Another Look at the Identification at Infinity of Sample Selection Models

Xavier d'Haultfoeuille
Arnaud Maurel

I Z A

Another Look at the Identification at Infinity of Sample Selection Models

Xavier d'Haultfoeuille
CREST-INSEE
and Université Paris I-Panthéon-Sorbonne

Arnaud Maurel
CREST-INSEE,
ENSAE, PSE and IZA

ABSTRACT

Another Look at the Identification at Infinity of Sample Selection Models[*]

It is often believed that without instrument, endogenous sample selection models are identified only if a covariate with a large support is available (see Chamberlain, 1986, and Lewbel, 2007). We propose a new identification strategy mainly based on the condition that the selection variable becomes independent of the covariates when the outcome, not one of the covariates, tends to infinity. No large support on the covariates is required. Moreover, we prove that this condition is testable. We finally show that our strategy can also be applied to the identification of generalized Roy models.

Corresponding author:

Arnaud Maurel
ENSAE
Bureau E12 - Timbre J120
3, Avenue Pierre Larousse
92245 Malakoff cedex
France
E-mail: maurel@ensae.fr

---

1 Introduction

Since the seminal work of Heckman (1974), the issue of endogenous selection has been an active topic of research in both applied and theoretical econometrics (see Vella, 1998, for a survey). The usual strategy to deal with this issue is to rely on instruments that determine selection but not the outcome. However, the search of a valid instrument may be difficult if not impossible in some applications. Another strategy, which has been sometimes advocated, relies on the fact that, loosely speaking, the selection problem becomes negligible "at the limit". Following this idea, Chamberlain (1986) proved that the effects of covariates on an outcome are identified under the linearity of the model and a large support assumption on at least one covariate.Lewbel (2007) generalized this result by proving that identification can be achieved without imposing any structure on the outcome equation, provided that a special regressor has a large support and under restrictions on the selection equation.[1]

The main drawback of the latter approach is that it requires the existence of a covariate with a large support. Thus, it breaks down when all covariates are discrete, a case which is fairly common in practice. In this paper, we consider another route for identifying the model at infinity. Intuitively, if selection is truly endogenous, then we can expect the effect of the outcome on the selection variable to dominate those of the covariates for large values of the outcome. Following this idea, our main identifying condition states that the selection variable is independent of the covariates at the limit, i.e., when the outcome tends to infinity. Under this condition, the model is identified without any large support condition on these covariates. Only an additive decomposition and a mild restriction on the residuals are required. Moreover, we show that the main condition is testable. Apart from the standard selection model, we apply our result to a generalization of the Roy model (1951) of self-selection accounting for non-pecuniary factors. We show that, in this framework, the effects of covariates on the outcomes are identified without exclusion restriction under a moderate dependence condition on the residuals.

The note is organized as follows. Section 2 presents the model and establishes the main identification result. Section 3 proves the testability of our main condition. Section 4 applies this result to generalized Roy models, and Section 5 concludes.

---

[1]These restrictions entail that the probability of selection tends to zero or one when the special regressor tends to infinity.

2 Main result

Let Y^* denote the outcome of interest, X denote covariates and D denote the dummy of selection. Let us consider the following selection model:

$$\begin{cases} Y^* &=& \psi(X) + \sigma(X)\varepsilon \\ D &=& \mathbb{1}\{g(X, Y^*) - \eta \geq 0\}. \end{cases} \tag{2.1}$$

The econometrician observes D, $Y = DY^*$ and X. Without loss of generality, we suppose that $\eta \perp\!\!\!\perp (X, Y^*)$ and $\eta \sim U[0,1]$. In this case, $g(X, Y^*) = P(D = 1|X, Y^*)$. We also make the innocuous normalizations that $\psi(x_0) = 0$ and $\sigma(x_0) = 1$ for a given $x_0 \in \text{Supp}(X)$ (where $\text{Supp}(T)$ denotes the support of the random variable T).[2] Our main result is based on the following assumptions.

Assumption 1 *(Additive decomposition)* $X \perp\!\!\!\perp \varepsilon$.

