Empirical Economics (2007) 33:339–357
DOI 10.1007/s00181-006-0106-1

ORIGINAL PAPER

Avoiding arbitrary exclusion restrictions using ratios of reduced-form estimates

Myoung-jae Lee · Pao-Li Chang

Received: 15 July 2005 / Accepted: 15 June 2006 / Published online: 5 October 2006

Abstract We show how to obtain *coherent structural-form (SF) exclusion restrictions using the reduced-form (RF) parameter ratios*. It will be shown that an *over-identified SF corresponds to a group of regressors sharing the same RF ratio value*; those regressors should be excluded jointly from the SF. If there is no group structure, then the SF is just-identified; in this case, however, it is no longer clear which regressor should be excluded. Hence, *just-identified SF's are more arbitrary than over-identified SF's* in terms of exclusion restrictions. This is in stark contrast to the notion that the former is less arbitrary than the latter, because the former excludes fewer regressors. We formalize these points, and then suggest to find the number of modes in the estimated RF ratios as a way to find groups in the ratios. For this purpose, an informal graphical method using a kernel nonparametric method and a formal modality test are employed. An empirical example with selling price in a residential real estate market and duration on the market as two endogenous variables is provided.

Keywords Simultaneous equations · Exclusion restrictions · Identification

JEL Classification C13 · C31

The authors are grateful to the editor and two anonymous referees for their comments.

M. Lee (✉)
Department of Economics, Korea University, Seoul 136-701, South Korea
e-mail: myoungjae@korea.ac.kr

P. Chang
School of Economics and Social Sciences, Singapore Management University,
90 Stamford Road, Singapore 178903, Singapore

1 Introduction

Consider the following two structural form (SF) equations:

$$\begin{array}{ll} SF1 : y_1 = \alpha_1 y_2 + x_1' \beta_1 + x_c' \gamma_1 + u_1, & \alpha_1 \neq 0 \\ SF2 : y_2 = \alpha_2 y_1 + x_2' \beta_2 + x_c' \gamma_2 + u_2, & \alpha_2 \neq 0 \end{array} \tag{1}$$

where y_1 and y_2 are the endogenous variables, x_j is a $k_j \times 1$ exogenous regressor vector, $j = 1, 2$, such that x_1 and x_2 share no common elements, x_c is a $k_c \times 1$ exogenous regressor vector common to the two SF's, u_1 and u_2 are error terms, and α_j, β_j, and γ_j are parameters $j = 1, 2$. By exogeneity, we mean zero correlation with the error terms. If $\alpha_1 \alpha_2 = 1$, then the system is singular; otherwise the system can be solved for the reduced forms (RF).

The model (1) occurs frequently in economics: for instance,

ex1: y_1 is price, and y_2 is quantity transacted;
ex2: y_1 is unemployment duration, and y_2 is new job wage;
ex3: y_1 is husband labor supply, and y_2 is wife labor supply.

Assuming that the rank condition for identification holds (see Lee and Kimhi (2005) for a new look at the identification issue in simultaneous equations), if x_1 (x_2) is a scalar with a non-zero coefficient, then y_2 (y_1) is 'just-identified'; otherwise, if x_1 (x_2) has multiple components with non-zero coefficients, then y_2 (y_1) is 'over-identified'. Exclusion restrictions that x_1 and x_2 share no common elements are necessary in identifying the SF's. Although exclusion restrictions are supposedly provided by the economic theory behind (1), in practice, usually the underlying economic theory gives little guideline on what to exclude and whether to exclude one or multiple variables. For (1), researchers often choose a single variable to exclude in a rather arbitrary way, and then examine whether to exclude more variables with some over-identification tests.

There are *two types of decisions involved in the exclusion restrictions* for (1) (hence, two types of arbitrariness):

GROUPING: Which variables to remove jointly (two groups of variables are needed);
ASSIGNING: Which group to remove from which equation.

For instance, in ex3, the exogenous regressors could be characteristics of the husband, wife, and some local economic conditions relevant to labor demand. In this case, we might use the husband characteristics for x_1, the wife characteristics for x_2, and the local economic conditions for x_c, because x_c is likely to affect both y_1 and y_2; this is grouping and assigning done together. In this example, the two decisions are fairly easy to make, but there are cases where they are not as in the following.

In testing empirically the trade protection theory of Grossman and Helpman (1994), Goldberg and Maggi (1999) derived one equation (Eq. (2) in p. 1139), obtained by maximizing a joint surplus function with respect to tariff. The equation includes three endogenous variables: tariff, 'import-demand elasticity', and 'import-penetration ratio'. Goldberg and Maggi argued that the import-demand elasticity should be on the left-hand side and then proceeded. But why this has to be the case is far from obvious. In this single equation with three endogenous variable case, both decisions—grouping and assigning—are ambiguous, with the economic theory suggesting hardly anything. The same equation was further used in Eicher and Osang (2002). Instead of using arbitrary SF's, some researchers choose to avoid SF's altogether. For instance, Russo et al. (2000) examine relationship between job vacancy duration and the number of applications arrived. Although these two variables are likely to be simultaneously related, Russo et al. (2000) could not come up with plausible exclusion restrictions, which prompted them to estimate only the RF's.

For this kind of situations where the economic theory fails to provide which SF should be formed with what exclusion restriction, between the two decisions of grouping and assigning, this paper will show that there exists an *empirically constructive way to do the grouping using only the RF estimates*; this paper, however, will offer little for the assigning.

