CRIX - a CRyptocurrency IndeX

Simon Trimborn

Wolfgang Karl Härdle

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de

Currencies - Cigarettes, USD, Cryptos

- Everything can be a currency

Figure 1: Cigarette trading in postwar germany (59)

- Everyone can offer a currency

Figure 2: Fridrich A. Hayek (59)

Digital Economy

- Amazon
- Paypal
- Google Wallet
- Cryptocurrencies
- Ripple

Cryptocurrencies

- Decentralized, virtual, low transaction costs

- NYSE, Andreesen Horowitz, DFJ: Coinbase funding (75 M$)
- Nasdaq: company-wide utilization of blockchain technology
- Citigroup: own coin development

Cryptocurrencies - Facts

- As of 20151011, CoinMarketCap.com
  - 636 cryptos
  - 2,022 exchange pairs
  - Market Cap 4.1 billion USD
- Barely derivatives
- Commodity Futures Trading Commission (USA)
  - Cryptos are commodities

Challenges

- What is the benchmark?
- How does the market evolve?
- Market index necessary to compare cryptos

What is the benchmark?

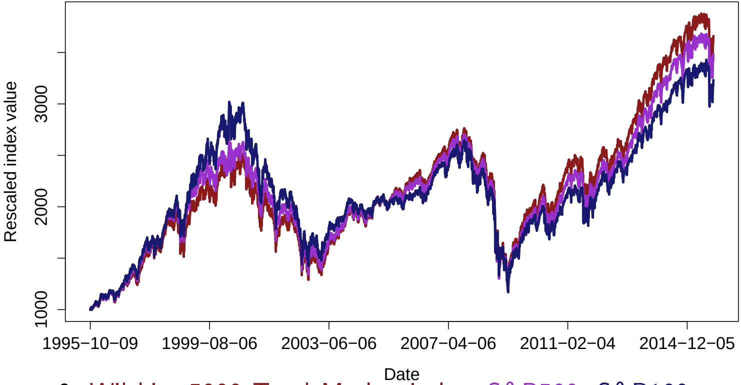


Figure 3: Wilshire 5000 Total Market Index, S&P500, S&P100, rescaled to a starting value of 1000 ▸ CRIX rules

CRIX - the benchmark

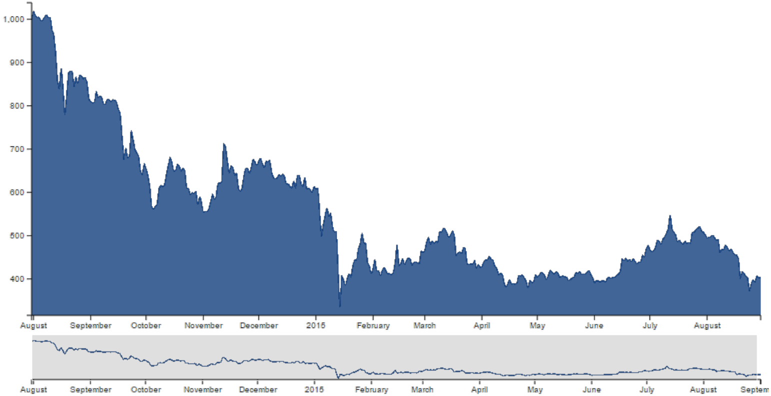


Figure 4: Screenshot: crix.hu-berlin.de (20150901)

Outline

1. Motivation ✓
2. Market Index - CRIX
3. CRIX family comparison
4. Simulation Study
5. Application to german stock market
6. Conclusion
7. Appendix

Data

- 194 cryptos
- Time period: 20140101 - 20150901
- Prices, capitalization, volumes
- CoinGecko

▸ Bitcoin

CRIX - Rules I

- Laspeyres' idea:

$$\text{CRIX}(k)_t = \frac{\sum_{j=1}^{k} MV_{jt}}{Divisor}$$

- MV_{jt}: market capitalization of crypto j
- k: number of constituents
- Liquidity rule:
  - Eligible if higher rank than 0.25 percentile
  - Measure regarding daily volume in USD and coins

