V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 1

Logistic Regression: MLE vs. OLS1 in Excel 2013

by
Milo Schield
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project

Slides and data at: www.StatLit.org/
pdf/2016-Schield-Logistic-MLE-OLS1-Excel2013-slides.pdf
pdf/2016-Schield-Logistic-MLE-OLS1-Excel2013-demo.pdf
Excel/2016-Schield-Logistic-MLE-OLS1-Excel2013.xlsx

V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 2

Background & Goals

Modeling a binary outcome requires a logistic model.
Doing logistic regression properly requires MLE.
Doing MLE in Excel is not easy. See Schield (2015)

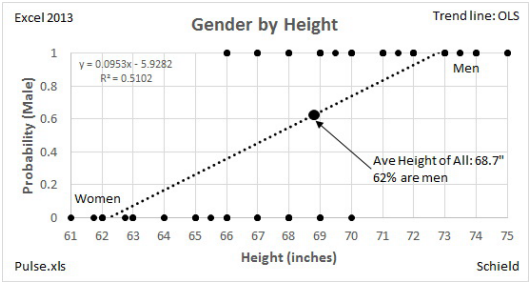
Schield created two logistic OLS models: OLS1+OLS3
OLS1: Model Ln[Odds(Pnudge)]. See Schield (2014a).
OLS2: Model Ln[Odds(Pgroup)]. See Schield (2016c)
OLS3: Use OLS to estimate logistic parameters.
See Schield (2014b)

These slides compare MLE with OLS1 logistic regression

V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 3

Model Gender by Height (OLS) Must use a logistic function

This linear trend-line goes outside the valid range



V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 4

Model using a Logistic Function

Range of Odds(p): Zero to infinity
Range of Ln[Odds(p)]: Minus infinity to infinity
Logistic model: Ln[Odds(p)] = a + b*X

V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 5

|Ln[Odds(p)]| = infinity if p = 0 or 1 Pnudge: =If(p=0, 0.001, 0.999)

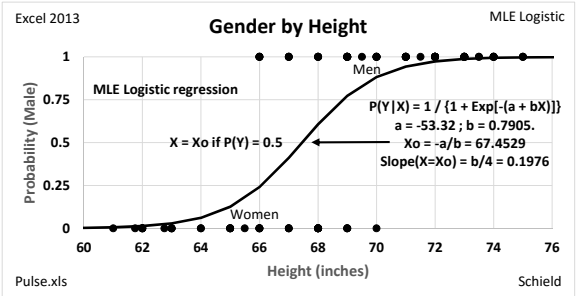
Source: Pulse dataset

A	B	C	D	E	
2	Height	Weight	Male	Male1	
3	68	150	1	0.999	=IF(D3=0, 0.001, 0.999)
4	69	145	1	0.999	
5	69	160	1	0.999	
6	72	145	1	0.999	
7	66	135	1	0.999	
	67	125	0	0.001	
	71	170	1	0.999	
	71	155	1	0.999	
	71.5	164	1	0.999	
	62	120	0	0.001	
	65.5	120	0	0.001	

OLS1 Model : Ln[Odds(Pnudge)] = Constant + b1*X1 + b2*X2 using Ordinary Least Squares.

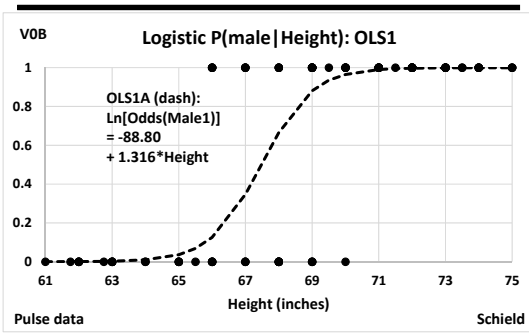
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 6

1a: Logistic P(male | Height) MLE



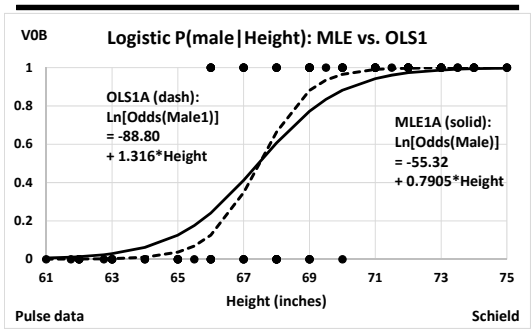
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 7

1b: Logistic P(male | Height) OLS1



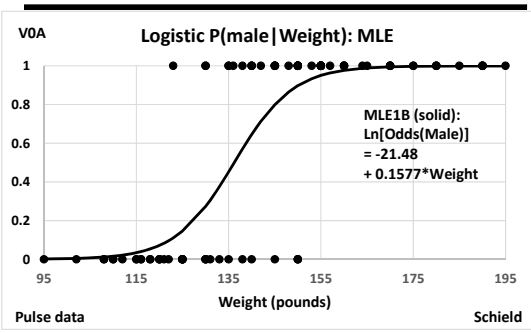
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 8

1c: Logistic P(male | Height) MLE vs. OLS1



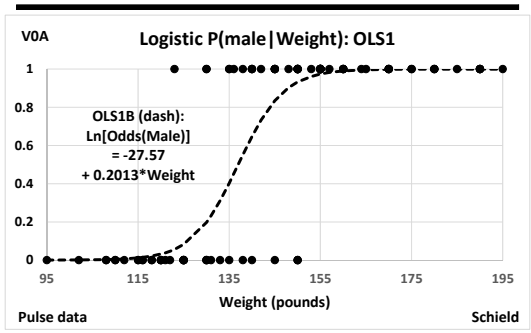
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 9

2a: Logistic P(male | Weight) MLE



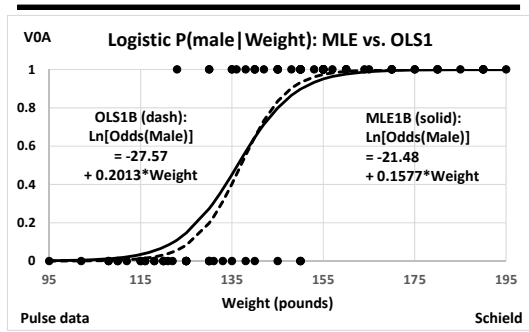
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 10

