A PROOFS OF THE THEOREMS

Whenever the context of an expectation operation is not clear, we disambiguate it by specifying the variable that the expectation is taken over and its distribution $\mathbb{E}_{\mathbf{x}\sim f(\boldsymbol{x})}[\mathbf{x}]$.

A.1 PROOF OF THEOREM 1

Proof. Given that the logarithm function is a strictly increasing function, we can omit it in the optimization; i.e., $\boldsymbol{\lambda}^\star = \operatorname{argmin}_{\boldsymbol{\lambda}} \mathbb{E}\left[\exp(\mathbf{g}^\top \boldsymbol{\lambda} + \ell)\right]$. Because this is an unconstrained optimization, the optimal solution occurs when the gradient is equal to zero.

$$\begin{aligned}\mathbb{E}[\mathbf{g}\exp\left(\mathbf{g}^\top \boldsymbol{\lambda}^\star + \ell\right)] &= \mathbf{0},\\ \mathbb{E}[\mathbf{g} w^\star(\mathrm{a},\mathbf{x})] &= \mathbf{0},\end{aligned} \tag{7}$$

where the last equation is due to the equation of the weights in the population optimization problem.

Using the definition for the $\mathbf{g}$ vector, Eq. (7) implies that $\mathbb{E}[w^\star(\mathrm{a},\mathbf{x})\mathrm{a}\boldsymbol{\phi}(\mathbf{x})] = \mathbf{0}$. Thus, we conclude that in the weighted population (with distribution $\widetilde{F}$), the a and $\boldsymbol{\phi}(\mathbf{x})$ are uncorrelated:

$$\mathbb{E}_{(\mathrm{a},\mathbf{x})\sim\widetilde{F}}[\mathrm{a}\boldsymbol{\phi}(\mathbf{x})] = \mathbf{0} \tag{8}$$

For every set $\mathcal{B} \subset \mathbb{A}\times\mathbb{X}$, we can write:

$$\widetilde{F}(\mathcal{B}) = \int_{\mathcal{B}} w^\star(a,\boldsymbol{x})\mathrm{d}F(a,\boldsymbol{x}). \tag{9}$$

The Radon-Nikodym theorem implies that $w^\star(a,\boldsymbol{x})$ is the Radon-Nikodym derivative:

$$w^\star(\boldsymbol{x},a) = \frac{\mathrm{d}\widetilde{F}(\boldsymbol{x},a)}{\mathrm{d}F(\boldsymbol{x},a)} = \frac{\widetilde{f}(\boldsymbol{x},a)}{f(\boldsymbol{x},a)} \tag{10}$$

$$= \frac{\widetilde{f}(\boldsymbol{x})\widetilde{f}(a) + \left\{\widetilde{f}(\boldsymbol{x},a) - \widetilde{f}(\boldsymbol{x})\widetilde{f}(a)\right\}}{f(\boldsymbol{x},a)}, \tag{11}$$

$$= w_{GSW}(a,\boldsymbol{x}) + \frac{\widetilde{f}(\boldsymbol{x},a) - \widetilde{f}(\boldsymbol{x})\widetilde{f}(a)}{f(\boldsymbol{x},a)} \tag{12}$$

Thus, using Eq. (8) and Assumptions 1 and 3 we can write

$$\sup_{a,\boldsymbol{x}} |w^\star(a,\boldsymbol{x}) - w_{GSW}(a,\boldsymbol{x})| \leq \delta_{\boldsymbol{\phi}_K}/c.$$

□

A.2 THEOREM 2

Proof. Given that the logarithm function is a strictly increasing function, we can omit it in the optimizations. Thus the sample and population solutions are:

$$\begin{aligned}\widehat{\boldsymbol{\lambda}}_n &= \operatorname*{argmin}_{\boldsymbol{\lambda}} \frac{1}{n}\sum_{i=1}^{n}\exp(\boldsymbol{g}_i^\top\boldsymbol{\lambda} + \ell_i),\\ \boldsymbol{\lambda}^\star &= \operatorname*{argmin}_{\boldsymbol{\lambda}} \mathbb{E}\left[\exp(\mathbf{g}^\top\boldsymbol{\lambda} + \ell)\right].\end{aligned}$$

The estimator is an M-estimator and given our sample-splitting, the proof follows the asymptotic normality of the estimator (Van der Vaart, 2000, Chapter 5.3).

$$\sqrt{n}\left(\widehat{\boldsymbol{\lambda}}_n - \boldsymbol{\lambda}^\star\right) \xrightarrow{d} \mathcal{N}(\mathbf{0},\boldsymbol{V}), \tag{13}$$

To obtain the value of $\boldsymbol{V_1}$, note that the optimal sample solution occurs at the solution of the following equation (Z-estimator equation):

$$\sum_{i=1}^{n}\boldsymbol{g}_i\exp\left(\boldsymbol{g}_i^\top\widehat{\boldsymbol{\lambda}}_n\right) = \mathbf{0}.$$

