independent, real valued, nondegenerate random variables with unknown distributions. The unknown constant c is real valued, finite, and nonzero.

ASSUMPTION A3: Either the median or the mean of U, V, or W is zero. The characteristic functions of U, V, and W do not vanish.

Assumption A3 is mainly used for identification of the distributions of U, V, and W, not for the identification of c. Kotlarski's Lemma requires some location normalization, as in Assumption A3. Evdokimov and White (2012) provide alternative conditions under which Kotlarski's Lemma holds even when the characteristic functions of U, V, and/or W can have zeros.

Kotlarski's Lemma assumes $c = 1$. We assume $c \neq 0$ because, if $c = 0$ then trivially we can only identify the distributions of W and of $V + U$. Moreover, we can immediately tell if $c = 0$, because in that case the distributions of X and Y will be independent.

For any random variables R and S, let $\sigma_R^2 = var(R)$ if this variance exists, and let $\sigma_{RS} = cov(R, S)$ if this covariance exists. Also let $\psi_R(t) = \ln E\left[\exp(itR)\right]$, the log characteristic function (also known as the cumulant generating function) of R, and similarly $\psi_{R,S}(t_1, t_2) = \ln E\left[\exp(it_1R + it_2S)\right]$.

We begin with a tiny Lemma:

LEMMA 1: Let Assumptions A1, A2, and A3 hold. If the constant c is point identified, then the distributions of U, V, and W are all also point identified.

equations give a bound on c (it must lie between zero and the coefficient of $t_1 t_2$), but this bound is tightened below.

Lemma 1 and Theorem 1 together show how to tell if V is normal or not, and show that Kotlarski's Lemma extends to point identification with an unknown factor loading c as long as V is non-normal.

Now consider the case where V is normal. For this case, we need some more notation. For a random variable R, define R's "largest normal factor" to be the variable $\widetilde{R}$ having the maximum variance such that $R = \widetilde{R} + \overline{R}$, where $\widetilde{R}$ and $\overline{R}$ are independently distributed and $\widetilde{R}$ is normally distributed. Without loss of generality, assume $\widetilde{R}$ has mean zero. Call $\overline{R}$ the non-normal factor. If no normal $\widetilde{R}$ exists, then R does not have a normal factor, and in this case we can let $\widetilde{R} = 0$ and $\overline{R} = R$. If R is normal then $\widetilde{R} = R - E(R)$ and $\overline{R} = E(R)$. See Schennach and Hu (2013) and Lewbel, Schennach, and Zhang (2020) for a similar use of normal factors. Reiersøl (1950) calls a normal factor a normal divisor.

Given a random variable R, the variance of $\widetilde{R}$ can be determined by

$$\sigma^2_{\widetilde{R}} = \sup\left\{\sigma^2 \in \mathbb{R}^+ : \phi_R(t) + t^2\sigma^2/2 \text{ is a log characteristic function}\right\}$$

If $\sigma^2_{\widetilde{R}} = 0$ then R does not have a normal factor, otherwise, $\sigma^2_{\widetilde{R}}$ given by this expression is the variance of the largest normal factor $\widetilde{R}$. This follows immediately from the definition of a characteristic function, since a positive $\sigma^2_{\widetilde{R}}$ means by construction that R equals the convolution of two independent random variables, one of which has the log characteristic function of a mean zero normal.[2] This means that if R has a known distribution, and hence a known characteristic function, we can determine if it has a normal factor or not, and we can point identify the distributions of $\widetilde{R}$ and $\overline{R}$.

THEOREM 2: Let Assumptions A1, A2, and A3 hold. Assume V is normally distributed. Then $\sigma_{\widetilde{X}\widetilde{Y}}$, $\sigma^2_{\widetilde{X}}$, and $\sigma^2_{\widetilde{Y}}$ are identified. If $\sigma_{\widetilde{X}\widetilde{Y}}/\sigma^2_{\widetilde{X}} = \sigma^2_{\widetilde{Y}}/\sigma_{\widetilde{X}\widetilde{Y}}$ then c is point identified by $c = \sigma_{\widetilde{X}\widetilde{Y}}/\sigma^2_{\widetilde{X}} = \sigma^2_{\widetilde{Y}}/\sigma_{\widetilde{X}\widetilde{Y}}$ and in this case neither W nor U have a normal factor. Otherwise, c is interval identified by $c \in \left[\sigma_{\widetilde{X}\widetilde{Y}}/\sigma^2_{\widetilde{X}}, \sigma^2_{\widetilde{Y}}/\sigma_{\widetilde{X}\widetilde{Y}}\right]$, and for each value of c in this interval, there is a corresponding, identified unique distribution for U, V, and W. This interval bound on c is sharp.

The fact that c is point identified when neither W nor U have a normal factor also appears in Reiersøl (1950). The identified sets in Theorem 2 are new, but are closely related to the Frisch (1934) bounds on mismeasured linear regressions. Taken together, Lemma 1, Theorem 1, and Theorem 2 completely characterize the identification of our model.