Add. Rules | Liquidity Rule | Unused Bitcoins

CRIX - Rules II

- Spine
  - Index members
  - Crucial for benchmark fit

$$CRIX(k)_t \overset{\min(k)}{\longrightarrow} \text{total market}_t$$

  - total $\text{market}_t = \frac{\sum_{j=1}^{J} MV_{jt}}{Divisor}$
  - Here: $J = 194$
- Quadratic loss function
- Sparse benchmark

CRIX - Rules III

1. Construct total market index: total market$_t = \frac{\sum_{j=1}^{194} MV_{jt}}{Divisor}$
2. Set $i = 1$
3. Construct $CRIX(k_i)$, $i = 1, 2, 3, \ldots,$ $k_1 < k_2 < k_3 < \cdots$
4. Compute $\varepsilon_{it} =$ total market$_t - CRIX(k_i)_t$
5. Kernel density estimation for density $f(\varepsilon_{it})$ with leave-one-out cross validation
6. Derive $BIC_i = -2 \log \prod_{t=1}^{n} f(\varepsilon_{it}) + k_i \log(n)$
7. If $i = 1$: Jump to 3., else 8.
8. If $BIC_{i-1} < BIC_i$: stop, else jump to 3. and $i = i + 1$

▸ KDE ▸ US indices

CRIX family

- CRIX
  - $k_1 = 5$
  - Step width: 5 constituents
  - Local optimum
- ECRIX
  - $k_1 = 1$
  - Step width: 1 constituents
  - Local optimum
- EFCRIX
  - $k_1 = 1$
  - Step width: 1 constituents
  - Optimum

Index members

- Compare last 3 M
- Amount used for next 3 M

Period	CRIX	ECRIX	EFCRIX	Maximum achievable
1	5	42	45	45
2	15	10	120	120
3	35	4	57	171
4	30	4	134	194

Table 1: Number of constituents in respective periods

Index Members

CRIX performance

Figure 5: CRIX CRIXindex

ECRIX performance

Figure 6: ECRIX ECRIXindex

EFCRIX performance

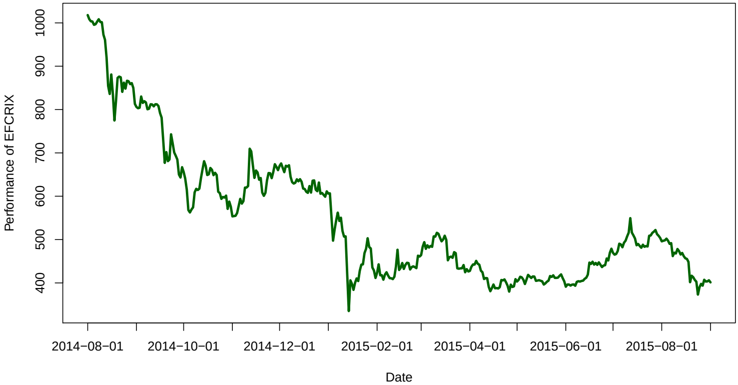


Figure 7: EFCRIX EFCRIXindex

Loss comparison I

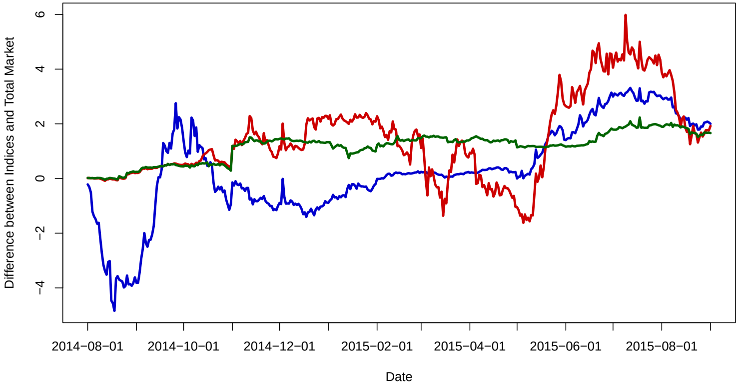


Figure 8: Realized difference between total market and CRIX, ECRIX, EFCRIX CRIXfamdiff