2b: Logistic P(male | Weight) OLS1



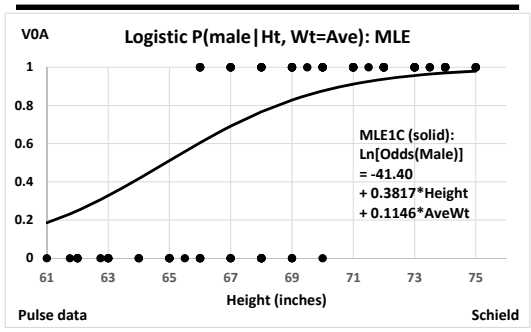
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 11

2c: Logistic P(male | Weight) MLE vs. OLS1



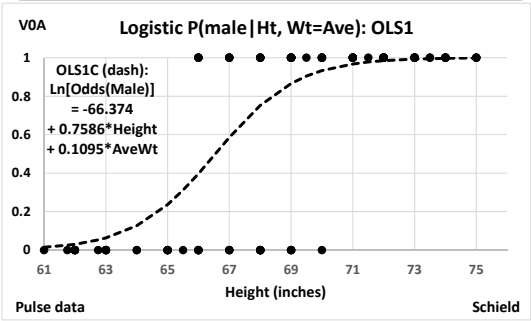
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 12

3a: Logistic P(male | Ht, Wt=Ave) MLE



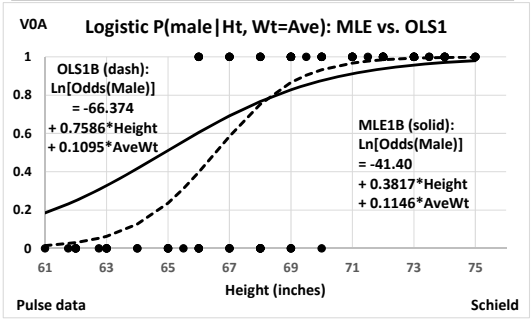
V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 13

3b: Logistic P(male | Ht, Wt=Ave) OLS1



V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 14

3c: Logistic P(male | Ht, Wt=Ave) MLE vs. OLS1



V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 15

Conclusions

1. Can use OLS with a logistic function to illustrate logistic regression since logistic-OLS is better than linear OLS.
2. OLS with 'nudge', a single predictor and logistic (OLS1A) is quite close to the MLE.
3. OLS with 'nudge' and logistic worsens as number of predictors increases (OLS1C) .
4. If higher accuracy is needed, use MLE (Excel or other) or consult a statistician.

V0B Schield MLE vs. OLS1-Based Logistic Excel 2013 16

References

Schield, Milo (2015). Copy at www.statlit.org/pdf/
2015-Schield-Logistic-MLE1A-Excel2013-Slides.pdf
2015-Schield-Logistic-MLE1B-Excel2013-Slides.pdf
2015-Schield-Logistic-MLE1C-Excel2013-Slides.pdf

Schield, Milo (2014a). Copy at www.statlit.org/pdf/
2014-Schield-Logistic-OLS1A-Excel2013-Demo.pdf
2014-Schield-Logistic-OLS1B-Excel2013-Demo.pdf
2014-Schield-Logistic-OLS1C-Excel2013-Demo.pdf

Schield, Milo (2014b). Copy at www.statlit.org/pdf/
2014-Schield-Logistic-MLE-OLS3-Excel2013-Demo.pdf

Logistic Regression: MLE vs. OLS1 in Excel 2013

by
Milo Schield
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project

Slides and data at: www.StatLit.org/
pdf/2016-Schield-Logistic-MLE-OLS1-Excel2013-slides.pdf
pdf/2016-Schield-Logistic-MLE-OLS1-Excel2013-demo.pdf
Excel/2016-Schield-Logistic-MLE-OLS1-Excel2013.xlsx

Background & Goals

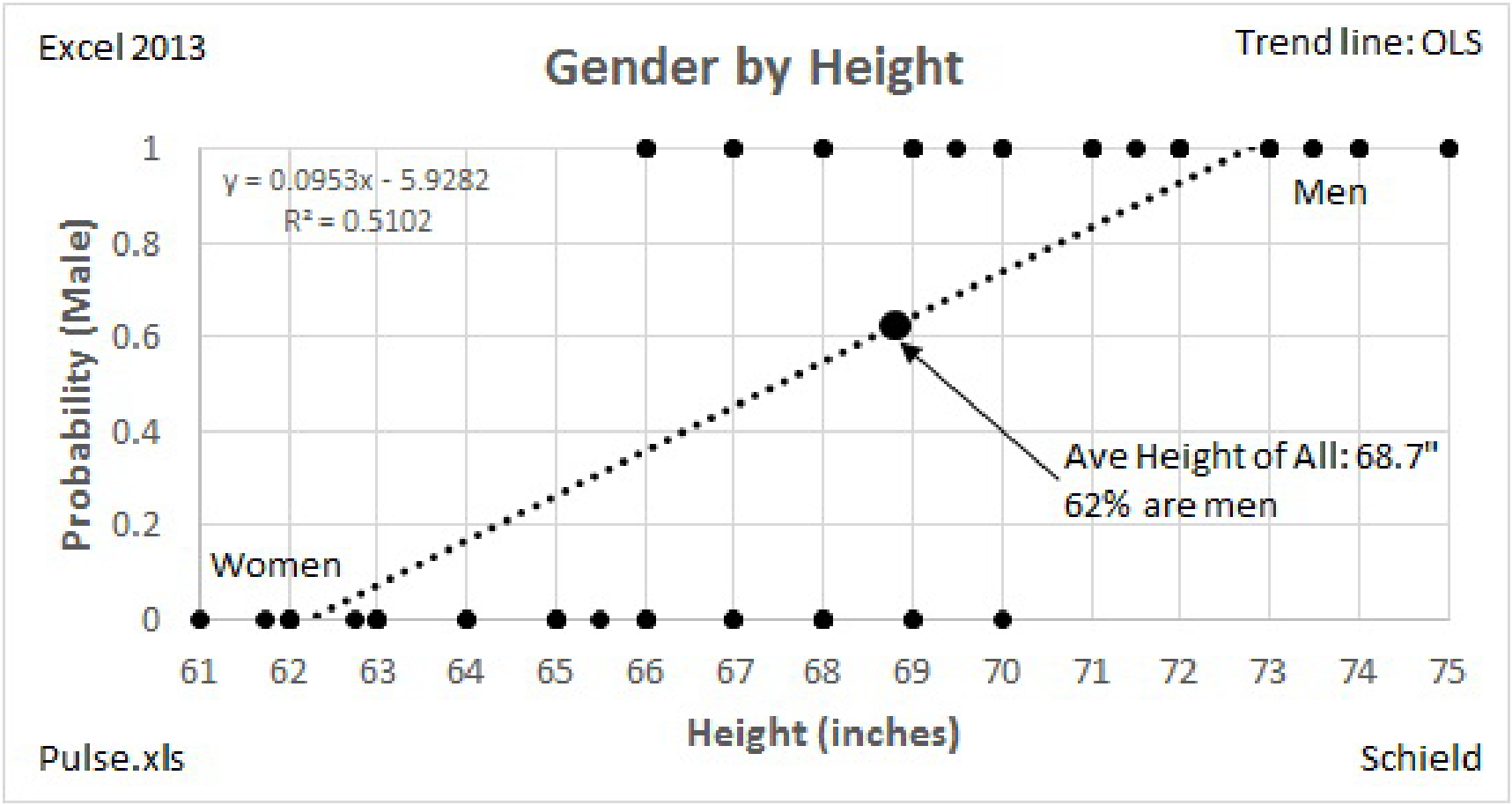
Modeling a binary outcome requires a logistic model.
Doing logistic regression properly requires MLE.
Doing MLE in Excel is not easy. See Schield (2015)

