Risk-Based Sampling: A Summary of Findings

Mark Powell

USDA Office of Risk Assessment and Cost Benefit Analysis

NAPPO Risk Based Sampling Symposium

26 June 2017

Baltimore, MD

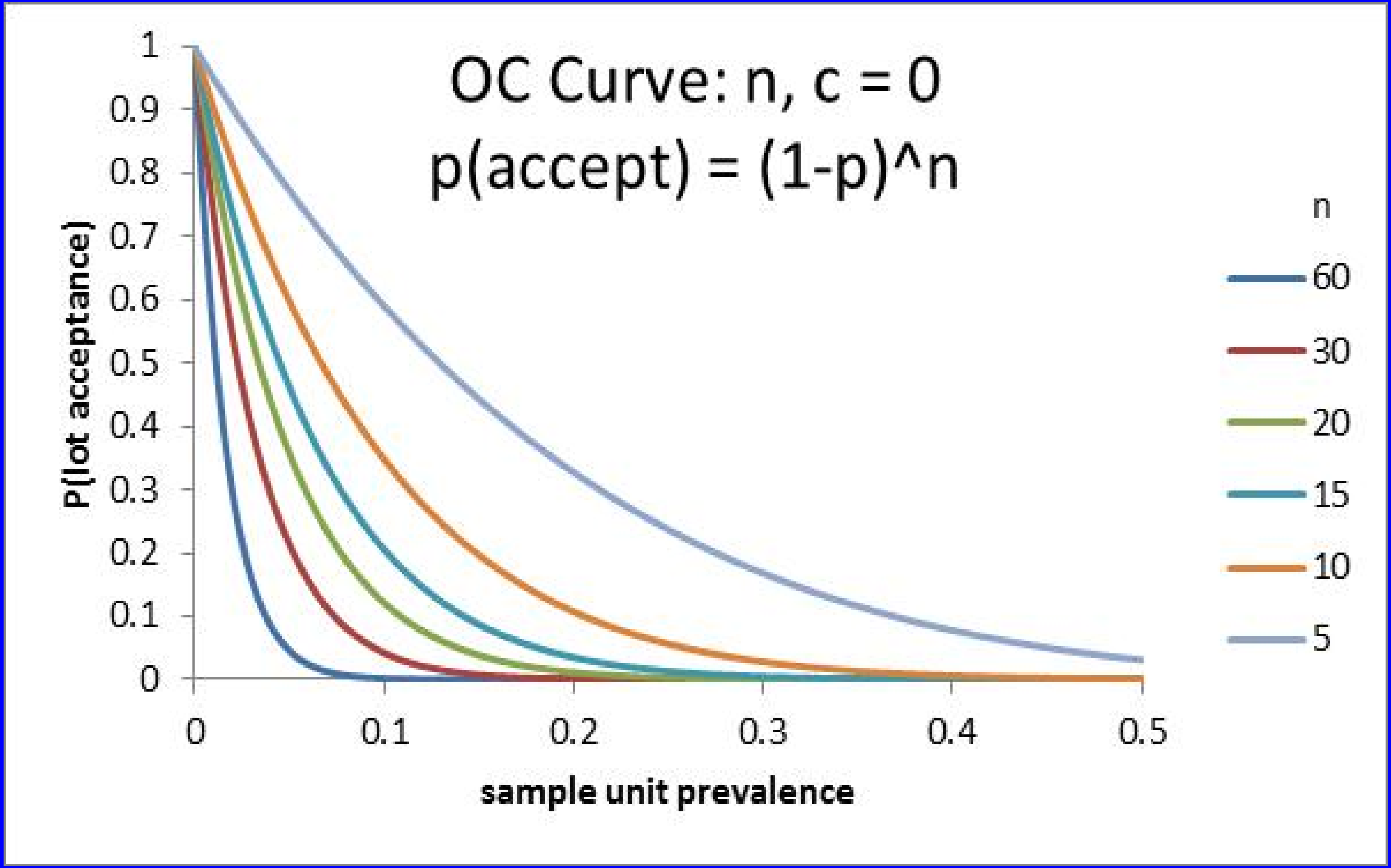
Risk-Based Sampling

- Powell, M. 2014. Optimal Food Safety Sampling Under a Budget Constraint. Risk Analysis. 34(1): 93-100.
- Powell, M. 2015. Risk-Based Sampling: I Don’t Want to Weight in Vain. Risk Analysis. 35(12):2172-2182.

Risk-Based Sampling

– Powell, M. 2014. Optimal Food Safety Sampling Under a Budget Constraint. Risk Analysis. 34(1): 93-100.

Conventional Lot Acceptance Sampling Plan Design



Simple Optimization Model

- $Max\ L_R = m[1 - q^n]$
  - $L_R =$ contaminated lots rejected
  - $m =$ lots
  - $n =$ samples per lot
  - $q = (1 - \text{p})$
  - $p =$ sample unit prevalence
  - 1-q^n = p(reject lot)
- *S.t.:* Budget constraint (C_T)

Simple Optimization Model

- $Max\ L_R = \frac{C_T}{C_l + nCn}[1 - q^n]$
  - $m = \frac{C_T}{C_l + nC_n}$ (budget constraint)
  - $C_T =$ budgeted total sampling cost ($)
  - $C_l =$ cost per lot ($)
  - $C_n =$ cost per sample ($)

$$n_{opt}(C_T, C_l, C_n, p) \rightarrow \frac{\delta(L_R)}{\delta(n)} = \frac{[C_l + nC_n][-C_T q^n ln(q)] - [C_T(1 - q^n)]C_n}{[C_l + nC_n]^2} = 0$$

Simple Optimization Model

- Obj Fxn: $L_R = m(1 - q^n) = f(m, n|q)$
- Constraint: $C_T \geq m(C_l + nC_n)$
- $L = f(m, n|q) + \lambda[C_T - m(C_l + nC_n)]$

Simple Optimization Model

1) $\frac{\delta L}{\delta m} = \frac{\delta f}{\delta m} - \lambda(C_l + nC_n) = 0$

2) $\frac{\delta L}{\delta n} = \frac{\delta f}{\delta n} - \lambda m C_n = 0$

3) $\frac{\frac{\delta f}{\delta m}}{\frac{\delta f}{\delta n}} = \frac{C_l + nC_n}{mC_n}$

4) $n = \frac{\frac{\delta f}{\delta m}}{\frac{\delta f}{\delta n}} m - \frac{C_l}{C_n}$

5) $\frac{\frac{\delta f}{\delta m}}{\frac{\delta f}{\delta n}} = \frac{(1-q^n)}{-mq^n ln(q)}$