Assumption 2 *(Tails of the residual)* $\sup(Supp(\varepsilon)) = +\infty$. *Moreover, there exists $\beta > 0$ such that $E(\exp(\beta\varepsilon)) < \infty$.*

Assumption 3 *(Independence at infinity)* *There exists $l > 0$ such that for all $x \in Supp(X)$, $\lim_{y\to\infty} g(x, y) = l$.*

Assumption 1 is usual in selection models and weaker than Chamberlain (1986)'s condition, since heteroskedasticity is allowed for here. Assumption 2 puts some weak restrictions on the tails of the distributions of ε. In the example of a wage equation where Y^* denotes the logarithm of the wage W, it is satisfied if $E[W^\beta] < \infty$ for a given $\beta > 0$.[3] Thus, it holds even if wages have very fat tails, Pareto-like for instance. Assumption 3 is the main condition here. It requires the probability of selection to be independent of X at the limit, i.e., for those who have a very large outcome. It holds for instance if the selection model is additive in Y^*, i.e., $D = \mathbb{1}\{Y^* + h(X) \geq \eta\}$, with $\eta \perp\!\!\!\perp (X, Y^*)$. In this latter case, $l = 1$, but this condition is not necessary in general. It would also hold in a more general model with $D = U\mathbb{1}\{Y^* + h(X) \geq \eta\}$ where $U \in \{0, 1\} \perp\!\!\!\perp (X, Y^*, \eta)$ is a random shock. For instance, such a framework is well suited to model participation to the labor market, with

---

[2]Unlike Heckman (1990) and Andrews & Schafgans (1998), we do not seek to identify the intercept of the model, which corresponds here to $E(\varepsilon)$. We conjecture that in our context, the intercept cannot be identified without further restriction.

[3]We also suppose that ε is unbounded. Identification is still possible otherwise, by using support variation. We do not consider this case here since it seems less relevant in practice, and does not really rely on our main condition, i.e., Assumption 3.

U denoting in that case an unobserved random shock related e.g. to health conditions that could prevent individuals from entering the labor market.

Theorem 2.1 *Under Assumptions 1-3, $\psi(.)$ and $\sigma(.)$ are identified.*

Proof: Subsequently, S_T denotes the survival function of the random variable T. Besides, we use the notation $f(y) \sim g(y)$ if there exists $r(.)$ such that $f(y) = g(y)(1 + r(y))$ with $\lim_{y \to \infty} r(y) = 0$. The result is based on the following lemma.

Lemma 2.1 *Let T be a real random variable such that $\sup(Supp(T)) = +\infty$ and $E(|T|) < \infty$. Suppose also that when $y \to \infty$, $S_T(y) \sim S_T(lf(y))$, where $\lim_{y \to \infty} f'(y) = 1$ and $l > 0$. Then $l = 1$.*

Proof of Lemma 2.1: Suppose that $l > 1$. Then there exists $\eta > 0$ such that $l > 1 + \eta$. Moreover, by assumption, there exists y_0 such that for all $y \geq y_0$,

$$S_T(y) < (1 + \eta)S_T(lf(y)).$$

Besides, $E(|T|) < \infty$ implies that $\int_0^\infty S_T(u)du < \infty$. Thus, for all $y > y_0$,

$$\int_y^\infty S_T(u)du < (1 + \eta) \int_y^\infty S_T(lf(u))du. \tag{2.2}$$

By assumption, the derivative of the function $g(y) = lf(y)$ tends to $l > 1$ when $u \to \infty$. Thus, there exists y_1 such that $g(y) > y$ and $g'(y) > 1 + \eta$ for all $y \geq y_1$. Hence, for all $y \geq y_1$, g is one to one and

$$
\begin{aligned}
\int_y^\infty S_T(lf(u))du &= \int_{g(y)}^\infty \frac{S_T(v)}{g'(g^{-1}(v))}dv \\
&< \frac{1}{1 + \eta} \int_{g(y)}^\infty S_T(v)dv \\
&< \frac{1}{1 + \eta} \int_y^\infty S_T(v)dv. \tag{2.3}
\end{aligned}
$$

Inequalities (2.2) and (2.3) imply that $\int_y^\infty S_T(u)du < \int_y^\infty S_T(u)du$ for all $y \geq \max(y_0, y_1)$, a contradiction. Similarly, one can show that $l < 1$ is impossible. Thus $l = 1$. $\square$