Schield created two logistic OLS models: OLS1+OLS3
OLS1: Model Ln[Odds(Pnudge)]. See Schield (2014a).
OLS2: Model Ln[Odds(Pgroup)]. See Schield (2016c)
OLS3: Use OLS to estimate logistic parameters.
See Schield (2014b)

These slides compare MLE with OLS1 logistic regression

Model Gender by Height (OLS) Must use a logistic function

This linear trend-line goes outside the valid range



Model using a Logistic Function

Range of Odds(p): Zero to infinity

Range of Ln[Odds(p)]: Minus infinity to infinity

Logistic model: Ln[Odds(p)] = a + b*X

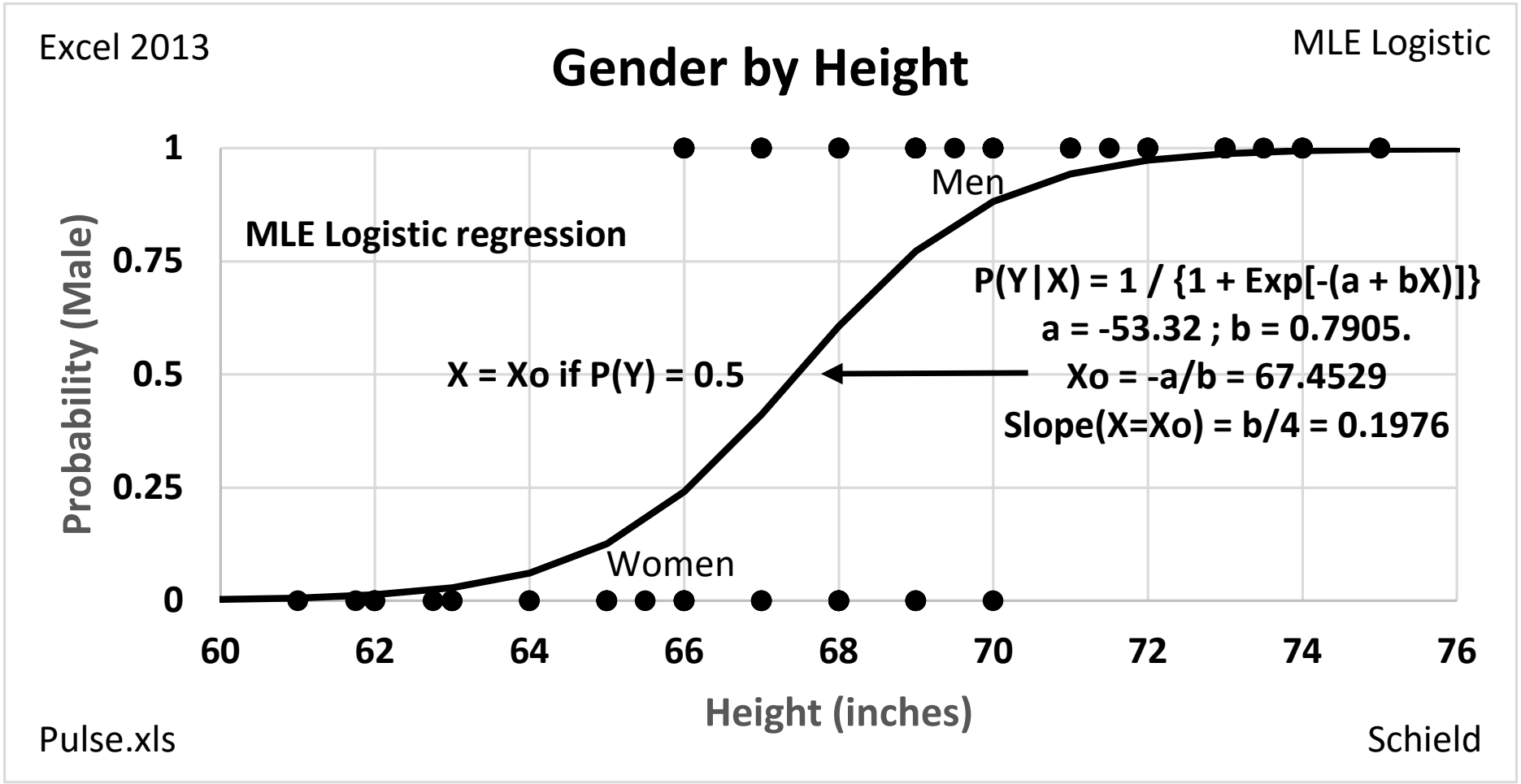
|Ln[Odds(p)]| = infinity if p = 0 or 1
Pnudge: =If(p=0, 0.001, 0.999)

