Predictive Analytics

Brian Hartman, PhD ASA
Assistant Professor of Statistics and Actuarial Program Director
Brigham Young University

March 20, 2018

ILTCI Mobile App Download Instructions

iPhone
iPad

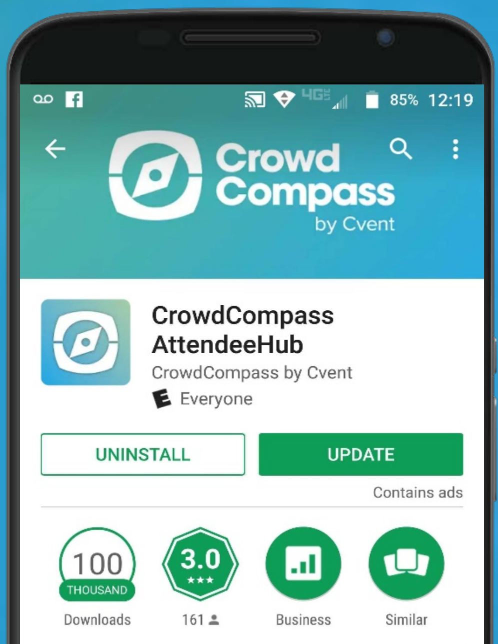
1) Type https://crowd.cc/s/1flyo in web browser
2) Click "Download iPhone/iPad App" to load Apple's App Store and download the app.

android

1) Type https://crowd.cc/s/1flyo in web browser
2) Click "Download Android App" to load the Google Play Store and download the app.

BlackBerry

1) You'll be using the web version of the app. Open the web browser, click the BlackBerry menu button, select "Go To" and type https://crowd.cc/s/1flyo.

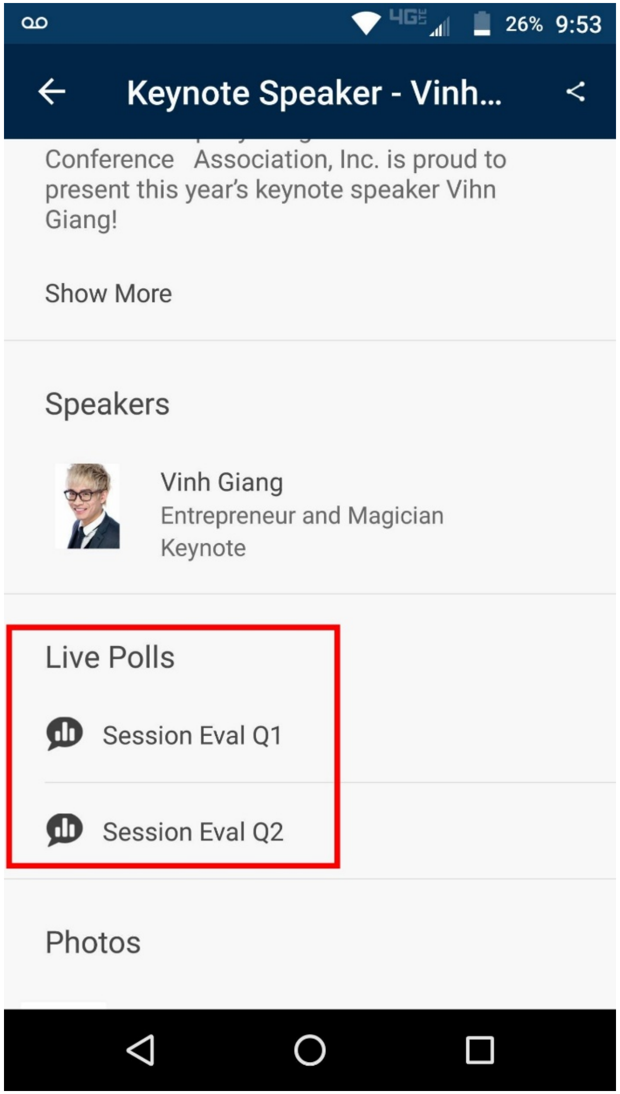


You can also just go to your app store and search 'AttendeeHub'. Once installed search 'ILTCI' and you'll find our app.

A Special Thank You to this year's Mobile App Sponsor

Nationwide®
is on your side

Session Survey Instructions



Once you are in the app go to the schedule and the session you are in.

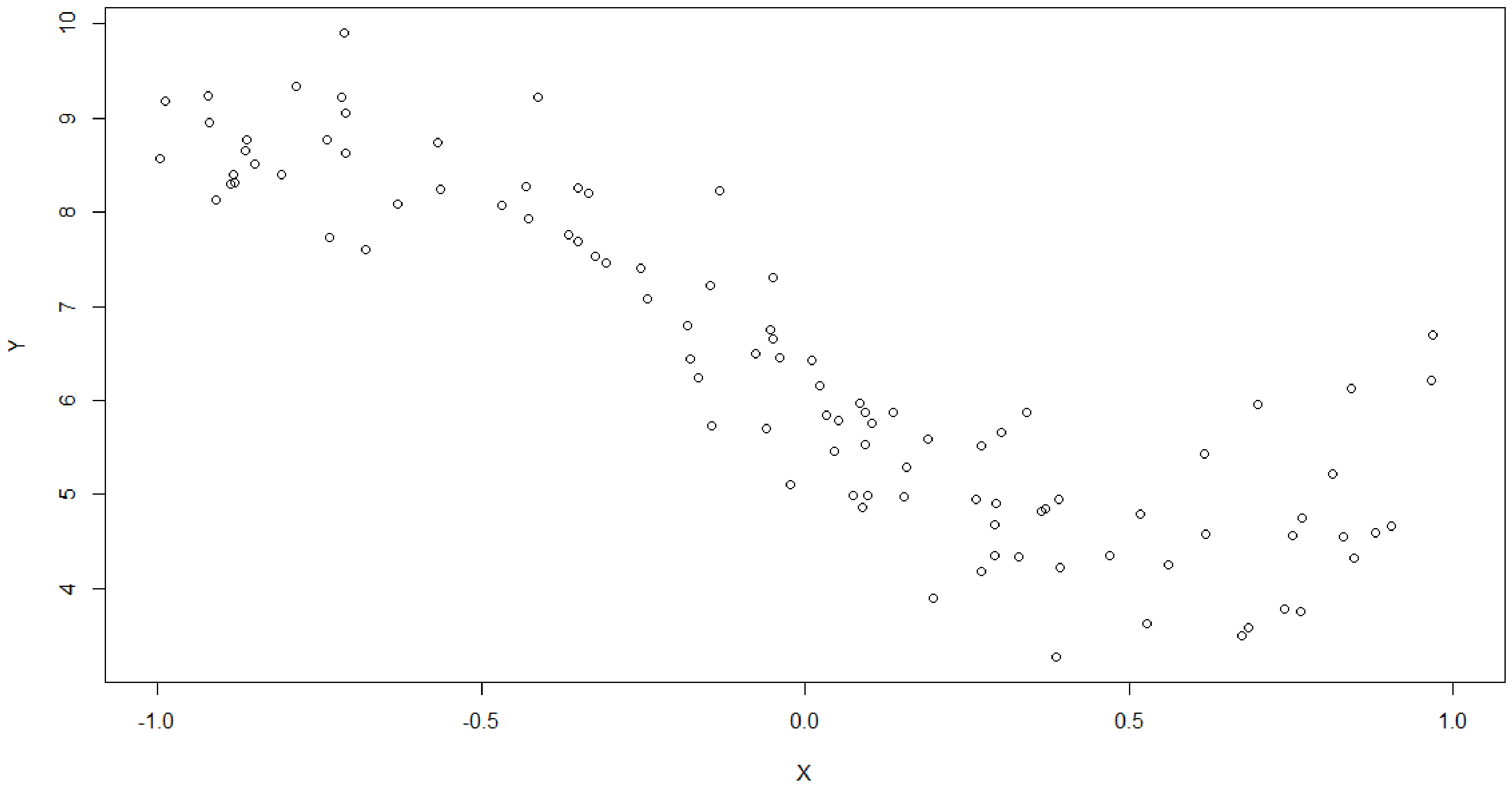
Scroll to the bottom to find the Live Polling questions.

This year the session survey questions can be found in this section and will take just a couple seconds to complete.

