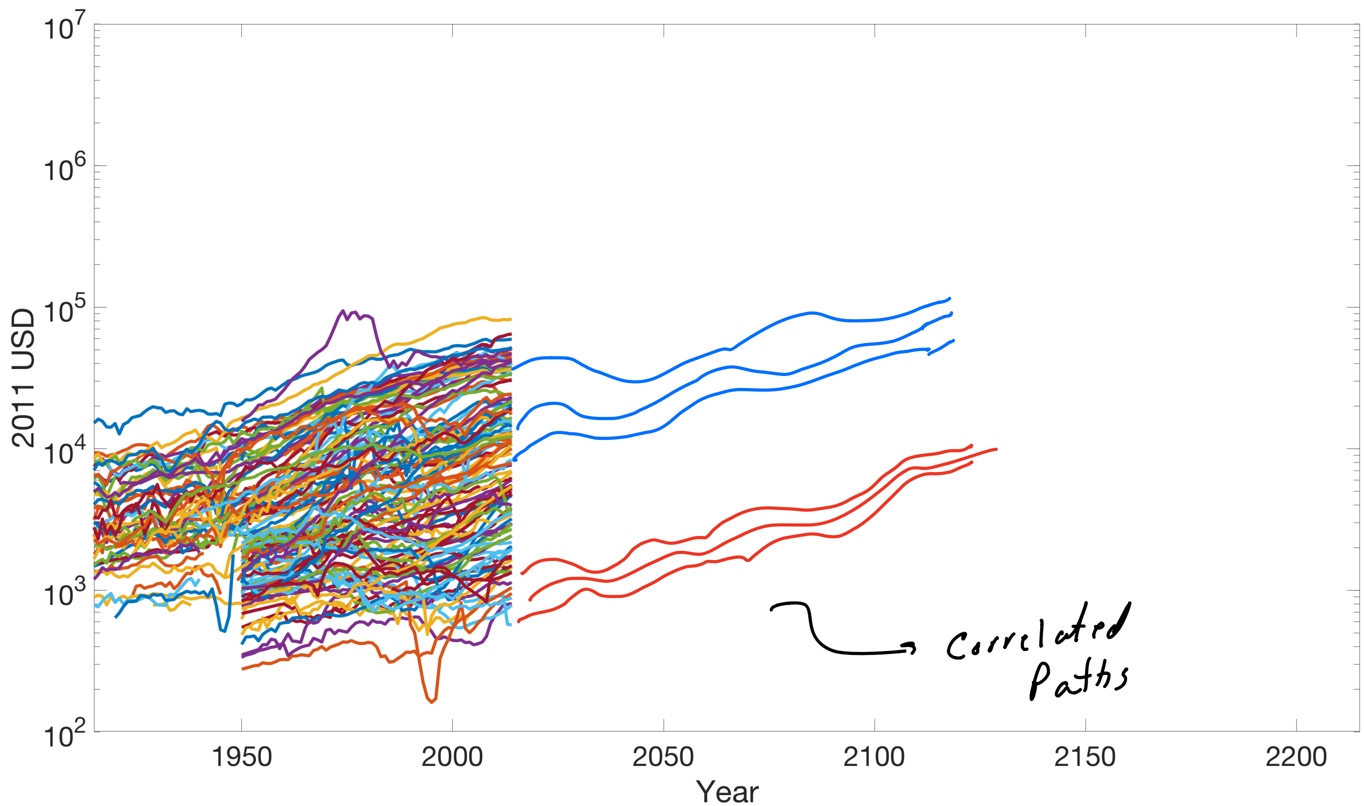
Convergence, persistence and comovement





10^7
10^6
10^5
10^4
10^3
10^2
2011 USD
1950
2000
2050
2100
2150
2200
Year
means of predictive distributions


10^7
10^6
10^5
10^4
10^3
10^2
2011 USD
1950
2000
2050
2100
2150
2200
Year
2 means of predictive
distributions



10^7
10^6
10^5
10^4
10^3
10^2
2011 USD
1950
2000
2050
2100
2150
2200
Year
Correlated Paths

Outline:

1. Look at the data to determine sensible features of a model.

2. Simplification: focus on 'long-run' variation/covariation.

3. Detailed description of model.

4. Estimation mechanics

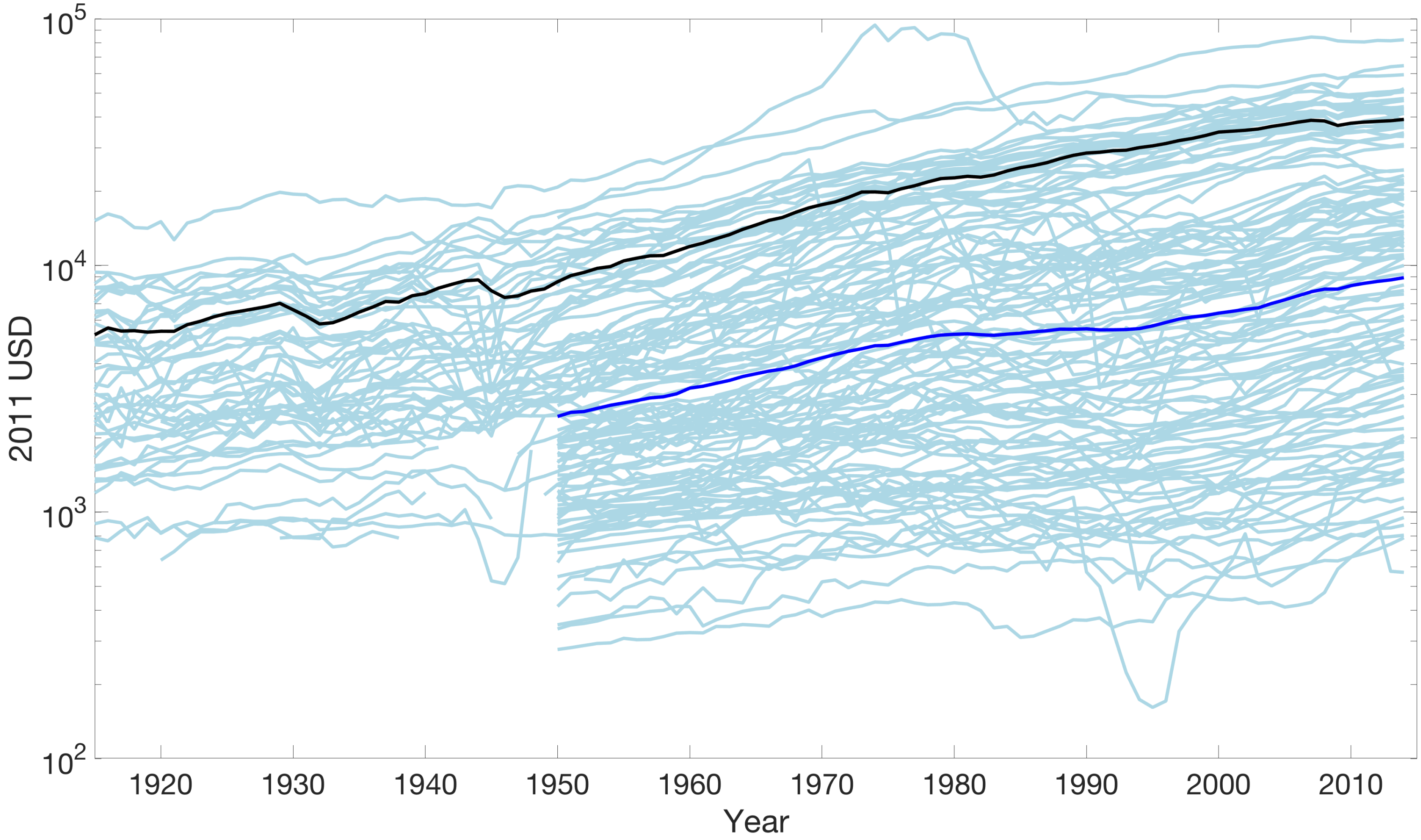
5. Results
   a. Convergence
   b. Long-run predictions

6. Different modelling choices

Notation: $\boldsymbol{Y_{it}}$ *= per-capita GDP for country i in year t.*

4 Features of the data and implications for modelling

Feature 1: "Common" Growth Factor



OECD and Average all countries

Model: $y_{it} = \ln(Y_{it})$

$$y_{it} = f_t + c_{it}$$

common global growth factor   country i factor

Feature 2: No reduction in cross-sectional dispersion

Medians, IQR and 90-10 range for histograms of y_{it}

Average value over	median	75th-25th	90th-10th
1950 - 1954	7.8	1.5	2.6
1960 - 1964	7.9	1.6	2.7
1985 - 1989	8.6	2.2	3.3
2010 - 2014	9.3	2.1	3.4

Model:

$$y_{it} = f_t + c_{it}$$

(long-run) variance of c_{it} is constant

(examined in more detail in *Different modeling choices* below)

Feature 3: Substantial persistence in cross section

Averages of y_{it} over 25+ year periods: Probability of moving from quartile i (1960-1987) to quartile j (1988-2014)

		Quartile in 1988-2014			
		1	2	3	4
Quartile in 1960-1987	1	0.786	0.214	0	0
	2	0.214	0.643	0.107	0.036
	3	0	0.143	0.714	0.143
	4	0	0	0.179	0.821

- Country in Q1; years until Prob(Q3 + Q4) > 0.25 $\approx$ 220 years
- Country in Q4: years until Prob(Q1 + Q2) > 0.25 $\approx$ 80 years

- Kremer, Onatski, Stock (2001) using 5 year transitions of relative income levels: Half-life = 285 years (Related: Quah (1993), Jones (1997, 2016))

Model:

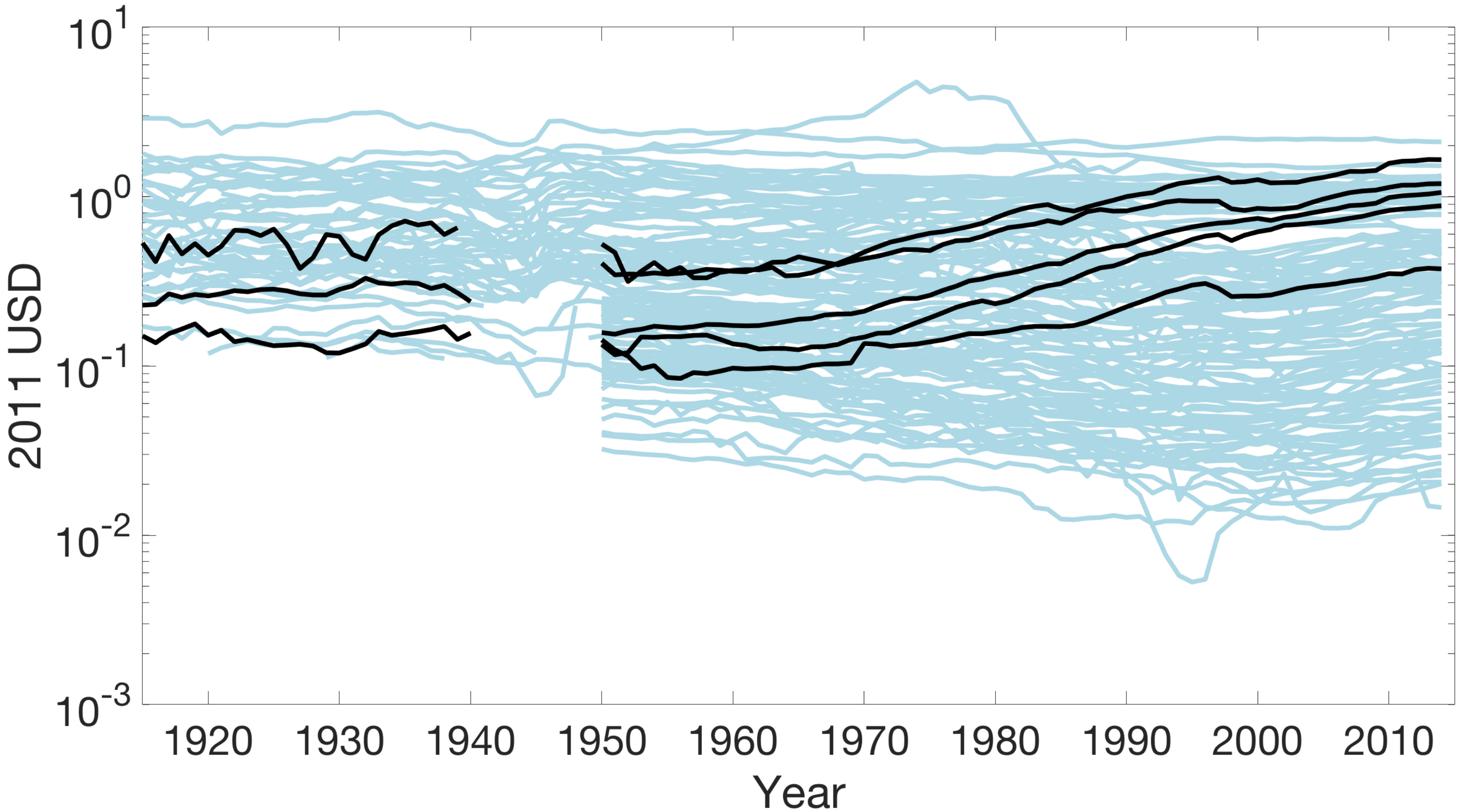
$$y_{it} = f_t + c_{it}$$

(long-run) variance of c_{it} is constant

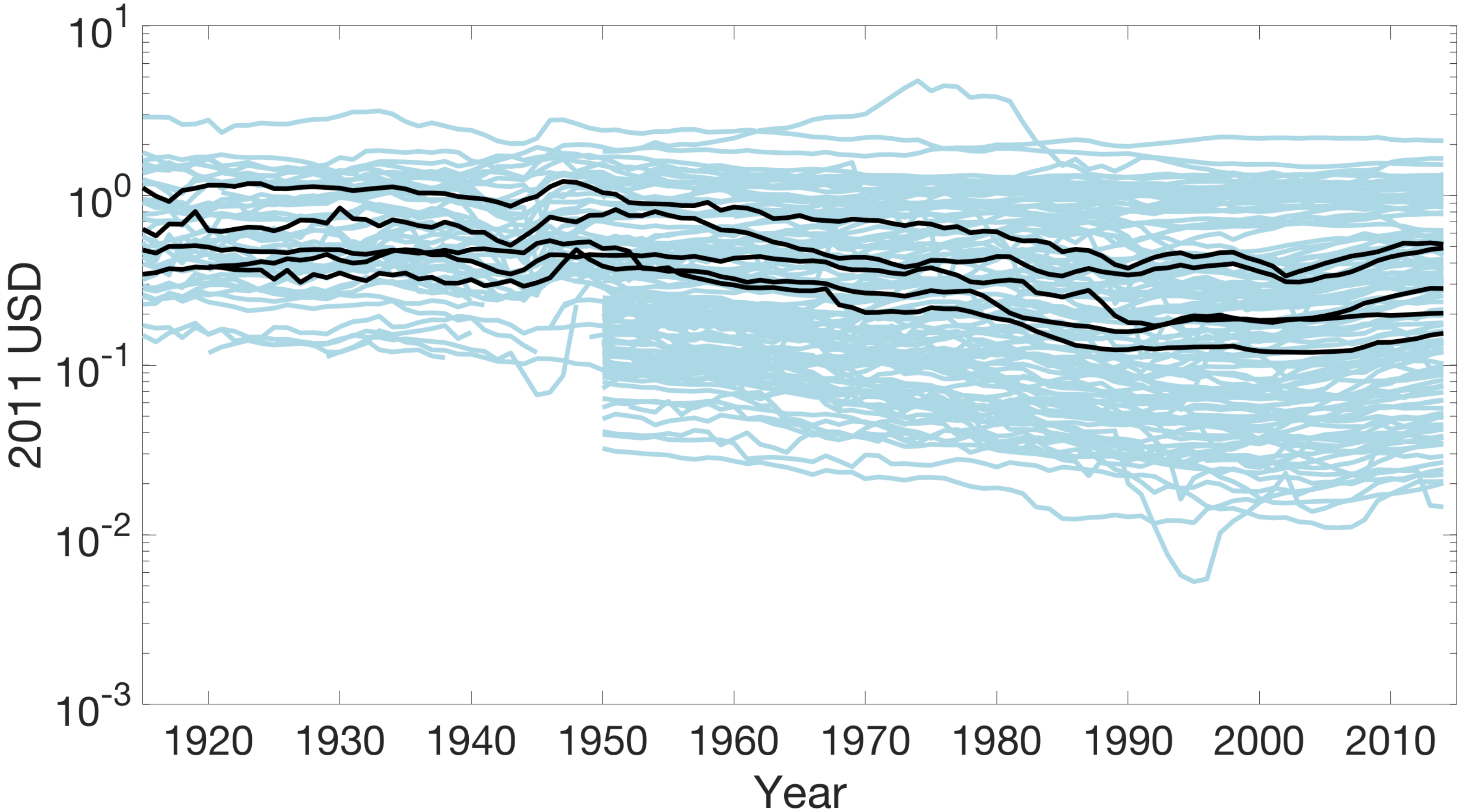
c_{it} is <u>very</u> persistent (but stationary)