Thus, the score function is $\boldsymbol{\psi}_{\boldsymbol{\lambda}} = \boldsymbol{g}_i \exp\left(\boldsymbol{g}_i^\top \boldsymbol{\lambda}\right)$. We denote the matrix of derivatives of the score function by $\dot{\boldsymbol{\psi}}_{\boldsymbol{\lambda}}$ whose elements are defined as $\dot{\psi}_{\boldsymbol{\lambda},kk'} = {}^{\partial \psi_{\boldsymbol{\lambda},k}}/_{\partial \lambda_{k'}}$. Using the theorem in (Van der Vaart, 2000, Chapter 5.3), we can write:

$$\boldsymbol{V} = \mathbb{E}[\dot{\boldsymbol{\psi}}_{\boldsymbol{\lambda}^\star}]^{-1} \mathbb{E}[\boldsymbol{\psi}_{\boldsymbol{\lambda}^\star} \boldsymbol{\psi}_{\boldsymbol{\lambda}^\star}^\top] \mathbb{E}[\dot{\boldsymbol{\psi}}_{\boldsymbol{\lambda}^\star}]^{-1}. \tag{14}$$

In the above equation we have assumed that $\mathbb{E}[\dot{\boldsymbol{\psi}}_{\boldsymbol{\lambda}^\star}]$ matrix is invertable. An unbiased sample estimation of $\boldsymbol{V}$ can be obtained by substituting $\widehat{\boldsymbol{\lambda}}_n$ in place of $\boldsymbol{\lambda}^\star$ and taking empirical expectations.

An application of the delta method on Eq. (13) yields:

$$\sqrt{n}\left(\frac{\exp\left(\boldsymbol{g}_i^\top \widehat{\boldsymbol{\lambda}}_n + \ell_i\right)}{\frac{1}{n}\sum_{i=1}^n \exp\left(\boldsymbol{g}_i^\top \widehat{\boldsymbol{\lambda}}_n + \ell_i\right)} - \frac{\exp\left(\boldsymbol{g}_i^\top \boldsymbol{\lambda}^\star + \ell_i\right)}{\frac{1}{n}\sum_{i=1}^n \exp\left(\boldsymbol{g}_i^\top \boldsymbol{\lambda}^\star + \ell_i\right)}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \sigma^2), \tag{15}$$

$$\sqrt{n}\left(\widehat{w}_n(a_i, \boldsymbol{x}_i) - \frac{\exp\left(\boldsymbol{g}_i^\top \boldsymbol{\lambda}^\star + \ell_i\right)}{\mathbb{E}[\exp(\mathbf{g}^\top \boldsymbol{\lambda}^\star + \lambda)]}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \sigma^2), \tag{16}$$

$$\sqrt{n}\left(\widehat{w}_n(a_i, \boldsymbol{x}_i) - w^\star(a_i, \boldsymbol{x}_i)\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \sigma^2), \tag{17}$$

where Eq. (16) is due to Slutsky's theorem and Eq. (17) is obtained by substitution of the definition for $w^\star(a_i, \boldsymbol{x}_i)$. The variance is obtained by defining the Softmax function $s(\boldsymbol{\lambda}) = \frac{\exp\left(\boldsymbol{g}_i^\top \boldsymbol{\lambda} + \ell_i\right)}{\frac{1}{n}\sum_{i=1}^n \exp\left(\boldsymbol{g}_i^\top \boldsymbol{\lambda} + \ell_i\right)}$. We denote the gradient of the Softmax function by $\nabla s(\boldsymbol{\lambda})$. We can write (Van der Vaart, 2000, Chapter 3):

$$\sigma^2(a_i, \boldsymbol{x}_i) = \nabla s(\boldsymbol{\lambda}^\star)^\top \, \boldsymbol{V} \, \nabla s(\boldsymbol{\lambda}^\star).$$

Substituting the value of $\boldsymbol{V}$ from (14), we conclude:

$$\boxed{\sigma^2(a_i, \boldsymbol{x}_i) = \nabla s(\boldsymbol{\lambda}^\star)^\top \, \mathbb{E}[\dot{\boldsymbol{\psi}}_{\boldsymbol{\lambda}^\star}]^{-1} \mathbb{E}[\boldsymbol{\psi}_{\boldsymbol{\lambda}^\star} \boldsymbol{\psi}_{\boldsymbol{\lambda}^\star}^\top] \mathbb{E}[\dot{\boldsymbol{\psi}}_{\boldsymbol{\lambda}^\star}]^{-1} \, \nabla s(\boldsymbol{\lambda}^\star).}$$

Note that the value of the softmax function depends on the value of $(a_i, \boldsymbol{x}_i)$ at each point. □

B Neural Network and Training Details

B.1 Details of the ℓ_θ Neural Network

The $\ell_{\boldsymbol{\theta}}$ network is defined as follows:

$$\ell_{\boldsymbol{\theta}}(z) = cz + \text{dense3}(\,\text{elu}(\,\text{layer\_norm}(\,\text{dense2}(\,\tanh(\,\text{dense1}(\,z\,))\,))\,)\,)$$

The linear term cz acts as a skip connection. The input and output dimensions for the dense linear layers are as follows:

$$\begin{aligned}\text{dense1} &: 1 \mapsto h,\\ \text{dense2} &: h \mapsto h,\\ \text{dense3} &: h \mapsto 1,\end{aligned}$$

where h denotes the hidden dimension. Because the softmax function is invariant to the constant shifts, we do not have any bias terms for dense3 and the skip connection. dense2 also does not have the bias because of the proceeding layer normalization. The dimension h has been tuned as a hyperparameter on a validation data and set to 10.

