Supplementary Material for “Measurement Error in Multiple Equations: Tobin’s q and Corporate Investment, Saving, and Debt”

Karim Chalak[1] Daniel Kim
University of Virginia BI Norwegian Business School

A Supplementary Material on Identification

A.1 Restricting the Correlations among the Disturbances

We extend A_6 to A’_6 which restricts the sign and/or magnitude of the correlation r_{\eta_j,\eta_h} between \eta_j and \eta_h.

**Assumption A’_6** Disturbance Correlation Restriction: \( \xi_{jh} \leq r_{\eta_j,\eta_h} \leq \bar{\xi}_{jh} \) where \(-1 \leq \xi_{jh} \leq \bar{\xi}_{jh} \leq 1\) for \( j, h = 1, ..., p \) and \( j < h \).

In particular, provided \( \sigma^2_{\tilde{Y}_j} \sigma^2_{\tilde{Y}_h} \neq 0 \), from the proof of Corollary A.1 we have that

\[
r_{\eta_j,\eta_h} = \frac{\rho\bar{\tilde{Y}}_j\tilde{Y}_h - \bar{\tilde{W}}_j\tilde{Y}_h - \bar{\tilde{W}}_j\tilde{Y}_h}{(\rho - R^2_{\tilde{W},\tilde{Y}_j})^{1/2}(\rho - R^2_{\tilde{W},\tilde{Y}_h})^{1/2}}.
\]

A’_6 may restrict the sign of \( r_{\eta_j,\eta_h} \) as encoded by the sign of the function

\[
S_{jh}(r) \equiv r \times r_{\tilde{Y}_j,\tilde{Y}_h} - r_{\tilde{W},\tilde{Y}_j} r_{\tilde{W},\tilde{Y}_h}.
\]

Further, A’_6 may restrict the magnitude of \( r_{\eta_j,\eta_h} \) (either \( r^2_{\eta_j,\eta_h} \leq c^2 \) or \( c^2 \leq r^2_{\eta_j,\eta_h} \)) as encoded by the sign of the function

\[
M_{jh}(r; c) \equiv (r \times r_{\tilde{Y}_j,\tilde{Y}_h} - r_{\tilde{W},\tilde{Y}_j} r_{\tilde{W},\tilde{Y}_h})^2 - c^2(r - R^2_{\tilde{W},\tilde{Y}_j})(r - R^2_{\tilde{W},\tilde{Y}_h}).
\]

As shown in the proof of Corollary A.1 when \( R^2_{\tilde{Y}_j,\tilde{Y}_h} \neq 1 \), the discriminant of the quadratic function \( M_{jh}(r; c) \) is given by

\[
\Delta_{jh}(c) \equiv c^2[R^4_{\tilde{W}(.\tilde{Y}_j,\tilde{Y}_h)}(1 - R^2_{\tilde{Y}_j,\tilde{Y}_h})^2 - (1 - c^2)(R^2_{\tilde{W},\tilde{Y}_j} - R^2_{\tilde{W},\tilde{Y}_h})^2],
\]

---

[1]Karim Chalak (corresponding author), Department of Economics, University of Virginia, chalak@virginia.edu. Daniel Kim, BI Norwegian Business School, daniel.s.kim@bi.no.
and, when $R_{Y_j,Y_h}^2 \neq c^2$, the roots of $M_{jh}(\cdot; c)$ are given by

$$
\tilde{\rho}_{jh}(c) \equiv \frac{F_{jh}(c) - \Delta_{jh}(c)^\frac{1}{2}}{2(R_{Y_j,Y_h}^2 - c^2)} \quad \text{and} \quad \rho_{jh}^+(c) \equiv \frac{F_{jh}(c) + \Delta_{jh}(c)^\frac{1}{2}}{2(R_{Y_j,Y_h}^2 - c^2)}
$$

where

$$
F_{jh}(c) \equiv -R_{W,(\tilde{Y}_j,\tilde{Y}_h)}^2(1 - R_{Y_j,Y_h}^2) + (1 - c^2)(R_{W,Y_j}^2 + R_{W,Y_h}^2).
$$

Corollary A.1 uses $S_{jh}(r)$ and $M_{jh}(r; c)$ to encode the sign and magnitudes restrictions in $A'_6$ and to express the identification region for $(\rho, \delta, \beta, \Gamma)$ under $A_1$-$A'_6$.

**Corollary A.1** Under the conditions of Theorem 3.1, $A_4$, $A_5$, and $A'_6$ for $j, h = 1, \ldots, p$ with $j < h$, $(\rho, \delta, \beta, \Gamma)$ is partially identified in the sharp set

$$
\mathcal{J}^{k,\tau,c} \equiv \{(r, D(r), B(r), G(r)) : 0 \leq G(r), \frac{1}{1 + \kappa} \leq r \leq 1, \frac{\sigma_{Y_j}^2}{\sigma_{W_j}^2}(1 - \tau_j) \leq G_{jj}(r), \quad \text{and} \quad \xi_{jh} \leq \frac{G_{jh}(r)}{[G_{jj}(r)G_{hh}(r)]^\frac{1}{2}} \leq \tau_{jh} \text{ for } j, h = 1, \ldots, p \text{ and } j < h \}.
$$

Further, $\rho$ is partially identified in the sharp set

$$
\mathcal{R}^{k,\tau,c} = [R_{W,Y_j}^2, 1] \cap \prod_{j,h=1}^p \left[ \frac{1}{1 + \kappa}, 1 \right] \cap \bigcap_{j<h} \mathcal{R}_{jh}^c,
$$

with

$$
\mathcal{R}_{jh}^c = \begin{cases}
S_{jh}(r) \leq 0 \quad \text{and} \quad M_{jh}(r; \xi_{jh}) \leq 0 \leq M_{jh}(r; \bar{c}_{jh}) & \text{if } \xi_{jh} \leq \bar{c}_{jh} \leq 0 \quad \text{and} \quad \sigma_{Y_j}^2 \sigma_{Y_h}^2 \neq 0 \\
\{S_{jh}(r) \leq 0 \quad \text{and} \quad M_{jh}(r; \xi_{jh}) \leq 0 \} \quad \text{or} \quad \{0 \leq S_{jh}(r) \quad \text{and} \quad M_{jh}(r; \bar{c}_{jh}) \leq 0 \} & \text{if } \xi_{jh} < 0 < \bar{c}_{jh} \text{ and } \sigma_{Y_j}^2 \sigma_{Y_h}^2 \neq 0 \\
0 \leq S_{jh}(r) \quad \text{and} \quad M_{jh}(r; \bar{c}_{jh}) \leq 0 \leq M_{jh}(r; \xi_{jh}) & \text{if } 0 \leq \xi_{jh} \leq \bar{c}_{jh} \text{ and } \sigma_{Y_j}^2 \sigma_{Y_h}^2 \neq 0 \\
r \in \emptyset & \text{if } 0 \not\in [\xi_{jh}, \bar{c}_{jh}] \text{ and } \sigma_{Y_j}^2 \sigma_{Y_h}^2 = 0.
\end{cases}
$$

where, provided $\sigma_{Y_j}^2 \sigma_{Y_h}^2 \neq 0$, we have

$$
0 \leq S_{jh}(r) \iff \begin{cases}
\frac{r_{W,Y_j}r_{W,Y_h}}{r_{Y_j,Y_h}} \leq r & \text{when } 0 < r_{Y_j,Y_h} \\
\frac{r_{W,Y_j}r_{W,Y_h}}{r_{Y_j,Y_h}} \leq 0 & \text{when } r_{Y_j,Y_h} = 0 \\
r \leq \frac{r_{W,Y_j}r_{W,Y_h}}{r_{Y_j,Y_h}} & \text{when } r_{Y_j,Y_h} < 0
\end{cases}
$$
and if $R_{Y_j,\tilde{y}_h}^2 = 1$ then $0 \leq M_{jh}(r; c) = (1 - c^2)(r - R_{W,Y_j}^2)^2$ whereas if $R_{Y_j,\tilde{y}_h}^2 \neq 1$ then

$$0 \leq M_{jh}(r; c) \iff \left\{ \begin{array}{ll}
-\infty < r < \infty & \text{when } 0 < c^2 < 1 - \frac{R_{W,(\tilde{y}_j,\tilde{y}_h)}^4}{(R_{W,Y_j}^2 - R_{W,Y_j}^2)^2} \\
\rho_{jh}^-(c) \cup [\rho_{jh}^+(c), \infty) & \text{when } c^2 = 0 < R_{Y_j,\tilde{y}_h}^2 \text{ or } 1 - \frac{R_{W,(\tilde{y}_j,\tilde{y}_h)}^4}{(R_{W,Y_j}^2 - R_{W,Y_j}^2)^2} \leq c^2 < R_{Y_j,\tilde{y}_h}^2 \\
\rho_{jh}^-(c) & \text{when } c^2 = R_{Y_j,\tilde{y}_h}^2 \text{ and } R_{W,(\tilde{y}_j,\tilde{y}_h)}^2 < R_{W,Y_j}^2 + R_{W,\tilde{y}_h}^2 \\
0 \leq (1 - R_{Y_j,\tilde{y}_h}^2) R_{W,Y_j}^2 - R_{W,\tilde{y}_h}^2 & \text{when } c^2 = R_{Y_j,\tilde{y}_h}^2 \text{ and } R_{W,(\tilde{y}_j,\tilde{y}_h)}^2 = R_{W,Y_j}^2 + R_{W,\tilde{y}_h}^2 \\
\frac{R_{W,(\tilde{y}_j,\tilde{y}_h)}^2 - (R_{W,Y_j}^2 + R_{W,\tilde{y}_h}^2)}{R_{W,Y_j}^2 - R_{W,\tilde{y}_h}^2} \leq r & \text{when } c^2 = R_{Y_j,\tilde{y}_h}^2 \text{ and } R_{W,Y_j}^2 + R_{W,\tilde{y}_h}^2 < R_{W,(\tilde{y}_j,\tilde{y}_h)}^2 \\
\rho_{jh}(c) & \text{when } R_{Y_j,\tilde{y}_h}^2 < c^2
\end{array} \right. $$

Last, $\delta$, $\beta$, and $\Gamma$ are partially identified in the sharp sets $\mathcal{D}^{k,\tau,c} = \{D(r) : r \in \mathcal{R}^{k,\tau,c}\}$, $\mathcal{B}^{k,\tau,c} = \{B(r) : r \in \mathcal{R}^{k,\tau,c}\}$, and $\mathcal{G}^{k,\tau,c} = \{G(r) : r \in \mathcal{R}^{k,\tau,c}\}$.

The bounds in Corollary A.1 correspond to those in Corollaries 3.2 and 3.3 when $(\ell_{jh}, \bar{c}_{jh})$ is set to $(-1, 0)$, $(0, 1)$, $(0, 0)$, or $(-1, 1)$. In particular, when $R_{Y_j,\tilde{y}_h}^2 \neq 0$, the proof of Corollary A.1 gives

$$\rho_{jh}^-(0) = \rho_{jh}^+(0) = \frac{r_{W,Y_j}^2 r_{W,\tilde{y}_h}^2}{r_{Y_j,\tilde{y}_h}^2} = \frac{\sigma_{W,Y_j}^2 \sigma_{W,\tilde{y}_h}^2}{\sigma_{W,Y_j,\tilde{y}_h}^2} \sigma_{W,Y_j} \sigma_{W,\tilde{y}_h} \sigma_{W,Y_j,\tilde{y}_h},$$

so that $0 \leq M_{jh}(\rho; 0) \iff \rho \in (-\infty, \infty)$ and $M_{jh}(\rho; 0) \leq 0 \iff \rho = \frac{\sigma_{W,Y_j} \sigma_{W,\tilde{y}_h}}{\sigma_{W,Y_j,\tilde{y}_h}}$. Also,

$$\rho_{jh}^-(1) = \rho_{jh}^+(1) = R_{W,(\tilde{y}_j,\tilde{y}_h)}^2 \quad \text{and} \quad \rho_{jh}^+(1) = \rho_{jh}^-(1) = 0.$$

Thus, when $R_{Y_j,\tilde{y}_h}^2 < 1$, $M_{jh}(\rho; 1) = M_{jh}(\rho; -1) \leq 0 \iff \rho \in (-\infty, 0] \cup [R_{W,(\tilde{y}_j,\tilde{y}_h)}^2, \infty)$. Since $R_{W,(\tilde{y}_j,\tilde{y}_h)}^2 \leq R_{W,Y_j}^2$, this magnitude restriction is not binding in $\mathcal{R}^{k,\tau,c}$. It follows that Corollary A.1 yields the same bound $\mathcal{R}^{k,\tau,c}$ from Corollary 3.3, with $\mathcal{R}_{jh}^{c}$ determined by the magnitude restriction encoded in $M_{jh}(\rho; 0) \leq 0$ when $\ell_{jh} = \bar{c}_{jh} = 0$ and by the sign restrictions, if any, encoded in $S_{jh}(r)$ otherwise.

