Supplementary Materials for Learning Models from Data with Measurement Error: Tackling Underreporting

Roy Adams [1] **Yuelong Ji** [2] **Xiaobin Wang** [2] **Suchi Saria** [1 3 4]

1. Proof of Corollary 1 for probit and cloglog models

In this section we prove Corollary 1 in Section 4.3 of the main paper for the cases of probit and cloglog regression.

Proof. Following the proof for logistic regression, a probit regression model violates the condition in Theorem 1 if

$$\sqrt{2}\,\text{erf}^{-1}\left(\alpha - 1 + \alpha\,\text{erf}\left(\frac{\phi x}{\sqrt{2}}\right)\right) = \phi' x$$

Similarly, a cloglog regression model violates the condition in Theorem 1 if

$$\log(-\log(1 - \alpha + \alpha\exp(-\exp(\phi x)))) = \phi' x$$

In both of these cases, the function on the left hand side is non-linear for $\alpha < 1$, so these equalities can only be true if ϕx and $\phi' x$ are constants which is true only when $A \perp X$. □

[1]Department of Computer Science, Johns Hopkins University [2]Center on the Life Origins of Disease, Department of Population, Family, and Reporductive Health, Johns Hopkins University Bloomberg School of Public Health [3]Department of Applied Math and Statistics, Johns Hopkins University [4]Bayesian Health. Correspondence to: Roy Adams <roy.james.adams@gmail.com>.

Proceedings of the 36^{th} International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.