

# The link between default and LGD

**Jon Frye**

**[Jon.Frye@chi.frb.org](mailto:Jon.Frye@chi.frb.org)**

**[http://www.chicagofed.org/webpages/people/frye\\_jon.cfm#](http://www.chicagofed.org/webpages/people/frye_jon.cfm#)**

**It is easier if you Google me  
Federal Reserve Bank of Chicago**

**PRMIA-NY**

**New York, 15 October 2013**

**Any views expressed are the author's and do not necessarily represent the views of the management of the Federal Reserve Bank of Chicago or the Federal Reserve System.**

# **An unusual tactic**

**I don't try to convince you**

- that I have superb insight into what's what,**
- that I have discovered an eternal truth,**
- or that what I say can't ever be improved upon.**

# This is a simple idea that works

The link between default and LGD introduces nothing new:

$$LGD = \Phi \left[ \Phi^{-1}[DR] - \frac{\Phi^{-1}[PD] - \Phi^{-1}[EL]}{\sqrt{1 - \rho}} \right] / DR$$

$\Phi[\cdot]$  = CDF of Standard Normal = Norm.S.Dist[ $\cdot$ ] in Excel

By contrast, regression has two unknowns,  $a$  and  $b$ :

$$LGD = a + b DR$$

# **Sidebar 1: Expectations and risk**

**Banks need to estimate PD and ELGD for many purposes.**

**I say NOTHING about how to do that.**

**PD (the expected default rate) might depend on many things.**

**ELGD (the expected LGD rate) might also.**

- **Detailed data, like offered by PECDC and others, helps.**

**I do say: Conditions that produce a high default rate  
also produce a high LGD rate.**

**If you have a model of the default rate, that's all you need.**

**Put the default rate into the LGD function; it tells the LGD rate.**

**No additional statistical work is needed.**

# Topics

**The problem of systematic LGD risk (4 slides)**

**Earlier models of LGD risk (1 slide)**

**Historical study of LGD function (2 slides)**

**Simulation study of LGD function (8 slides)**

**Quarterly conditional loss forecasts (3 slides)**

**Conclusion (1 slide)**

# Always: Loss Rate = DR \* LGD

## A portfolio containing 10 loans of equal exposure

Loan #:	1	2	3	4	5	6	7	8	9	10
Exposure:	\$10	\$10	\$10	\$10	\$10	\$10	\$10	\$10	\$10	\$10
Default?:	0	0	0	1	0	1	0	0	1	0
Loss:	0	0	0	\$1	0	\$2	0	0	\$3	0
LGD:	---	---	---	10%	---	20%	---	---	30%	---

