## A Supplemental figures

1) Estimate initial prediction functions 
$$\hat{g}(t,\vec{x}) \approx \mathbf{E}\left(Y \mid T=t,\vec{X}=\vec{x}\right)$$

$$\hat{m}(t,\vec{x}) \approx P\left(T=t \mid \vec{X}\right)$$
2) Define a new covariate 
$$\hat{H}(T,\vec{X}) = \frac{\mathbb{I}(T=t)}{\hat{m}(t,\vec{X})}$$
Sample split: Carry out step 3 in a different sample from steps 1 and 2

3) Regress Y on the new covariate with an offset 
$$\mathbf{E}(Y \mid T,\vec{X}) \approx \hat{g}(T,\vec{X}) + \hat{H}(T,\vec{X})\beta$$
Offset Clever Coefficient term covariate (from 1) (from 2) 
$$\mathbf{f}(T,\vec{X}) = \hat{g}(T,\vec{X}) + \hat{H}(T,\vec{X})\beta$$
Target the prediction function 
$$\mathbf{f}(T,\vec{X}) = \hat{g}(T,\vec{X}) + \hat{H}(T,\vec{X})\beta$$
Original prediction rule optimized for the way we will aggregate prediction 
$$\mathbf{f}(T,\vec{X}) = \hat{g}(T,\vec{X}) + \hat{H}(T,\vec{X})\beta$$
Targeted prediction rule optimized for the way we will aggregate prediction 
$$\mathbf{f}(T,\vec{X}) = \hat{g}(T,\vec{X}) + \hat{H}(T,\vec{X})\beta$$
Original prediction rule optimized for disaggregate prediction function 
$$\hat{\mathbf{E}}(Y(t)) = \frac{1}{n} \sum_{i=1}^{n} \hat{g}'(t,\vec{X}_i)$$

Fig. 12. Targeted learning with a continuous outcome. This method is analogous to the method for a binary outcome presented in Figure 11. We include the continuous version here for comparison with double machine learning (Appendix Figure 13).

**Fig. 13.** Double machine learning. This figure presents double machine learning (Chernozhukov et al., 2018) using the notation of targeted learning (Van der Laan and Rose, 2018) in order to emphasize the parallels between the two methods. For targeted learning with a continuous outcome, see Appendix Figure 12. For targeted learning with a binary outcome, see Figure 11.