Source: Pulse dataset

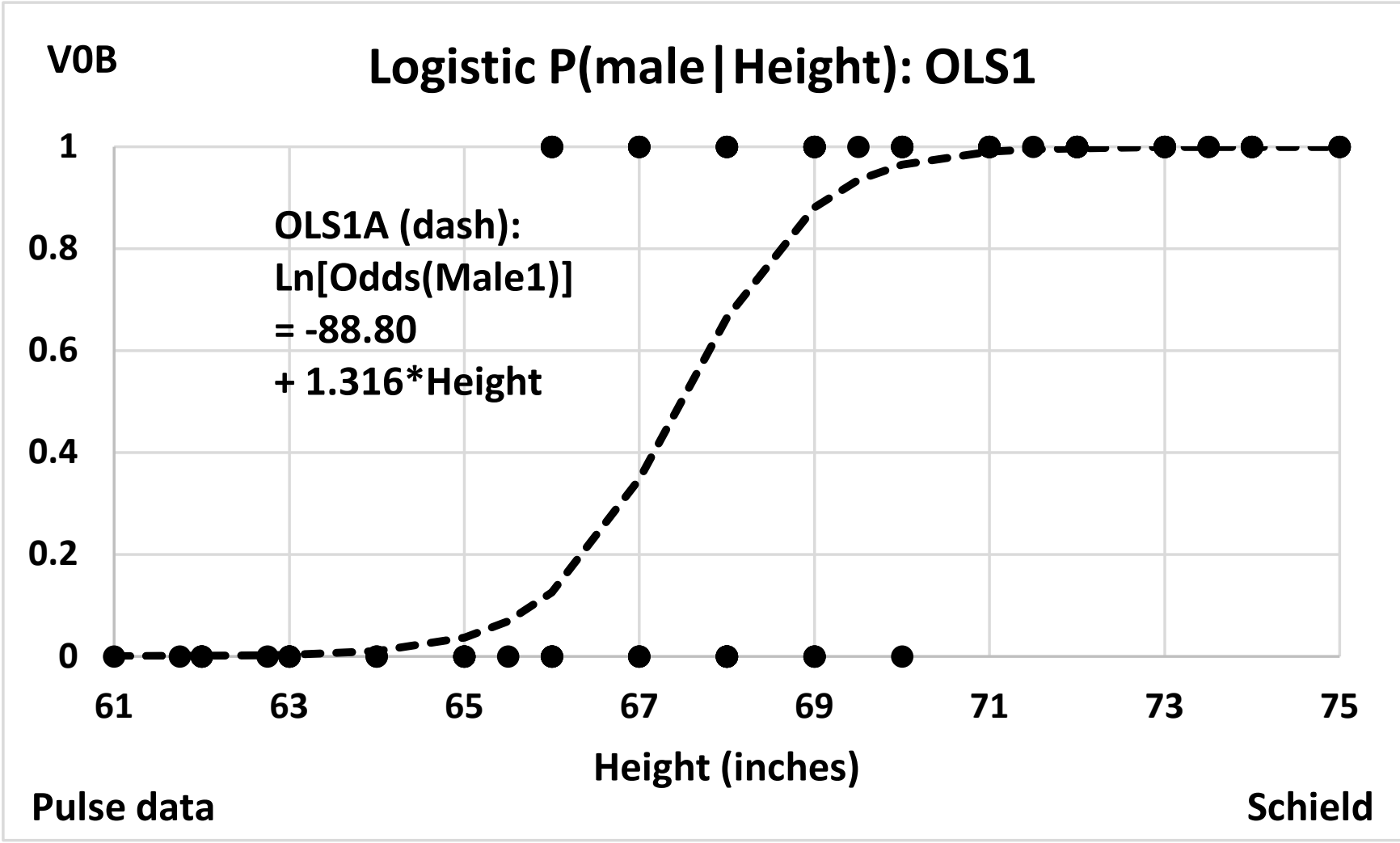
A	B	C	D	E	
2	Height	Weight	Male	Male1	
3	68	150	1	0.999	=IF(D3=0, 0.001, 0.999)
4	69	145	1	0.999	
5	69	160	1	0.999	
6	72	145	1	0.999	
7	66	135	1	0.999	
	67	125	0	0.001	
	71	170	1	0.999	
	71	155	1	0.999	
	71.5	164	1	0.999	
	62	120	0	0.001	
	65.5	120	0	0.001	

OLS1 Model : Ln[Odds(Pnudge)] = Constant + b1*X1 + b2*X2 using Ordinary Least Squares.