Proof of Theorem 2: Separating Y and X into their normal and non-normal factors, we have $Y = \widetilde{Y} + \overline{Y}$ and $X = \widetilde{X} + \overline{X}$. Similarly, Separating W and U into normal and non-normal factors, we also have $Y = cV + \widetilde{W} + \overline{W}$ and $X = V + \widetilde{U} + \overline{U}$. When V is normal, this implies $\widetilde{Y} = cV + \widetilde{W}$, $\overline{Y} = \overline{W}$, $\widetilde{X} = V + \widetilde{U}$ and $\overline{X} = \overline{U}$. This in turn means that, with V

[2] An explicit mathematical expression for "being a characteristic function" and hence defining $\sigma^2_{\widetilde{R}}$ can be obtained from Bochner's Theorem, e.g., Theorem 4.2.2 in Lukacs (1970).

normal, $\overline{X}$ and $\overline{Y}$ are independent of each other and of the joint distribution of $\widetilde{Y}$ and $\widetilde{X}$. Since the marginal distributions of $\overline{Y}$ and $\overline{X}$ are identified, we can identify the left side of

$$\psi_{Y,X}(t_1,t_2) / \psi_{\overline{Y}}(t_1) / \psi_{\overline{X}}(t_2) = \psi_{\widetilde{Y},\widetilde{X}}(t_1,t_2)$$

And therefore the joint normal distribution of the mean zero variables $\widetilde{Y}$ and $\widetilde{X}$ is identified. In particular, this means that $\sigma^2_{\widetilde{Y}}$, $\sigma^2_{\widetilde{X}}$, and $\sigma_{\widetilde{X}\widetilde{Y}}$ are identified.

The remaining step now borrows heavily from the Frisch (1934) bounds on mismeasured linear regression. From the identified second moments of $\widetilde{Y}$ and $\widetilde{X}$, we have $\sigma^2_{\widetilde{Y}} = c^2\sigma^2_V + \sigma^2_{\widetilde{W}}$, $\sigma^2_{\widetilde{X}} = \sigma^2_V + \sigma^2_{\widetilde{U}}$, and $\sigma_{\widetilde{X}\widetilde{Y}} = c\sigma^2_V$, which provides three equations in the four unknown constants $\sigma^2_{\widetilde{U}}$, $\sigma^2_{\widetilde{W}}$, σ^2_V, and c. The only constraints on these parameter values are that $c \neq 0$, $\sigma^2_{\widetilde{U}}$ and $\sigma^2_{\widetilde{W}}$ must be non-negative (either can be zero if the corresponding normal factor doesn't exist), and σ^2_V must be positive. These being the only constraints is what makes the corresponding bounds be sharp. The equation $\sigma_{\widetilde{X}\widetilde{Y}} = c\sigma^2_V$ means that the sign of c equals the sign of $\sigma_{\widetilde{X}\widetilde{Y}}$ to ensure $\sigma^2_V > 0$. Then $\sigma^2_{\widetilde{U}} \geq 0$ requires $\sigma^2_{\widetilde{X}} - \sigma_{\widetilde{X}\widetilde{Y}}/c \geq 0$ and $\sigma^2_{\widetilde{W}} \geq 0$ requires $\sigma^2_{\widetilde{Y}} - c\sigma_{\widetilde{X}\widetilde{Y}} \geq 0$. Therefore, either $\sigma_{\widetilde{X}\widetilde{Y}} > 0$ and $\sigma_{\widetilde{X}\widetilde{Y}}/\sigma^2_{\widetilde{X}} \leq c \leq \sigma^2_{\widetilde{Y}}/\sigma_{\widetilde{X}\widetilde{Y}}$, or $\sigma_{\widetilde{X}\widetilde{Y}} < 0$ and $\sigma^2_{\widetilde{Y}}/\sigma_{\widetilde{X}\widetilde{Y}} \leq c \leq \sigma_{\widetilde{X}\widetilde{Y}}/\sigma^2_{\widetilde{X}}$. Either way c lies in the interval between $\sigma_{\widetilde{X}\widetilde{Y}}/\sigma^2_{\widetilde{X}}$ and $\sigma^2_{\widetilde{Y}}/\sigma_{\widetilde{X}\widetilde{Y}}$,

Frisch, R. (1934), "Statistical confluence analysis by means of complete regression systems," Vol. 5, Universitetets Økonomiske Instituut.

Kotlarski, I. I. (1967), "On characterizing the gamma and normal distribution," Pacific Journal of Mathematics, 20, 69–76.

Lewbel, A. (1997), "Constructing Instruments for Regressions With Measurement Error When No Additional Data are Available, With an Application to Patents and R&D," Econometrica, 65(5), 1201-1213.

Lewbel, A., S. M. Schennach, and L. Zhang (2020), "Identification of a Triangular Two Equation System Without Instruments," Centre for microdata methods and practice working paper CWP41/20.

Li, S. and X. Zheng, (2019), "A Generalization of Lemma 1 in Kotlarski (1967)," Rice University unpublished manuscript.

Li, T. and Q. Vuong (1998), “Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators,” Journal of Multivariate Analysis, 65, 139–165.

Lukacs, E. (1970), “Characteristic Functions”, Second edition, Griffin, London.

Navarro, S. and J. Zhou, (2017) "Identifying agent’s information sets: An application to a lifecycle model of schooling, consumption and labor supply," Review of Economic Dynamics, 25, 58-92.

Reiersøl, O. (1950), “Identifiability of a linear relation between variables which are subject to error,” Econometrica, 18, 375-389.

Schennach, S. M. and Y. Hu (2013), "Nonparametric Identification and Semiparametric Estimation of Classical Measurement Error Models Without Side Information," Journal of the American Statistical Association, 108, 177-186.

GJ Székely, G. J. and C. R. Rao (2000), "Identifiability of distributions of independent random variables by linear combinations and moments" Sankhyā: The Indian Journal of Statistics, Series A, 62(2), 193-202.