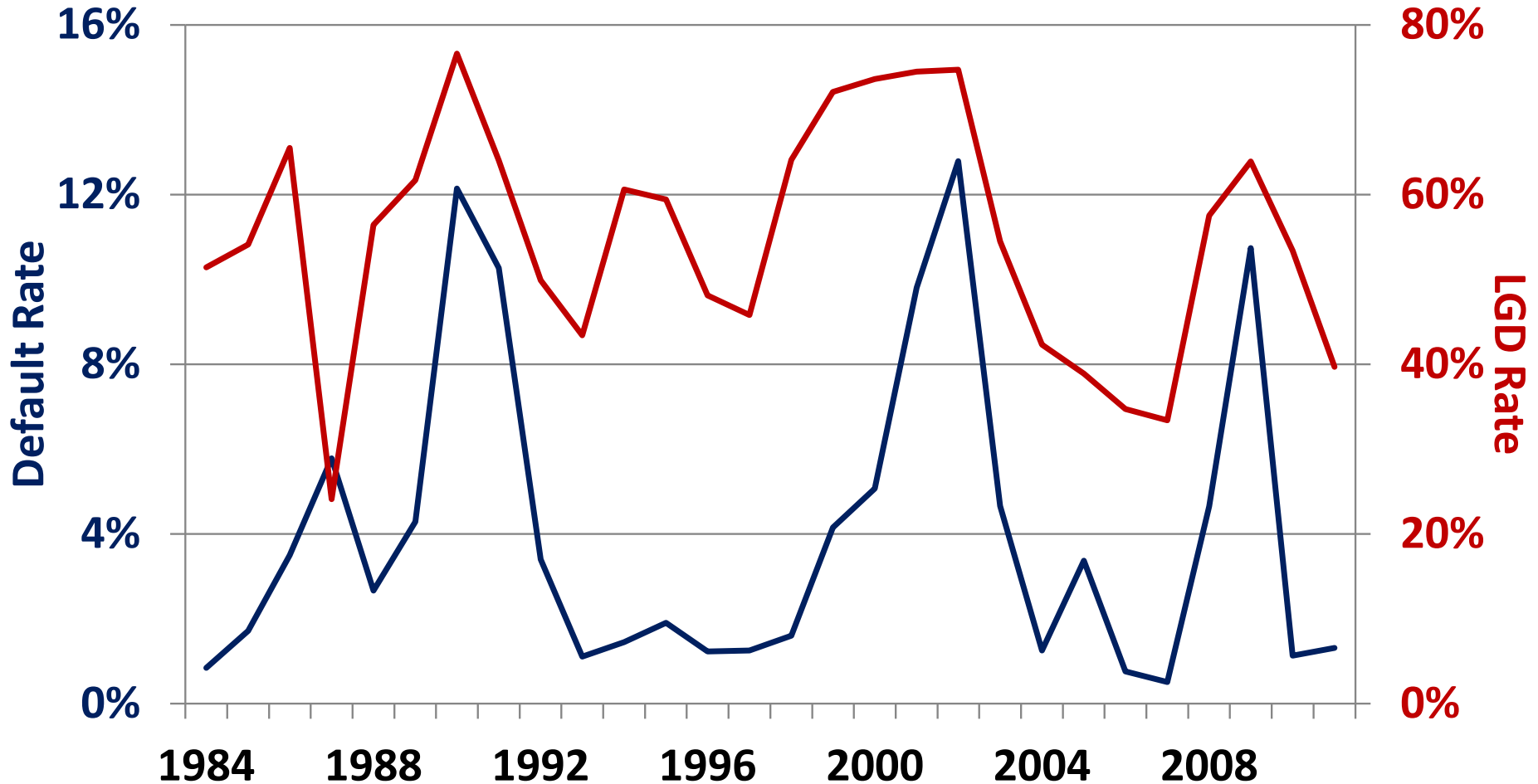
**Loss rate = \$6 / \$100 = 6%**

**Default rate = 30%; LGD rate = 20%**

**Loss rate = DR \* LGD = 30% \* 20% = 6%**

**Taking expectations, EL = PD \* ELGD**

# When DR is high, LGD is high



# Sidebar 2: The Data Problem

**The previous chart goes back 28 years to 1984.**

**Before 1984 there weren't many bonds, defaults or LGDs.**

**Nobody kept track of loans. Earlier data doesn't tell much.**

**Of the 28 years, 23 of them are low-default years.**

**Then, portfolio LGD is the average of a few noisy losses.**

- **Portfolio LGD might be anything. Such a number doesn't tell us much.**

**The 5 high-default years come from only 3 recessions.**

**Because of serial dependence, information content is less than 5.**

**Therefore, what we know about the link between default and LGD is based on very little actual information.**

**It is very unlikely that we can improve on the simplest LGD model.**



# Risk affects everything

## Economic capital

The more LGD moves, the more capital you need.