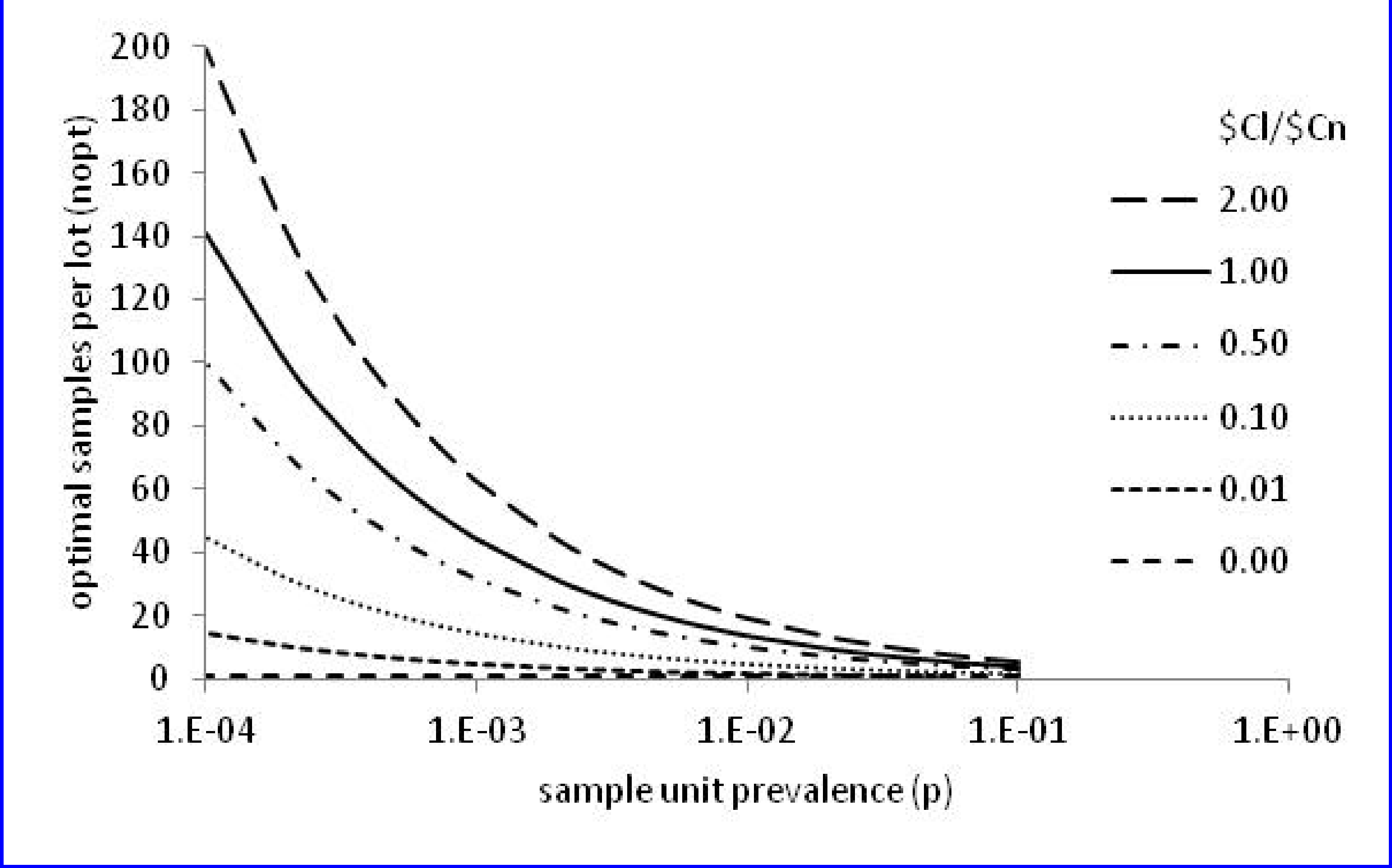
6) $n + \frac{(1-q^n)}{q^n ln(q)} + \frac{C_l}{C_n} = 0$