A.2 Relaxing the Classical Measurement Error Assumption

Assume $A_1$ and $A_2$ and consider removing the classical measurement error assumption $A_3$ to allow $\varepsilon$ to be correlated with $U$ or $\eta$:

$$Y' = U\delta + \eta \quad \text{and} \quad W = U + \varepsilon \quad \text{where} \quad \text{Cov}(U, \eta) = 0.$$
Here, we dispense with $X$ for simplicity - if $\text{Cov}[X, (\varepsilon, \eta)'] = 0$ then the analysis proceeds analogously after projecting on $X$. We can express the moments in $\text{Var}[(Y', W')]$ by

$$\sigma_{W}^2 = \sigma_{U}^2 + \sigma_{\varepsilon}^2 + 2\sigma_{U, \varepsilon}, \quad \sigma_{W,Y} = \sigma_{U}^2 \delta + \sigma_{U, \varepsilon} \delta + \sigma_{\eta, \varepsilon}, \quad \text{and} \quad \sigma_{Y}^2 = \delta' \sigma_{U}^2 \delta + \sigma_{\eta}^2.$$ 

Let $W$ and $U$ be nondegenerate. Dividing the first equation by $0 < \sigma_{W}^2$ gives

$$1 = \rho_u + \rho_{\varepsilon} + 2\frac{\sigma_{U, \varepsilon}}{\sigma_{W}^2} \quad \text{where} \quad \rho_u \equiv \frac{\sigma_{U}^2}{\sigma_{W}^2} \text{ and } \rho_{\varepsilon} \equiv \frac{\sigma_{\varepsilon}^2}{\sigma_{W}^2}.$$ 

Further, normalizing the second and third moments by $\sigma_{W}^2$, defining $\zeta \equiv \frac{\sigma_{U, \varepsilon}}{\sigma_{W}^2}$ and $\Gamma \equiv \frac{\sigma_{\eta}^2}{\sigma_{W}^2}$, and substituting for $\frac{\sigma_{U, \varepsilon}}{\sigma_{W}^2} = \frac{1}{2}(1 - \rho_u - \rho_{\varepsilon})$ gives the nonlinear system of equations

$$b_{Y,W} \equiv \frac{\sigma_{W,Y}}{\sigma_{W}^2} = \rho_u \delta + \frac{1}{2}(1 - \rho_u - \rho_{\varepsilon}) \delta + \zeta \quad \text{and} \quad \frac{\sigma_{Y}^2}{\sigma_{W}^2} = \delta' \rho_u \delta + \Gamma,$$

where the dimension of the unknowns ($\rho_u, \rho_{\varepsilon}, \zeta, \delta, \Gamma$) exceeds the number of equations. The system’s unknowns must also obey that $\text{Var}[(U, \varepsilon, \eta')]$ is positive semi-definite. Nevertheless, without additional assumptions, these restrictions do not identify the elements of $\delta$. This holds even if $\eta$ and $\varepsilon$ are assumed to be uncorrelated so that $\zeta \equiv \frac{\sigma_{U, \varepsilon}}{\sigma_{W}^2} = 0$.

In particular, if $\zeta \equiv \frac{\sigma_{U, \varepsilon}}{\sigma_{W}^2} = 0$ then (we let $1 + \rho_u \neq \rho_{\varepsilon}$ - otherwise, $b_{Y,W} = \frac{1}{2}(1 + \rho_u - \rho_{\varepsilon}) \delta = 0$ does not help identify $\delta$)

$$\delta = D(\rho_u, \rho_{\varepsilon}) \equiv \frac{2}{1 + \rho_u - \rho_{\varepsilon}} b_{Y,W} \quad \text{and} \quad \Gamma = G(\rho_u, \rho_{\varepsilon}) \equiv \frac{\sigma_{Y}^2}{\sigma_{W}^2} - \frac{4\rho_u}{(1 + \rho_u - \rho_{\varepsilon})^2} b'_{Y,W} b_{Y,W}.$$ 

Further, since $\sigma_{\eta, \varepsilon} = 0$, $\text{Var}[(U, \varepsilon, \eta')]$ is block-diagonal and therefore $0 \preceq \text{Var}[(U, \varepsilon, \eta')]$ if and only if $0 \preceq \text{Var}[(U, \varepsilon)']$ and $0 \preceq \text{Var}[(\eta)]$. The first constraint $0 \preceq \text{Var}[(U, \varepsilon)']$ holds if and only if (where we use $\sigma_{U, \varepsilon}^2 \leq \sigma_{U}^2 \sigma_{\varepsilon}^2$)

$$0 \leq \rho_u, \quad 0 \leq \rho_{\varepsilon}, \quad \text{and} \quad \frac{1}{4}(1 - \rho_u - \rho_{\varepsilon})^2 \leq \rho_u \rho_{\varepsilon}.$$ 

Further, since (recall that we let $0 < \rho_u$)

$$0 \leq \frac{(1 + \rho_u - \rho_{\varepsilon})^2}{4\rho_u} = \frac{(1 - \rho_u - \rho_{\varepsilon})^2}{4\rho_u} + \frac{4(1 - \rho_{\varepsilon})\rho_u}{4\rho_u} \leq \frac{4\rho_u \rho_{\varepsilon}}{4\rho_u} + \frac{4(1 - \rho_{\varepsilon})\rho_u}{4\rho_u} = 1,$$

it follows from the proof of Corollary 3.3 that the second constraint $\Gamma \equiv \frac{\sigma_{\eta}^2}{\sigma_{W}^2} \succeq 0$ holds if and only if

$$R_{W,Y}^2 \leq \frac{(1 + \rho_u - \rho_{\varepsilon})^2}{4\rho_u} = \left(\frac{1 + \rho_u - \rho_{\varepsilon}}{2}\right)^2 \frac{1}{\rho_u} \leq 1.$$
Clearly, $b_{y,W} = 0$ if and only if $\delta_j = 0$. However, more generally, these constraints fail to bound $\delta_j$. In particular, for any value $d \in \mathbb{R}\setminus\{0\}$ for $\delta_j$ there exists a value $(r_u, r_\varepsilon)$ for $(\rho_u, \rho_\varepsilon)$ such that $d = D_j(r_u, r_\varepsilon)$ and the constraints on $(\rho_u, \rho_\varepsilon)$ in $0 \leq \text{Var}([U, \varepsilon])$ hold. Specifically, let $r \equiv \frac{1}{d}b_{y,W}$ and set $(r_u, r_\varepsilon)$ such that $r^2 \leq r_u \leq \frac{r^2}{R_{W,Y}}$ and $r_\varepsilon = 1 + r_u - 2r$. Then $D_j(r_u, r_\varepsilon) = \frac{2}{1+r_u-r_\varepsilon}b_{y,W} = \frac{1}{r}b_{y,W} = d$. Further, the $0 \leq \text{Var}([U, \varepsilon])$ constraint holds since $0 < r^2 \leq r_u$ and

$$
\begin{align*}
\delta_u = 1 + r_u - 2r \geq 1 + r^2 - 2r = (1-r)^2 \geq 0, \\
\delta_u - \frac{1}{4}(1-r_u - \delta_u) = r_u(1 + r_u - 2r) - \frac{1}{4}(1 - r_u - (1 + r_u - 2r))^2 = r_u - r^2 \geq 0.
\end{align*}
$$

Last, the $\Gamma \equiv \frac{\sigma^2}{\sigma^2_{W,Y}} \geq 0$ constraint holds since $R_{W,Y}^2 \leq \frac{r^2}{r_u} = \left(\frac{1 + r_u - r_\varepsilon}{2}\right)^2 \frac{1}{r_u} \leq 1$.

This paper’s analysis maintains the classical measurement error assumption $A_3$. In this case, $\zeta \equiv \frac{\sigma_{u,\varepsilon}}{\sigma^2_W} = 0$ and $\frac{\sigma_{u,\varepsilon}}{\sigma^2_W} = 0$ and thus $\rho_u = 1 - \rho_u$. This reduces the dimension of the unknown parameters in the equations for $\frac{\sigma^2_{W,Y}}{\sigma^2_W}$ and $\frac{\sigma^2}{\sigma^2_W}$ by $p + 1$, from $(\rho_u, \rho_\varepsilon, \zeta, \delta, \Gamma)$ to $(\rho_u, \delta, \Gamma)$, and yields two-sided bounds for the elements of $\delta$. It is of interest to derive analytical expressions for the sharp identification regions for $(\rho_u, \rho_\varepsilon, \zeta, \delta, \Gamma, \beta)$ without $A_3$ under restrictions analogous to $A_4$-$A_6$. To keep the scope of the paper manageable, we leave tackling this problem in more detail to other work.

## B Supplementary Material on Inference

### B.1 Algorithm for Inference on $\rho$

In order to apply only one algorithm that delivers $\hat{\rho}_u(\lambda; 1-\alpha_{21})$, $\hat{\rho}_u(\lambda; 1-\alpha_{21})$, and $CI_{1-\alpha_{21}}^p(\lambda)$, it is useful to adopt the following notation. For $r \in [0, 1]$, we let

$$
g^v(\pi; r, \lambda) = (g^v_1(\pi; r, \lambda), \ldots, g^v_M(\pi; r, \lambda)) \quad \text{where} \quad g^v_1(\pi; r, \lambda) \equiv r - \rho^v_u(\lambda) \quad \text{for} \quad v = 1, \ldots, M, \quad \text{and}$$

$$
g^u(\pi; r, \lambda) = (g^u_1(\pi; r, \lambda), \ldots, g^u_M(\pi; r, \lambda)) \quad \text{where} \quad g^u_1(\pi; r, \lambda) \equiv \rho^u_u(\lambda) - r \quad \text{for} \quad v = 1, \ldots, M.
$$

Thus, $\rho_u^v(\lambda) = -g^v_0(\pi; 0, \lambda)$ and $\rho_u^0(\lambda) = g^u_0(\pi; 0, \lambda)$. Further, we collect all the lower and upper bounds, denoted by $g^v(\pi; r, \lambda)$ for $v = 1, \ldots, 2M$, into

$$
g^v(\pi; r, \lambda) = (g^v_1(\pi; r, \lambda)', g^u_1(\pi; r, \lambda)')'.
$$

$^2$We employ $g^v(\pi; 0, \lambda)$ to transform a lower bound for $\rho$ into an upper bounds for $-\rho$. We then use a single algorithm (for an upper bound) when estimating the lower and upper bounds for $\rho$. 

5
We estimate \(g^e(\pi; r, \lambda)\) using the consistent plug-in estimator \(g^e(\hat{\pi}; r, \lambda)\). Using the delta method, the linearly independent subset \(g^e(\hat{\pi}; r, \lambda)\) of \(g^e(\hat{\pi}; r, \lambda)\) (recall that some of bounds in \(g^e(\pi; r, \lambda)\) are constant or linearly dependent, e.g. in the single equation case or under the diagonal variance restriction in \(A_6\)) is asymptotically normally distributed:

\[
\sqrt{n}(g^e(\hat{\pi}; r, \lambda) - g^e(\pi; r, \lambda)) \xrightarrow{d} N(0, \nabla_{\pi}g^e(\pi; r, \lambda)\Sigma_{\pi}g^e(\pi; r, \lambda)')
\]

Note that \(\nabla_{\pi}g^e(\pi; r, \lambda)\) does not depend on \(r\). Section B.2 collects the expressions for \(g^e(\pi; r, \lambda)\), and \(\nabla_{\pi}g^e(\pi; r, \lambda)\).

Next, for each \(\ell \in \Lambda_{1-\alpha,2}\), we implement algorithm 1 in Chernozhukov, Lee, and Rosen (2013). To compute, \(C\widetilde{I}_{1-\alpha,1}(\ell)\), we invert a test statistic and perform a grid search over \((0, 1]\). For a thorough discussion of the algorithm\(^3\) we refer the reader to Chernozhukov, Lee, and Rosen (2013) and Chernozhukov, Kim, Lee, and Rosen (2015).

1. Let \(\alpha \leq \frac{1}{2}\) and \(V^c \equiv V^d \cup V^u \equiv \{1, ..., M\} \cup \{M+1, ..., 2M\}\).

   If the target output is:

   (a) \(\hat{\rho}_o^d(\ell; 1-\alpha)\) or \(\hat{\rho}_o^u(\ell; 1-\alpha)\) then set \(m = l\) or \(u\) and \(r = 0\).

   (b) \(C\widetilde{I}_{1-\alpha}(\ell)\) then set \(m = c\) and \(r \in (0, 1]\).

2. Set \(\tilde{\gamma} = 1 - \frac{0.1}{\log n}\). Simulate \(S\) draws \(Z_1, ..., Z_S\) from \(N(0, I_{2M})\).

3. For each \(v \in V^c\), compute\(^4\) \(\hat{h}(v; \ell) = [\mathbf{1}(v = 1), ..., \mathbf{1}(v = 2M)][\nabla_{\pi}g^e(\hat{\pi}; r, \ell)\Sigma_{\pi}g^e(\hat{\pi}; r, \ell)']^{1/2}\) and set \(se(v; \ell) = \frac{1}{\sqrt{n}} \|\hat{h}(v; \ell)\|\).

4. Define \(V^m_+ = \{v \in V^m : se(v; \ell) \neq 0\}\). Compute

\[
c_{V^m}(\tilde{\gamma}; \ell) = \tilde{\gamma}\text{-quantile of } \left\{ \sup_{v \in V^m_+} \frac{\hat{h}(v; \ell)Z_s}{\|\hat{h}(v; \ell)\|}, s = 1, ..., S \right\}
\]

and

\[
\hat{V}^m = \{v \in V^m_+ : g^m_v(\hat{\pi}; 0, \ell) \leq \min_{v \in V^m_+} [g^m_v(\hat{\pi}; 0, \ell) + c_{V^m}(\tilde{\gamma}; \ell)se(v; \ell)] + 2c_{V^m}(\tilde{\gamma}; \ell)se(v; \ell)\}.
\]

\(^3\)We adjust the algorithm in Chernozhukov, Lee, and Rosen (2013) slightly since some of our bounds are deterministic (e.g. \(\rho \leq 1\)). Specifically, we use the estimated bounds to calculate the critical value. Then we report the smallest upper bound among the precision-corrected estimators and the deterministic bounds.