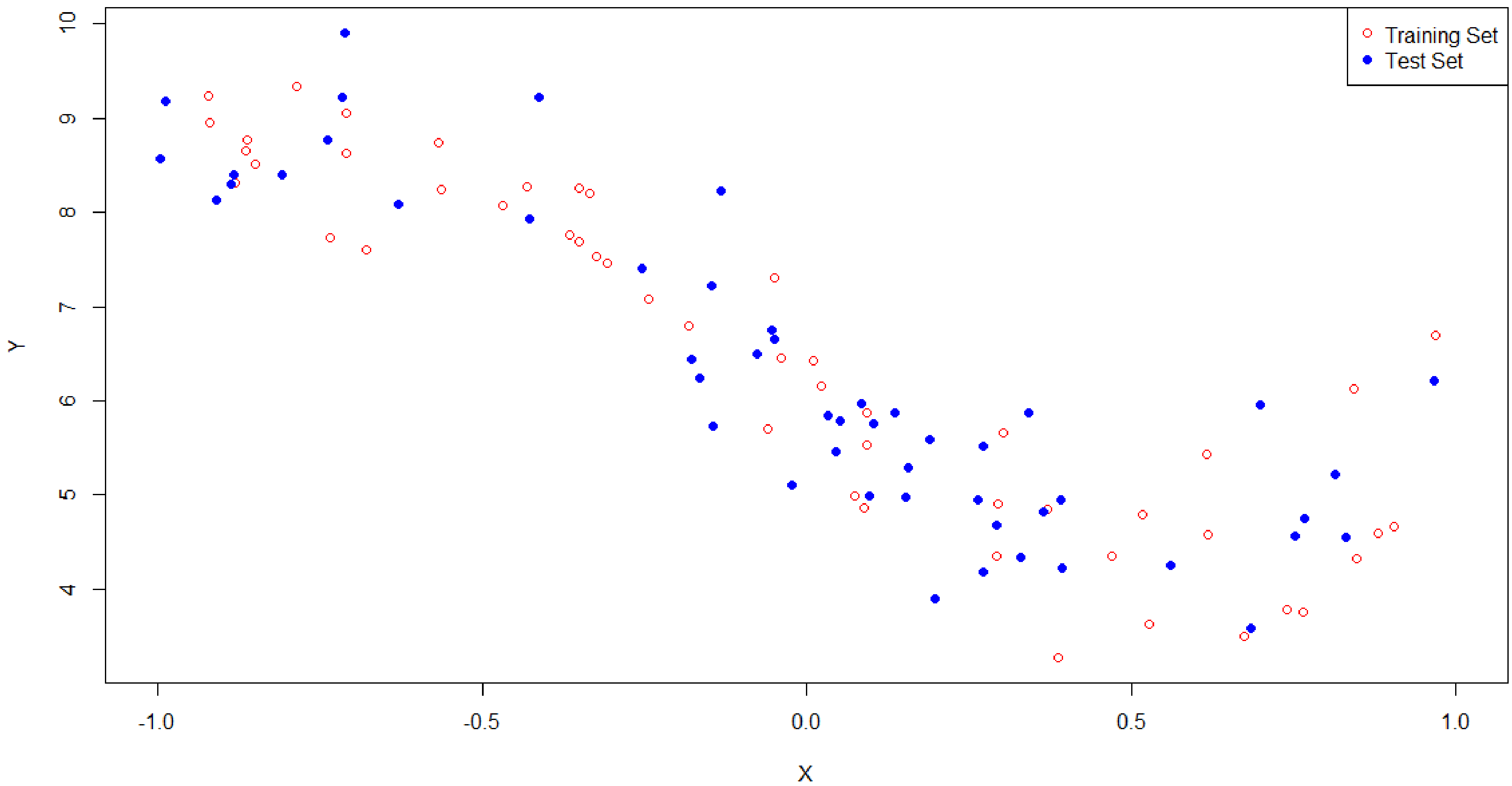
Motivating Example

Bias vs. Variance

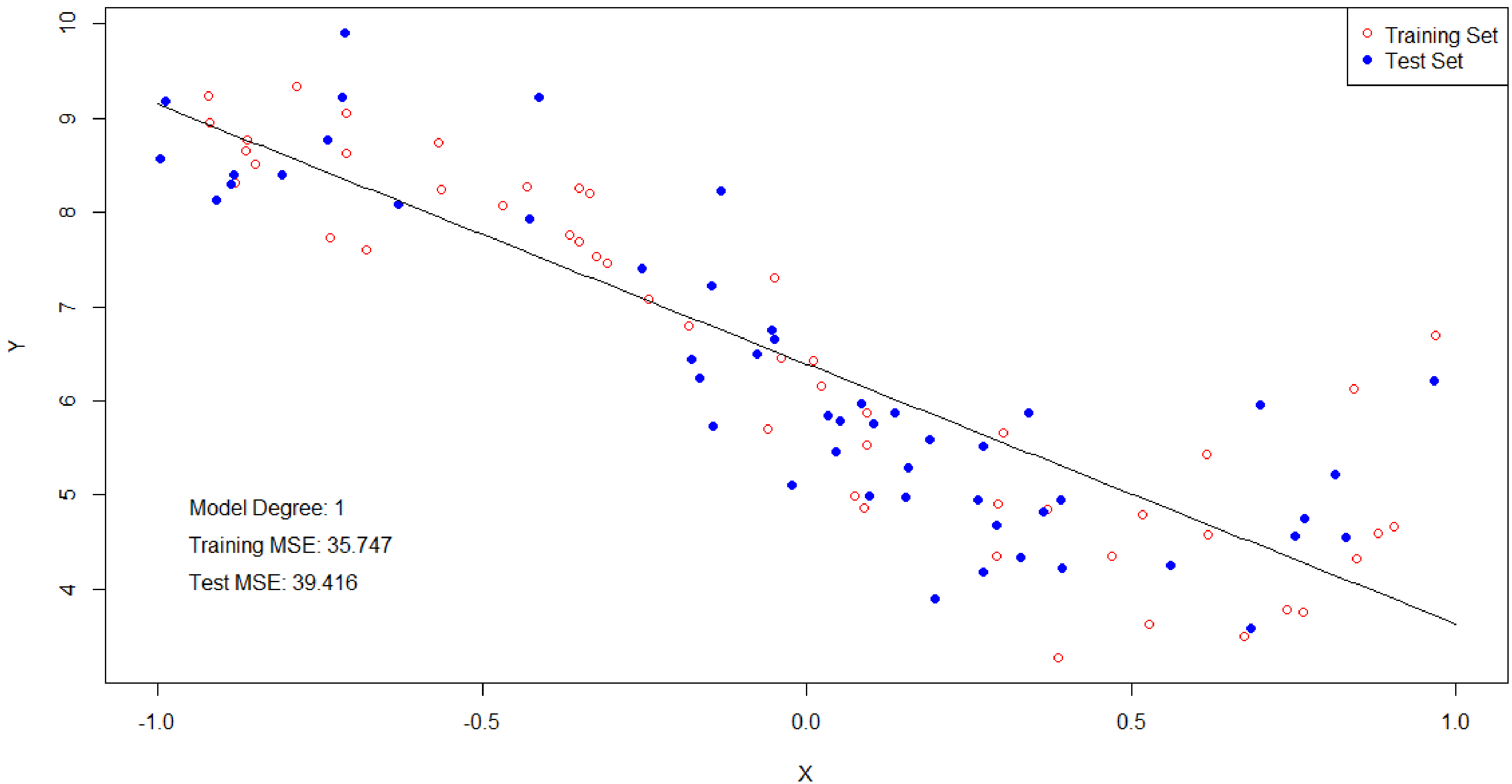
Motivating Example



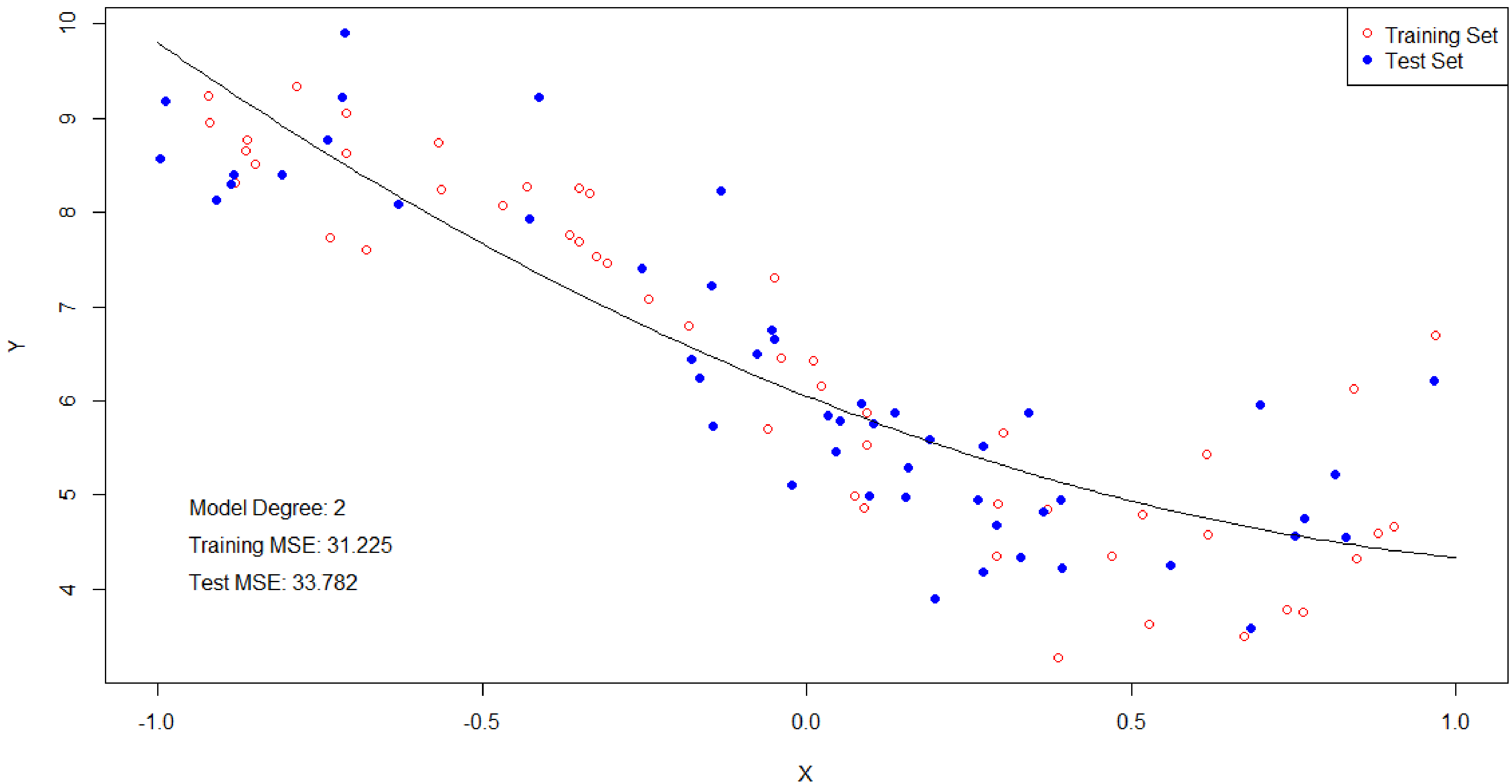
Motivating Example



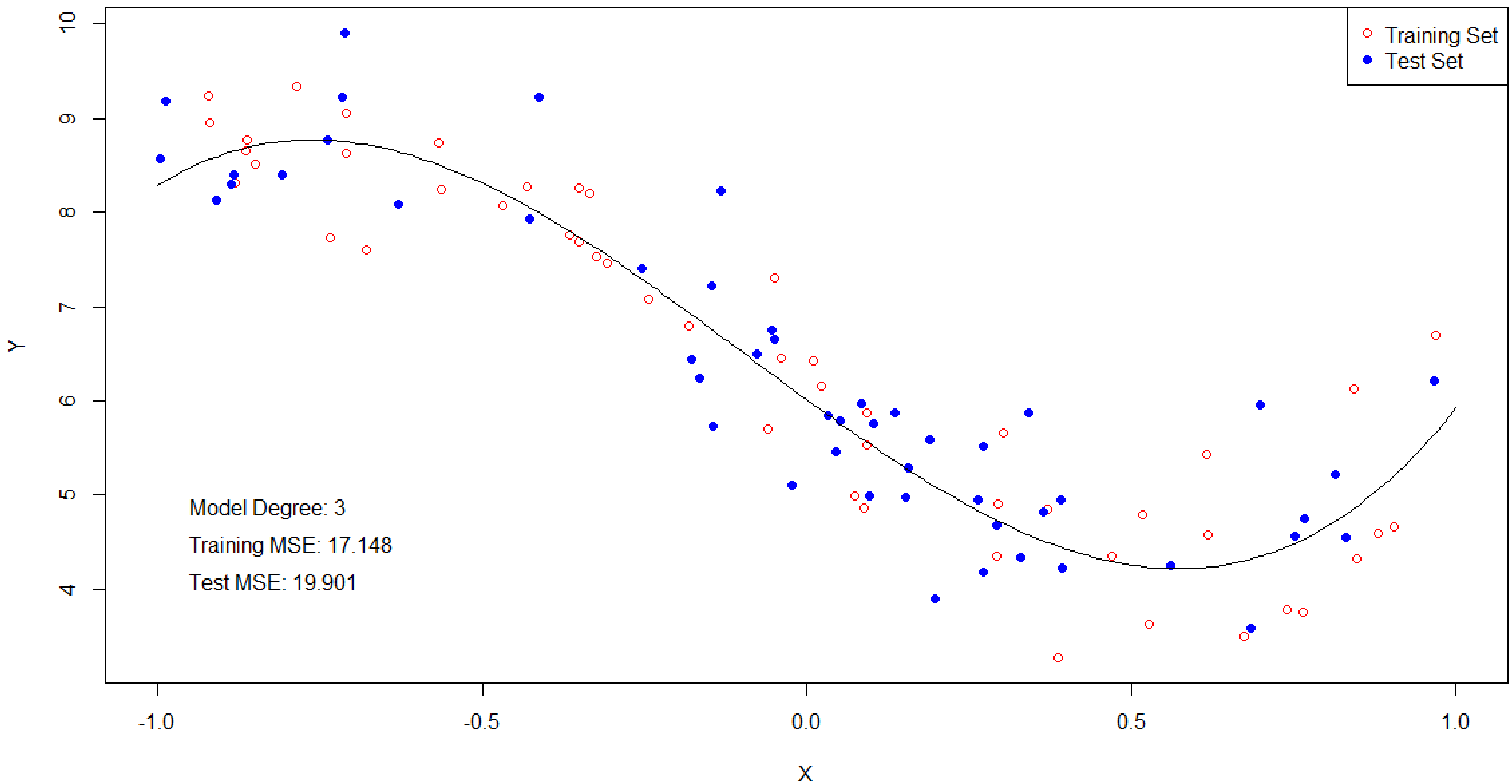
Motivating Example



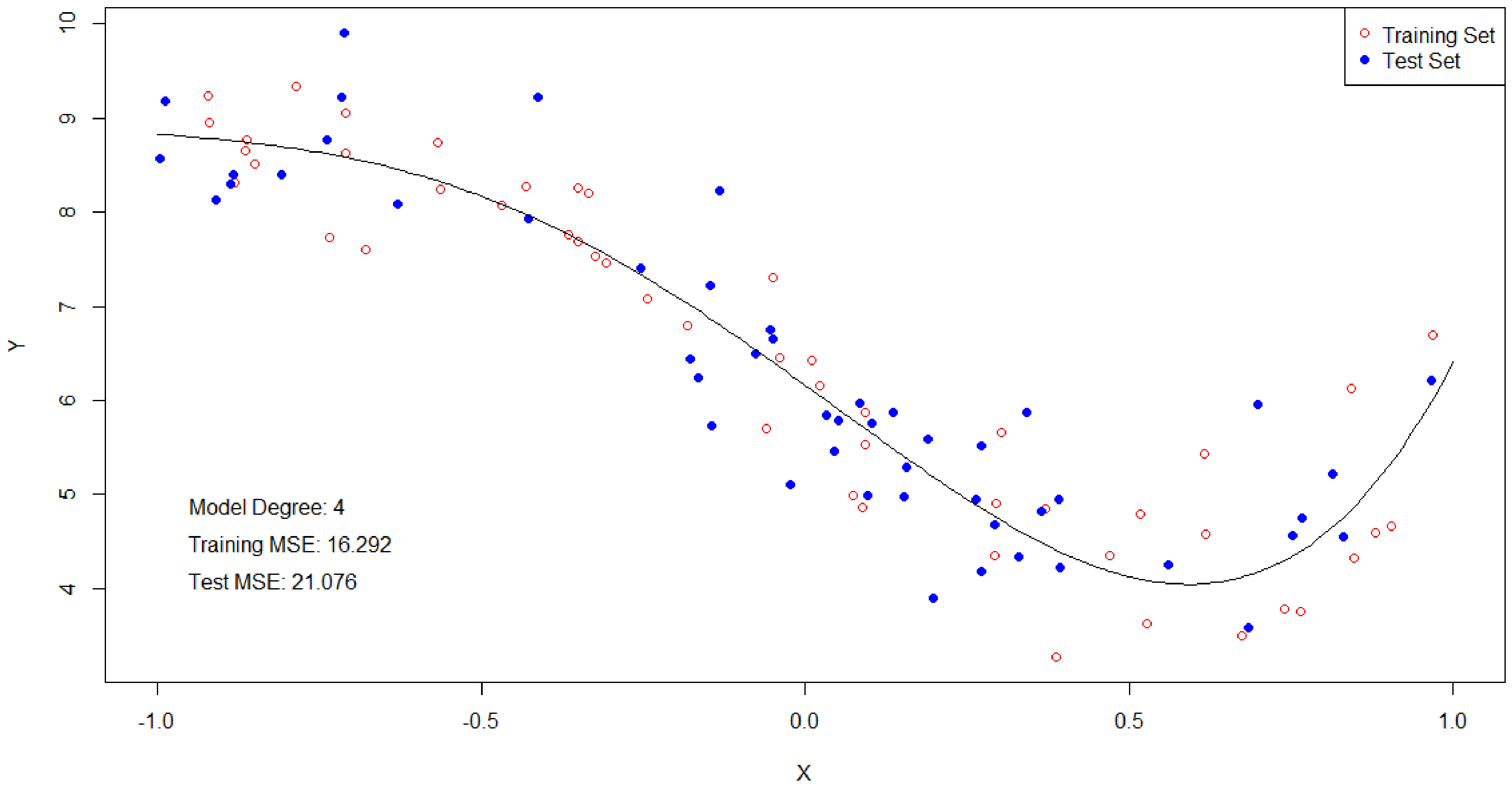
Motivating Example



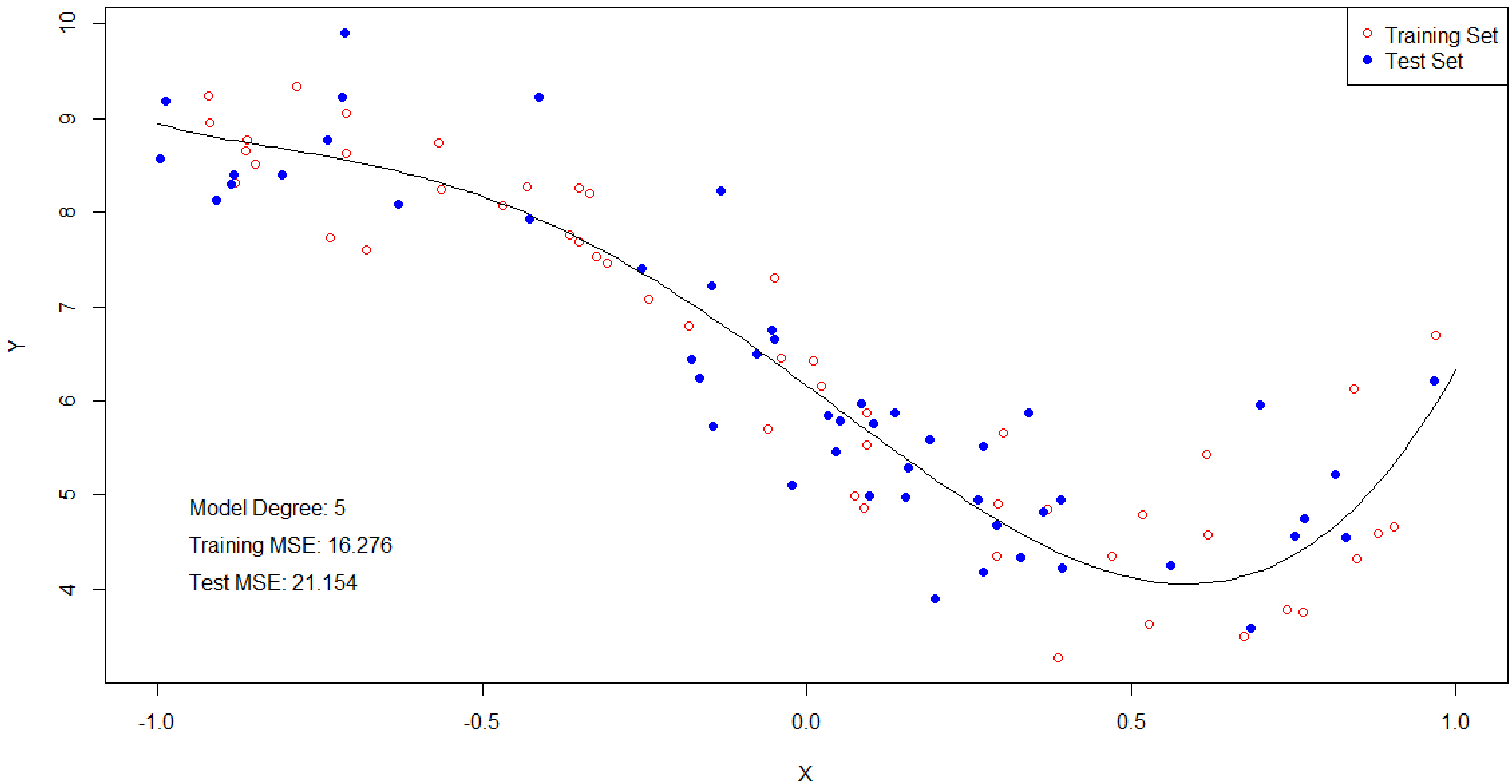
Motivating Example



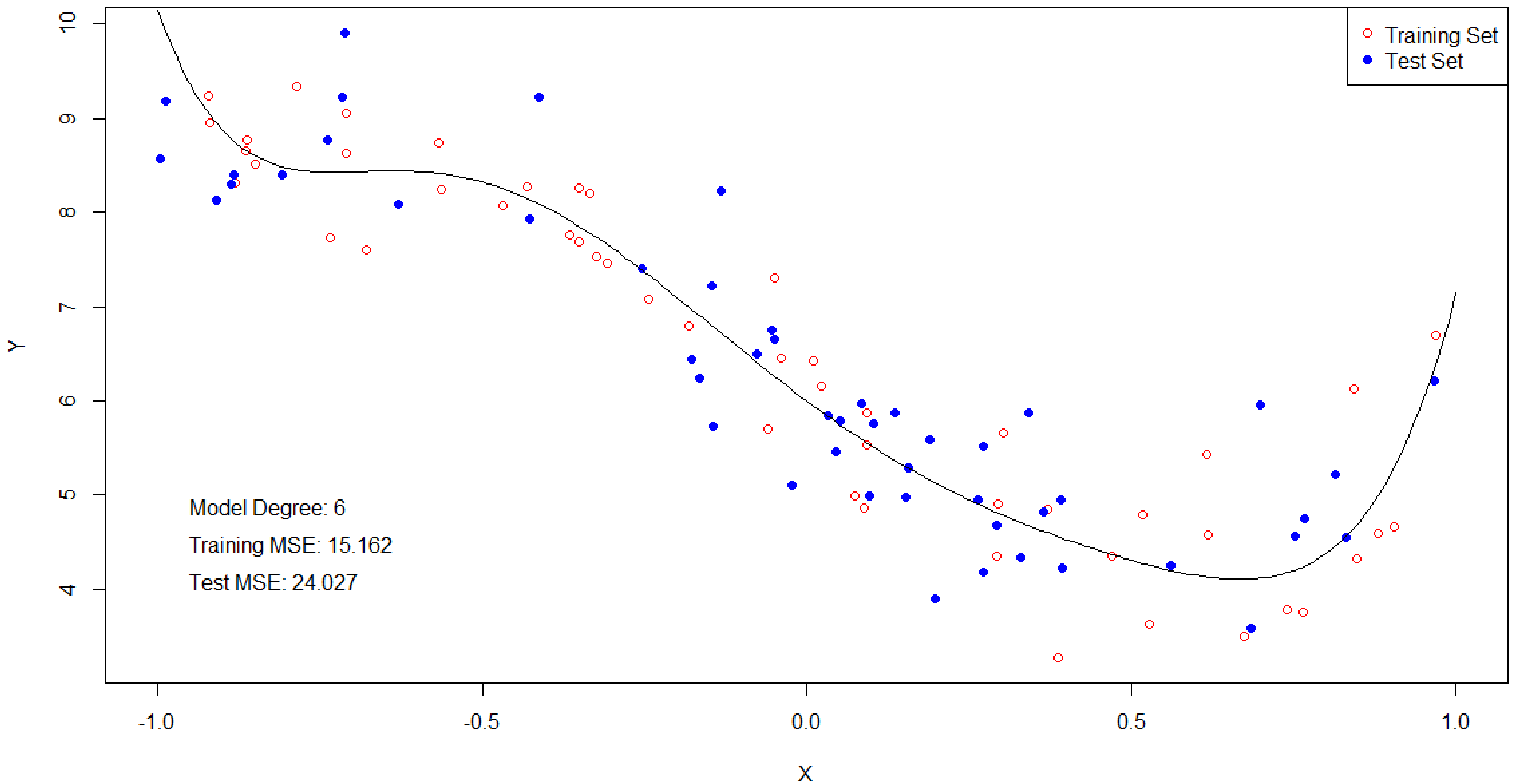
Motivating Example



