Quantitative Fairness

Mahmood Hikmet, Partha Roop, Prakash Ranjitkar
University of Auckland

THE UNIVERSITY OF AUCKLAND NEW ZEALAND

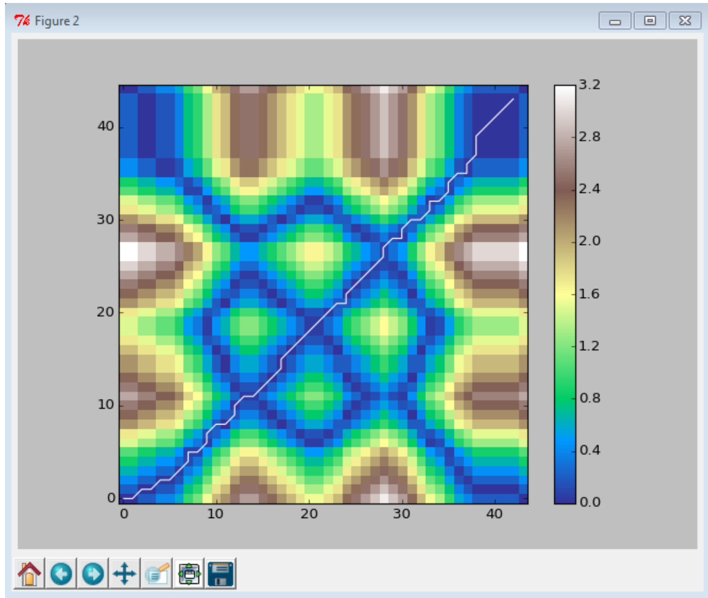
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

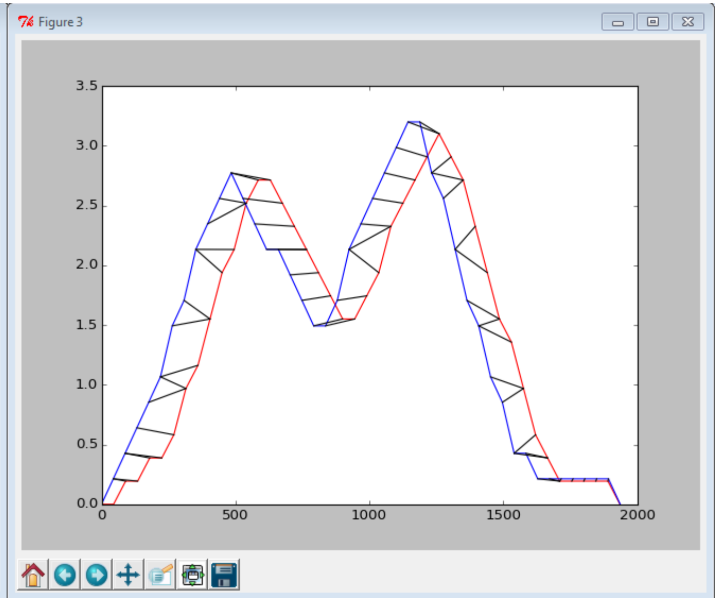
Who Am I?

- Mahmood Hikmet – 26 years old
- Born in Iraq in 1989 – left during the Gulf War
- Living in New Zealand since 1996
- Bachelor of Engineering in Computer Systems Engineering
    - Now Studying PhD
- Interests:
    - Cooking
    - Baking
    - Brewing Beer
    - Beekeeping
    - Game Development
    - Poetry

Work

- Research and Development Engineer at HMI Technologies
- Projects:
    - Web-based Electronic Road Sign Control using GPRS
    - Bike-Loop: Inductive loop vehicle classification using Speech-Based Algorithms





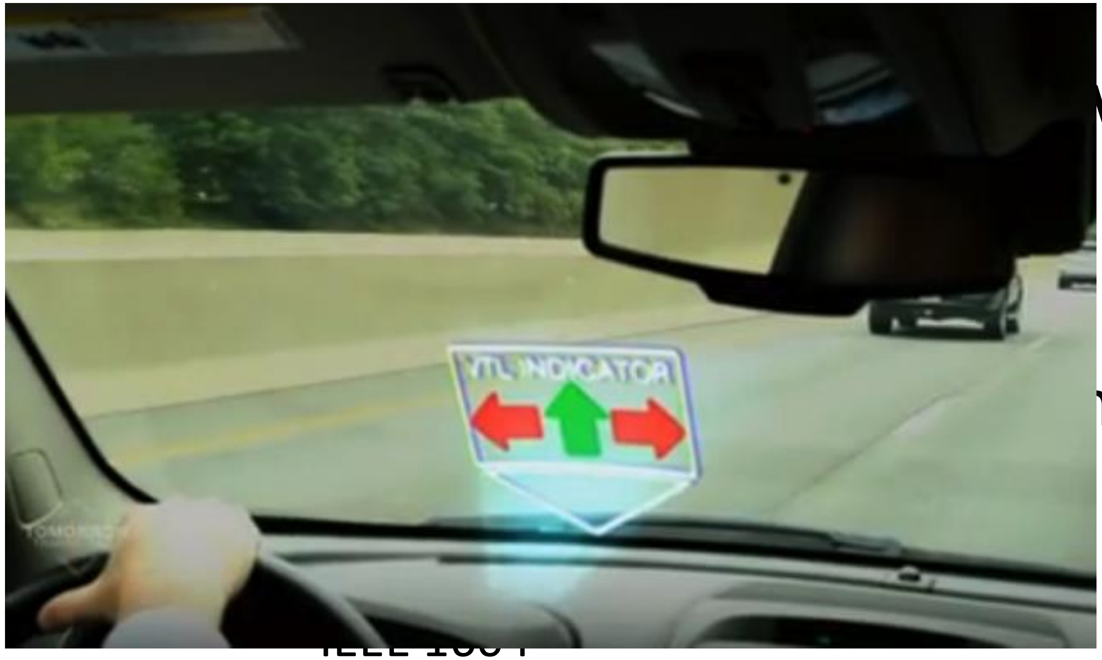
Intelligent Transport Systems

A method by which to intelligently optimise Transportation

Safety

Efficiency

Environmental

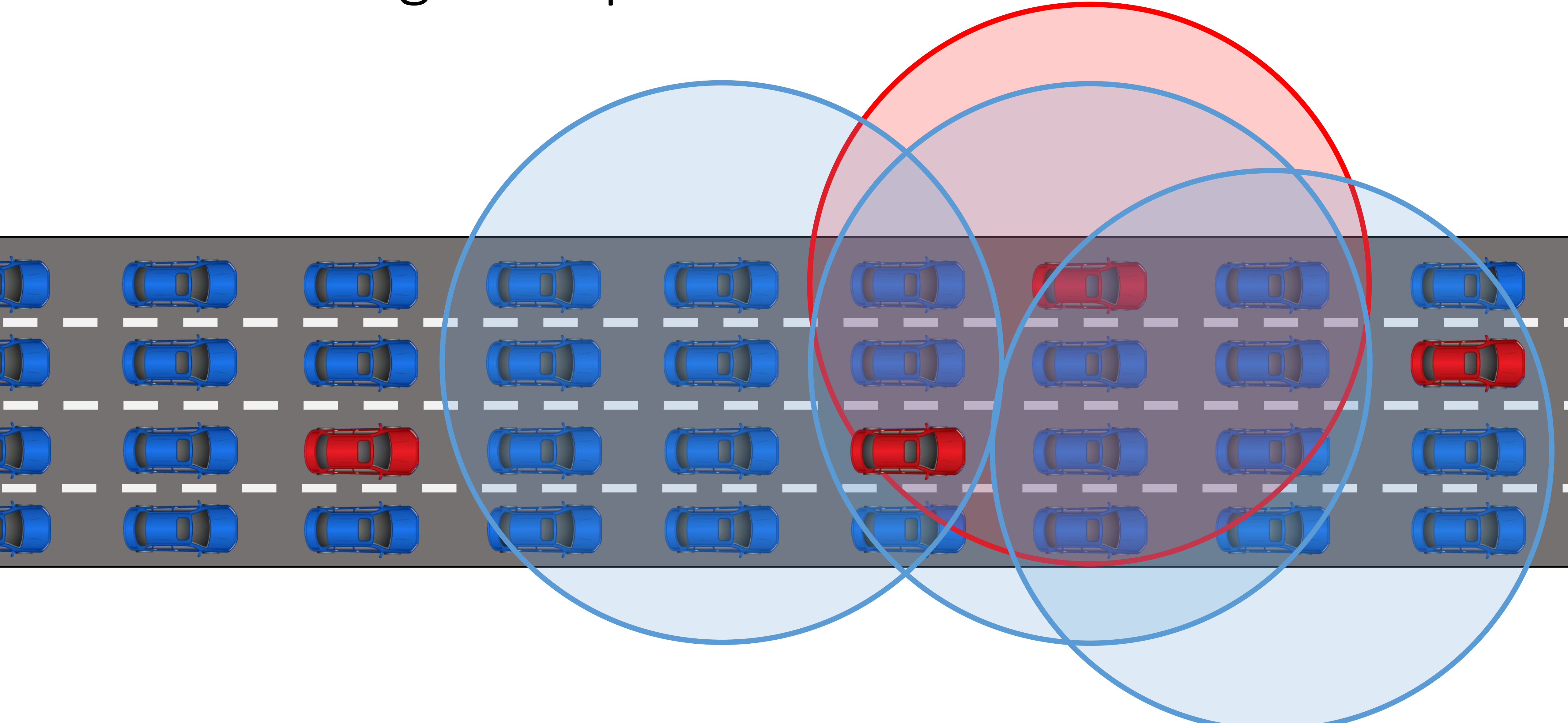

VTL INDICATOR


TOPLINE
SCANIA


Do Not Pass
Warning


COLLISION
ALERT

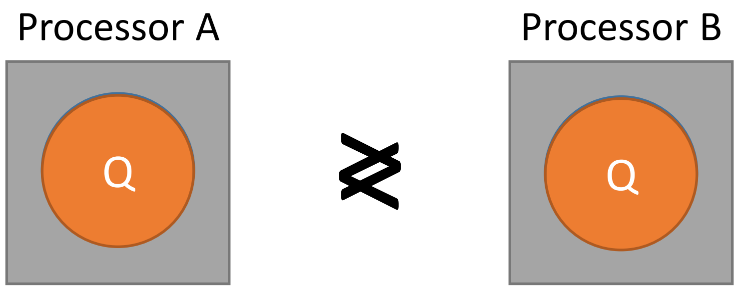
Motivating Example



Access to the Medium

- Data Link Layer
  - MAC protocol
  - CSMA/CA = current standard
- Impossible to know where a safety message will come from
  - Assume that it can come from anywhere
- Prioritised access for safety messages is required
  - What if there many safety messages competing against each other?
- All vehicles require equal access to the medium

The Issue of Fairness



- Multiple papers boast “Better Fairness”
  - Lack of quantifiable measure.
  - Comparisons across research papers become difficult
- Draws many parallels with the issue of finding Time Predictability
  - “Any quantifiable measure for Time Predictability is susceptible to changing based on application, environment and a multitude of other factors”
- Assuming that we concede the above criteria – we can gather an idea of how far particular protocols operate
  - One test will not give the answer, but many tests will give a general idea

Schoeberl, Martin. "Is time predictability quantifiable?." *Embedded Computer Systems (SAMOS), 2012 International Conference on*. IEEE, 2012.

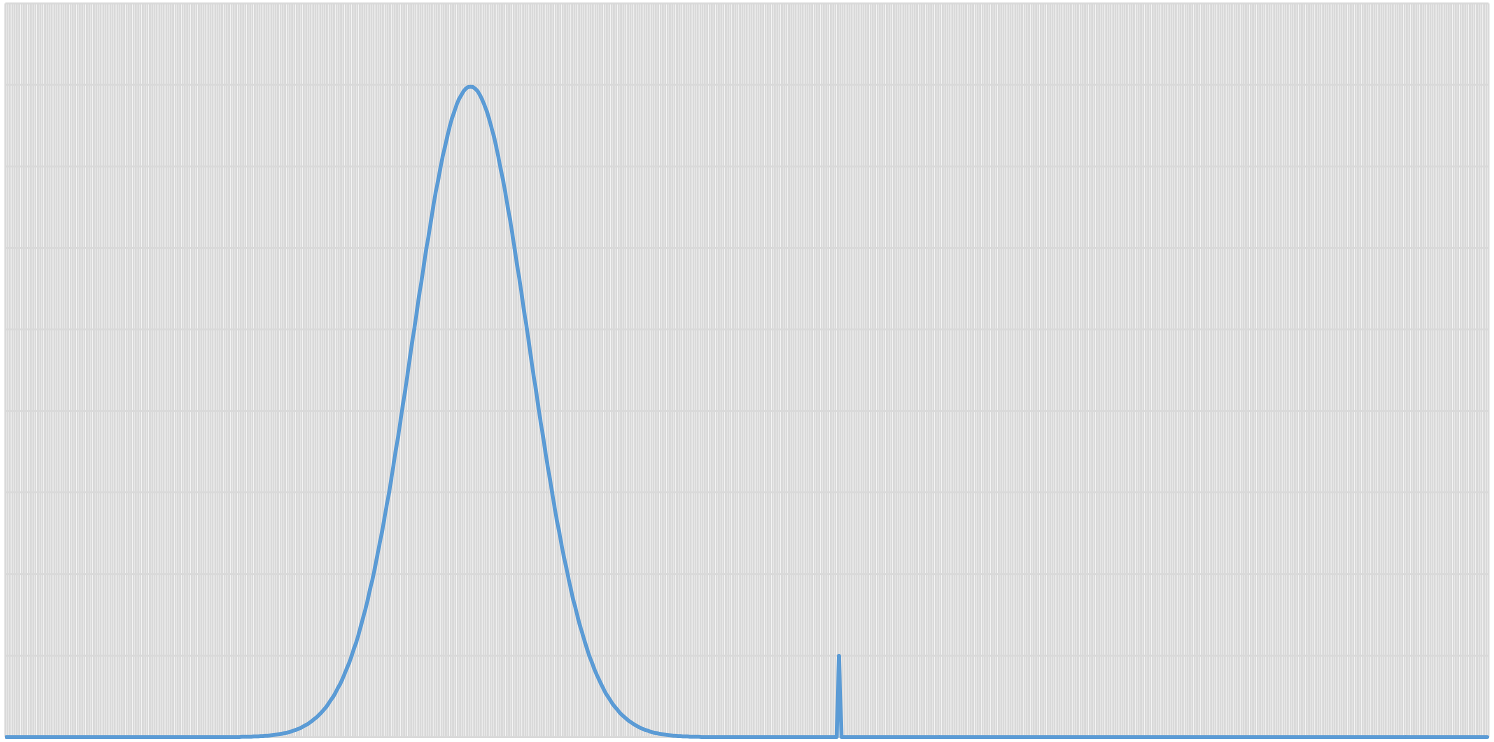
Definition of “Delay”