Note: $n_{opt} = f\left(p, \frac{C_l}{C_n}\right)$

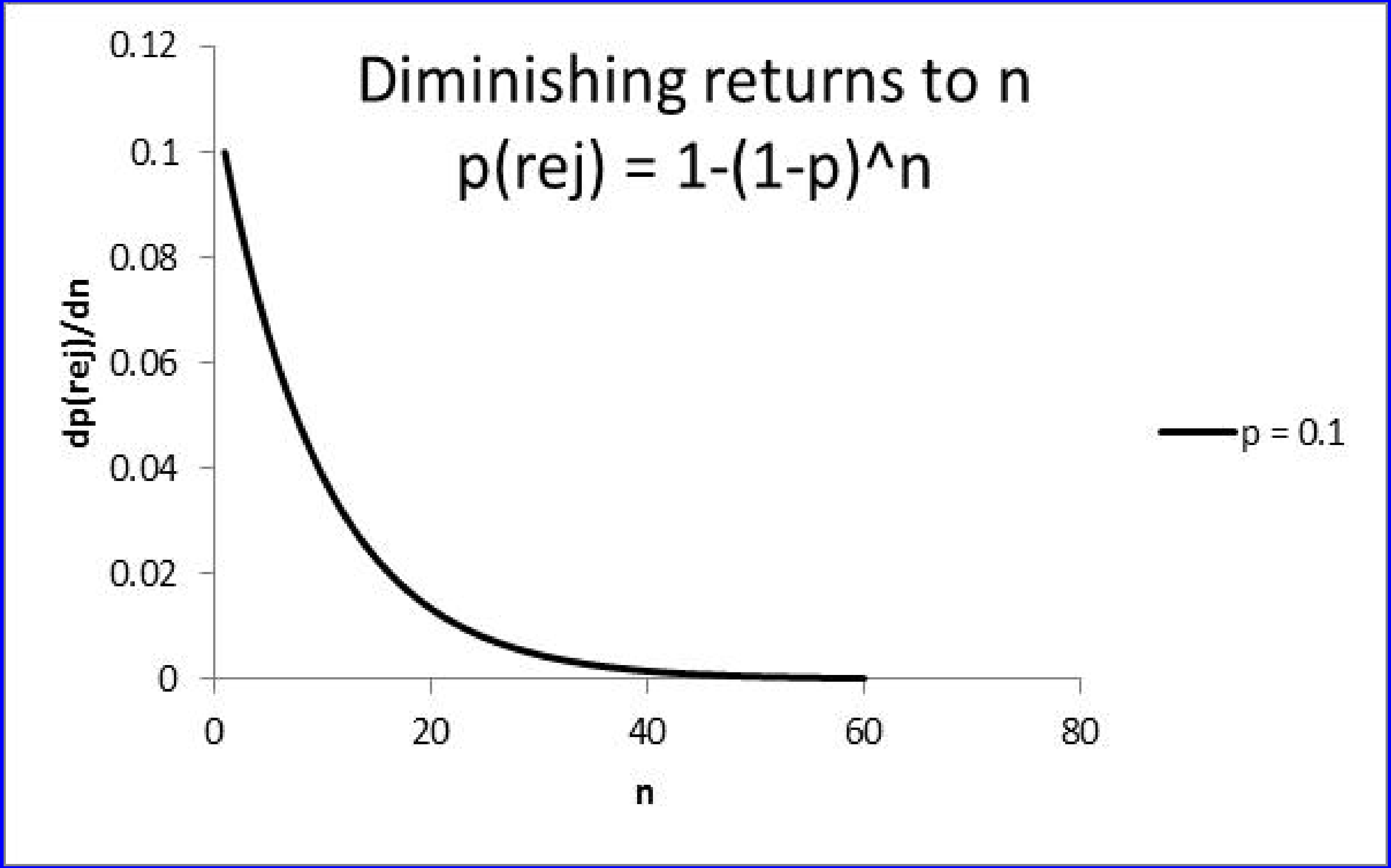
Results

- If budget constraint does not permit testing 100% of lots, n_{opt} for a given sample unit prevalence (*p*) depends only on the cost ratio (C_l/Cn).
- The budget constraint (C_T) determines absolute number of lots tested in a budget period (*m*) or the frequency of lot inspection ($1/m$)

Results


200
180
160
140
120
100
80
60
40
20
0
optimal samples per lot (nopt)
1.E-04
1.E-03
1.E-02
1.E-01
1.E+00
sample unit prevalence (p)
$Cl/$Cn
2.00
1.00
0.50
0.10
0.01
0.00

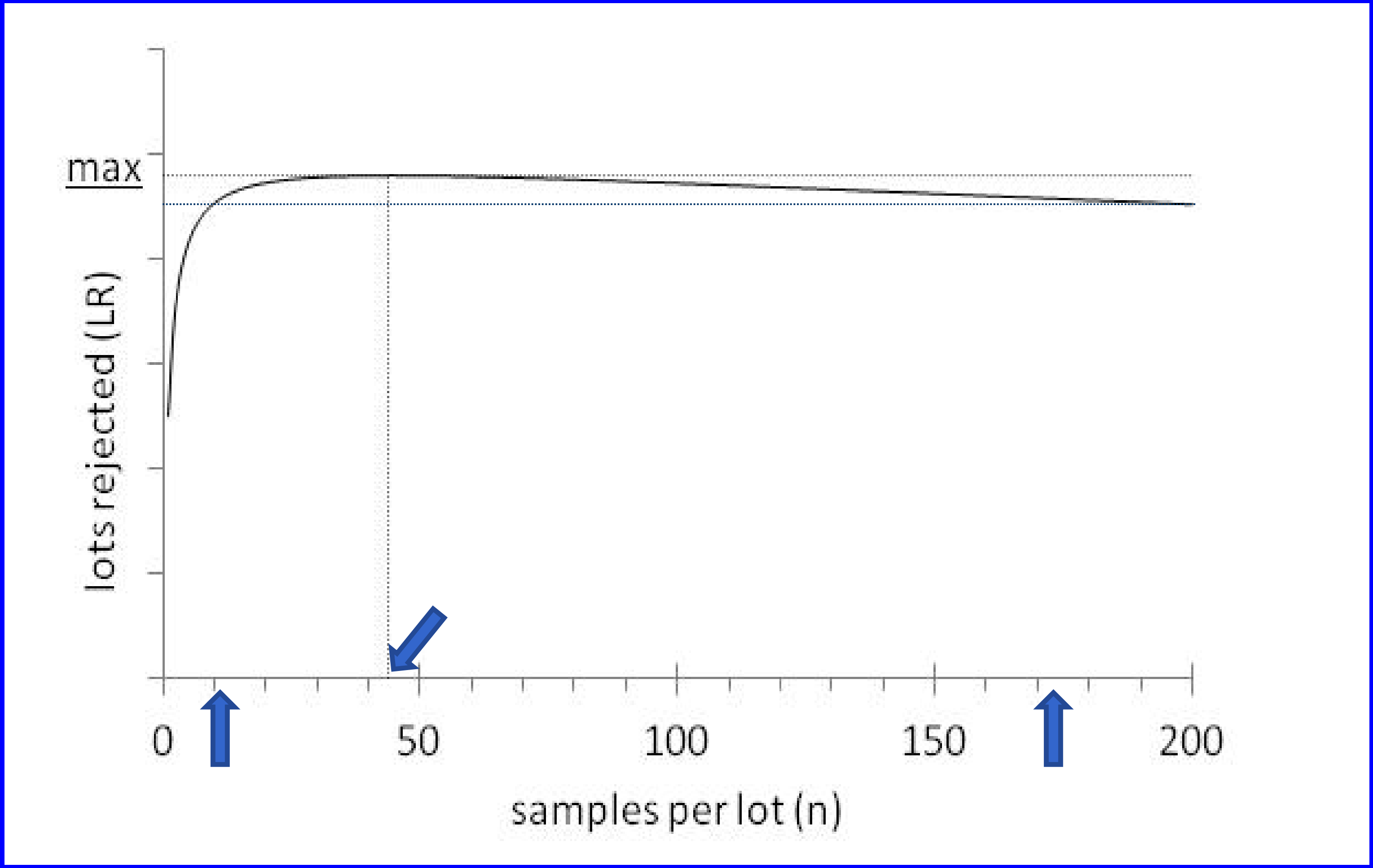
Results



Diminishing returns to n
p(rej) = 1-(1-p)^n
0.12
0.1
0.08
0.06
0.04
0.02
0
dp(rej)/dn
0
20
40
60
80
n
p = 0.1

Results

$\$C_l/\$C_n = 1$ and $p = 10^{-3}$



Conclusion

- National Research Council (1985): sampling plans based on "sound statistical concepts" need to "achieve a high degree of confidence in the acceptability of a lot."
- Economic design of measures is not new.
- Scarce resources should force us to consider the tradeoff between depth (n) and coverage (m).
- Multiple, competing objectives for sampling.