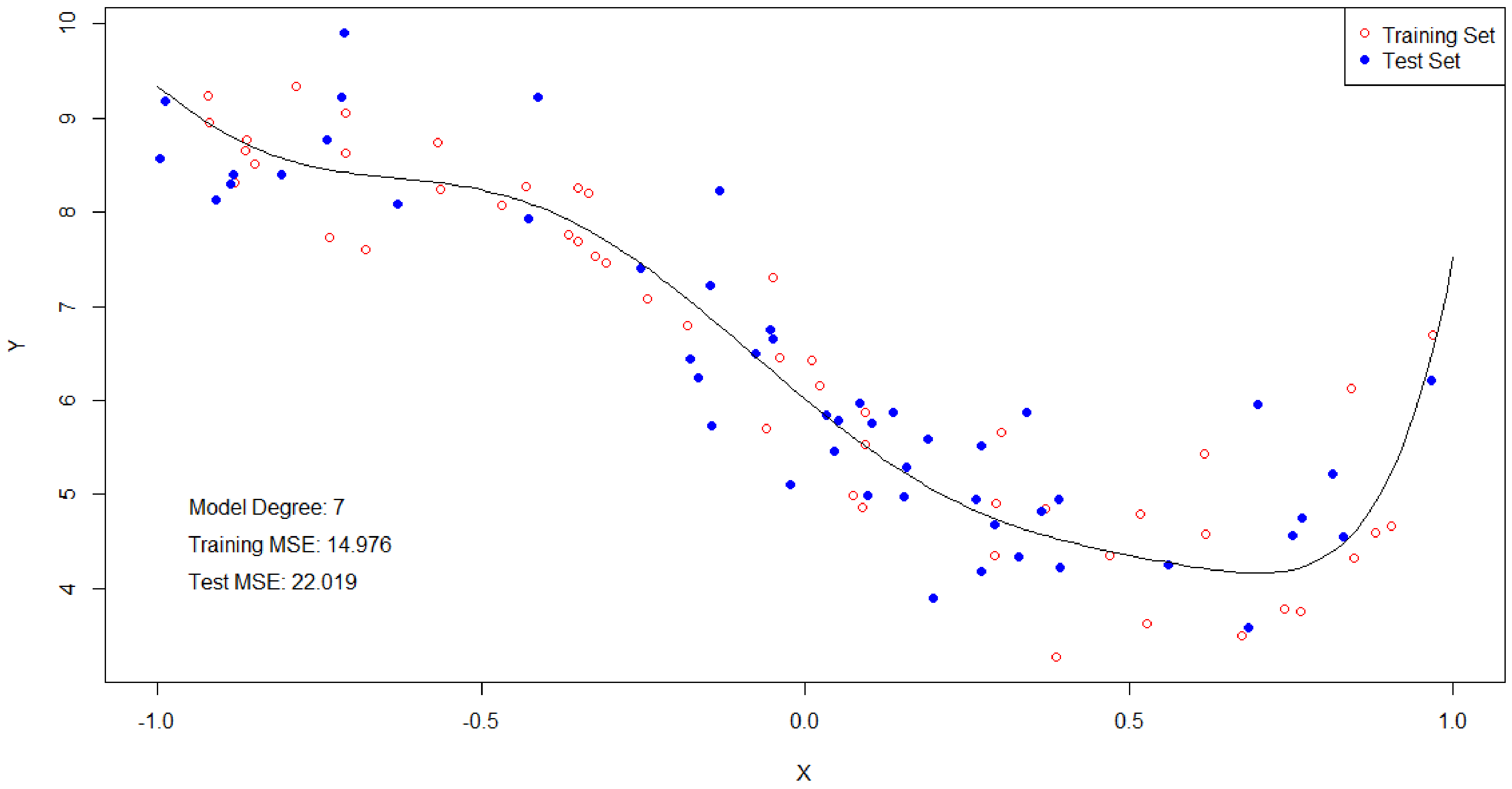
Motivating Example



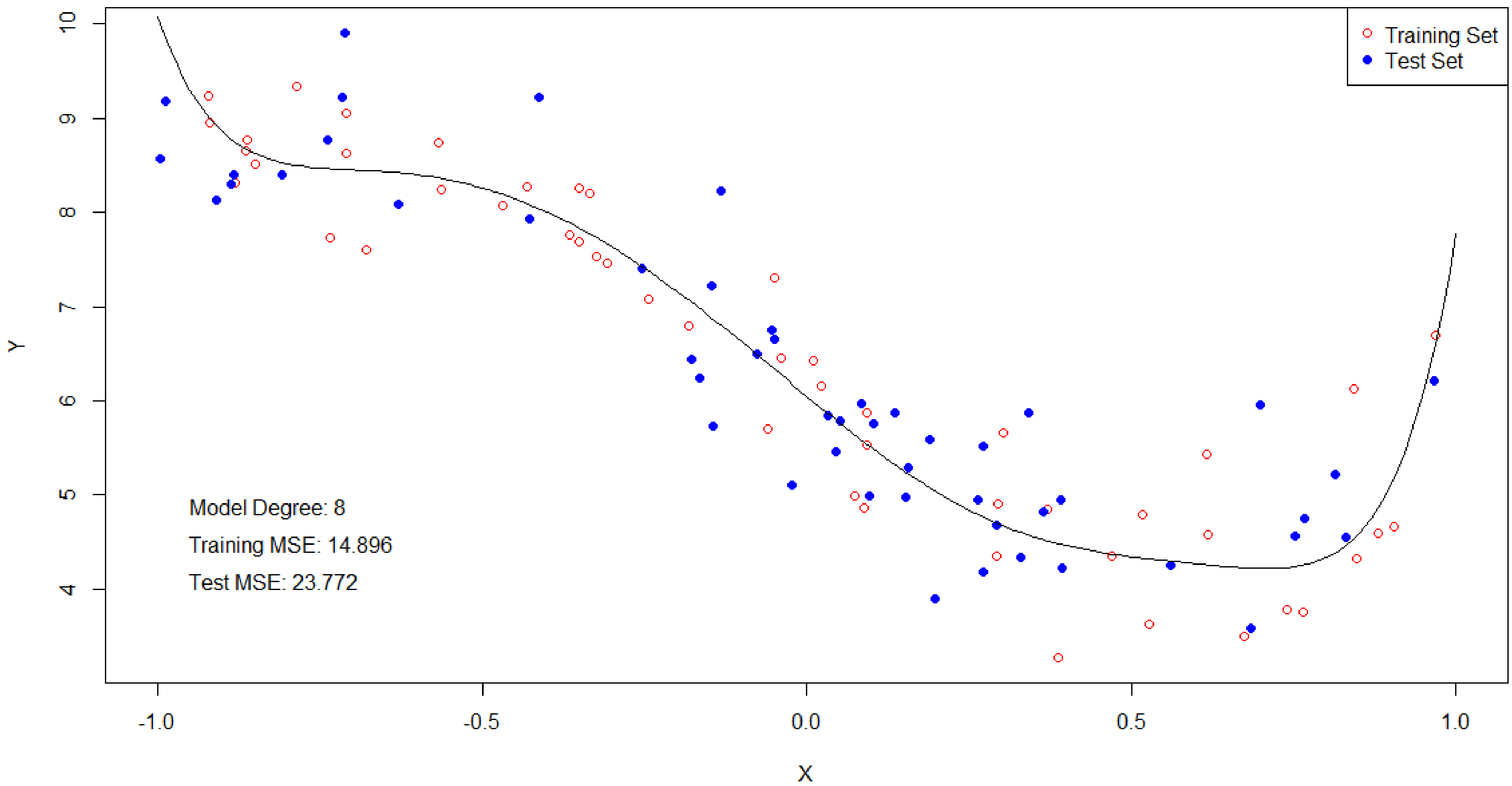
Motivating Example



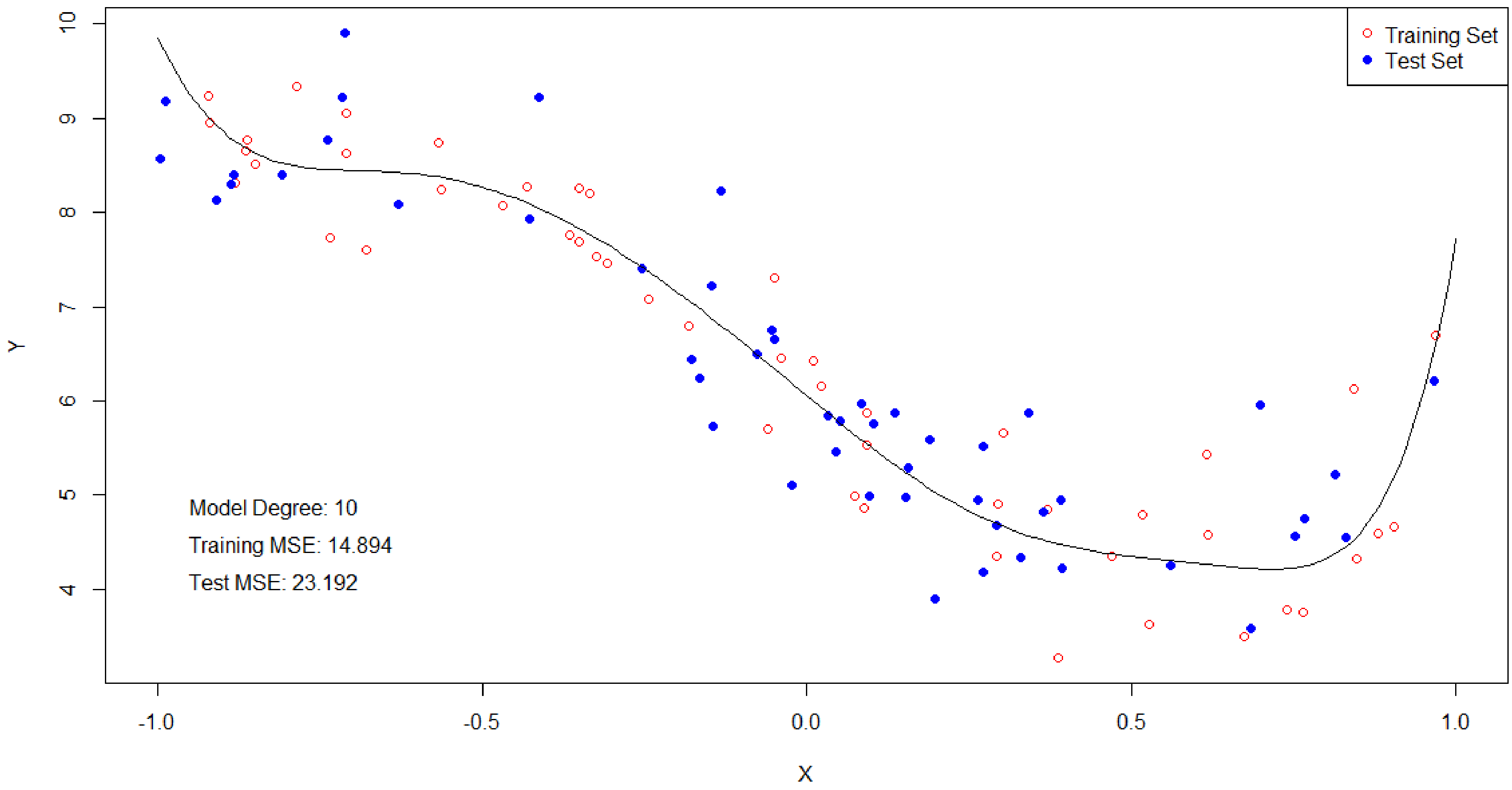
Motivating Example



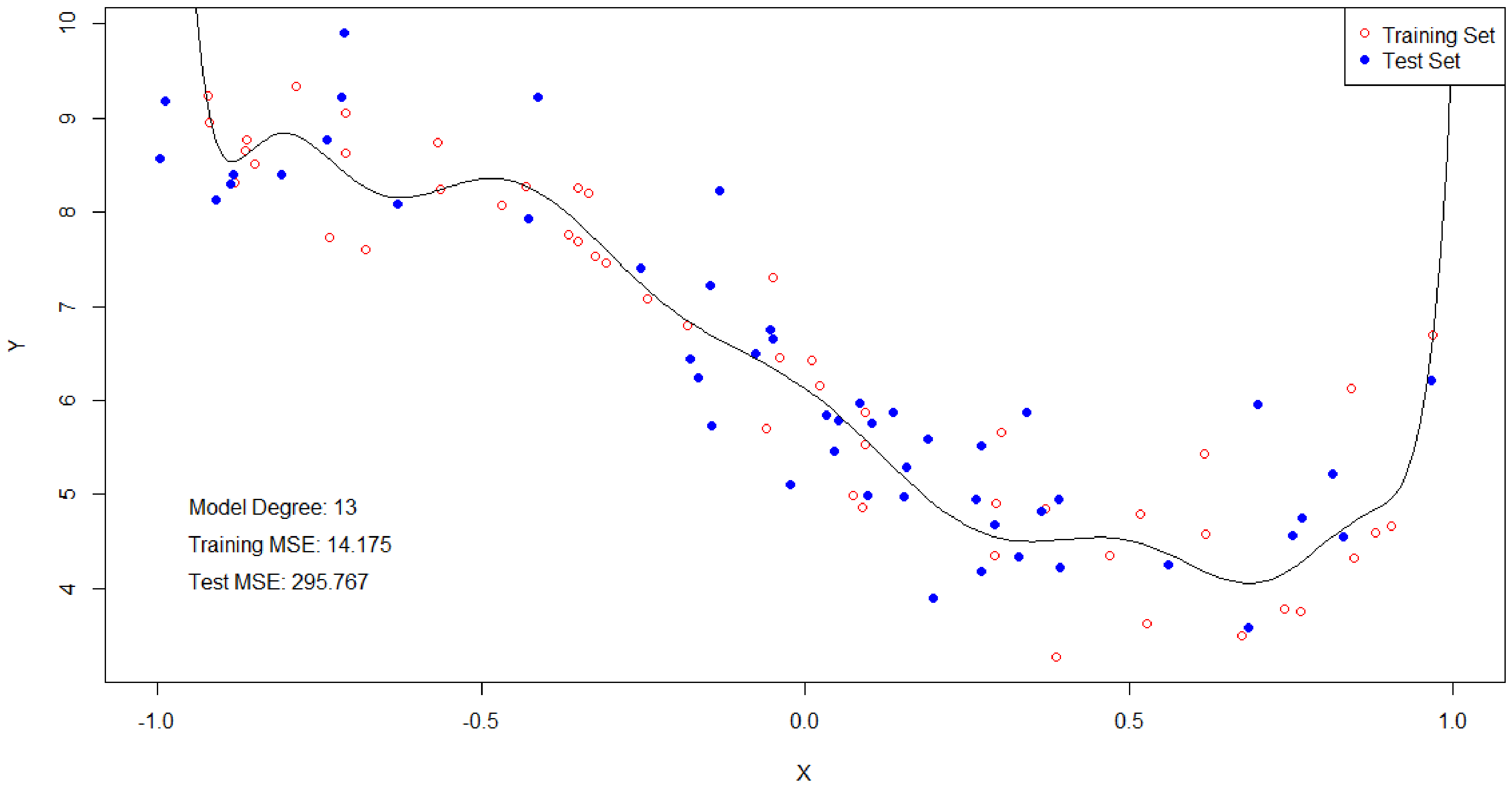
Motivating Example



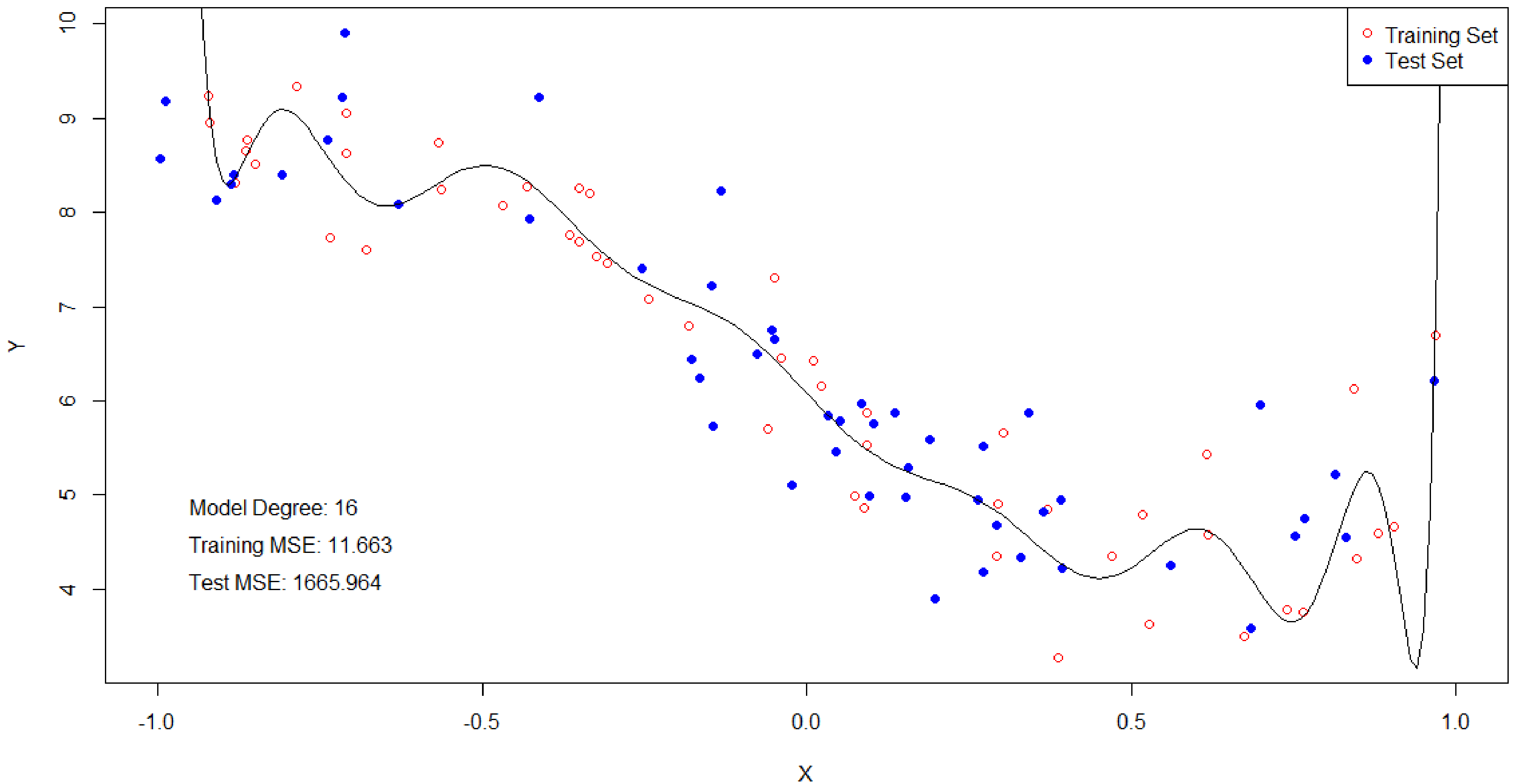
Motivating Example



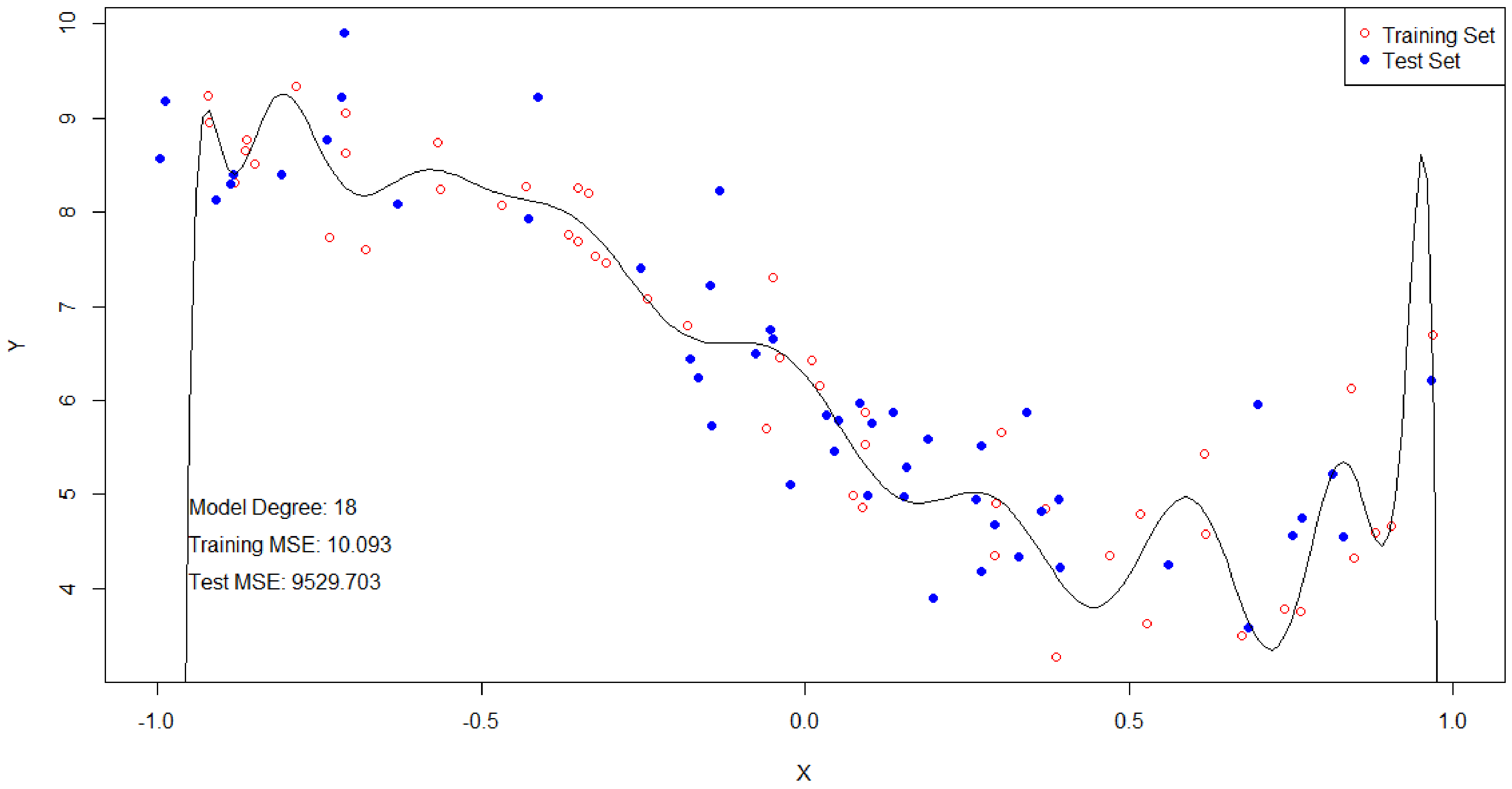
Motivating Example



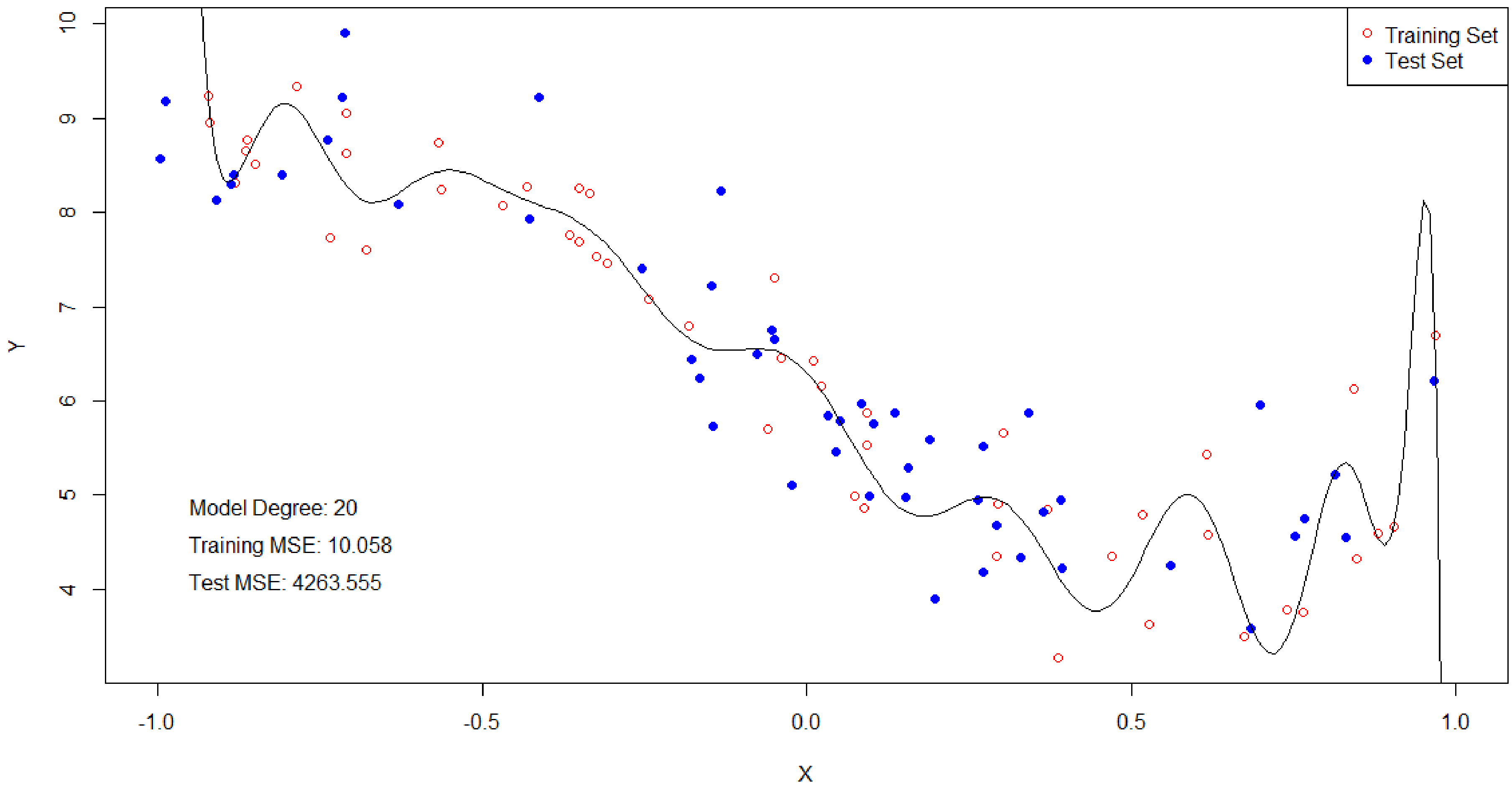
Motivating Example