Now let us prove Theorem 2.1. Let $q(y, x) = P(D = 1, Y \geq y | X = x)$. We have

$$q(y, x) = \int_y^\infty g(x, u)dP^{Y^*|X=x}(u)$$

4

By Assumption 3, as $u \to \infty$, we have $g(x, u) \to l > 0$. Thus, using standard results on integrals, we get as $y \to \infty$,

$$q(y, x) \sim l P(Y^* \geq y | X = x).$$

By Assumption 1, $P(Y \geq y | X = x) = S_\varepsilon((y - \psi(x))/\sigma(x))$, where $S_\varepsilon(.)$ denotes the survival function of ε. Thus,

$$q(y, x) \sim l S_\varepsilon \left(\frac{y - \psi(x)}{\sigma(x)} \right). \tag{2.4}$$

Similarly,

$$q(y, x_0) \sim l S_\varepsilon(y) \tag{2.5}$$

In other words,

$$q(y, x) \sim q \left(\frac{y - \psi(x)}{\sigma(x)}, x_0 \right) \tag{2.6}$$

Now, let us show that actually, as $y \to \infty$, for all $s > 0$ and $u \in \mathbb{R}$,

$$q(y, x) \sim q(sy + u, x_0) \implies (s, u) = \left(\frac{1}{\sigma(x)}, -\frac{\psi(x)}{\sigma(x)} \right) \tag{2.7}$$

Because the function q is identified, this implies that $\sigma(x)$ and $\psi(x)$ are identified. If $q(y, x) \sim q(sy + u, x_0)$, then by (2.4) and (2.5),

$$S_\varepsilon(t(y + v)) \sim S_\varepsilon(y), \tag{2.8}$$

where $t = s\sigma(x)$ and $v = (1/\sigma(x))(\psi(x) + u/s)$. Besides, by Assumption 2, $\sup(\mathrm{Supp}(\varepsilon)) = +\infty$ and $E(|\varepsilon|) < \infty$. Thus, by Lemma 2.1, $t = 1$, i.e. $s = 1/\sigma(x)$. Thus, $\sigma(x)$ is identified. Besides, by (2.8),

$$S_{e^{\beta\varepsilon}}(wy) \sim S_{e^{\beta\varepsilon}}(y),$$

where β is defined in Assumption 2 and $w = \exp(\beta v)$. Because $E(\exp(\beta\varepsilon)) < \infty$, we can apply Lemma 2.1 once more. This yields $w = 1$, or $u = -\psi(x)/\sigma(x)$. Thus, $\psi(x)$ is identified. $\square$

3 Testability

The main identifying condition in the setting above is Assumption 3, so one may wonder whether this assumption is refutable or not. The answer turns out to be affirmative. Indeed, together with Assumption 1, this condition implies (2.6), which can be stated as

$$\forall x \in \mathrm{Supp}(X), \; \exists(s(x), u(x)) \in \mathbb{R}^{*+} \times \mathbb{R} : \; q(y, x) \sim q(s(x)y - u(x), x_0). \tag{3.1}$$

This condition can be directly tested in the data. The following theorem shows that actually, under a slight reinforcement of Assumption 2 and another mild condition, Condition (3.1) and Assumption 3 are equivalent.

Theorem 3.1 *Suppose that Assumption 1 holds, ε has unbounded support, there exists $\alpha > 1$, $\beta > 0$ such that $E[\exp(\beta|\varepsilon|^\alpha)] < \infty$ and there exists $l(x) > 0$ such that*

$$\lim_{y \to \infty} g(x, y) = l(x). \tag{3.2}$$

Then Assumption 3 is equivalent to Condition (3.1).

Proof: We shall first prove a result similar to the one of Lemma 2.1.

Lemma 3.1 *Let T be a real random variable such that $\sup(Supp(T)) = +\infty$ and $E(|T|) < \infty$. Suppose also that when $y \to \infty$, $S_T(y) \sim lS_T(f_\delta(y))$, where $l > 0$ and $f_\delta(.)$ is strictly increasing for y large enough and satisfies (i) $f'_\delta(y) \to 0$ if $\delta < 0$, (ii) $f'_0(y) \to C > 0$ and (iii) $f'_\delta(y) \to \infty$ if $\delta > 0$. Then $\delta = 0$. Moreover, if $f_0(y) = y$, then $l = 1$.*