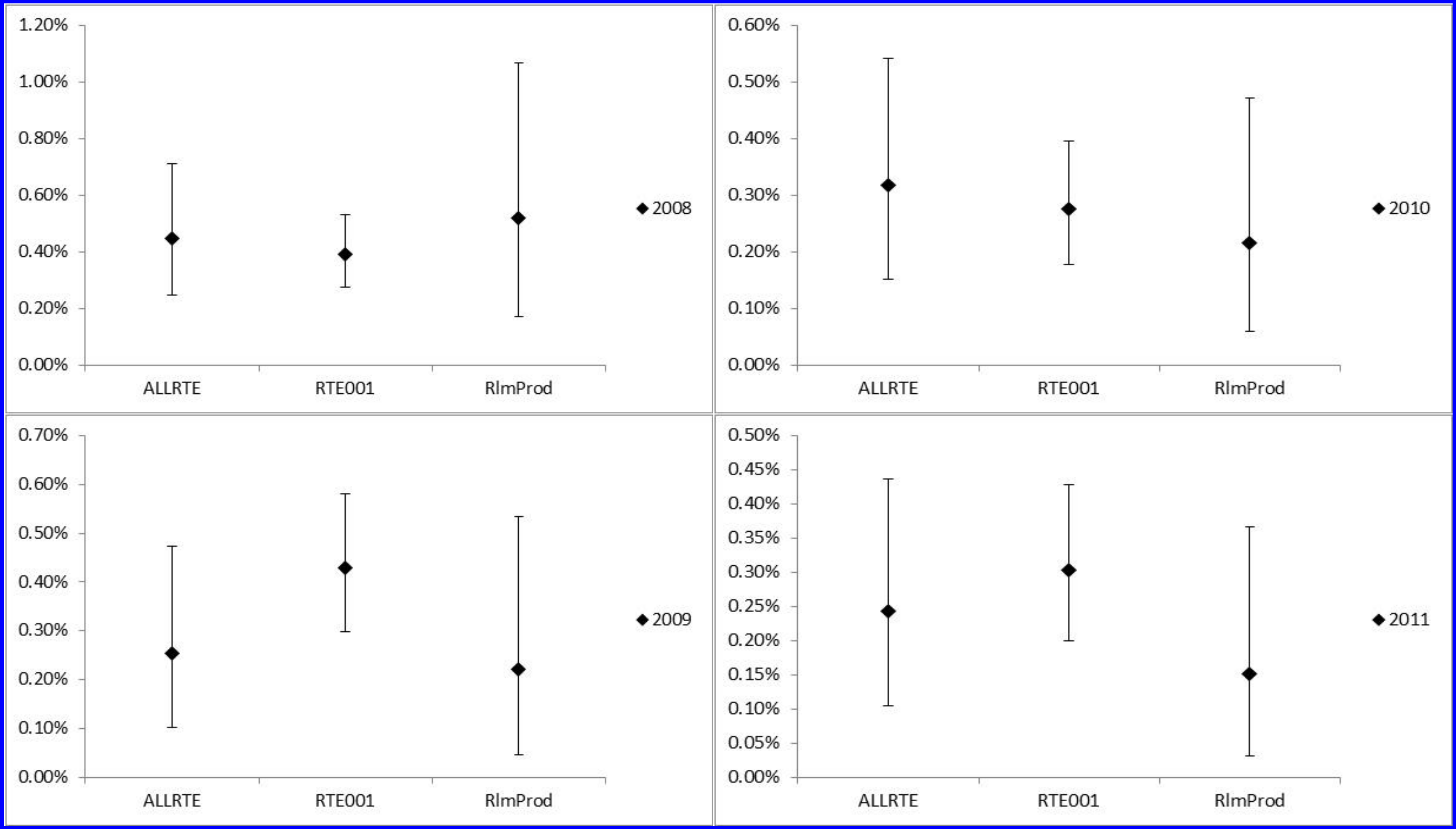
Risk-Based Sampling

– Powell, M. 2015. Risk-Based Sampling: I Don’t Want to Weight in Vain. Risk Analysis. 35(12):2172-2182.

Risk Based Inspection

- GAO (General Accounting Office) (1992). Food Safety and Quality: Uniform, Risk-based Inspection System Needed to Ensure Safe Food Supply.
- GAO (General Accounting Office) (1994). Risk-Based Inspections and Microbial Monitoring Needed for Meat and Poultry.
- USDA/OIG (U.S. Department of Agriculture, Office of Inspector General). (2007). Issues Impacting the Development of Risk-Based Inspection at Meat and Poultry Processing Establishments
- IOM (Institute of Medicine) (2009). Review of Use of Process Control Indicators in the FSIS [Food Safety and Inspection Service] Public Health Risk-Based Inspection System.
- NRC (National Research Council) (2009). Letter Report on the Development of a Model for Ranking FDA Product Categories on the Basis of Health Risks.
- NRC (National Research Council) (2009). Letter Report on the Review of the Food Safety and Inspection Service Proposed Risk-Based Approach to and Application of Public-Health Attribution.
- NRC (National Research Council) (2009). Review of the Methodology Proposed by the Food Safety and Inspection Service for Follow-up Surveillance of In-Commerce Businesses.
- NRC (National Research Council) (2009). Review of the Methodology Proposed by the Food Safety and Inspection Service for Risk-Based Surveillance of In-Commerce Activities.
- NRC (National Research Council) (2010). Enhancing Food Safety: The Role of the Food and Drug Administration.
- NRC (National Research Council) (2011). A Risk-Characterization Framework for Decision-Making at the Food and Drug Administration.

Food Safety Example: *Listeria monocytogenes* in Ready-to-Eat Meat and Poultry



Risk-Based Animal Health Surveillance

- "The rapid rate of acceptance of this core concept of risk-based surveillance has outpaced the development of its theoretical and practical bases."
  - Stark et al. 2006. "Concepts of risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches." BMC Health Services Research. 6:20.

Risk Portfolio Analysis

- Prattley et al. 2007. Application of portfolio theory to risk-based allocation of surveillance resources in animal populations. Preventive Veterinary Medicine 81: 56–69.
- Cannon. 2009. Inspecting and monitoring on a restricted budget—where best to look? Preventive Veterinary Medicine 92:163–174.
- Cox. 2009. What's Wrong with Hazard-Ranking Systems? An Expository Note. Risk Analysis 29(7): 940-948.