Motivating Example



Motivating Example



Bias vs. Variance (*Hastie et al.* 2009)

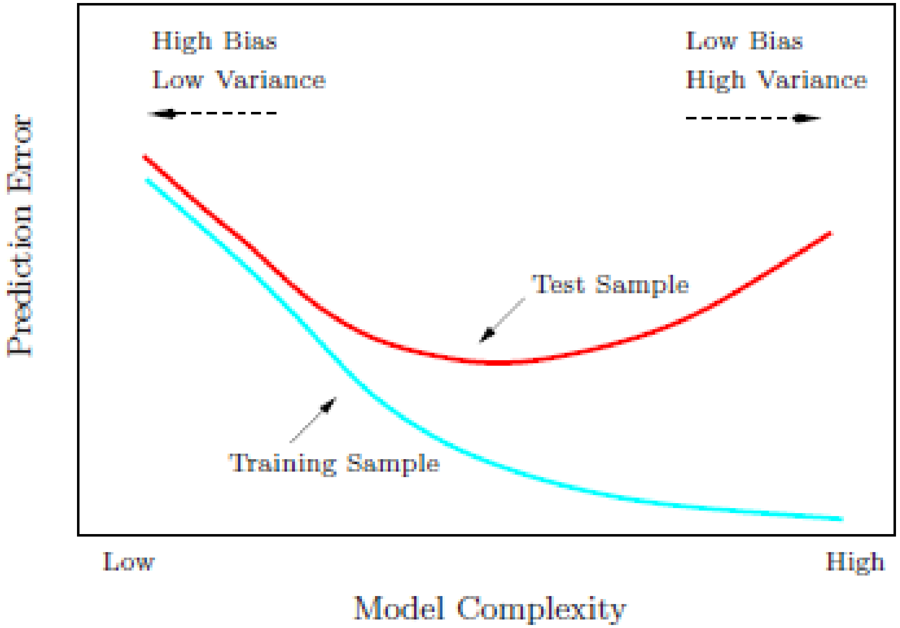
The expected squared prediction error is:

$$E\left[\left(Y-\hat{f}(x)\right)^2\right]$$

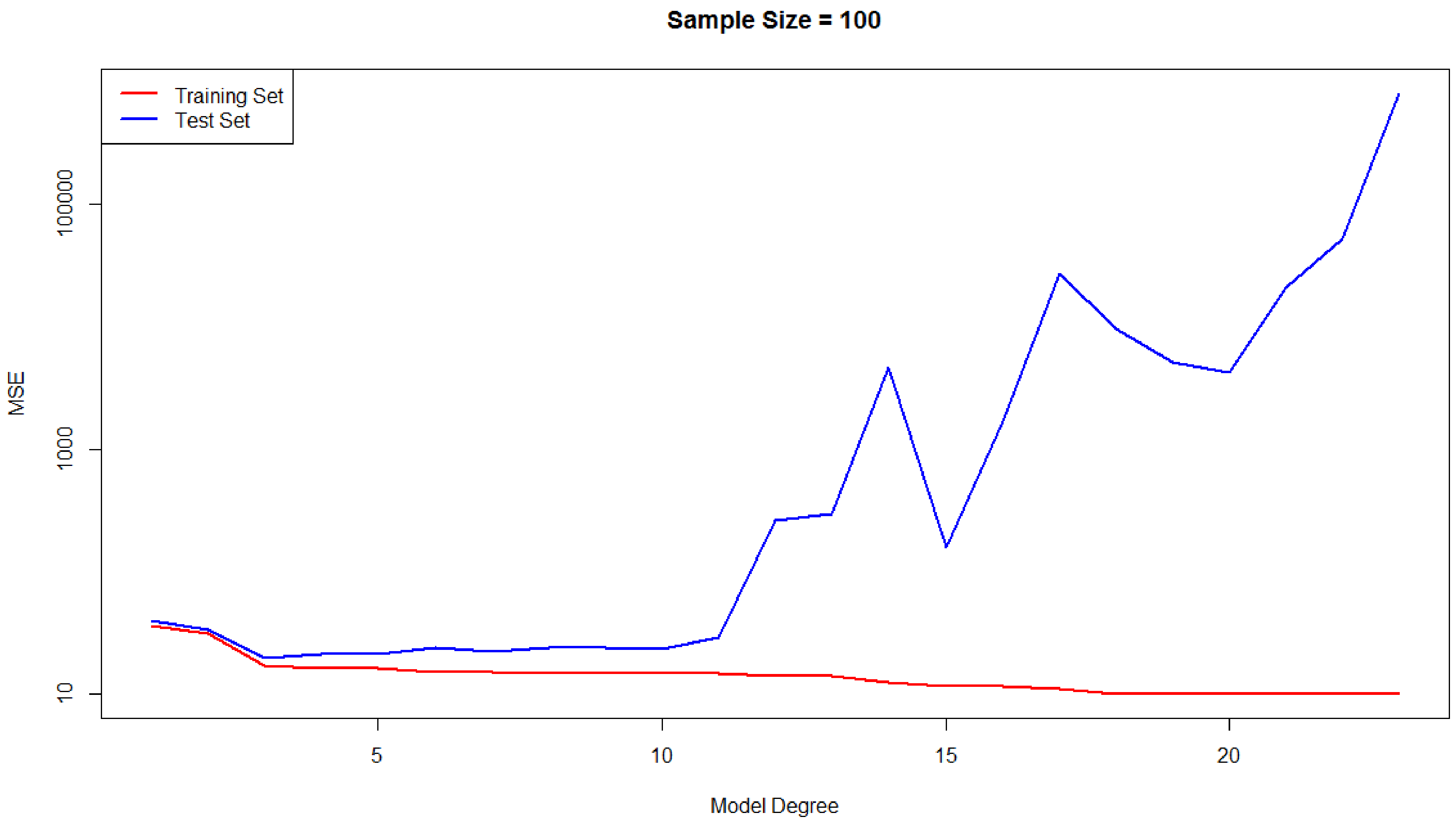
$$\left(E\left[\hat{f}(x)\right]-f(x)\right)^2+E\left[\hat{f}(x)-E\left[\hat{f}(x)\right]\right]^2+\sigma_e^2$$

$$\text{Bias}^2 \quad + \quad \text{Variance} \quad + \quad \text{Irreducible Error}$$