The main idea is examining the ratios of the two RF estimates (and then making linear combinations of the two RF's to recover the SF's). We will say that two exogenous regressors belong to the same *group* if their RF ratios are the same, and call an exogenous regressor a "*loner*" if it does not belong to any group. As will be shown later, the RF ratios satisfy the following:

- First, if both SF's are over-identified, then there are two groups and a loner/loners in the RF ratios.
- Second, if one SF is over-identified and the other SF is just-identified, then there are only one group and a loner/loners.
- Third, if both SF's are just-identified, then there are only loners.

Hence, examining the RF ratios provide valuable information for the grouping part of the exclusion restrictions. As for the assigning part, there is no clear-cut rule to apply. But we will provide some practical answers, if not the solutions, through an empirical illustration.

In short, the goal of this paper is to show how to build SF's empirically from the RF's, avoiding the arbitrariness involved in SF exclusion restrictions to some extent. Being able to build SF's with the information in the RF's is advantageous, for we can construct the RF's better using various specification tests. The idea of learning about SF specifications using the RF is not new; see, for instance, Maasoumi (1990) and Spanos (1990). But, in the literature, there has been no discussion on avoiding arbitrary exclusion restrictions in the way we will do in this paper.

Although we will deal with linear simultaneous equations, our approach is also good for simultaneous equations with limited dependent variables where the two endogenous variables are fully observed in a subset of the data. One example is 'simultaneous tobit' which is appropriate for ex2 and ex3 above when one endogenous variable is censored (duration in ex2 and wife labor supply in ex3); see Blundell and Smith (1994), Lee (1995), and the references therein.

The rest of the paper is organized as follows. Section 2 introduces informally our main idea of examining RF ratios to avoid grouping arbitrariness. Section 3 formalizes our idea and lays out 'modality'-based procedures to detect groups in the RF ratios. Section 4 provides an empirical example using a real estate market data set; how to avoid assigning arbitrariness is discussed here as well, albeit limitedly. Section 5 draws conclusions.

2 Implications of exclusion restrictions on RF ratios

Define

$$\theta \equiv 1 - \alpha_1\alpha_2.$$

Assuming $\alpha_1\alpha_2 \neq 1$, solve the SF's in (1) for the RF's:

$$y_1 = \frac{1}{\theta}\left\{x_1'\beta_1 + x_c'(\gamma_1 + \alpha_1\gamma_2) + x_2'\alpha_1\beta_2 + u_1 + \alpha_1 u_2\right\}, \tag{2}$$

$$y_2 = \frac{1}{\theta}\left\{x_2'\beta_2 + x_c'(\gamma_2 + \alpha_2\gamma_1) + x_1'\alpha_2\beta_1 + u_2 + \alpha_2 u_1\right\}$$

$$\Longrightarrow \text{RF1: } y_1 = w'\eta_1 + v_1 \quad \text{and} \quad \text{RF2: } y_2 = w'\eta_2 + v_2, \tag{3}$$

defining w, η_1, η_2, v_1, and v_2 appropriately such that w is the system exogenous regressor vector, η_1 and η_2 are the RF parameters, v_j's are the RF error terms.

From (2) and (3), the ratios of the coefficients of RF1 and RF2 corresponding to x_1 is α_2^{-1}, and the ratios for x_2 is α_1; as for x_c, the ratios can be anything. Hence, we get the following structure on the RF ratios:

$$x_1 : \alpha_2^{-1}, \quad x_2 : \alpha_1, \quad x_c : \text{possibly all different.} \tag{4}$$

Thus, examining the ratios of RF1 and RF2 is informative for learning about the SF exclusion restrictions. In the following, we examine various possibilities of (4) assuming that there are six exogenous regressors $w_1, \ldots, w_6$. The number six is only to simplify exposition; no loss of generality here.

2.1 Two groups and some loners

Suppose that the ratios of the RF coefficients are

$$0.5, 0.5, \quad 3, 3, \quad 10, 20. \tag{5}$$

Then, owing to (4), the first group (w_1, w_2) must be in one SF, the second group (w_3, w_4) in the other SF, and the third (w_5, w_6) in both. Depending on the value of γ_1 and γ_2, some variables in x_c may show ratios close to 0.5 or 3. But so long as they are not exactly 0.5 nor 3, we should be able to get a grouping like (5) in large enough samples.

Turning to assigning, if we have a prior knowledge that (w_1, w_2) should be excluded from SF1, then the following SF's are obtained from two linear combinations of the RF's in (3):

$$y_1 - 0.5y_2 = \ldots \text{ no } (w_1, w_2) \quad \text{and} \quad y_1 - 3y_2 = \ldots \text{ no } (w_3, w_4) \tag{6}$$

$$\Longrightarrow y_1 = 0.5y_2 +, \ldots \text{ no } (w_1, w_2) \quad \text{and} \quad y_2 = \frac{1}{3}y_1 +, \ldots \text{ no } (w_3, w_4). \tag{7}$$

Hence, $\alpha_1 = 0.5$ and $\alpha_2 = 1/3$ in (1). Equations (7) show that it is wrong to start with a SF excluding, for instance, w_1 and w_3 together. Excluding (w_1, w_2) from SF1 is equivalent to solving the first equation in (6) for y_1, not for y_2 ; if we exclude (w_1, w_2) from SF2, we would solve the first equation for y_2.