\(^4\)\(\nabla_{\pi}g^e(\hat{\pi}; r, \ell)\Sigma_{\pi}g^e(\hat{\pi}; r, \ell)’\) may be positive semi-definite and its matrix square root is computed using a singular value decomposition.
Recall that the nuisance parameters \( \lambda \) of \( \beta \) once in \( \Delta \) method
\( \ell \) depends only on the signs (negative, zero, or positive) of the first \( T \) components
\( \rho \). From Corollary 3.2 we have that if \( \ell \) is such that \( (c_{jh}, \bar{c}_{jh}) \neq (-1, 1), \ sgn(-\ell_{T+t}) \notin \{c_{jh}, \bar{c}_{jh}\}, \) and \( \ell_t=0 \) then \( R_{s_{jh}}(\ell)=\emptyset. \) As such, we drop these \( \ell \) values from \( \Lambda_{1-\alpha_2} \) since
they have no effect on \( CI_{1-\alpha}^{\rho}(\ell) = \{r \in (0, 1] : g_0^c(\hat{\pi}; r, \ell) \geq 0 \}. \)

In the single equation bounds or when \( A_6 \) is not in force, the value \( \ell \) of the nuisance
parameters does not affect the bounds. Otherwise, let \( t = 1, \ldots, T \) enumerate the \( T = \frac{1}{2}p(p-1) \) \( (j_t, h_t) \) pairs, \( j_t, h_t = 1, \ldots, p \) with \( j_t < h_t \), that correspond to the first \( T \) components
of \( \lambda \). From Corollary 3.2 we have that if \( \ell \) is such that \( (c_{jh}, \bar{c}_{jh}) \neq (-1, 1), \ sgn(-\ell_{T+t}) \notin \{c_{jh}, \bar{c}_{jh}\}, \) and \( \ell_t=0 \) then \( R_{s_{jh}}(\ell)=\emptyset. \) As such, we drop these \( \ell \) values from \( \Lambda_{1-\alpha_2} \) since
they have no effect on \( CI_{1-\alpha_1}^{\rho}(\ell) = \{r \in (0, 1] : g_0^c(\hat{\pi}; r, \ell) \geq 0 \}. \)

Otherwise, if \( m = c \) then report
\( CI_{1-\alpha}^{\rho}(\ell) = \{r \in (0, 1] : g_0^c(\hat{\pi}; r, \ell) \geq 0 \}. \)

B.2 Delta Method

Recall that the nuisance parameters \( \lambda \equiv G^\lambda(\pi) \), the vector of lower and upper bounds
\( g_0^c(\pi; r, \lambda) \) in the intersection bounds algorithm for inference on \( \rho \), and the parameters \( \delta_j, \beta_j, \) and \( \Gamma_{jk}, \ j, h = 1, \ldots, p \) and \( l = 1, \ldots, k \), (written in the form \( \theta \equiv H(\pi; \rho) \)) can all be expressed as functions of the vector of estimands

\[
\pi'_{1B} \equiv \left( \pi_1', \pi_2', \pi_3', \pi_4', \pi_5', \pi_6', \pi_7' \right)_{1 \times p(k+1)} \left( 1 \times (p+k) \right)_{1 \times (p+k)} \left( 1 \times k \right)_{1 \times p} \left( 1 \times 1 \right)_{1 \times \frac{1}{2}p(p-1)}
\]

\[
= \left[ vec(b_{Y,W}(Y,X))', b_{W'Y}(Y,X)', b_{W,Y}(Y,W)' \right. \ldots b_{W,Y}(Y,W)' \ldots b_{W,Y}(Y,W)' , vec(b_{Y,X})', b_{W,Y}' \sigma_{W}^{-2}(\sigma_{Y_1}^2, \ldots, \sigma_{Y_p}^2), \sigma_{W}^{-2}(\sigma_{Y_1}^2, \ldots, \sigma_{Y_p}^2). \right]
\]
Since the plug-in estimator \( \hat{\pi} \) satisfies \( \sqrt{n}(\hat{\pi} - \pi) \xrightarrow{d} N(0, \Sigma) \), the delta method gives

\[
\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, \nabla_\pi g^\lambda(\pi) \Sigma \nabla_\pi g^\lambda(\pi)'),
\]

\[
\sqrt{n}(g_c^*(\hat{\pi}; r, \lambda) - g_c^*(\pi; r, \lambda)) \xrightarrow{d} N(0, \nabla_\pi g_c^*(\pi; r, \lambda) \Sigma \nabla_\pi g_c^*(\pi; r, \lambda)'),
\]

\[
\sqrt{n}(H(\hat{\pi}; r) - H(\pi; r)) \xrightarrow{d} N(0, \nabla_\pi H(\pi; r) \Sigma \nabla_\pi H(\pi; r)'),
\]

for any \( r \in (0, 1] \). In what follows, we provide expressions for \( g^\lambda, \nabla_\pi g^\lambda(\pi), g_c^*(\pi; r, \lambda), \nabla_\pi g_c^*(\pi; r, \lambda), H(\pi; r) \) and \( \nabla_\pi H(\pi; r) \).

**B.2.1 Nuisance Parameters**

The \( 2T = p(p - 1) \) nuisance parameters are collected in

\[
\lambda = (\lambda_1, \cdots, \lambda_{2T})' = g^\lambda(\pi) \equiv (\sigma_W^{-1}\sigma_{\bar{Y}_1, \bar{Y}_2}, \cdots, \sigma_W^{-2}\sigma_{\bar{Y}_p, \bar{Y}_2, \bar{Y}_W}, \cdots, \sigma_W^{-1}\sigma_{\bar{Y}_{p-1}, \bar{Y}_W}, b_{\bar{Y}_1}, b_{\bar{Y}_2}, \cdots, b_{\bar{Y}_{p-1}}, b_{\bar{Y}_W})'.
\]

It follows that, for \( t = 1, \ldots, T \), the components of \( \nabla_\pi g^\lambda(\pi) \) are given by

\[
\nabla_\pi g^\lambda_t(\pi) = \left[ \begin{array}{c} 0 \\ 1 \times [p(k+1)+(p+k)+p k+k+p] \\ \frac{i_t'}{2 p(p-1)} \end{array} \right],
\]

where \( i_t \) is the unit vector with 1 in the \( t \)th position and 0 elsewhere, and for \( t = T + 1, \ldots, 2T \)

\[
\nabla_\pi g^\lambda_{2T+t}(\pi) = \left[ \begin{array}{c} 0 \\ 1 \times (p+k) \cdot 2 p(p-1) \end{array} \right] + \left[ \begin{array}{c} \bar{b}_{\bar{Y}_{2T+t}} \bar{W} \\ 0 \\ 1 \times 2 p(p-1) \end{array} \right].
\]

**B.2.2 Lower and Upper Intersection bounds**

Consider the joint equation bounds with \( \lambda = \ell^* \) with \( (\ell_{j, t}, \ell_{j, t}^*) \in \{(-1, 0), (0, 1)\} \) and \( \text{sgn}(\ell_t^*) \in [\ell_{j, t}, \ell_{j, t}^*]\{0\} \) for \( t = 1, \ldots, T \). In this case, we have \( g^\ell(\pi; r, \lambda) = (g^\ell(\pi; r, \lambda)', g^{\ell^*}(\pi; r, \lambda)')' \) with

\[
g^\ell(\pi; r, \ell^*) = \left[ \begin{array}{c} r - R_W^2 \bar{Y} \\ r - \frac{R_W^2}{\bar{Y}_1} \\ r - \frac{R_W^2}{\bar{Y}_2} \bar{Y}_1 \\ \vdots \\ r - \frac{R_W^2}{\bar{Y}_p} \bar{Y}_{p-1} \\ r - \frac{b_{\bar{Y}_1} b_{\bar{Y}_2} \bar{W}}{\sigma_W^2 \sigma_{\bar{Y}_1, \bar{Y}_2}} \\ \vdots \\ r - \frac{b_{\bar{Y}_{p-1}} b_{\bar{Y}_p} \bar{W}}{\sigma_W^2 \sigma_{\bar{Y}_{p-1}, \bar{Y}_p}} \end{array} \right] \quad \text{and} \quad g^{\ell^*}(\pi; r, \ell^*) = \left[ \begin{array}{c} 1 - r \\ 1 - r \\ 1 - r \\ \vdots \\ 1 - r \\ \infty \\ \infty \end{array} \right]
\]
where \( M \equiv 2 + p + T \) (recall \( T \equiv \frac{1}{2}p(p - 1) \)) and

\[
R_{W,Y}^2 = b_{Y,W} b_{W,Y} = \sum_{h=1}^{p} b_{Y_h,(W,X')_h,b_{W,(Y,X')_h}} \quad \text{and} \quad R_{W,Y}^2 = b_{Y,W} b_{W,Y} = b_{Y_j,(W,X')_j,b_{W,(Y,X')_j}}.
\]

The components \( \nabla g^c(\pi; r, \ell^*) \) are then given by

1. For \( v = 1 \)

\[
\nabla g^c_{1}(\pi; r, \ell^*) = \left[ \sum_{h=1}^{p} v_h \otimes \left[ \begin{array}{c} -b_{W,(Y,X')_h} \ 0 \\ 0 \end{array} \right] \right] = \left[ \begin{array}{c} -b_{Y,W} \ 0 \\ 0 \end{array} \right]
\]

2. for \( v = 2 \),

\[
\nabla g^c_{2}(\pi; r, \ell^*) = 0,
\]

3. for \( v = 2 + j \) and \( j = 1, \ldots, p \)

\[
\nabla g^c_{v}(\pi; r, \ell^*) = \left[ v_j \otimes \left[ \begin{array}{c} -b_{W,Y} \ 0 \\ 0 \end{array} \right] \right] = \left[ \begin{array}{c} -b_{Y,W} \ 0 \\ 0 \end{array} \right]
\]

4. for \( v = 2 + p + t \) and \( t = 1, \ldots, T \),

\[
\nabla g^c_{v}(\pi; r, \ell^*) = \left[ \begin{array}{c} \nabla g^c_{1}(\pi; r, \ell^*) \ 0 \end{array} \right] = \left[ \begin{array}{c} \nabla g^c_{1}(\pi; r, \ell^*) \ 0 \end{array} \right],
\]

where

\[
\nabla g^c_{1}(\pi; r, \ell^*) = \nabla g^c_{1}(r - \frac{b_{Y_h,W} b_{Y_h,W}}{\sigma^2_W \sigma_{Y_h,Y_h}}) = (\sigma^2_W \sigma_{Y_h,Y_h})^{-1} \left[ \begin{array}{c} v_j \otimes \left[ \begin{array}{c} -b_{Y_h,W} \ 0 \\ 0 \end{array} \right] \right] + \left[ \begin{array}{c} v_j \otimes \left[ \begin{array}{c} -b_{Y_h,W} \ 0 \\ 0 \end{array} \right] \right]
\]

and

\[
\nabla g^c_{2}(\pi; r, \ell^*) = \nabla g^c_{2}(r - \frac{b_{Y_h,W} b_{Y_h,W}}{\sigma^2_W \sigma_{Y_h,Y_h}}) = \left[ v_j \otimes \left[ \begin{array}{c} -b_{Y_h,W} \ 0 \\ 0 \end{array} \right] \right]
\]

5. for \( v = 2 + p + T + 1, \ldots, 2(2 + p + T) \)

\[
\nabla g^c_{v}(\pi; r, \ell^*) = 0
\]

Above, we set \( \lambda = \ell^* \) where \( (\xi_{ij}, \tau_{ij}) \in \{(-1,0),(0,1)\} \) and \( \text{sgn}(\ell^*) \in \{\xi_{ij}, \tau_{ij}\}\} \) for \( t = 1, \ldots, T \equiv \frac{1}{2}p(p - 1) \). More generally, we consider an arbitrary \( \ell \in \Lambda_{1-\alpha_2} \) and define the matrix \( P_{2M \times 2M} \) \( (M \equiv 2 + p + T) \) to operationalize how the nuisance parameters \( \lambda \)
determines whether $\mathcal{R}_{jh}^c$ contains an upper bound, lower bound, neither, or both according to Corollary 3.3 (recall that we have already dropped from $\Lambda_{1-\alpha_2}$ the values $\ell$ such that $\mathcal{R}_{jh}^c(\ell) = \emptyset$). In particular, let

$$g^c(\pi; \ell, r) = Pg^c(\pi; \ell^*, r) \quad \text{and} \quad \nabla_\pi g^c(\pi; \ell, r) = P\nabla_\pi g^c(\pi; \ell^*, r)$$

where we set the $i^{th}$ row $P_v$ of $P$ as follows, for $t = 1, ..., \frac{1}{2}p(p - 1)$:

1. Set $P = I_{2M \times 2M}$.
2. If $(\xi_{jh}, \overline{\xi}_{jh}) = (0, 0)$ and $\ell_t \neq 0$ then change $P_{M+(2+p+t)}$ to $-i_{2+p+t}$.
3. If $(\xi_{jh}, \overline{\xi}_{jh}) \in \{(-1, 0), (0, 1)\}$ and $\text{sgn}(\ell_t) \not\in [\xi_{jh}, \overline{\xi}_{jh}]$ then change (a) $P_{2+p+t}$ to $i_{M+(2+p+t)}$ and (b) $P_{M+(2+p+t)}$ to $-i_{2+p+t}$.
4. If $(\xi_{jh}, \overline{\xi}_{jh}) \in \{(-1, 0), (0, 1)\}$ and $\text{sgn}(\ell_t) \in [\xi_{jh}, \overline{\xi}_{jh}] \setminus \{0\}$ then keep (a) $P_{2+p+t}$ as $i_{2+p+t}$ and (b) $P_{M+(2+p+t)}$ as $i_{M+(2+p+t)}$.
5. Otherwise, change $P_{2+p+t}$ to $i_{M+(2+p+t)}$.

Moreover, for the $j^{th}$ single equation bounds, $P$ mutes the irrelevant bounds as follows:

1. Change $P_1$ to $i_{M+(2+p+1)}$
2. For $h = 1, ..., p$, if $h \neq j$ then change $P_{2+h}$ to $i_{M+(2+p+h)}$
3. For $t = 1, ..., \frac{1}{2}p(p - 1)$, change $P_{2+p+t}$ to $i_{M+(2+p+t)}$.