Loss comparison II

	MSE	MDA
CRIX vs. Total Market	3.0441	0.9949
ECRIX vs. Total Market	4.4467	0.9975
EFCRIX vs. Total Market	1.6541	1.0000

Table 2: Comparison of CRIX, ECRIX, EFCRIX against total market CRIXfamdiffloss

Simulation I

- 300 simulated time series
- Prices log normal distributed
- 3 groups of 100 time series each
- Number of shares/coins constant over time, simulated with uniform distribution
- BIC computation quarterly
- Index members exchange quarterly

Simulation II

- σ_{ij}^2 variance in period i, group j
- 5 periods
  1. 12 month
     - $\sigma_{11}^2 = 0.005, \sigma_{12}^2 = 0.01, \sigma_{13}^2 = 0.015$
  2. 3 month
     - $\sigma_{21}^2 = \sqrt{0.005}, \sigma_{22}^2 = \sqrt{0.01}, \sigma_{23}^2 = \sqrt{0.015}$
  3. 6 month
     - $\sigma_{31}^2 = 0.005, \sigma_{32}^2 = 0.01, \sigma_{33}^2 = 0.015$
  4. 3 month
     - $\sigma_{41}^2 = \sqrt{0.005}, \sigma_{42}^2 = \sqrt{0.01}, \sigma_{43}^2 = \sqrt{0.015}$
  5. 6 month
     - $\sigma_{51}^2 = 0.005, \sigma_{52}^2 = 0.01, \sigma_{53}^2 = 0.015$

Simulation III

Period	Decision period	Applied period
1	calm	calm
2	calm	calm
3	calm	turbulent
4	turbulent	calm
5	calm	calm
6	calm	turbulent
7	turbulent	calm
8	calm	calm

Table 3: Behavior of market in the periods and the behavior of the market to which the amount of constituents is applied

▸ Empirical Quantiles

Simulation IV

Period	Mean members
1	14.62
2	15.16
3	13.14
4	13.79
5	17.73
6	15.76
7	16.03
8	19.79

Table 4: Mean number of constituents in respective periods for simulated time series

Simulation V

- Less index members in more equally divided market
- Market dominator causes number of index members to rise
- In calmer market environment tendency to more index members

CRIX methdodology & German stock market

- German Prime Standard
- Basis for DAX, MDAX, SDAX, TecDAX
- DAX often interpreted as market indicator
- DAXCRIX
  - CRIX methodology applied to Prime Standard companies
  - Time period: 20040101 - 20130430
  - Yearly constituent list of Prime Standard
  - BIC computation yearly
  - Index members exchange quarterly

Index members DAXCRIX

Period	DAXCRIX	DAX	Maximum achievable
1	20	30	390
2	15	30	339
3	25	30	310
4	5	30	329
5	30	30	328
6	30	30	285
7	45	30	245
8	20	30	172
9	30	30	223

Table 5: Number of constituents in respective periods for DAX and DAX-CRIX

Loss comparison DAX & DAXCRIX

	MSE	MDA
DAXCRIX vs. Total Market	2612.62	0.94
DAX vs. Total Market	21148.60	0.78

Table 6: Comparison of DAX with CRIX methodology (DAXCRIX) and rescaled DAX against total market DAXCRIXloss

Conclusion

- CRIX represents market very good
- EFCRIX best but too many constituents
- Choice of CRIX well in terms of MSE and MDA
- Methodology enhances fit to German stock market
  - But strategy may cause high transaction costs
  - Use analysis as lower bound of index members

CRIX - a CRyptocurrency IndeX

Simon Trimborn

Wolfgang Karl Härdle

Ladislaus von Bortkiewicz Chair of Statistics

C.A.S.E. – Center for Applied Statistics and Economics

Humboldt–Universität zu Berlin

http://lvb.wiwi.hu-berlin.de

http://www.case.hu-berlin.de

Bibliography

Simon Trimborn and Wolfgang Karl Härdle (2015)
CRIX or evaluating Blockchain based currencies
Oberwolfach Report No. 42/2015 "The Mathematics and Statistics of Quantitative Risk"
DOI: 10.4171/OWR/2015/42