## Risk and reward

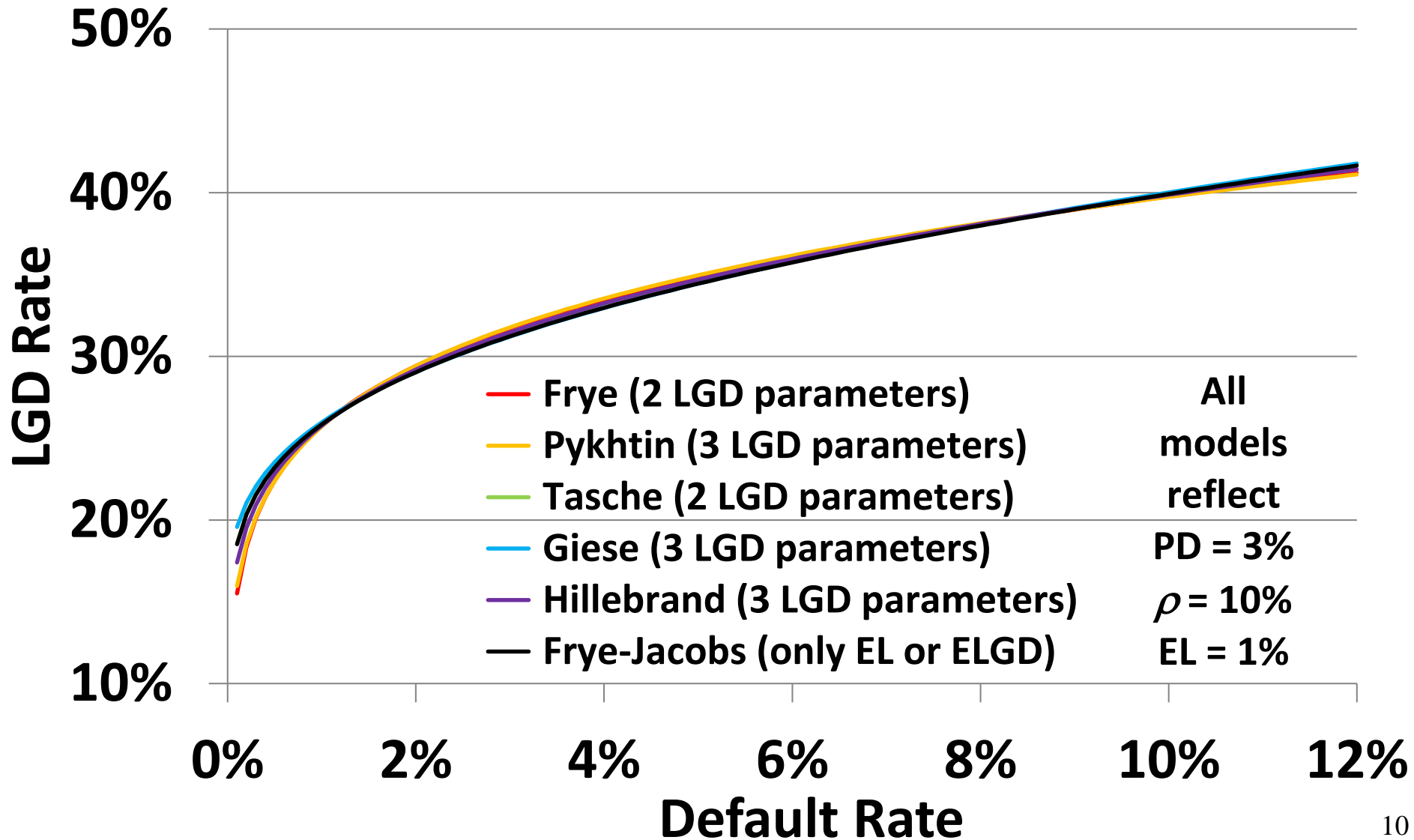
If a loan has more LGD risk, a lender wants more reward.

## Pricing

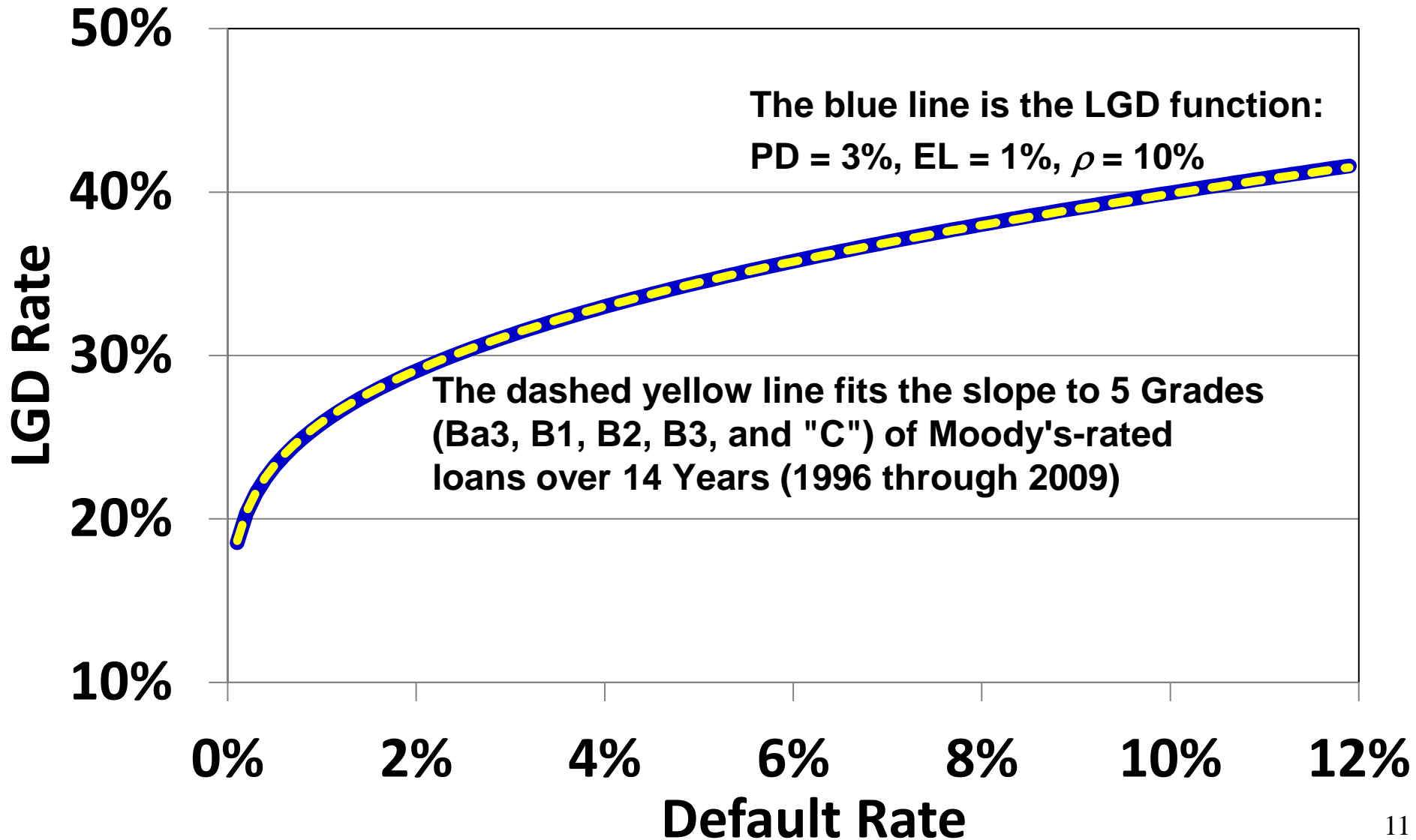
Even if expected LGDs are equal, different loans can default in different conditions having different LGDs.