Proof of Lemma 3.1: Suppose that $\delta > 0$. By assumption, there exists $l' > 0$ and y_0 such that for all $y \geq y_0$,

$$S_T(y) < l'S_T(f_\delta(y)). \tag{3.3}$$

Besides, there exists y_1 such that $f_\delta(.)$ is one to one on $[y_1, \infty)$, with $f'_\delta(y) > l'$ and $f_\delta(y) > y$ for all $y \geq y_1$. Thus, for all $y \geq y_1$,

$$
\begin{aligned}
\int_y^\infty S_T(f_\delta(u))du &= \int_{f_\delta(y)}^\infty \frac{S_T(v)}{f'_\delta(f_\delta^{-1}(v))}dv \\
&< \frac{1}{l'}\int_{f_\delta(y)}^\infty S_T(v)dv \\
&< \frac{1}{l'}\int_y^\infty S_T(v)dv. \tag{3.4}
\end{aligned}
$$

Inequalities (3.3) and (3.4) imply that $\int_y^\infty S_T(u)du < \int_y^\infty S_T(u)du$ for all $y \geq \max(y_0, y_1)$, a contradiction. The proof that $\delta < 0$ is impossible follows similarly. Thus $\delta = 0$. Finally, if $f_0(y) = y$, then $S_T(y) \sim lS_T(y)$, which implies directly that $l = 1$. $\square$

Now let us prove Theorem 3.1. By the proof of Theorem 2.1, Assumption 3 implies Condition (3.1). Thus, it suffices to prove that Condition (3.1) implies Assumption 3. For all $x \in \text{Supp}(X)$, by a similar reasoning as in the previous proof,

$$q(y, x) \sim l(x)S_\varepsilon\left(\frac{y - \psi(x)}{\sigma_{(x)}}\right).$$

6

The same holds for $q(y, x_0)$. Thus, by Condition (3.1), there exists $s > 0$ and $u \in \mathbb{R}$ such that

$$S_\varepsilon(y) \sim lS_\varepsilon(sy + u), \tag{3.5}$$

where $l = l(x)/l(x_0)$. This implies that

$$S_{\exp(\beta\varepsilon)}(y) \sim lS_{\exp(\beta\varepsilon)}(\exp(\beta u)y^s).$$

By assumption, $E[\exp(\beta\varepsilon)] < \infty$. Thus, by applying Lemma 3.1 to $f_\delta(y) = \exp(\beta u)y^{\exp(\delta)}$ (with $\delta = \ln s$), we get $s = 1$. Hence, by (3.5),

$$S_{\exp(\beta\varepsilon^\alpha)}(\exp(\beta y^\alpha)) \sim lS_{\exp(\beta\varepsilon^\alpha)}(\exp(\beta(y + u)^\alpha)).$$

After some manipulations, we obtain

$$S_{\exp(\beta\varepsilon^\alpha)}(y) \sim lS_{\exp(\beta\varepsilon^\alpha)}(f_u(y)),$$

where

$$f_u(y) = y^{\left(1 + u\left(\frac{\beta}{\ln y}\right)^{1/\alpha}\right)^\alpha}.$$

Some computations show that f_u is strictly increasing for y large enough and (i) $f_u'(y) \to 0$ if $u < 0$, (ii) $f_0(y) = y$ and (iii) $f_u'(y) \to \infty$ if $u > 0$. Thus, by Lemma 3.1, $u = 0$ and $l = 1$. In other terms, $l(x) = l(x_0)$ for all $x \in \text{Supp}(X)$, which proves that Assumption 3 holds. □

Consider for instance the case where selection is exogenous and $x \mapsto g(x, y) = P(D = 1 | X = x)$ is nonconstant. In this case, Condition (3.2) is satisfied with $l(x) = P(D = 1 | X = x)$. Thus, by Theorem 3.1, Condition (3.1) fails to hold, and one is able to reject the "independence at infinity" assumption.

4   Application to generalized Roy Models

Let us consider a class of generalized Roy models where each individual chooses the sector $D \in \{0, 1\}$ that provides him with the higher utility. Suppose that the utility U_i associated with each sector $i \in \{0, 1\}$ is the sum of the log-earnings[4] $\ln Y_i = \psi_i(X) + \varepsilon_i$ and a random nonpecuniary component $G_i(X) + \eta_i$. Thus, $D = \mathbb{1}\{\ln Y_1 > \ln Y_0 + G(X) + \eta\}$ with $G(X) = G_0(X) - G_1(X)$ and $\eta = \eta_0 - \eta_1$, and the econometrician only observes $\ln Y = D \ln Y_1 + (1 - D) \ln Y_0$, as well as D and X. Without loss of generality, we assume

---

[4]For the sake of simplicity, we shall consider an homoskedastic model on the outcome in the following.

that there exists $x_0 \in \text{Supp}(X)$ such that $\psi_0(x_0) = \psi_1(x_0) = 0$. The standard Roy model, in which the chosen sector is the one yielding the higher earnings, corresponds to $\eta = 0$ and $G(X) = 0$. This framework also encompasses Heckman (1974)'s model of labor market participation. In this latter case, Y_1 corresponds to the logarithm of the potential wage, $G_1(X) = \eta_1 = 0$, $\ln Y_0 = 0$ and $G_0(X)$ (resp. η_0) is the observable (resp. unobservable) part of the logarithm of the reservation wage.