Bitcoin

- Counteract inflation
  - Fixed max amount
- Anonymity
- Needs of users
  - Decision on structure

▸ Movie: Bitcoin - Made simple

Anonymity - Black market

- Wallets are anonym
- Transactions are anonym
- Blockchain core feature
- Causes problems

Figure 9: US government warning

Source: www.wikipedia.org

The Blockchain - Spine of Bitcoin

- Transaction list
- Transaction processors called miners
- Miners collect & publish transactions
- Order is invariable

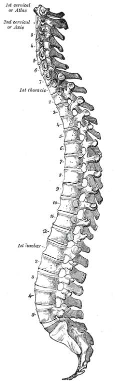


Figure 10: Spine

The Blockchain

- Sometimes parallel chains
  - Due to e.g. internet lag
- Green block: first block (Genesisblock)
- Black blocks: main chain
- Purple blocks: parallel chains

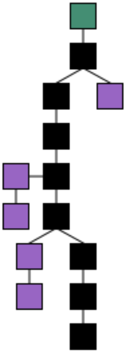


Figure 11: Blockchain

Source: www.wikipedia.de

The Blockchain - Lag

- Assume: 2 blocks mined simultaneously
  - Miner 1: Australia
  - Miner 2: Germany
- Effect of lag:
  - Some receive australian block
  - Some receive german block
- Parallel chain
- For next block:
  - Check which chain contains the most difficult to find blocks
  - Becomes main chain

Process of Transactions

- Users organize process
- Some users (miners) create transaction list
  - Next block of blockchain
- Blocks have strict order, ensured by signature
- Miners search for signature
- Signature encrypted by cryptography

Transaction example

Who accepts Bitcoin?

- Overstock - Retailer
- Dell
- University of Cumbria
- Expedia - Travel booking agency
- Republican Party of Louisiana

back: Index Construction

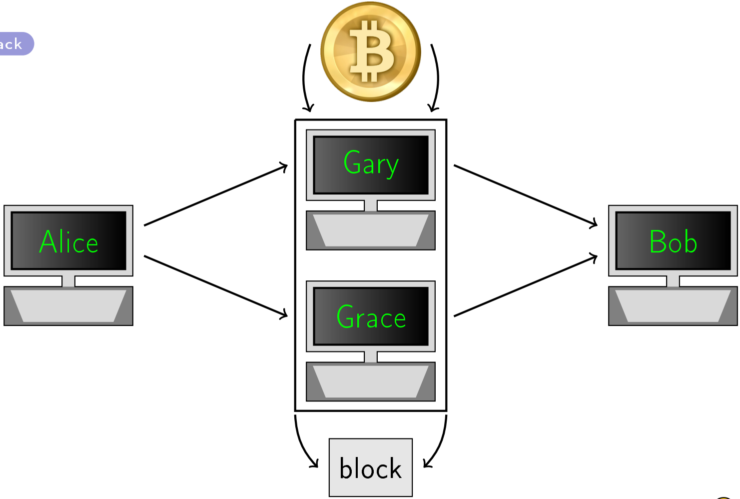
Bitcoin - The System I

- Take 4 people
  - Alice, Bob
  - Gary, Grace
- 2 special users (miners)
  - Gary
  - Grace
- Alice buys a rare book from Bob and pays with Bitcoin
- Gary and Grace process this transaction

back

Bitcoin - The System II

back



Bitcoin - The System III

- Gary OR Grace receive Bitcoins for service
- BOTH add transaction to list
- BOTH compute hash value (trial and error)
- Click for online hash generator
- List: one block of the blockchain
- Hash value: gives position of block in blockchain
- Contains part of hash value of last block

back

CRIX - Add. Rules

- High volatility: weights recalculated 1 M
- Maximum weight of CRIX member: 50 %
- Crypto made insensitive if trading stops