Feature 4: Comovement of y_{it} within cross section

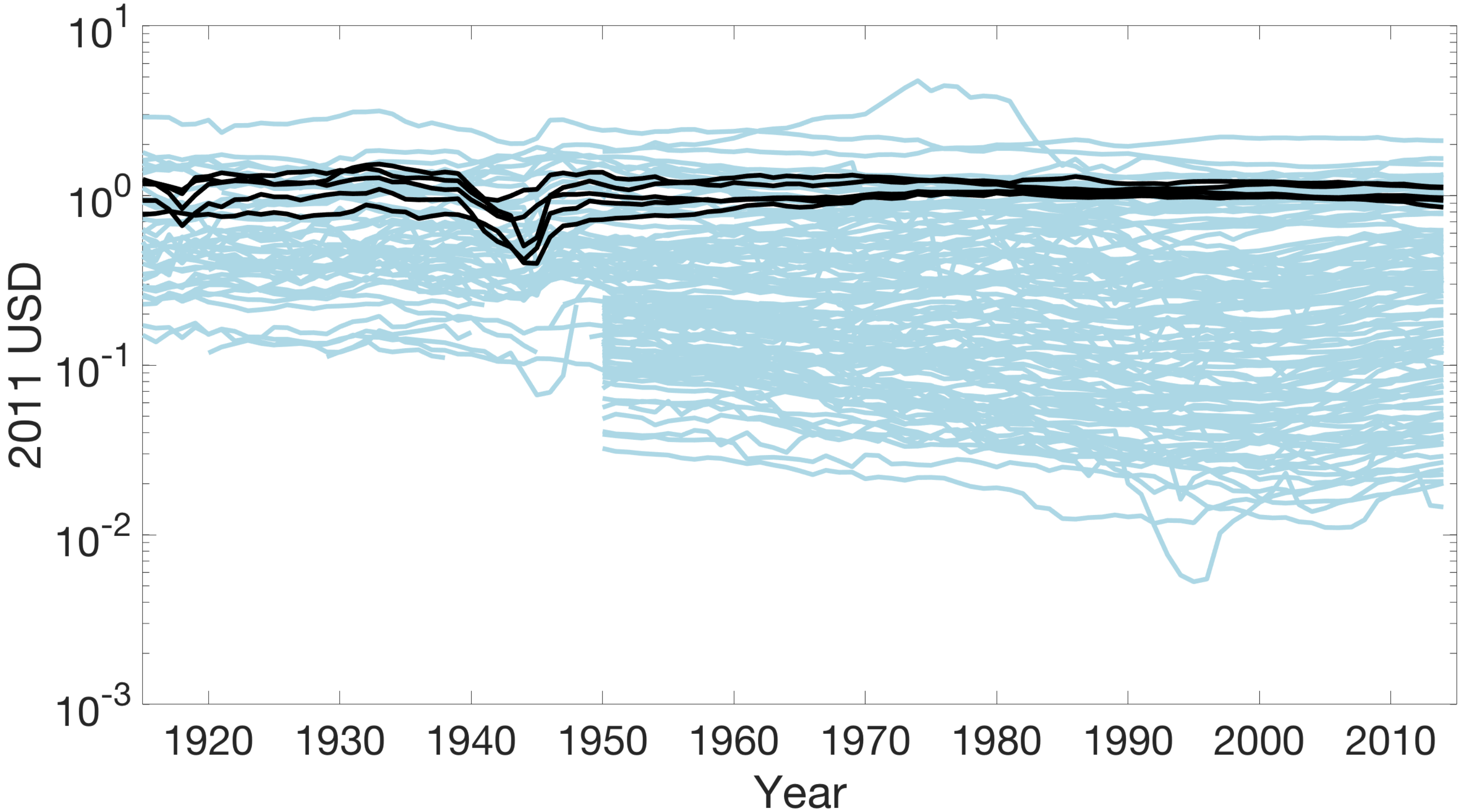
Examples:
(a) Hong Kong, South Korea, Singapore, Taiwan, Thailand



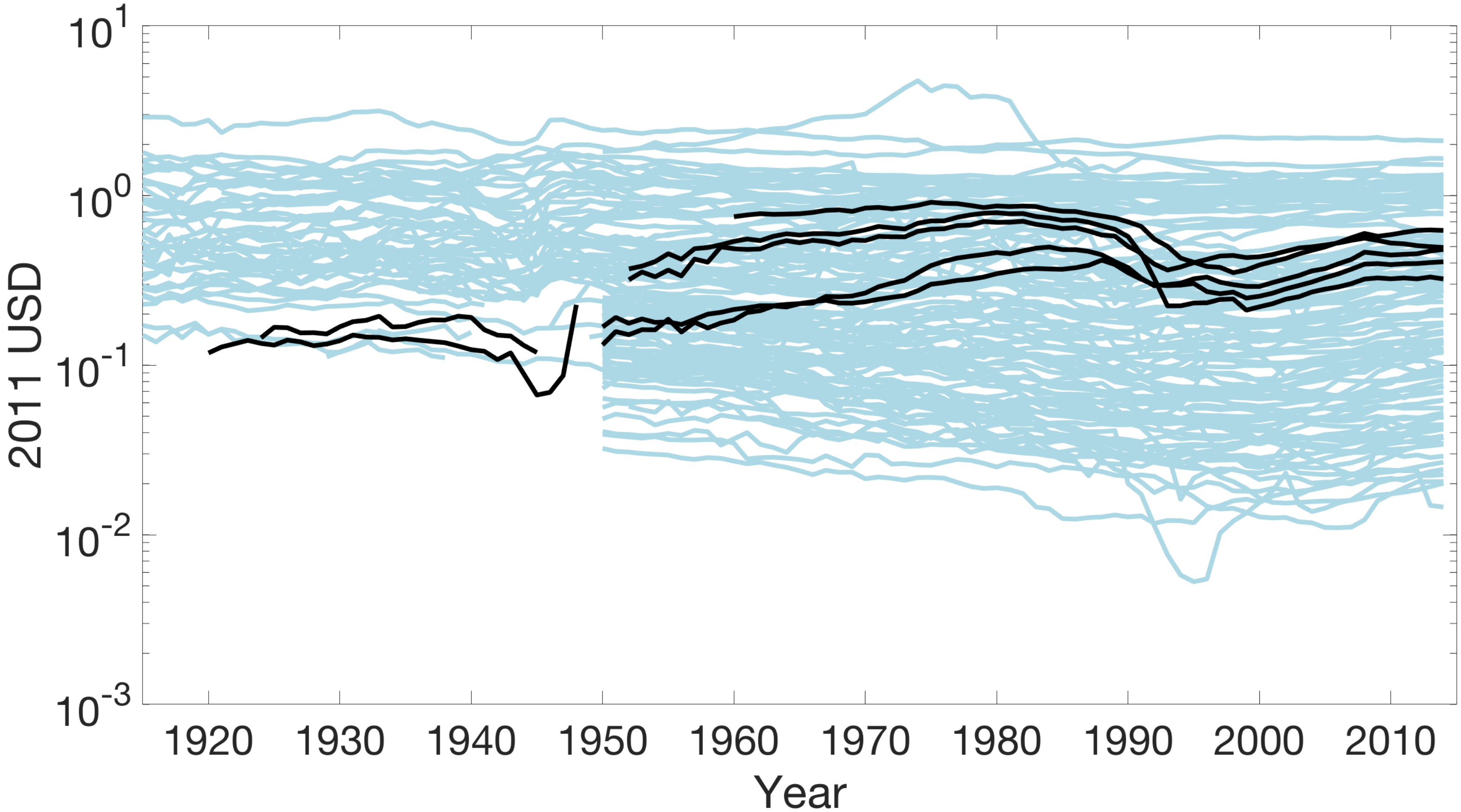
(b) Argentina, Bolivia, El Salvador, Uruguay, Peru



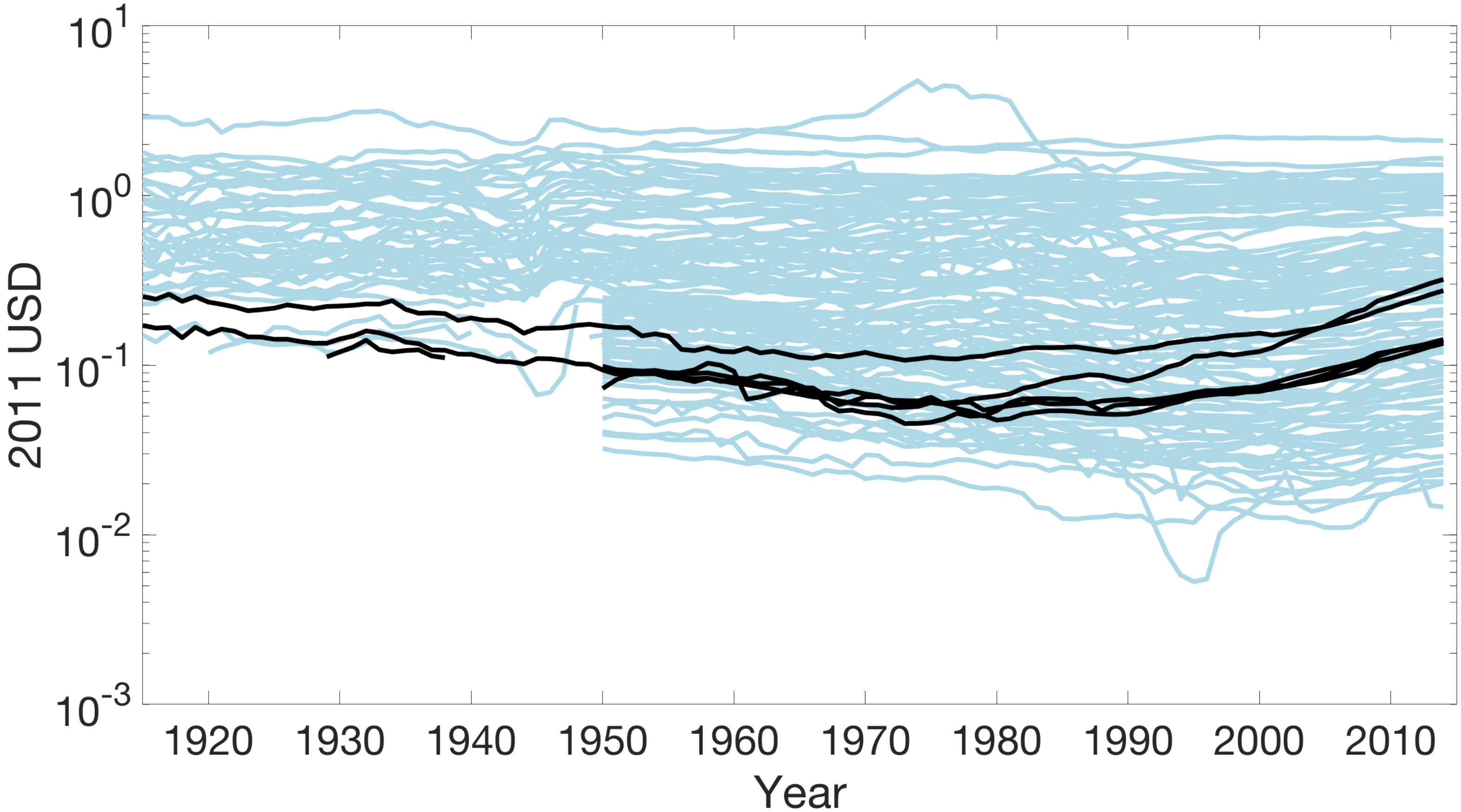
(c) Belgium, France, Italy, Netherlands, Denmark



(d) Bulgaria, Croatia, Russia, Serbia, Romania



(e) China, India, Laos, Sri Lanka, Vietnam



Model:

$$y_{it} = f_t + c_{it}$$

(long-run) variance of c_{it} is constant

c_{it} is very persistent (but stationary)

c_{it} is correlated within "groups"

Outline:

1. ~~Look at the data to determine some sensible features of a model.~~

2. **Simplify problem: focus on 'long-run' variation/covariation.**

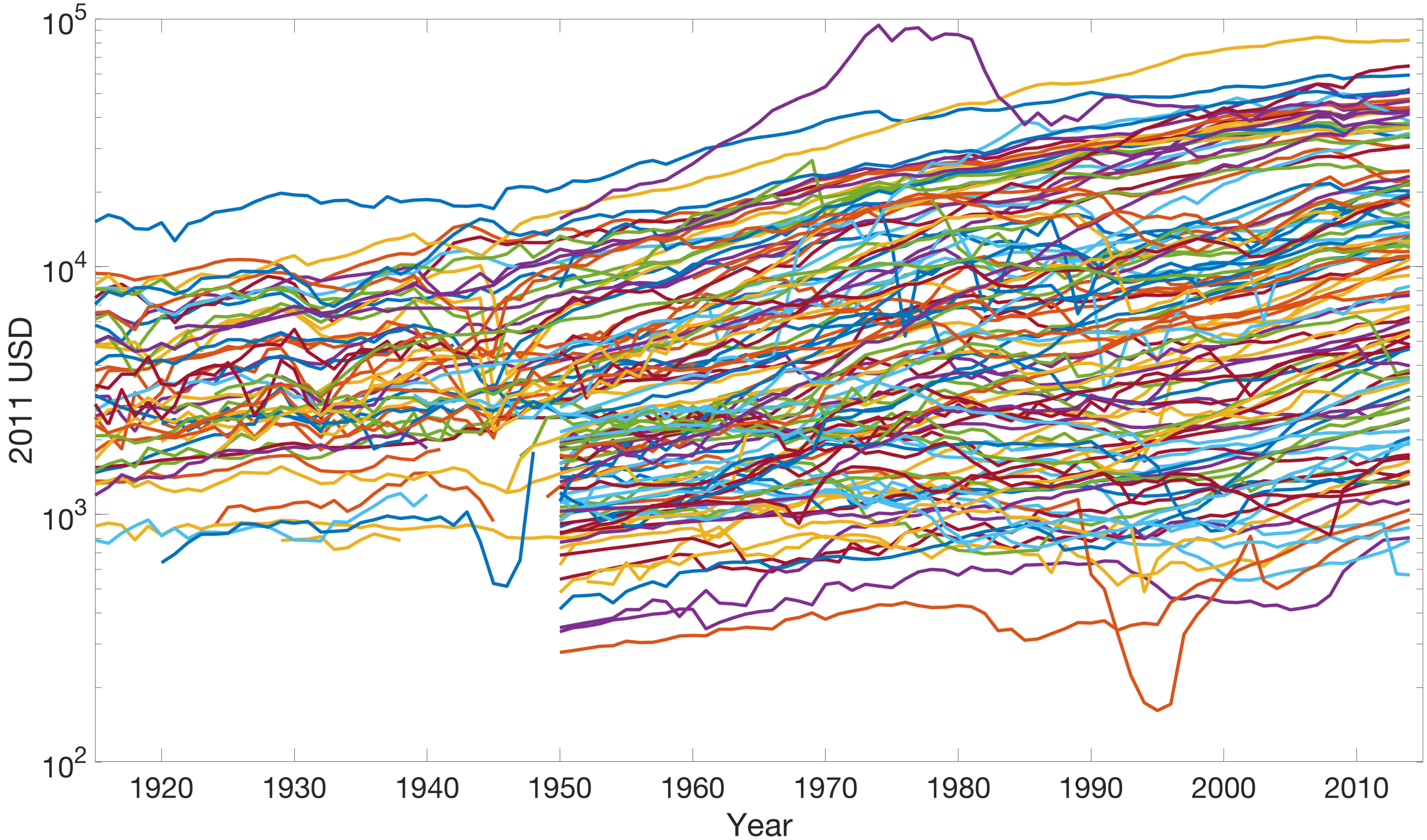
3. Detailed description of model.

4. Estimation mechanics

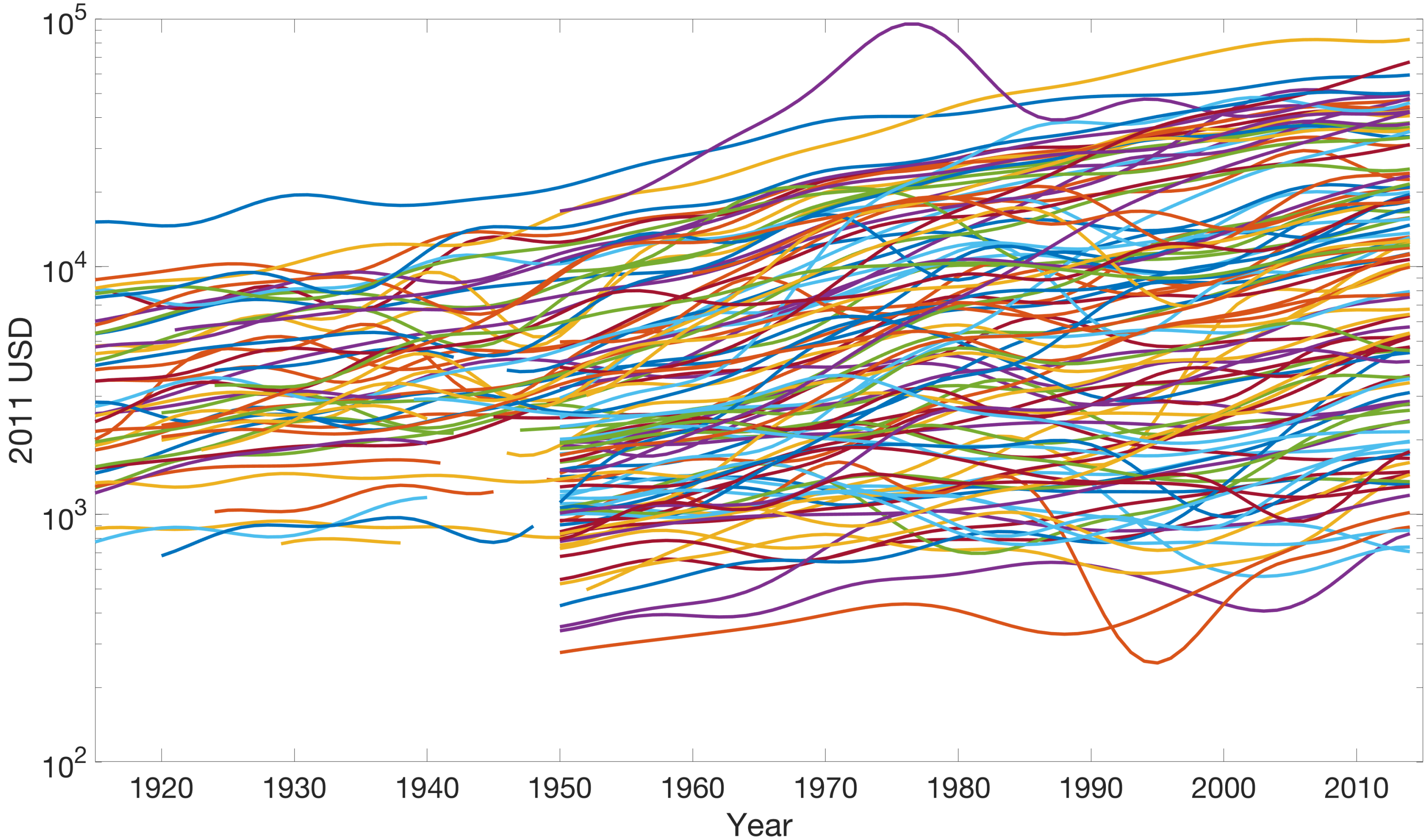
5. Results
   a. Convergence
   b. Long-run predictions