The generalized Roy models we consider in this section can be used in a broad range of economic settings. Basically, these models are well suited for most of the situations in which self-selection between two alternatives is driven both by the relative pecuniary and non-pecuniary returns. This framework can be used for instance to model the decision to attend higher education after graduating from high school, thus extending Willis & Rosen (1979) by accounting for non-pecuniary factors entering the schooling decision (see, e.g., Carneiro et al., 2003, and D'Haultfœuille & Maurel, 2009). Other examples of applications include occupational choice (see, e.g., Dagsvik & Strøm, 2006 for the choice between private and public sector) as well as migration decisions (see, e.g., Borjas, 1987 and Bayer et al., 2008) accounting for non-pecuniary factors.[5] Theorem 2.1 can be applied to provide identification of (ψ_0, ψ_1) without exclusion restriction nor any large support condition on the covariates, as the following result shows.

Corollary 4.1 *Suppose that $(\varepsilon_0, \varepsilon_1, \eta) \perp\!\!\!\perp X$, ε has unbounded support and there exists $\beta_0, \beta_1 > 0$ such that $E[\exp(\beta_i \varepsilon_i)] < \infty$ for $i \in \{0, 1\}$ and*

$$\lim_{u \to \infty} P(\varepsilon_i + (1 - 2i)\eta \le a + u | \varepsilon_{1-i} = u) = l_{1-i} > 0 \tag{4.1}$$

for all $a \in \mathbb{R}$ and $i \in \{0, 1\}$. Then ψ_0 and ψ_1 are identified.

Proof: Since $(\varepsilon_0, \varepsilon_1, \eta) \perp\!\!\!\perp X$, Condition (4.1) implies that

$$\lim_{u \to \infty} P(\ln Y_1 > \ln Y_0 + G(X) + \eta | \ln Y_1 = u, X = x) = l_1.$$

In other words,

$$\lim_{u \to \infty} P(D = 1 | \ln Y_1 = u, X = x) = l_1.$$

Thus, we can apply Theorem 2.1 to $(D, \ln Y_1, X)$ and ψ_1 is identified. The same result holds for ψ_0. □

---

[5] Note that this generalized Roy model is also used as a structural underlying framework for the estimation of treatment effects, with D corresponding in that case to the treatment status and $G + \eta$ to the cost of receiving treatment (see Heckman & Vytlacil, 2005).

To the best of our knowledge, this is the first identification result on the effects of co-variates in generalized Roy models without exclusion restriction. Identification without exclusion restriction of the competing risk model, which is strongly related to the standard Roy model, has already been considered in the literature by Heckman & Honore (1989),[6] Abbring & van den Berg (2003), Lee (2006) and Lee & Lewbel (2008).[7] However, all of the strategies proposed in these papers break down when turning to generalized Roy models.

Identification of (ψ_0, ψ_1) is obtained in Corollary 4.1 under rather mild restrictions on the unobservables. In particular, Condition (4.1) can be understood as a moderate dependence assumption between the unobservables. It is automatically satisfied for instance if ε_0, ε_1 and η are independent. It also holds if $(\varepsilon_0, \varepsilon_1, \eta)$ is gaussian, provided that

$$|Cov(\varepsilon_i, \varepsilon_{1-i} + (2i-1)\eta)| < V(\varepsilon_i), \quad i \in \{0, 1\}.$$

Noteworthy, this condition does not put drastic restrictions on the dependence between unobservables. For instance, it will be satisfied in the standard Roy model if $V(\varepsilon_0) = V(\varepsilon_1)$, as long as $(\varepsilon_0, \varepsilon_1)$ is not degenerated. It is also satisfied for instance in Heckman (1974)'s empirical application to labor market participation of married women, although the estimated correlation between ε and η is quite large (0.83).