### B.2.3 $\delta_j, \beta_{jl},$ and $\Gamma_{jh}$

Letting $\pi$ enter explicitly in $D_j$, we have that, for $j = 1, ..., p$,

$$\delta_j = D_j(\pi; r) = \frac{1}{r}b_{Y_j,\bar{Y}} \quad \text{and} \quad \nabla_\pi D_j(\pi; r) = \left[ \begin{array}{c} \ell_j^* \otimes \left[ \frac{1}{r} \quad 0 \right] \quad 0 \\ 0 \quad 0 \end{array} \right].$$

Similarly, for $j = 1, ..., p$ and $l = 1, ..., k$, we have that

$$\beta_{jl} = B_{jl}(\pi; r) = b_{Y_{j,l}} - b_{W_{j,l}X} \frac{1}{r}b_{Y_j,\bar{Y}} \quad \text{and} \quad \nabla_\pi B_{jl}(\pi; r) = \left[ \begin{array}{c c c c} \ell_j^* \otimes \left[ -\frac{1}{r}b_{W_{j,l}} \quad 0 \right] \end{array} \right].$$
Last, for \( j, h = 1, \ldots, p \) and \( j \leq h \), \( \Gamma_{jh} \) is given by:

\[
\Gamma_{jh} = G_{jh}(\pi; r) = \sigma_W^{-2}\sigma_{\hat{Y}_j, \hat{Y}_h} - b_{\hat{Y}_j, \hat{Y}_h}r^{-1}b_{\hat{Y}_j, \hat{Y}_h}. 
\]

Letting \( \iota'_{(j,h)} \) take the value 1 at the entry \((j, h)\) corresponding to \( \sigma_W^{-2}\sigma_{\hat{Y}_j, \hat{Y}_h} \), we have

\[
\nabla_\pi G_{jh}(\pi; r) = \begin{bmatrix}
\iota'_j \otimes \begin{bmatrix}
-\frac{1}{r} b_{\hat{Y}_j, \hat{Y}_h} & 0 \\
0 & 1 \times k
\end{bmatrix} + \iota'_h \otimes \begin{bmatrix}
-\frac{1}{r} b_{\hat{Y}_j, \hat{Y}_h} & 0 \\
0 & 1 \times k
\end{bmatrix} & 0 \\
0 & \iota'_{(j,h)}
\end{bmatrix}.
\]

**Joint Confidence Regions**

We sometimes construct a confidence region for the vector of parameters \( \beta_t = (\beta_{1t}, \ldots, \beta_{p,t})' \) associated with the variable \( X_t \) in the system of \( Y \) equations. First, we construct a \( 1 - \alpha_2 \) confidence region \( CR^\rho_{1-\alpha_2} \) for \( \rho \in \mathcal{R}^{k \times c} \). For each \( r \in CR^\rho_{1-\alpha_2} \), the delta method gives the asymptotic distribution of the plug-in estimator \( B_l(\hat{\pi}; r) \) for \( B_l(\pi; r) = b_{Y,X,l} - b_{W,X,l}r^{-1}b_{\hat{Y},\hat{W}} \). Specifically, we have that

\[
\sqrt{n}(B_l(\hat{\pi}; r) - B_l(\pi; r)) \xrightarrow{d} N(0, \Sigma_{B_l}(r)) \quad \text{where} \quad \Sigma_{B_l}(r) = \nabla_\pi B_l(\pi; r)\Sigma B_l(\pi; r)' 
\]

and \( \nabla_\pi B_l(\pi; r) \) stacks the expressions \( \nabla_\pi B_{jl}(\pi; r) \) for \( j = 1, \ldots, p \) derived above. We construct a \( 1 - \alpha_1 \) confidence region \( CR^{B_l}_{1-\alpha_1} \) for \( B_l(\pi; r) \), by inverting the following Wald statistic which has an asymptotic \( \chi^2_p \) distribution:

\[
CR^{B_l}_{1-\alpha_1}(r) = \{ b_t \in \mathcal{B}_t^* : \sqrt{n}(B_l(\hat{\pi}; r) - b_t)'\Sigma_{B_t}(r)^{-1}\sqrt{n}(B_l(\hat{\pi}; r) - b_t) \leq q_{1-\alpha_1} \}.
\]

Here, \( q_{1-\alpha_1} \) is the \( 1 - \alpha_1 \) quantile of \( \chi^2_p \) and we perform a grid search over an initial neighborhood \( \mathcal{B}_t^* \). For instance, we let \( \mathcal{B}_t^* \) be a cube that contains each of the \( p \) unidimensional \( 1 - \alpha_1 \) confidence regions (e.g. \( c = 1.5 \times c' \) where \( c' \) is the \( 1 - \frac{\alpha_1}{2} \) quantile of a standard normal random variable):

\[
\mathcal{B}_t^* = \{ (b_{1t}, \ldots, b_{pt}) : B_{jl}(\hat{\pi}; r) - c*se(B_{jl}(\hat{\pi}; r)) \leq b_{jl} \leq B_{jl}(\hat{\pi}; r) + c*se(B_{jl}(\hat{\pi}; r)) \text{ for } j = 1, \ldots, p \}.
\]

Last, we construct the confidence region \( CR^{\beta_t}_{1-\alpha_1-\alpha_2} \) for \( \beta_t \) by forming the union:

\[
CR^{\beta_t}_{1-\alpha_1-\alpha_2} = \bigcup_{r \in CR^{\rho}_{1-\alpha_2}} CR^{B_l}_{1-\alpha_1}(r)
\]

and use \( CR^{\beta_t}_{1-\alpha_1-\alpha_2} \) to form decisions regarding a null hypothesis for \( (\beta_{1t}, \ldots, \beta_{pt}) \).
C Extension of the Framework to Panel Data

Consider the unbalanced panel equations with firm fixed effects $\gamma_i$:

$$Y_{it} = \gamma_i' + X_{it}' \beta + U_{it} \delta + \eta_{it}' \quad \text{and} \quad W_{it} = U_{it} + \varepsilon_{it} \quad \text{for} \ i = 1, \ldots, n \ \text{and} \ t \in S_i.$$ 

We assume that the data is missing at random from certain time periods. Specifically, we let

$$X_{it} = \begin{cases} 1 & \text{if} \ i \in S_i \\ 0 & \text{otherwise} \end{cases}$$

When time fixed effects are included, this imposes “strict exogeneity” across time periods, as is common when applying a within transformation. Given that $\bar{A}_it \equiv A_{it} - \bar{A}_i$, we obtain the panel analogue of assumption $A_1$:

$$\bar{Y}_{it} = \bar{X}_i \beta + \bar{U}_{it} \delta + \bar{\eta}_{it} \quad \text{and} \quad \bar{W}_{it} = \bar{U}_{it} + \bar{\varepsilon}_{it} \quad \text{for} \ i = 1, \ldots, n.$$ 

Let the binary indicator $I_{is}$ denote whether the observation $(Y_{it}, X_{it}, W_{it})$ is missing (at random). Let $I_i$ stack $I_{it}$ for $t = 1, \ldots, T$. Let

$$\sigma_{\bar{A}_it, \bar{B}_i} \equiv E(\bar{A}_{it}' \bar{B}_i) = E(\sum_{t \in S_i} \bar{A}_{it}' \bar{B}_i) = \sum_{t=1}^{T} E(I_{it} \bar{A}_{it}' \bar{B}_i) = \sum_{t=1}^{T} E(I_{it}) E(\bar{A}_{it}' \bar{B}_i).$$ 

The number of time periods $T$ should not be confused with the dimension of the nuisance parameter $\lambda_{2T \times 1}$ in Online Appendix B.

12
In particular, we have $\sigma_{X_i,\tilde{y}_i} = 0$ and $\sigma_{X_i,\tilde{e}_i} = 0$. Further, let

$$b_{\tilde{A},B_i} \equiv \sigma_{B_i}^{-2} \sigma_{B_i,\tilde{A}} \quad \text{and} \quad \epsilon'_{\tilde{A}_{it},B_{it}} \equiv \tilde{A}_{it}' - \tilde{B}_{it}' b_{\tilde{A},B_i}.$$ 

Then, provided $\sigma_{X_i}^2$ is nonsingular,

$$\beta = b_{\tilde{Y}_i,\tilde{X}_i} - b_{\tilde{W}_i,\tilde{X}_i} \delta.$$ 

Let $\tilde{A}_{it} \equiv \epsilon_{\tilde{A}_{it},\tilde{X}_i}$ and $\tilde{A}_i = [\tilde{A}_{i1}, ..., \tilde{A}_{iT_i}]'$. By A1-A3, we obtain

$$\tilde{Y}_i = \tilde{U}_i \delta + \tilde{\eta}_i \quad \text{and} \quad \tilde{W}_i = \tilde{U}_i + \tilde{e}_i.$$ 

Further, we have

$$\sigma_{W_i}^2 = \sigma_{U_i}^2 + \sigma_{\tilde{e}_i}^2, \quad \sigma_{W_i,\tilde{Y}_i} = \sigma_{W_i,\tilde{U}_i} \delta = \sigma_{U_i}^2 \delta, \quad \text{and} \quad \sigma_{\tilde{Y}_i}^2 = \delta' \sigma_{U_i}^2 \delta + \sigma_{\tilde{\eta}_i}^2.$$ 

Provided $\sigma_{W_i}^2$ is nonsingular, we have

$$b_{\tilde{W}_i,\tilde{Y}_i} = \sigma_{W_i}^{-2} \sigma_{\tilde{W}_i,\tilde{Y}_i} = \rho \delta \quad \text{and} \quad \sigma_{\tilde{W}_i}^{-2} \sigma_{Y_i}^2 = \delta' \rho \delta + \sigma_{\tilde{W}_i}^{-2} \sigma_{\tilde{\eta}_i}^2,$$

where

$$\rho = \sigma_{W_i}^{-2} \sigma_{U_i}^2 = E(\sum_{t \in S_i} \tilde{W}_{it} \tilde{W}_i')^{-1} E(\sum_{t \in S_i} \tilde{U}_{it} \tilde{U}_i').$$

Given $\rho \neq 0$, we obtain the representation from Theorem 3.1 and we apply the results of the paper to the transformed variables. For inference, we use the robust standard errors that are clustered at the firm level. For example, we estimate $b_{\tilde{A},B_i}$ and $\epsilon_{\tilde{A},B_i} = (\epsilon'_{\tilde{A}_{it},B_i}, ..., \epsilon'_{\tilde{A}_{iT_i},B_i})'$ using their plug in sample analogues

$$\hat{b}_{\tilde{A},B_i} \equiv \left(\frac{1}{n} \sum_{i=1}^{n} \tilde{B}_i' \tilde{B}_i\right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \tilde{B}_i' \tilde{A}_i\right) \quad \text{and} \quad \hat{\epsilon}'_{\tilde{A}_{it},B_{it}} = \tilde{A}_{it}' - \tilde{B}_{it}' \hat{b}_{\tilde{A},B_i}$$

and estimate the asymptotic variance of $\sqrt{n}(\hat{b}_{\tilde{A},B_i} - b_{\tilde{A},B_i})$ by

$$\left(\frac{1}{n} \sum_{i=1}^{n} \tilde{B}_i' \tilde{B}_i\right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \tilde{B}_i' \tilde{\epsilon}_{\tilde{A}_i,B_i}\right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \tilde{B}_i' \tilde{B}_i\right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \tilde{B}_i' \tilde{B}_i\right)^{-1}.$$

Note that the interpretation of A4-A6 applies to the stacked and within-transformed variables. In particular, A4 assumes that

$$\sigma_{\tilde{e}_i}^2 = E(\sum_{t \in S_i} \tilde{e}_{it}^2) = \sum_{t=1}^{T} E(I_{it})E(\tilde{e}_{it}^2) \leq \kappa \sigma_{\tilde{U}_i}^2 = \kappa E(\sum_{t \in S_i} \tilde{U}_{it}^2) = \kappa \sum_{t=1}^{T} E(I_{it})E(\tilde{U}_{it}^2).$$
For this to hold, it suffices that \( E(\varepsilon_{it}^2) \leq \kappa E(\tilde{U}_{it}^2) \) for \( t = 1, \ldots, T \). A_5 assumes that

\[
R_{Y_{ij}, \tilde{U}_{it}}^2 = 1 - \frac{\sigma_{\tilde{\eta}_{jit}}^2}{\sigma_{Y_{ij}}^2} = 1 - \frac{E(\sum_{t \in S_i} \tilde{\eta}_{jit}^2)}{E(\sum_{t \in S_i} \tilde{Y}_{jit}^2)} = 1 - \frac{\sum_{t=1}^T E(I_{it}) E(\tilde{\eta}_{jit}^2)}{\sum_{t=1}^T E(I_{it}) E(\tilde{Y}_{jit}^2)} \leq \tau_j,
\]

and it suffices for this that \( R_{Y_{ij}, \tilde{U}_{it}}^2 = 1 - \frac{\sigma_{\tilde{\eta}_{jit}}^2}{\sigma_{Y_{ij}}^2} \leq \tau_j \) for \( t = 1, \ldots, T \). And A_6 assumes that

\[
\zeta_{jh} \leq r_{\tilde{\eta}_{jit} \tilde{\eta}_{hit}} = \frac{E(\sum_{t \in S_i} \tilde{\eta}_{jit} \tilde{\eta}_{hit})^2}{E(\sum_{t \in S_i} \tilde{\eta}_{jit}^2) E(\sum_{t \in S_i} \tilde{\eta}_{hit}^2)} = \frac{\sum_{t=1}^T E(I_{it}) E(\tilde{\eta}_{jit} \tilde{\eta}_{hit})}{[\sum_{t=1}^T E(I_{it}) E(\tilde{\eta}_{jit}^2)][\sum_{t=1}^T E(I_{it}) E(\tilde{\eta}_{hit}^2)]^{\frac{1}{2}}} \leq \tau_{jh},
\]

which holds if one imposes the same sign restriction on \( Cov(\tilde{\eta}_{jit}, \tilde{\eta}_{hit}) \) for \( t = 1, \ldots, T \).

The panel analysis without fixed effects proceeds similarly but omits the within transformation (i.e. it sets \( \gamma_i = \gamma \) for \( i = 1, \ldots, n \) and \( \tilde{A}_{it} = A_{it} - \frac{1}{n} \sum_{i=1}^n \sum_{t \in S_i} A_{it} \)).