Within the scope of this research, “delay” will refer to the amount of time between two transmissions (TBT) from a single node

$$TBT = \frac{seconds}{packets}$$

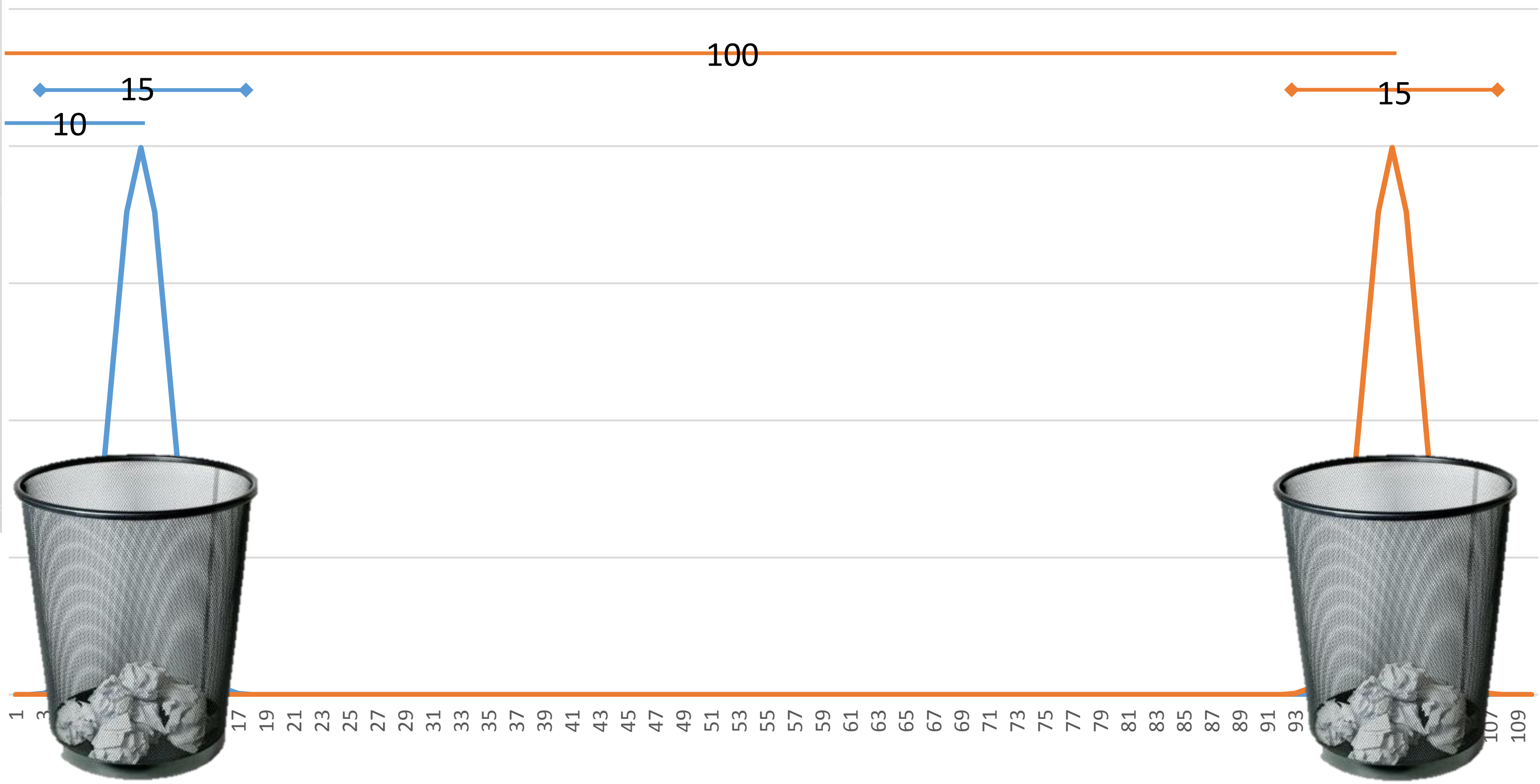
Desired Qualities of a Quantitative Fairness Measure

Sensitivity to Outliers



Desired Qualities of a Quantitative Fairness Measure

Diminishing Sensitivity at Larger Values



Desired Qualities of a Quantitative Fairness Measure

Bounded

$$completely\ fair \ \leq\ x\ \leq completely\ unfair$$

Mean (μ)

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

"Average" Value

Desired Quality	Satisfied?
Sensitivity to Outliers	✗
Diminishing Sensitivity at Larger Values	✗
Bounded	✗

Standard Deviation (σ)

Average Distance from Mean

Desired Quality	**Satisfied?**
Sensitivity to Outliers	~
Diminishing Sensitivity at Larger Values	✗
Bounded	✗

Coefficient of Variation

$$c_v = \frac{\sigma}{\mu}$$

Standard Deviation divided by Mean

Desired Quality	Satisfied?
Sensitivity to Outliers	~
Diminishing Sensitivity at Larger Values	✓
Bounded	✗

Jain Index

$$f(x) = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \sum_{i=1}^{n} x_i^2}$$

Fraction of Population who have received their "Fair Share"

Desired Quality	Satisfied?
Sensitivity to Outliers	✓
Diminishing Sensitivity at Larger Values	✓
Bounded	✓

Jain Index

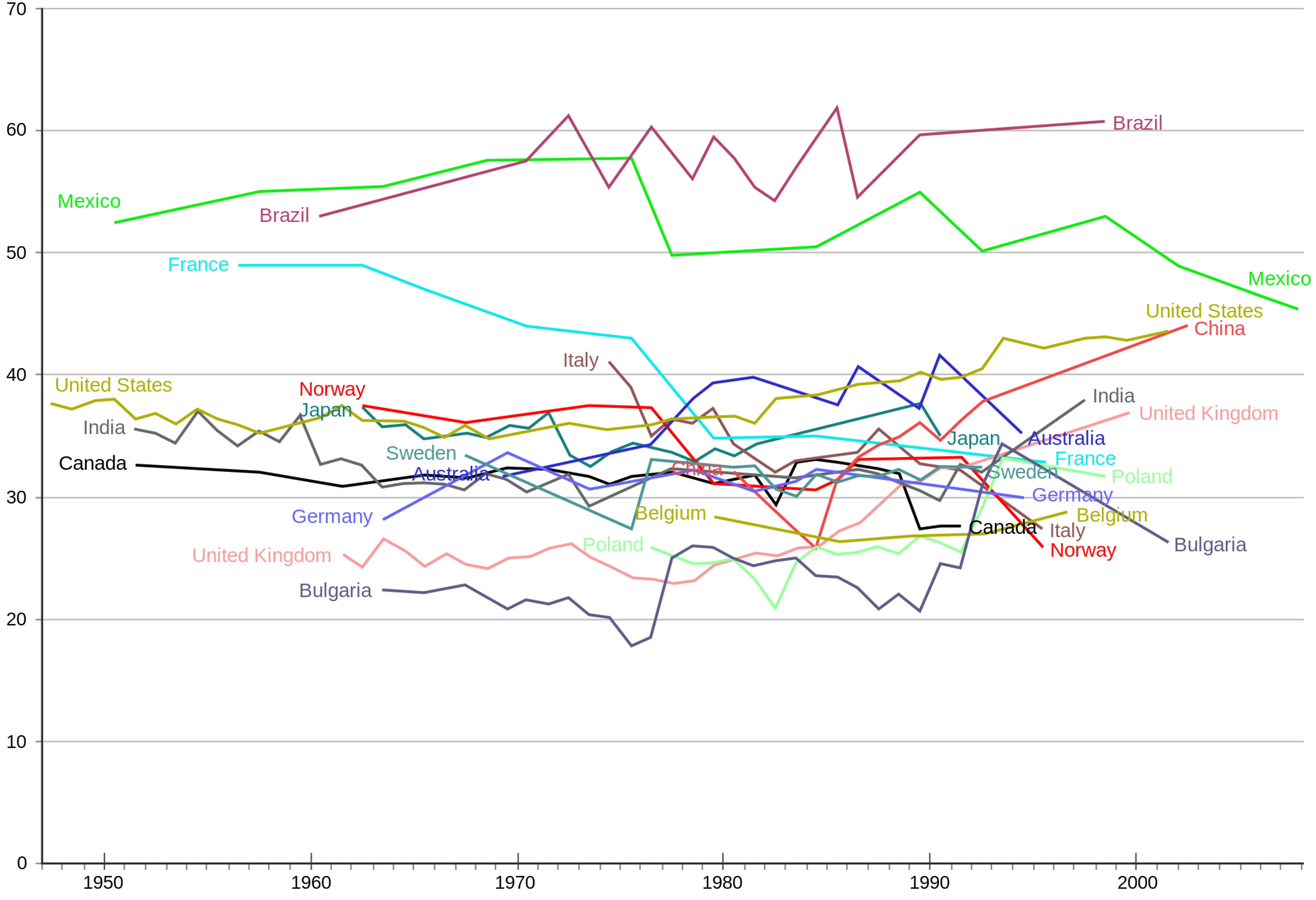
- If Jain Index is 0.2, then 20% of the Population have received their fair share
    - Only holds when there are only 2 different delays suffered by the system
- Consider this case:
    - 20 nodes ($n = 20$)
    - Incremental Delay ($x_1 = 1$, $x_2 = 2$… $x_{20} = 20$)
- Resulting Jain Coefficient is 0.7683
    - 76.8% of our 20 nodes are receiving their fair share
    - 76.8% of 20 is 15.37
- Jain’s Explanation is not always intuitive

Gini Coefficient

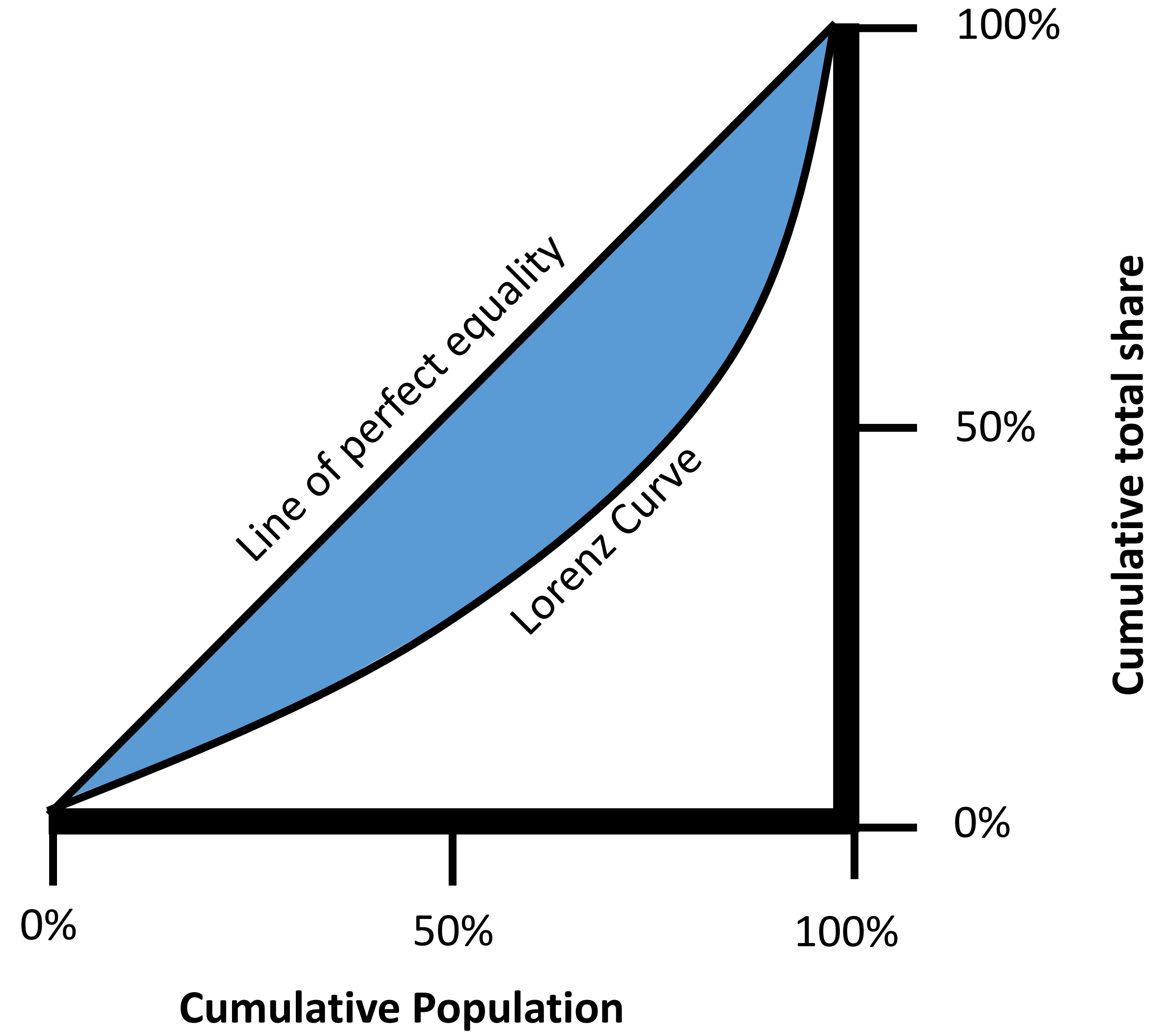
- Developed by Corrado Gini in 1912
  - Italian Statistician/Sociologist
- Used as an indicator of the distribution of wealth within a nation
- Value between 0 $\rightarrow$ 1
  - 1 = Absolutely unfair
  - 0 = Absolutely fair
- Example:
  - Everyone has the same amount of money = 0
  - No one has money except for one person $\approx$ 1

Gini Index - Income Disparity since World War II

where 0 is perfect equality, and 100 is perfect inequality (i.e., one person has all the income)