The question: how much LGD risk exists?

# All models can agree



# LGD function fits Moody's data



# Summary: It works

**Compared to the LGD function, other functions do not exhibit statistical significance**

- when fit to all loans, to all bonds, or to all instruments;
- when fit to each combination of grade and seniority; or
- when using any of four different alternative LGD functions.

**Stating it as a positive:**

**The LGD function fits Moody's historical data.**

# Simulation study

**A recent study\* simulates data under perfect conditions.**

**The data are then analyzed:**

- with a linear regression model, and
- with the nonlinear LGD function.

**The LGD function performs better.**

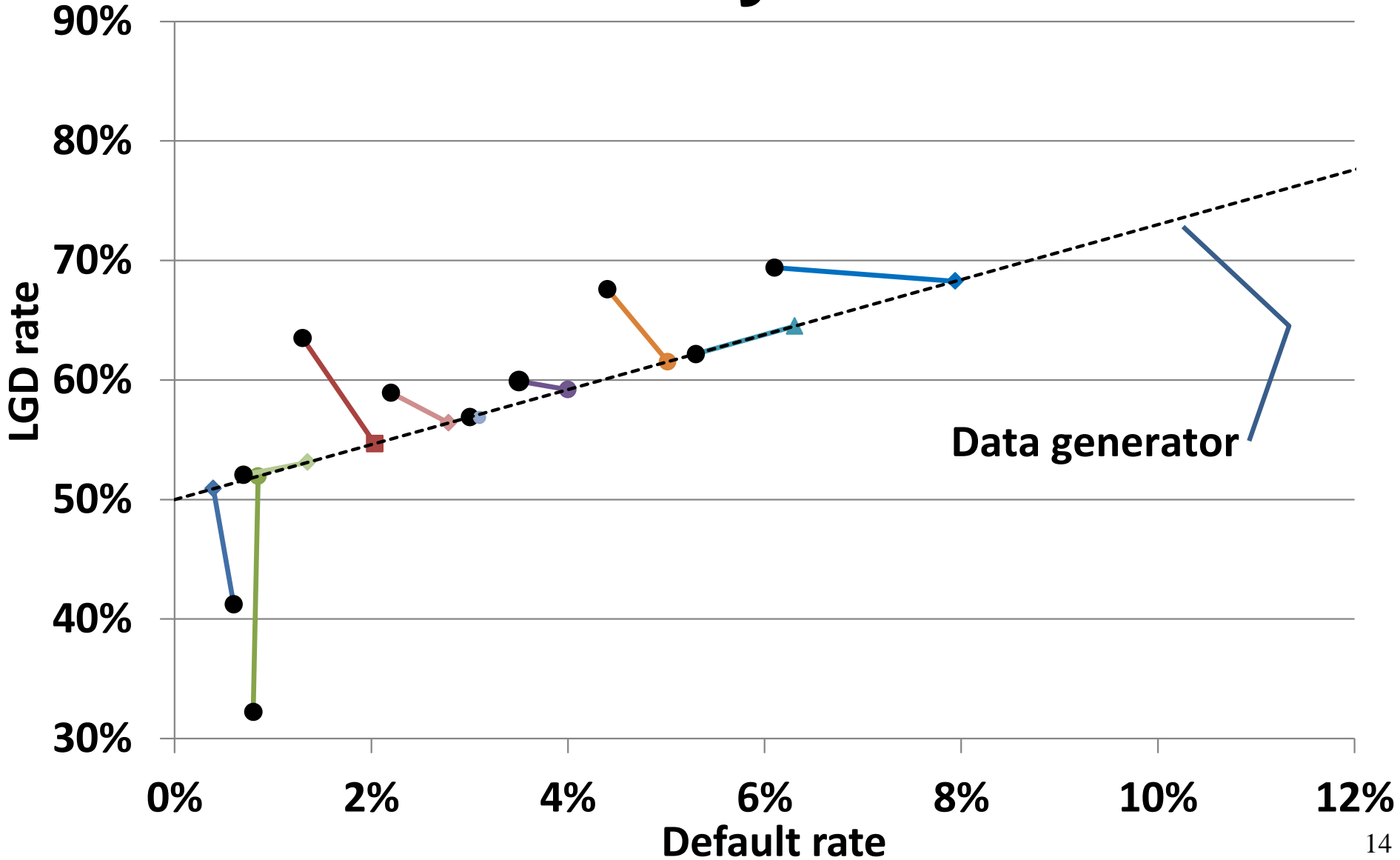
**Reason: Regression does not have enough data to work well.**