Modern Portfolio Theory

- Markowitz (1952) Mean Variance Optimization (MVO):

$$Min_w \ \boldsymbol{w}^T \Sigma \boldsymbol{w} = \sigma_p^2$$

$$\boldsymbol{w}^T \boldsymbol{\mu} = \mu_p^*$$

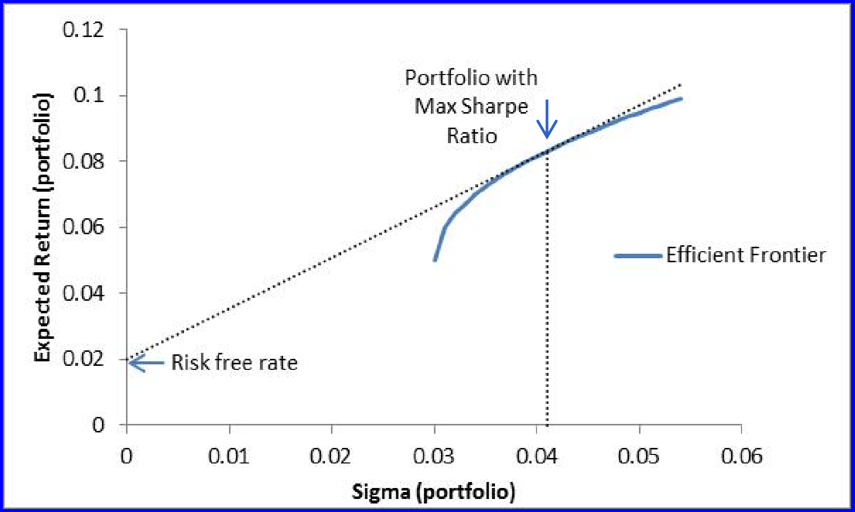
$$\boldsymbol{w}^T \mathbf{1}_N = 1$$

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}$$

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_N \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \theta_1^2 & \cdots & \theta_{1N} \\ \vdots & \ddots & \vdots \\ \theta_{N1} & \ldots & \theta_N^2 \end{bmatrix}$$

$$\text{Sharpe Ratio (S)} = \frac{\mu_p}{\sigma_p}$$



Limitations of Mean-Variance Optimization

- Extreme asset weights
- Weights sensitive to small changes in inputs
- Poor out-of-sample performance
- Assumes stationary process
- # parameters (μ_i, σ_i^2, σ_{ij}) = 2N +(N*(N-1)/2)
- Model uncertainty

Limitations of Mean-Variance Optimization

- Sensitivity of the model to input errors
  - Frankfurter et al. 1971. "Portfolio Selection: The Effects of Uncertain Mean, Variances, and Covariances" The Journal of Financial and Quantitative Analysis 6(5): 1251-1262.
  - Hodges and Brealey. 1972. "Portfolio Selection in a Dynamic and Uncertain World" Financial Analysts Journal 28(6): 58-69.

Limitations of Mean-Variance Optimization

Example of sensitivity to estimation error

- Obj: max $S = \mu_p/\sigma_p$
- Consider two identical assets, A and B:
  - $\mu_A=\mu_B=10\%$; $\sigma_A=\sigma_B=5\%$; $\rho_{AB} = 0.9$
  - true optimal weights: $w_A = w_B = 0.5$
- Assume μ_A estimated with 10% error:
  - $\bar{x}_A = 11\%$
  - est. optimal weights: $\hat{w}_A = 0.95$; $\hat{w}_B = 0.05$

Limitations of Mean-Variance Optimization

- Naïve portfolio formation rules, such as the equal weight (1/N) rule, can outperform MVO.
  - Bloomfield, T., R. Leftwich, and J. Long. 1977. Portfolio Strategies and Performance. Journal of Financial Economics 5:201–18.

Limitations of Mean-Variance Optimization

- Very long history of returns needed to estimate mean excess return accurately
- Even if long time series are available, may not be reasonable to assume the parameters were stationary over that long a period
  - Merton. 1980. On Estimating the Expected Return on the Market. NBER Paper 444.

Limitations of Mean-Variance Optimization

- Given estimation uncertainty, optimal portfolio not well-defined; statistically equivalent portfolios with very different asset weights
  - Michaud. 1989. "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" Financial Analysts Journal. 45: 31-42.

DeMiguel et al (2007) Optimal versus Naïve Diversification

- Evaluated out of sample performance of sample-based MVO and its extensions designed to reduce effects of estimation error.
- 14 optimizing portfolio models compared to naïve diversification (1/N)
- 7 datasets of monthly returns
- Optimization models -10 year moving estimation window to predict next month's performance.
- 3 criteria: Sharpe ratio, CEQ return, and turnover

DeMiguel et al (2007) Optimal versus Naïve Diversification

- None of the portfolio optimization models performed consistently better than 1/N portfolio
- Estimation window needed for MVO and its extensions to outperform 1/N:
  - ~ 3000 months for N = 25 portfolio
  - ~ 6000 months for N = 50 portfolio

Optimal vs. Simple: Simulation

- Let $i = 1, ..., 27$ producers; $j = 1, ..., 20$ yrs.

Factor	hi	med	low
Mean lot prevalence (μ)	0.01	0.005	0.001
CV lot prevalence (σ/μ)	2	1	0.5
Volume (L_t, lots/yr)	100,000	10,000	1,000