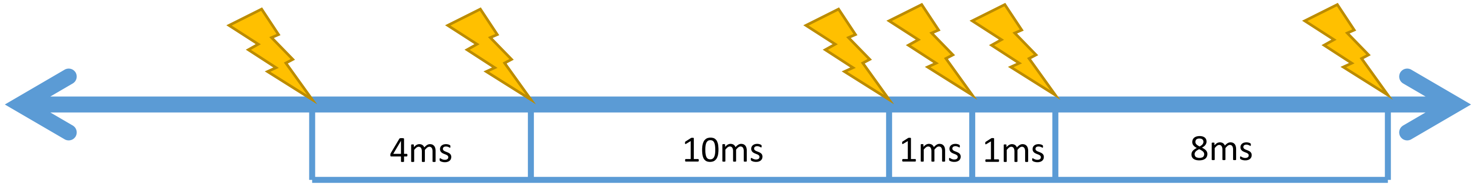
Gini Coefficient



Gini Coefficient

for measuring Distribution of Delay Between Transmissions

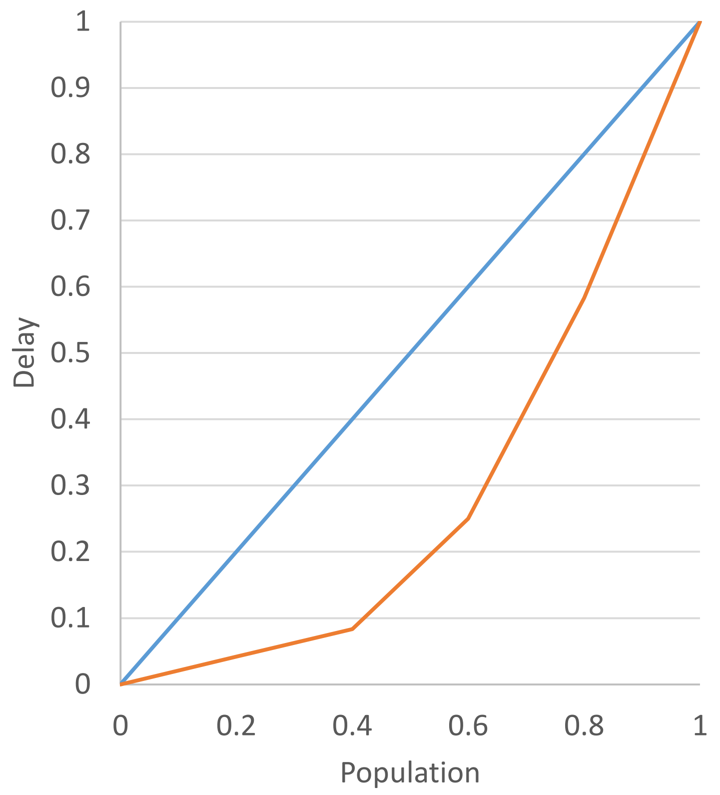
- Rather than population, treat “time between transmissions” as a member of population
- Use time rather than income as measure of “wealth”



Time Between Transmissions	Normalised	Accumulation of Normalised Values
1ms	0.0417	0.0417
1ms	0.0417	0.0833
4ms	0.1667	0.2500
8ms	0.3333	0.5833
10ms	0.4167	1.0000

Gini Coefficient

Time Between Transmissions	Normalised	Accumulation of Normalised Values
1ms	0.0417	0.0417
1ms	0.0417	0.0833
4ms	0.1667	0.2500
8ms	0.3333	0.5833
10ms	0.4167	1.0000



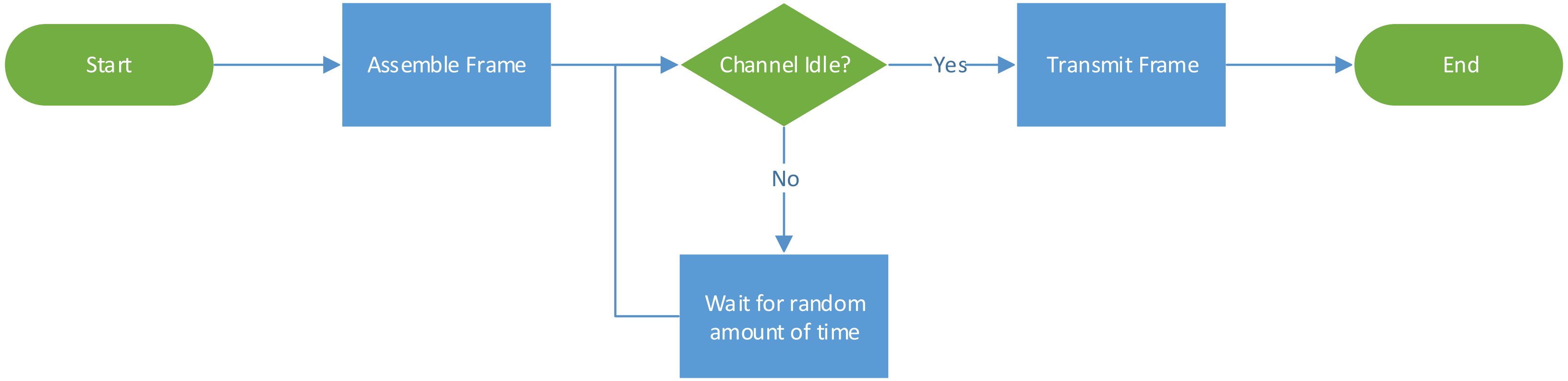
Experimental Set-up

- 600 vehicles spread across 4 lanes
- Each node is loaded with 1000 x 400 Byte packets
- The experiment is run until all packets have been transmitted
- Time between transmissions is recorded for every node
- Vehicle Densities
    - 140 vehicles/lane/km
    - 70 vehicles/lane/km
    - 30 vehicles/lane/km
    - 7 vehicles/lane/km
- MAC Protocols
    - CSMA/CA
    - TDMA
    - STDMA

Assumptions

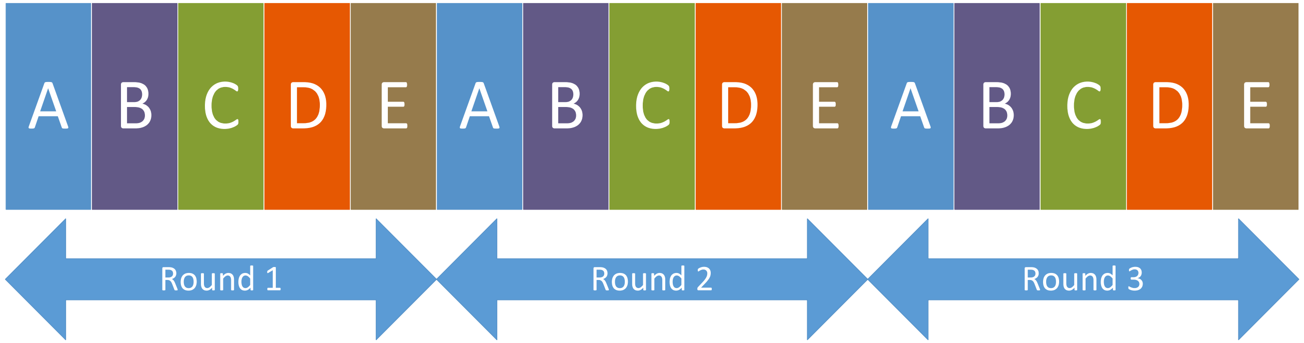
- All nodes are identical in terms of capability
  - Application
  - Networking Layers
  - Devices
  - Priority
  - Load
- No congestion control
- Using DSRC Control Channel for communication
  - 5.9GHz 802.11p WiFi

Carrier-Sensing Multiple-Access with Collision Avoidance (CSMA/CA)

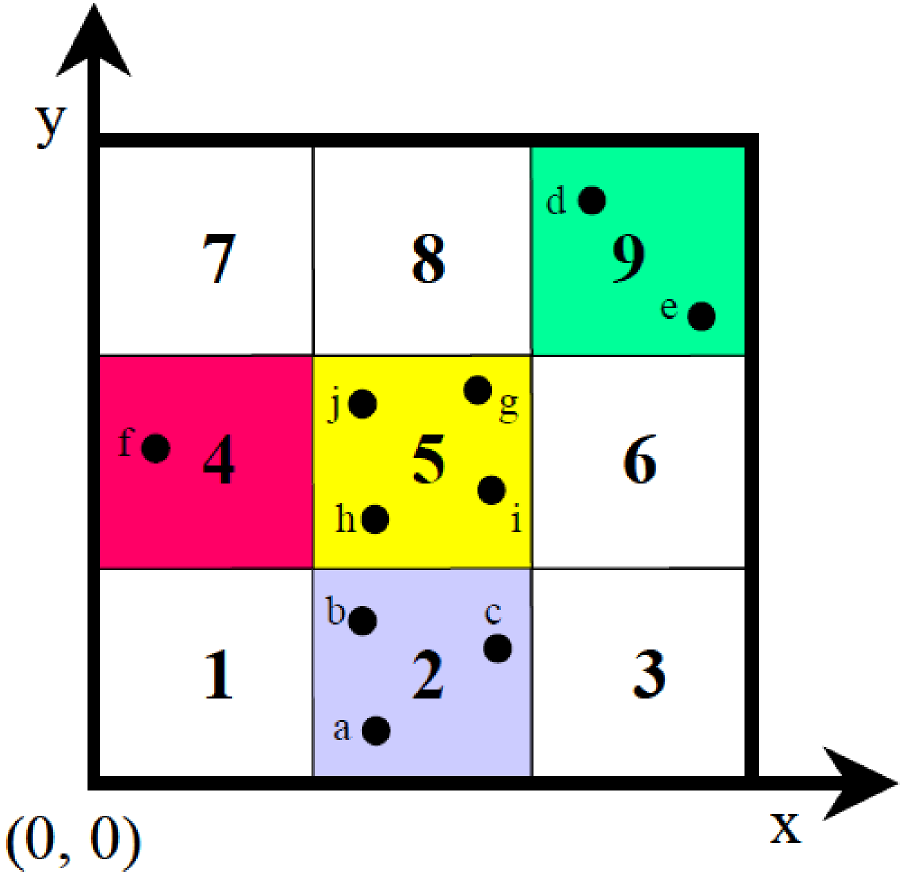


Time-Division Multiple-Access (TDMA)

- All nodes are time-synchronised
- Each node is assigned a slot
  - Node may only transmit during its slot
- All slots together form a “round”



Self-Organising TDMA (STDMA)