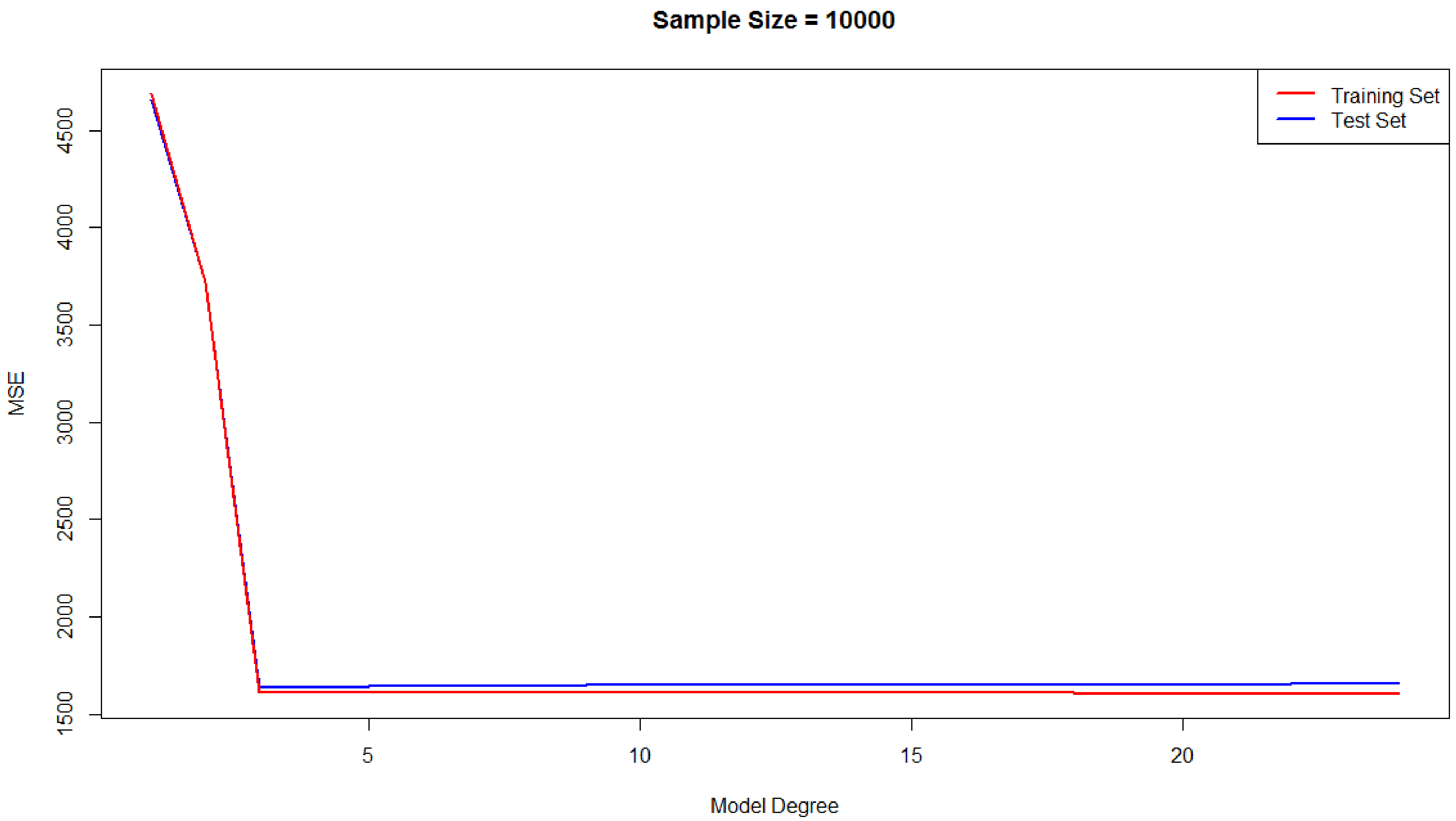
A perfect model and infinite data would reduce the first two terms to zero, but with finite data and imperfect models, we will need to choose between minimizing bias and minimizing variance.



Bias vs. Variance in our Example



Increasing the Sample Size



Predictive Analytics

John Murdzek, FSA
Senior Experience Studies Actuary
Genworth Financial

March 20, 2018

Disclaimer and Data

- Disclaimer: The views and opinions expressed in this presentation are those of the presenter and do not reflect the official policy or position of Genworth Financial, Inc. or any of its subsidiaries.

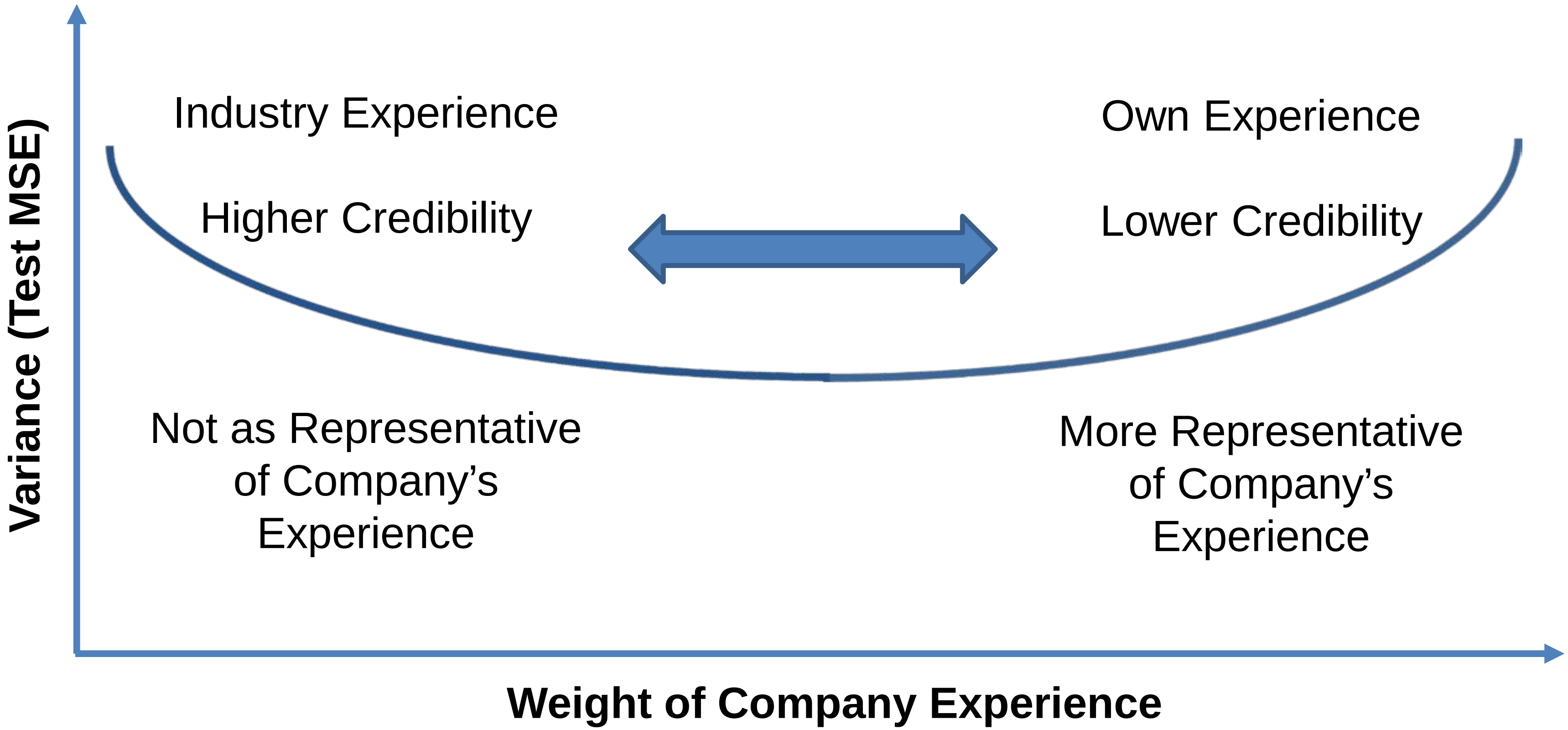
- Data: All data used to support this presentation (including "Own") comes from the SOA LTC Claim Termination Rates Database 2000-2011

Overview

- Goal: Observe the traversing of the BVT using hold-out data from publicly available LTCI CTR data

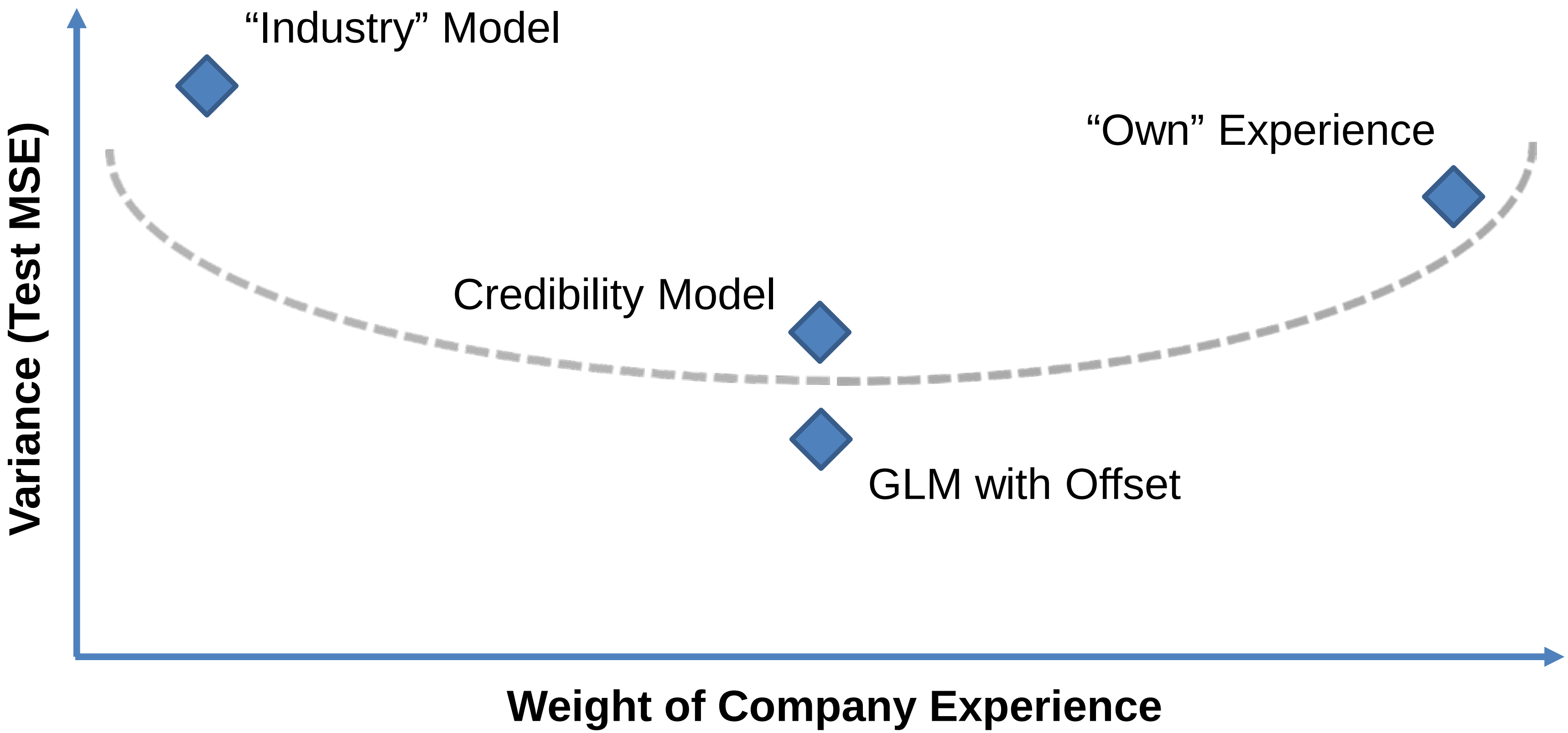
- Classically fit industry and own experience
- A/E and Limited fluctuation credibility views
- GLM views
- Validation using hold-out data
- Test MSE and BVT concepts

Traversing the BVT



- Credibility approaches help address the trade-off of Bias and Variance between Industry and Company experience, and so can Predictive Analytical techniques

BVT Overview on Hold-Out Data



- Credibility and GLM approaches produce a reasonable mid-range result

Note: Graphs within the presentation are approximate

Basic CTR Modeling Approach

- Classical fit to industry experience
  - Base table varying by gender and claim type
  - Relativities for claim duration and age
- Data
  - From 2015 SOA LTC Experience CTRs
  - “Industry” Data = 12,449 terminations
  - “Own” Data = 2,646 terminations (not GNW)
    - Females = 80% of “Industry” CTRs
    - Males = 130% of “Industry” CTRs
  - Hold-Out Data = 1,073 terminations

Classical Fit Process

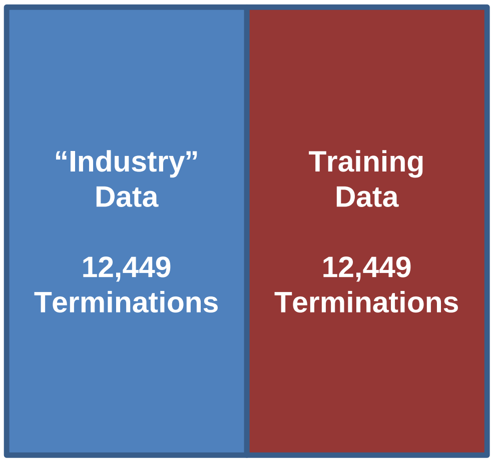
	Data:		Modeling:	Testing:
I:	"Industry" Data 12,449 Terminations	Training Data 12,449 Terminations	Base Table + Two Relativities Use Entire Sample to Train	Apply Model to Hold-Out "Own" Data Observe Test MSE
II:	"Own" Data 2,646 Terminations	Training Data 1,573 Terminations Hold-Out Data 1,073 Terminations	CTRs = Raw Termination Data Use Training Sub-Sample	Apply Model to Hold-Out "Own" Data Observe Test MSE

Note: All data comes from the SOA LTC Claim Termination Rates Database 2000-2011

Classical Fit Process

Data: | **Modeling:** | **Testing:**

I:



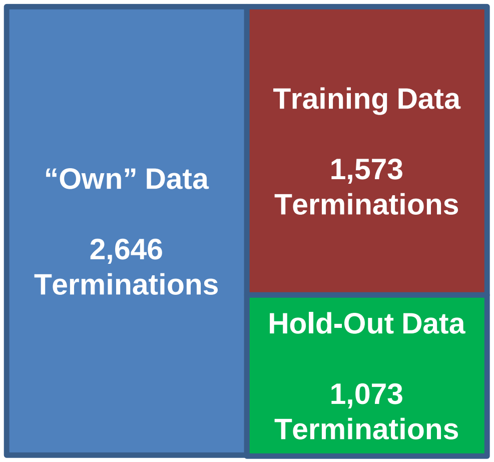
Base Table + Two Relativities

Use Entire Sample to Train

Apply Model to Hold-Out “Own” Data

Observe Test MSE

II:



CTRs = Raw Termination Data

Use Training Sub-Sample

Apply Model to Hold-Out “Own” Data

Observe Test MSE

Note: All data comes from the SOA LTC Claim Termination Rates Database 2000-2011

Classical Fit to “Industry” Training Data

Base Table:

Gender	Situs	CTR	Termination Count
Female	ALF	1.6%	1,663
Female	HHC	2.6%	3,488
Female	NH	2.8%	2,452
Male	ALF	2.6%	898
Male	HHC	3.3%	2,240
Male	NH	3.8%	1,708