- It might take a century or more for regression to win.

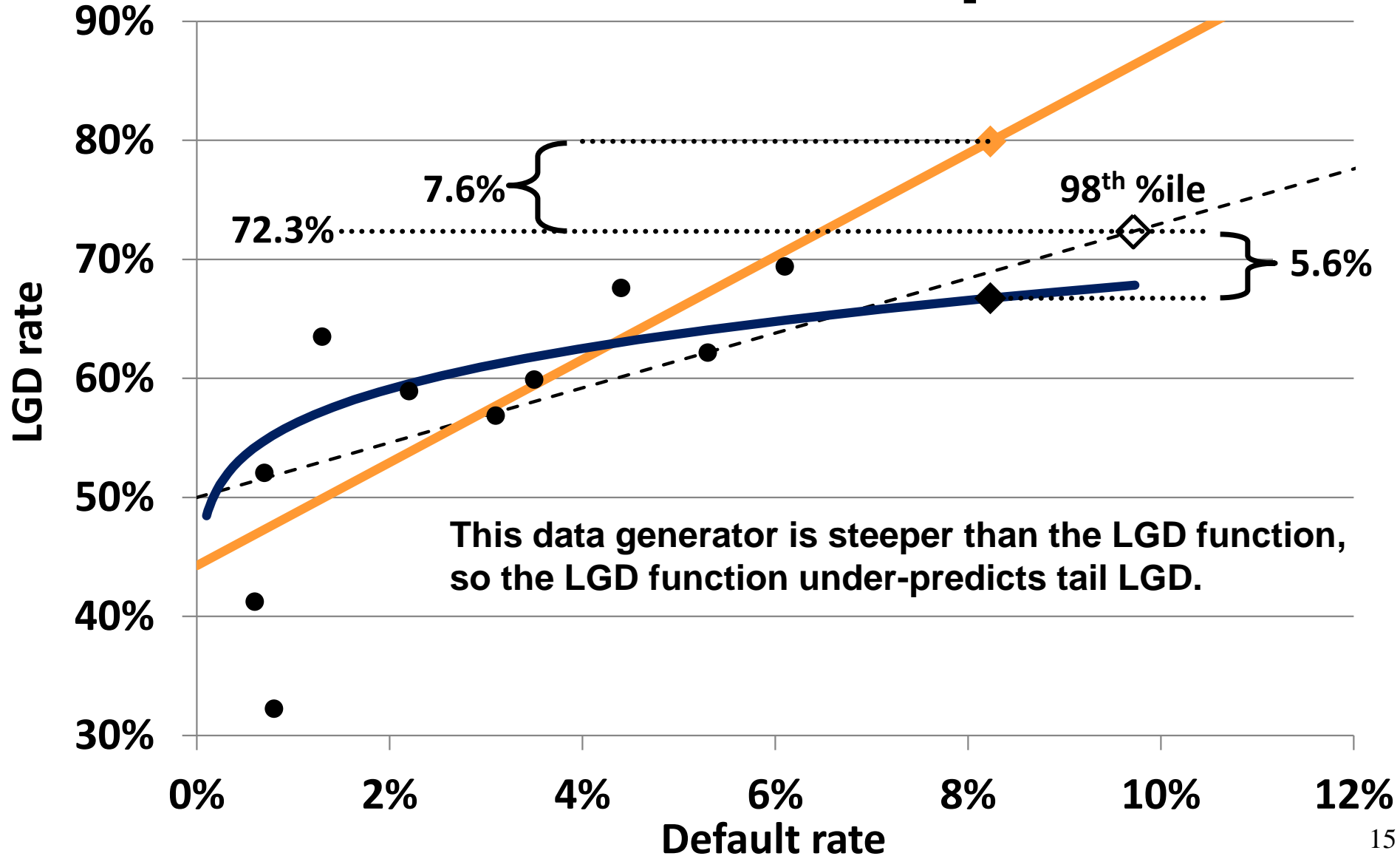
**Here's an example of one simulation run...**