**D Mathematical Proofs**

**Proof of Theorem 3.1** By A_2-A_3, \( Cov[(\eta', \varepsilon'), X] = 0 \). Since \( Var(X) \) is nonsingular, \( A_1 \) gives

\[
\beta = b_{Y,X} - b_{W,X} \delta.
\]

A_2-A_3 also give \( \sigma_{\tilde{U}, \varepsilon} = 0 \) and \( \sigma_{\tilde{U}, \eta} = \sigma_{\varepsilon, \eta} = 0 \). Using \( \tilde{\varepsilon} = \varepsilon - E(\varepsilon) \) and \( \tilde{\eta} = \eta - E(\eta) \) together with \( \tilde{Y}' = \tilde{U} \delta + \tilde{\eta}' \) and \( \tilde{W} = \tilde{U} + \tilde{\varepsilon} \), we obtain

\[
\sigma_{\tilde{W}}^2 = \sigma_{\tilde{U}}^2 + \sigma_{\tilde{\varepsilon}}^2, \quad \sigma_{\tilde{W}, \tilde{Y}} = \sigma_{\tilde{W}, \tilde{U}} \delta = \sigma_{\tilde{U}}^2 \delta, \quad \text{and} \quad \sigma_{\tilde{Y}}^2 = \delta' \sigma_{\tilde{U}}^2 \delta + \sigma_{\eta}^2.
\]

Since \( Var[(X', U)'] \) is nonsingular, \( \sigma_{\tilde{U}}^2 \neq 0 \). Thus, \( \sigma_{\tilde{W}}^2 \neq 0 \) and

\[
b_{\tilde{Y}, \tilde{W}} \equiv \sigma_{\tilde{W}}^{-2} \sigma_{\tilde{Y}, \tilde{W}} = \rho \delta \quad \text{and} \quad \sigma_{\tilde{W}}^{-2} \sigma_{\tilde{Y}}^2 = \delta' \rho \delta + \Gamma.
\]

Since \( \rho \neq 0 \), we obtain

\[
\delta = D(\rho) \equiv \frac{1}{\rho} b_{\tilde{Y}, \tilde{W}}
\]

\[
\beta = B(\rho) \equiv b_{Y,X} - b_{W,X} D(\rho) = b_{Y,X} - b_{W,X} \frac{1}{\rho} b_{\tilde{Y}, \tilde{W}}, \quad \text{and}
\]

\[
\Gamma = G(\rho) \equiv \sigma_{\tilde{W}}^{-2} \sigma_{\tilde{Y}}^2 - D(\rho)'\rho D(\rho) = \sigma_{\tilde{W}}^{-2} \sigma_{\tilde{Y}}^2 - b_{\tilde{Y}, \tilde{W}}' \frac{1}{\rho} b_{\tilde{Y}, \tilde{W}}.
\]

**Lemma D.1** Under the conditions of Theorem 3.1 \( R_{Y_{ij}, \tilde{W}}^2 \leq R_{Y_{ij}, \tilde{U}}^2 \).
Proof of Lemma D.1: If \( \sigma_{Y_j}^2 = 0 \), set \( R_{Y_j,W}^2 = R_{Y_j,U}^2 = 0 \). If \( 0 < \sigma_{Y_j}^2 \), we have

\[
R_{Y_j,W}^2 = \frac{\sigma_W^2}{\sigma_{Y_j}^2} h_{Y_j,W} = \frac{\sigma_W^2}{\sigma_{Y_j}^2} (\delta_j \rho)^2
\]

and

\[
R_{Y_j,U}^2 = 1 - \frac{\sigma_{Y_j}^2}{\sigma_{Y_j}^2} = \frac{1}{\sigma_{Y_j}^2} (\sigma_{Y_j}^2 - \sigma_{Y_j}^2) = \frac{1}{\sigma_{Y_j}^2} \delta_j^2 \sigma_{Y_j}^2 = \frac{\sigma_W^2}{\sigma_Y^2} \delta_j^2 \rho
\]

It follows that

\[
R_{Y_j,U}^2 - R_{Y_j,W}^2 = \frac{\sigma_W^2}{\sigma_Y^2} (\delta_j^2 \rho - \delta_j^2 \rho^2) = \frac{\sigma_W^2}{\sigma_Y^2} \rho (1 - \rho) \delta_j^2 \geq 0.
\]

Proof of Corollary 3.2: The identification set \( \mathcal{J}^{k,r,e} \) obtains from \( A_1-A_6 \) and the \( (\text{Var}[(\tilde{Y}', \tilde{W}')] \) moments given by (in)equalities (4-7), using the expressions in Theorem 3.1. To show that \( \mathcal{J}^{k,r,e} \) is sharp, let \( d = D(r), b = B(r), \) and \( g = G(r) \). We show that for each \((r, d, b, g) \in \mathcal{J}^{k,r,e}\) there exist random variables \((U^*, \eta'^*, \varepsilon^*)\) such that \( Y' = X'b + U^*d + \eta'^* \) and \( W = U^* + \varepsilon^* \) that satisfy \( A_2-A_6 \). Specifically, \((X, U^*, \varepsilon^*, \eta'^*)\) satisfy \( A_2-A_3, \text{Cov}[^*], \text{Cov}[\varepsilon, (X', U'^*)] = 0, \text{Cov}[\varepsilon^*, (\eta'^*, X', U')] = 0 \). Further, \( \sigma_{W}^2 \sigma_{Y_j}^2 = r \) and thus \( A_4 \) holds, \( \sigma_{Z_j}^2 \leq \kappa \sigma_{Y_j}^2 \). Last, \( G(r) = \sigma_W^2 \sigma_{Y_j}^2 \) and therefore \( A_5 \) holds when \( \sigma_{Y_j}^2 \neq 0 \),

\[
1 - \frac{\sigma_{Y_j}^2}{\sigma_{Y_j}^2} = \frac{\sigma_W^2}{\sigma_Y^2} (\frac{\sigma_{Y_j}^2}{\sigma_W^2} - G_{j}(r)) \leq \frac{\sigma_W^2}{\sigma_Y^2} [\frac{\sigma_{Y_j}^2}{\sigma_W^2} - \frac{\sigma_{Y_j}^2}{\sigma_W^2} (1 - \tau_j)] = \tau_j,
\]

and \( A_6 \) holds since \( \tau_{j} \leq sgn(G_{j}(r)) \leq \bar{\tau}_{j} \).

To construct these variables we proceed similarly to Chalak and Kim (2019, proof of corollary 3.2). In particular, we let \( V \) be any random variable such that \( \tilde{V} \equiv \varepsilon_{V,X} \) is nondegenerate and satisfies

\[
\sigma_{\tilde{V}W'} = \sqrt{r} \sigma_{\tilde{V}W'} \quad \text{and} \quad \sigma_{\tilde{V}Y'} = \frac{1}{\sqrt{r}} \sigma_{\tilde{V}W'} \frac{\sigma_{\tilde{V}Y'}}{\sigma_{\tilde{V}W'}}.
\]

Note that these covariance restrictions are coherent. Specifically,

\[
\text{Var}(\tilde{V}, \tilde{W}, \tilde{Y}') = \begin{bmatrix}
\sigma_{V}^2 & \sqrt{r} \sigma_{V} \sigma_{W'} & \frac{\sigma_{V} \sigma_{W}}{\sqrt{r}} \\
\frac{\sigma_{W} \sigma_{V}}{\sqrt{r}} & \sigma_{W}^2 & \frac{\sigma_{W} \sigma_{W'}}{\sqrt{r}} \\
\frac{\sigma_{W} \sigma_{V}}{\sqrt{r}} & \frac{\sigma_{W} \sigma_{W'}}{\sqrt{r}} & \sigma_{W'}^2
\end{bmatrix}
\]

is positive semi-definite because \( 0 < \sigma_{V}^2 \) and its Schur complement

\[
0 \leq \sigma_{(\tilde{W}, \tilde{Y}')'} - \sigma_{(\tilde{W}, \tilde{Y}')'} \sigma_{(\tilde{W}, \tilde{Y}')'} ^{-1} \sigma_{(\tilde{W}, \tilde{Y}')'} = \begin{bmatrix}
(1 - r) \sigma_{W}^2 & 0 \\
0 & \sigma_{W}^2 G(r)
\end{bmatrix}
\]

15
is positive semi-definite since it is block diagonal with $0 \leq (1 - r)\sigma_W^2$ and $0 \leq G(r)$.

For instance, to construct $V$, set $\sigma_\tilde{V}$ to some value (e.g. $\sigma_\tilde{V} = 1$) and let $\vartheta$ be any random variable that is uncorrelated with $(X', W, Y')'$ (e.g. a residual from a regression on $(X', W, Y')$). When $\sigma_{\tilde{W},\tilde{Y}}^2$ is nonsingular, one can use the above restrictions on $\sigma_{\tilde{W},\tilde{V}}$ and $\sigma_{\tilde{Y},\tilde{V}}$ to construct $b_{\tilde{V},(\tilde{W},\tilde{Y})'}$ and the scalar
\[\kappa = \left\{ \frac{1}{\sigma_\tilde{V}^2} [\sigma_\tilde{V}^2 - b'_{\tilde{V},(\tilde{W},\tilde{Y})'} \sigma_{\tilde{W},\tilde{Y}}^2 b_{\tilde{V},(\tilde{W},\tilde{Y})'}] \right\}^{\frac{1}{2}} \]

($\kappa$ is set such that the variance of the generated $\tilde{V}$ is $\sigma_\tilde{V}^2$) in order to generate
\[\tilde{V} = (\tilde{W}, \tilde{Y}) b_{\tilde{V},(\tilde{W},\tilde{Y})'} + \kappa \vartheta.\]

If $\sigma_{\tilde{W},\tilde{Y}}^2$ is singular, one can generate $\tilde{V}$ by omitting the redundant $\tilde{Y}$ components from the above regression construction. Last, $V = X' b_{V,X} + \tilde{V} + E[V - X' b_{V,X}]$ obtains by setting $b_{V,X}$ and $E(V)$ to some value (e.g. zero).

Then it suffices to construct $U^*, \varepsilon^*$, and $\eta^*$ as follows
\[W \equiv (X', V) b_{W,(X',V)'} + \{ \varepsilon_{W,(X',V)'} + E[W - (X', V)b_{W,(X',V)'}] \} \equiv U^* + \varepsilon^*,\]
and, if $r \neq 1$,
\[Y \equiv (X', V, \varepsilon^*) b_{Y,(X',V,\varepsilon^*)'} + \{ \varepsilon_{Y,(X',V,\varepsilon^*)'} + E[Y - (X', V, \varepsilon^*)b_{Y,(X',V,\varepsilon^*)'}] \} \equiv (X', V, \varepsilon^*) b_{Y,(X',V,\varepsilon^*)'} + \eta^*\]

whereas if $r = 1$ then $r_{\tilde{W},\tilde{V}} = 1$ and $\varepsilon_{W,(X',V)'} = \varepsilon_{\tilde{W},\tilde{V}} = 0$ and
\[Y = (X', V) b_{Y,(X',V)'} + \{ \varepsilon_{Y,(X',V)'} + E[Y - (X', V)b_{Y,(X',V)'}] \} \equiv (X', V) b_{Y,(X',V)'} + \eta^.*\]

In particular, $(X, U^*, \varepsilon^*, \eta^*)$ satisfy $A_2-A_3$ since by construction $Cov[\varepsilon^*, (X', U^*)'] = 0$ and $Cov[\varepsilon^*, (\eta^*, X', U^*)'] = 0$. To verify that $A_1$ holds, note that if $r \neq 1$,
\[Y = V b_{\tilde{Y},\tilde{V}} + X'(b_{V,X} - b_{V,X} b_{\tilde{Y},\tilde{V}}) + \varepsilon^* b_{Y,\varepsilon^*} + \{ \varepsilon_{Y,(X',V,\varepsilon^*)'} + E[Y - (X', V, \varepsilon^*)b_{Y,(X',V,\varepsilon^*)'}] \}
\begin{eqnarray*}
&=& V b_{\tilde{W},\tilde{V}} d + X'(b_{V,X} - b_{V,X} b_{\tilde{W},\tilde{V}})d + X'(b_{V,X} - b_{V,X} d) + \varepsilon^* b_{Y,\varepsilon^*} + \eta^* \\
&=& (X', V) b_{W,(X',V)'} d + X'b + \eta^* \\
&=& U^* d + X'b + \eta^* \\
\end{eqnarray*}\]

where the first equality uses $Cov[\varepsilon^*, (X', V)'] = 0$ and partitioned regression, the second equality makes use of
\[b_{\tilde{Y},\tilde{V}} = \sigma_{\tilde{W}}^{-2} \sigma_{\tilde{Y},\tilde{V}} = \sigma_{\tilde{W}}^{-2} \frac{1}{\sqrt{r}} \sigma_{\tilde{W}} \sigma_{\tilde{V}} \sigma_{\tilde{Y},\tilde{W}} = \sigma_{\tilde{W}}^{-2} \frac{1}{\sqrt{r}} = \sigma_{\tilde{W}}^{-2} \sigma_{\tilde{W},\tilde{V}} b_{\tilde{Y},\tilde{V}} = b_{\tilde{W},\tilde{V}} d,\]

16
and the third equality uses partitioned regression, \( b = b_{Y,X} - b_{W,X}d \), and

\[
\begin{align*}
\sigma^2_{\hat{Y},\bar{W}} &= \frac{\sigma^2_{\hat{Y},\bar{W},(X',V)'} \sigma^2_{\bar{W},(X',V)'}}{\sigma^2_{\bar{W},(X',V)'}} = \frac{1}{\sigma^2_{\bar{W},\bar{V}}} \Cov(\hat{Y}, \bar{W} - \hat{V}b_{\hat{W},\hat{V}}) \\
&= \frac{1}{(1 - r)\sigma^2_{\bar{W}}} [\sigma^2_{\hat{Y},\bar{W}} - \frac{1}{\sqrt{\sigma^2_{\hat{W}} \sigma^2_{\bar{V}}}} \sqrt{\sigma^2_{\bar{V}}} \sigma^2_{\bar{W}}] = 0.
\end{align*}
\]

If \( r = 1 \), a similar calculation gives,

\[
\begin{align*}
Y &= (X', V)b_{Y,(X',V)'} + \{\epsilon_{Y,(X',V)'} + E[Y - (X', V)b_{Y,(X',V)'}]\} \\
&= (X', V)b_{W,(X',V)'}d + X'b + \eta^* \equiv U^*d + X'b + \eta^*.
\end{align*}
\]

Last, to verify that \( A_4 - A_6 \) hold, it suffices to verify that

\[
\sigma^2_{\theta,\bar{V}} = \frac{\Var(\hat{V}b_{\hat{W},\bar{V}})}{\sigma^2_{\bar{V}}} = \frac{\sigma^2_{\hat{V},\bar{V}}}{\sigma^2_{\bar{V}}}, \text{ and} \\
G(r) = \sigma^2_{\theta,\bar{V}} - b'_{\theta,\bar{V}}b_{\theta,\bar{V}} = \sigma^2_{\theta}(d'\sigma^2_{\theta}d + \sigma^2_{\eta^*}) - b'_{\theta,\bar{V}}b_{\theta,\bar{V}} = \sigma^2_{\theta} - \sigma^2_{\eta^*}.
\]

**Proof of Corollary 3.3.** We start by deriving the identification region \( R^{k,r,c} \) for \( \rho \). First, we show that \( R^2_{W,Y} \leq \rho \leq 1 \). If \( \sigma^2_{\hat{Y},\hat{W}} = 0 \) or \( \sigma^2_{\bar{V}} = 0 \) then set \( R^2_{W,Y} = 0 \leq \rho \leq 1 \). Suppose that \( \sigma^2_{\hat{Y},\hat{W}} \neq 0 \). Since \( 0 < \rho \) and \( 0 < \Gamma \) then for any vector \( x \in \mathbb{R}_+ \), we have

\[
0 \leq \rho x'\sigma^2_{\hat{W}} \sigma^2_{\hat{Y}}x - x'b'_{\hat{W},\hat{V}}b_{\hat{W},\hat{V}}x.
\]

Suppose that \( \sigma^2_{\hat{Y}} \) is positive definite so that \( 0 < \sigma^2_{\hat{W}} \sigma^2_{\hat{Y}} \sigma^2_{\hat{W}} \) (this is without loss of generality since we can drop the redundant \( \hat{Y} \) components otherwise). In particular, for \( x = \sigma^2_{\hat{W}} \sigma_{\hat{Y},\hat{W}} \), we obtain

\[
R^2_{\hat{Y},\hat{W}} = \sigma^2_{\hat{Y},\hat{W}} \sigma^2_{\hat{Y},\hat{W}} - \frac{(\sigma^2_{\hat{Y},\hat{W}} \sigma^2_{\hat{Y},\hat{W}} \sigma^2_{\hat{W}} \sigma^2_{\hat{V}} \sigma^2_{\hat{Y},\hat{W}})(\sigma^2_{\hat{Y},\hat{W}} \sigma^2_{\hat{Y},\hat{W}} \sigma^2_{\hat{W}} \sigma^2_{\hat{V}} \sigma^2_{\hat{Y},\hat{W}})}{(\sigma^2_{\hat{W}} \sigma^2_{\hat{Y},\hat{W}})(\sigma^2_{\hat{W}} \sigma^2_{\hat{Y},\hat{W}} \sigma^2_{\hat{V}} \sigma^2_{\hat{Y},\hat{W}})} \leq \rho \leq 1.
\]

Second, by \( A_4 \), we have \( 1 - \rho = \frac{\sigma^2_{\hat{Y}}}{\sigma^2_{\hat{Y}}} \leq \kappa \frac{\sigma^2_{\hat{Y}}}{\sigma^2_{\hat{Y}}} = \kappa \rho \) and thus \( \rho \in [\frac{1}{1 + \kappa}, 1] \). Third, by \( A_5 \), we have that for \( j = 1, \ldots, p, R^2_{\hat{Y},\hat{W}} = (1 - \frac{\sigma^2_{\eta_j}}{\sigma^2_{\hat{Y}}}) \leq \tau_j \) (recall that if \( \sigma^2_{\eta_j} = 0 \) then we set \( R^2_{\hat{Y},\hat{W}} = 0 \)). Multiplying by \( \sigma^2_{\hat{Y}} \) and substituting for \( \Gamma_{jj} \) we obtain

\[
\frac{b'_{\hat{Y},\hat{W}}b_{\hat{Y},\hat{W}}}{\rho^2_{\hat{Y},\hat{W}}} = \frac{\sigma^2_{\hat{Y}}}{\sigma^2_{\hat{W}}} - (\sigma^2_{\hat{W}} \sigma^2_{\hat{Y}} - b'_{\hat{Y},\hat{W}}b_{\hat{Y},\hat{W}}) \leq \frac{\tau_j}{\sigma^2_{\hat{Y}}}
\]
and thus \( \frac{1}{\tau_j} R^2_{W_j,W_j} = \frac{1}{\tau_j} b^2_{\tilde{Y}_j,W} \sigma^2_{\tilde{W}} \leq \rho \leq 1 \). Last, \( R^c_{jh} \) obtains since \( 0 < \rho \) and \( \Gamma_{jh} = G_{jh}(\rho) = \sigma^{-2}_W \sigma_{\tilde{Y}_j,\tilde{Y}_h} - b_{\tilde{Y}_j,W} \frac{1}{\rho} b_{\tilde{Y}_h,W} \) so that

\[
G_{jh}(\rho) \leq 0 \text{ if and only if } \begin{cases}
\frac{b_{\tilde{Y}_j,W} b_{\tilde{Y}_h,W}}{\sigma^2_{\tilde{W}} \sigma_{\tilde{Y}_j,\tilde{Y}_h}} \leq \rho & \text{when } \sigma^{-2}_W \sigma_{\tilde{Y}_j,\tilde{Y}_h} < 0 \\
0 \leq b_{\tilde{Y}_j,W} b_{\tilde{Y}_h,W} & \text{when } \sigma^{-2}_W \sigma_{\tilde{Y}_j,\tilde{Y}_h} = 0 \\
\rho \leq \frac{b_{\tilde{Y}_j,W} b_{\tilde{Y}_h,W}}{\sigma^2_{\tilde{W}} \sigma_{\tilde{Y}_j,\tilde{Y}_h}} & \text{when } \sigma^{-2}_W \sigma_{\tilde{Y}_j,\tilde{Y}_h} > 0
\end{cases}
\]

Combining the results, we have \( \rho \in R^{k,\tau,c} = [R^2_{W,Y}, 1] \cap [\frac{1}{1+\kappa}, 1] \cap \prod_{j=1}^p \left[ \frac{1}{\tau_j} R^2_{W_j,Y_j} \right] \cap \prod_{j,h=1}^p \mathcal{R}^c_{jh} \).

To show that \( R^{k,\tau,c} \) is sharp, it suffices to show that every \( r \in R^{k,\tau,c} \) corresponds to a point \( (r,d,b,g) \in \mathcal{J}^{k,\tau,c} \). Let \( r \in R^{k,\tau,c} \). First, we show that \( 0 < G(r) \). If \( R^2_{W,Y} = 0 \) then \( G(r) = \sigma^{-2}_W \sigma^2_{\tilde{Y}} \geq 0 \). Otherwise, note that

\[
G(1) = \sigma^{-2}_W \sigma^2_{\tilde{Y}} - b'_{Y,\tilde{W}} b_{\tilde{Y},W} = \sigma^{-2}_W \sigma_{\tilde{Y},W} \sigma^{-2}_W \sigma_{\tilde{W},\tilde{Y}} = \sigma^{-2}_W E(\epsilon_{\tilde{Y},W} \epsilon'_{\tilde{Y},W}) \geq 0.
\]

Further, when \( R^2_{W,Y} \neq 0 \), \( 0 < \sigma^4_W R^2_{W,Y} \) and

\[
\sigma^4_W R^2_{W,Y} G(R^2_{W,Y}) = (R^2_{W,Y} \sigma^2_{\tilde{W}}) \sigma^2_{\tilde{Y}} - \sigma_{\tilde{Y},W} \sigma_{\tilde{W},\tilde{Y}} = \text{Var}(b'_{W,Y} \tilde{Y}) \sigma^2_{\tilde{Y}} - \sigma_{\tilde{Y},W} \sigma_{\tilde{W},\tilde{Y}} \geq 0
\]

since and for any vector \( x \in \mathbb{R}^p \), applying the Cauchy–Schwarz inequality gives

\[
x' \text{Var}(b'_{W,Y} \tilde{Y}) \sigma^2_{\tilde{Y}} x - x' \sigma_{\tilde{Y},W} \sigma_{\tilde{W},\tilde{Y}} x = \text{Var}(b'_{W,Y} \tilde{Y}) \text{Var}(x' \tilde{Y}) - [\text{Cov}(x' \tilde{Y}, \tilde{W})]^2
\]

\[
= \text{Var}(b'_{W,Y} \tilde{Y}) \text{Var}(x' \tilde{Y}) - [\text{Cov}(x' \tilde{Y}, b'_{W,Y} \tilde{Y})]^2 \geq 0
\]

where we make use of \( \tilde{W}' = \tilde{Y}' b_{W,Y} + \epsilon'_{\tilde{Y},W} \) and \( \text{Cov}(\tilde{Y}, \epsilon_{\tilde{W},\tilde{Y}}) = 0 \) in the last equality. Since \( r \in R^{k,\tau,c} \subseteq [R^2_{W,Y}, 1] \), there exists \( 0 \leq \lambda \leq 1 \) such that \( \frac{1}{r} = \lambda + (1 - \lambda) \frac{1}{R^2_{W,Y}} \) and it follows that

\[
0 \leq G(r) = \lambda G(1) + (1 - \lambda) G(R^2_{W,Y}).
\]

Clearly, \( \frac{1}{1+\kappa} \leq r \leq 1 \). Further, for \( j = 1, \ldots, p \), if \( \sigma^2_{\tilde{Y}_j} \neq 0 \) then \( \frac{1}{\tau_j} b^2_{\tilde{Y}_j,W} \sigma^2_{\tilde{W}} = \frac{1}{\tau_j} R^2_{W_j,Y_j} \leq r \)

implies that \( \frac{1}{\sigma^2_{\tilde{W}}} (1 - \tau_j) \leq \frac{\sigma^2_{\tilde{Y}_j}}{\sigma^2_{\tilde{W}}} - b^2_{\tilde{Y}_j,W} \frac{1}{r} = G_{jj}(r) \) (if \( \sigma^2_{\tilde{Y}_j} = 0 \) then \( 0 = \frac{1}{\sigma^2_{\tilde{W}}} (1 - \tau_j) \leq G_{jj}(r) = 0 \)). Last, from the expression for \( G_{jh}(r) = \sigma^{-2}_W \sigma_{\tilde{Y}_j,\tilde{Y}_h} - b_{\tilde{Y}_j,W} \frac{1}{r} b_{\tilde{Y}_h,W} \), we have that

\[
\sigma_{\tilde{Y}_j,\tilde{Y}_h} \leq \text{sgn}(G_{jh}(r)) \leq \tilde{c}_{jh} \text{ for } j, h = 1, \ldots, p \text{ with } j < h.
\]

The sharp bounds \( D^{k,\tau}, B^{k,\tau}, \) and \( G^{k,\tau} \) for \( \delta, \beta, \) and \( \Gamma \) follow from the mappings \( D(\cdot), B(\cdot), \) and \( G(\cdot) \) in Theorem 3.1.
Proof of Theorem 5.1 First, for random column vectors $A$ and $B$, we collect the regression intercept and slope estimands as follows

$$A' = [E(A)' - E(B)'b_{AB}] + B'b_{AB} + \epsilon'_{AB} \equiv (1, B')b'_{AB} + \epsilon'_{AB}.$$ 

Given observations $\{A_i, B_i\}_{i=1}^n$, denote the linear regression intercept ($\hat{b}_{AB}^0$) and slope ($\hat{b}_{AB}^{*}$) estimators and the sample residual ($\hat{\epsilon}_{AB,i}$) by:

$$\hat{b}_{AB} = (\hat{b}_{AB}^0, \hat{b}_{AB}^{*})' \equiv (-\frac{1}{n} \sum_{i=1}^{n}(1, B_i')'(1, B_i'))^{-1}(-\frac{1}{n} \sum_{i=1}^{n}(1, B_i')'A_i')$$ and $\hat{\epsilon}_{AB,i} \equiv A_i' - (1, B_i')\hat{b}_{AB}$.

