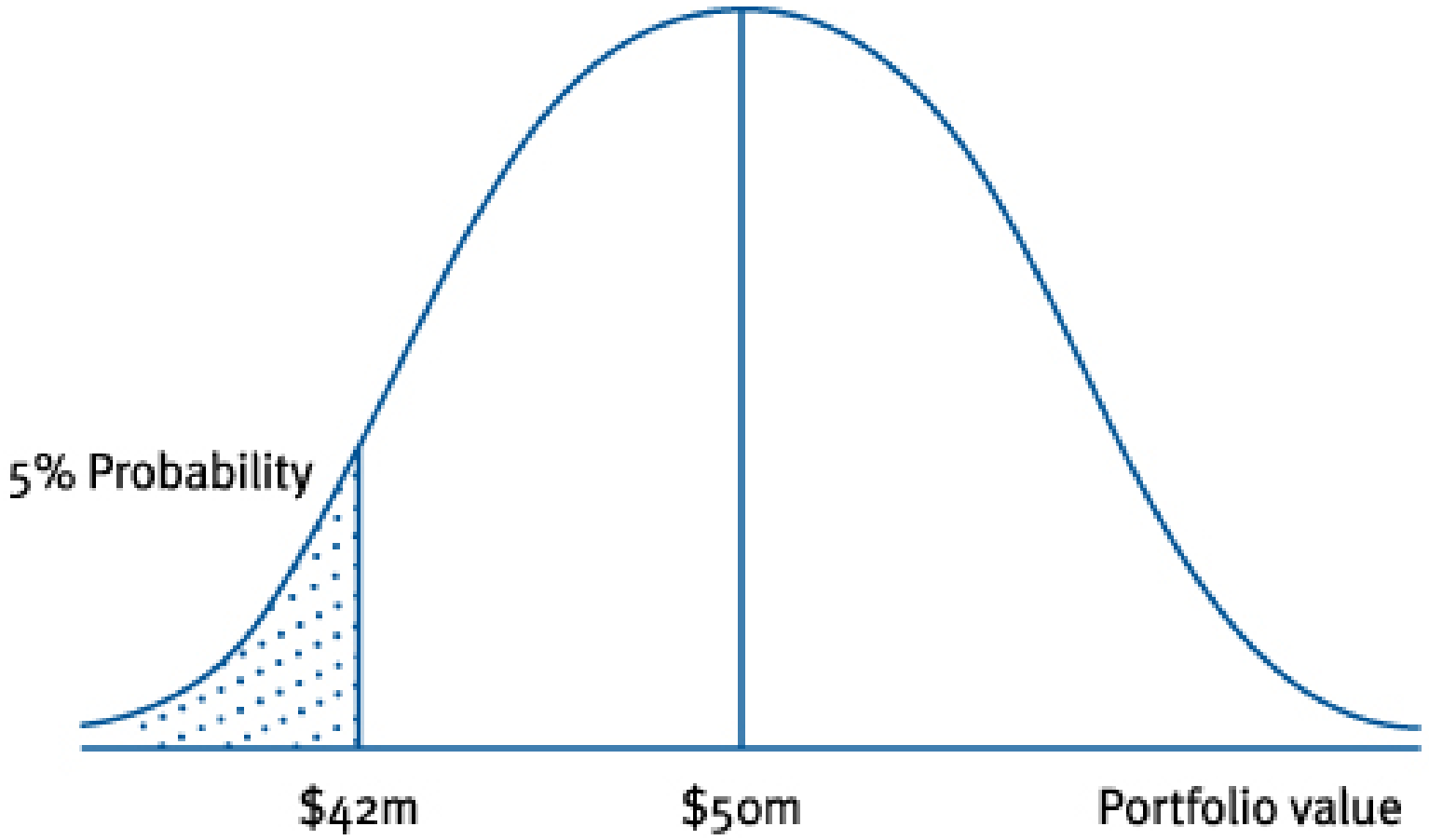
Quantifying and Mitigating Turnover-Induced Knowledge Loss

Rigby, Zhou, Donadelli, Mockus
Concordia University
University of Tennessee

Quantifying Turnover Risk

- Damage from past losses
  - Measures: Knowledge at Risk (KaR) and ES
- Potential for future losses
  - Historical and Monte Carlo Simulations
- Suggest possible successors
  - Developer-to-File and File-to-File matrices
- Adapt financial risk measures to software

Value at Risk



Data

- Version history of Chrome and Avaya project
- Can’t use number of commits
- Extract the blame information in each quarter
  - Which devs own which lines
- File is abandoned when 90% or more of the owners have left

Avaya Loss Distribution

- Mean 209
- Median 130
- KaR_95 = 642
- ES = 797
- Not normally distributed
- Unexpected loss is 3.8 times larger than mean