**\* Frye, "LGD as a function of the default rate", under review**

# The lucky bank



# The LGD function outperforms



# Sidebar 3: Regression

**Regression fits the data better than the LGD function.**

This is always the case!

- The LGD function needs only one thing in order to work: ELGD.
- The regression needs two things: slope and intercept.

If you put in more things, your model always works better.

**Regression is being misled by the data.**

The regression thinks that the data are a good guide to the data generator, but they aren't. Random data are too steep in this case.

**Regression "over-fits" the data.**

The regression line fits the historical data, but future data is more likely to show up near the data generator, which is much lower.



# Sidebar 4: The LGD function

The LGD function does not predict that data is nearby.

If data were always near the LGD function, the data itself would show a moderate, positive relationship like the function does.

- Regression would "see" this relationship and do an adequate job.

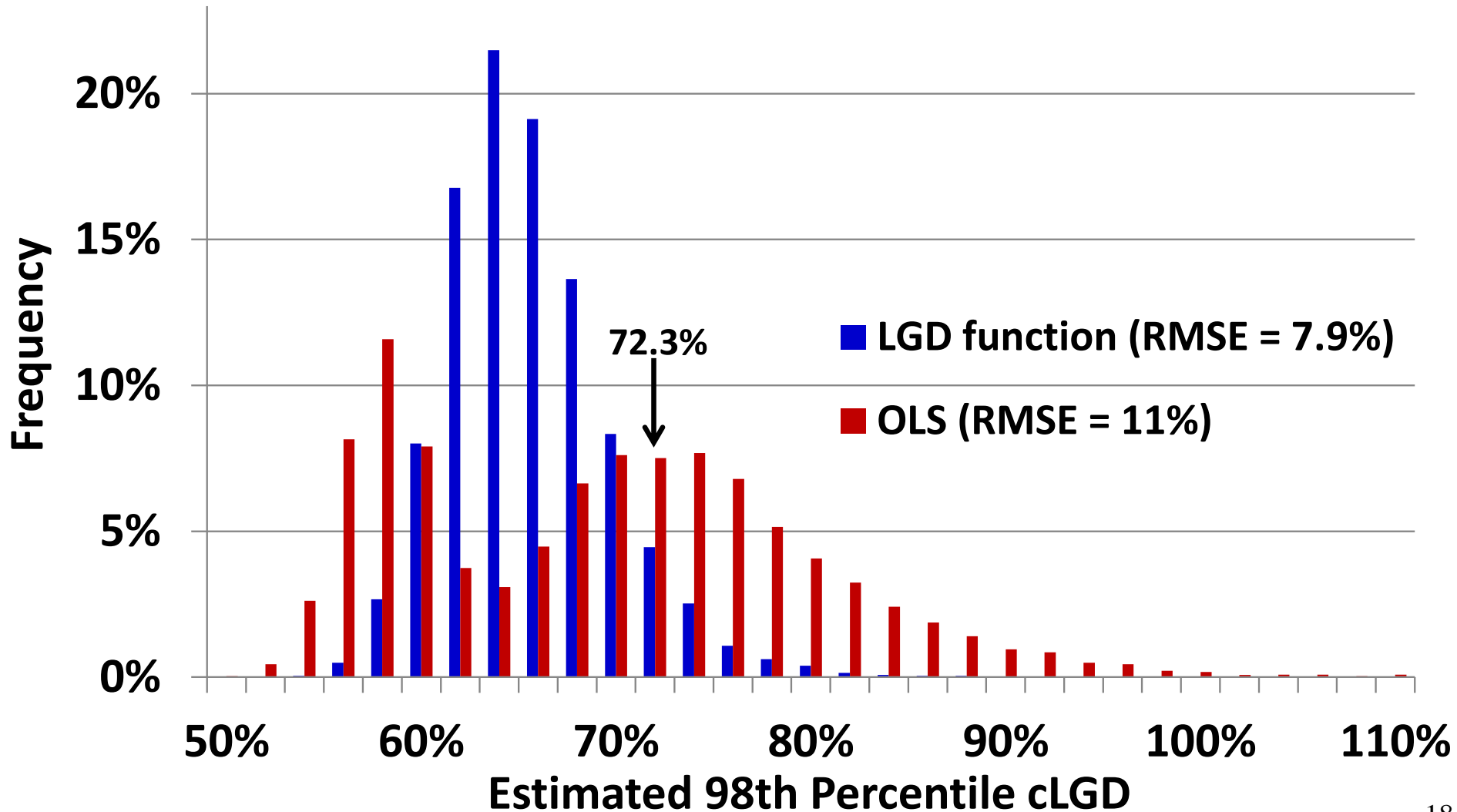
The LGD function contributes something when the data do not look like the LGD function, as was illustrated.