6. Different modelling choices

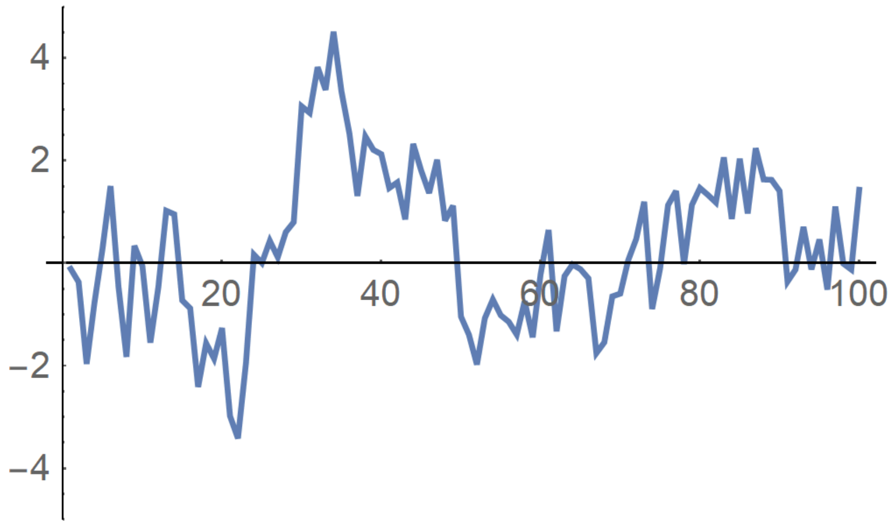
Original Data

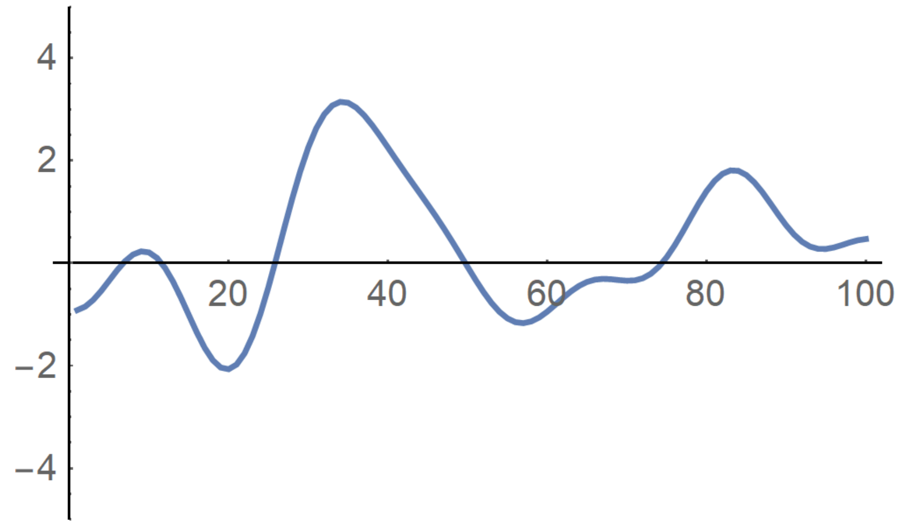


Low-Frequency Transformed Data

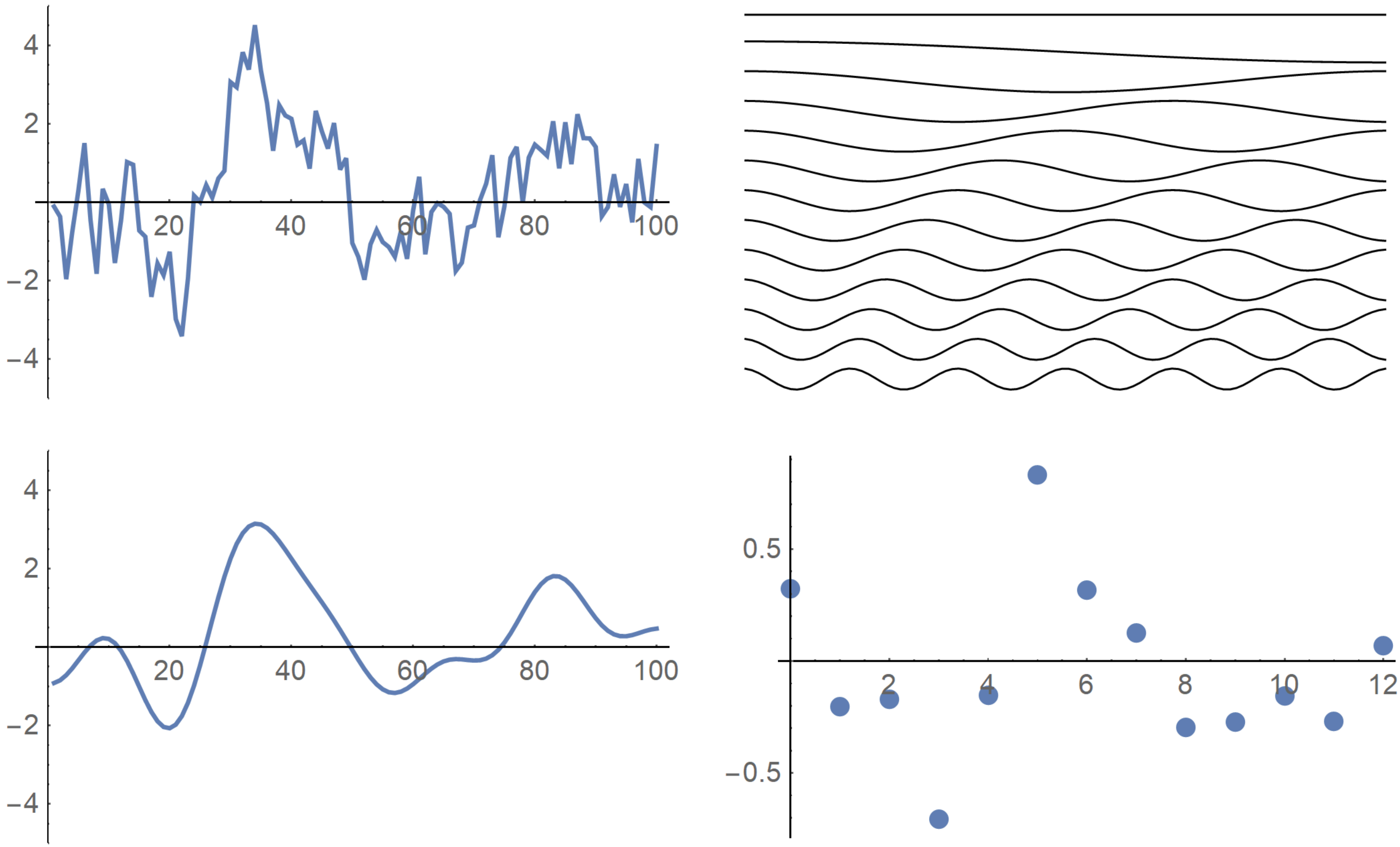


A Simplification: Focus on low-frequency variability in data

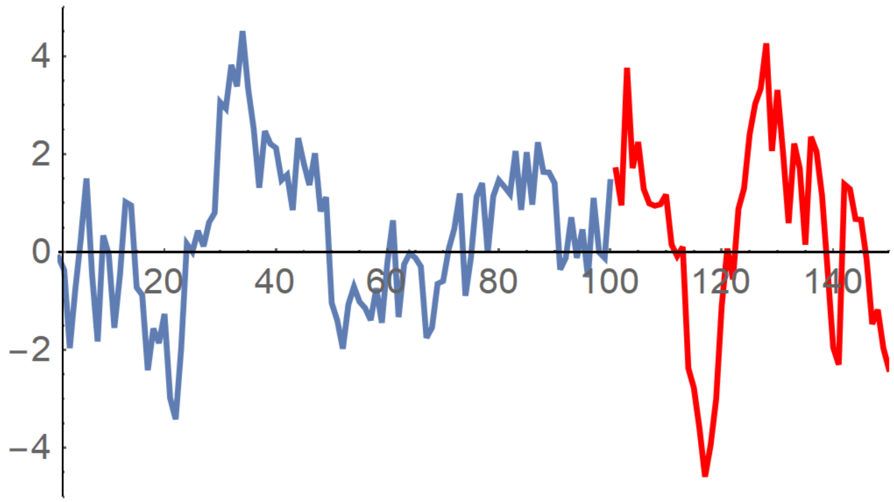


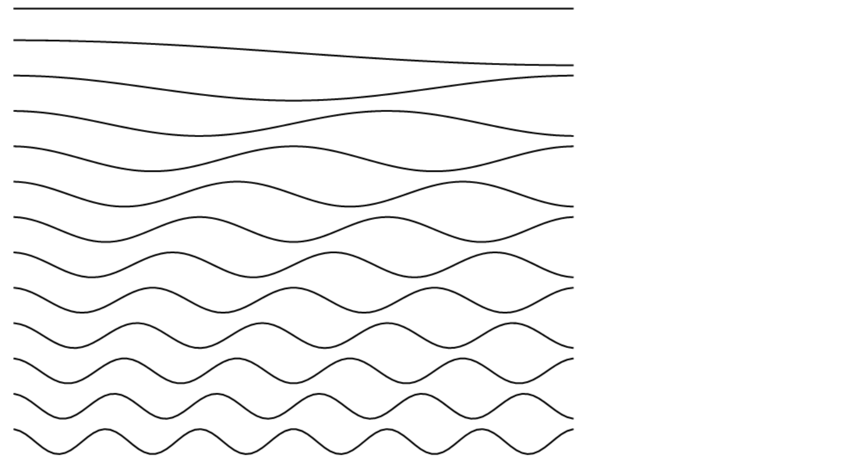


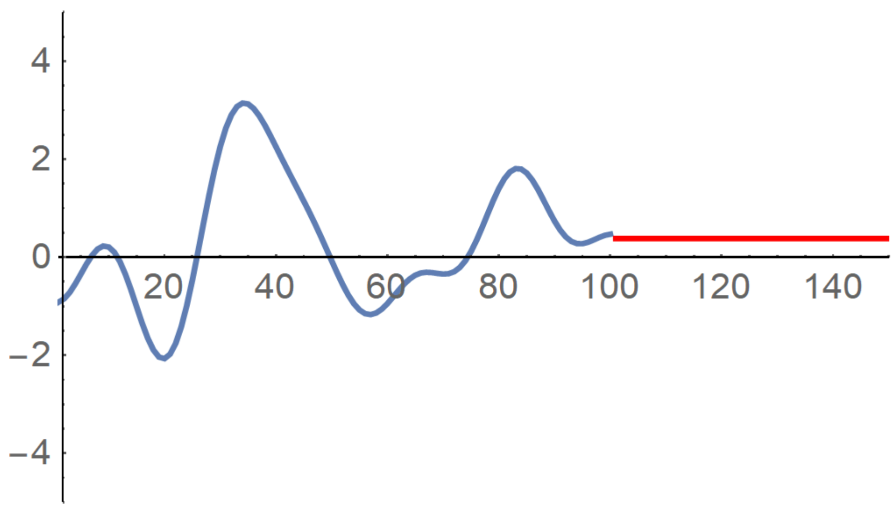
Low-frequency data compression

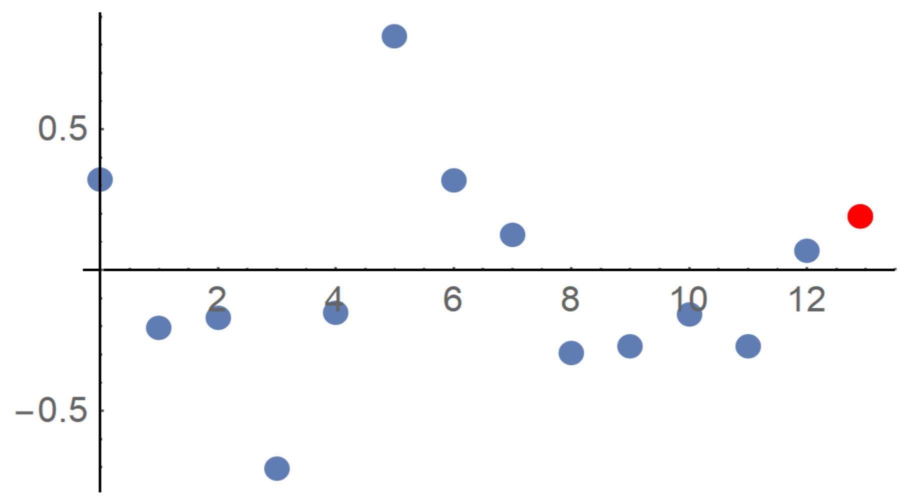


Implications for long-run forecasting:









Simplification: Focus on low-frequency variability in data

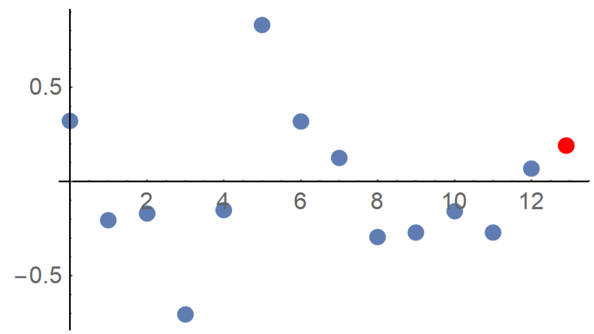
Selected Literature

- I(0): classic time series work on periodogram analysis, band-spectrum regression (Engle (1974)), etc.

- More recent:

    - Müller (2004): HAR/HAC inference ('Student-*t* inference', etc.)

    - Müller and Watson (2008), (2013), (2016), (2018)



Why is this a simplification ?

- Number of observations: (fewer dots than time series observations)

- Dots are "averages" of data $\Rightarrow$ *Normally distributed*
  - o Rationalizes Gaussian likelihood
  - o Prediction of future (red dot) from past (blue dots)

- Modelling: only low-frequency features of model matter

- Inference: $Y \sim \mathrm{N}(0,\Sigma(\theta))$ … inference about parameters of covariance matrix of normal

Outline:

1. ~~Look at the data to determine some sensible features of a model.~~

2. ~~Simplify problem: focus on 'long-run' variation/covariation.~~

3. **Detailed description of model**

4. Estimation mechanics

5. Results
   a. Convergence
   b. Long-run predictions

6. Different modelling choices

Details of model: Cross-country covariation:

$$y_{it} = f_t + c_{it}$$

$$c_{i,t} = \mu + \lambda_{c,i}\, g_{J(i),t} + u_{c,i,t}$$

$$g_{j,t} = \lambda_{g,j}\, h_{K(j),t} + u_{g,j,t}$$

"Clustered" factor model for c_{it} (Frühwirth-Schnatter and Kaufmann (2008), Hamilton and Owyang (2012), etc.) with added hierarchical structure.

Details of model: **Cross-country covariation**

$$y_{it} = f_t + c_{it}$$

$$c_{it} = \mu + \lambda_{c,i}\, \mathbf{g}_{\mathbf{J(i),t}} + u_{c,i,t}$$

- $\mathbf{g}_{\mathbf{J(i),t}}$ is a "group factor"

- Each country is a member of 1 group

$$g_{j,t} = \lambda_{g,j}\, h_{K(j),t} + u_{g,j,t}$$

Details of model: **Cross-country covariation**

$$y_{it} = f_t + c_{it}$$

$$c_{it} = \mu + \lambda_{c,i}\, \boldsymbol{g_{J(i),t}} + u_{c,i,t}$$

- $\boldsymbol{g_{J(i),t}}$ is a "group factor"
- Each country is a member of 1 group

$$\boldsymbol{g_{j,t}} = \lambda_{g,j}\, \boldsymbol{h_{K(j),t}} + u_{g,j,t}$$

- Correlation across groups
- $\boldsymbol{h_{K(j),t}}$ is a "group-of-group factor"
- Each group is a member of 1 group-of-group

Details of model: **Temporal Covariation**

(Note: Only low-frequency characteristics of model matter.)

$$y_{it} = f_t + c_{it}$$

$$f_t = f_0 + \underbrace{m_t}_{\text{local growth rate}} \times t + \underbrace{a_t}_{\text{deviation from local trend}}$$

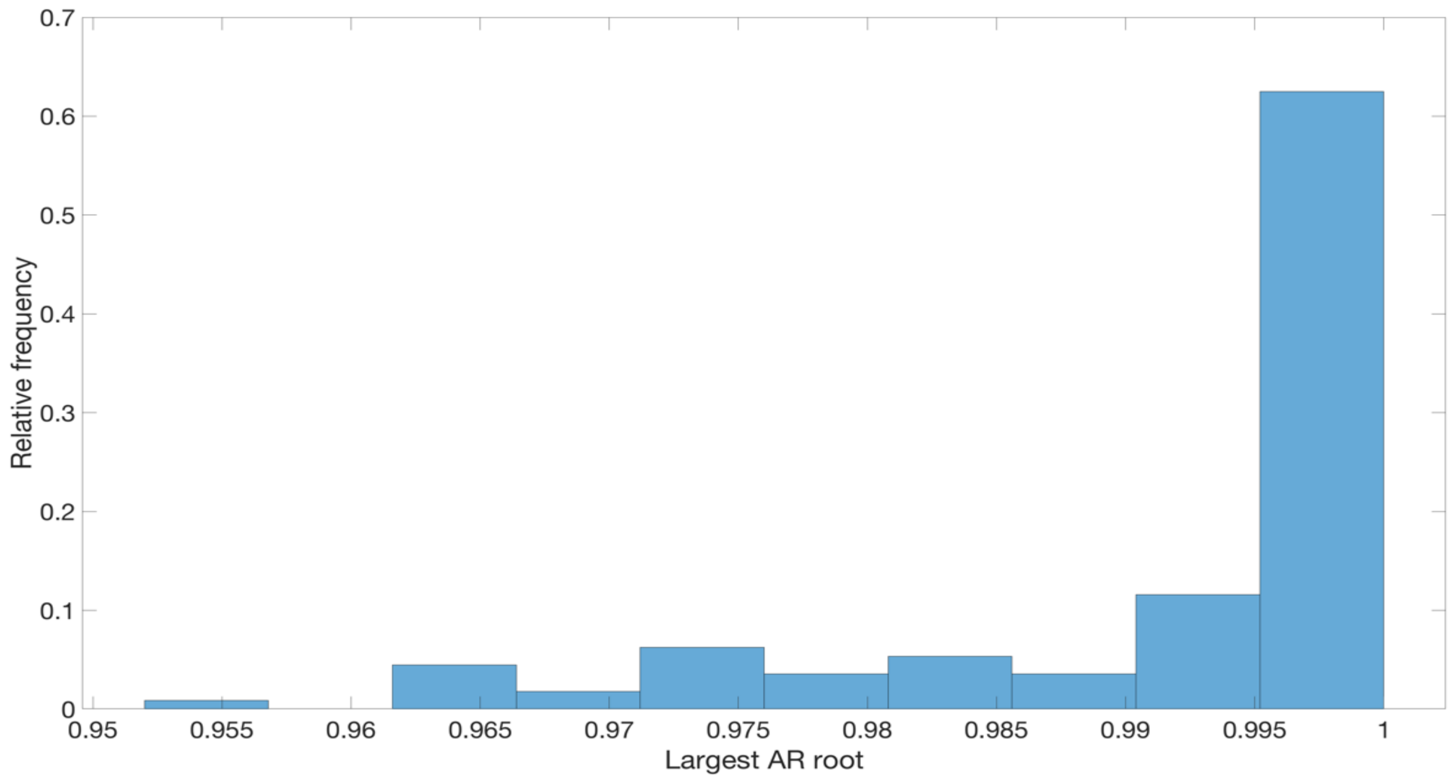
m_t, a_t are independent Gaussian random walks: $\Delta m_t = \varepsilon_{m,t}$, $\Delta a_t = \varepsilon_{a,t}$

With var($\varepsilon_{m,t}$) $\ll$ var($\varepsilon_{a,t}$), f_t evolves like a random walk with drift, but with a slowly varying drift term (m_t). ("local-level" model for f_t).

$$y_{it} = f_t + c_{it}$$

- c_{it} is "very" persistent
- $c_{it} - \bar{c}_{i,1:T}$ is persistent
    - ADF^{μ} statistics
        - Fraction of ADF_tstats < –2.57 (10% CV) = 0.17
        - Fraction of ADF_tstats < –2.86 (5% CV) = 0.10

Histogram of 112 median unbiased estimate of largest AR root from ADF$^{\mu}$ statistics (Stock (19xx)).