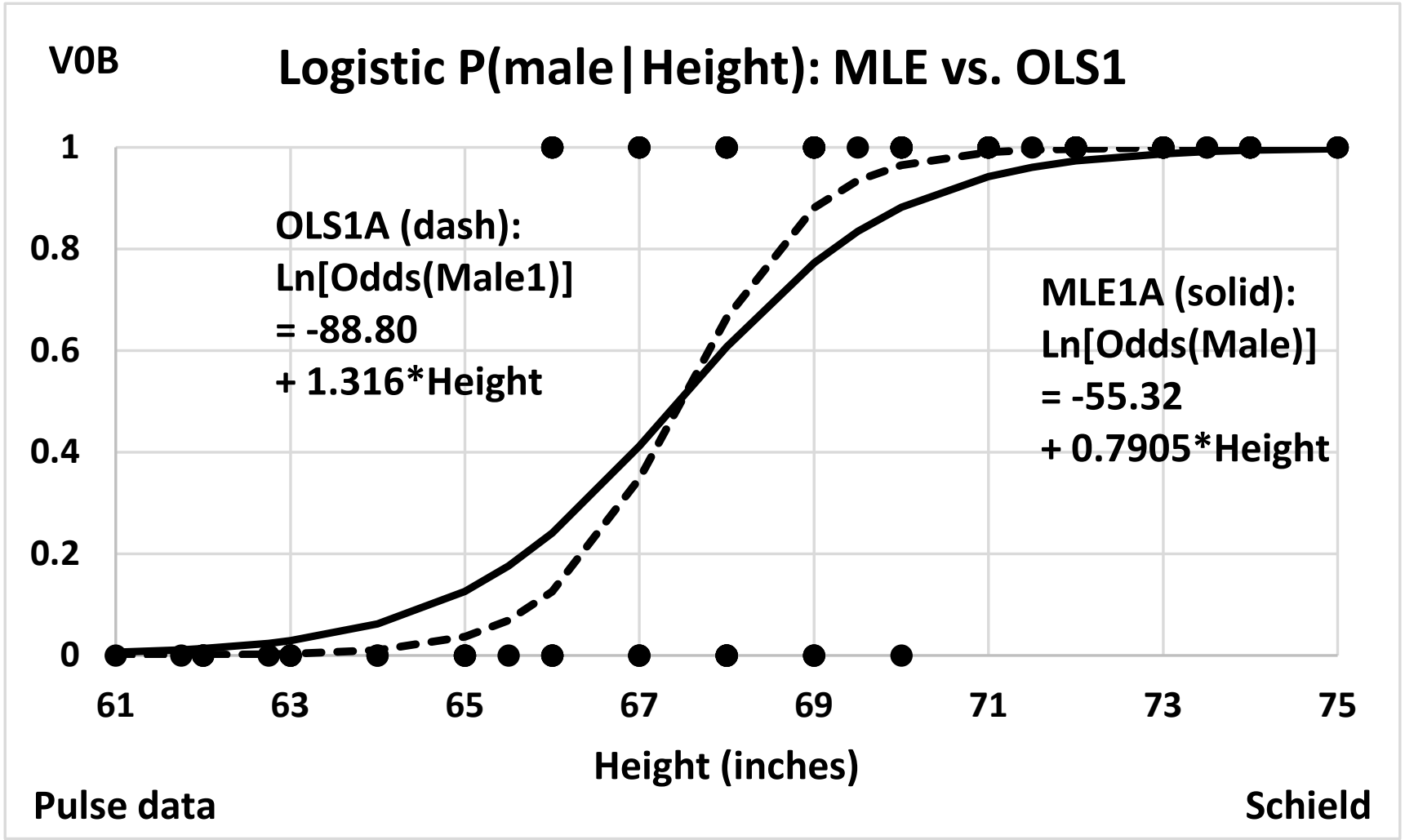
1a: Logistic P(male | Height) MLE



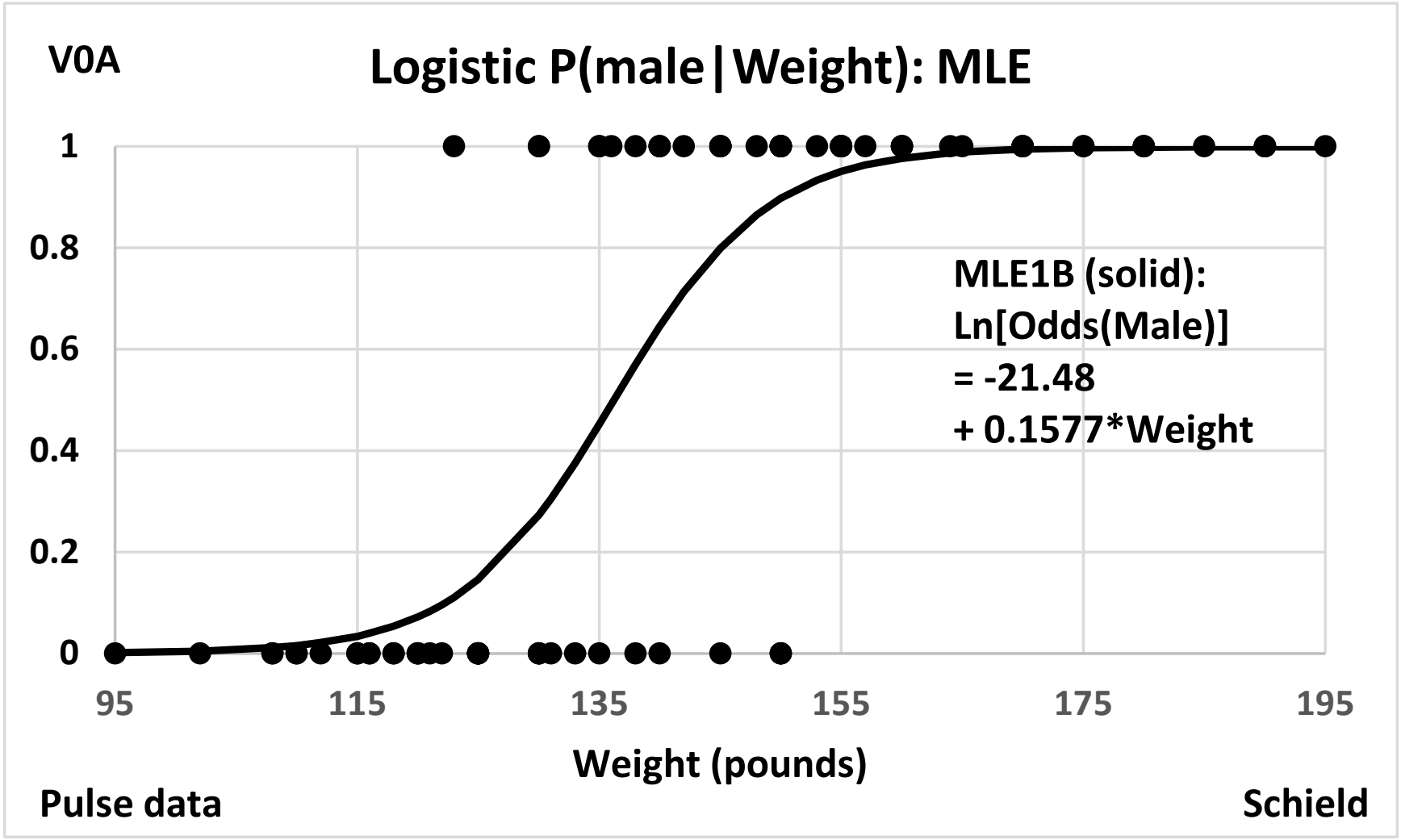
1b: Logistic P(male|Height) OLS1



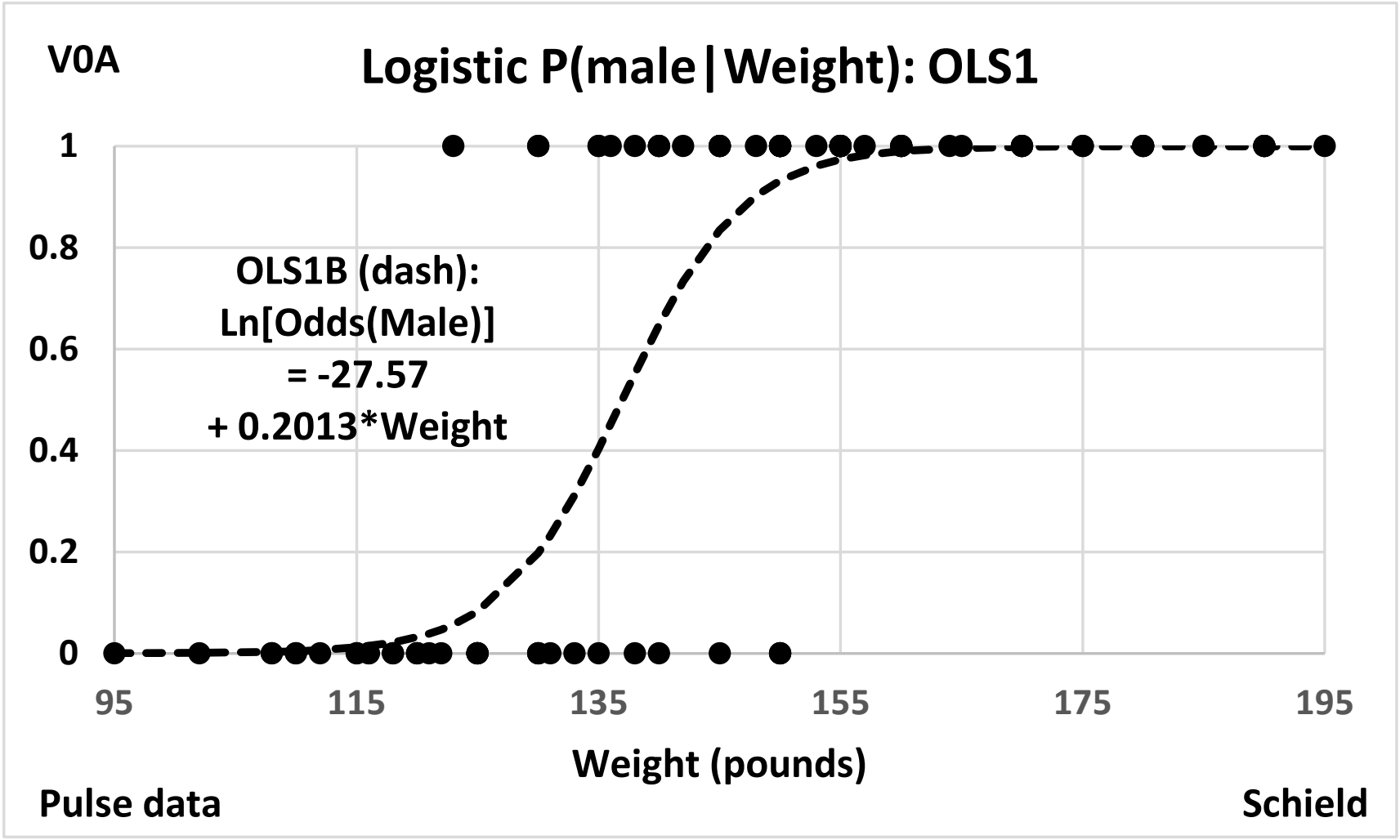
1c: Logistic P(male|Height) MLE vs. OLS1