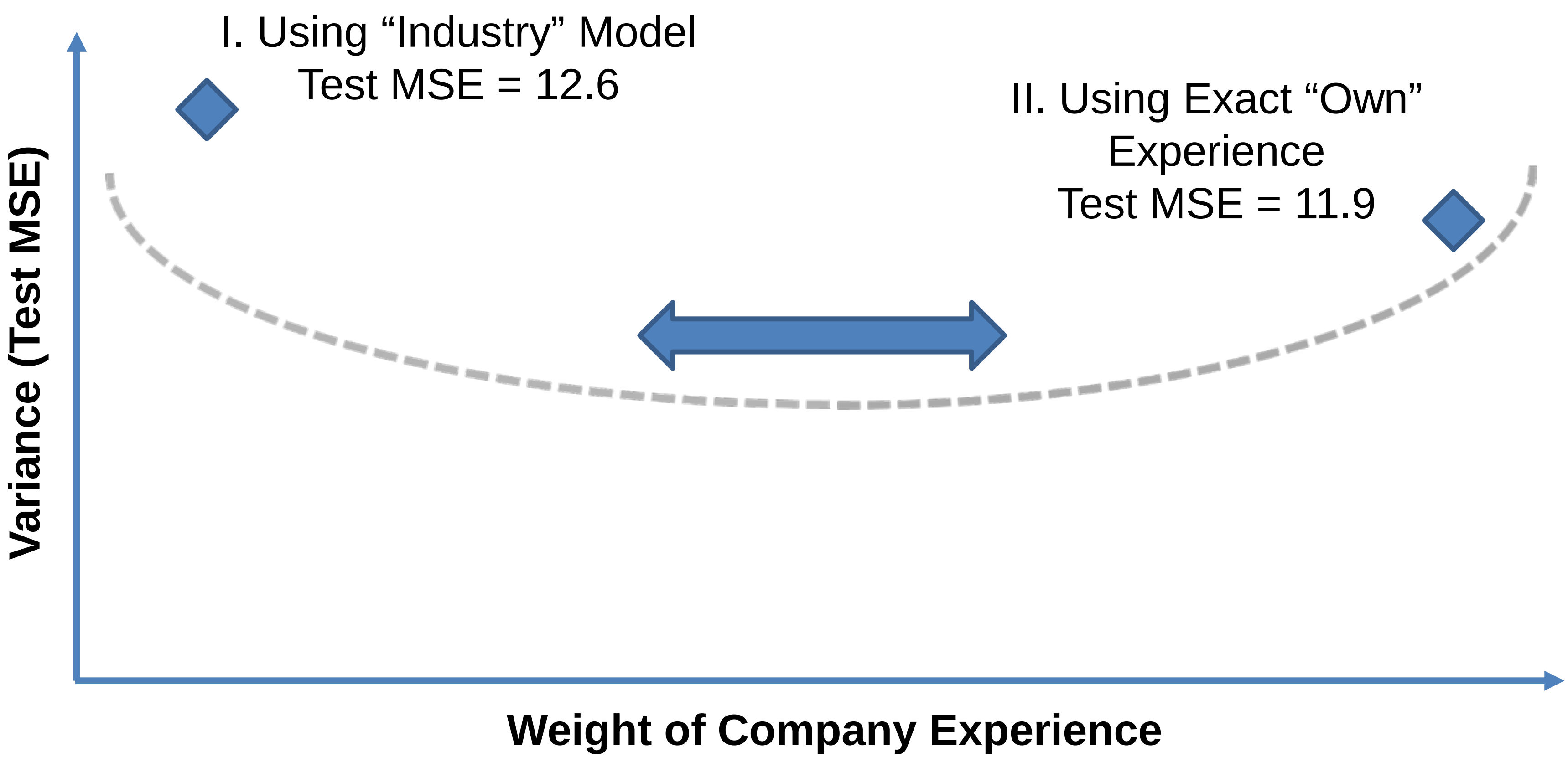
Factor Tables:

Claim Duration	Situs	CTR Factor	Termination Count
Mos 1-3	ALF	174%	103
Mos 4-12	ALF	81%	604
Yrs 2+	ALF	105%	1,854
Mos 1-3	HHC / NH	361%	1,197
Mos 4-12	HHC / NH	128%	3,821
Yrs 2+	HHC / NH	74%	4,870

Claim Age	Gender	Situs	CTR Factor	Termination Count
18 to 74	All	ALF	88%	468
75 to 84	All	ALF	97%	1,410
85+	All	ALF	117%	683
18 to 74	Female	HHC / NH	119%	1,952
75 to 84	Female	HHC / NH	90%	2,784
85+	Female	HHC / NH	99%	1,204
18 to 74	Male	HHC / NH	98%	1,279
75 to 84	Male	HHC / NH	98%	1,913
85+	Male	HHC / NH	111%	756

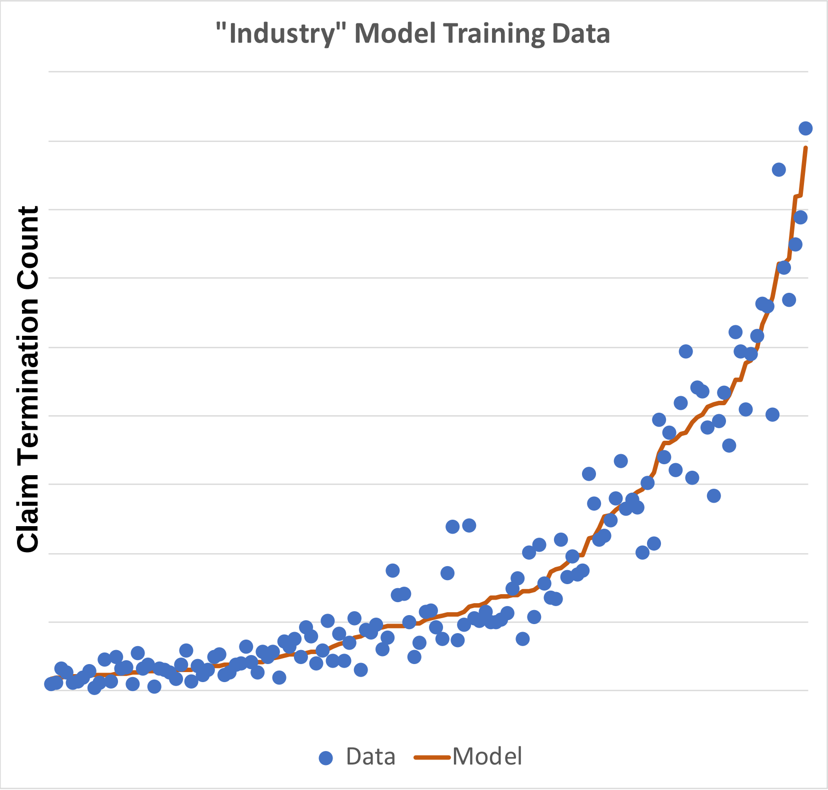
- “Industry” CTR = Base Table x
  Duration-Situs Factor x
  Age-Gender-Situs Factor

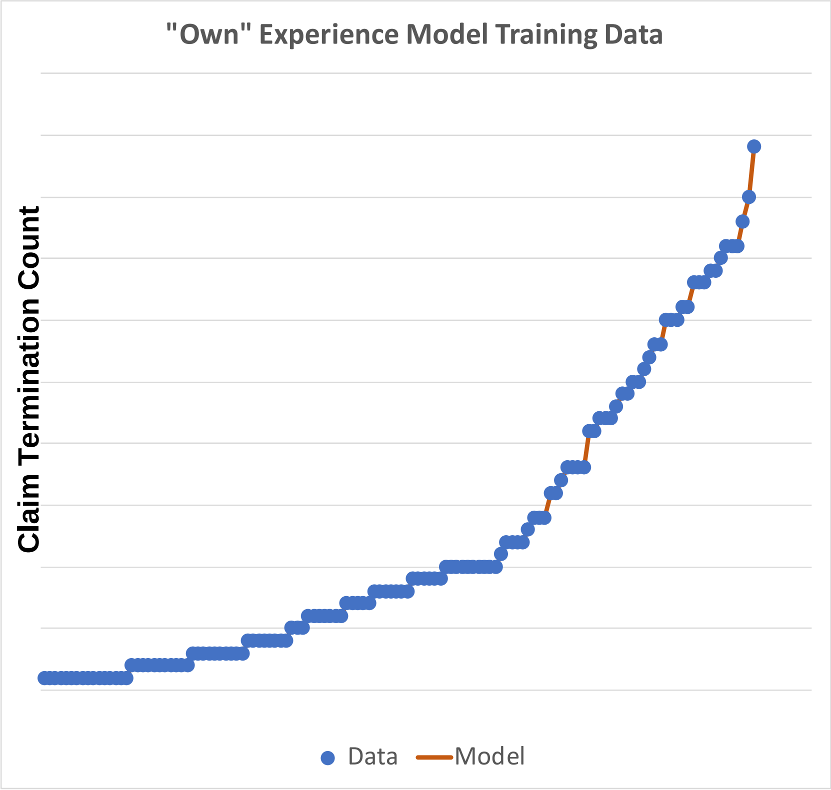
Classical BVT Using Hold-Out Data



- Test MSE = $\frac{1}{N}\sum_{\substack{Hold-Out \\ Data}}(Term.Count_{Actual} - Term.Count_{Model})^2$

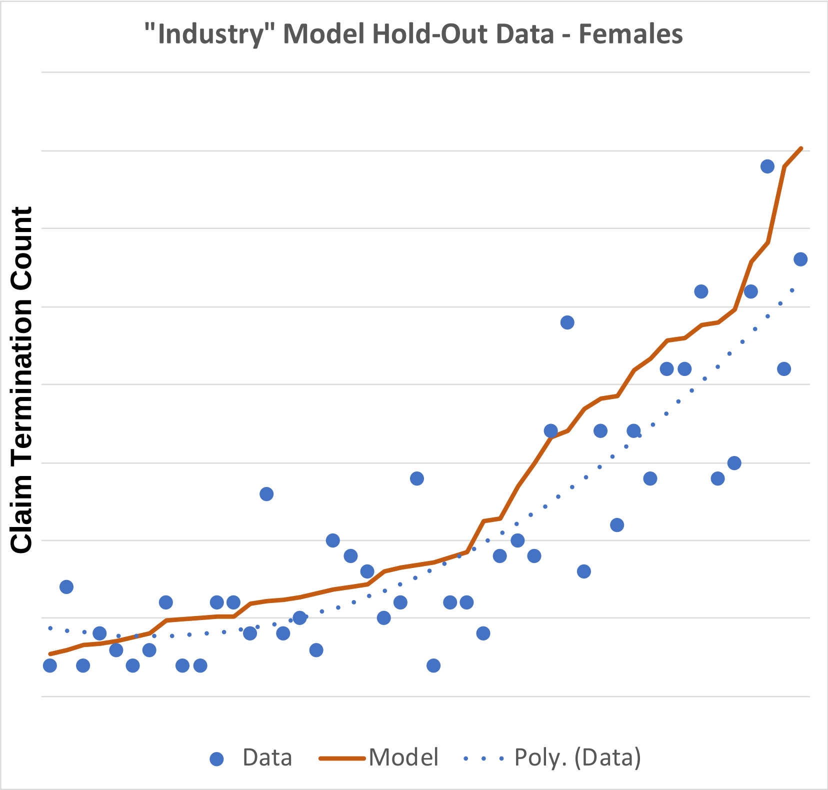
Why the Classical Fits Can Be Improved

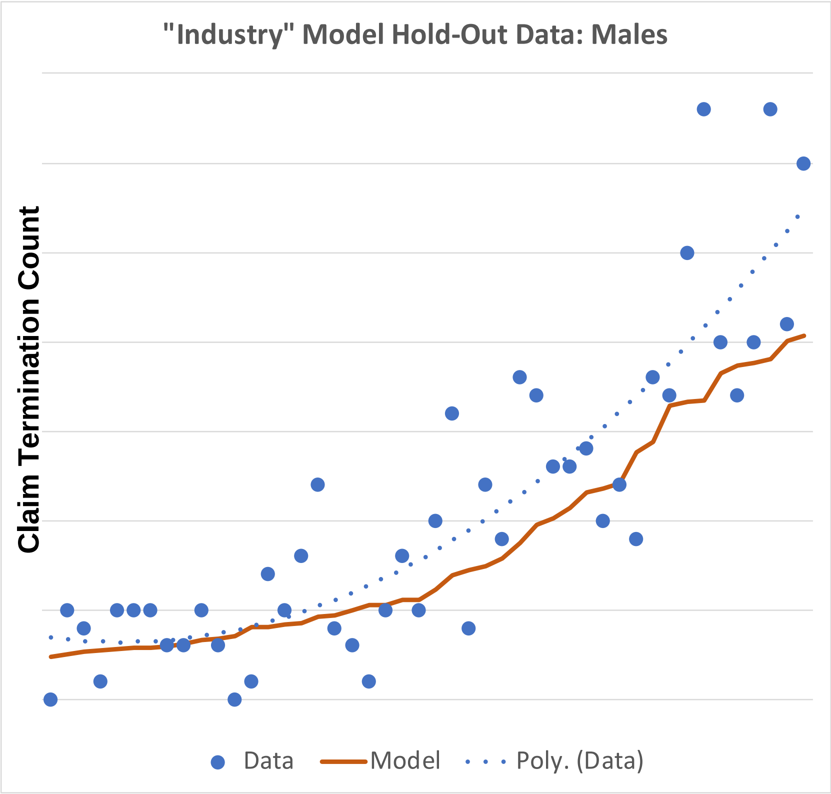




- “Industry” model looks good, except that we know it will be biased vs. the “Own” Experience
- “Own” Experience model is purposely over-fit

Hold-Out Results: “Industry” Model





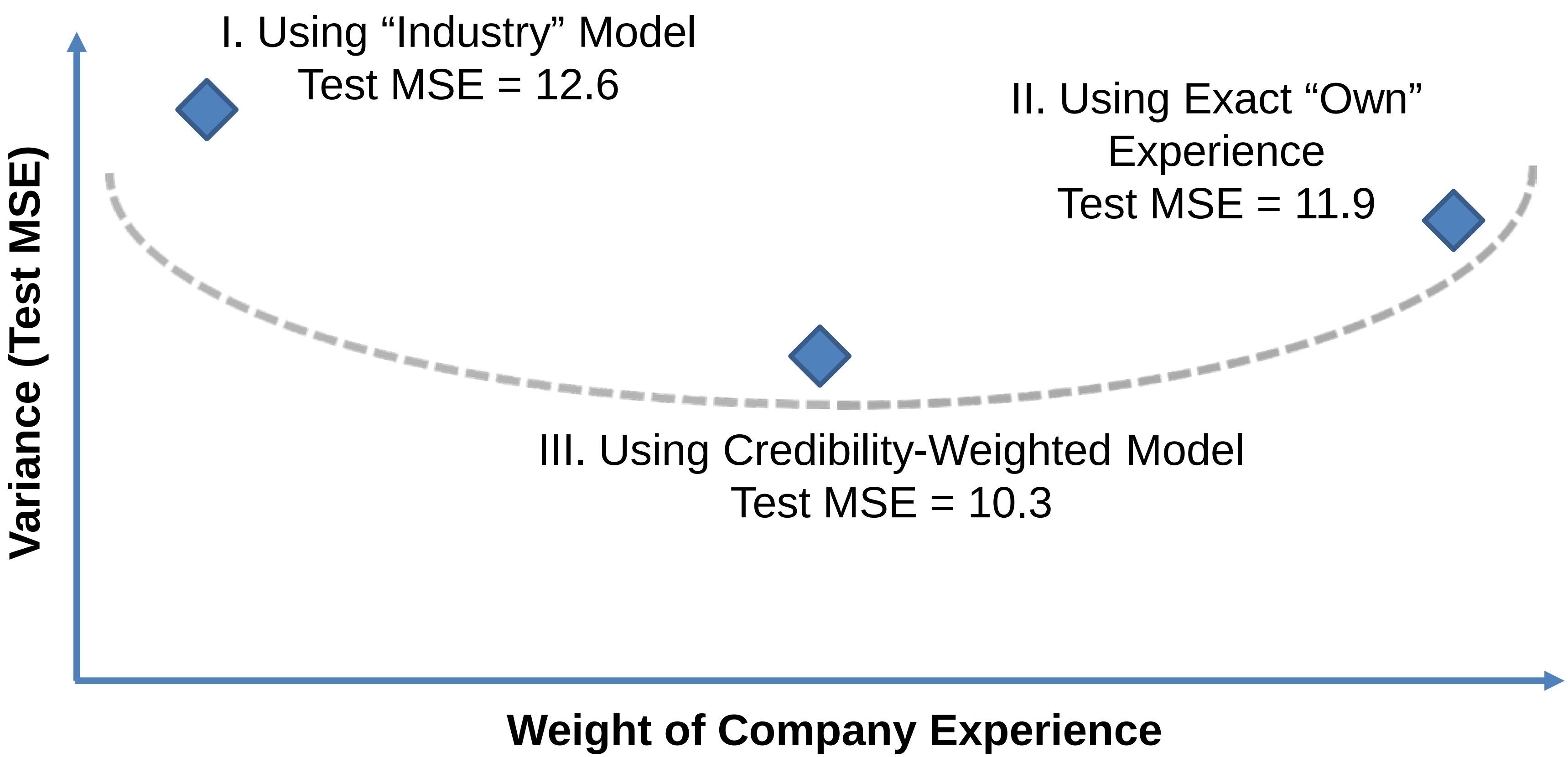
- The “Industry” model bias can be seen in the Hold-Out data when viewed by gender, and can be addressed using Credibility or GLM approach