Back

Liquidity Rule I

- Rely often on turnover

$$Turnover = \frac{Volume}{Floating\ \ Coins}$$

- Floating Coins for cryptos unclear
- Rule motivated by STOXX Japan 600 & AEX Family
- Measure relative to asset universe
- Small trading volume in USD but high traded coins taken into account

back

Liquidity Rule II

- Liquidity rule (one of these):
  1. 0.25 percentile of ADTV (Average Daily Trading Volume):

$$ADTV_i \geq ADTV_{0.25}$$

  2. 0.25 percentile of ADRTC (Average Daily Relative Traded Coins):

$$ADRTC_i \geq ADRTC_{0.25}$$

- Checked every month

back

Usage of Bitcoins

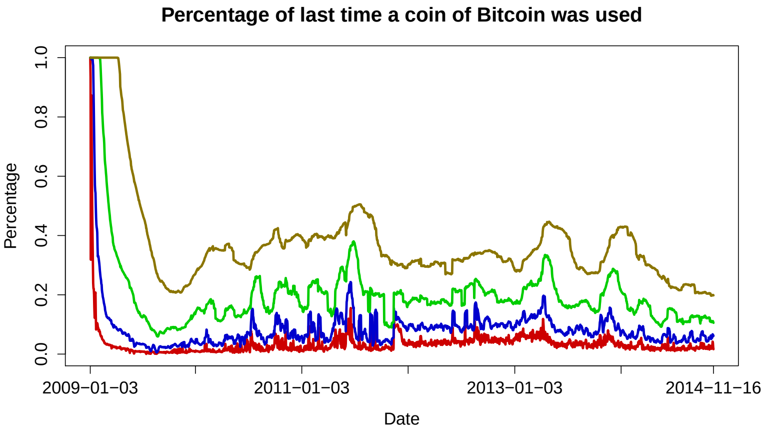


Figure 12: one day, one week, one month, 1-3 month source: John Radcliff

back

Methodology

- GARCH(1,1)

$$\sigma_t^2 = \alpha_0 + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

- Full variance for time horizon T

$$\sigma_T^2 = \sum_{t=1}^{T} \sigma_t^2$$

back

GARCH parameters

	alpha0	alpha	beta
CRIX	0.00007572	0.08457491	0.88392853
DAX	0.00000257	0.09926597	0.89304179

Table 7: Parameters of the GARCH(1,1) models without mean for CRIX and DAX

back

Rolling window GARCH parameters

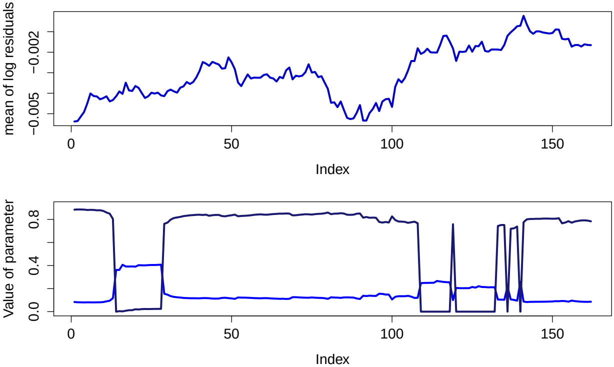


Figure 13: Upper plot: Mean of CRIX log residuals in 250 data windows, Lower plot: α, β GARCH(1,1) parameters for 250 data points rolling windows

back

Index members

Period	1	2	3	4	5	6	7	8	9	10
1	btc	ltc	msc	nxt	nmc	doge	mec	wdc	ftc	zet
2	btc	xrp	ltc	ppc	doge	nxt	msc	nmc	pts	qrk
3	btc	xrp	ltc	ppc	doge	pts	xpm	mec	vtc	ybc
4	slr	blk	pts	hbn	pot	prt	efl	zeit	cbx	fair
5	slr	pts	safe	rdd	prt	grcx	pnd	cbx	karm	mona
6	btc	ltc	nxt	ppc	xc	zet	ftc	mec	ifc	ybc
7	btc	ltc	nxt	xrp	ppc	dash	doge	nmc	msc	blk
8	btc	ltc	xrp	bts	nxt	ppc	dash	doge	nmc	safe
9	btc	ltc	xrp	bts	nxt	doge	ppc	dash	nmc	safe
10	btc	xrp	ltc	bts	doge	nxt	ppc	dash	nmc	safe
11	btc	xrp	ltc	bts	doge	nxt	ppc	xcp	safe	dash
12	btc	xrp	ltc	bts	safe	doge	nxt	str	ppc	xcp
13	btc	mtc	xrp	ltc	xpy	bts	safe	str	doge	nxt
14	btc	xrp	ltc	bts	doge	sync	safe	dash	xpy	str
15	btc	xrp	ltc	bts	dash	doge	nxt	safe	str	xpy
16	btc	xrp	ltc	dash	bts	str	doge	safe	nxt	banx
17	btc	xrp	ltc	dash	str	doge	bts	nxt	safe	banx
18	btc	xrp	ltc	bts	doge	str	dash	nxt	banx	ppc