Deviation from country-specific means have half – life of 140 years:

$(0.995)^{140} \approx 0.5$

AR component process: $\mathrm{AR}^{\mathrm{Comp}}(\rho_1, \rho_2)$

$$x_t = x_{1,t} + x_{2,t}$$

$$x_{1,t} = \rho_1 x_{1,t-1} + e_{1,t}$$

$$x_{2,t} = \rho_2 x_{2,t-1} + e_{2,t}$$

$$\rho_2 < \rho_1 < 1$$

An alternative model: $(1-\rho \mathrm{L})^d x_t = e_t$

Parameterization: separating persistence and variability

$$y_{it} = f_t + c_{it}$$

$$c_{it} = \mu + \lambda_{c,i}\, g_{J(i),t} + u_{c,i,t}$$

$$g_{j,t} = \lambda_{g,j}\, h_{K(j)} + u_{g,j,t}$$

$$u_{c,i,t} = s_{c,i} w_{c,i,t}$$

$$u_{g,j,t} = s_{g,j} w_{g,j,t}$$

$$h_{k,t} = s_{h,k} w_{h,k,t}$$

where $w_{\cdot,\cdot,t}$ are independent $\mathrm{AR}^{\mathrm{C}}(\rho_{\cdot 1}, \rho_{\cdot 2})$ processes with unit variance and $s_{\cdot}$ are scale factors.

Outline:

1. ~~Look at the data to determine some sensible features of a model.~~

2. ~~Simplify problem: focus on 'long-run' variation/covariation.~~

3. ~~Detailed description of model~~

4. **Estimation mechanics**

5. Results
   a. Convergence
   b. Long-run predictions

6. Different modelling choices

Estimation:

$$y_{it} = f_t + c_{it}$$

$$c_{i,t} = \mu + \lambda_{c,i}\, g_{J(i),t} + u_{c,i,t}$$

$$g_{j,t} = \lambda_{g,j}\, h_{K(j)} + u_{g,j,t}$$

Many parameters:

- f: $(m_0, f_0,\ \sigma_{\Delta m},\ \sigma_{\Delta a})$
- group factors: 25 g-factors, 10 h-factors, (112 - $\lambda_{c,i}$, 25 - $\lambda_{g,j}$)
- persistence: 112 + 25 + 10 values of $(\rho_1, \rho_2, \sigma_1/\sigma_2)$
- variability: 112 + 25 + 10 values of s.

Observations: (number of dots) = $N_{Countries} \times N_{dots/country} \approx 112 \times 10.5$

Estimation by Bayes methods: Some priors will matter

Parameters with priors that don't matter much:

(1) f_t:

- Shrinkage toward OECD: $\overline{y}_t^{OECD} = f_t + \overline{c}_t^{OECD}$ with $\overline{c}_t^{OECD} \sim N(0,\text{small})$

- f_0, m_0, and overall scale (uninformative priors)

(2) c_{it}:

- mean μ, 'average' value of scales s. (Uninformative priors)

- exchangeable hierarchical priors on relative scales and factor loadings (λ_i) (shrunk toward uniform with sensible support).

Parameters with prior that matter:

(1) m_t is local average annual growth rate of f_t:

$\sigma(m_{t+h} - m_t) = \boldsymbol{\sigma_{\Delta m}} \times h^{1/2}$

- $h = 50$
  - Very large value of $\sigma_{\Delta m} \Rightarrow \sigma(m_{t+h} - m_t) = 2\%$
  - Very small value $\sigma_{\Delta m} \Rightarrow \sigma(m_{t+h} - m_t) = 0\%$
  - Prior with linearly decreasing weights between these two values. Mean yields $\sigma_{\Delta m}\sqrt{h}$ = (2/3)% for $h = 50$.

(2) $(\rho_{i,1}, \rho_{i,2}, \sigma_{i,1}/\sigma_{i,2})$: for each of the 112+25+10 components. These are exchangeable with hierarchical prior that is shrunk toward a prior with 'half-life' distributions given below:

half-life : h such that $\mathrm{cor}(x_t, x_{t+h}) = ½$

Percentile	0.10	0.25	0.50	0.75	0.90
h	45	83	193	371	539

Estimation: Practical details

(1) Gaussian Likelihood ... *dots* ~ $N(0,\Sigma(\theta))$,

$$\Sigma(\theta) = \Sigma_1(\theta_1) + \Sigma_2(\theta_2) + \Sigma_3(\theta_3) \ldots + \Sigma_N(\theta_N)$$

(2) Handful of parameters with standard diffuse priors, analytic posterior

(3) Other parameters specified on grid. ($\Sigma_i(\theta_i)$ can be precomputed)

(4) Exchangeable (over countries, factors, etc.) Dirichlet (multinomial) prior on grid of values.

(5) UM computes a zillion draws in 3 minutes.

Outline:

1. ~~Look at the data to determine some sensible features of a model.~~

2. ~~Simplify problem: focus on 'long-run' variation/covariation.~~

3. ~~Detailed description of model~~

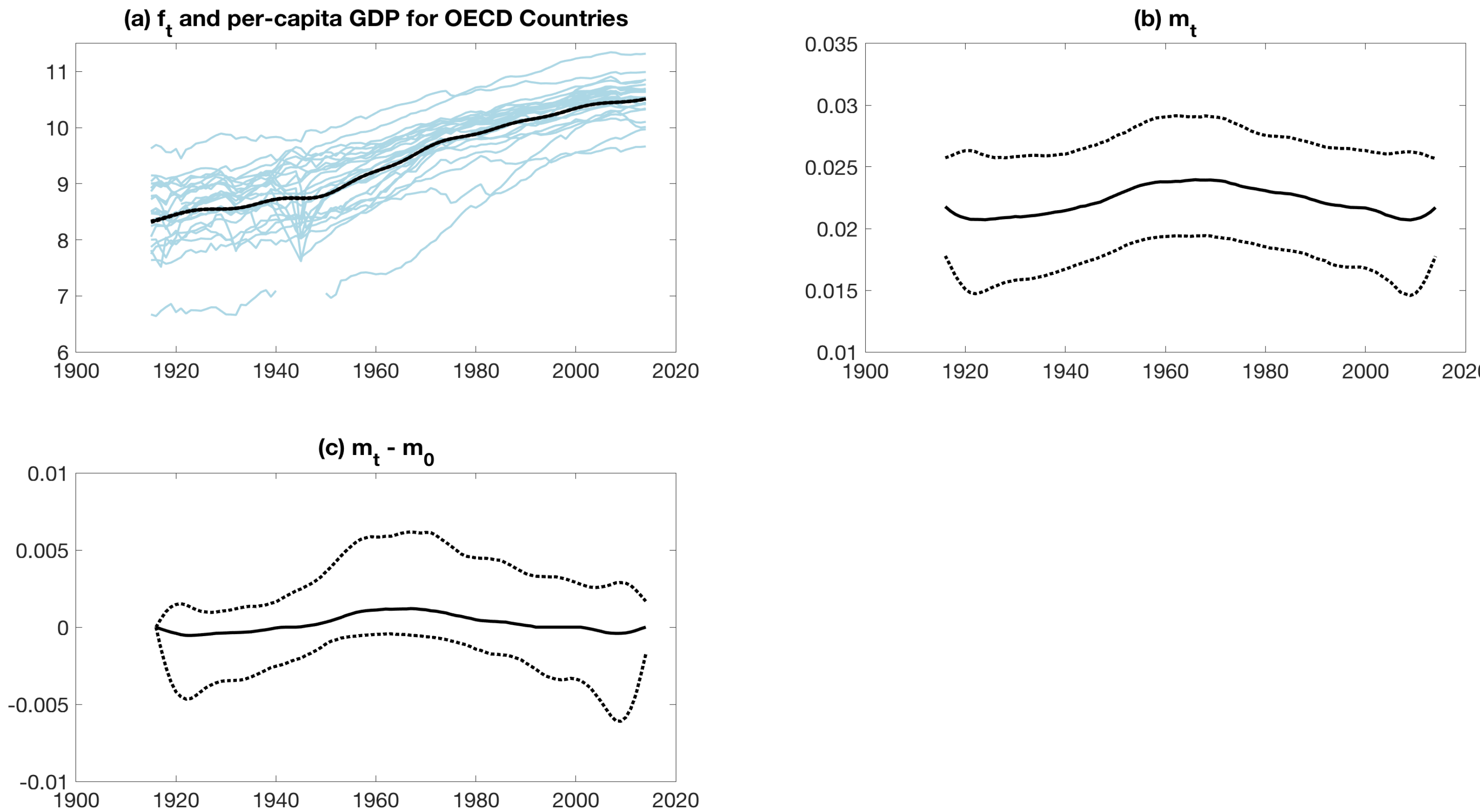
4. ~~Estimation mechanics~~

5. **Results**
   a. Convergence
   b. Long-run predictions

6. Different modelling choices

Selected Results: *f*-factor

$$y_t = f_t + c_{i,t} \qquad f_t = f_0 + m_t \times t + a_t$$

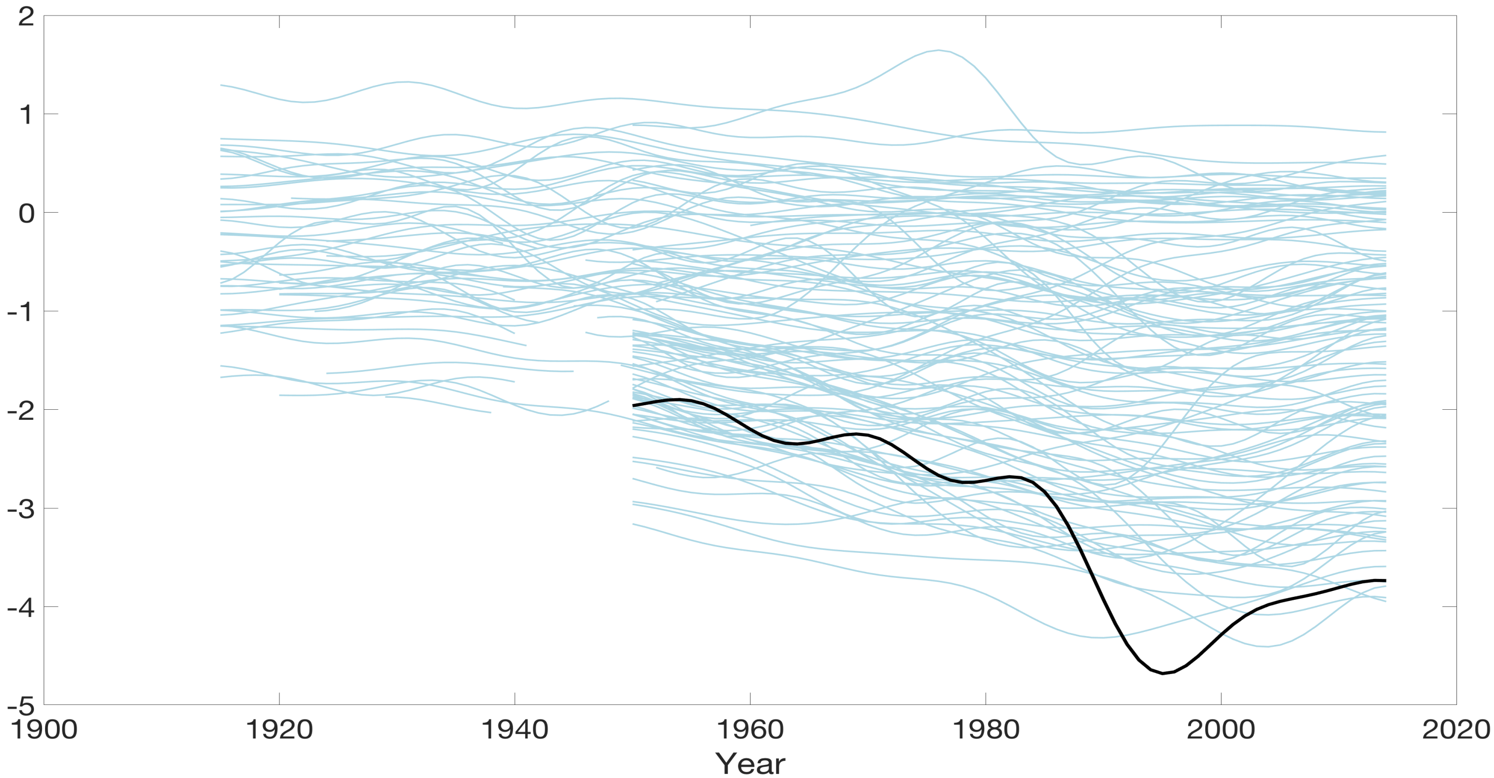


Median and 68% pointwise credible set

(WIP . narrowing of bands in (c) at end of sample)

Selected Results: Persistence and variance of c_{it}

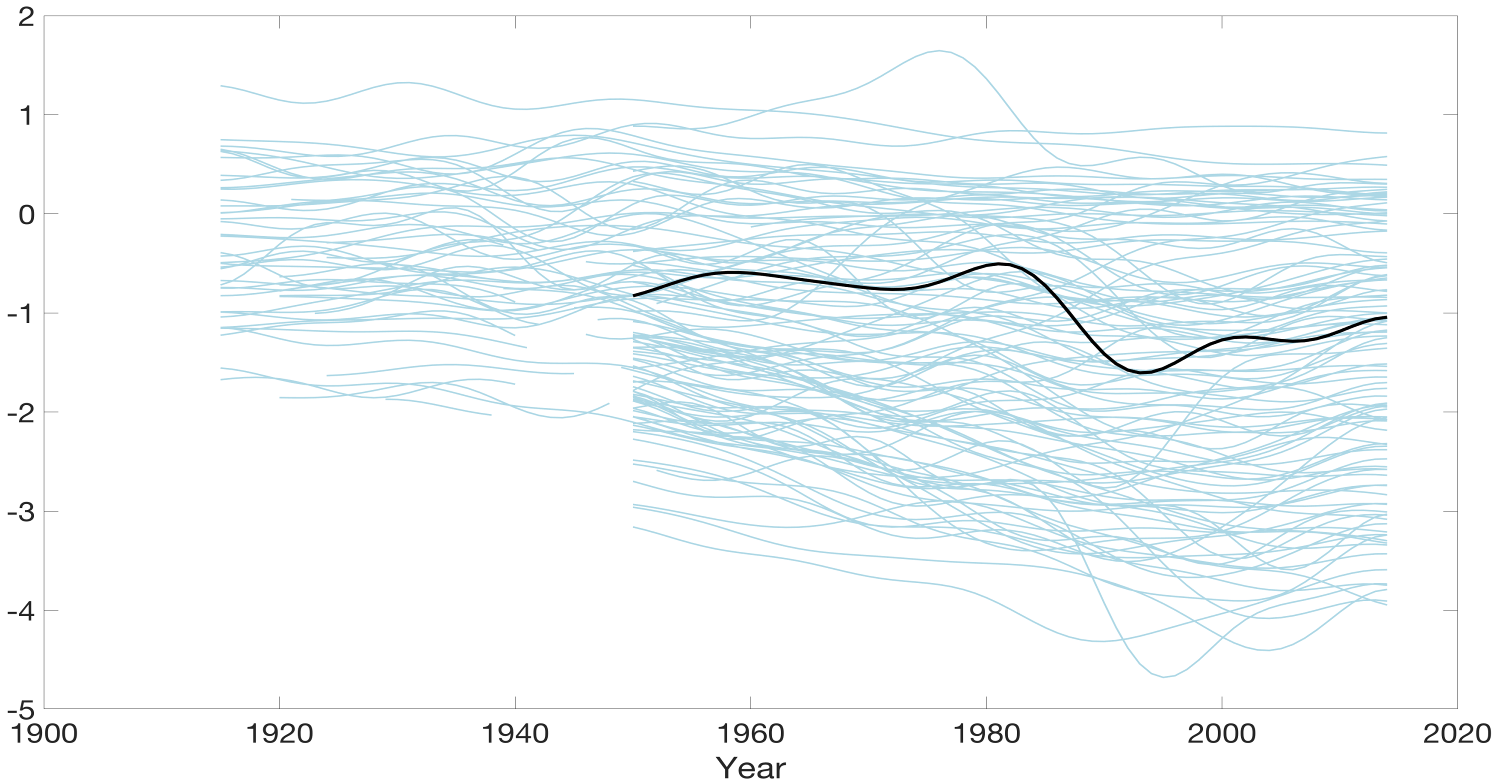
Posterior means of c_{it}: Liberia



Percentiles of posterior

	0.05	0.16	0.50	0.84	0.95
half-life	37	44	63	98	136
σ_c	1.3	1.4	1.5	1.6	1.7
$\sigma_{\Delta_{50}c_{it}}$	1.1	1.2	1.4	1.6	1.7