Very steep or very shallow data are likely to come from pure randomness, and the LGD function doesn't care about that.

- Only ELGD matters to the LGD function.

The next slide summarizes 10,000 simulation runs...

# On 10,000 runs, the LGD function performs better on average



# Other simulation experiments

A skeptic wonders whether the LGD function would outperform regression under other conditions.

I tried different values of all eight control variables.

Only two variables matter to the conclusion:

- the slope of the data generator,
- and more importantly, the number of years of data.

In all the experiments, the data generator is linear.

If the data generator were curved (like the LGD function or earlier models) regression would do worse in comparison.

# How to beat the LGD function

## Have more LGD risk.

The LGD function under-predicts if the data generator is very steep.

- This gives regression a chance to win.

## Have more years of data.

Beware: "more" can mean more than 100 years.

- The data requirement depends on the steepness of the data generator.

And: "years" means independently drawn data points.

- Real-world data are not independent. They are serially dependent.
- In the real world it takes longer for regression to win.

# **Dodd-Frank annual stress tests**

**There is interest in using the LGD function for DFAST.**

**This involves a new wrinkle, if the LGD function were determined appropriate for this supervisory purpose.**

## **Recall:**

- The views I express are mine.**
- They do not necessarily represent the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.**

# Loss depends on the economy

**The new wrinkle: The economy changes slowly.**

At a quarterly rate, it changes more slowly.

In smoothed scenario, it changes more slowly yet.

**There is little dispersion in a quarterly model, so  $\rho \approx 0$ .**

**To find credit loss in a future quarter:**

(1) In the LGD function, set PD and ELGD to their values today.

- Set  $\rho$  equal to zero.

(2) Use your model to find "stressed PD" in the future quarter.