Further, we collect into $\pi^*$ the following estimands

$$\pi^* \equiv [\text{vec}(b_{Y_i,X_i}')', b_{Y_i,X_i'}^{*'}, b_{W_i,Y_i,X_i'}^{*'}, ..., b_{W_i,Y_{ip},X_i'}^{*'}, \text{vec}(b_{Y,X}')', b_{Y,X}^{*'}, \sigma^{-2}_W \text{vec}(\sigma^2_{Y})']',$$

and into $\tilde{\pi}$ the corresponding estimators:

$$\tilde{\pi} \equiv [\text{vec}(\tilde{b}_{Y_i,X_i}')', \tilde{b}_{Y_i,X_i'}^{*'}, \tilde{b}_{W_i,Y_i,X_i'}^{*'}, ..., \tilde{b}_{W_i,Y_{ip},X_i'}^{*'}, \text{vec}(\tilde{b}_{Y,X}')', \tilde{b}_{Y,X}^{*'}, \tilde{\sigma}^{-2}_W \text{vec}(\tilde{\sigma}^2_{Y})']'.$$

Last, let $\tilde{\mu}^2_A = \frac{1}{n} \sum_{i=1}^n A_i A_i'$, $\hat{Q} \equiv \text{diag}\{ I \otimes \hat{\mu}^2_{(1,W,X')}, \hat{\mu}^2_{(1,Y',X')}, \hat{\mu}^2_{(1,Y_{ip},X')}, ..., I \otimes \hat{\mu}^2_{(1,Y')}, \hat{\mu}^2_{(1,Y')}, I \otimes 1_T \otimes \hat{\sigma}^2_W \}$.

and

$$L \equiv \frac{1}{n} \sum_{i=1}^{n} [\text{vec}((1, W_i, X_i')'\epsilon_{Y_i(X')}'i), (1, Y_i', X_i')\epsilon_{W_i(Y',X')}'i, (1, Y_{ip}, X_i')\epsilon_{W_i(Y_{ip},X')}'i, \text{vec}((1, X_i')'\epsilon_{Y_{ip}(X')}', (1, X_i')\epsilon_{W,X,i}, \text{vec}(\epsilon_{Y,X,i}\epsilon_{Y_{ip},X,i}^-\sigma^2_Y)]'. $$

Recall that $Q$ is finite (by A1(i)) and nonsingular. For a symmetric matrix $C$ and a vector $D$, let $C_1$ denote the submatrix that removes the last $\frac{1}{2}p(p+1)$ rows and columns of $C$ and let $D_1$ be the subvector that removes the last $\frac{1}{2}p(p+1)$ rows of $D$. Then

$$\sqrt{n}(\tilde{\pi} - \pi^*_1) = \hat{Q}_1^{-1} \sqrt{n}L_1 = (\hat{Q}_1^{-1} - Q_1^{-1})\sqrt{n}L_1 + Q_1^{-1}\sqrt{n}L_1.$$ 

Since (i) gives $\hat{Q}_1^{-1} - Q_1^{-1} = o_p(1)$ and (ii) gives $\sqrt{n}L_1 \rightarrow_d N(0, \Xi_1)$, we obtain that $\sqrt{n}(\pi_1 - \pi^*_1) = Q_1^{-1} \sqrt{n}L_1 + o_p(1) \rightarrow_d N(0, \Sigma^*_1)$. Moreover, it follows from $\hat{\mu}^2_{(1,Y')}$, $\hat{\mu}^2_{(1,Y_{ip})}$, $\sqrt{n}(\tilde{b}_{Y_{ip},X} - b_{Y_{ip},X}^*) = O_p(1)$, and $\frac{1}{n} \sum_{i=1}^{n} \epsilon_{Y_{ip},X,i}(1, X_i')' = E[\epsilon_{Y_{ip},X}(1, X')]' + o_p(1) = o_p(1)$ for $j = 1, ..., p$
Thus, since \( n \) is a proof in Corollary 3.2 implies that \( J \) have the moments given by (in)equalities (4-7), using the expressions in Theorem 3.1. The sharpness proof in Corollary A.1: Multiplying the numerator and denominator by 0.

\[
\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} \epsilon_{Y,j,x,i} \epsilon_{Y_h,x,i} \right) = (\sigma_{\hat{W}})^{-1} n^{-1/2} \sum_{i=1}^{n} \epsilon_{Y,j,x,i} \epsilon_{Y_h,x,i} + o_p(1).
\]

Similarly, by (i), we obtain that

\[
\frac{1}{n} \sum_{i=1}^{n} \epsilon_{Y,j,x,i} \epsilon_{Y_h,x,i} = E(\epsilon_{Y,j,x} \epsilon_{Y_h,x}) + o_p(1) = \sigma_{\hat{Y}_j, \hat{Y}_h} + o_p(1) \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} \epsilon_{W,x,i}^2 = \sigma_W^2 + o_p(1).
\]

Thus, since \( n^{-1/2} \sum_{i=1}^{n} \epsilon_{Y,j,x,i} \epsilon_{Y_h,x,i} \) is \( o_p(1) \) by (ii), we have that for \( j, h = 1, \ldots, p \)

\[
\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} \epsilon_{Y,j,x,i} \epsilon_{Y_h,x,i} \right) = (\sigma_{\hat{W}})^{-1} n^{-1/2} \sum_{i=1}^{n} \epsilon_{Y,j,x,i} \epsilon_{Y_h,x,i} + o_p(1).
\]

Together with \( \sqrt{n}(\hat{\pi}_1 - \pi_1) = Q_1^{-1} \sqrt{n} L_1 + o_p(1) \), we obtain by (i) and (ii) that

\[
\sqrt{n}(\hat{\pi} - \pi^*) = Q^{-1} \sqrt{n} L + o_p(1) \xrightarrow{d} N(0, \Sigma^*)
\]

and therefore that the subvector \( \sqrt{n}(\hat{\pi} - \pi)^d \xrightarrow{d} N(0, \Sigma) \).

**Proof of Corollary A.1** The identification set \( J^{k,\pi,e} \) obtains from \( A_1-A_6 \) and the \((Var[(\hat{Y}, \hat{W})])\) the moments given by (in)equalities (47), using the expressions in Theorem 3.1. The sharpness proof in Corollary 3.2 implies that \( J^{k,\pi,e} \) is sharp. Specifically, since \( G(r) = \sigma_W^{-2} \sigma_{\eta^*}^2 \), we have that \( \xi_{jh} \leq r_{\hat{\eta}_j, \hat{\eta}_h} \leq \hat{c}_{jh} \).

To derive \( R^{k,\pi,e} \), for \( j, h = 1, \ldots, p \) and \( j < h \), consider the restriction

\[
\xi_{jh} \leq \Gamma_{jh} = \frac{G_{jh}(\rho)}{[G_{jj}(\rho)G_{hh}(\rho)]^{1/2}} \leq \frac{\sigma_{\hat{W}}^{-2} \sigma_{\hat{Y}_j, \hat{Y}_h} - b_{\hat{Y}_j, \hat{W}} \frac{1}{\rho} b_{\hat{Y}_h, \hat{W}}}{(\sigma_{\hat{W}}^{-2} \sigma_{\hat{Y}_j} - \frac{1}{\rho^2} b_{\hat{Y}_j, \hat{W}})^{1/2}(\sigma_{\hat{W}}^{-2} \sigma_{\hat{Y}_h} - \frac{1}{\rho^2} b_{\hat{Y}_h, \hat{W}})^{1/2}} \leq \check{c}_{jh}.
\]

If \( \sigma_{\hat{Y}_j}^2 = 0 \) or \( \sigma_{\hat{Y}_h}^2 = 0 \) then \( \sigma_{\eta_j}^2 = 0 \) or \( \sigma_{\eta_h}^2 = 0 \) and \( \xi_{jh} \leq \sigma_{\eta_j, \eta_h} \leq \check{c}_{jh} \) is either incorrect (if \( 0 \notin [\xi_{jh}, \check{c}_{jh}] \)) or uninformative about \( \rho \) (if \( 0 \in [\xi_{jh}, \check{c}_{jh}] \)). Suppose that \( \sigma_{\hat{Y}_j}^2 \neq 0 \) and \( \sigma_{\hat{Y}_h}^2 \neq 0 \). Multiplying the numerator and denominator by \( 0 < \rho \sigma_{\hat{W}}^{-2} \sigma_{\eta^*}^{-1} \sigma_{\hat{Y}_j}^{-1} \) gives

\[
\xi_{jh} \leq \frac{\rho r_{\hat{Y}_j, \hat{Y}_h} - r_{\hat{W}, \hat{Y}_j} r_{\hat{W}, \hat{Y}_h}}{(\rho - R_{\hat{W}, \hat{Y}_j})^{1/2}(\rho - R_{\hat{W}, \hat{Y}_h})^{1/2}} \leq \check{c}_{jh}.
\]
The expression for $\mathcal{R}_{jh}^e$ then follows from encoding the sign of $r_{\eta_j,\eta_h}$ via the function

$$S_{jh}(r) \equiv r \times r\tilde{Y}_j\tilde{Y}_h - r\tilde{W}_j\tilde{W}_h$$

and the magnitude of $r_{\eta_j,\eta_h}$ ($r_{\eta_j,\eta_h}^2 \leq c^2$ or $c^2 \leq r_{\eta_j,\eta_h}^2$) via the quadratic function

$$M_{jh}(r; c) \equiv (r \times r\tilde{Y}_j\tilde{Y}_h - r\tilde{W}_j\tilde{W}_h)^2 - c^2(r - R_{\tilde{W}_j\tilde{Y}_h}^2)(r - R_{\tilde{W}_j\tilde{Y}_h}^2).$$

By Corollary 3.3, we obtain that $\rho \in \mathcal{R}_{jh}^{k,\tau,e} = [R_{\tilde{W}_j\tilde{Y}_h}^2, 1] \cap \left[ \frac{1}{1+\kappa}, 1 \right] \cap \bigcap_{j=1}^{p} \left[ \frac{1}{\tau_j} R_{\tilde{W}_j\tilde{Y}_h}^2, 1 \right] \cap \mathcal{R}_{jh}^e$, with $j, h = 1, \ldots, p$.

In addition, $\mathcal{R}_{jh}^{k,\tau,e}$ is sharp since every $r \in \mathcal{R}_{jh}^{k,\tau,e}$ corresponds to a point $(r, d, b, g) \in \mathcal{J}_{k,\tau,e}$. Specifically, if $r \in \mathcal{R}_{jh}^{k,\tau,e}$ then $\frac{1}{1+\kappa} \leq r \leq 1$, $0 \leq G(r)$, and $R_{\tilde{W}_j\tilde{Y}_h}^2 \leq \tau_j$ for $j = 1, \ldots, p$ by Corollary 3.3. Further, since $r \in \mathcal{R}_{jh}^{k,\tau,e} \subseteq \mathcal{R}_{jh}^e$, from the sign and magnitude restrictions in $S_{jh}(r)$ and $M_{jh}(r; c)$, we have that $\zeta_{jh} \leq \frac{G_{jh}(r)}{[G_{jh}(r)G_{jh}(r)]^{\frac{1}{2}}} \leq \bar{c}_{jh}$ for $j, h = 1, \ldots, p$ with $j < h$.

Next, we examine the behavior of $S_{jh}(r)$ and $M_{jh}(r; c)$ when $\sigma_{\tilde{Y}_j,\tilde{Y}_h}^2 \neq 0$. First, we have

$$0 \leq S_{jh}(r) \iff \begin{cases} \frac{r\tilde{W}_j\tilde{Y}_h - r\tilde{W}_j\tilde{Y}_h}{r\tilde{Y}_j\tilde{Y}_h} \leq r & \text{when } 0 < r_{\tilde{Y}_j,\tilde{Y}_h} \\ \frac{r\tilde{W}_j\tilde{Y}_h - r\tilde{W}_j\tilde{Y}_h}{r\tilde{Y}_j\tilde{Y}_h} \leq 0 & \text{when } r_{\tilde{Y}_j,\tilde{Y}_h} = 0 \\ r \leq \frac{r\tilde{W}_j\tilde{Y}_h - r\tilde{W}_j\tilde{Y}_h}{r\tilde{Y}_j\tilde{Y}_h} & \text{when } r_{\tilde{Y}_j,\tilde{Y}_h} < 0 \end{cases}.$$

Further, if $R_{\tilde{Y}_j,\tilde{Y}_h}^2 = 1$ then

$$M_{jh}(r; c) = (1 - c^2)(r - R_{\tilde{W}_j\tilde{Y}_h}^2)^2 \geq 0.$$

Suppose instead that $R_{\tilde{Y}_j,\tilde{Y}_h}^2 \neq 1$. We obtain

$$M_{jh}(r; c) = r^2 R_{\tilde{Y}_j,\tilde{Y}_h}^2 + R_{\tilde{W}_j\tilde{Y}_h}^2 R_{\tilde{W}_j\tilde{Y}_h}^2 - 2r \times r\tilde{Y}_j\tilde{Y}_h r\tilde{W}_j\tilde{W}_h - c^2 r^2 + c^2 r(R_{\tilde{W}_j\tilde{Y}_h}^2 + R_{\tilde{W}_j\tilde{Y}_h}^2) - c^2 R_{\tilde{W}_j\tilde{Y}_h}^2 R_{\tilde{W}_j\tilde{Y}_h}^2$$

$$= r^2(R_{\tilde{Y}_j,\tilde{Y}_h}^2 - c^2) + r[-2r\tilde{Y}_j\tilde{Y}_h r\tilde{W}_j\tilde{Y}_h + c^2(R_{\tilde{W}_j\tilde{Y}_h}^2 + R_{\tilde{W}_j\tilde{Y}_h}^2)] + (1 - c^2)R_{\tilde{W}_j\tilde{Y}_h}^2 R_{\tilde{W}_j\tilde{Y}_h}^2$$

$$= r^2(R_{\tilde{Y}_j,\tilde{Y}_h}^2 - c^2) + r[R_{\tilde{W}_j\tilde{Y}_h}^2(1 - R_{\tilde{Y}_j,\tilde{Y}_h}^2) - (1 - c^2)(R_{\tilde{W}_j\tilde{Y}_h}^2 + R_{\tilde{W}_j\tilde{Y}_h}^2)]$$

$$+ (1 - c^2)R_{\tilde{W}_j\tilde{Y}_h}^2 R_{\tilde{W}_j\tilde{Y}_h}^2,$$

where the last equality makes use of

$$R_{\tilde{W}_j\tilde{Y}_h}^2(1 - R_{\tilde{Y}_j,\tilde{Y}_h}^2) = \frac{R_{\tilde{W}_j\tilde{Y}_h}^2 + R_{\tilde{W}_j\tilde{Y}_h}^2 - 2r\tilde{W}_j\tilde{Y}_h r\tilde{Y}_j\tilde{Y}_h}{1 - R_{\tilde{Y}_j,\tilde{Y}_h}^2}.$$
If \( c^2 = R_{Y_j, \tilde{y}_h}^2 \) then \( M_{jh}(\cdot; c) \) is a linear function

\[
M_{jh}(r; r_{Y_j, \tilde{y}_h}) = r[R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2) - (1 - R_{Y_j, \tilde{y}_h}^2)(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})] \\
+ (1 - R_{Y_j, \tilde{y}_h}^2)R_{W,\tilde{Y}_j}^2 R_{W,\tilde{y}_h} \\
= (1 - R_{Y_j, \tilde{y}_h}^2)\{r[R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2 - (R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})] + R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h}\}
\]

and

\[
0 \leq M_{jh}(r; c) \iff \begin{cases} 
\frac{-R_{W,\tilde{Y}_j}^2 R_{W,\tilde{y}_h}}{R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2 - (R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2} & \text{when } c^2 = R_{Y_j, \tilde{y}_h}^2 \text{ and } R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2 > R_{W,\tilde{Y}_j}^2 + R_{W,\tilde{y}_h}^2 \\
0 \leq (1 - R_{Y_j, \tilde{y}_h}^2)R_{W,\tilde{Y}_j}^2 R_{W,\tilde{y}_h} & \text{when } c^2 = R_{Y_j, \tilde{y}_h}^2 \text{ and } R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2 = R_{W,\tilde{Y}_j}^2 + R_{W,\tilde{y}_h}^2 \\
\frac{-R_{W,\tilde{Y}_j}^2 R_{W,\tilde{y}_h}}{R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2 - (R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2} \leq r & \text{when } c^2 = R_{Y_j, \tilde{y}_h}^2 \text{ and } R_{W,\tilde{Y}_j}^2 + R_{W,\tilde{y}_h}^2 < R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2
\end{cases}
\]