back

Table 8: First 10 CRIX constituents in the respective periods

Heston Nandi GARCH(1,1) Option Pricing model for European Options

- Log returns X_t follow process:

$$X_t = r_f + (\mu - \frac{1}{2})\sigma_t^2 + \sigma_t Z_t \tag{1}$$

  with r_f risk free rate, $Z_t \sim N(0,1)$, μ the mean and σ_t^2 a GARCH(1,1) process.

- GARCH(1,1) process:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha(Z_{t-1} - \gamma\sigma_{t-1})^2 \tag{2}$$

back

Risk neutrality

- Transform (1) and (2)
- Receive risk neutral versions

$$X_t = r_f + (\mu - \frac{1}{2})\sigma_t^2 + \sigma_t Z_t^*$$

and

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha(Z_{t-1}^* - \gamma^*\sigma_{t-1})^2$$

with $Z_t^* = Z_t + \mu\sigma_t$ and $\gamma^* = \gamma + \mu$.

back

Log-linear generator function

$$f(\theta) = S_t^{\theta} \exp\left(A_{t;T,\theta} + B_{t;T,\theta}\sigma_{t+1}^2\right)$$

where

$$A_{t;T,\theta} = A_{t+1;T,\theta} + \theta r_f + B_{t+1;T,\theta}\omega - \frac{1}{2}\log(1 - 2\alpha B_{t+1,T,\theta})$$

$$B_{t;T,\theta} = \theta(\mu - \frac{1}{2} + \gamma) - \frac{1}{2}\gamma^2 + \beta B_{t+1;T,\theta} + \frac{1/2(\theta - \gamma)^2}{1 - 2\alpha B_{t+1;T,\theta}}$$

with terminal conditions

$$A_{T;T,\theta} = 0$$

$$B_{T;T,\theta} = 0$$

back

Heston Nandi Options

- European Call option

$$
\begin{aligned}
C_t = {} & \frac{1}{2} S_t \\
& + \frac{\exp\{-r(T-t)\}}{\pi} \int_0^\infty \Re\left\{ \frac{K^{-i\theta} f^*(i\theta+1)}{i\theta} \right\} d\theta \\
& - K\exp\{-r(T-t)\}\left[\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re\left\{ \frac{K^{-i\theta} f^*(i\theta)}{i\theta} \right\} d\theta \right]
\end{aligned}
$$

where C Call price, S_t asset price, K strike price and $\Re$ real part of a complex number. f^* the risk neutral version of f.

- Put-Call parity for Put price:

$$
P_t = C_t + K\exp\{-r(T-t)\} - S_t
$$

back

Kernel Density Estimation (KDE)