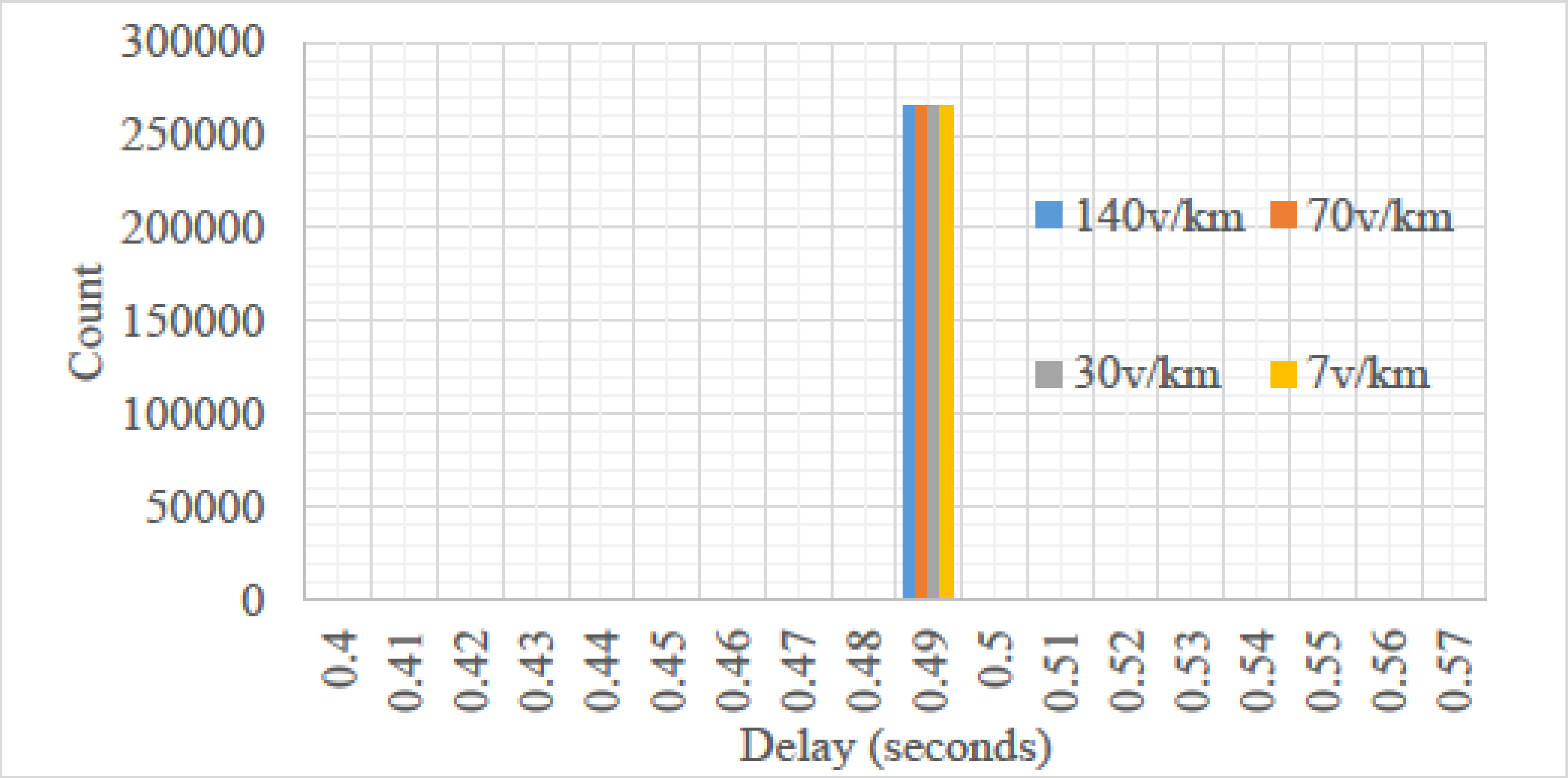
slots

rounds

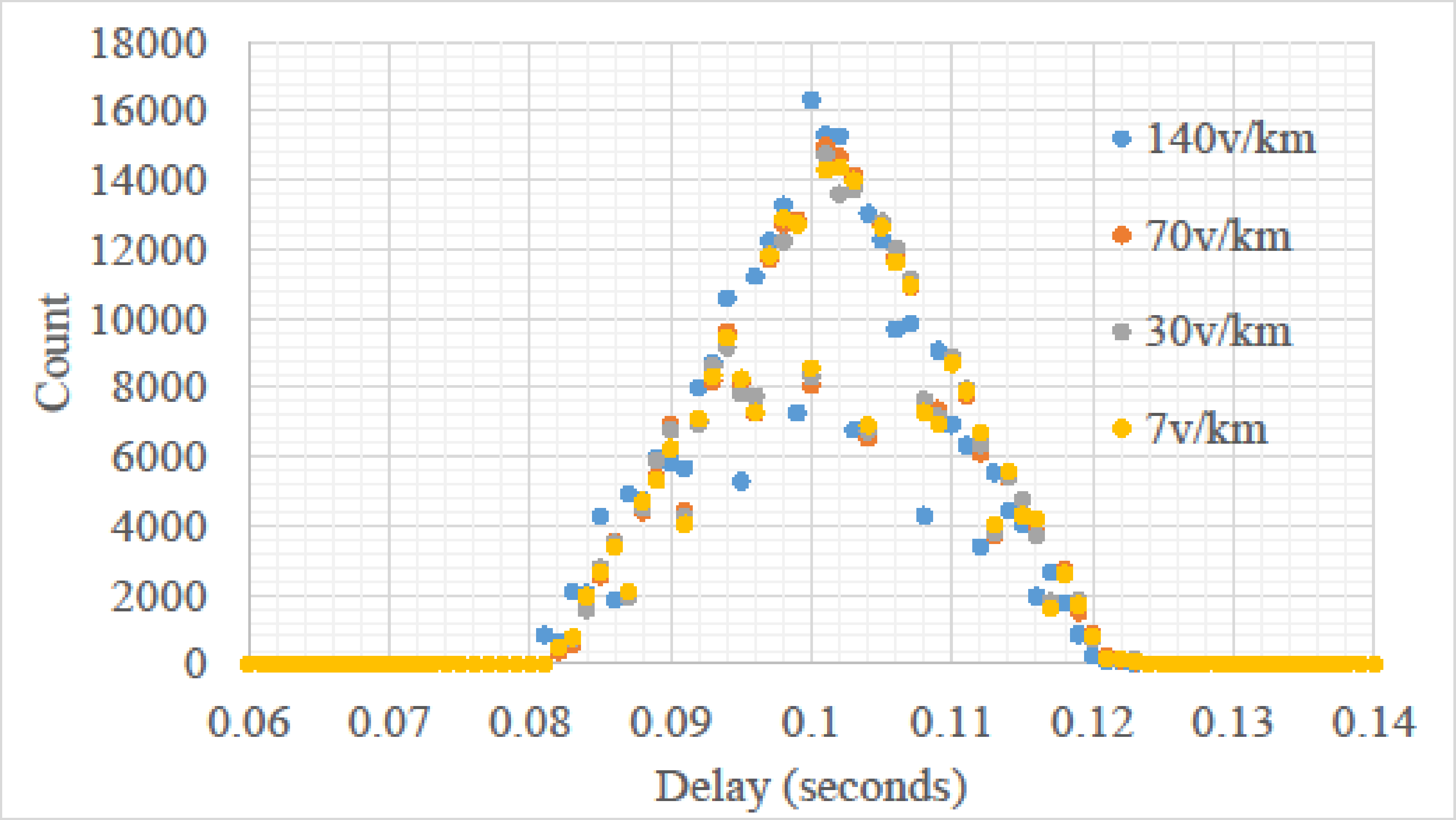
1	c	3	f	g	6	7	8	e
1	a	3	f	h	6	7	8	d
1	b	3	f	i	6	7	8	e
1	c	3	f	j	6	7	8	d
1	a	3	f	g	6	7	8	e
1	b	3	f	h	6	7	8	d
1	c	3	f	i	6	7	8	e
1	a	3	f	j	6	7	8	d

Time Between Transmissions

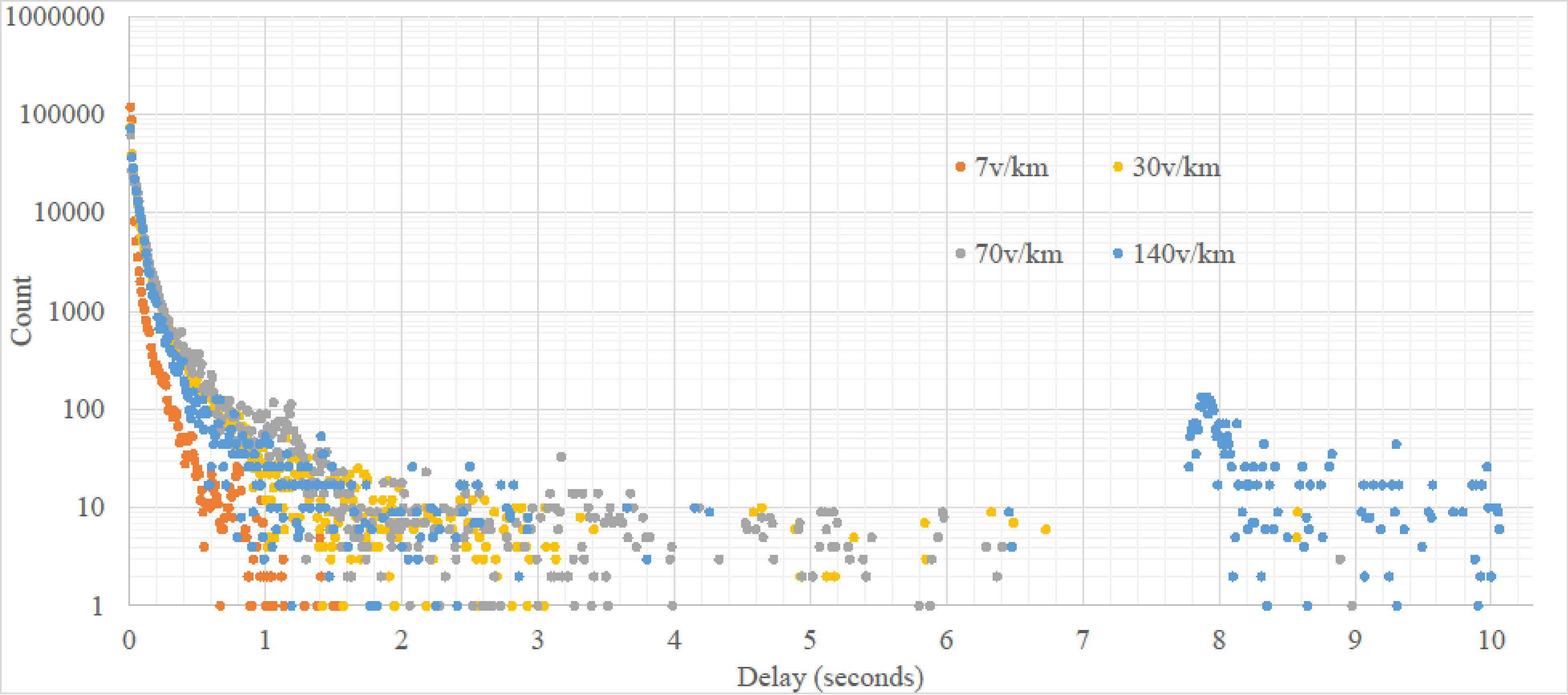
TDMA



Time Between Transmissions
STDMA

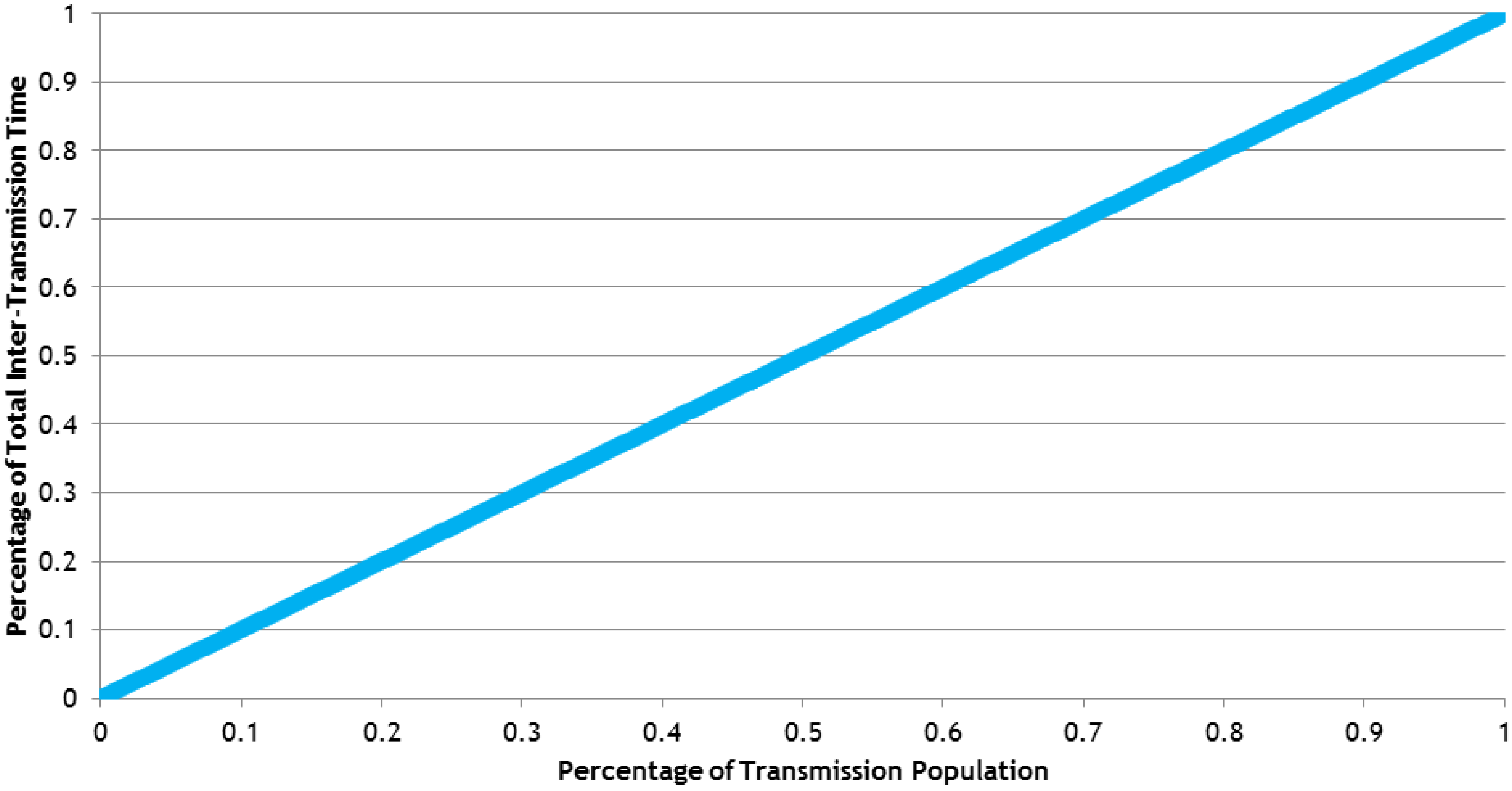


Time Between Transmissions
CSMA/CA



Gini Coefficient

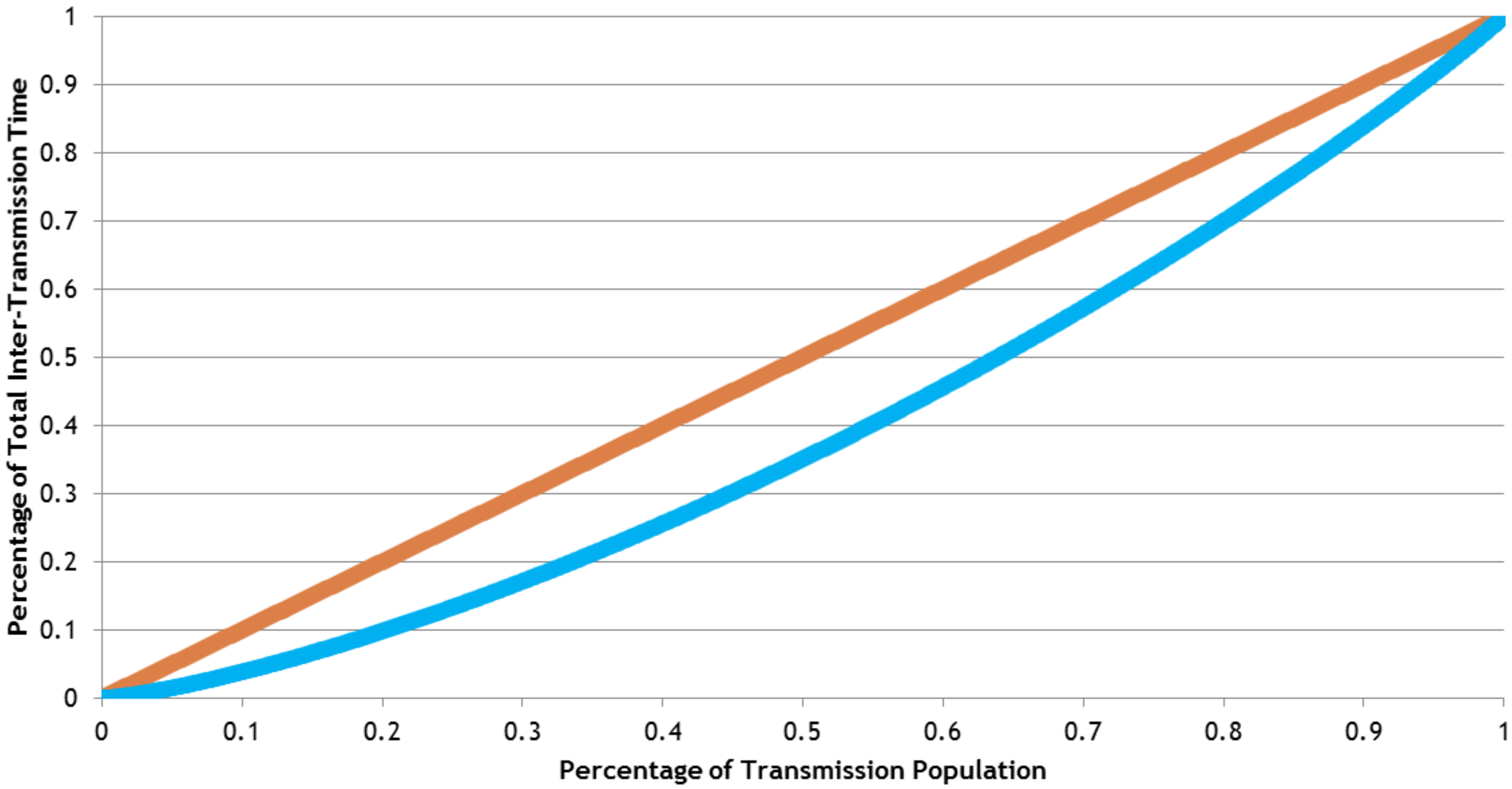
140v/km/lane – 4 lanes – TDMA



Vehicles/km	TDMA		
	Gini	Jain	Jain-1
7	0.0000	1.0000	0.0000
30	0.0000	1.0000	0.0000
70	0.0000	1.0000	0.0000
140	0.0000	1.0000	0.0000