Condition (4.1) is appealing because of its simple interpretation in terms of dependence between the unobservables. Nevertheless, ψ_0 and ψ_1 may be identified even if it fails, as soon as the "independence at infinity" conditions hold in this context, namely as soon as for all $x \in \text{Supp}(X)$, $\lim_{u \to \infty} P(D = 0 | \ln Y_0 = u, X = x) = l_0 > 0$ and $\lim_{u \to \infty} P(D = 1 | \ln Y_1 = u, X = x) = l_1 > 0$. Furthermore, one can apply Theorem 3.1 to this generalized Roy model, thus implying that the latter identifying conditions can be tested.

5   Concluding remarks

This note shows that identification of sample selection models can be achieved without instrument by letting the outcome, not a covariate, tend to infinity. The main condition, apart from the additive separability, is the "independence at infinity" of the selection variable and the covariates. In particular, unlike Chamberlain (1986) and Lewbel (2007), our

---

[6]Heckman & Honore (1989) use exclusion restrictions but only to identify the distribution of the underlying durations. Their proof shows that the effects of covariates are identified without such restrictions. Besides, identification of the standard Roy model has also been considered by Heckman & Honore (1990), but in the presence of exclusion restrictions.

[7]Interestingly, these last two papers do not rely on identification at the limit.

identification strategy does not rely on the existence of a covariate with a large support. Noteworthy, our identification proof is constructive, and an estimator of $\psi(.)$ and $\sigma(.)$ could be based on (2.7) for instance. One possible route for estimation would be to use trimmed means, in a similar spirit as Andrews & Schafgans (1998). In this case, we conjecture that the rate of convergence would depend on the thickness of the tail of the distribution of the outcome, as in Andrews & Schafgans (1998) and Khan & Tamer (2009). The derivation of the estimators and their properties seems quite intricate, however, and we leave this issue for future research.

References

Abbring, J. & van den Berg, G. (2003), 'The identifiability of the mixed proportional hzards competing risks model', *J.R. Statist. Soc. B* **65**, 701–710.

Andrews, D. K. & Schafgans, M. (1998), 'Semiparametric estimation of the intercept of a sample selection model', *Review of Economic Studies* **65**, 497–517.

Bayer, P. J., Khan, S. & Timmins, C. (2008), Nonparametric identification and estimation in a generalized roy model. NBER Working Paper.

Borjas, G. (1987), 'Self-selection and the earnings of immigrants', *American Economic Review* **77**, 531–553.

Carneiro, P., Hansen, K. & Heckman, J. (2003), 'Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice', *International Economic Review* **44**, 361–422.

Chamberlain, G. (1986), 'Asymptotic efficiency in semiparametric model with censoring', *Journal of Econometrics* **32**, 189–218.

Dagsvik, J. & Strøm, S. (2006), 'Sectoral labour supply, choice restrictions and functional form', *Journal of Applied Econometrics* **21**, 803–826.

D'Haultfœuille, X. & Maurel, A. (2009), Inference on a generalized roy model with exclusion restrictions. Working Paper.

Heckman, J. J. (1974), 'Shadow prices, market wages, and labor supply', *Econometrica* **42**, 679–694.

Heckman, J. J. (1990), 'Varieties of selection bias', *The American Economic Review* **80**, 313–318.

Heckman, J. J. & Honore, B. (1989), 'The identifiability of competing risks models', *Biometrika* **76**, 325–330.

Heckman, J. J. & Honore, B. (1990), 'The empirical content of the roy model', *Econometrica* **58**, 1121–1149.

Heckman, J. & Vytlacil, E. (2005), 'Structural equations, treatment effects, and econometric policy evaluation', *Econometrica* **73**, 669–738.

Khan, S. & Tamer, E. (2009), Irregular identification, support conditions and inverse weight estimation. Working Paper.

Lee, S. (2006), 'Identification of a competing risks model with unknown transformations of latent failure times', *Biometrika* **93**, 996–1002.

Lee, S. & Lewbel, A. (2008), Identification in accelerated failure time competing risks models. Working Paper.

Lewbel, A. (2007), 'Endogenous selection or treatment model estimation', *Journal of Econometrics* **141**, 777–806.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford Economic Papers(New Series)* **3**, 135–146.

Vella, F. (1998), 'Estimating models with sample selection bias: a survey', *Journal of Human Ressources* **33**, 127–169.

Willis, R. & Rosen, S. (1979), 'Education and self-selection', *Journal of Political Economy* **87**, S7–S36.