- Compute pdf with KDE

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

  with $K(u)$ Gaussian kernel, h bandwidth

- Bandwidth selection with Wand & Jones plug-in estimator

back

Simulation Study I

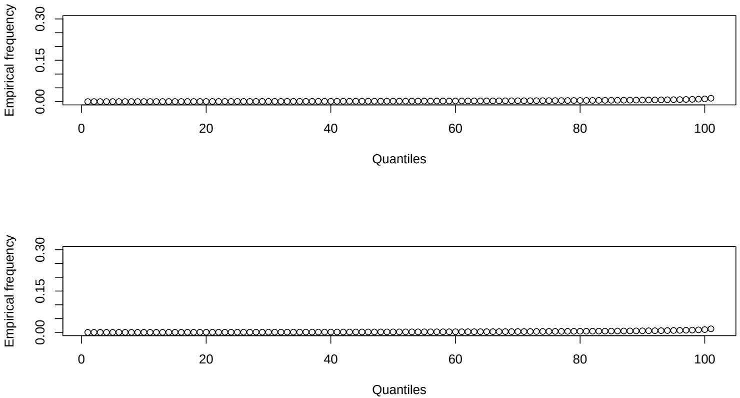


Figure 14: Empirical Average Quantiles of the simulations in the first and second period ▸ Back

Simulation Study II

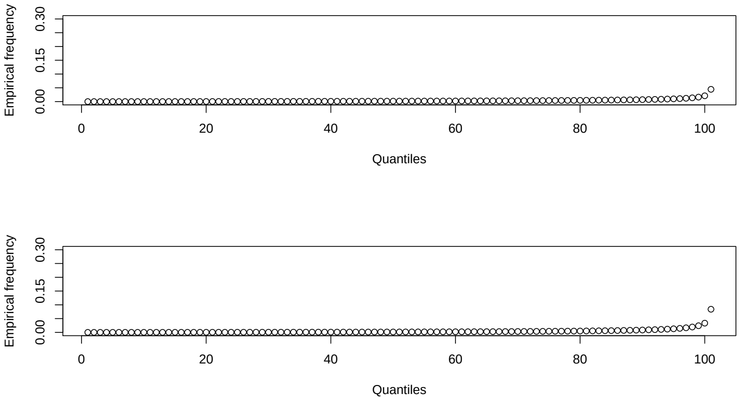
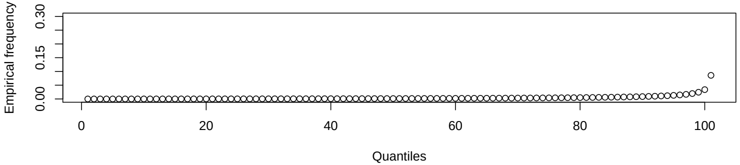


Figure 15: Empirical Average Quantiles of the simulations in the third and fourth period Back

Simulation Study III



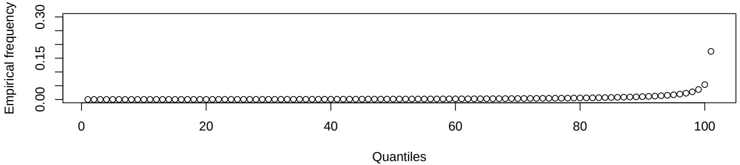


Figure 16: Empirical Average Quantiles of the simulations in the fifth and sixth period Back

Simulation Study IV

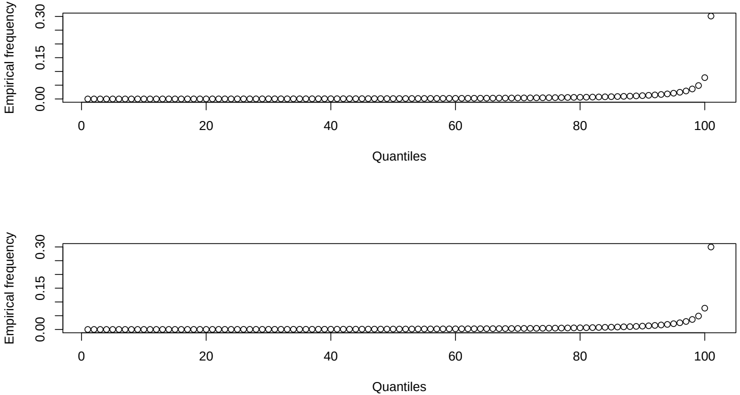


Figure 17: Empirical Average Quantiles of the simulations in the seventh and eight period Back

Attributions

- Cigarette trading in postwar germany, Bundesarchiv, Bild 183-R79014 / CC-BY-SA
- By The original uploader was DickClarkMises at English Wikipedia [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons