Appendix B

Lemma 1. Let $\hat{\alpha}_1$ be the OLS estimator for $\alpha_1$ in the two-level model in (23). If the true functional form relationship between potential outcomes and the treatment assignment score is correctly specified in the model, then, $\hat{\alpha}_1$ is a consistent estimator for $\alpha_1$. Furthermore, as the number of units, $n$, increases to infinity in (23) and for fixed $m$, $\hat{\alpha}_1$ converges to a normal distribution with variance:

$$
\text{AsyVar}_R^T(\hat{\alpha}_1^T) = \frac{1}{p(1-p)(1-\rho_{TS}^2)} \left[ \frac{\sigma_t^2}{n} + \frac{\sigma_{\delta}^2}{nm} \right],
$$

where $\rho_{TS}$ is the correlation between $T_i$ and $\text{Score}_i$. A comparable expression can be obtained for the aggregated model in (8) by setting $\sigma_t^2 = 0$ and replacing $\sigma_{\delta}^2$ with $\sigma_{\eta}^2$.

Proof. Write (23) in terms of centered random variables as follows:

$$
(B.2.2) \quad \begin{align*}
\tilde{w}_{ij} &= \alpha_1 \tilde{T}_{ij} + \alpha_2 \tilde{S}_{ij} + (\tau_i^* + \delta_j^*),
\end{align*}
$$

where $w_{ij} = w_{ij} - E(w_{ij}), \quad \tilde{T}_{ij} = T_{ij} - p, \quad \tilde{S}_{ij} = \text{Score}_i - E(\text{Score}_i), \quad \tau_i^* = \tau_i - E(\tau_i)$ and $\delta_j^* = \delta_j - E(\delta_j).$ Let $\tilde{\tilde{w}}_{ij}, \tilde{\tilde{T}}_{ij}, \tilde{\tilde{S}}_{ij}, \tilde{\tau}_i$ and $\tilde{\delta}_j$ be respective empirically centered variables. If $Z_i^* = (T_i^* S_i^*)$ and $\tilde{Z}_i = (\tilde{T}_i, \tilde{S}_i)$, then the OLS estimator for the parameters in (B.2.2) is as follows:

$$
\begin{pmatrix}
\hat{\alpha}_1 \\
\hat{\alpha}_2
\end{pmatrix} = \left[ \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{Z}_i \tilde{Z}_j \right]^{-1} \left[ \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{Z}_i \tilde{\tilde{w}}_{ij} \right].
$$

Standard asymptotic arguments can be used to prove that as $n$ approaches infinity,

$$
(B.2.3) \quad \begin{align*}
\begin{pmatrix}
\hat{\alpha}_1 \\
\hat{\alpha}_2
\end{pmatrix} \xrightarrow{p} \left[ mE(Z_i^* Z_i^*) \right]^{-1} E(mZ_i^* w_{ij}^*) = \left( \begin{pmatrix}
\alpha_1 \\
\alpha_2
\end{pmatrix} \right) + \left[ mE(Z_i^* Z_i^*) \right]^{-1} E \left[ mZ_i^* (\tau_i^* + \delta_j^*) \right].
\end{align*}
$$

In this expression,

$$
(B.2.4) \quad \left[ mE(Z_i^* Z_i^*) \right]^{-1} = \begin{pmatrix}
mp(1-p) & m\sigma_{TS} \\
m\sigma_{TS} & m\sigma_{\delta}^2
\end{pmatrix}^{-1} = \begin{pmatrix}
\sigma_{\tau}^2 & -\sigma_{TS} \\
-\sigma_{TS} & \sigma_{\delta}^2
\end{pmatrix}^{-1} \left[ \begin{pmatrix}
1 \\
mp(1-p) - \sigma_{TS}^2
\end{pmatrix}
\right].
$$

and

$$
E(mZ_i^* [\tau_i^* + \delta_j^*]) = \begin{pmatrix}
m\sigma_{\tau} \\
m\sigma_{\delta}^2
\end{pmatrix},
$$

Appendix B

B-1
where $\sigma_{r\tau}$ is the covariance between $T_{iR}^{RD}$ and $\tau_i$, and $\sigma_{S_i}$ is the covariance between $Score_i$ and $\tau_i$. Note that the covariance between $Z_i^*$ and $\delta_j^*$ is zero because $T_{iR}^{RD}$ and $Score_i$, do not vary within schools. Thus, after some algebra, it can be seen that as $n$ approaches infinity,

$$
(B.2.5) \quad \hat{\alpha_i} \overset{p}{\rightarrow} \alpha_i + \left( \frac{\sigma_{S_i}^2 \sigma_{r\tau} - \sigma_{TS_i} \sigma_{S_i}}{\sigma_{S_i}^2 p(1-p) - \sigma_{TS_i}^2} \right).
$$

The second term on the right-hand-side of (B.2.5) is zero because it is the coefficient estimate on $T_{iR}^{RD}$ when $\tau_i$ is regressed on $T_{iR}^{RD}$ and $Score_i$. This conditional expectation is zero, because controlling for $Score_i$, there is no variation in treatment status. (Note that this result does not hold if the model is specified incorrectly and $\tau_i$ contains omitted score variables.) Thus, $\hat{\alpha_i}$ is asymptotically unbiased.

To obtain the asymptotic distribution of the two-level OLS estimator, we can rewrite (B.2.3) as follows:

$$
\sqrt{n} \left( \hat{\alpha}_i - \alpha_i \right) = n^{-1/2} \left[ mE(Z_i' Z_i^*) \right]^{-1} \sum_{i=1}^{m} \sum_{j=1}^{m} Z_i' (\tau_i^* + \delta_j^*) + o_p(1),
$$

where $o_p(1)$ denotes a term that converges in probability to zero. Thus, using (B.2.4), we find after some algebra that

$$
(B.2.6) \quad \sqrt{n}(\hat{\alpha}_i - \alpha_i) = n^{-1/2} \frac{1}{m \left( \sigma_{S_i}^2 p(1-p) - \sigma_{TS_i}^2 \right)} \sum_{i=1}^{m} (m\tau_i^* + \sum_{j=1}^{m} \delta_j^*) \left( \sigma_{S_i}^2 T_i^* - \sigma_{TS_i} S_i^* \right) + o_p(1).
$$

Because $E \left[ (m\tau_i^* + \sum_{j=1}^{m} \delta_j^*) \left( \sigma_{S_i}^2 T_i^* - \sigma_{TS_i} S_i^* \right) \right] = 0$, a simple application of the central limit theorem (see, for example, Rao 1973) can be used to show that $\hat{\alpha}_i$ is asymptotically normally distributed with mean zero and the following variance:

$$
(B.2.7) \quad \text{AsyVar}(\hat{\alpha}_i) = \frac{E(m\tau_i^* + \sum_{j=1}^{m} \delta_j^*)^2 E(\sigma_{S_i}^2 T_i^* - \sigma_{TS_i} S_i^*)^2}{nm^2 \left[ \sigma_{S_i}^2 p(1-p) - \sigma_{TS_i}^2 \right]^2} = \frac{\sigma_{S_i}^2 (\sigma_{r\tau}^2 + [\sigma_{S_i}^2 / m])}{m \sigma_{S_i}^2 p(1-p) - \sigma_{TS_i}^2}.
$$

The expressions in (B.2.7) and (B.2.1) are equivalent because $\sigma_{TS_i}^2 = \sigma_{S_i}^2 p(1-p) \rho_{r\tau}^2$. 