2.2 One group and some loners

Suppose, instead of (5), we have the following RF ratios:

$$0.5, 0.5, \quad 3, 10, 20, 30. \tag{8}$$

In this case, (w_1, w_2) should be removed together from a SF that is over-identified, and one variable out of $(w_3, \ldots, w_6)$ should be removed from the other SF that is just-identified. Differently from (5), however, now it is not clear at all which one to choose among $w_3, \ldots, w_6$. In this sense, *just-identified SF's are more arbitrary than over-identified SF's*, although the former may look less restrictive excluding only a single variable than the latter. Depending on the arbitrary choice of a single variable to exclude, the just-identified SF parameter will be different.

Suppose we exclude (w_1, w_2) from SF1 and w_6 from SF2. Then we get the following SF's:

$$\begin{aligned} & y_1 - 0.5y_2 = \ldots \text{ no } (w_1, w_2) \quad \text{and} \quad y_1 - 30y_2 = \ldots \text{ no } w_6 \\ \Longrightarrow\ & y_1 = 0.5y_2 +, \ldots \text{ no } (w_1, w_2) \quad \text{and} \quad y_2 = \frac{1}{30}y_1 +, \ldots \text{ no } w_6. \end{aligned} \tag{9}$$

2.3 All loners

Now suppose we have

$$0.5,\ 1,\ 3,\ 10,\ 20,\ 30 \tag{10}$$

with no group whatsoever. In this case, we can obtain two just-identified SF's by removing any two variables. Depending on which ones are removed in (10), the SF coefficients will be different.

Suppose we remove w_1 and w_2 to get

$$\begin{aligned} y_1 - 0.5y_2 = \ldots \text{ no } w_1 \quad \text{and} \quad y_1 - y_2 = \ldots \text{ no } w_2 \\ \Longrightarrow y_1 = 0.5y_2+, \ldots \text{ no } w_1 \quad \text{and} \quad y_1 = 1 \cdot y_2+, \ldots \text{ no } w_2. \end{aligned} \tag{11}$$

Imagine that w_1 is a policy variable which can be controlled to attain some target level of y_1 or y_2. The $y_1 - 0.5y_2$ equation simply shows that we cannot hit the two targets (y_1 and y_2) freely with one "tool" w_1: no matter how we choose w_1, still y_1 and y_2 will maintain the relationship $y_1 - 0.5y_2$ because w_1 is absent in the $y_1 - 0.5y_2$ equation. This seems rather trivial; it would be more interesting to see the relationship $y_1 - 0.5y_2$ between y_1 and y_2 being undisturbed even if we control multiple tools, which will be the case if the tools are removable together. That is, *viewed from the RF's, over-identified SF's are meaningful while just-identified SF's are not.* Put it differently, removing one regressor can be always done with a linear combination of the y_1 and y_2 RF as in

$$y_1 + \lambda y_2 = w'(\eta_1 + \lambda \eta_2) + v_1 + \lambda v_2,$$

which is thus trivial, whereas removing multiple regressors with a single linear combination requires a group structure in the RF ratios, which is thus nontrivial.

2.4 One group only without any loner

Suppose we have

$$0.5,\ 0.5,\ 0.5,\ 0.5,\ 0.5,\ 0.5:$$

only one group without any loners. In this case, the only possible linear combination of y_1 and y_2 is $y_1 - 0.5y_2$ to yield "no exogenous-variable relation"

$$y_1 - 0.5y_2 = v_1 - 0.5v_2. \tag{12}$$

If we try to get two SF's out of this, the only way is solving this single equation twice for y_1 and y_2, respectively, which then yields the following singular system with $\alpha_1\alpha_2 = 1$:

$$y_1 = 0.5y_2 + v_1 - 0.5v_2 \quad \text{and} \quad y_2 = 2y_1 + (v_2 - 2v_1).$$

Note that the correlation coefficient between the two SF error terms is -1. Analogously to the remarks after (11), this singular system is interesting in that no matter how we control all exogenous variables, still we cannot change the structural relationship $y_1 - 0.5y_2$ in the singular system. Equation (12) can be viewed as an 'equilibrium (or stable)' relationship between y_1 and y_2: regardless of the w variables, y_1 and y_2 will maintain (12).

In fact, the concept of singular relationship can be applied also to (8) above. Suppose we are not sure of the assigning decisions in (8). Then the only sensible thing to do in (8) is to remove w_1 and w_2 and present the result as

$$y_1 - 0.5y_2 = \cdots \text{ no } (w_1, w_2), \ldots + v_1 - 0.5v_2 \tag{13}$$

without trying to solve this for y_1 nor for y_2. This display depicts a 'conditionally (on $w_3, \ldots, w_6$) stable relationship' between y_1 and y_2, showing that, even if we control the two exogenous variables w_1 and w_2, we cannot hit the two targets y_1 and y_2 freely. This is because the two control variables affect y_1 and y_2 in the (proportionally) same way. If we solve (13) twice first for y_1 and then for y_2, we will get two singular SF's. *Singular SF's are not useless; we just have to take "one half" of them and interpret it as a stable relation between the response variables that is invariant to certain regressor changes.*