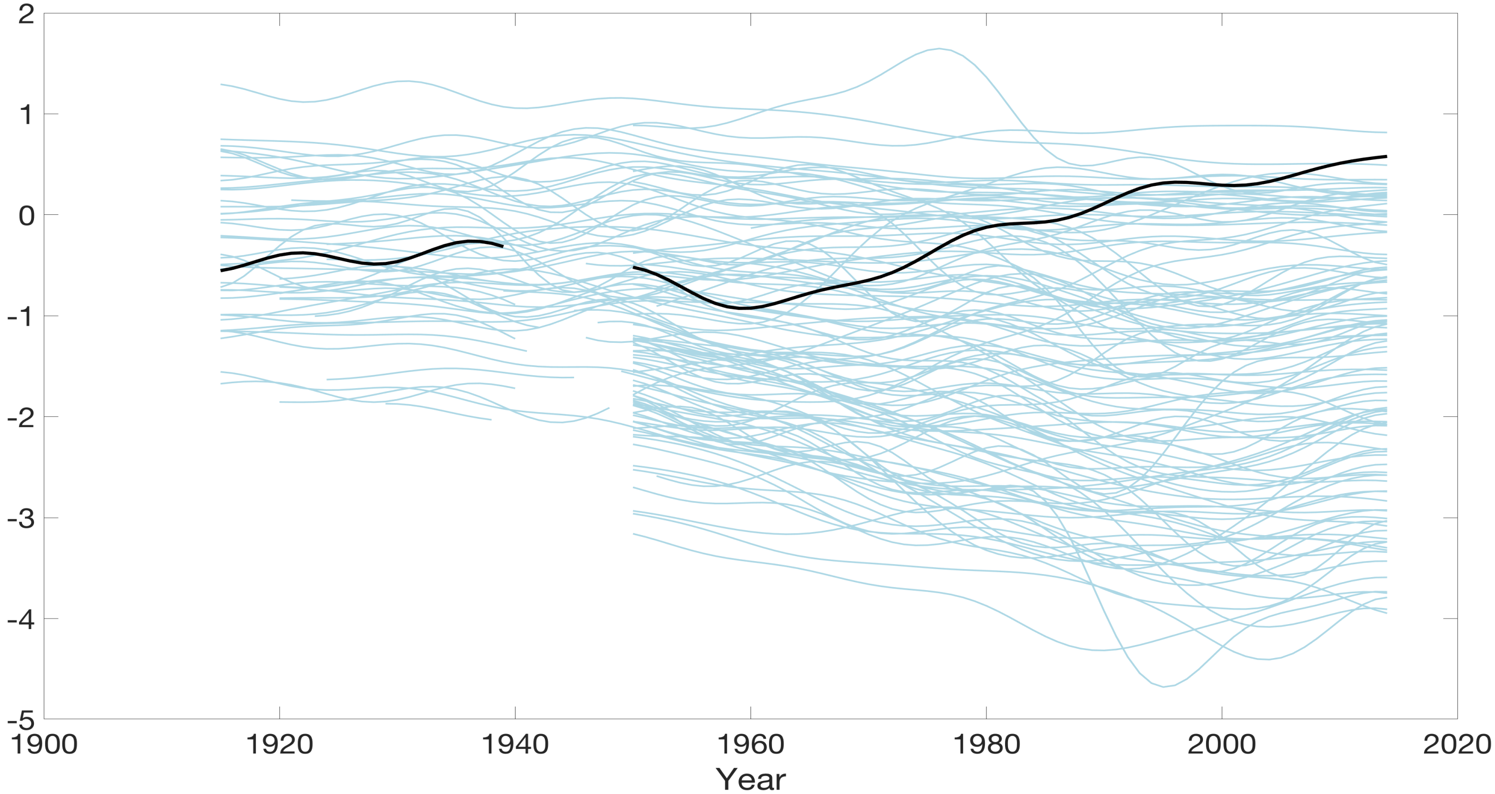
Posterior means of c_{it}: Iraq



Percentiles of posterior

	0.05	0.16	0.50	0.84	0.95
half-life	38	50	85	158	229
σ_c	0.8	0.9	1.2	1.5	1.6
$\sigma_{\Delta_{50}c_{it}}$	0.7	0.8	1.0	1.2	1.3

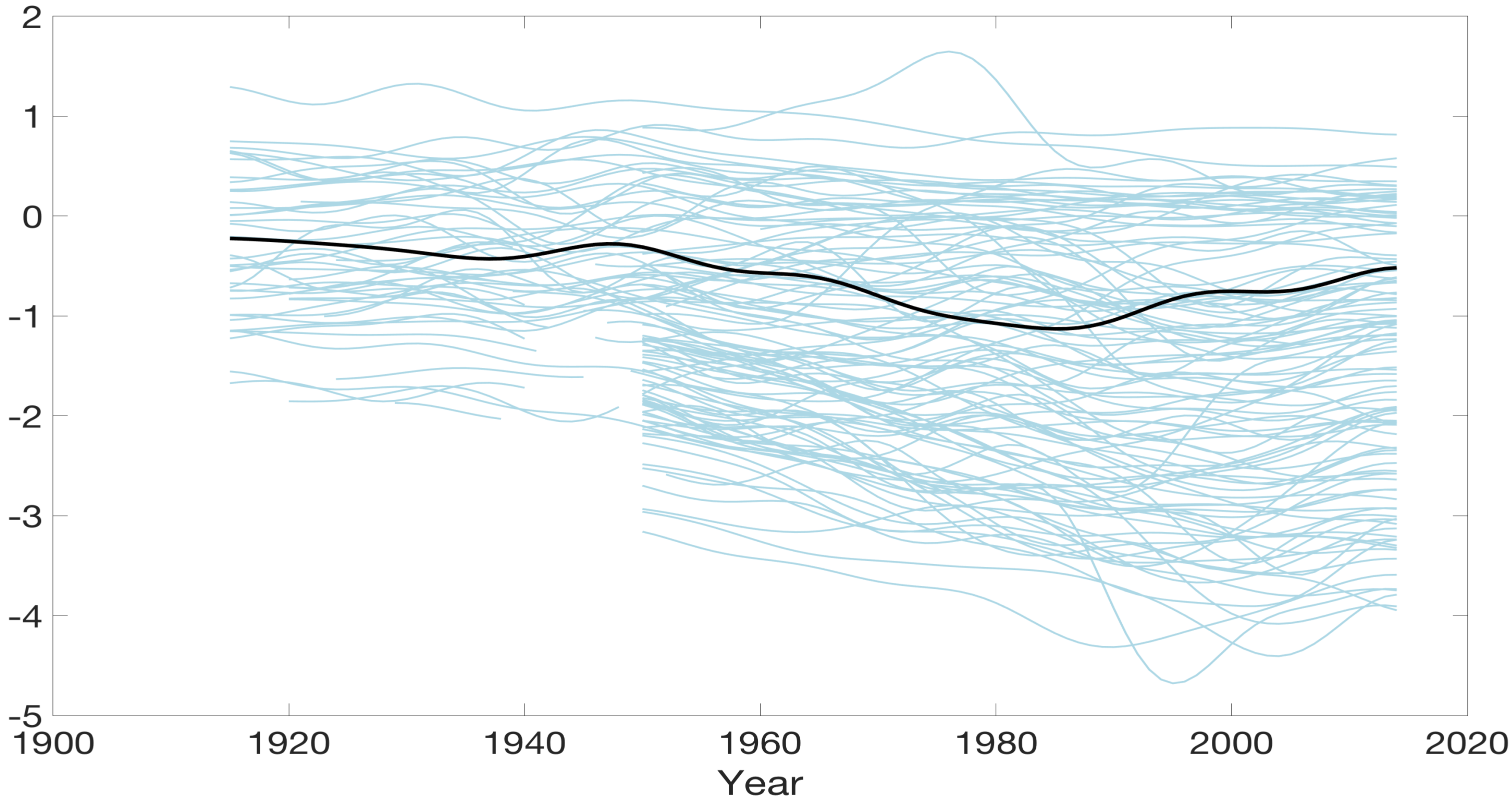
Posterior means of c_{it}: Singapore



Percentiles of posterior

	0.05	0.16	0.50	0.84	0.95
half-life	62	90	163	277	370
σ_c	0.8	0.9	1.1	1.4	1.5
$\sigma_{\Delta_{50} c_{it}}$	0.6	0.6	0.8	1.0	1.1

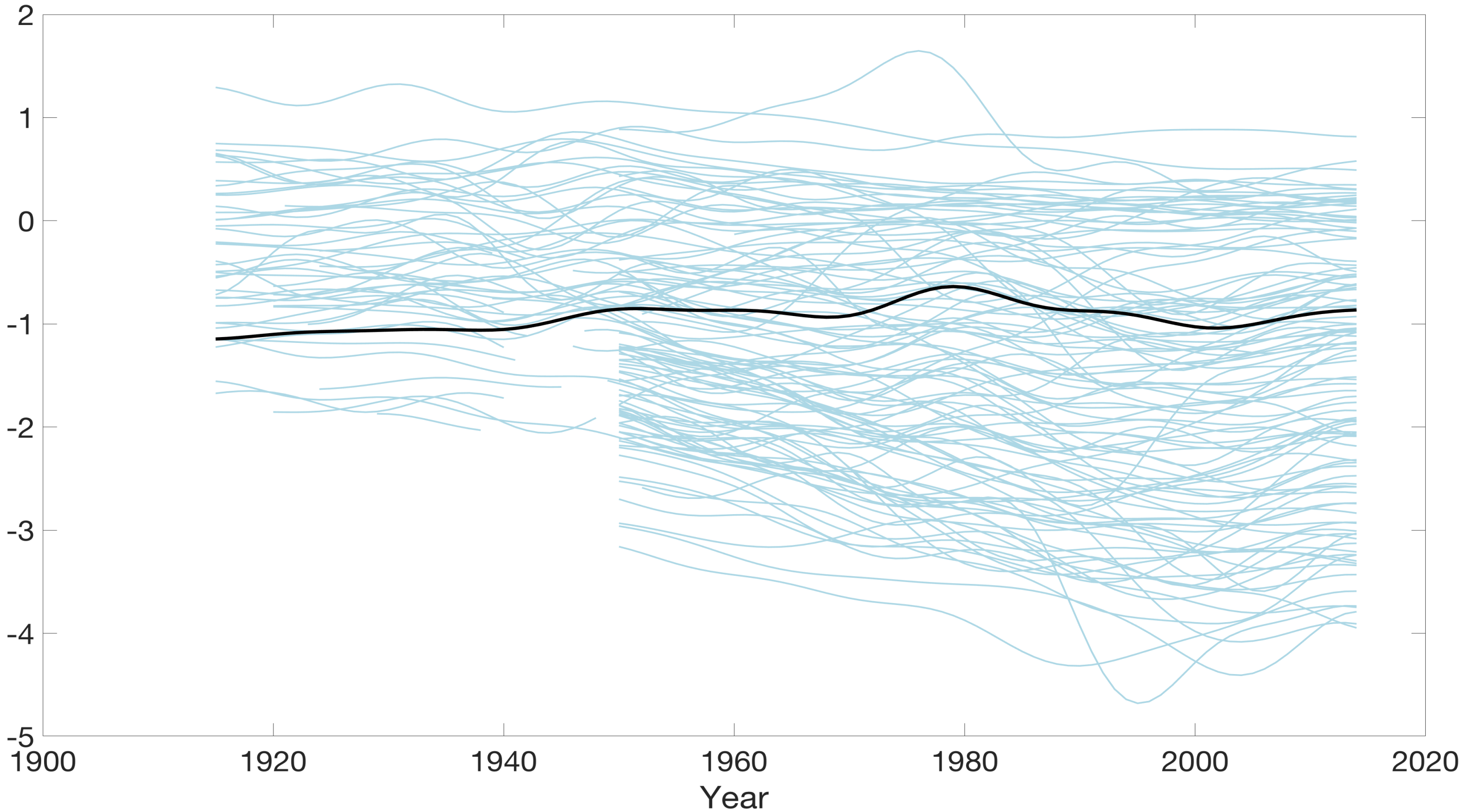
Posterior means of c_{it}: Chile



Percentiles of posterior

	0.05	0.16	0.50	0.84	0.95
half-life	117	168	270	416	387
σ_c	0.9	1.0	1.2	1.4	1.5
$\sigma_{\Delta_{50}c_{it}}$	0.5	0.6	0.7	0.8	0.9

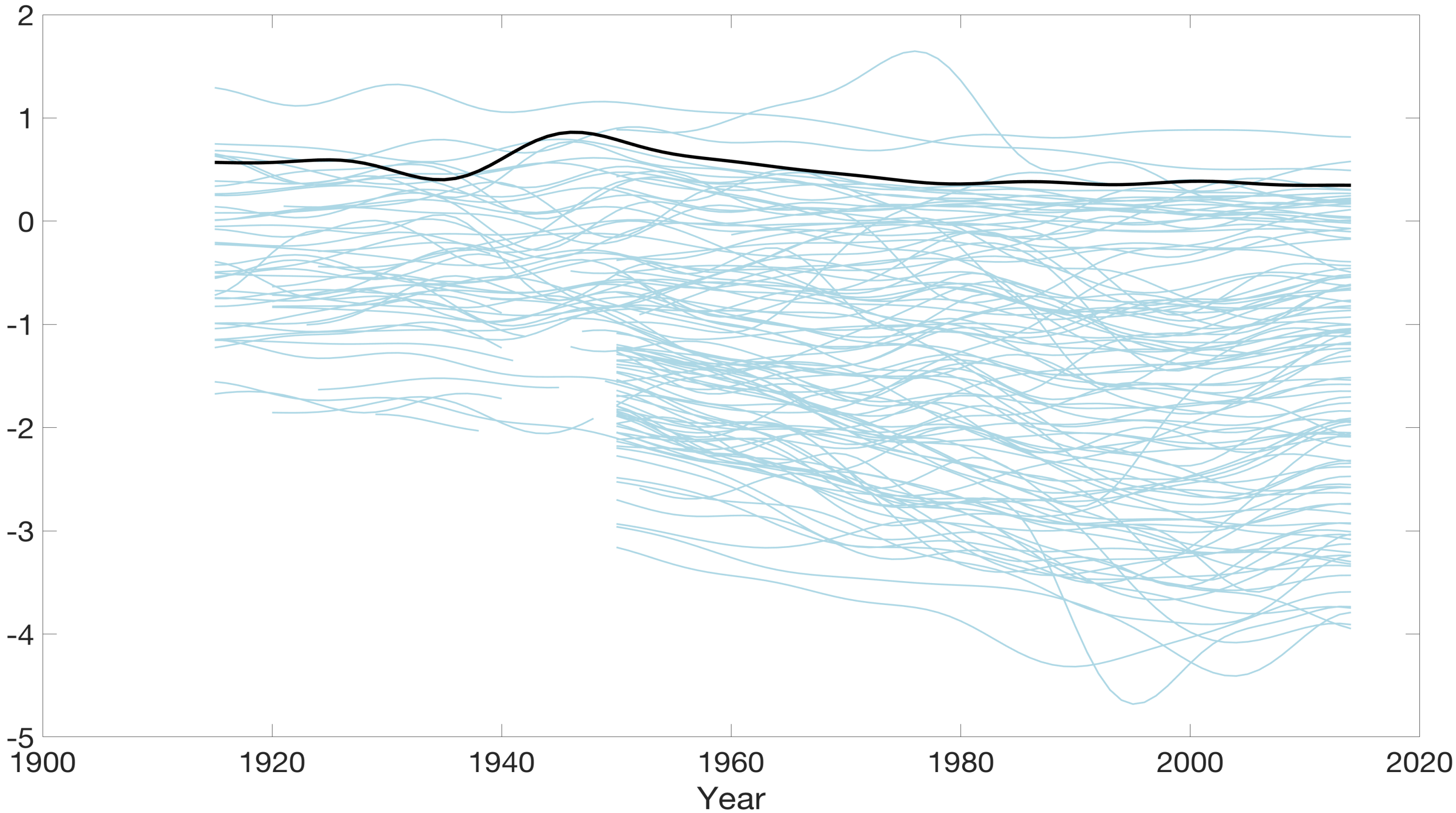
Posterior means of c_{it}: Brazil



Percentiles of posterior

	0.05	0.16	0.50	0.84	0.95
half-life	155	206	313	441	523
σ_c	0.7	0.7	0.9	1.1	1.2
$\sigma_{\Delta_{50}c_{it}}$	0.6	0.7	0.8	0.9	1.0

Posterior means of c_{it}: United States



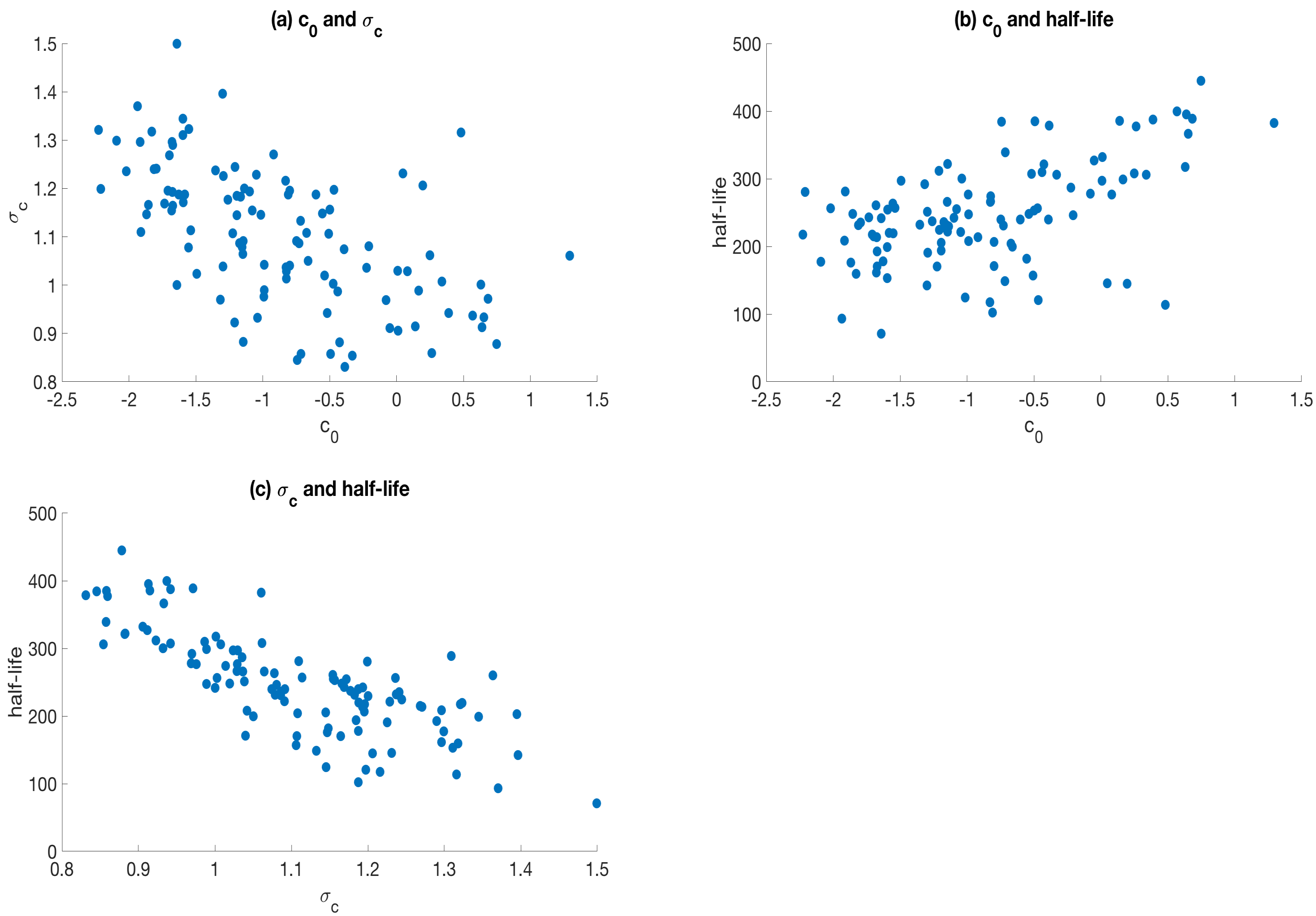
Percentiles of posterior

	0.05	0.16	0.50	0.84	0.95
half-life	218	277	396	527	599
σ_c	0.7	0.8	0.9	1.1	1.3
$\sigma_{\Delta_{50} c_{it}}$	0.3	0.4	0.4	0.5	0.5