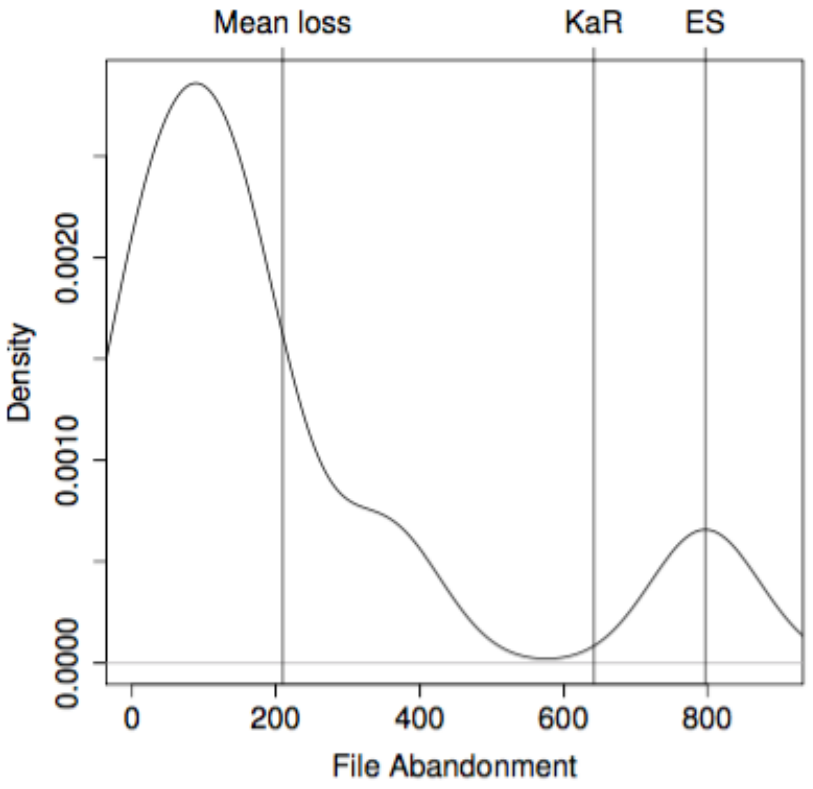


Figure 1: Avaya Loss distribution

Historical Simulation

1. the knowledge distribution of each quarter
2. num developers who leave for each quarter
3. stratified random sampling to choose the same number of leavers, rnd(lc,lnc)
4. actual knowledge loss compared to simulated loss