3 Formal procedures

3.1 Identification of groups

Instead of starting with the SF's (1) with exclusion restriction explicit, consider the SF's with the system regressors w (of dimension $k \times 1$) and no exclusion restrictions:

$$y_1 = \alpha_1 y_2 + w'\delta_1 + u_1 \quad \text{and} \quad y_2 = \alpha_2 y_1 + w'\delta_2 + u_2,$$

where δ_1 and δ_2 are the SF parameters. Solve these for y_1 and y_2 to get, under $(1 - \alpha_1\alpha_2 =)\ \theta \neq 0$,

$$y_1 = w'\eta_1 + v_1 \quad \text{and} \quad y_2 = w'\eta_2 + v_2 \quad \text{where}$$
$$\eta_1 = \frac{\delta_1 + \alpha_1\delta_2}{\theta}, \quad \eta_2 = \frac{\delta_2 + \alpha_2\delta_1}{\theta}, \quad v_1 = \frac{u_1 + \alpha_1 u_2}{\theta}, \quad v_2 = \frac{u_2 + \alpha_2 u_1}{\theta}.$$

Let

$$w = (w_1, \ldots, w_k)', \quad \delta_1 = (\delta_{11}, \ldots, \delta_{1k})', \ \ \delta_2 = (\delta_{21}, \ldots, \delta_{2k})'$$
$$\eta_1 = (\eta_{11}, \ldots, \eta_{1k})' \quad \text{and} \quad \eta_2 = (\eta_{21}, \ldots, \eta_{2k})'.$$

For an element w_j of w, we have

$$\eta_{1j} = \frac{\delta_{1j} + \alpha_1\delta_{2j}}{\theta} \quad \text{and} \quad \eta_{2j} = \frac{\delta_{2j} + \alpha_2\delta_{1j}}{\theta}.$$

We examine three cases in the following: $(\eta_{1j} = 0, \eta_{2j} = 0)$, $(\eta_{1j} \neq 0, \eta_{2j} = 0)$, and $(\eta_{1j} \neq 0, \eta_{2j} \neq 0)$. For the last case, we look at two elements w_j and $w_{j'}$ of w to examine their RF ratios.

Firstly, suppose $\eta_{1j} = \eta_{2j} = 0$. Both η_{1j} and η_{2j} being zero implies both δ_{1j} and δ_{2j} being zero (and vice versa), which can be seen by solving the preceding display for δ_{1j} and δ_{2j}. That is, plug $\delta_{1j} = \eta_{1j}\theta - \alpha_1\delta_{2j}$ from the η_{1j} equation into the η_{2j} equation to get

$$\eta_{2j} = \frac{\delta_{2j} + \alpha_2(\eta_{1j}\theta - \alpha_1\delta_{2j})}{\theta} = \delta_{2j} + \alpha_2\eta_{1j} \implies \delta_{2j} = \eta_{2j} - \alpha_2\eta_{1j}$$
$$\implies \delta_{1j} = \eta_{1j}\theta - \alpha_1(\eta_{2j} - \alpha_2\eta_{1j}) = \eta_{1j} - \alpha_1\eta_{2j}.$$

Hence, any component w_j with $\eta_{1j} = \eta_{2j} = 0$ is irrelevant to the simultaneous system and should be discarded. This is feasible as η_{1j} and η_{2j} are identified; if necessary, redefine w after removing such components. From now on, *assume that w_j should be at least in one SF to ignore the case $\eta_{1j} = \eta_{2j} = 0 (\iff \delta_{1j} = \delta_{2j} = 0)$.*

Secondly, suppose $\eta_{1j} \neq 0$ and $\eta_{2j} = 0$. That is,

$$(a) : \delta_{1j} + \alpha_1\delta_{2j} \neq 0 \quad \text{but} \quad (b) : \delta_{2j} + \alpha_2\delta_{1j} = 0.$$

There are three subcases to consider:

$$(i) : (\delta_{1j} = 0, \delta_{2j} \neq 0), \quad (ii) : (\delta_{1j} \neq 0, \delta_{2j} = 0), \quad (iii) : (\delta_{1j} \neq 0, \delta_{2j} \neq 0).$$

Subcase (i) contradicts (b), and thus can be ignored. Subcase (ii) implies none in (a) and $\alpha_2 = 0$ in (b). Subcase (iii) implies $\alpha_1 \neq -\delta_{1j}/\delta_{2j}$ in (a) and $\alpha_2 = -\delta_{2j}/\delta_{1j}$ in (b). Subcases (ii) and (iii) can be summed up as $\alpha_2 = -\delta_{2j}/\delta_{1j}$; $\alpha_1 \neq -\delta_{1j}/\delta_{2j}$ is non-informative because $\alpha_1\alpha_2 \neq 1$ has been already assumed. In view of this, one of the two actions may be taken: one is following (ii) to assign w_j to SF1 and verify later $\alpha_2 = 0$, and the other is keeping w_j in both SF's and verify later $\alpha_2 = -\delta_{2j}/\delta_{1j}$. The latter action is more general, but the former may be adopted if there are not enough exclusion restrictions available at the end, or if w_j should accompany some other variables in assigning. When $\eta_{1j} = 0$ and $\eta_{2j} \neq 0$, the main finding can be summed up as $\alpha_1 = -\delta_{1j}/\delta_{2j}$, and analogous actions may be taken.