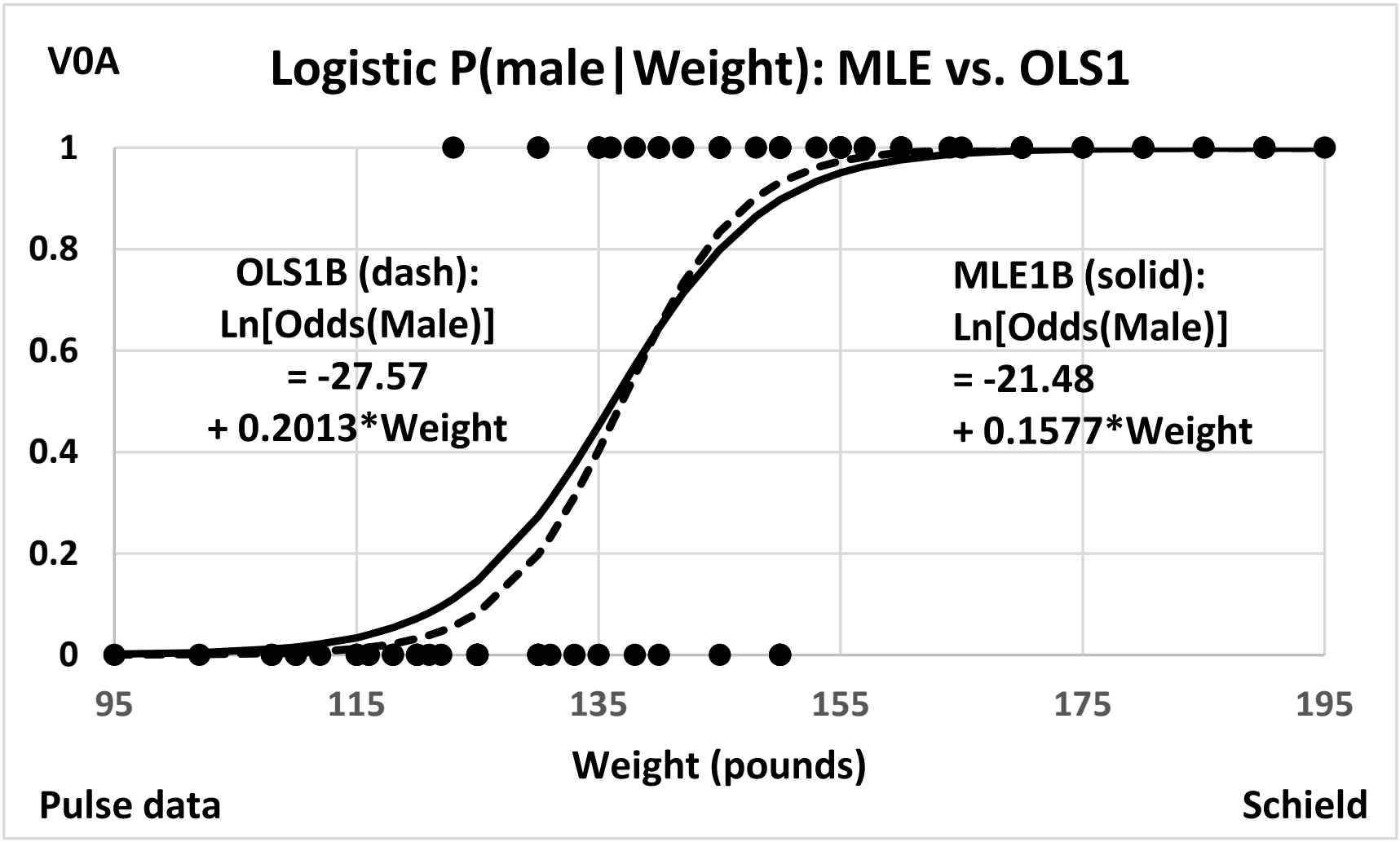
2a: Logistic P(male|Weight) MLE



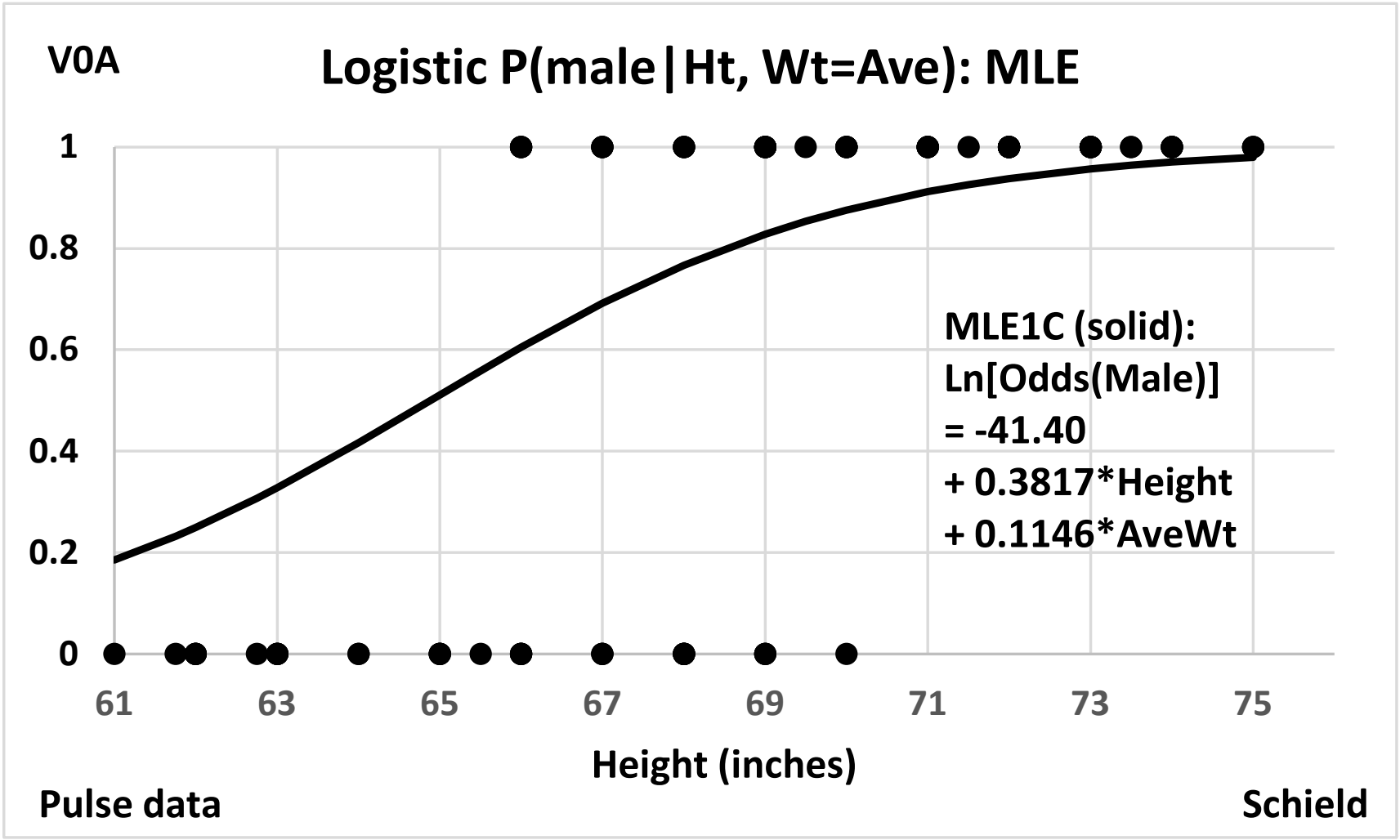