Distribution of Posterior Means Across 112 Countries

	Percentile				
	0.05	0.16	0.50	0.84	0.95
Half-life	120	171	242	321	386
σ_c	0.86	0.94	1.11	1.27	1.35
$\sigma_{\Delta_{50} c_{it}}$	0.40	0.48	0.66	0.84	0.97

Selected Results: Initial Conditions, σ_c and half-life



Selected Results: **Covariability**

Posterior Means of pairwise correlations

	$\Delta_{50}\, y_{i,t}$		$\Delta_{50}\, c_{i,t}$
average	0.37		0.08
largest	0.95 (France, Netherlands)		0.92 (France, Netherlands)
smallest	0.12 (Liberia, Saudi Arabia)		0.00 (Fraction < 0.01 = 0.39)

Average Pairwise Correlations of $\Delta_{50}c_{it}$ (Posterior means) in Selected 5-country groups

Countries					Correlation
China	India	Laos	Sri Lanka	Vietnam	0.71
Hong Kong	Korea	Singapore	Taiwan	Thailand	0.67
Cent. African Rep.	Guinea	Haiti	Senegal	Madagascar	0.63
Belgium	Denmark	France	Italy	Netherlands	0.59
Benin	Bangladesh	Kenya	Nepal	Tanzania	0.53
Bulgaria	HRV	ROU	Russia	Serbia	0.51
Australia	Canada	Great Britain	New Zealand	United States	0.47
Burkino FAso	Ghana	Mozambique	Chad	Uganda	0.45
Brazil	Costa Rica	Dominincan Rep.	Ecuador	Poland	0.41
Cote d'Ivoire	Mauritania	Niger	Togo	Zambia	0.41
Argentina	Bolivia	Peru	El Salvador	Uruguay	0.40
Switzerland	Finland	Norway	Portugal	Sweden	0.36

Selected Results: Long-run Forecasts

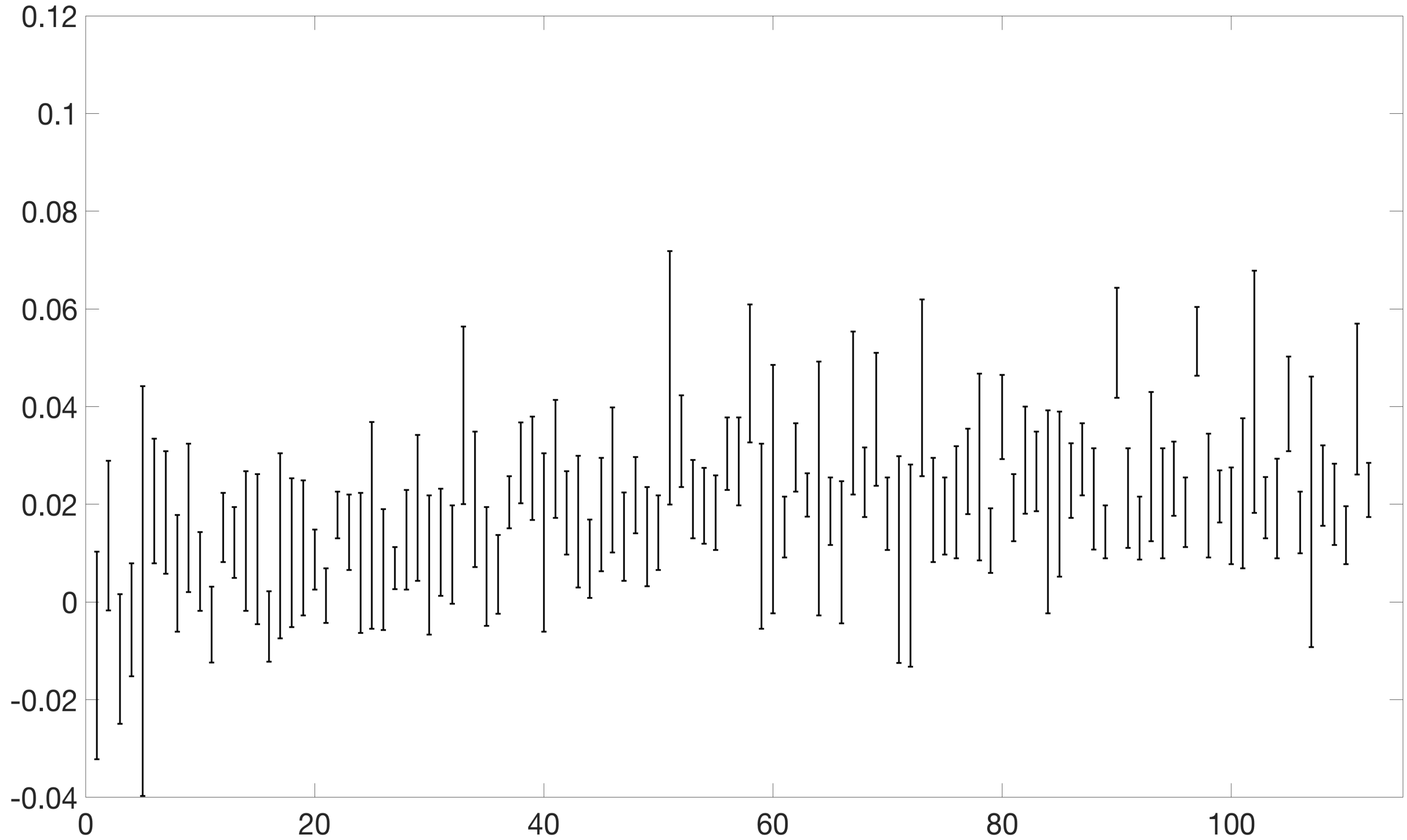
Average growth over next h years: $\left(y_{i,T+h} - y_{i,T}\right)/\,h$ for $h = 50, 100$

Univariate Benchmarks (location, scale, equivariant prediction intervals):

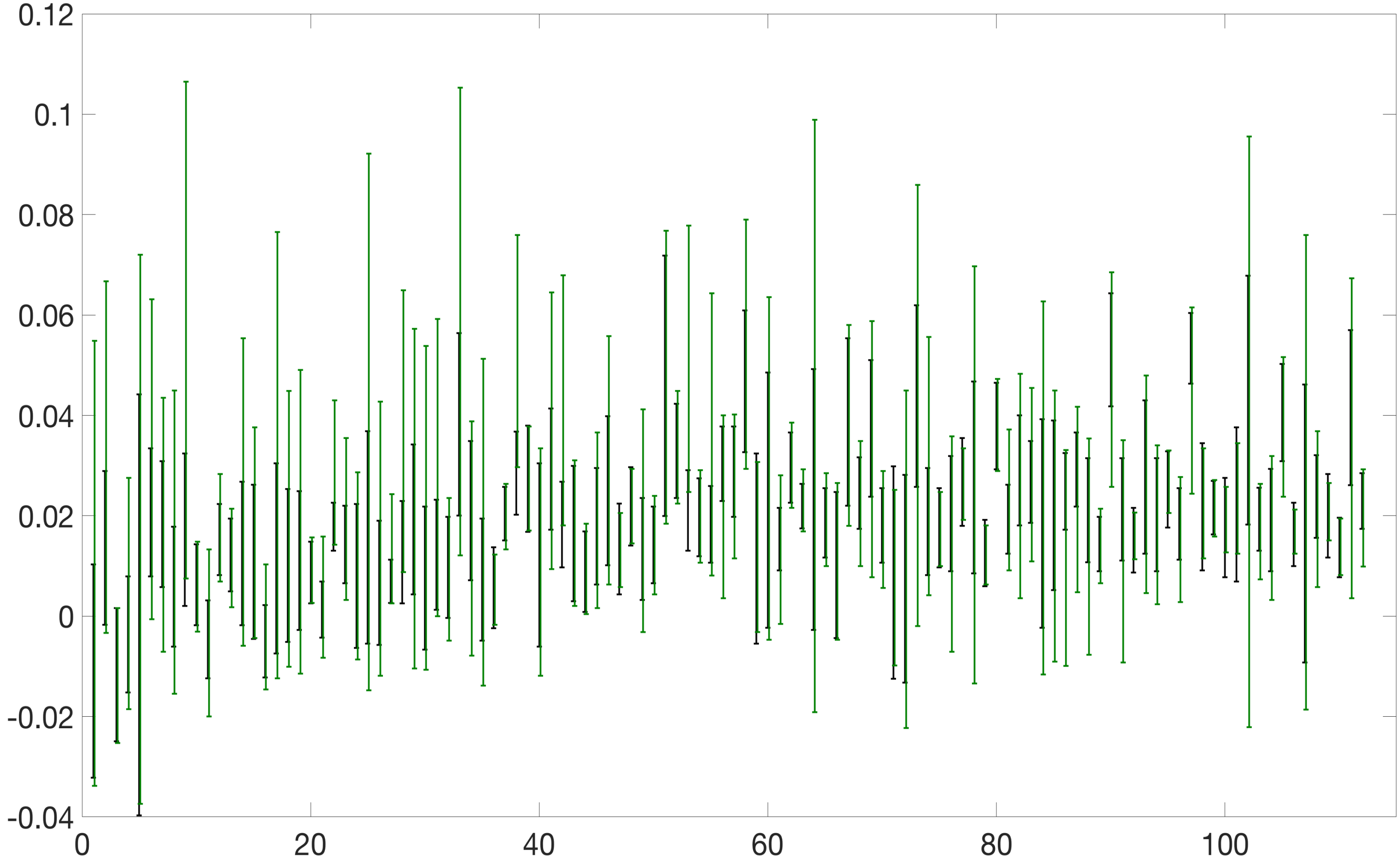
- $(1\text{-L})y_{it} = \mu + u_{it}$
- $(1\text{–L})^{1+\boldsymbol{d}}y_{it} = \mu + u_{it}$ ($d \sim$ U(–0.4, 1.0))

Univariate benchmarks: $(1\text{-}L)y_{it} = \mu + u_{it}$

67% prediction intervals for average growth over next 100 years. Countries ordered from poorest to richest (2010-2014)

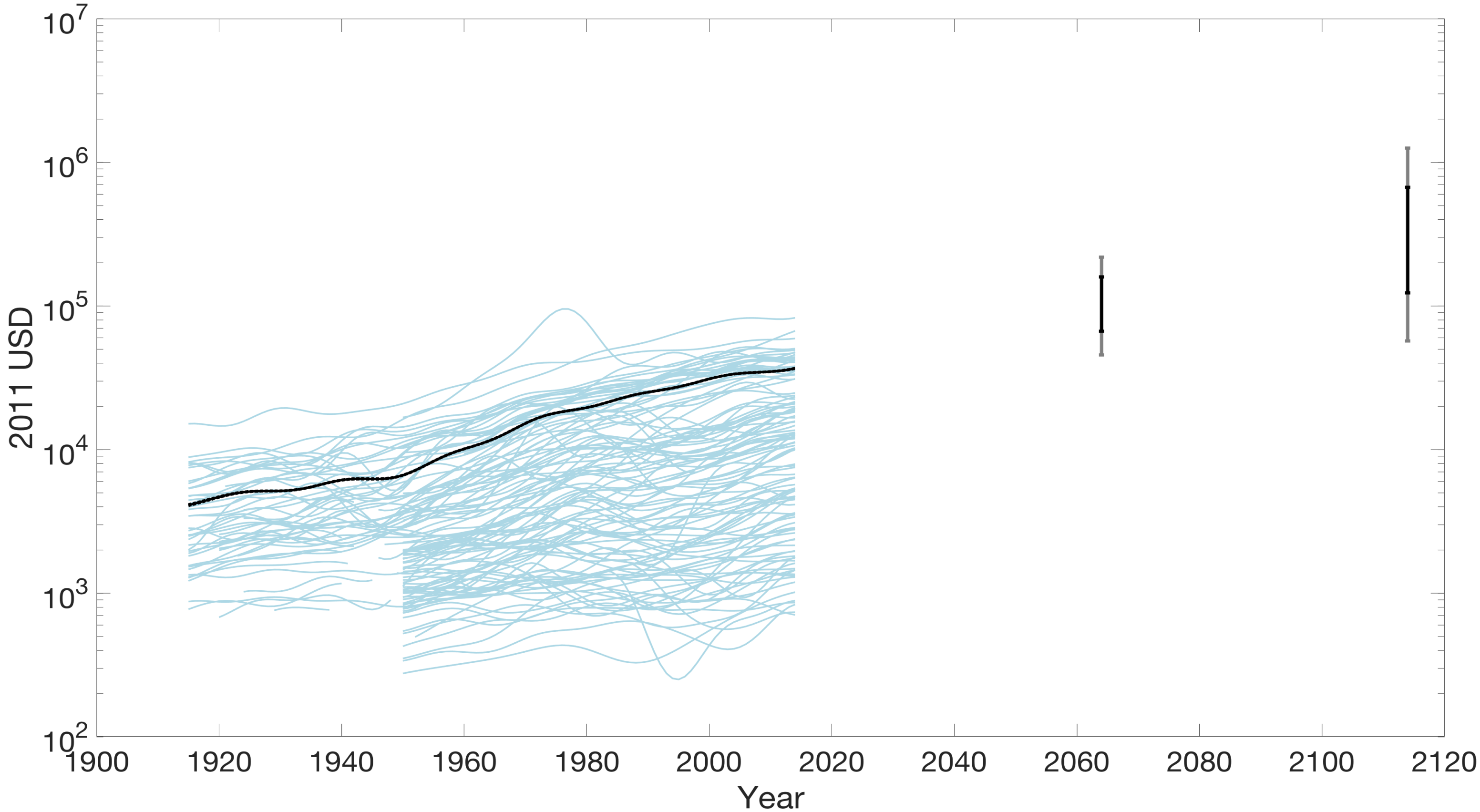


Univariate benchmarks: $(1-L)y_{it} = \mu + u_{it}$ and $(1-L)^{1+d}y_{it} = \mu + u_{it}$