Thirdly, consider now only the components of w with $\eta_{1j} \neq 0$ and $\eta_{2j} \neq 0$. Pick two such components, say j and j', to form the ratios $\rho_j = \eta_{1j}/\eta_{2j}$ and $\rho_{j'} = \eta_{1j'}/\eta_{2j'}$. Suppose $\rho_j = \rho_{j'}$:

$$\frac{\delta_{1j} + \alpha_1\delta_{2j}}{\delta_{2j} + \alpha_2\delta_{1j}} = \frac{\delta_{1j'} + \alpha_1\delta_{2j'}}{\delta_{2j'} + \alpha_2\delta_{1j'}}$$
$$\iff (\delta_{1j} + \alpha_1\delta_{2j})(\delta_{2j'} + \alpha_2\delta_{1j'}) = (\delta_{1j'} + \alpha_1\delta_{2j'})(\delta_{2j} + \alpha_2\delta_{1j})$$
$$\iff \delta_{1j}\delta_{2j'} + \alpha_1\delta_{2j}\delta_{2j'} + \alpha_2\delta_{1j}\delta_{1j'} + \alpha_1\alpha_2\delta_{1j'}\delta_{2j}$$
$$= \delta_{1j'}\delta_{2j} + \alpha_1\delta_{2j}\delta_{2j'} + \alpha_2\delta_{1j}\delta_{1j'} + \alpha_1\alpha_2\delta_{1j}\delta_{2j'}$$
$$\iff \delta_{1j}\delta_{2j'} + \alpha_1\alpha_2\delta_{1j'}\delta_{2j} = \delta_{1j'}\delta_{2j} + \alpha_1\alpha_2\delta_{1j}\delta_{2j'}.$$

If $\delta_{1j} = 0$ (and thus necessarily $\delta_{2j} \neq 0$), then the last equation becomes $\delta_{1j'}\delta_{2j}\theta = 0 \implies \delta_{1j'} = 0$. If $\delta_{2j} = 0$ (and thus necessarily $\delta_{1j} \neq 0$), then we get $\delta_{1j}\delta_{2j'}\theta = 0 \implies \delta_{2j'} = 0$. In short, *for two elements w_j and $w_{j'}$ with $\eta_{1j} \neq 0$, $\eta_{2j} \neq 0$, $\eta_{1j'} \neq 0$, and $\eta_{2j'} \neq 0$, if the RF ratios η_{1j}/η_{2j}* and *$\eta_{1j'}/\eta_{2j'}$ are equal, then one* *element excluded from a SF implies the other excluded from the same SF.* The RF ratios carry only the common measurement units for y_1 and y_2; as the ratios are not affected by any regressor measurement units, they are comparable "pure numbers" not subject to measurement unit manipulation.

Some remarks are in order for the second case above. Firstly, if $\alpha_2 = 0$ due to $\delta_{1j} \neq 0$ and $\delta_{2j} = 0$, then there will be no more simultaneity although the endogeneity problem can be still present through the correlation between u_1 and u_2; an analogous statement can be made for $\alpha_1 = 0$ due to $\delta_{1j} = 0$ and $\delta_{2j} \neq 0$. Secondly, in either subcase (ii) or (iii)—(i) was ruled out already—w_j cannot be used as an instrument for the α -parameters. Thirdly, testing for $H_o : (\eta_{1j} \neq 0, \eta_{2j} = 0)$ or $H_o : (\eta_{1j} = 0, \eta_{2j} \neq 0)$ could be difficult as the hypotheses involve both simple and composite components. Given the way hypotheses are tested in practice (i.e., only rejecting a null hypothesis makes sense), relying on the third case $(\eta_{1j} \neq 0, \eta_{2j} \neq 0)$ seems sensible.

The remainder of this subsection recasts the third case above using matrices; we owe this part to a referee. On one hand, this matrix-based exposition looks more elegant, and on the other hand, it may facilitate generalization for more than two SF's although we do not pursue this possibility in this paper.

Define

$$y \equiv (y_1, y_2)', \quad u \equiv (u_1, u_2)',$$
$$\Lambda \equiv \begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix}, \quad \Delta \equiv \begin{bmatrix} \delta_{11}, \ldots, \delta_{1k} \\ \delta_{21}, \ldots, \delta_{2k} \end{bmatrix}, \quad \Pi \equiv \Lambda^{-1}\Delta, \quad \text{and } v \equiv \Lambda^{-1}u.$$

to get the SF's and RF's:

$$\Lambda y = \Delta w + u \implies y = \Lambda^{-1}\Delta w + \Lambda^{-1}u = \Pi w + v.$$

Pick the last two elements of w to define

$$w_* \equiv (w_{k-1}, w_k)', \quad w_{-*} \equiv (w_1, \ldots, w_{k-2})', \quad \text{and} \quad \Pi = (\Pi_{-*}, \Pi_*)$$

where the last 2×2 submatrix Π_* is

$$\Pi_* \equiv \frac{1}{\theta} \begin{bmatrix} 1 & \alpha_1 \\ \alpha_2 & 1 \end{bmatrix} \cdot \begin{bmatrix} \delta_{1,k-1} & \delta_{1k} \\ \delta_{2,k-1} & \delta_{2k} \end{bmatrix}.$$

If the RF ratio for the $(k-1)$th element is the same as the RF ratio for the kth element, then the first and second rows are linearly dependent, which implies $\det(\Pi_*) = 0$, which in turn implies

$$\det\left(\begin{bmatrix} \delta_{1,k-1} & \delta_{1k} \\ \delta_{2,k-1} & \delta_{2k} \end{bmatrix}\right) = \delta_{1,k-1}\delta_{2k} - \delta_{1k}\delta_{2,k-1} = 0.$$