Gini Coefficient

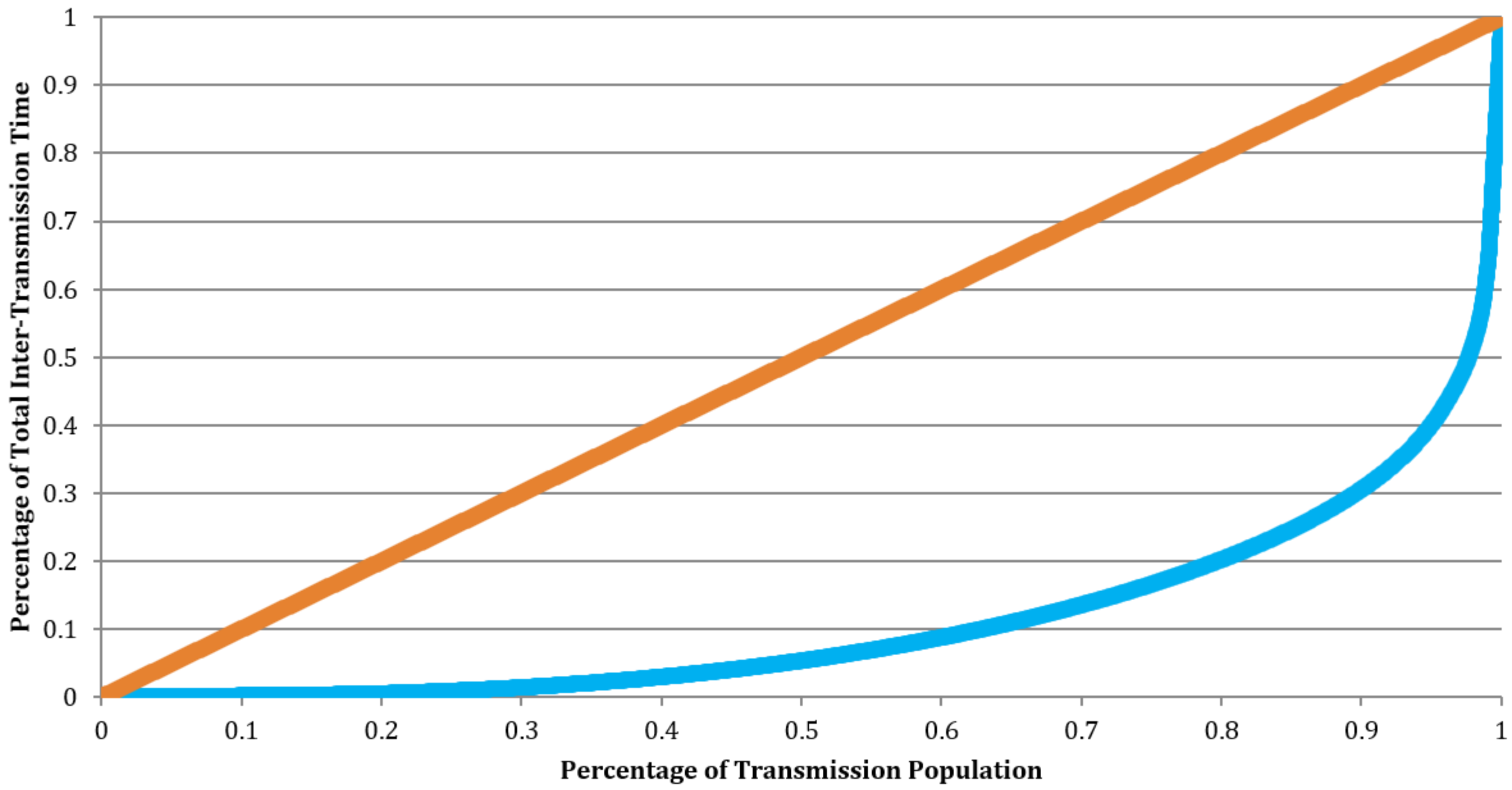
140v/km/lane – 4 lanes – STDMA



Vehicles/km	STDMA		
	Gini	Jain	Jain-1
7	0.0464	0.9934	0.0066
30	0.0463	0.9934	0.0066
70	0.0462	0.9935	0.0065
140	0.0464	0.9935	0.0065

Gini Coefficient

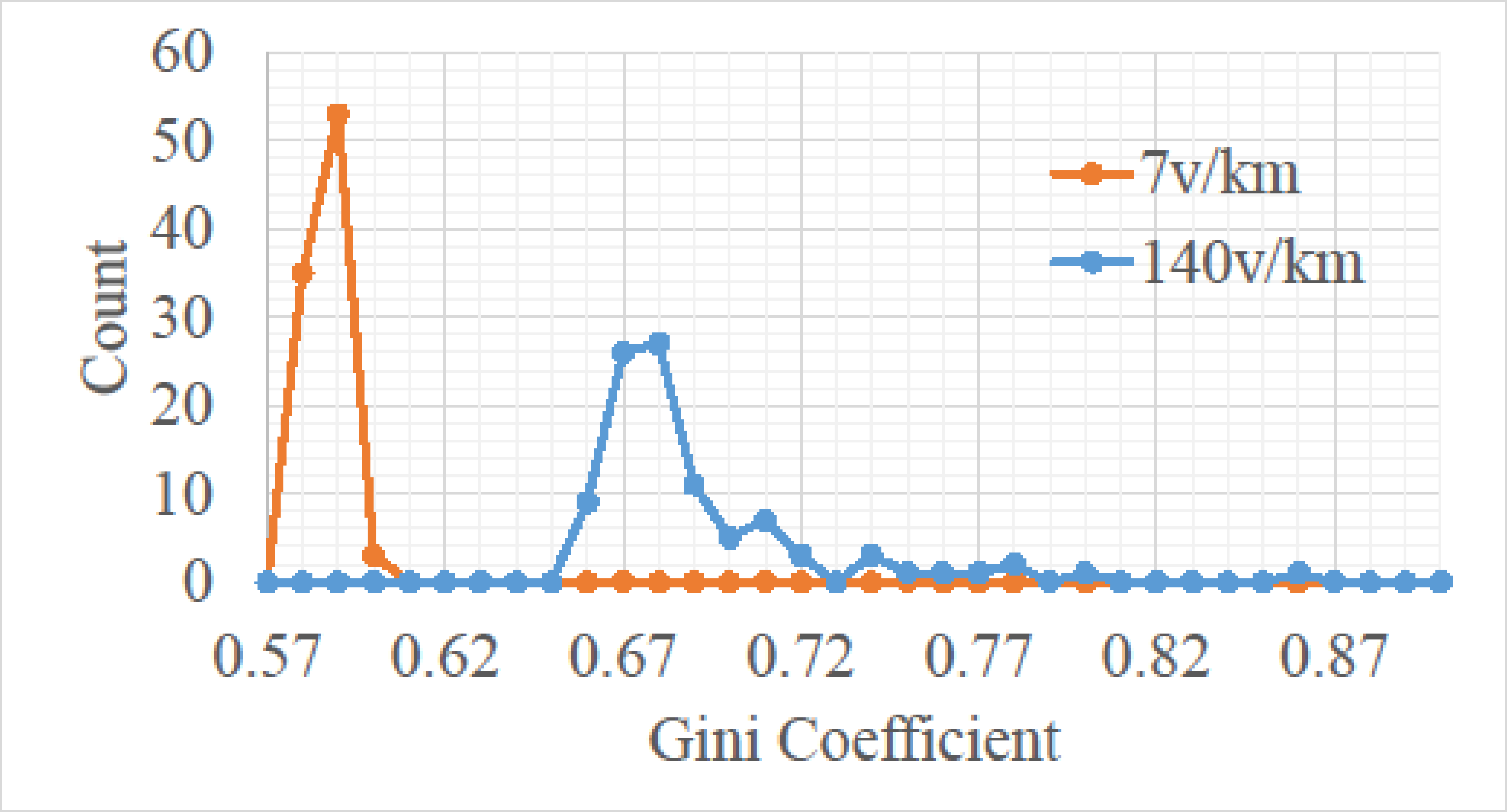
140v/km/lane – 4 lanes – CSMA/CA



Vehicles/km	CSMA/CA		
	Gini	Jain	Jain-1
7	0.6199	0.2795	0.7205
30	0.6894	0.1434	0.8566
70	0.6921	0.1266	0.8734
140	0.8544	0.0323	0.9677

Gini Coefficient Spread

CSMA/CA – 100 Simulations



Mathematically Obtaining the Worst Case Gini Coefficient

- Simulations may never produce the theoretical worst-case Gini-Coefficient
- A mathematically-obtained Worst Case Gini Coefficient will never be exceeded assuming that the application remains the same
- Let us only assume delays are either individually Best Case (smallest) or Worst Case (largest)

Definitions

- D_{BC} | D_{WC}
  - Best Case Delay (shortest time) | Worst Case Delay (longest time)
- P_{BC} | P_{WC}
  - Best Case Proportion | Worst Case Proportion
  - If P_{WC} is 0.2, then 20% of the population suffer D_{WC}, 80% of the population (P_{BC}) suffer D_{BC}
- TD_{BC} | TD_{WC}
  - Total Best Case Delay | Total Worst Case Delay
  - The total amount of delay suffered by each respective proportion of the population
    - $TD_{BC} = P_{BC} * D_{BC}$