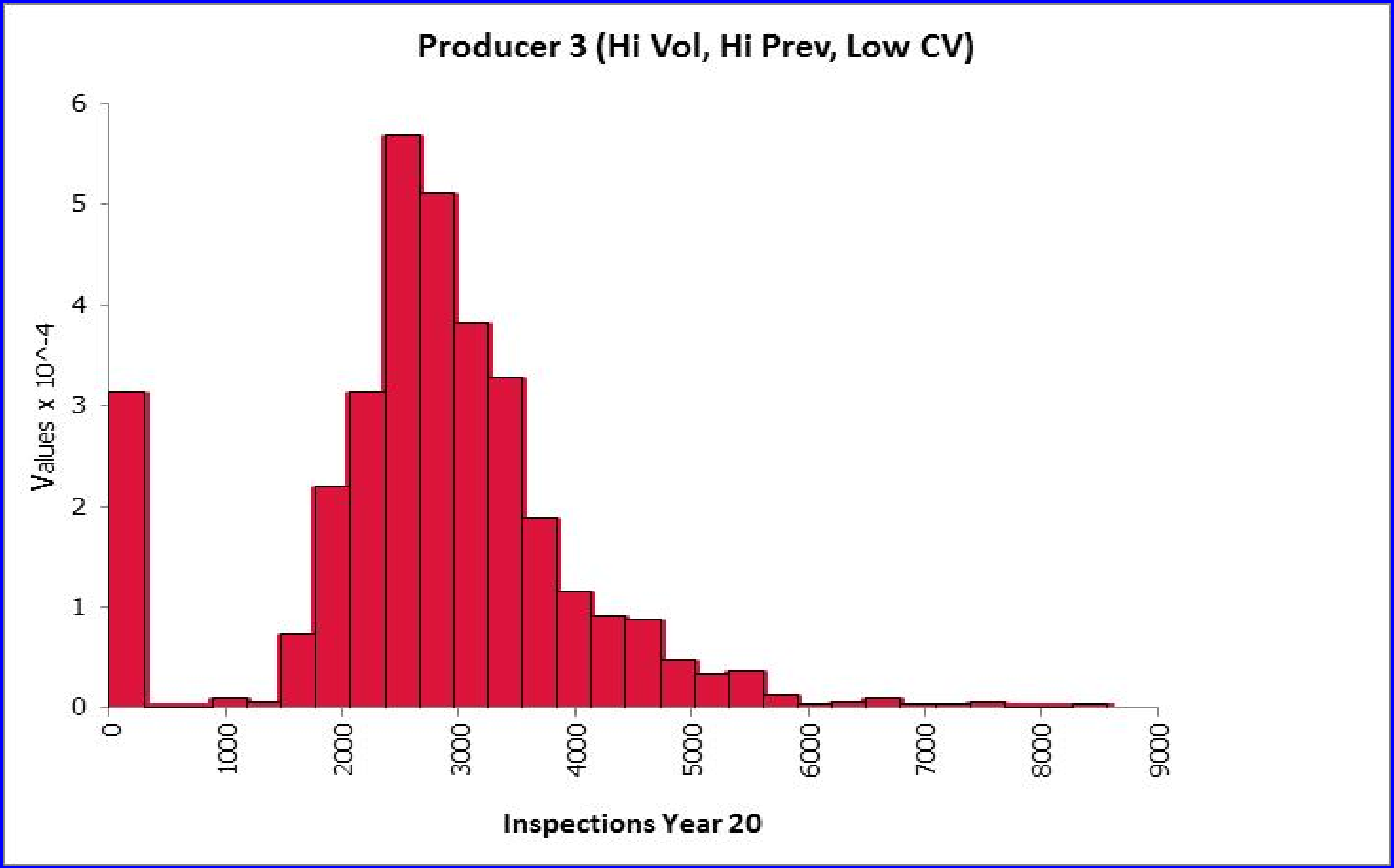
- Lot prevalence (p_i) ~ Beta(μ_i, σ_i)
- Freq. of lot inspxn ~ 1% (budget constraint)
- $\Sigma_i I_{ij}$ (lot inspxns)/yr. ~ 9,990
- p_{detxn} = 78.5% (e.g., $p_{w/in}$ = 5%, n/lot = 30)

Optimal vs. Simple: Simulation

- Simulation of Contaminated Lots & Inspection
  - #contam lots (c_{ij}) ~ Binomial(L_i, p_i)
  - #contam lot inspected (ci_{ij}) ~ Hypergeo(I_{ij}, c_{ij}, L_i)
  - #contam lots detected (x_{ij}) ~ Binomial(ci_{ij}, p_{detxn})
- Optimal Allocation: $I_{ij} \propto L_i * \hat{p}_{ij}$
  - s.t. $I_{i1} = 9990/27 = 370$; $1 \le I_{ij} \le L_i$ for j = 2, …, 20
  - $\hat{p}_{ij} = \frac{\sum_{t=1}^{j-1} x_i}{\sum_{t=1}^{j-1} I_i}$
- Simplified Allocation: $I_i \propto L_i$