Hence if $\delta_{1,k-1} = 0$, then $\delta_{1k} = 0$ because $\delta_{2,k-1} \neq 0$: w_{k-1} and w_k are excluded together from SF1. Analogously, if $\delta_{2k} = 0$, then $\delta_{2,k-1} = 0$ because $\delta_{1k} \neq 0$:

w_{k-1} and w_k are excluded together from SF2. Therefore, the RF ratio for the $(k-1)$th element being the same as the RF ratio for the kth element implies that w_k and w_{k-1} should be excluded jointly if they are to be excluded from any SF. If the $(k-2)$th element of w also has the same RF ratio, then we can apply the same logic to the $(k-2)$th and kth elements. Hence, the exclusion restriction applies jointly to all components of w that share the same RF ratio.

3.2 Modality tests for RF-ratio groups

Turning to how to implement the above procedure in practice, suppose the RF's have been estimated and the regressors with zero coefficients in both RF's have been removed as they are irrelevant. What we should look for in the RF ratios is "close numbers" as we have only estimates, not the true ratios. With estimates, close numbers mean bunching up of numbers. As bunching increases the frequency of RF ratios around some points, this feature may be best detected by *modality tests* when the RF ratios are taken as data points.

An informal way to detect modality is a graphical analysis: estimate the density of the RF ratios nonparametrically and see the number and locations of the modes. A formal way to detect modality is employing a multimodality test. Several tests for modality have appeared in the literature as can be seen, for instance, in Cheng and Hall (1998). Among them, the Silverman (1981) test seems to be the easiest to implement, which is described in the following and will be used later in our empirical section. The test is a no stranger to economists, as it has been used for 'income convergence' in Bianchi (1997), Kang and Lee (2005), Zhu (2005) and so on. Note that the so-called 'clustering', 'discrimination', and 'classification' methods in the statistics literature are not appropriate in finding group structure in the RF ratios as these methods would assign each ratio to a group whereas we desire to isolate only one or two groups without disturbing "loners".

Consider a sample $y_1, \ldots, y_N$ from a density f. In testing for 'H_o: μ modes' versus 'H_a: more than μ modes' for $\mu = 1, 2, \ldots$, define the 'μ-critical bandwidth'

$$h_{\text{crit}} \equiv \inf\{h : \hat{f}(y_o, h) \text{ has at most } \mu \text{ modes}\},$$
$$\text{where } \hat{f}(y_o, h) \equiv \frac{1}{Nh} \sum_i \phi\left(\frac{y_i - y_o}{h}\right);$$

h is a 'bandwidth' converging to 0 as $N \to \infty$, and ϕ is the $N(0,1)$ density playing the role of 'kernel' (weight) in the kernel nonparametric density estimator $\hat{f}(y_o, h)$ of f at $y = y_o$. Due to ϕ, as Silverman (1981) proves, it holds that

$$\hat{f}(y_o, h) \text{ has more than } \mu \text{ modes iff } h < h_{\text{crit}}.$$

The main idea is that if h_{crit} is "large", then the H_o is rejected, because it takes a "deliberate" over-smoothing with the large h_{crit} to get only μ (or fewer) modes.

The p-value for this test can be obtained by a bootstrap from the data: sample $w_1^*, \ldots, w_N^*$ with replacement from $y_1, \ldots, y_N$ and get

$$y_i^* = \bar{w}^* + \frac{w_i^* - \bar{w}^* + h_{\text{crit}}\varepsilon_i}{\{1 + (h_{\text{crit}}/s_y)^2\}^{1/2}}, \quad \text{where } \varepsilon_i\text{'s } iid\ N(0,1)$$
$$\text{and } s_y^2 = \frac{1}{N-1}\sum_i (y_i - \bar{y})^2,$$

to see whether

$$\hat{f}^*(y_o, h_{\text{crit}}) \equiv \frac{1}{N \cdot h_{\text{crit}}} \sum_i \phi\left(\frac{y_i^* - y_o}{h_{\text{crit}}}\right)$$

has more than μ modes or not. Repeating this, say B times, the bootstrap p-value is $B^{-1}\sum_{j=1}^{B} 1[h^*_{\text{crit},j} > h_{\text{crit}}]$ where $1[A] = 1$ if A holds and 0 otherwise and $h^*_{\text{crit},j}$ is the μ-critical bandwidth for the jth pseudo sample. Equivalently, the p-value is

$$\frac{1}{B}\sum_{j=1}^{B} 1[\hat{f}^*(y_o, h_{\text{crit}}) \text{ has more than } \mu \text{ modes in the } j\text{th pseudo sample}].$$

A few remarks are given in the remainder of this subsection for the Silverman's test in general (the first and second remarks below) and its application to RF ratios in particular (the third).

Firstly, if h gets too small, then $\hat{f}(y_o, h)$ becomes too jagged showing (misleadingly) too many modes; at the extreme, every observation may be taken as a mode. This means that the Silverman's test will break down for large values of μ. This problem with the Silverman's test has been avoided in the literature by proceeding sequentially: start with μ versus more than μ modes, $\mu = 1, 2, \ldots$, to proceed to $\mu + 1$ only if μ modes are rejected; otherwise stop to conclude μ modes. This sequential procedure will be also adopted in the empirical example below.