Univariate and Multivariate

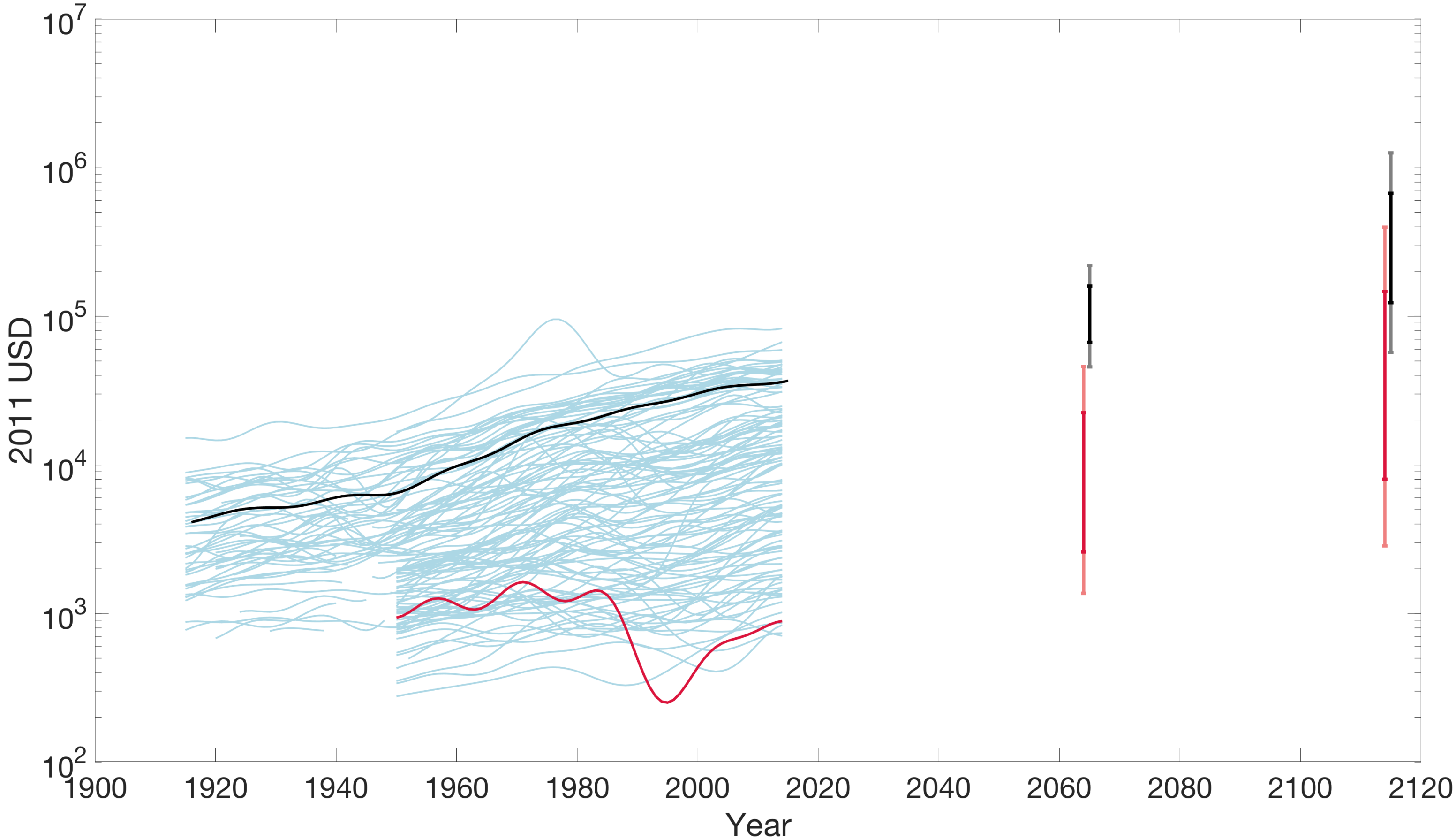
50 and 100 year forecasts: *f*-factor



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5

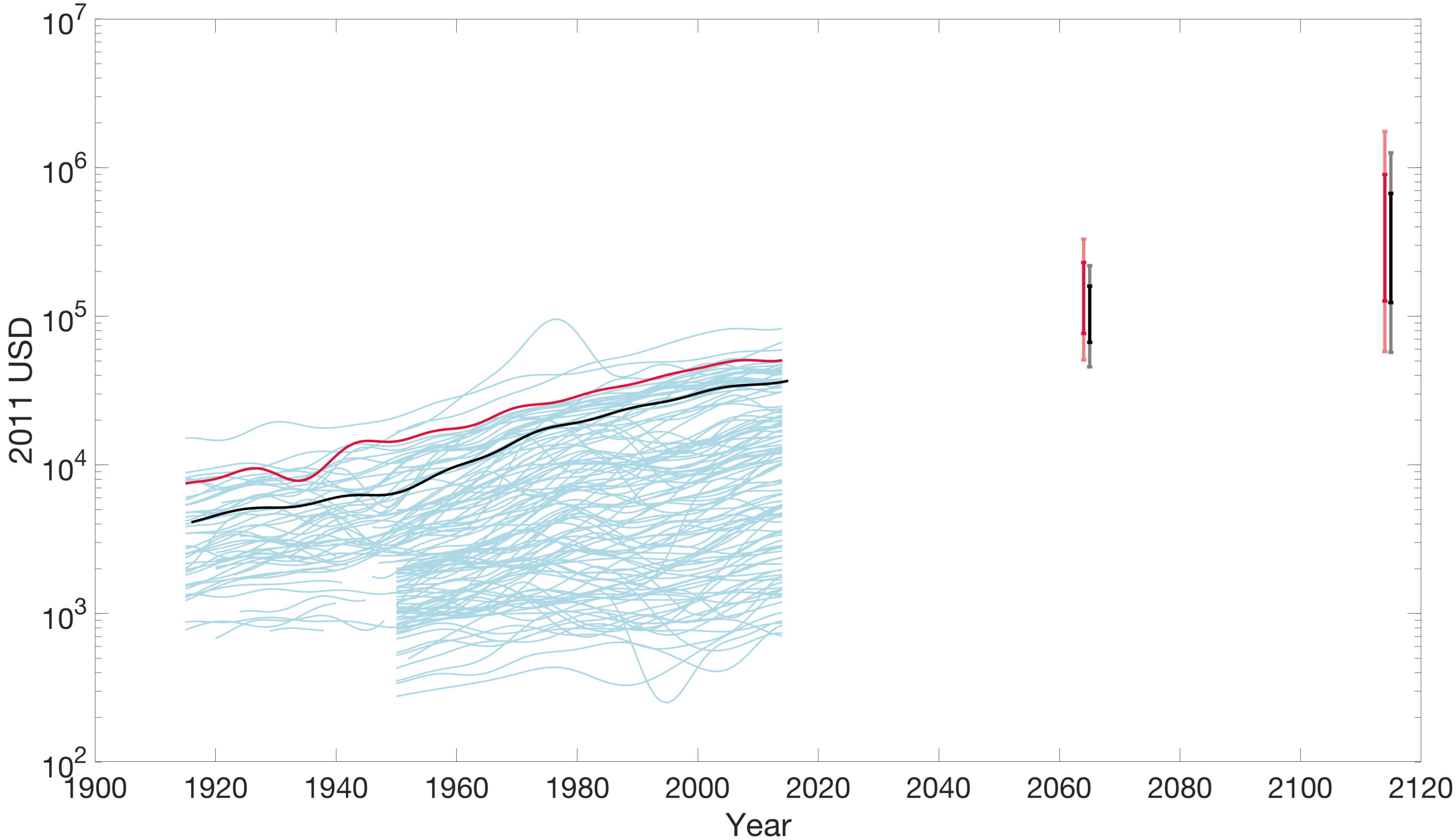
50 and 100 year forecasts: Liberia



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
Liberia	1.1	2.2	3.6	5.0	6.0

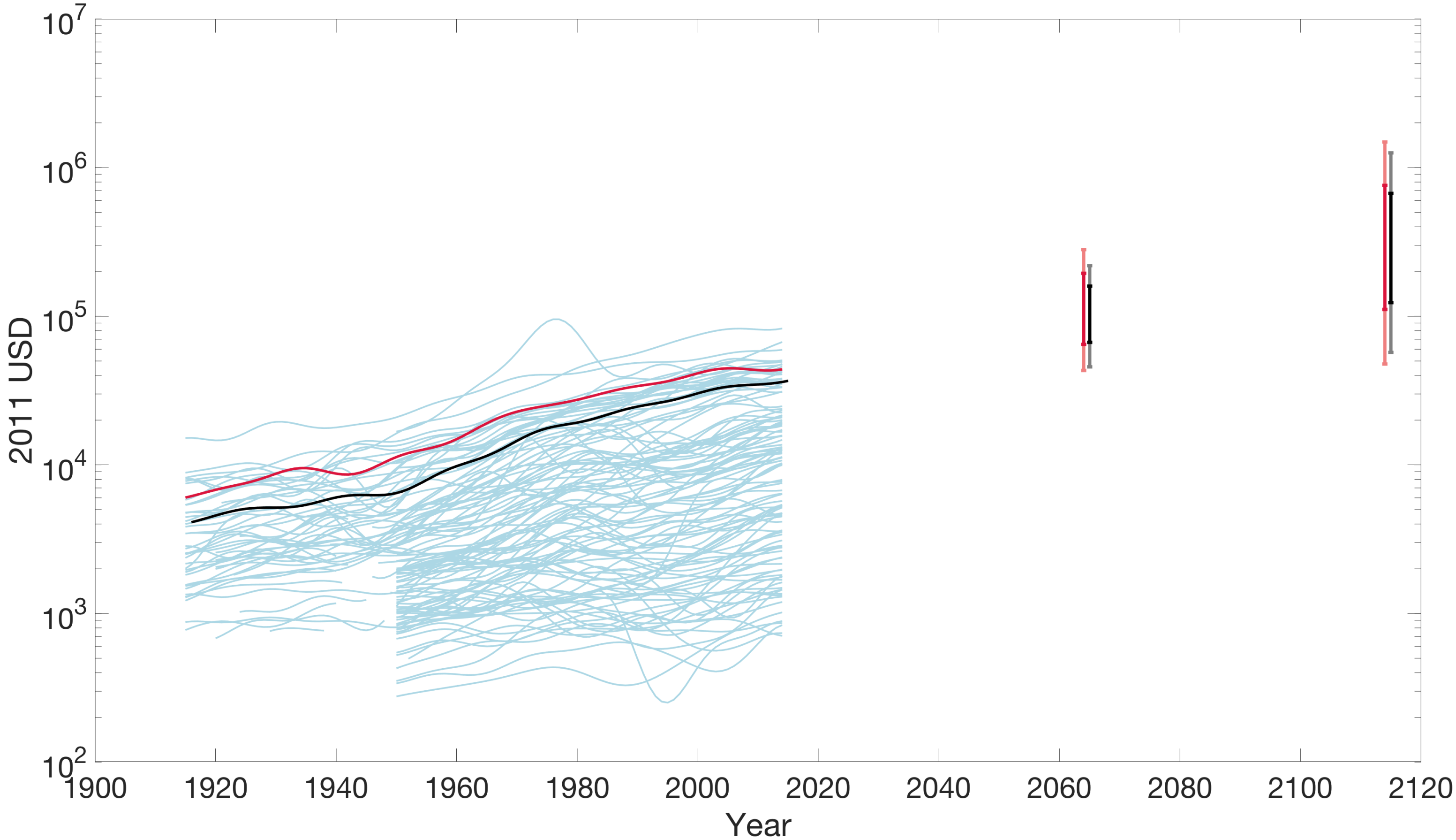
50 and 100 year forecasts: USA



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
USA	0.1	0.9	1.9	2.9	3.5

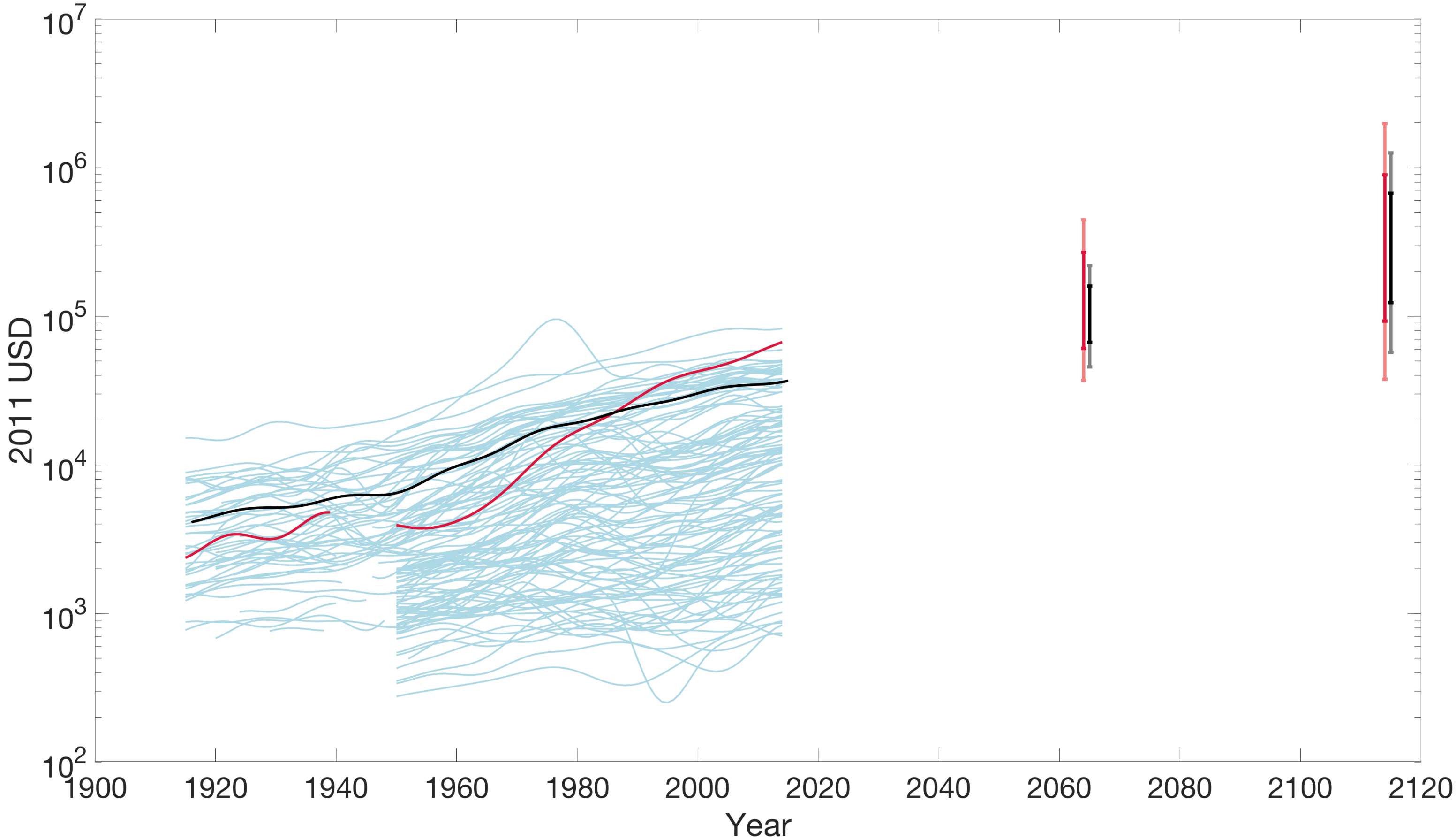
50 and 100 year forecasts: Denmark



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
Denmark	0.1	0.9	1.9	2.9	3.5

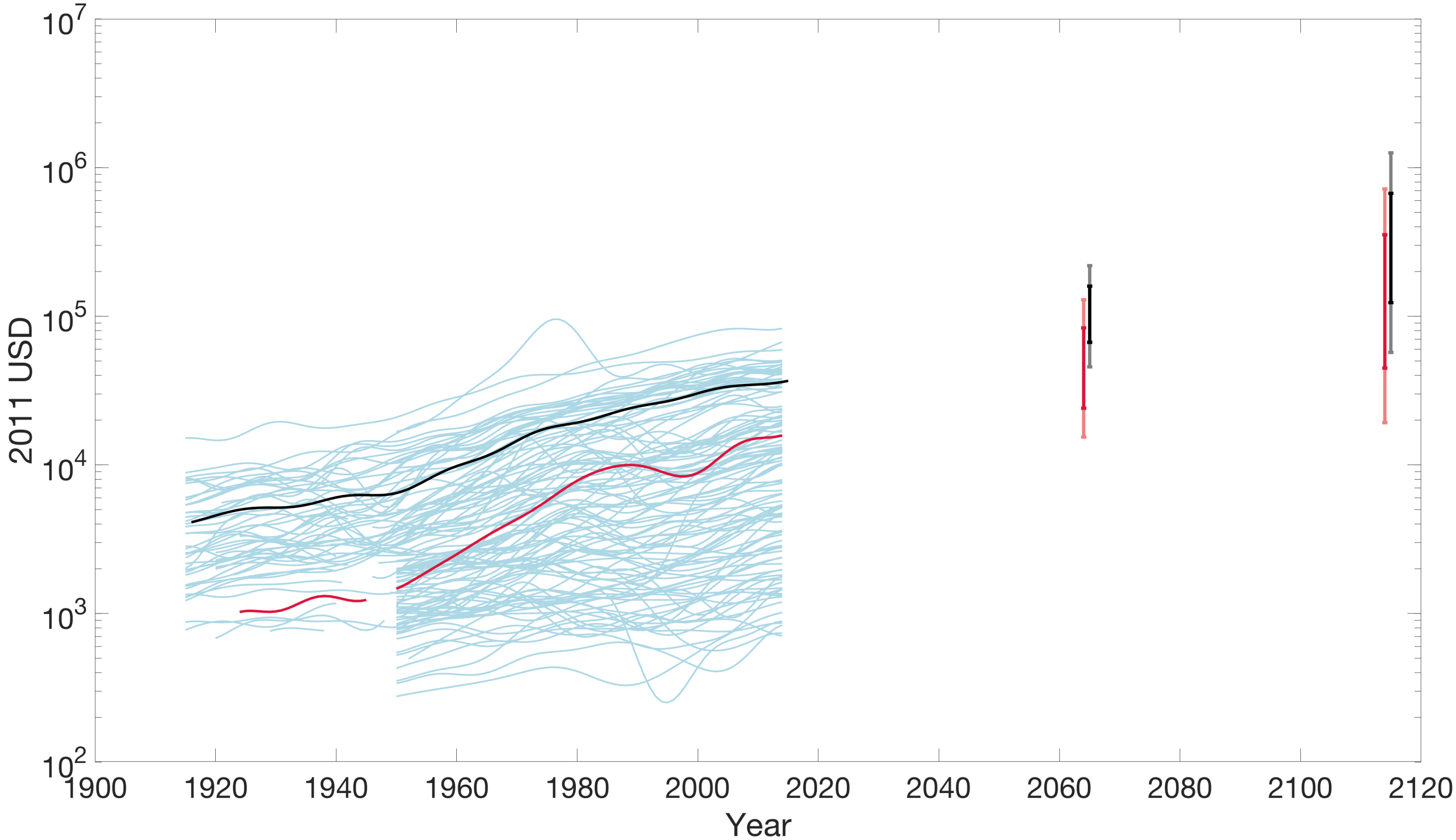
50 and 100 year forecasts: Singapore



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
Singapore	–0.5	0.4	1.5	2.6	3.4

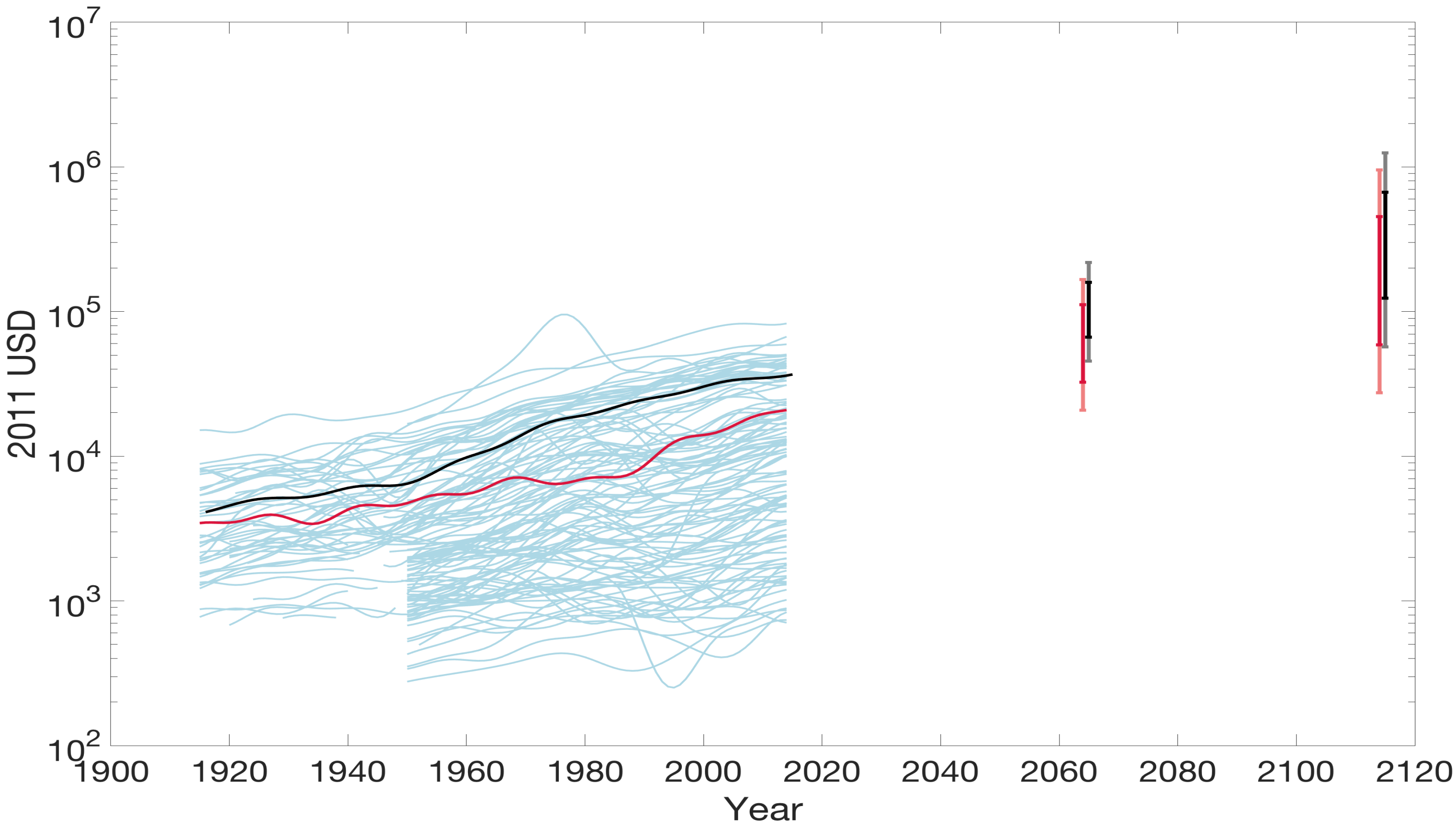
50 and 100 year forecasts: Bulgaria



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
Bulgaria	0.2	1.0	2.1	3.1	3.8