Variables

Best to Worst Case Ratio

$$BWR = \frac{D_{BC}}{D_{WC}}$$

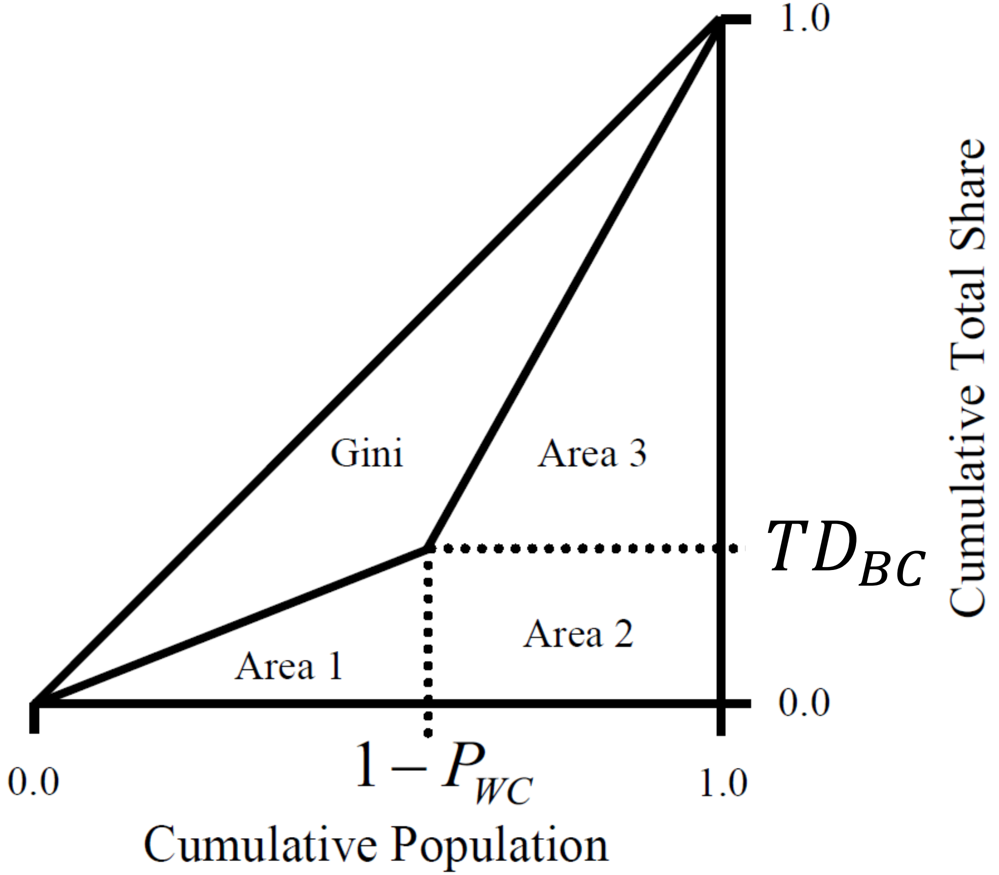
$$0 \le D_{BC} \le D_{WC}$$

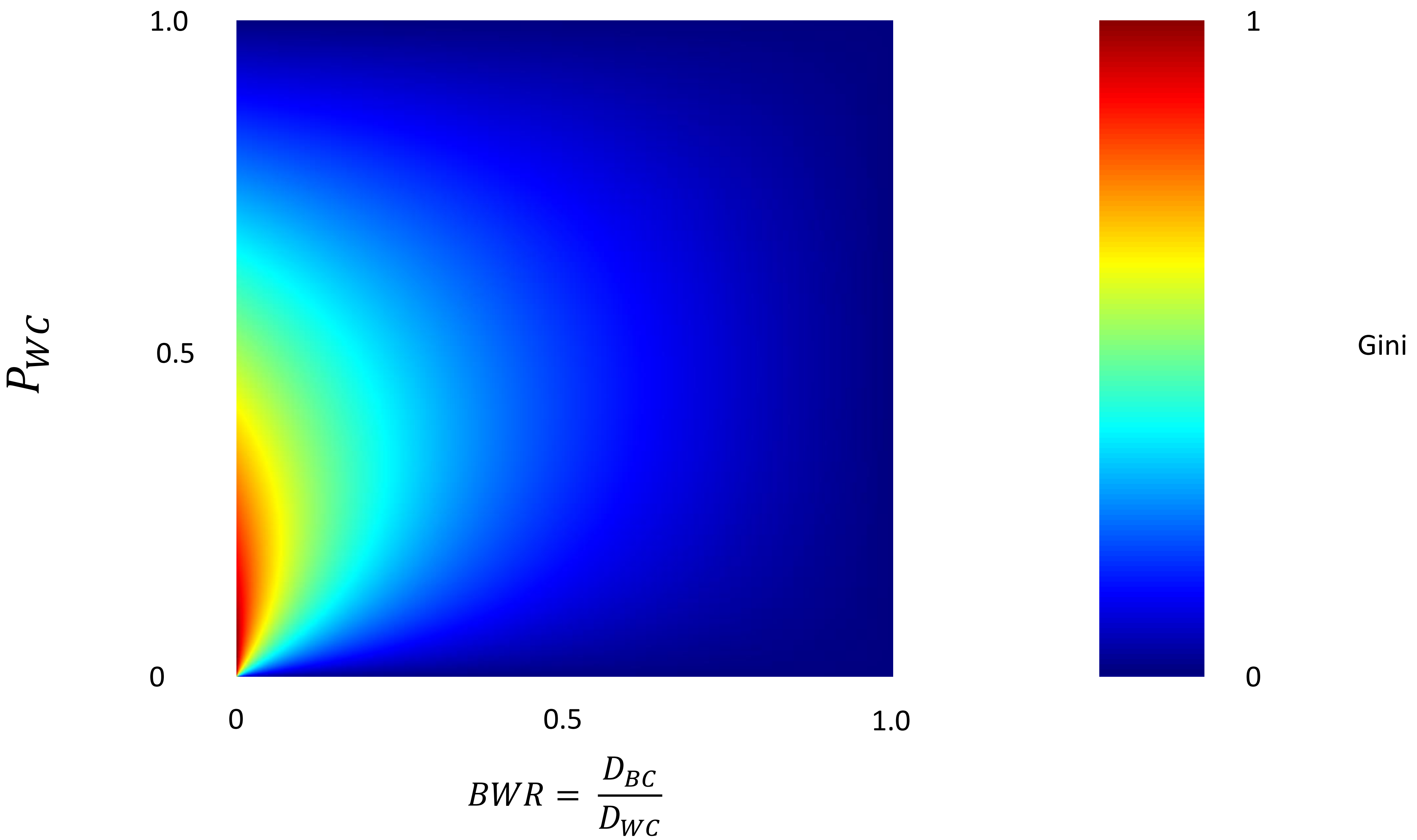
$$0 \le BWR \le 1$$

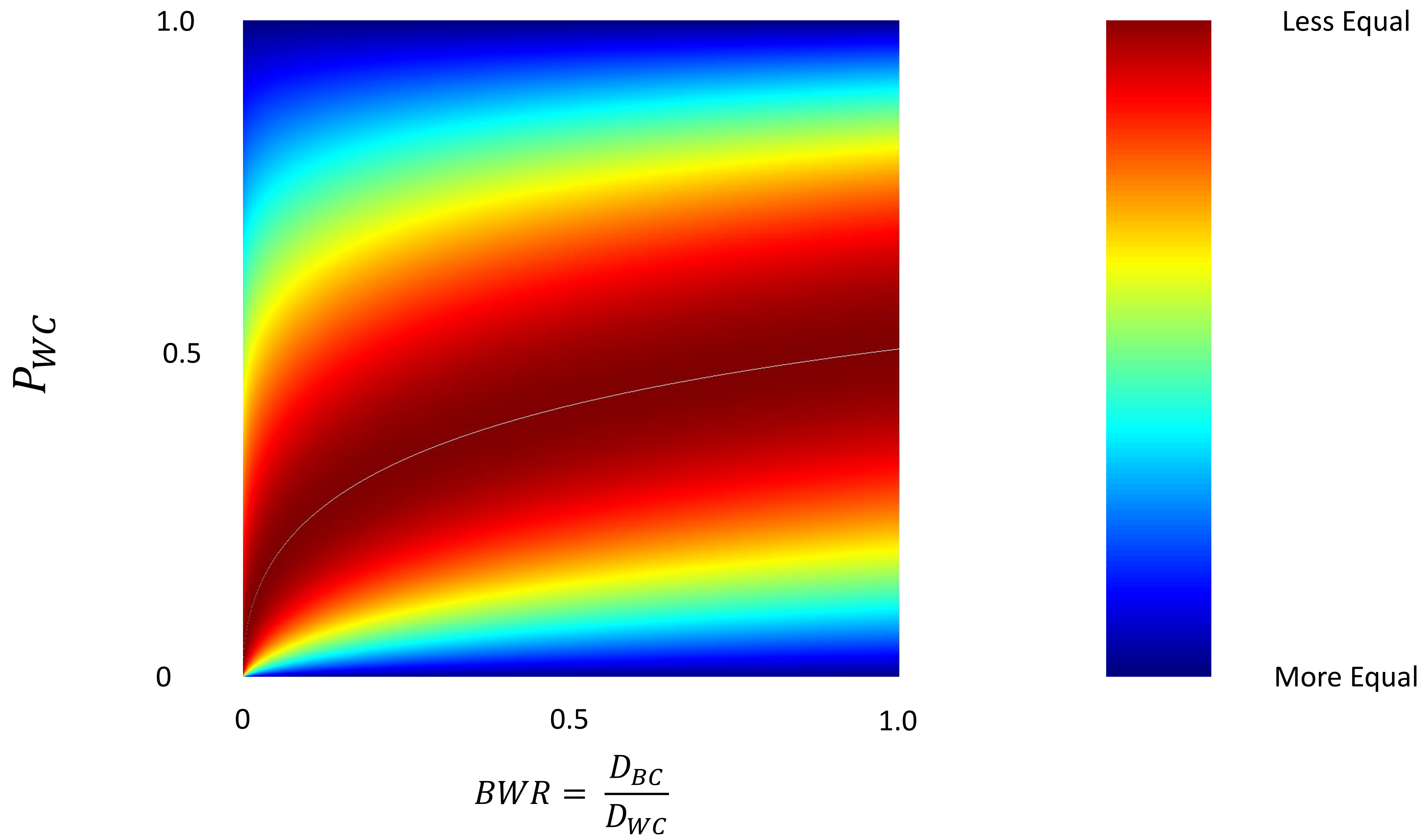
Worst Case Proportion

$$P_{WC}$$

Worst Case Gini Coefficient




1.0
0.5
0
P_{WC}
0
0.5
1.0
$BWR = \frac{D_{BC}}{D_{WC}}$
1
0
Gini



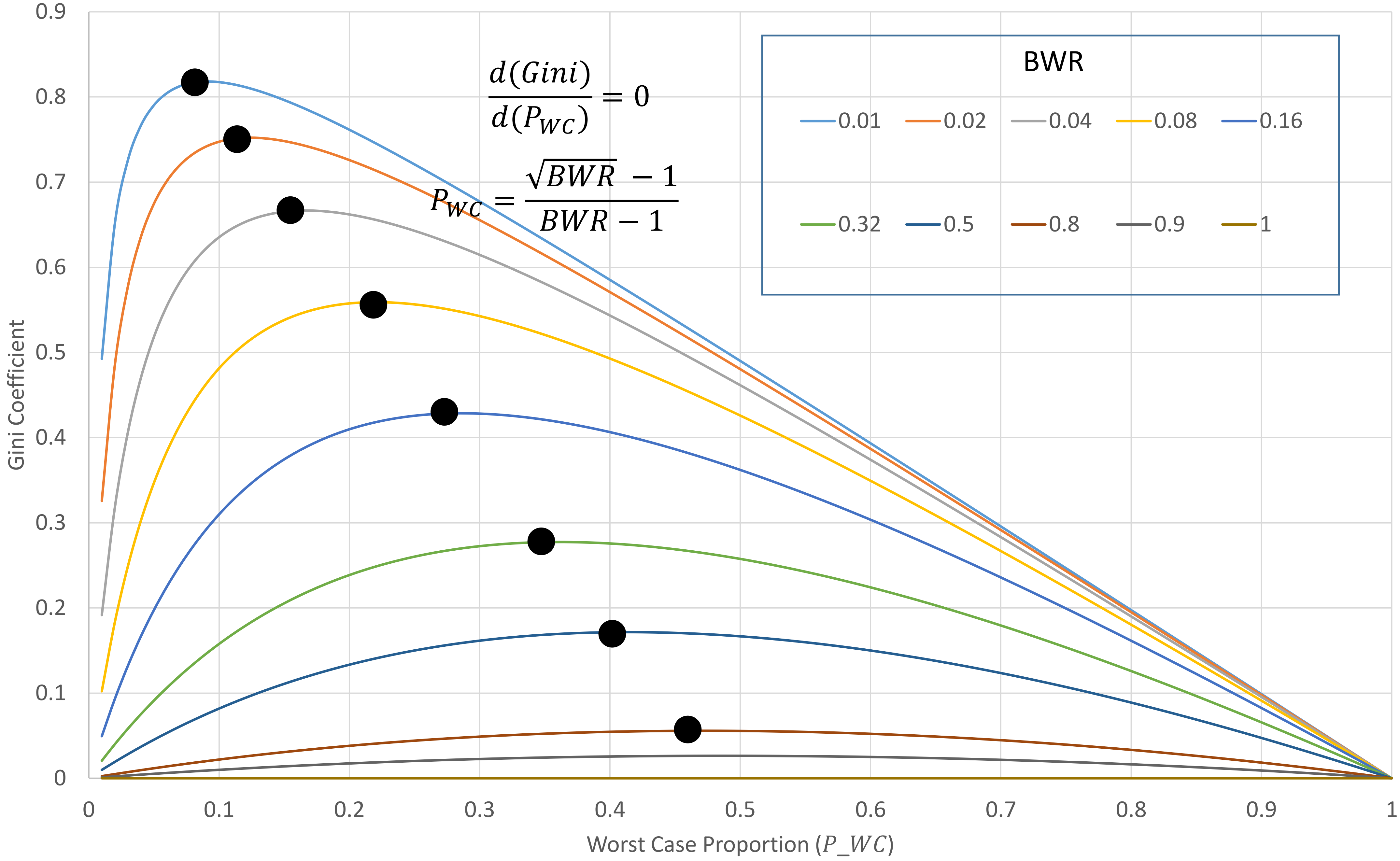
1.0
0.5
0
P_{WC}
0
0.5
1.0
$BWR = \frac{D_{BC}}{D_{WC}}$
Less Equal
More Equal

System Identification of P_{WC}

- Orthogonal Least Squares with Cross-Validation

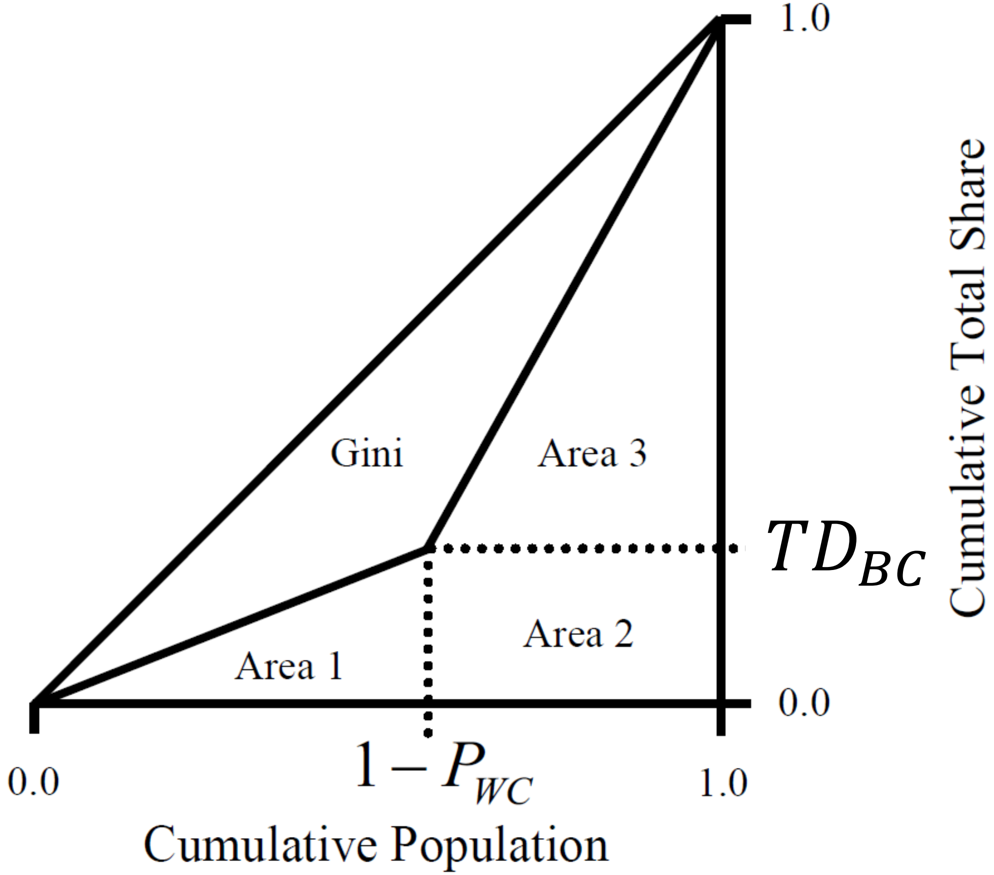
$$P_{WC} = \theta_0 + \theta_1 BWR + \theta_2 BWR^2 + \theta_3 BWR^3$$

	Values			
Resolution	θ_0	θ_1	θ_2	θ_3
1×10^{-2}	-0.6957	0.2460	-0.0342	-0.0016
5×10^{-3}	-0.6945	0.2457	-0.0357	-0.0018
2×10^{-3}	-0.6952	0.2489	-0.0337	-0.0016
1×10^{-3}	-0.6936	0.2473	-0.0349	-0.0018
5×10^{-4}	-0.6934	0.2478	-0.0345	-0.0017
2×10^{-4}	-0.6934	0.2479	-0.0345	-0.0017
1×10^{-4}	-0.6935	0.2477	-0.0346	-0.0017
5×10^{-5}	-0.6934	0.2478	-0.0345	-0.0017
Mean	-0.6941	0.2474	-0.0346	-0.0017
Variance	8.584×10^{-7}	1.11×10^{-6}	3.28×10^{-7}	5.71×10^{-9}



$$\frac{d(Gini)}{d(P_{WC})} = 0$$
$$P_{WC} = \frac{\sqrt{BWR} - 1}{BWR - 1}$$
BWR
0.01
0.02
0.04
0.08
0.16
0.32
0.5
0.8
0.9
1
Gini Coefficient
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Worst Case Proportion (P_WC)