Secondly, Hall and York (2001) suggest two ways to minor-modify the Silverman's test, after showing that it systematically under-rejects. Between the two, the simpler is replacing $B^{-1}\sum_{j=1}^{B} 1[h^*_{crit,j} > h_{crit}]$ with $B^{-1}\sum_{j=1}^{B} 1[h^*_{crit,j} > h_{crit}1.13]$ for nominal level 5% test, which makes the p-value smaller (and thus the test rejects more easily). The number 1.13 comes from equation (4.1) of Hall and York (2001). This modification will be also used in our empirical analysis.

Thirdly, RF ratios can be formed in two different ways: e.g., η_{1j}/η_{2j} or η_{2j}/η_{1j}. So long as $\eta_{1j} \neq 0$ and $\eta_{2j} \neq 0$, which one is used does not matter, because bunching of numbers in one way is equivalent to bunching of numbers in the other way. But a caution is in order in regard to outliers. Suppose there are outliers in η_{1j}/η_{2j}, $j = 1, 2, \ldots$. Then in the reverse ratios, those outliers will gather around zero; this applies even to two outliers at the opposite ends, one positive and one negative. That is, "false" modes may be found at the tails or at zero. In fact, this problem occurs because either $\eta_{2j} \simeq 0$ or $\eta_{1j} \simeq 0$. Variables with this feature fall in the second case of the preceding subsection, and they should not be considered in the RF ratio comparison.

4 An empirical analysis: price and duration

4.1 Data and RF estimation to find groups

This section provides an empirical illustration. We are interested in two variables in a residential real estate market: first, the percentage discount (DISC) of the selling price (SP) relative to the initial listing price (LP) which is approximately $100 \times \ln(LP/SP)$, and second, the duration (T) of the property on the market until sale. LP and SP are measured in \$1,000, and T is measured in days. DISC and T are simultaneously related and depend on the characteristics of the house, its surroundings, the real estate broker, and etc. If a property's list price is higher (which tends to accompany a larger sales discount), then it might attract fewer potential buyers and thus take longer to sell. On the other hand, if the property has stayed on the market for a longer period, its seller might be more likely to offer a larger discount in order to make a deal.

Alternatively to DISC, we may consider using SP as an endogenous variable. But applying Least Squares Estimation of SP on LP with the data set to be described below, we obtained (t values in $(\cdot)$)

$$\ln \mathrm{SP} = \underset{(1.68)}{0.09} + \underset{(87.0)}{0.97} \cdot \ln \mathrm{LP}, \quad R^2 = 0.97; \tag{14}$$

that is, most of SP is explained by LP with the slope coefficient almost equal to 1. Taking $-100\{\ln(SP) - \ln(LP)\}$ as the response variable, we explain the small unexplained part of SP. The reason for using log-transformation for LP and SP, as well as for T in the estimation that follows, is to remove the asymmetry in the variables and potential heteroskedasticity.

A data set of size 467 was collected from the Multiple Listing Service of the State College District in Pennsylvania for the year 1991. State College is a small college town with the permanent population of about 50,000. The houses sold during the year were sampled and their duration since the first listing was recorded. The following is a list of exogenous variables available: years built minus 1,900 (YR), number of rooms (ROOM), number of bathrooms (BATH), dummy for heating by electricity (ELEC), property tax in \$1,000 (TAX), dummy for spring listing (L1), dummy for summer listing (L2), dummy for fall listing (L3), sale month interest rate (RATE), dummy for sale by a big broker (BIGS), and the number of houses on the market in the month when the house is listed (SUPPLY). Also lot-size that is highly correlated with ROOM was available but not used, for the variable has too many missings. The summary statistics are given in Table 1.

Table 1 Summary statistics

Variables	Mean	SD	Min	LQ	Med	UQ	Max
T	188.20	149.70	1.00	92.00	133.00	239.00	851.00
SP	114.86	57.70	19.00	78.40	97.50	134.00	400.00
LP	124.11	64.90	26.00	82.50	104.50	142.50	418.00
BATH	2.02	0.67	1.00	1.50	2.00	2.50	4.50
ELEC	0.52	0.50	0.00	0.00	1.00	1.00	1.00
ROOM	7.09	1.70	3.00	6.00	7.00	8.00	12.00
TAX	1.38	0.65	0.11	0.93	1.22	1.59	4.52
YR	73.03	15.09	21.00	63.00	77.00	86.00	91.00
L1	0.29	0.45	0.00	0.00	0.00	1.00	1.00
L2	0.31	0.46	0.00	0.00	0.00	1.00	1.00
L3	0.19	0.39	0.00	0.00	0.00	0.00	1.00
BIGS	0.78	0.42	0.00	1.00	1.00	1.00	1.00
RATE	9.33	0.32	8.50	9.24	9.47	9.58	9.64
SUPPLY	62.41	19.27	38.00	48.00	56.00	90.00	91.00