2b: Logistic P(male|Weight) OLS1



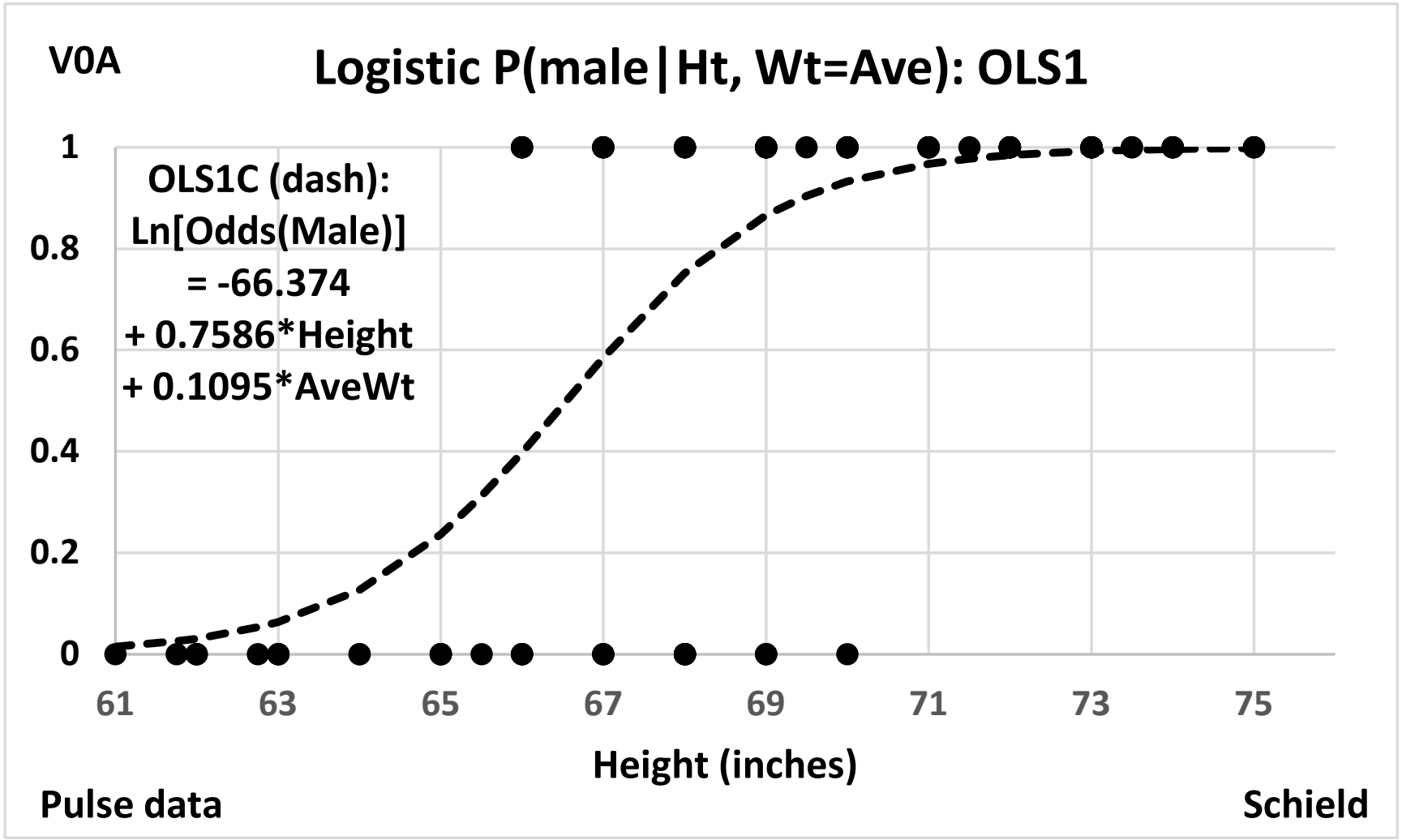
2c: Logistic P(male|Weight) MLE vs. OLS1