Historical Simulations

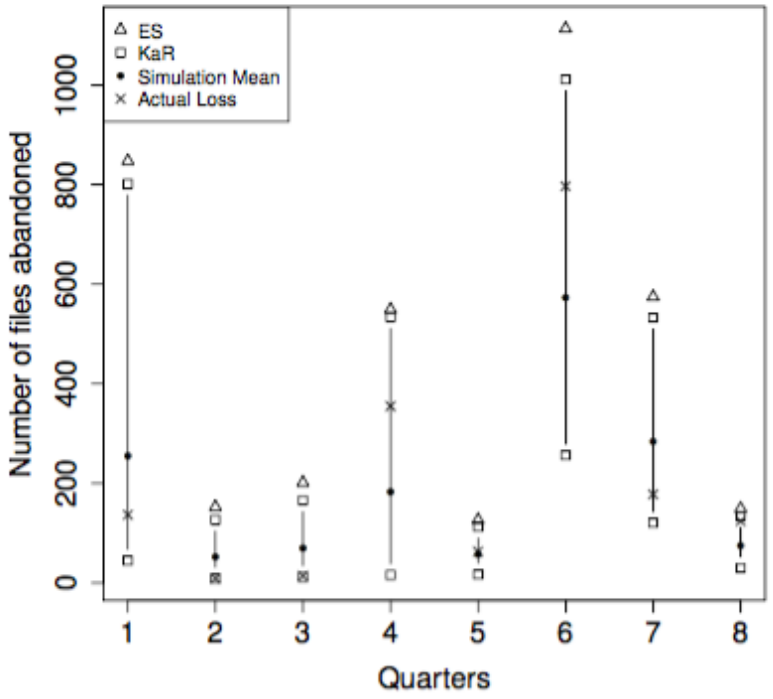


Figure 3: Historical Simulation for Avaya

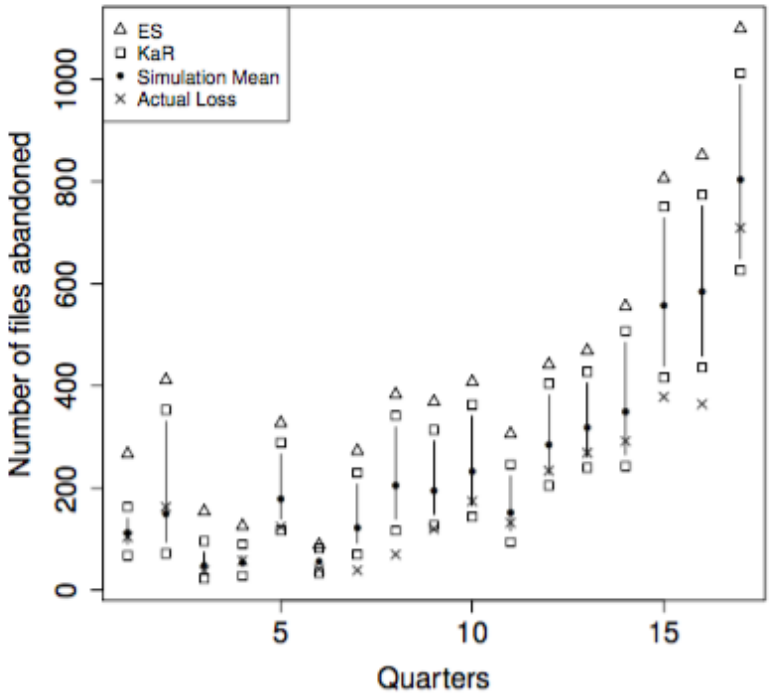


Figure 4: Historical Simulation for Chrome

Monte Carlo Simulations

1. knowledge distribution of the most recent quarter
2. vary the group size, g, of developers who leave from 1 to 200 people
3. we select groups of developers to leave at random, rnd(g). We do 1000 runs for each group size.
4. compare to actual and historical