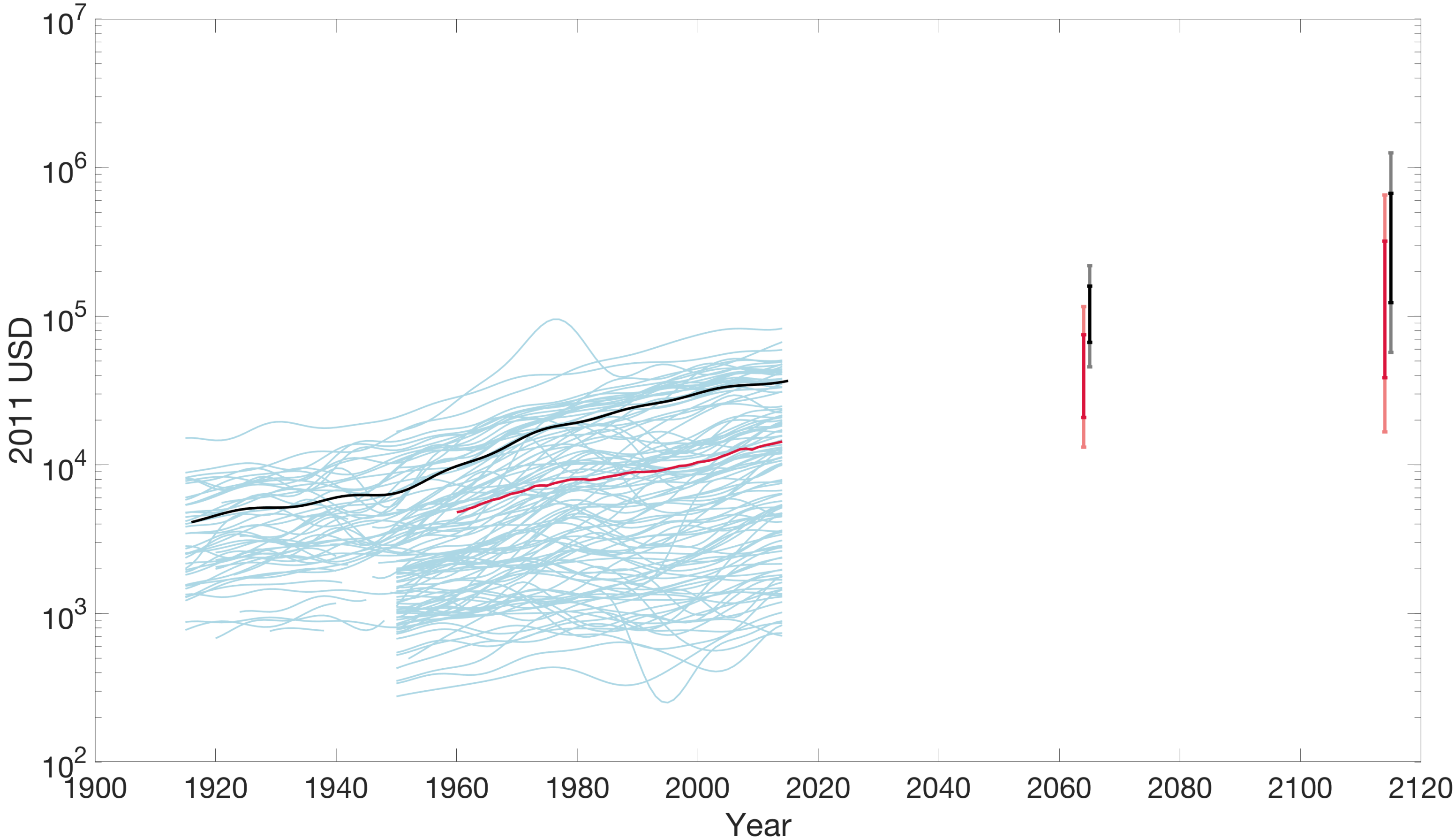
50 and 100 year forecasts: Chile



Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
Chile	0.3	1.0	2.1	3.1	3.8

50 and 100 year forecasts: global average (2014 population weights)

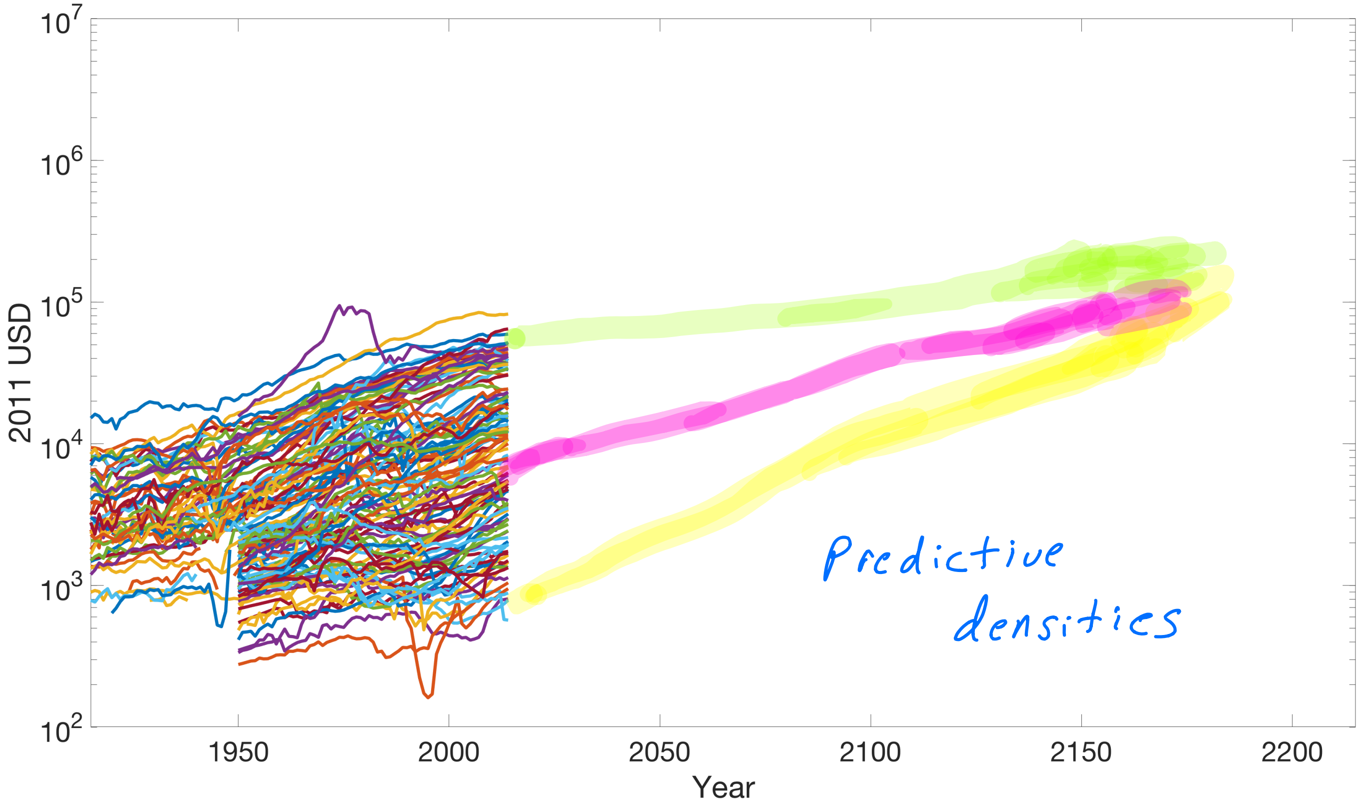


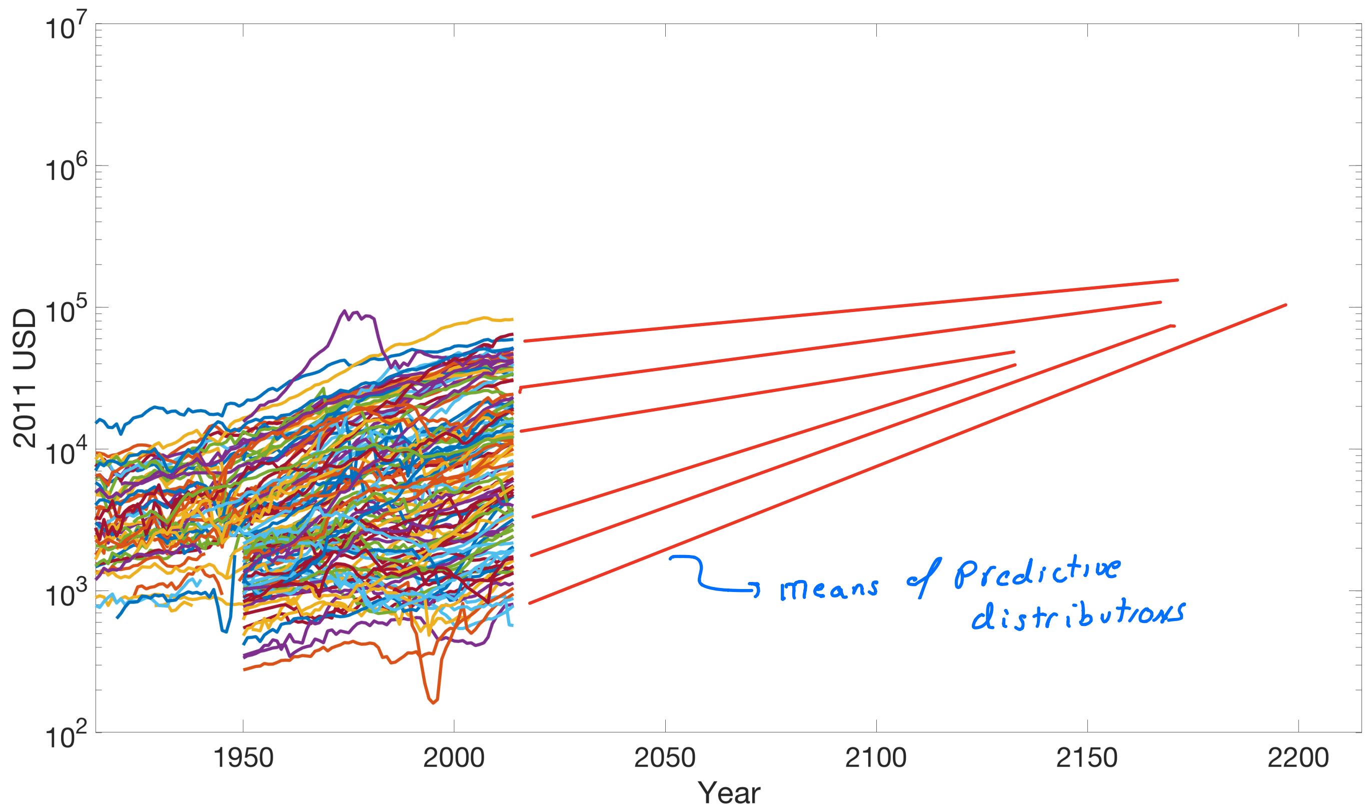
Percentiles of Predictive Distribution: 100-year average growth rate (PAAR)

	5%	16%	50%	84%	95%
f-factor	0.4	1.2	2.1	2.9	3.5
Global avg.	0.5	1.3	2.3	3.2	3.8

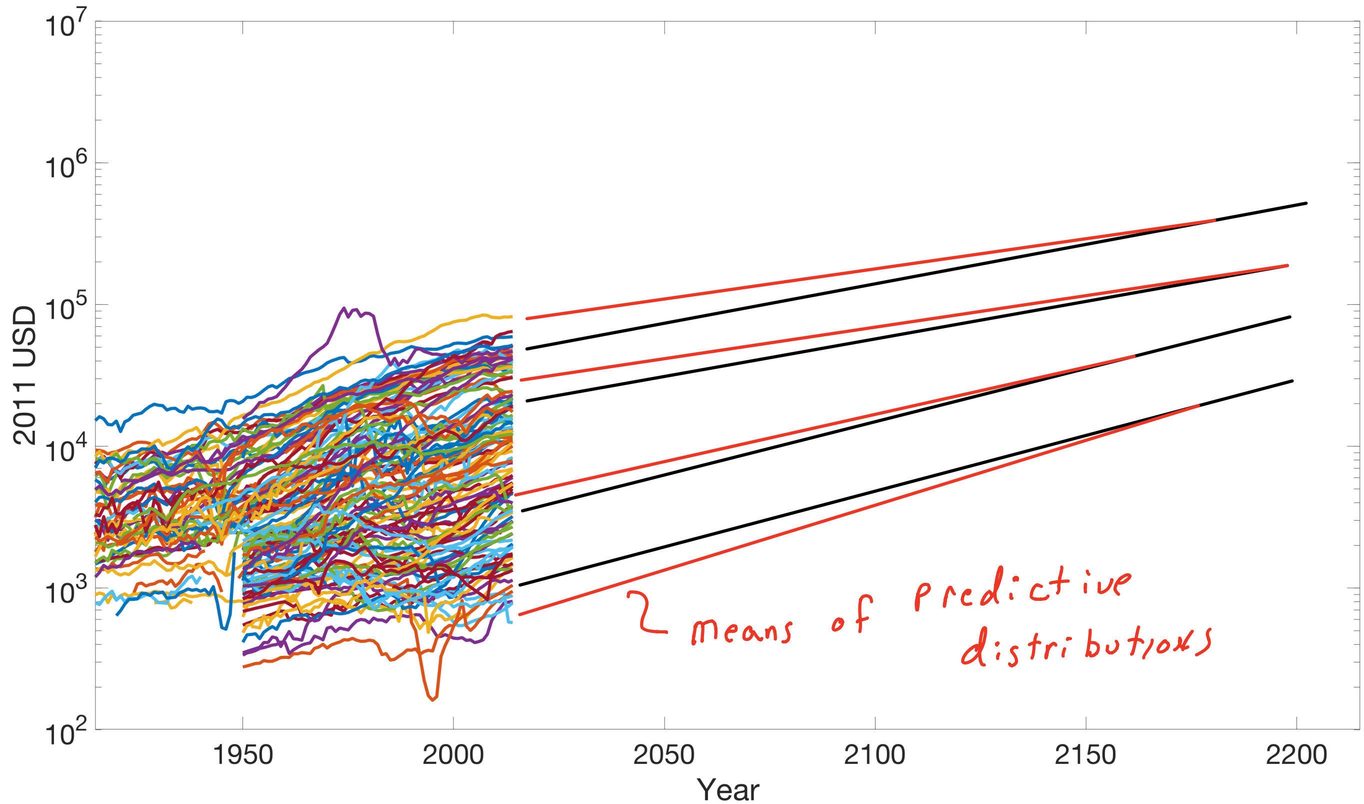
Summary

Convergence, persistence and comovement

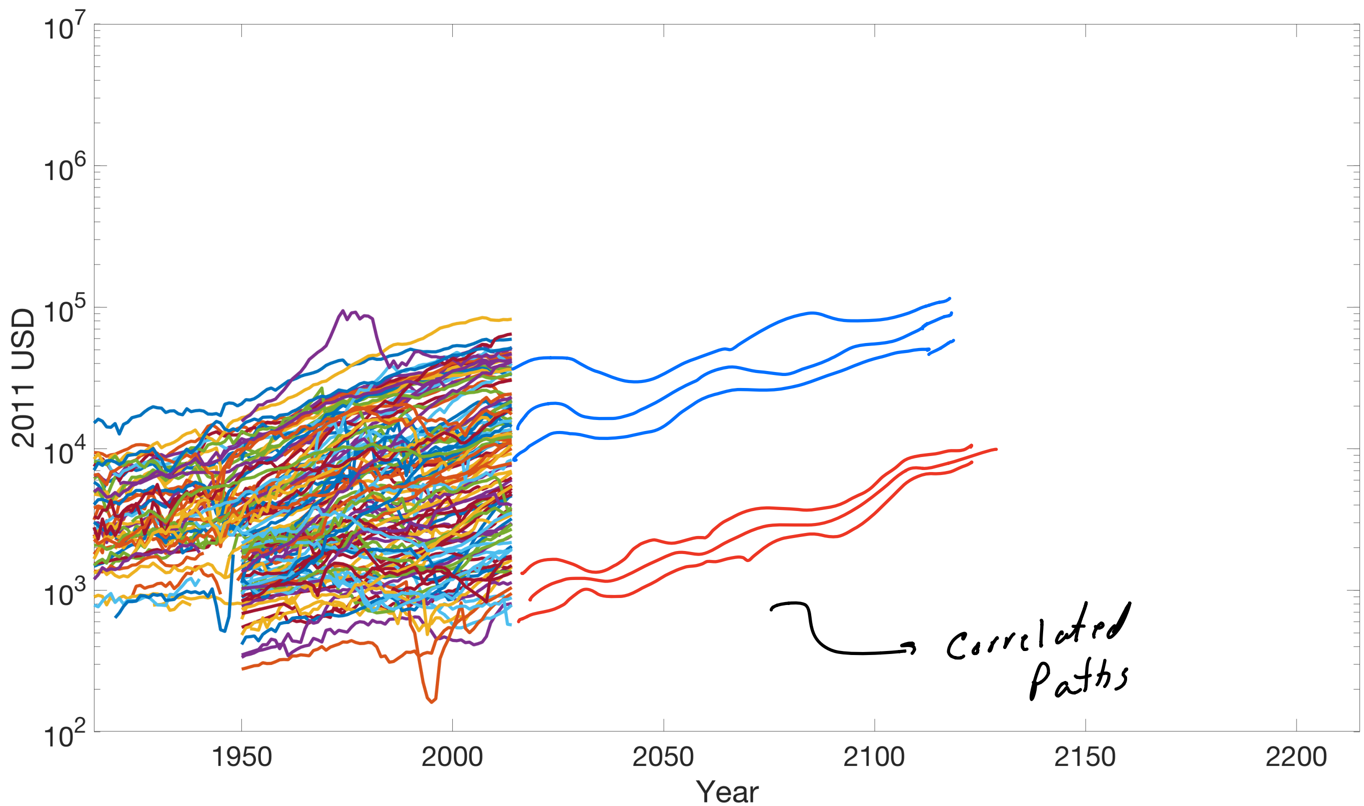




10^7
10^6
10^5
10^4
10^3
10^2
2011 USD
1950
2000
2050
2100
2150
2200
Year
means of predictive distributions



10^7
10^6
10^5
10^4
10^3
10^2
2011 USD
1950
2000
2050
2100
2150
2200
Year
2 means of predictive distributions



10^7
10^6
10^5
10^4
10^3
10^2
2011 USD
1950
2000
2050
2100
2150
2200
Year
Correlated Paths

That's it so far …