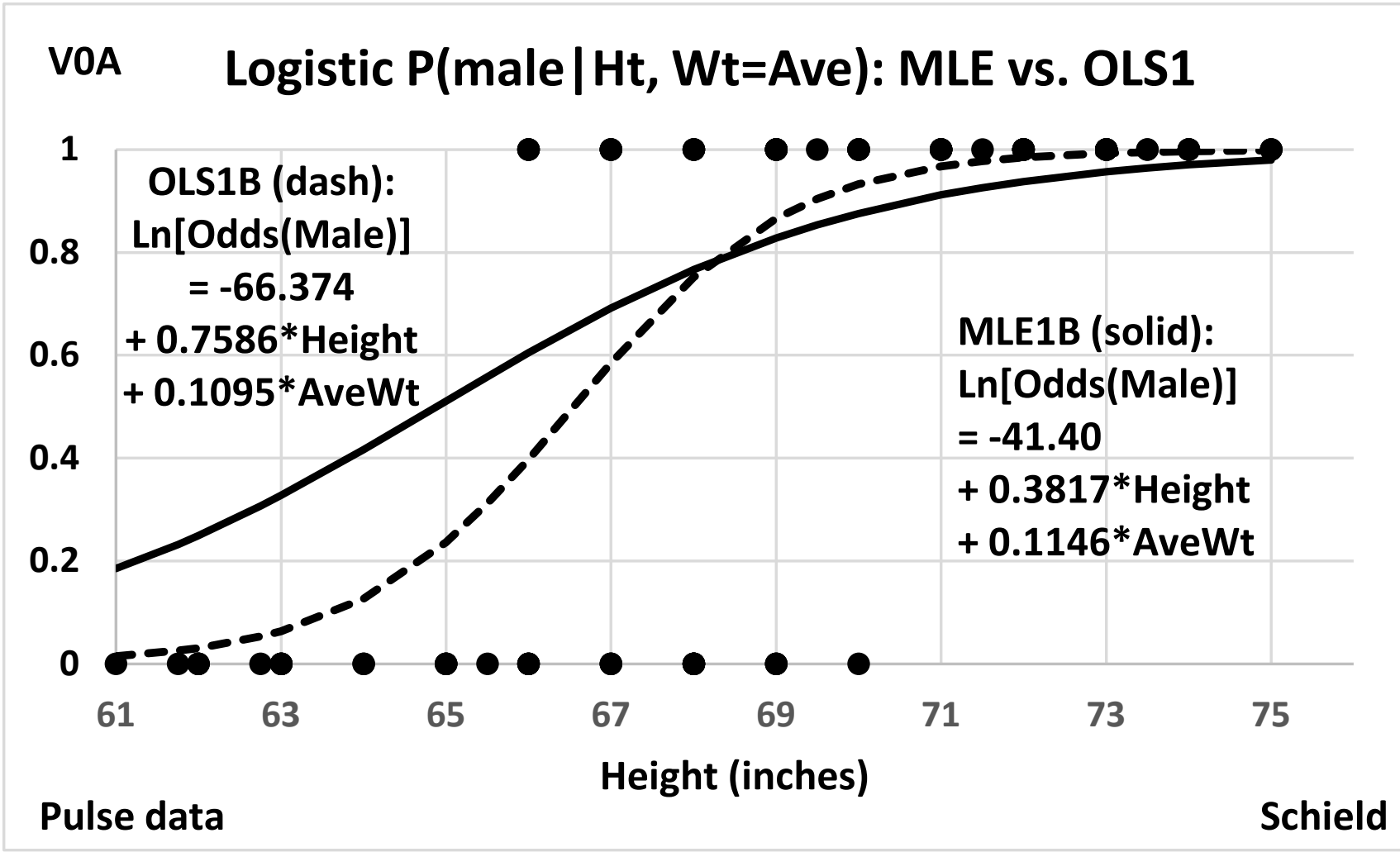
3a: Logistic P(male | Ht, Wt=Ave) MLE



3b: Logistic P(male|Ht, Wt=Ave) OLS1



3c: Logistic P(male|Ht, Wt=Ave) MLE vs. OLS1



Conclusions

1. Can use OLS with a logistic function to illustrate logistic regression since logistic-OLS is better than linear OLS.
2. OLS with ‘nudge’, a single predictor and logistic (OLS1A) is quite close to the MLE.
3. OLS with ‘nudge’ and logistic worsens as number of predictors increases (OLS1C) .
4. If higher accuracy is needed, use MLE (Excel or other) or consult a statistician.

References

Schield, Milo (2015). Copy at www.statlit.org/pdf/
2015-Schield-Logistic-MLE1A-Excel2013-Slides.pdf
2015-Schield-Logistic-MLE1B-Excel2013-Slides.pdf
2015-Schield-Logistic-MLE1C-Excel2013-Slides.pdf

Schield, Milo (2014a). Copy at www.statlit.org/pdf/
2014-Schield-Logistic-OLS1A-Excel2013-Demo.pdf
2014-Schield-Logistic-OLS1B-Excel2013-Demo.pdf
2014-Schield-Logistic-OLS1C-Excel2013-Demo.pdf

Schield, Milo (2014b). Copy at www.statlit.org/pdf/
2014-Schield-Logistic-MLE-OLS3-Excel2013-Demo.pdf