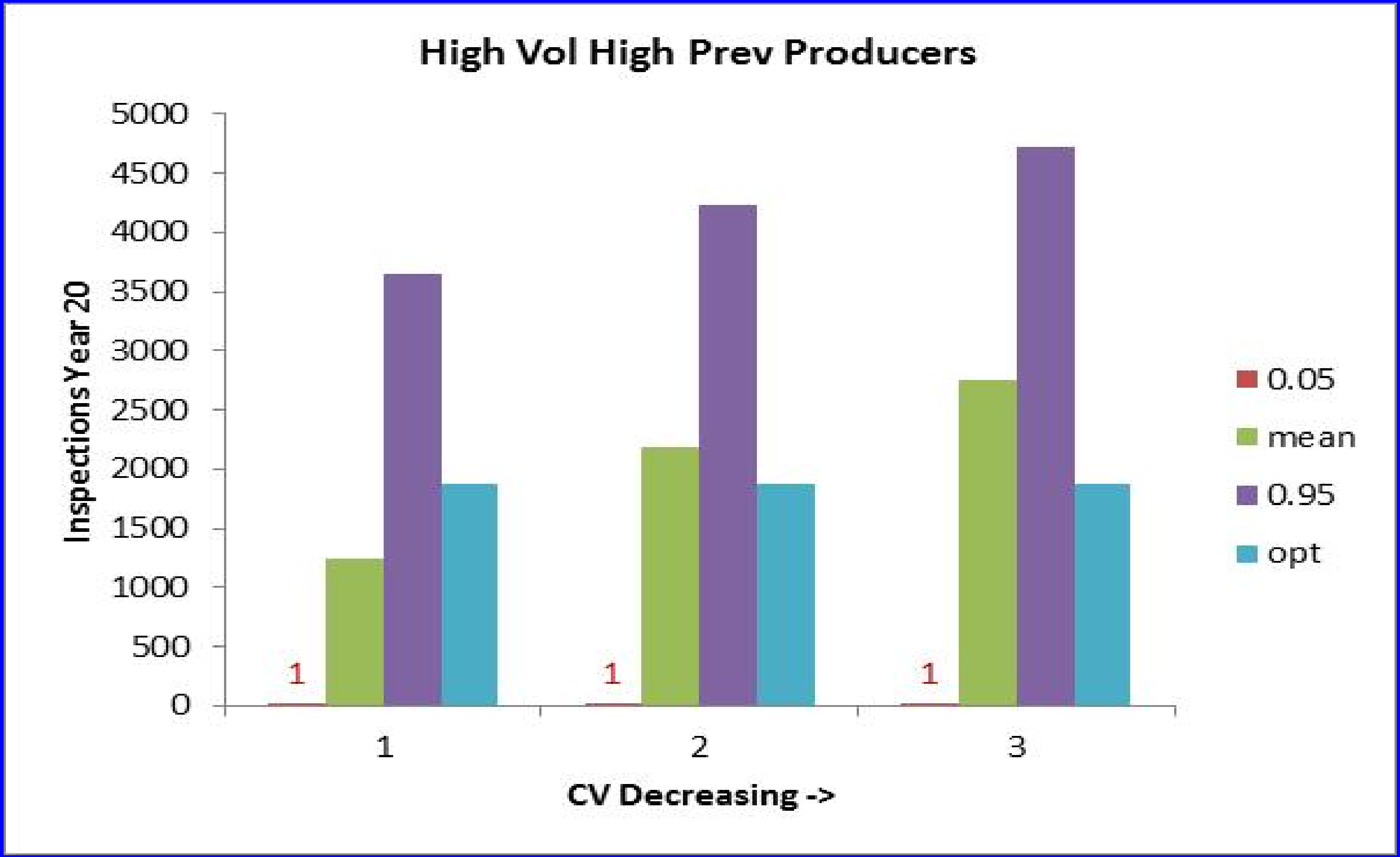
Optimal vs. Simple: Simulation

Optimized Allocation



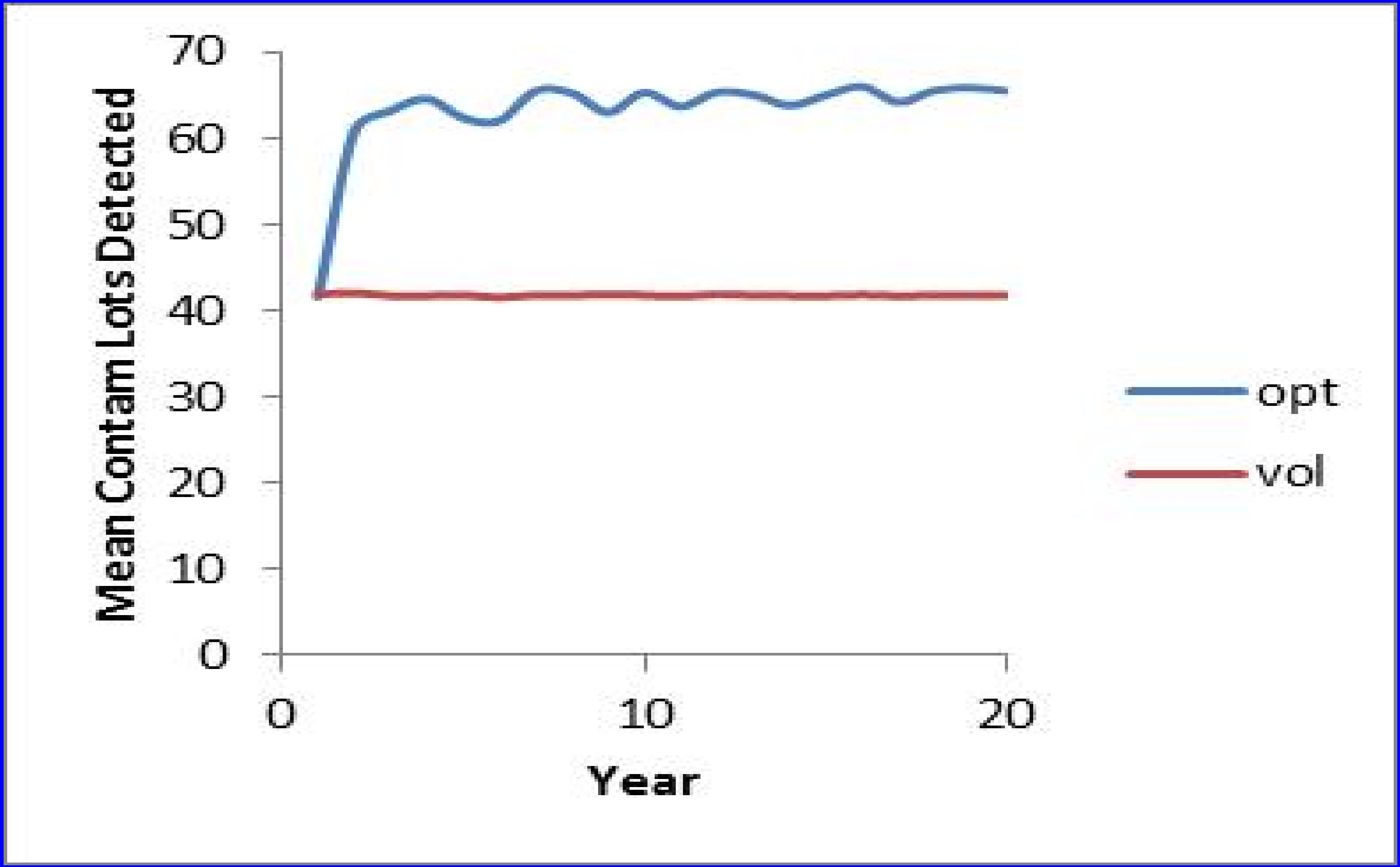
Optimal vs. Simple: Simulation

Optimized Allocation



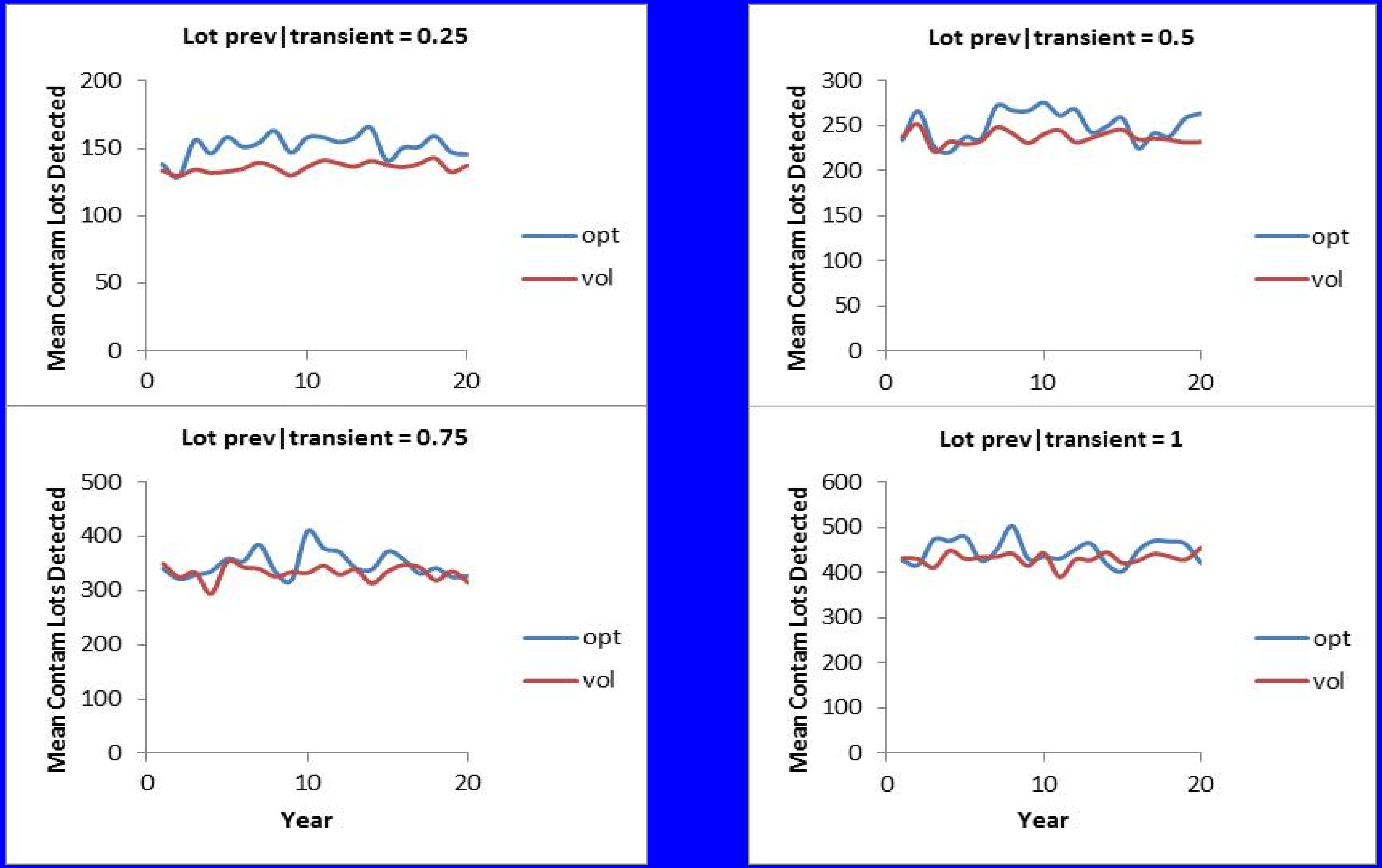
Optimal vs. Simple: Simulation

Process Stationary over 20 years



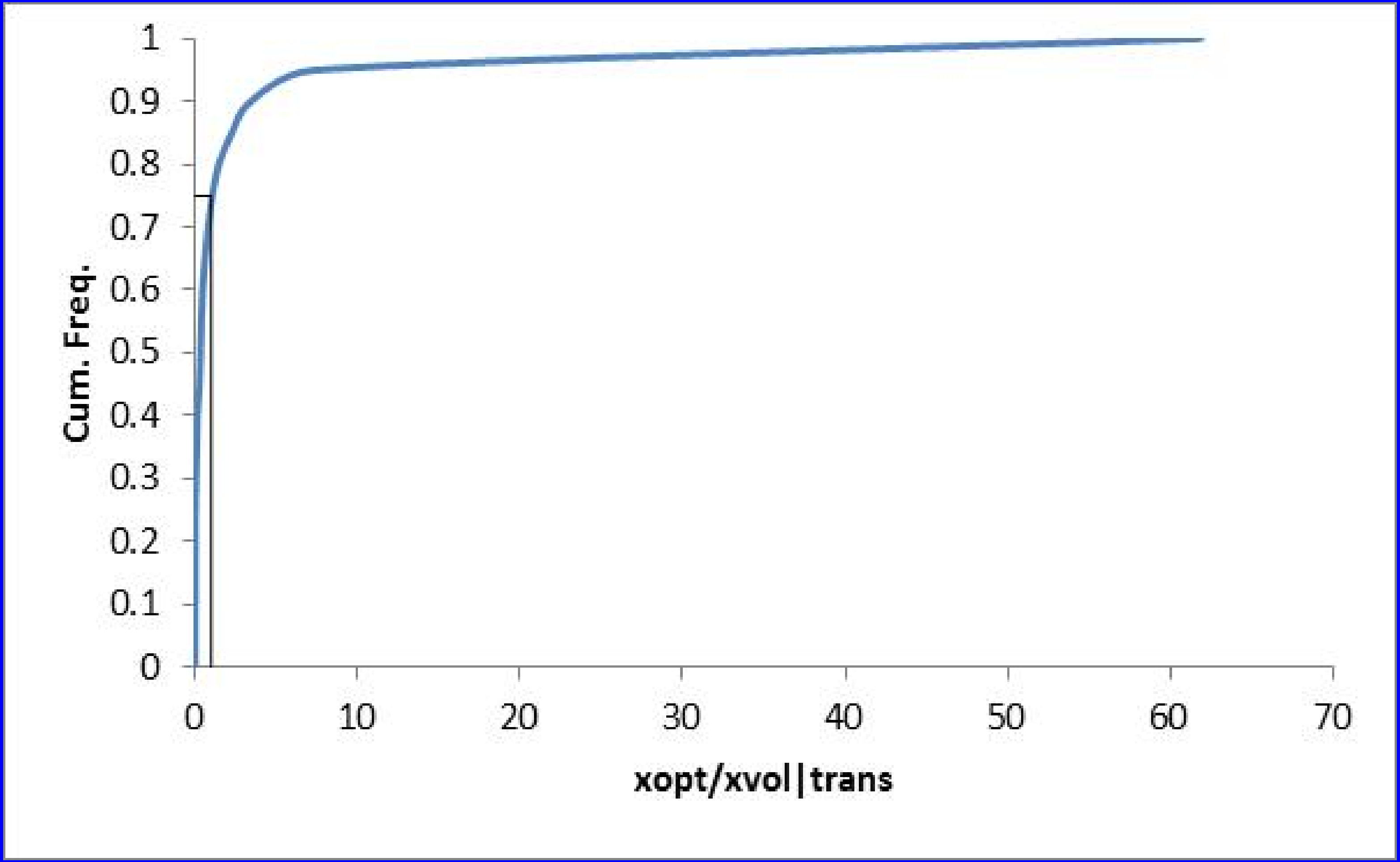
Optimal vs. Simple: Simulation

Producer's annual p(transient) = 0.05/yr. (1/20 years)



Optimal vs. Simple: Simulation

Producer's annual p(transient) = 0.05/yr. (1/20 years)



Optimal vs. Simple

- Gigerenzer. 2011. Heuristic decisionmaking. Annual Review of Psychology. 62: 451-482.
- Pflug et al. 2012. The 1/N investment strategy is optimal under high model ambiguity. Journal of Banking & Finance. 36: 410-417.
- Not all simple heuristics will outperform optimization.

Disclaimers

The opinions expressed herein are the views of the author and do not necessarily reflect the official policy or position of the United States Department of Agriculture. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government.

Thank you