B.2 Details of the Propensity Score Computation for IPW

We model both $f(a)$ and $f(a|\boldsymbol{x})$ as univariate normal distributions. This is the correct assumption in our synthetic data. The marginal distribution $f(a)$ is estimated by simply finding the mean and standard deviation of the observed treatment values. For the conditional distribution, we write $\mathrm{a}|\mathbf{x} \sim \mathcal{N}(\mu_a(\mathbf{x}), \sigma^2_{a|\boldsymbol{x}})$, where $\mu_a(\mathbf{x})$ is modeled using a feedforward neural network with two layers and $\sigma^2_{a|\boldsymbol{x}}$ is estimated using the residuals of the neural network predictions. The dimension of the neural network has been tuned as a hyperparameter on validation data and set to 30.

B.3 FURTHER TRAINING DETAILS

We used PyTorch to implement E2B. For reproducibility purposes, we provide the final settings used for training:

- Learning algorithm: Adam with learning rate 0.001, no AMSGrad.
- Batch size: 100
- Max epochs: 400
- Weight decay: 2.5×10^{-5}.
- Validation on a dataset of size 400, every 10 steps.

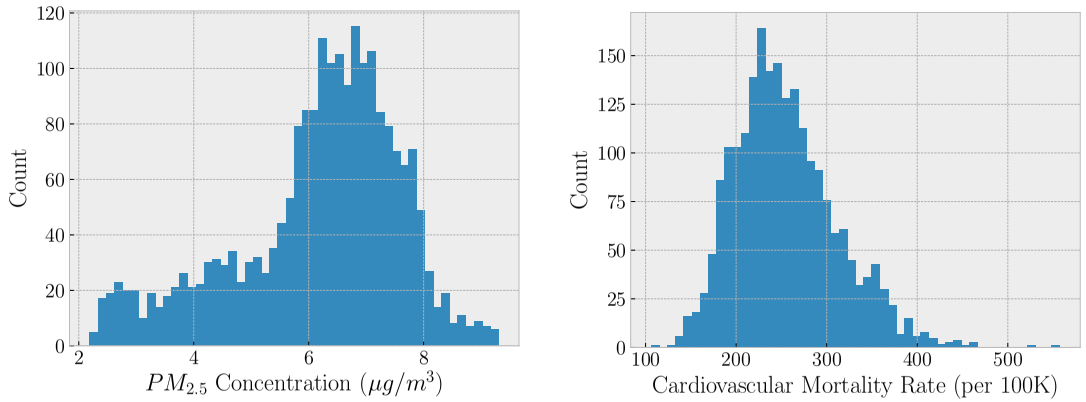
B.4 DETAIL OF PERMUTATION WEIGHTING

We created a stacked data by stacking $\{(\boldsymbol{x}_i, a_i, a_i \odot \boldsymbol{x}_i)\}_{i=1}^n$ and $\{(\boldsymbol{x}_i, \widetilde{a}_i, \widetilde{a}_i \odot \boldsymbol{x}_i)\}_{i=1}^n$, where $\widetilde{a}$ are permutations of the original treatments. We trained a random forest classifier to predict whether each data is from the permuted or the original set. We tried both random forests and neural networks and obtained better results with the former. We also calibrated the predicted probabilities of the classifier before computation of the weights.

C DATA AND PREPROCESSING DESCRIPTION

Table 2: NSAPH Data Description

	PM2.5	CMR	healthfac	population	ses	unemploy	HH_inc	femaleHH	vacant	owner_occ	eduattain	pctfam_pover
count	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0	2132.0
mean	6.17	255.25	0.18	10.78	0.0	7.85	10.69	11.92	14.25	71.44	35.03	11.25
std	1.45	56.76	0.5	1.26	0.96	2.83	0.24	3.94	8.71	7.76	7.07	5.2
min	2.19	106.14	-2.85	6.2	-1.84	0.0	9.91	2.1	3.8	19.3	9.4	0.0
25%	5.51	215.38	0.0	10.04	-0.67	6.0	10.54	9.3	8.8	67.7	30.4	7.6
50%	6.43	248.16	0.14	10.62	-0.14	7.6	10.67	11.2	11.65	72.7	35.4	10.55
75%	7.15	288.77	0.34	11.46	0.47	9.3	10.83	13.6	16.6	76.7	39.9	13.82
max	9.3	557.43	3.33	16.07	6.46	30.9	11.66	38.0	74.0	89.7	54.6	44.9



(a) Histogram of $PM_{2.5}$

(b) Histogram of Cardiovascular Mortality Rate

Figure 3: The Histograms of Data