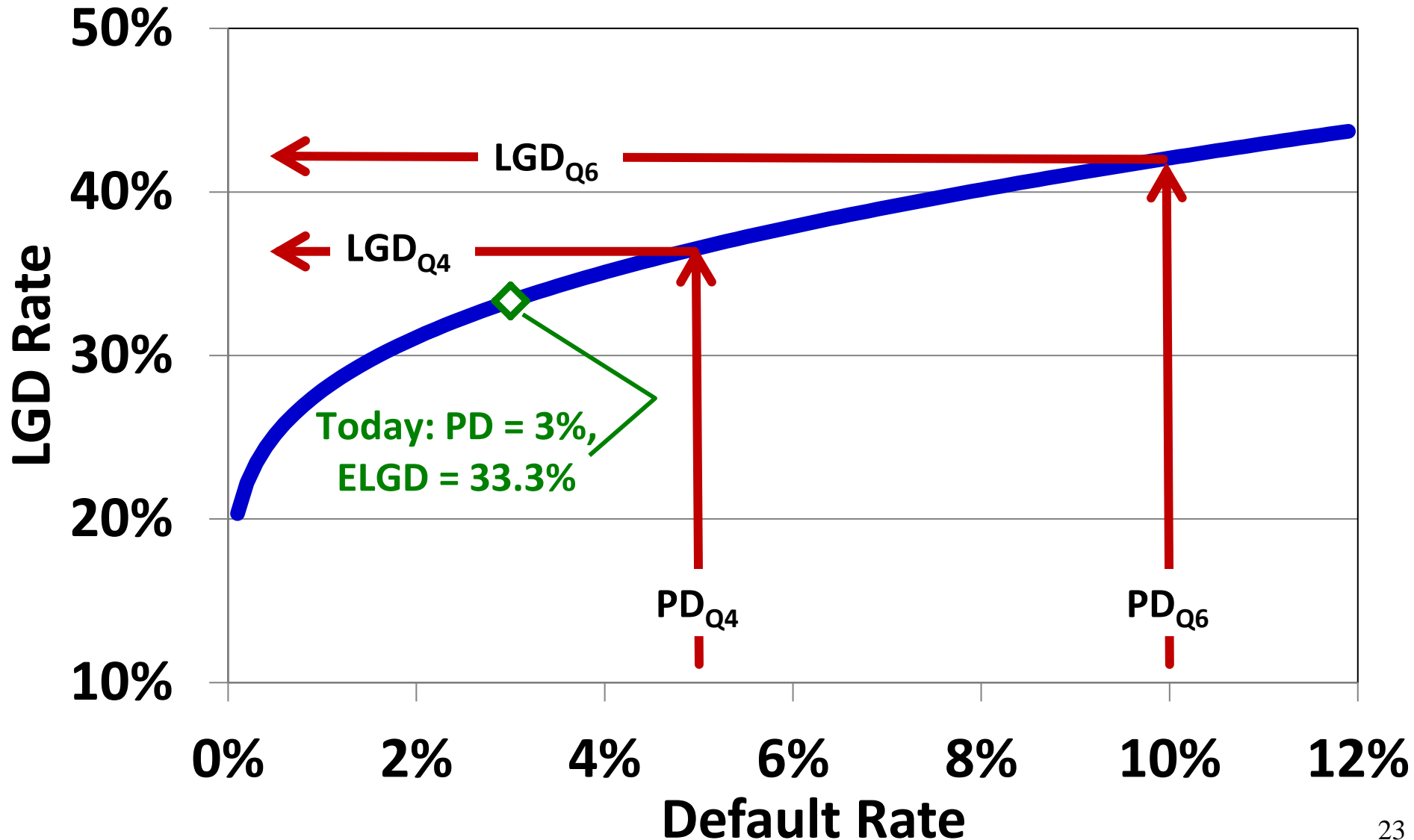
(3) Plug this into the LGD function where you see "DR".

(4) The function tells LGD in that quarter; Loss = DR \* LGD.

**Net, the loss scenario rides along a fixed LGD function.**

The LGD function stays where it is from quarter to quarter...

# Example of quarterly projections



# LGD as a function of the default rate

The LGD function is easy to apply.

It has no new parameters.

It attributes moderate LGD risk to every exposure.

It works well.

It is consistent with historical data on loans and bonds.

It outperforms linear regression in simulation experiments.

- It is likely to continue to outperform for a long time.

Don't imagine that a risk model is significant, unless it demonstrates significance in a statistical test.

If the elaborate model is not significant, use the simple one.

- Otherwise, you are guessing.



# References

Frye, "LGD as a function of the default rate,"

[http://www.chicagofed.org/webpages/people/frye\\_jon.cfm#](http://www.chicagofed.org/webpages/people/frye_jon.cfm#)

Frye and Jacobs, "Credit loss and systematic loss given default," *Journal of Credit Risk* (1–32) Volume 8/Number 1, Spring 2012.

Frye, Jon, "Depressing Recoveries," *Risk*, November 2000, 108-111.

Giese, Guido, "The impact of PD/LGD correlations on credit risk capital," *Risk*, April 2005, pages 79-84.

Hillebrand, Martin. "Modeling and estimating dependent loss given default", *Risk*, September 2006, pages 120-125.

Pykhtin, Michael. "Unexpected recovery risk," *Risk*, August 2003, pages 74-78.

Tasche, Dirk. "The single risk factor approach to capital charges in case of correlated loss given default rates," Working Paper, 2004, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=510982](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=510982)

# Questions?

**Thank you for your attention**