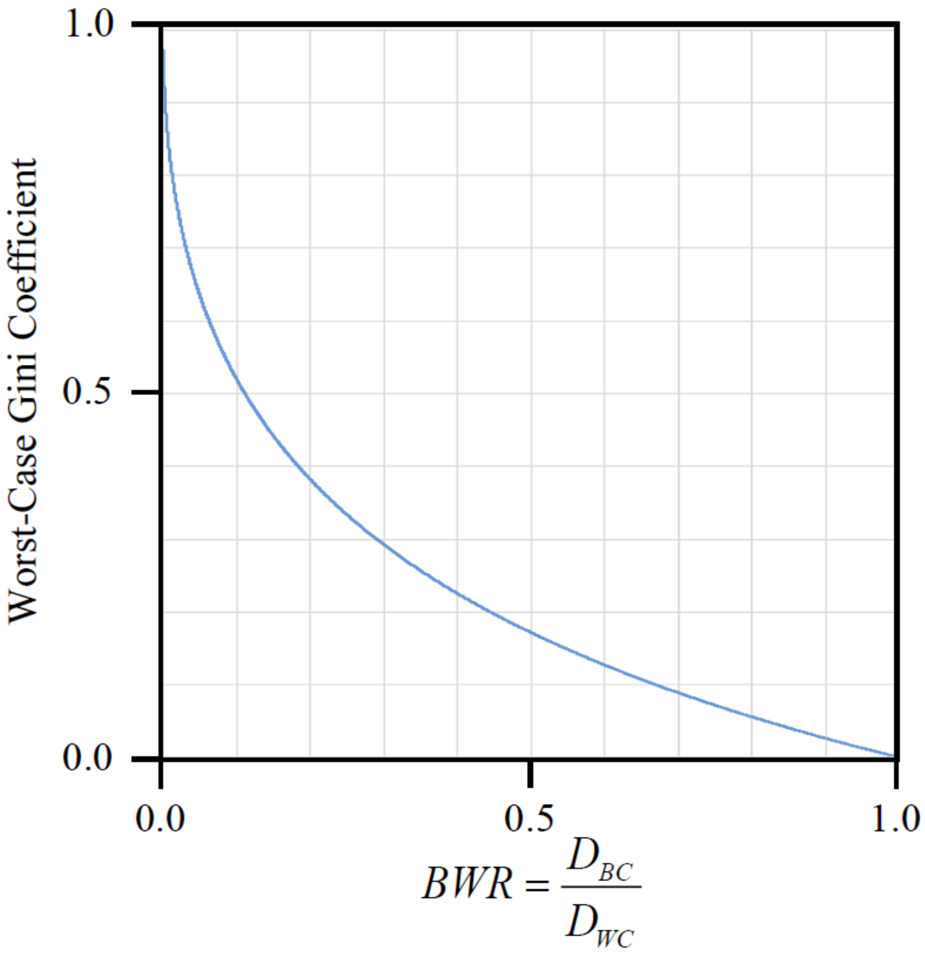
Worst Case Gini Coefficient



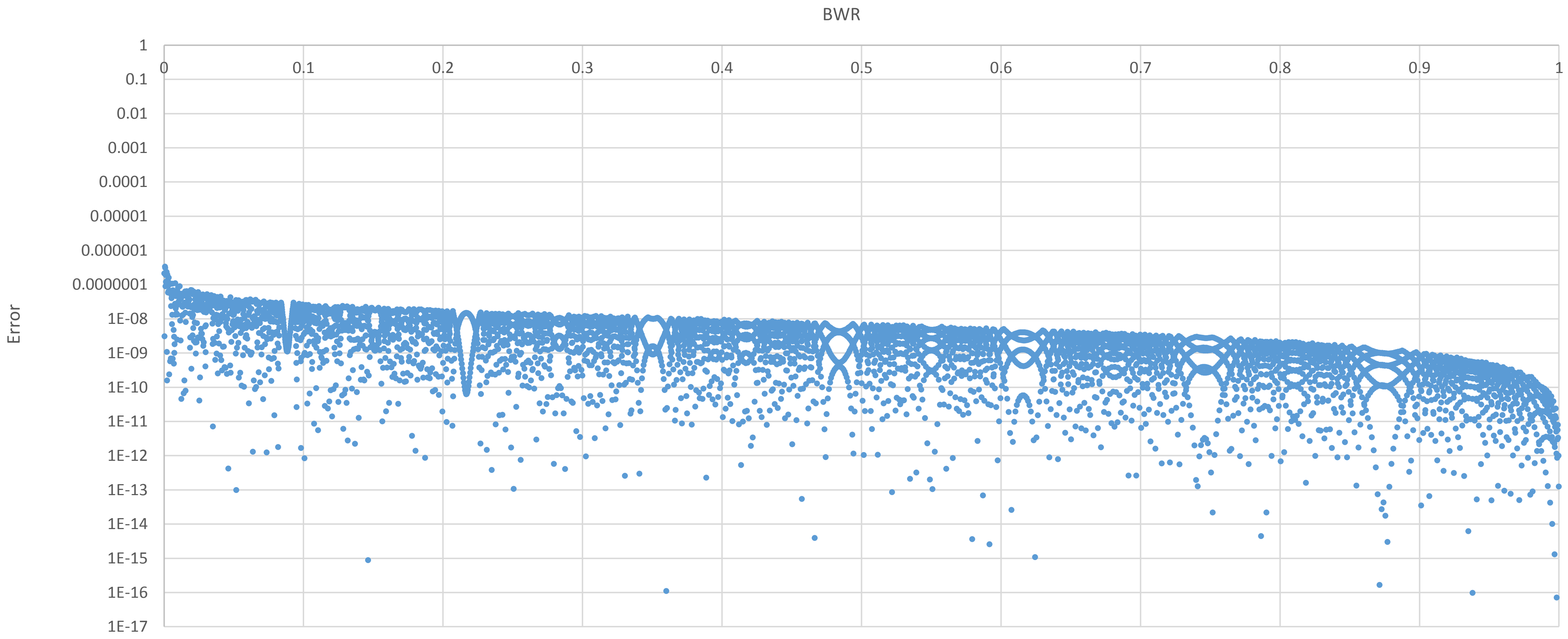
Worst Case Gini Coefficient

$$WCGC = 1 - \frac{2\sqrt{BWR}}{\sqrt{BWR} + 1}$$

Worst Case Gini Coefficient

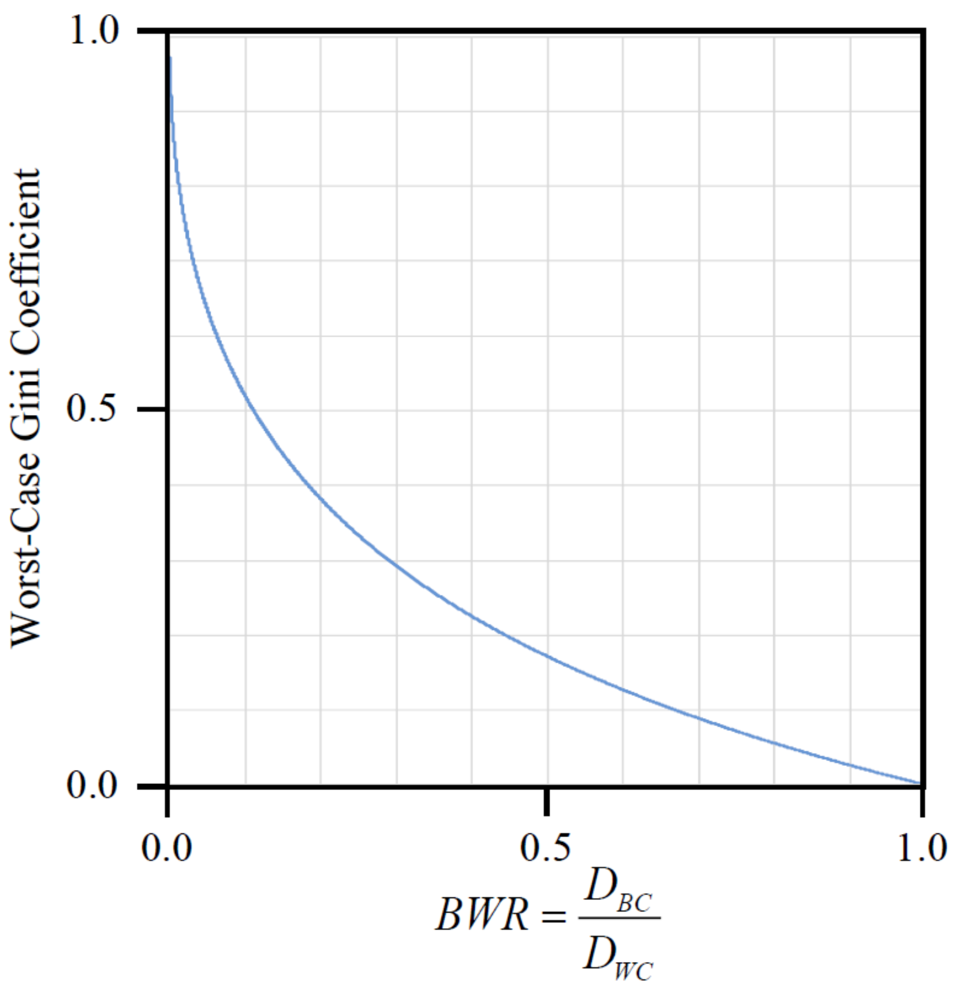


Error Between System Identified WCGC and Mathematically Derived WCGC



Worst Case Gini Coefficient

	TDMA	STDMA	CSMA/CA	CSMA/CA (theoretical)
D_{BC}	0.49s	0.8037	0.0032	0.0032
D_{WC}	0.49s	0.1223	10.054	∞
BWR	1.00	0.6572	0.0003	0
WCGC	0.00	0.1046	0.9650	1

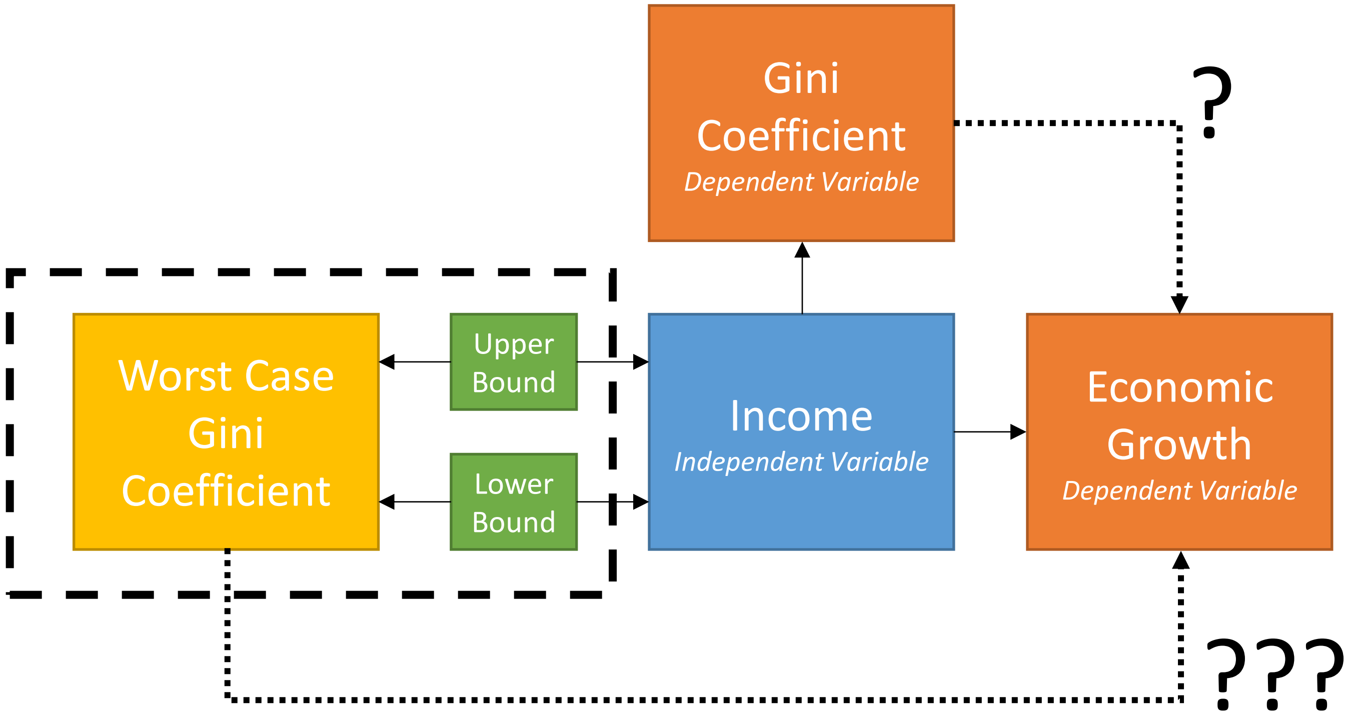


Worst Case Gini Coefficient

A bounded quantifiable measure for fairness of a distribution

Wherever the upper and lower bounds of delay can be guaranteed, the Worst Case Gini Coefficient can be equally guaranteed

Example of System Setup



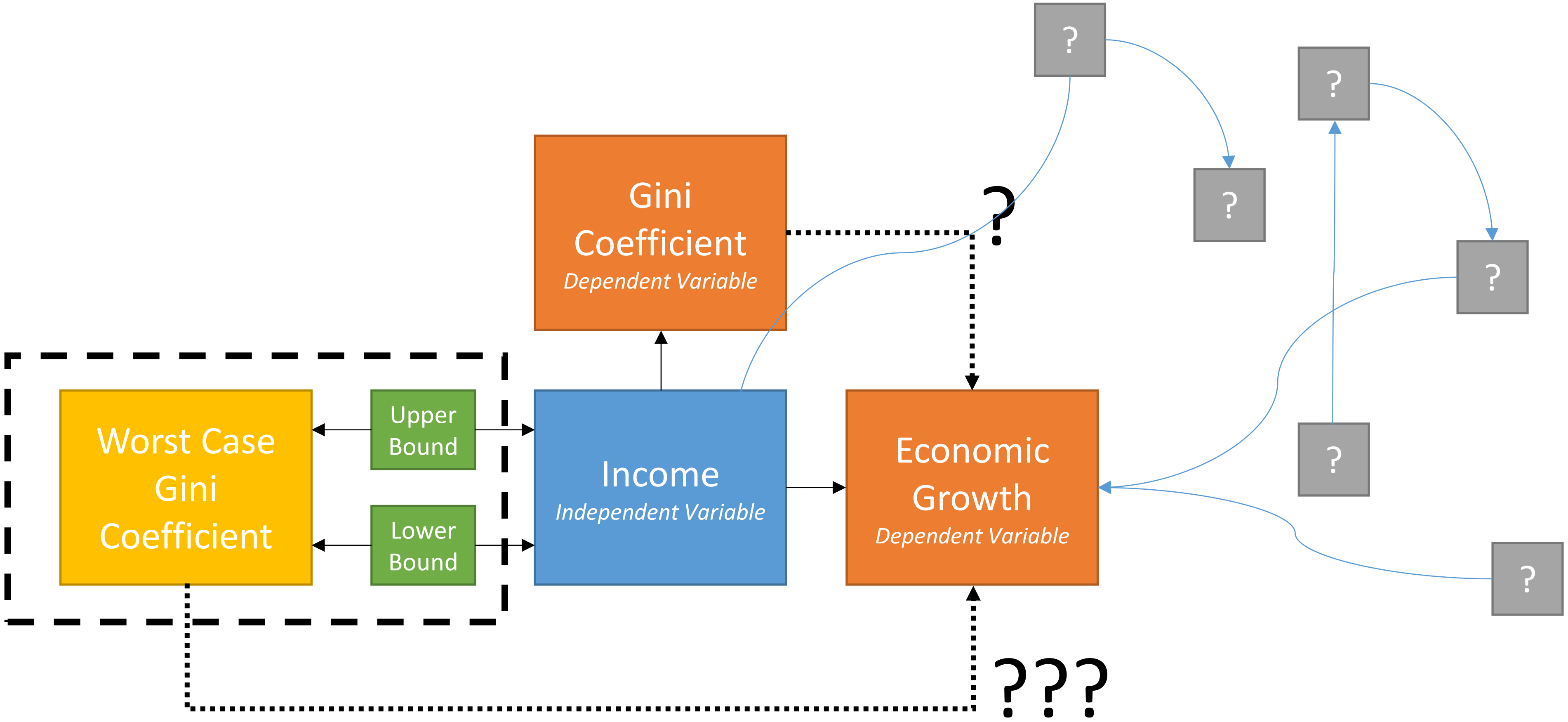
A	B
0	0
0	1
~~1~~	~~0~~
1	1

$$(A \to B) \nrightarrow (B \to A)$$

- Just because we can control fairness, does not mean we can control our dependent variable
- For example:
  - Fajnzlber found that there is a positive correlation between Gini inequality in income and violent crime [2]
  - If we bound income we can lower Worst Case Gini
  - However, since Gini is lower, this does not mean crime will be lower.
  - There is a possibility of a third factor which impacts income which in turn impacts the Gini. Or there could be a combination of factors.
  - In this case, we will have treated the symptom, but not the cause.

