We would like to thank the reviewers for their useful, detailed feedback! We will update the paper with the suggested minor revisions regarding typos and presentation improvements, and respond to individual reviewers comments below.

Reviewer #1:

1. "Originality": we are aware of no paper looking at policy evaluation in the setting we consider where we only have proxies for the confounders, but the latent model is identifiable. We cite the relevant papers on effect estimation from proxies and on policy evaluation in unidentifiable settings.
2. "Quality": In order to actually evaluate proxy methods out-of-sample, data *has to be simulated*. The supplement already includes results from a range of link functions ("more complex function of Y"). Changing the dimensions of X, Z gave similar comparative performance (see R4 pt2 for more detail). We can also add the Twins dataset from Louizos et al. [26] to the supplement but it is also simulated (simulated treatment assignment T and proxies X).
3. "Significance":
    (a) Regarding the assumption of an identified latent confounder model: we reference the rich literature on the conditions and methods for identifying such a model, and we focus on the downstream task of policy evaluation, where there is a lack of previous literature. Therefore our research should be seen as complementary to this existing work.
    (b) It is true that, in finite samples, an estimated latent model will have errors, but our algorithm uses this model only to approximate the Q matrix and this error can be made to vanish as $n \to \infty$. In particular, Thm 3 (convergence) trivially holds also if we use any $\hat{\phi} \to \phi$ in L_1 in probability by Slutsky's Theorem (we will update the theorem and proof accordingly).
    (c) All of the assumptions in Section 3.3 are assumptions about the choice of function class $\mathcal{F}$, not about the data distribution, and we show they are specified by particular choices (RKHS), so they absolutely do not limit the practical applicability of our method as none of those assumptions has to be tested/verified. One substantial assumption is specification ($\mu \in \mathcal{F}$), which is common and necessary for consistency. We can make the method nonparametric and use a universal kernel (e.g., RBF) to avoid this; a trivial corollary to Thm 3 using universal approximators (such as RBF RKHS) would give slower but specification-free $o_p(1)$ convergence (because we can get an error bound of $\epsilon + O_p(1/\sqrt{n})$ for any $\epsilon > 0$ using Thm 3 and universal L_∞ approximation of μ, where the O_p term's constant can depend on ϵ).

Reviewer #3:

1. "Cons": Our convergence theory in Sec 3 is general and easily extends easily to neural nets with weight decay, as they immediately satisfy all of the assumptions (we will update to make this more explicit). Regarding a concrete implementation of such a NN-based algorithm, that would be very involved (optimizing a GAN-like adversarial neural net objective, which is known to be challenging), and is far beyond the scope of our paper, but we think it would be a promising direction for future work.
2. "Improvements": See R1 pt2, R4 pt2. In order to actually evaluate proxy methods out-of-sample, data *has to be simulated*. We can add results from the Twins dataset from Louizos et al. [26] to the supplement with the other extra results but note that it is also simulated.

Reviewer #4:

1. Thanks for catching. This is any positive definite diagonal covariance matrix. We will update paper to define this.
2. Our pilot experiments with different dimensions of X and Z indicated similar comparative performance. We chose the dimensions for our experiments for reasons of convenience (as this setup gave us low sample complexity and allowed running many replications of interesting experiments on fewer data points). We will provide some additional numbers in the supplement. The code (on GitHub) allows anyone to try and tinker with higher dimensions. (Note there is already a plethora of additional results in supplemental Sec C.)
3. We like your suggestion on exploring the effect of the strength of relationship between Z and X, and are going to run an additional experiment (to place in supplement) to explore this effect. Concretely, we can measure strength of relationship by the fraction of X variance explained by Z, so we can just vary σ_X on line 245 and run the same experiment over an array of different relationship strengths.
4. You are correct. In our theory the only part of Assumption 1 that we actually use is that for every t, $Y(t)$ is conditionally independent of (X, T), given Z, so there is nothing prohibiting an arrow from X to T in the DAG in Figure 1. We will update Assumption 1 and Figure 1 accordingly. (Note that all our math stays exactly the same. All that will change is adding $X \to T$ in Figure 1 and rephrasing Assumption 1.)