Otherwise, if \( c^2 \neq R_{Y_j, \tilde{y}_h}^2 \), the discriminant of \( M_{jh}(\cdot; c) \) is

\[
\Delta_{jh}(c) = [R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2) - (1 - c^2)(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2 - 4(1 - c^2)(R_{Y_j, \tilde{y}_h}^2 - c^2) R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h} \\
= [R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2) - (1 - c^2)(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2] \\
- (1 - c^2)4R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h}^2 R_{W,\tilde{Y}_j}^2 + 4c^2(1 - c^2) R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h} \\
= [R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2) - (1 - c^2)(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2] \\
- (1 - c^2)[R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2) - (R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2] + 4c^2(1 - c^2) R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h} \\
= c^2 R_{W,(\tilde{Y}_j, \tilde{y}_h)}^4(1 - R_{Y_j, \tilde{y}_h}^2)^2 - c^2(1 - c^2)(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2 + 4c^2(1 - c^2) R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h} \\
= c^2\{R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2)^2 - (1 - c^2)[(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2 - 4R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h}]\} \\
= c^2[R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2)^2 - (1 - c^2)(R_{W,\tilde{Y}_j} - R_{W,\tilde{y}_h})^2].
\]

In particular, \( \Delta_{jh}(c) < 0 \) if and only if

\[
0 < c^2 < 1 - \frac{R_{W,(\tilde{Y}_j, \tilde{y}_h)}^4(1 - R_{Y_j, \tilde{y}_h}^2)^2}{(R_{W,\tilde{Y}_j}^2 - R_{W,\tilde{y}_h}^2)^2}.
\]

Further, we have that

\[
1 - \frac{R_{W,(\tilde{Y}_j, \tilde{y}_h)}^4(1 - R_{Y_j, \tilde{y}_h}^2)^2}{(R_{W,\tilde{Y}_j}^2 - R_{W,\tilde{y}_h}^2)^2} \leq R_{Y_j, \tilde{y}_h}^2 \text{ since if } c^2 = R_{Y_j, \tilde{y}_h}^2 \text{ then}
\]

\[
\Delta_{jh}(c) = [R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2(1 - R_{Y_j, \tilde{y}_h}^2) - (1 - c^2)(R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2] - 4(1 - c^2)(R_{Y_j, \tilde{y}_h}^2 - c^2) R_{W,\tilde{Y}_j} R_{W,\tilde{y}_h} \\
= (1 - R_{Y_j, \tilde{y}_h}^2)^2[R_{W,(\tilde{Y}_j, \tilde{y}_h)}^2 - (R_{W,\tilde{Y}_j} + R_{W,\tilde{y}_h})^2] \geq 0
\]

22
and if \( R^2_{Y_jY_h} = 0 \) then

\[
1 - \frac{R^4_{W,(\tilde{Y}_j,\tilde{Y}_h)}(1 - R^2_{Y_jY_h})^2}{(R^2_{W,Y_j} - R^2_{W,Y_h})^2} = 1 - \frac{(R^2_{W,Y_j} + R^2_{W,Y_h})^2}{(R^2_{W,Y_j} - R^2_{W,Y_h})^2} \leq 0 = R^2_{Y_jY_h}.
\]

It follows that if \( 0 < c^2 < 1 - \frac{R^4_{W,(\tilde{Y}_j,\tilde{Y}_h)}(1 - R^2_{Y_jY_h})^2}{(R^2_{W,Y_j} - R^2_{W,Y_h})^2} \) then \( c^2 < R^2_{Y_jY_h} \) and

\[0 \leq M_{j,h}(r; c) \iff -\infty < r < \infty.
\]

If \( c^2 \neq R^2_{Y_jY_h} \) and \( 0 \leq \Delta_{j,h}(c) \) then define

\[
F_{j,h}(c) \equiv -R^2_{W,(\tilde{Y}_j,\tilde{Y}_h)}(1 - R^2_{Y_jY_h}) + (1 - c^2)(R^2_{W,Y_j} + R^2_{W,Y_h}),
\]

so that \( M_{j,h}(\rho; c) \) has the two roots

\[
\rho_{j,h}^-(c) \equiv \frac{F_{j,h}(c) - \Delta_{j,h}(c)^{\frac{1}{2}}}{2(R^2_{Y_jY_h} - c^2)} \quad \text{and} \quad \rho_{j,h}^+(c) \equiv \frac{F_{j,h}(c) + \Delta_{j,h}(c)^{\frac{1}{2}}}{2(R^2_{Y_jY_h} - c^2)}.
\]

We then have that

\[0 \leq M_{j,h}(r; c) \iff \begin{cases} r \in (-\infty, \rho_{j,h}^-(c)] \cup [\rho_{j,h}^+(c), \infty) & \text{when } c^2 < R^2_{Y_jY_h} \\ r \in [\rho_{j,h}^-(c), \rho_{j,h}^+(c)] & \text{when } R^2_{Y_jY_h} < c^2 \end{cases}.
\]

Combining these results, yields the equivalence between \( 0 \leq M_{j,h}(r; c) \) and the range of \( r \).

The sharp bounds \( D^{k,\tau,c} \), \( B^{k,\tau,c} \), and \( G^{k,\tau,c} \) follow from the mappings \( D(\cdot) \), \( B(\cdot) \), and \( G(\cdot) \) in Theorem 3.1.
Figure 4: 50% (dark) and 95% (light) confidence regions for $\beta_{j1}$ (cash flow) for $j = 1, 2, 3$ (investment, saving, and debt) from year 1970 to 2017, when $X$ includes asset tangibility.

We consider the regions $B_{j1}$, $B_{j1}^{c*}$, and $B_{j1}^{c*,\tau,c}$ where $c^* = 0$, $\kappa$ and $\tau$ are such that $\hat{\kappa}^* = 0.5$ and $\hat{\tau}^* = (0.9, 0.9, 0.9)'$, and $c$ is such that $(\zeta_{12}, \zeta_{12}) = (\zeta_{23}, \zeta_{23}) = (-1, 0)$ and $(\zeta_{13}, \zeta_{13}) = (0, 1)$. The shaded vertical bars indicate years in which the maintained assumptions are rejected.
Table 7: Bounds on the Cash Flow Coefficients in the Investment, Saving, and Debt Equations Using the Full Panel and Accounting for Asset Tangibility

<table>
<thead>
<tr>
<th>$S^k_{1,t}$</th>
<th>$J^k_{1,t}$</th>
<th>$J^k_{1,t},c$</th>
<th>$J^k_{1,t},c^*$</th>
<th>$b_{Y,(W,X')}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Results without fixed effects for $\kappa = \infty$ and $\tau = (1,1,1)'$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>$[-0.172, 0.145]$</td>
<td>$[-0.049, 0.145]$</td>
<td>$[0.118, 0.146]$</td>
<td>$[0.118, 0.141]$</td>
</tr>
<tr>
<td>&amp; $(-0.213, 0.151)$ &amp; $(-0.074, 0.151)$ &amp; $(0.111, 0.152)$ &amp; $(0.110, 0.148)$ &amp; $(0.136, 0.150)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>$[-0.540, 0.124]$</td>
<td>$[-0.029, 0.124]$</td>
<td>$[0.102, 0.124]$</td>
<td>$[0.101, 0.121]$</td>
</tr>
<tr>
<td>&amp; $(-0.621, 0.130)$ &amp; $(-0.048, 0.130)$ &amp; $(0.094, 0.131)$ &amp; $(0.094, 0.128)$ &amp; $(0.114, 0.129)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>$[-0.424, 1.865]$</td>
<td>$[-0.424, -0.224]$</td>
<td>$[-0.425, -0.390]$</td>
<td>$[-0.421, -0.389]$</td>
</tr>
<tr>
<td>&amp; $(-0.438, 2.229)$ &amp; $(-0.424, -0.189)$ &amp; $(-0.442, -0.372)$ &amp; $(-0.438, -0.371)$ &amp; $(-0.436, -0.400)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em><em>Results without fixed effects for $\hat{\kappa}^</em> = 0.5$ and $\hat{\tau}^</em> = (0.9,0.9,0.9)'$**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>$[0.125, 0.145]$</td>
<td>$[0.125, 0.145]$</td>
<td>$[0.124, 0.146]$</td>
<td>$[0.124, 0.141]$</td>
</tr>
<tr>
<td>&amp; $(0.119, 0.151)$ &amp; $(0.119, 0.151)$ &amp; $(0.117, 0.152)$ &amp; $(0.117, 0.148)$ &amp; $(0.136, 0.150)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>$[0.106, 0.124]$</td>
<td>$[0.106, 0.124]$</td>
<td>$[0.106, 0.124]$</td>
<td>$[0.106, 0.121]$</td>
</tr>
<tr>
<td>&amp; $(0.101, 0.130)$ &amp; $(0.101, 0.130)$ &amp; $(0.099, 0.131)$ &amp; $(0.099, 0.128)$ &amp; $(0.114, 0.129)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>$[-0.424, -0.396]$</td>
<td>$[-0.424, -0.396]$</td>
<td>$[-0.425, -0.395]$</td>
<td>$[-0.421, -0.395]$</td>
</tr>
<tr>
<td>&amp; $(-0.438, -0.382)$ &amp; $(-0.424, -0.378)$ &amp; $(-0.442, -0.372)$ &amp; $(-0.438, -0.371)$ &amp; $(-0.436, -0.400)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Results with year and firm fixed effects for $\kappa = \infty$ and $\tau = (1,1,1)'$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>$[-0.719, 0.132]$</td>
<td>$[-0.549, 0.132]$</td>
<td>$[-0.565, 0.132]$</td>
<td>$[-0.564, -0.154]$</td>
</tr>
<tr>
<td>&amp; $(-0.758, 0.138)$ &amp; $(-0.581, 0.138)$ &amp; $(-0.606, 0.139)$ &amp; $(-0.605,-0.136)$ &amp; $(0.122, 0.137)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>$[-2.913, 0.174]$</td>
<td>$[-0.329, 0.174]$</td>
<td>$[-0.342, 0.174]$</td>
<td>$[-0.341, -0.031]$</td>
</tr>
<tr>
<td>&amp; $(-3.142, 0.182)$ &amp; $(-0.369, 0.182)$ &amp; $(-0.392, 0.184)$ &amp; $(-0.391, -0.006)$ &amp; $(0.159, 0.181)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>$[-0.374, \infty]$</td>
<td>$[-0.374, -0.288]$</td>
<td>$[-0.374, -0.285]$</td>
<td>$[-0.374, -0.285]$</td>
</tr>
<tr>
<td>&amp; $(-\infty, \infty)$ &amp; $(-0.388, -0.237)$ &amp; $(-0.402, -0.219)$ &amp; $(-0.402, -0.220)$ &amp; $(-0.383, -0.353)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em><em>Results with year and firm fixed effects for $\hat{\kappa}^</em> = 0.5$ and $\hat{\tau}^</em> = (0.9,0.9,0.9)'$**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>$[0.081, 0.132]$</td>
<td>$[0.081, 0.132]$</td>
<td>$[0.081, 0.132]$</td>
<td>$-0.129$</td>
</tr>
<tr>
<td>&amp; $(0.075, 0.138)$ &amp; $(0.075, 0.138)$ &amp; $(0.073, 0.139)$</td>
<td>$- (0.122, 0.137)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>$[0.133, 0.174]$</td>
<td>$[0.133, 0.174]$</td>
<td>$[0.132, 0.174]$</td>
<td>$-0.170$</td>
</tr>
<tr>
<td>&amp; $(0.124, 0.182)$ &amp; $(0.124, 0.182)$ &amp; $(0.121, 0.184)$</td>
<td>$- (0.159, 0.181)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>$[-0.374, -0.358]$</td>
<td>$[-0.374, -0.358]$</td>
<td>$[-0.374, -0.358]$</td>
<td>$-0.368$</td>
</tr>
<tr>
<td>&amp; $(-0.385, -0.346)$ &amp; $(-0.385, -0.346)$</td>
<td>$(-0.388, -0.342)$</td>
<td>$(-0.383, -0.353)$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sample is an unbalanced panel of 161,959 firm-year observations. $Y_1$, $Y_2$, and $Y_3$ denote Investment, Saving, and Debt respectively and $X = [\text{Cash Flow}, \text{Firm Size}, \text{Asset Tangibility}]$. When year fixed effects are included, $X$ also includes year indicator variables. When firm fixed effects are included, the equations' variables undergo a within transformation. $c$ sets $(c_{12},c_{12}) = (c_{23},c_{23}) = (-1,0)$ and $(c_{13},c_{13}) = (0,1)$ whereas $c^* = 0$. Robust standard errors for $\pi$ are clustered by firm. 50% and 95% confidence regions are in brackets and parentheses respectively.