[2] Fajnzlber, Pablo, Daniel Lederman, and Norman Loayza. "Inequality and violent crime." *JL & Econ.* 45 (2002): 1.

Example of System Setup



More Fair ≢ Better

- As can be seen from previous examples, having higher levels of fairness does not always equate to high levels of "something good"
- With some exceptions, more fair is *generally* better than more unfair
- For simple or abstracted systems, we can use Fairness-Based Control
  - Small Local Systems
  - Abstracted or Simple Distributed Systems
  - Networks?
- For complex systems, Fairness-Based Control is much more difficult
  - Sociological Issues
  - Economies
  - Massive Distributed Systems

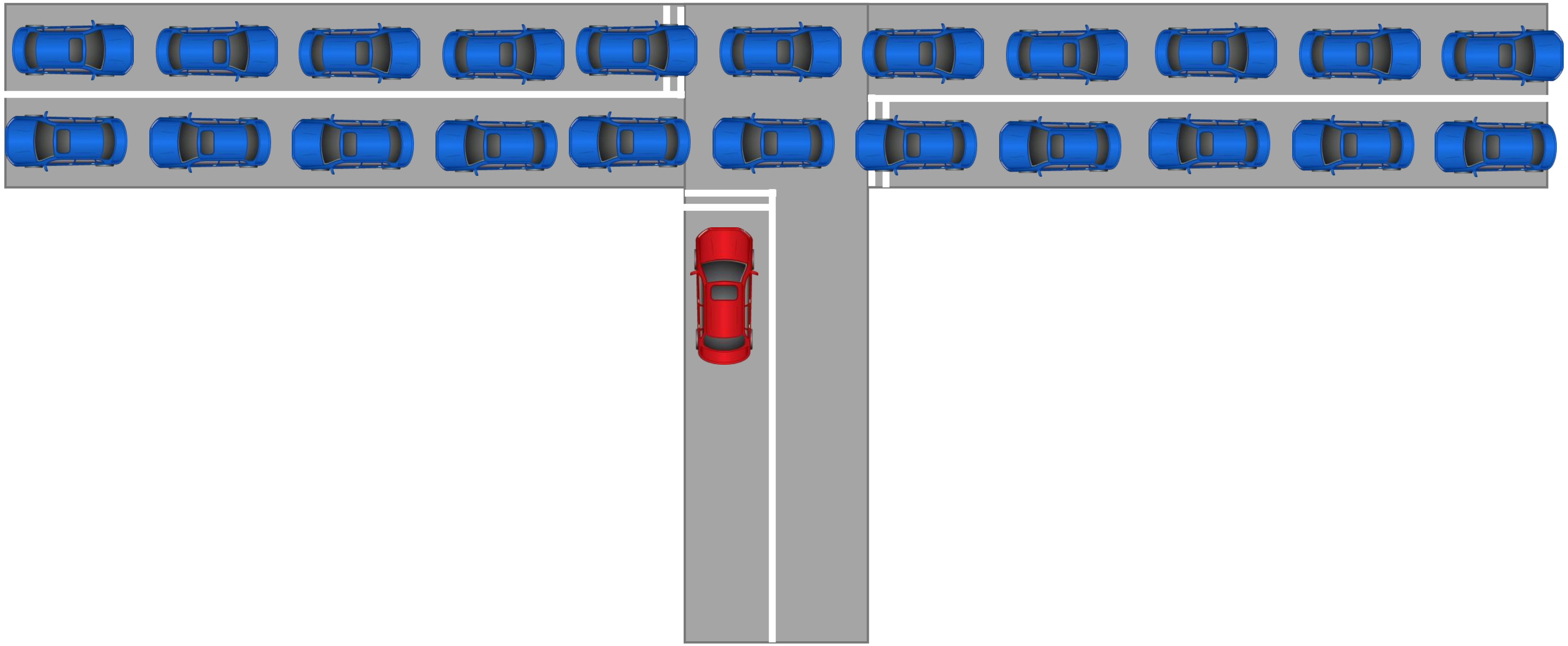
Fairness Trade-offs

- Fairness will usually come at the cost of something else
  - In networking it will come at the cost of throughput
  - In processing it will come at the cost of efficiency
  - In economics (income) it will come at the cost of economic diversity
- A trade-off evaluation should be conducted between fairness and the “throughput equivalent” to not lose sight of the initial purpose of the system
  - Sediq et. Al. [3] performed such an evaluation between efficiency and the Jain Index for Resource Allocation of Wireless Systems

[3] Bin Sediq, A.; Gohary, R.H.; Yanikomeroglu, H., "Optimal tradeoff between efficiency and Jain's fairness index in resource allocation," *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on* , vol., no., pp.577,583, 9-12 Sept. 2012

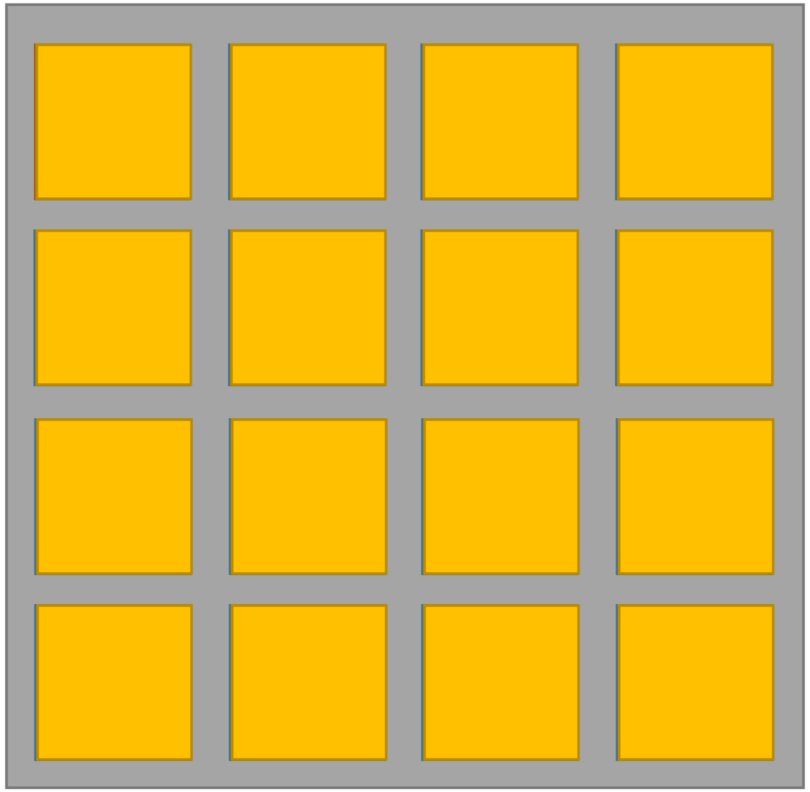
Where Can Fairness Help?

- Traffic Optimisation



Where Can Fairness Help?

- Workload Distribution



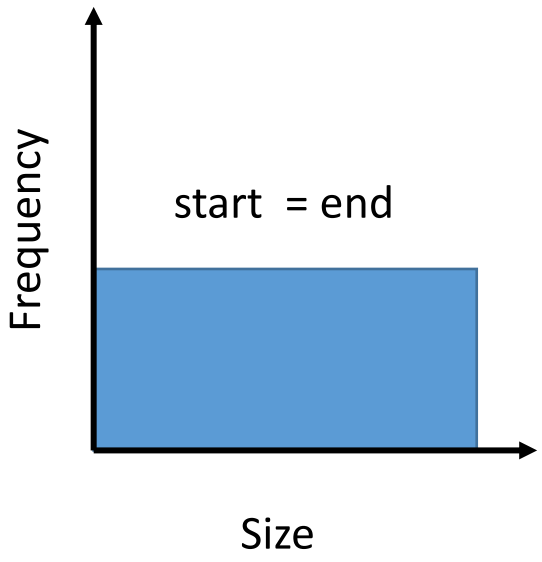
Further Research

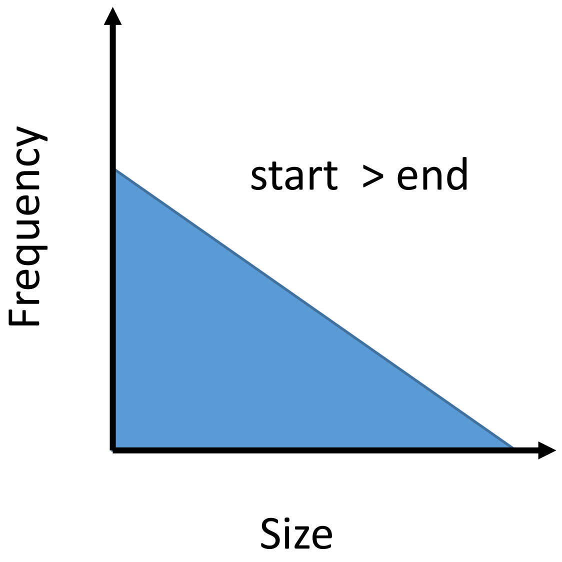
- Gini-Based Elimination (i.e. Fairness-Based Scheduling)
  - Given a population, which member should be removed/serviced in order to have the highest impact on Gini
- "Effective" Gini
  - If we are aware of the potential distribution of a population, can we also know its Gini
    - Formulas?

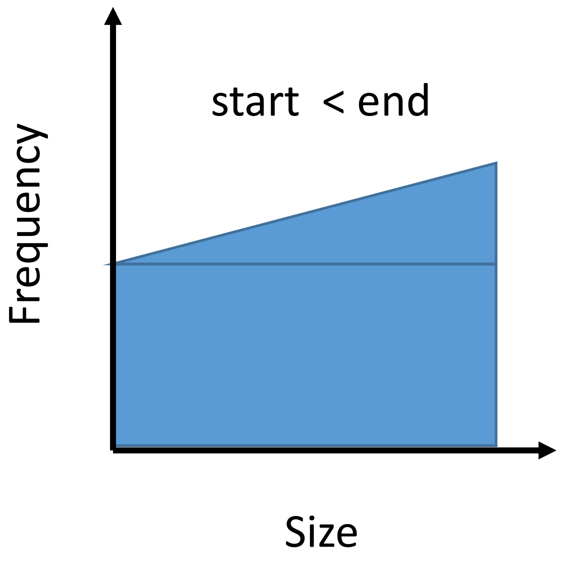
“Effective” Gini – Work in Progress

- Worst Case Gini will give us the **absolute theoretical maximum** of Gini for that particular BWR
- In *most* cases this value will not be hit
- If we know what our distribution looks like – are we able to predict the range of “effective” Gini?

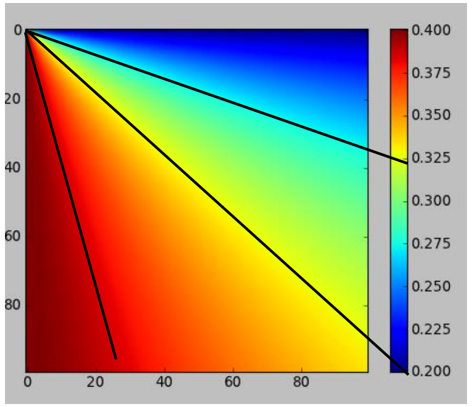
Trapezoidal Distributions



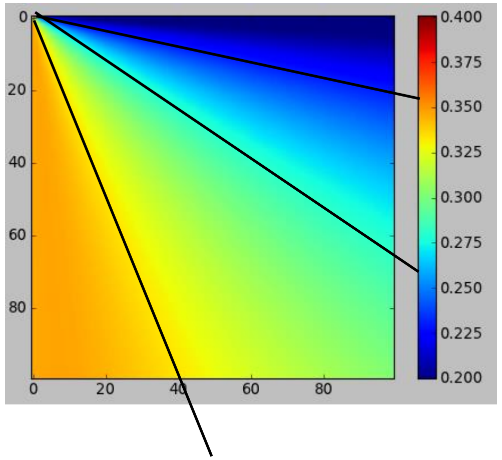




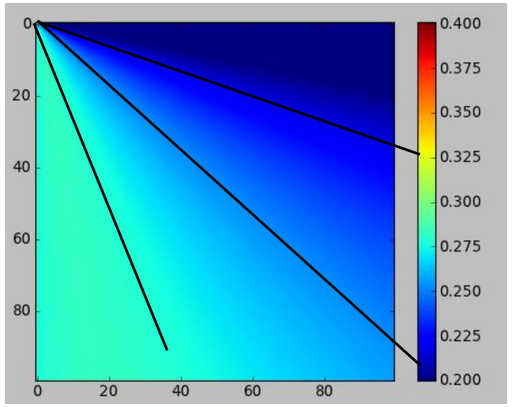
At Different BWR's



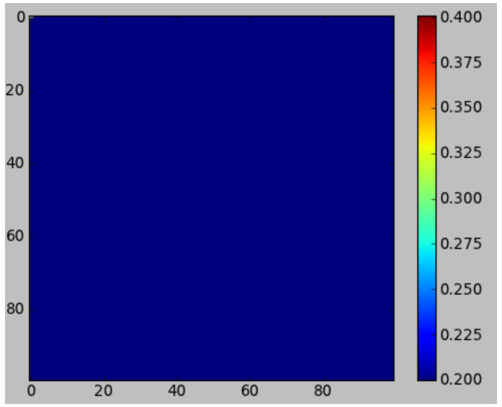
BWR = 0



BWR = 0.05



BWR = 0.13



BWR = 0.29

$$Gini = 0.4 - \frac{0.2\tan^{-1}\left(\frac{end}{start}\right)}{0.5\pi}$$