Monte Carlo Simulations

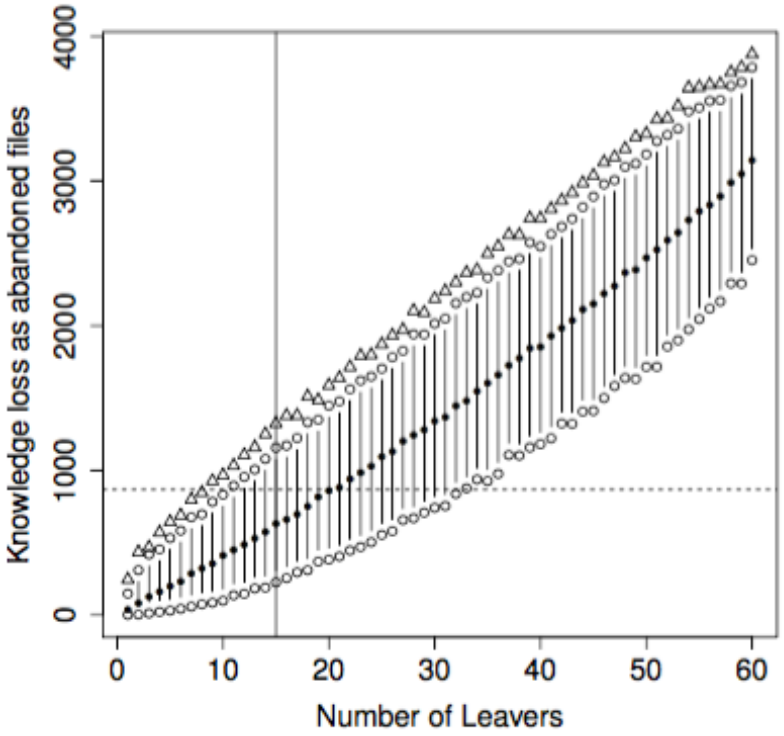


Figure 5: Truck Factor Simulation Avaya

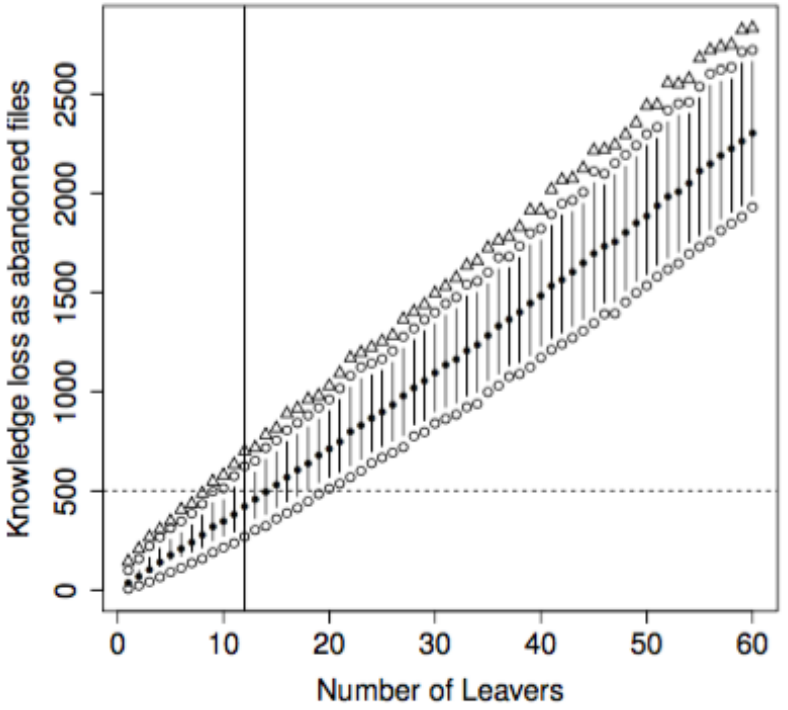


Figure 6: Truck Factor Simulation Core Developers on Chrome

Possible Successors

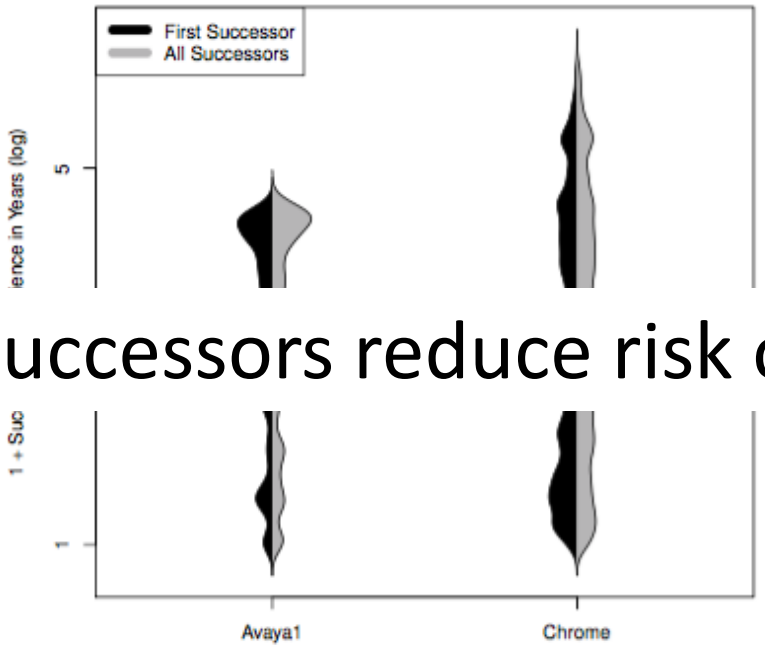


Figure 7: Experience of actual successors

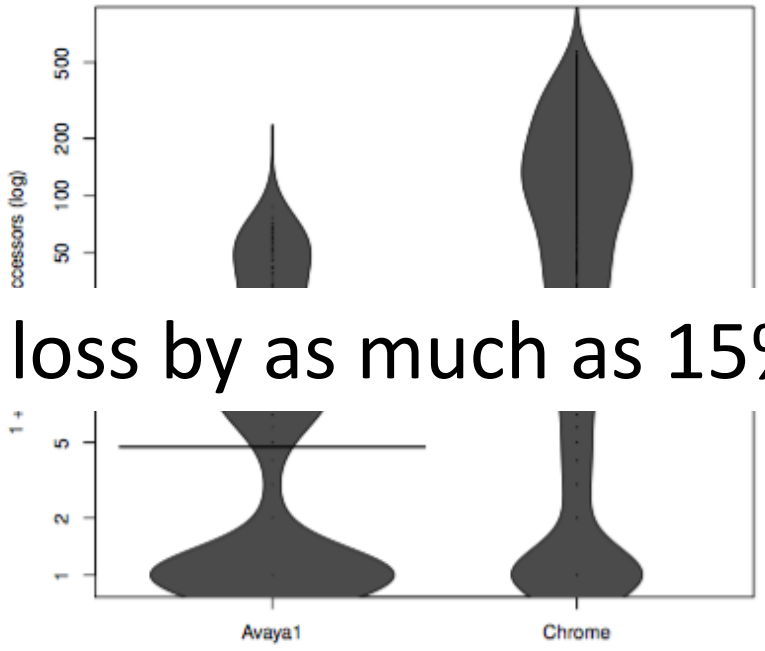


Figure 8: Number of potential successors for an abandoned file

Successors reduce risk of loss by as much as 15%

Correct successor: 48% and 34% of the time, Avaya and Chrome

Future Work

- Which changes to the knowledge distribution would lead to the greatest reduction in risk
  - ie who should work with whom to reduce risk?
- Software project stress testing through simulation

Conclusions

- Adapt financial risk measures to software
- Assess the amount of Knowledge at Risk (KaR) from turnover at Avaya and Chrome
  - 3 times larger than expected loss
- Historical and Monte Carlo simulations of loss
  - “Truckfactor” exaggerates loss
- Suggest possible successors
  - Successors reduce risk of loss by as much as 15%