Credibility View

- Credibility-Weighting Procedure (III.)
  - Limited fluctuation credibility
  - Credibility-Weighted CTR =
    Z x "Own" + (1 – Z) x "Industry"
  - "Own" CTR is the raw experience CTR
  - "Industry" is the model

- 1,082 terminations = full credibility
  - $Z = \text{Min}[1, (N / 1{,}082)^{1/2}]$
  - N = number of terminations

BVT Using Hold-Out Data with Cred View



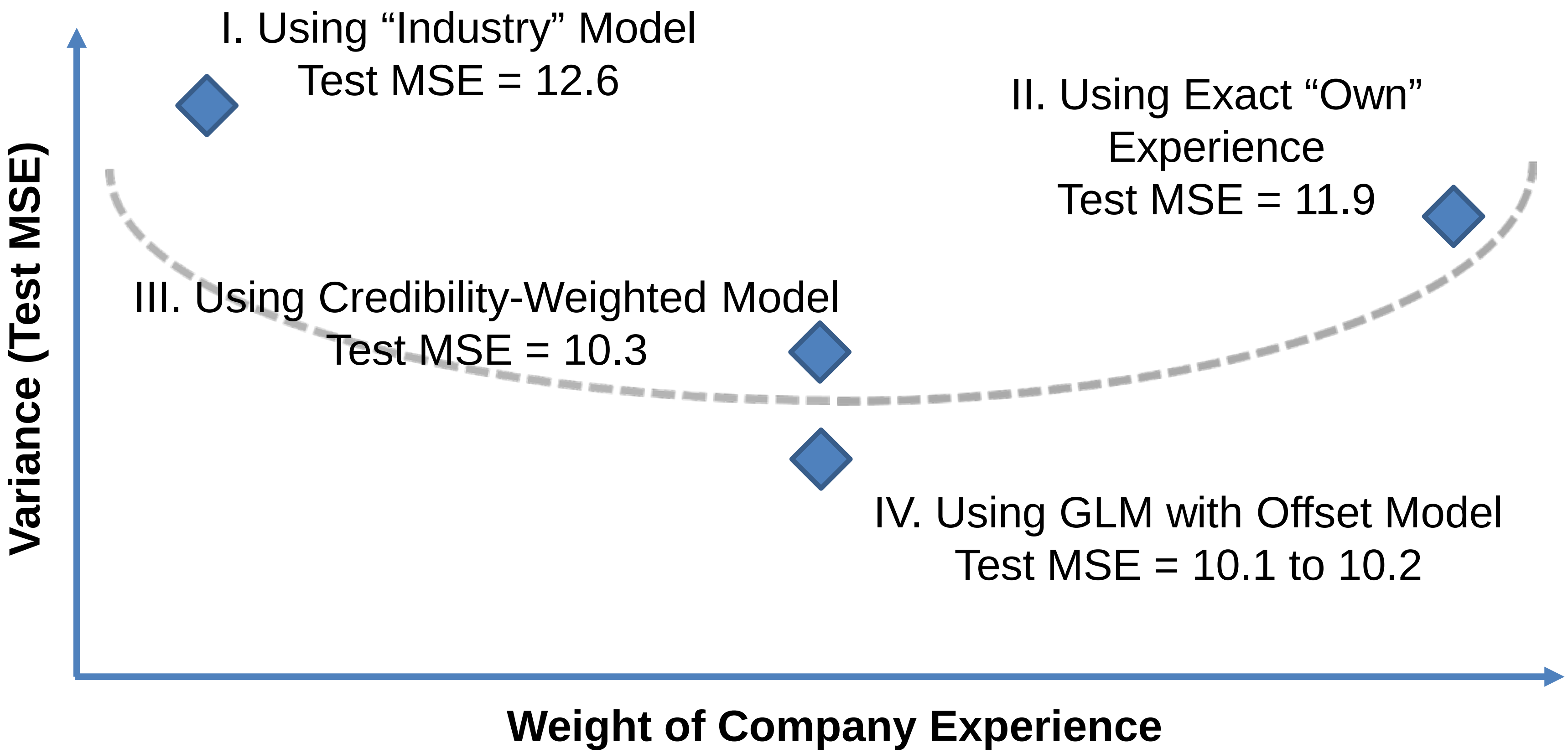
- Credibility weighting borrows the better aspects of each model, and performs better out-of-sample than either model alone

GLM View

- Generalized Linear Model (IV.)
  - Normal distribution with a log link
  - Trained on Own Experience data (the 1,573)
  - Offset to the "Industry" model

- Equation
  - $Terminations_{Own} = Expected_{Industry} \times \; exp(\sum_i \beta_i x_i)$
  - Same factors as "Industry" model with slightly less granularity by claim duration

BVT Using Hold-Out Data with GLM View



- The GLM slightly improves upon the limited fluctuation credibility result

Concluding Remarks

- GLM vs. Credibility approach
  - Less tied to subjective assumptions
  - Results appear to be as good or slightly better
  - Stepping stone to more sophisticated approaches (e.g., Penalized GLM)
- Hold-Out Data approach
  - Relatively simplistic way to reduce the chance of over-fitting or under-fitting
  - Another stepping stone to more sophisticated approaches (e.g., k-fold Cross-Validation)

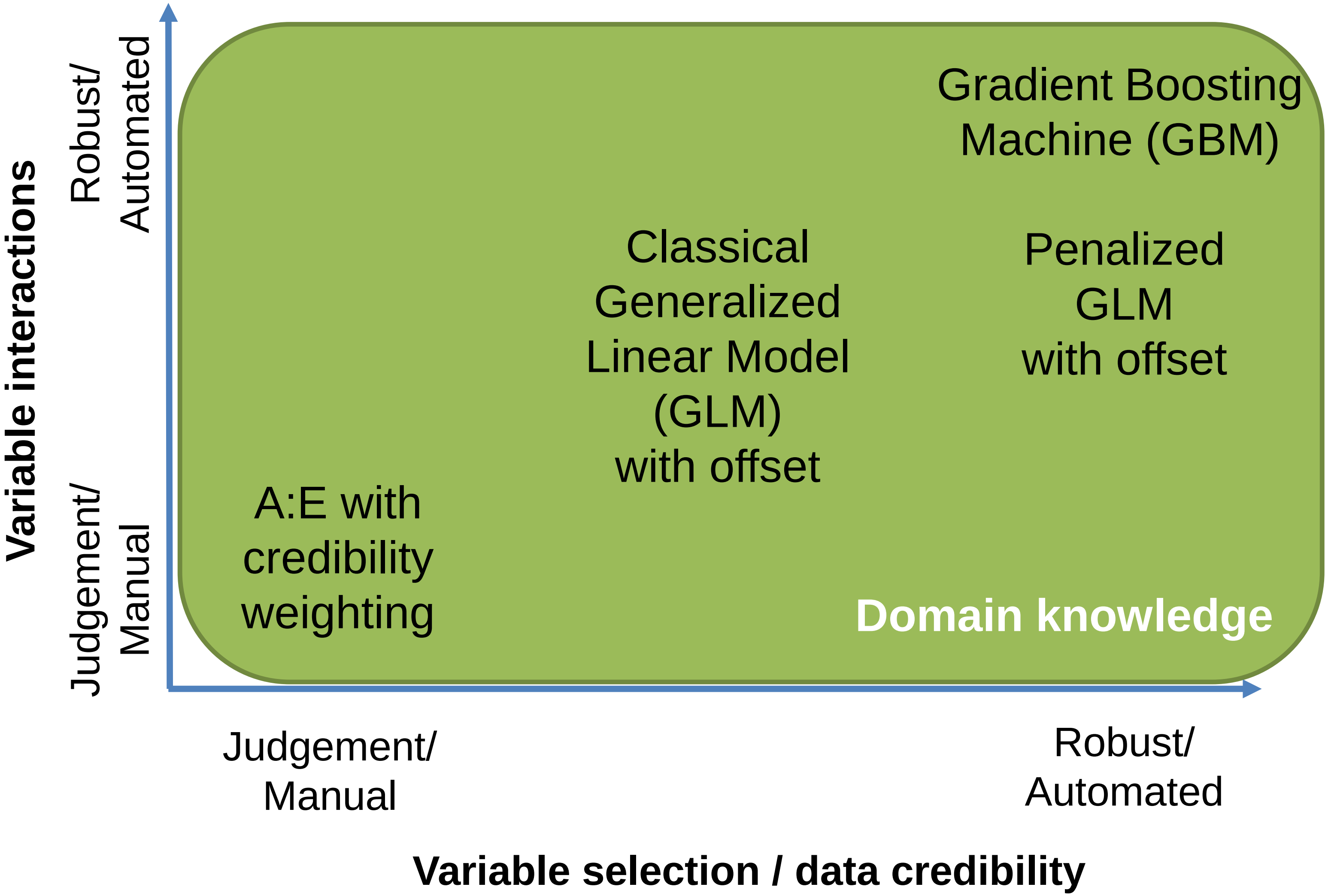
Predictive Analytics

Missy Gordon, FSA, MAAA
Principal and Consulting Actuary
Milliman, Minneapolis

March 20, 2018

ILTCI

Traversing BVT



Traversing BVT

Traversing BVT	A:E with credibility weighting	Classical GLM with Offset	Penalized GLM with Offset	GBM
Data credibility	Judgement	Judgement	Cross validation	Cross validation to tune hyper-parameters to control for overfitting
Variable selection	Judgement	In-sample tests of fit	Cross validation	Automated process to minimize prediction error
Interactions	Judgement	Judgement	Judgement	Automated process to minimize prediction error

Domain knowledge

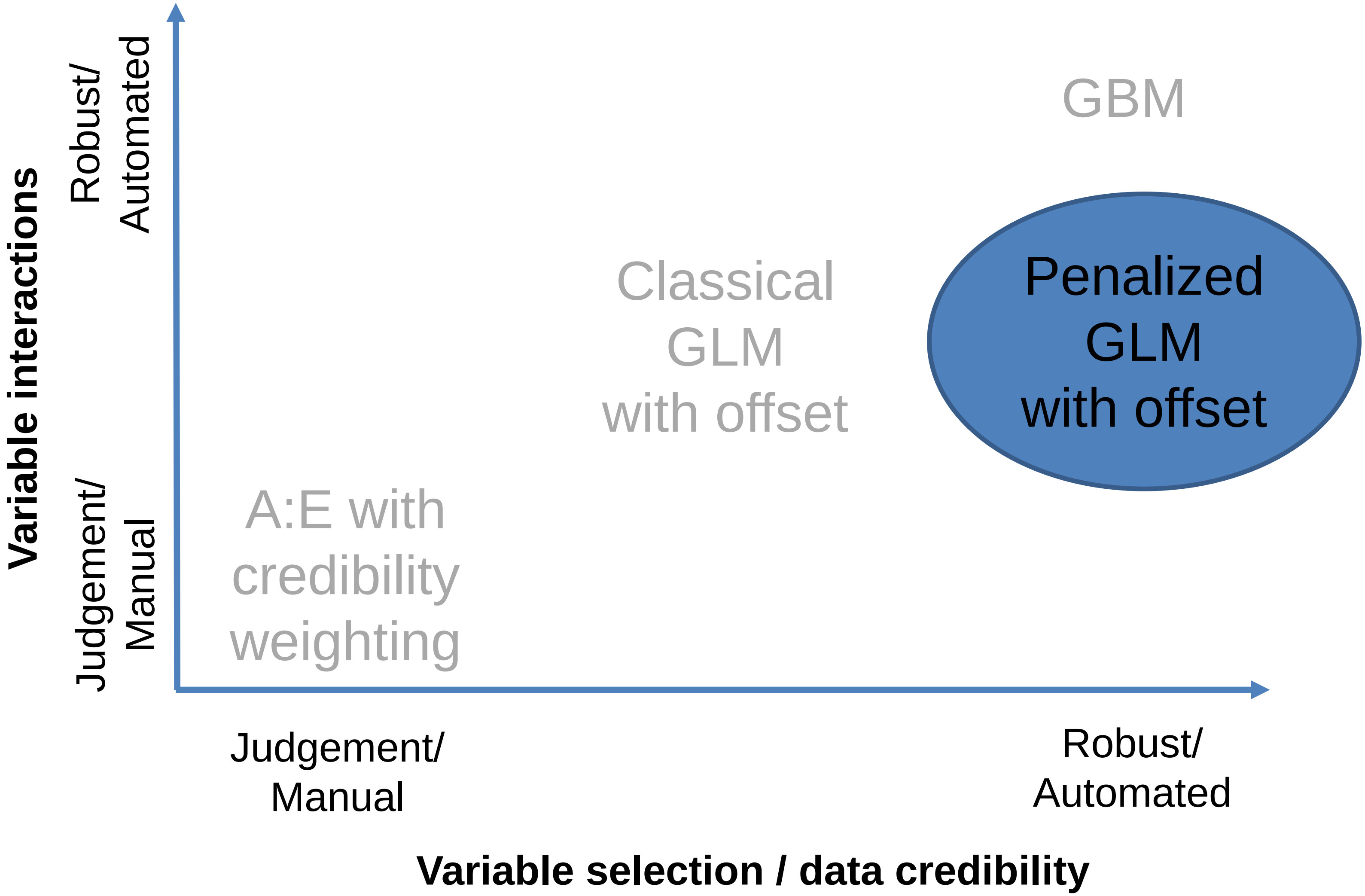
Classical GLM: challenges traversing BVT

1. Gives full credibility to data, unless using judgement

2. Violating underlying GLM assumptions may produce misguided conclusions relative to variable selection

3. Judgement to determine interactions and doesn't handle multicollinearity well

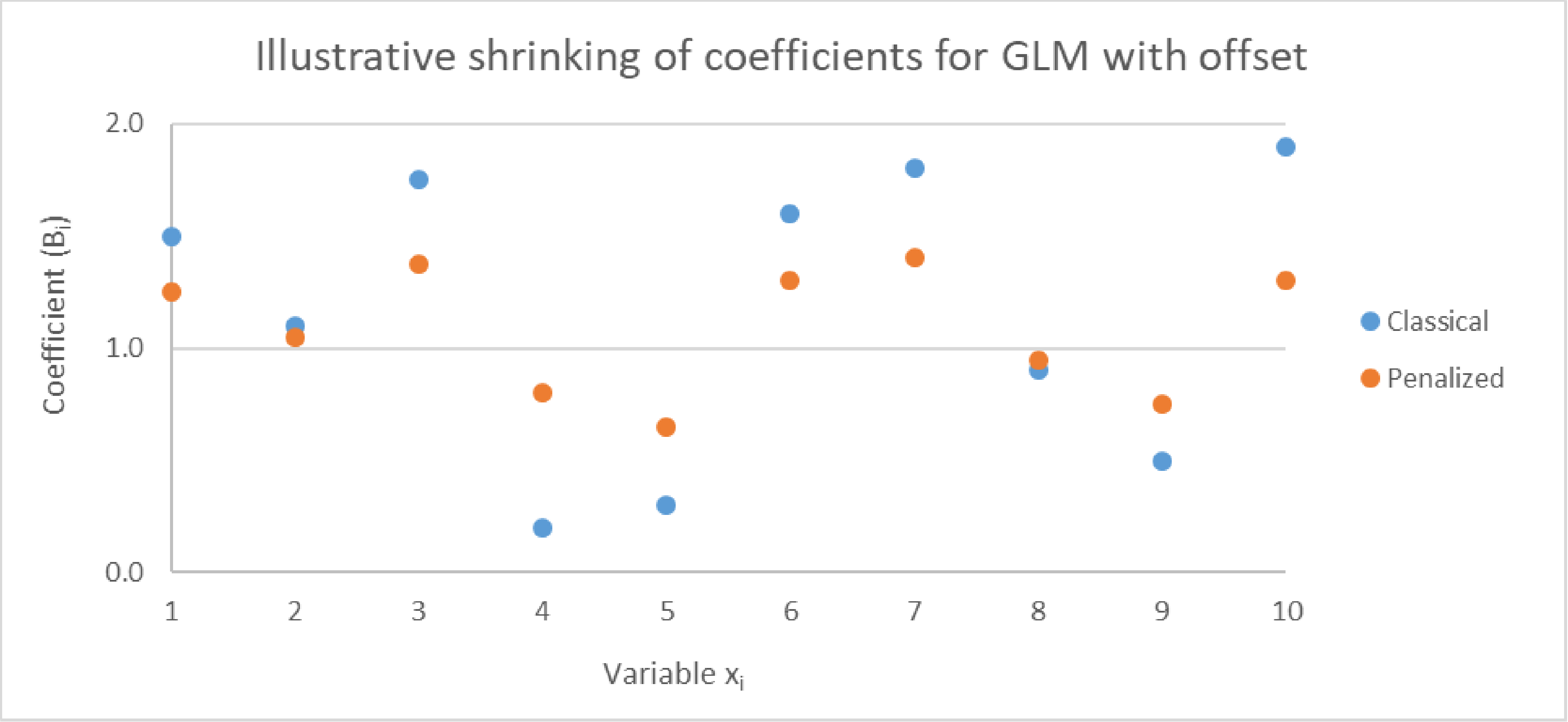
Traversing BVT



Penalized GLM: how it works

- Develops coefficients using GLM with offset
  - Similar to simultaneous A:E adjustments
- Penalizes (shrinks) coefficients
  - Similar to credibility weighting in A:E study
  - Controls for overfitting
  - No penalty (full data weight) = Classical GLM

Penalized GLM: coefficients after shrinking



Penalized GLM: how it traverses BVT

- Data credibility and variable selection by shrinking coefficients
  - Automates decision by minimizing the cross validation prediction error
- Judgement to determine interactions
  - Better handling of multicollinearity
  - Challenge remains of navigating complex interactions

Cross validation: automates traversing BVT

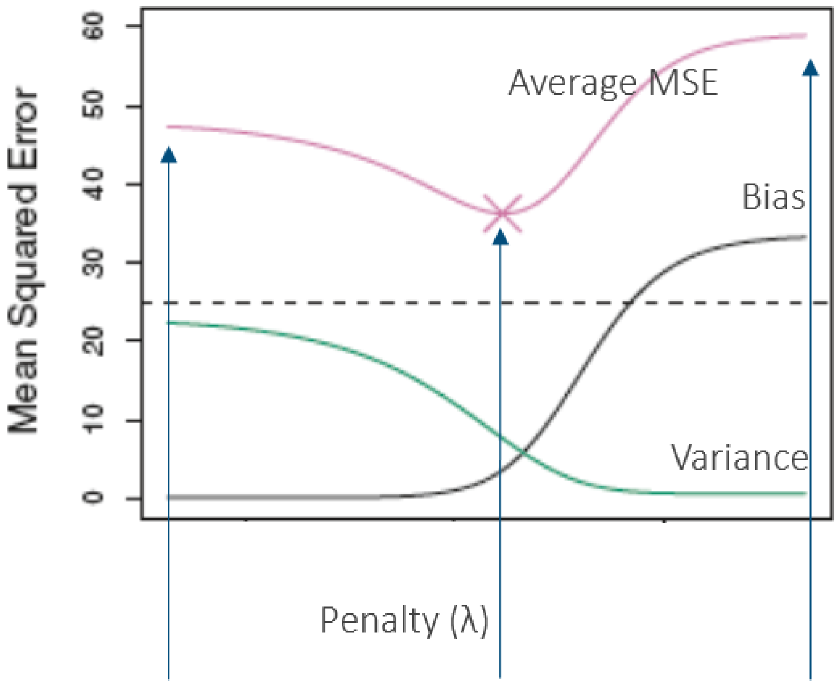
3-Fold	Test 1	Test 2	Test 3	Average
1 33%	1 Holdout	1 Use	1 Use	Test 1
2 33%	2 Use	2 Holdout	2 Use	Test 2
3 33%	3 Use	3 Use	3 Holdout	Test 3
Calibration data	100%	100%	100%	

MSE on holdout data

Automated process!

K-fold cross-validation
- Use subset of data to develop coefficients
- Calculate error of predicted values on holdout data
- Average error across the k tests

Penalized GLM: how it traverses BVT



Overfitting	Balanced	Underfitting
No penalty	Minimize error	Full penalty
Fully trust data	Credibility of data	Don't trust data
(Classical GLM)		(Benchmark)

- Test range of penalties (data credibility)
- Chose penalty that minimizes prediction error
- **Automated process tests thousands of models with a few lines of code!**

Minimizing prediction error

- Objective to minimize MSE or SSE
- Classical GLM: $SSE = \sum(\boldsymbol{Y} - \boldsymbol{x\beta})^2$
- Ridge: $SSE + \lambda * \sum \boldsymbol{\beta}^2$
  - Shrinks coefficients, but remains > 0
  - Helps with multicollinearity

Minimizing prediction error

- LASSO: $SSE + \lambda * \sum|\boldsymbol{\beta}|$
  - Can shrink coefficients to 0
  - Provides automatic feature (variable) selection

- Elastic net: $SSE + \lambda * (\boldsymbol{\alpha} * \sum|\boldsymbol{\beta}| + (\mathbf{1} - \boldsymbol{\alpha}) * \sum \boldsymbol{\beta}^2)$
  - Blend of Ridge and LASSO
  - Helps with multicollinearity and provides feature selection
    - $\boldsymbol{\alpha}$ controls the blend
    - $\boldsymbol{\alpha}$ = 0 then Ridge, $\boldsymbol{\alpha}$ ∈ (0, 1) then Elastic net, $\boldsymbol{\alpha}$ = 1 then LASSO

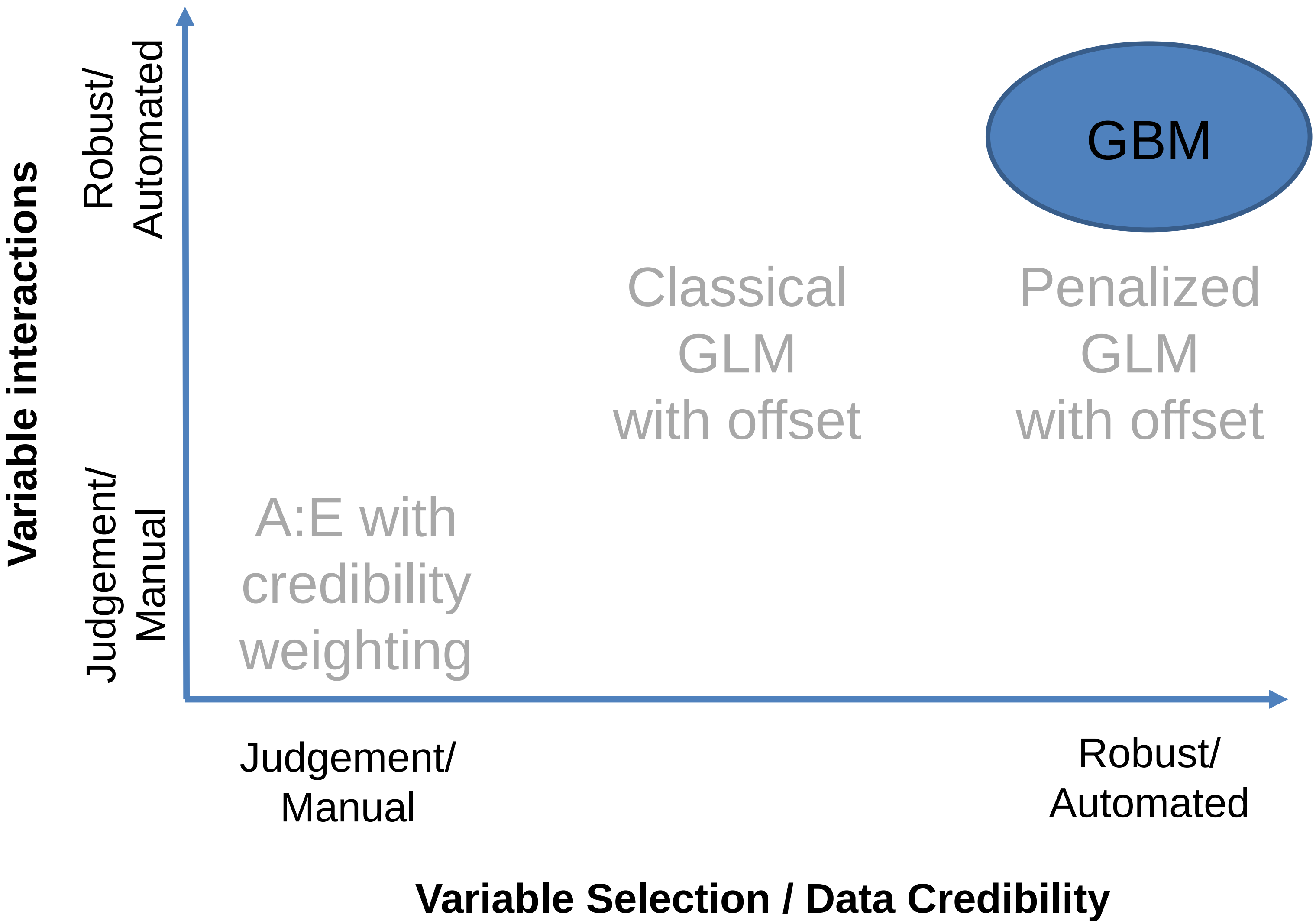
In-sample vs. out-of-sample tests

- Model objective
  - Minimize prediction error on future data
- Training error
  - Optimistic and decreases by adding variables
- Two fixes
  - In-sample tests: theoretical formula increases training error based on number of variables
  - Out-of-sample tests: directly estimates prediction error

In-sample vs. out-of-sample tests

	In-sample tests	Out-of-sample tests
	AIC, BIC, Adjusted R^2 p-values to prune parameters	Separate train/test datasets k-fold cross validation
Pros	- Model selection using all data - Fast to calculate	- No theoretical formulas - Compare across algorithms
Cons	- Relies on theoretical formulas - May misguide if assumptions violated - Harder (or not possible) to compare across algorithms	- Computationally expensive - Potential to misuse if not setup properly (information leak)

Traversing BVT



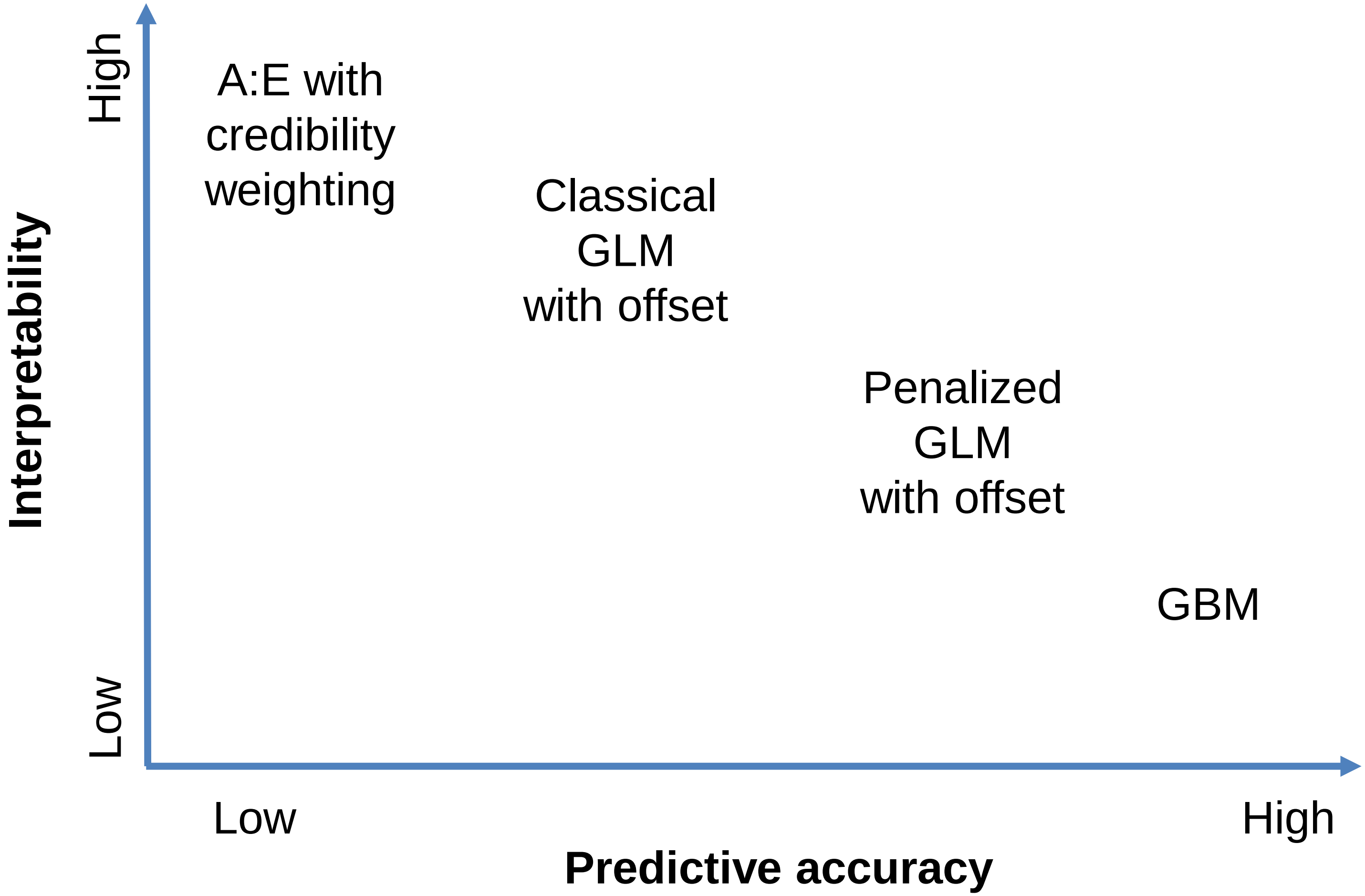
GBM: how it works

- Develops layers of “A:E” adjustments
- Layers of decision trees to minimize error
  - Slices data to create variable buckets
  - At each point tests every variable and possible slice to minimize error

GBM: how it traverses BVT

- Variable selection and interactions
  - Non-parametric model
  - Automates decisions by minimizing the prediction error
  - **Handles complex interactions and provides information on variable importance**
- Data credibility incorporated using cross validation to tune hyperparameters that control for overfitting

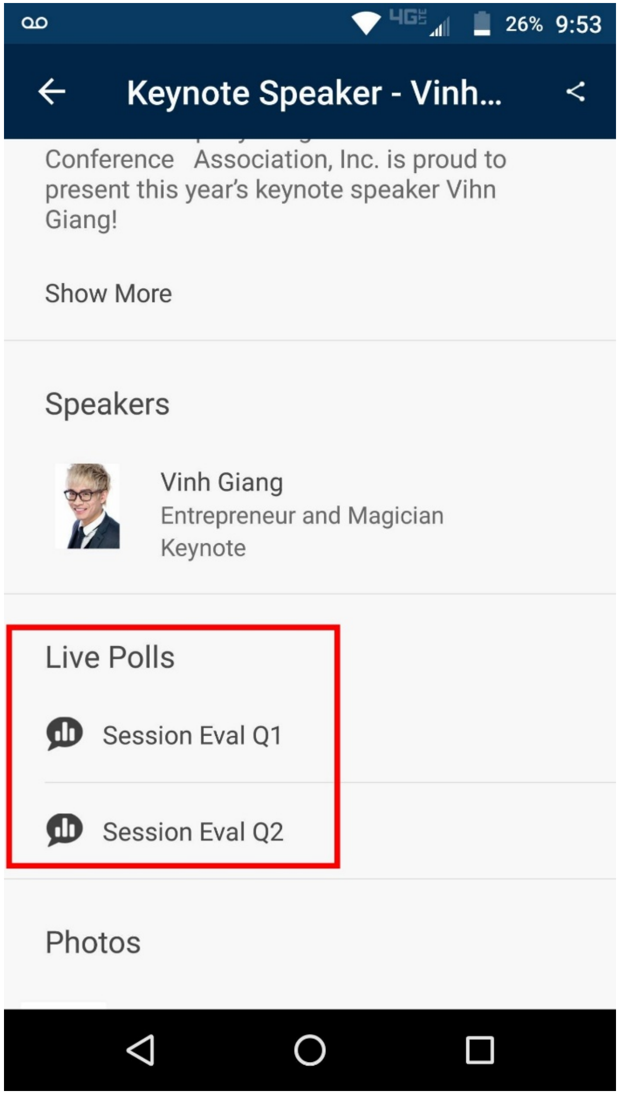
Model interpretability vs. accuracy



Why might we want lower accuracy?

- Stepping stone
  - Isolate changes from one model to the next
- Assumption format
  - Higher inference: multiplicative factors
  - Lower inference: sets of tables or seriatim
- Purpose / materiality

Session Survey Instructions



Once you are in the app go to the schedule and the session you are in.

Scroll to the bottom to find the Live Polling questions.

This year the session survey questions can be found in this section and will take just a couple seconds to complete.

Q&A