Table 2 Reduced form estimation

	RF1 for DISC	RF2 for ln T	RF1/RF2
	Est. (t-value)	Est. (t-value)	
1	4.51 (0.29)	−1.158 (−0.58)	−3.9
BATH	−3.21 (−1.46)	0.011 (0.04)	−306.1
BATH^2	0.94 (1.83)	0.017 (0.24)	55.8
ELEC	2.62 (3.34)	0.184 (2.45)	14.2
ROOM	−0.29 (−0.99)	−0.036 (−1.08)	8.1
TAX	−5.69 (−1.79)	−0.593 (−1.90)	9.6
TAX^2	0.72 (1.11)	0.109 (1.62)	6.6
YR	−0.24 (−2.11)	−0.002 (−0.27)	144.1
L1	−2.71 (−2.48)	−0.288 (−2.74)	9.4
L2	−3.96 (−3.72)	−0.275 (−2.29)	14.4
L3	−2.35 (−2.29)	−0.169 (−1.21)	13.9
lnLP	8.85 (3.36)	0.592 (2.73)	15.0
BIGS	−12.93 (−1.51)	−0.318 (−0.66)	40.6
BIGSYR	0.14 (1.27)	0.003 (0.48)	46.1
RATE	−1.22 (−0.97)	0.448 (2.87)	−2.7
SUPPLY	0.05 (2.39)	0.003 (1.16)	20.1
R^2 (s)	0.18 (6.95)	0.11 (0.75)	

Estimating the RF's for DISC and ln T, we get Table 2 where the asymptotic variances were estimated in the usual heteroskedasticity-robust way. Table 2 reveals that the two variables, BATH and YR, could be those variables with their RF1 coefficients non-zero and RF2 coefficients zero. We can keep the two variables in both SF's, but as shown below, BATH^2 will be assigned to a SF. Hence if we are to assign BATH and BATH^2 together, then the right SF is SF1 as the second case in Sect. 3.1 suggests. Assigning BATH to SF1 means that, again according to the second case in Sect. 3.1, we are entertaining the possibility of $\alpha_2 = 0$ at this early stage.

The following is the RF ratios in the last column of Table 2 in the increasing order except BATH and YR:

1	RATE	TAX^2	ROOM	$L1$	TAX	$L3$
−3.9	−2.7	6.6	8.1	9.4	9.6	13.9
ELEC	L2	lnLP	SUPPLY	BIGS	BIGSYR	$BATH^2$
14.2	14.4	15.0	20.1	40.6	46.1	55.8

Figure 1 shows two kernel nonparametric density estimates for these ratios over the domain $[-4, 56]$, with the left panel somewhat under-smoothed and the right panel somewhat over-smoothed. The figures point out either two modes (at about 10 and 45) or just one at about 10.

Conducting Silverman's (1981) modality test with bootstrap repetition number 5,000, we obtained the p-value 0.303 for 'H_o: 1 mode' versus 'H_a: more than 1 modes' when the level 5% modification in Hall and York (2001) was applied. This formal test fails to reject unimodality (i.e., only one group in the RF ratios). But given the apparent bimodality in the figures, we also tested for 'H_o: 2 modes' versus 'H_a: more than 2 modes' to obtain the p-value 0.238. The test does not seem to have much power with the small number of ratios (only 14 with two ratios excluded). We will show our analyses for two groups as well as for one group only.

Although it is hard to put down the cutoff points exactly, two groups according to the figures are:

"Long Group": (6.6, 8.1, 9.4, 9.6, 13.9, 14.2, 14.4, 15.0)
for (TAX^2, ROOM, L1, TAX, L3, ELEC, L2, lnLP)

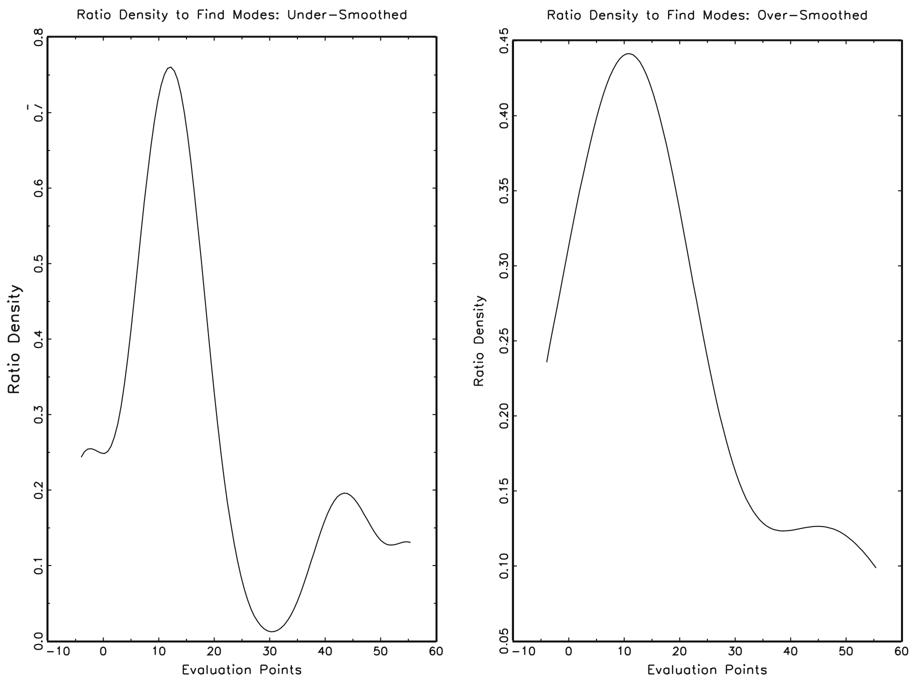

Fig. 1 Kernel ratio–density estimates

"Short Group": (40.6, 46.1, 55.8) for (BIGS, BIGSYR, BATH^2)

If we go a little conservative, we might shorten the list as well. Our discussion in the preceding sections was done with the population parameters; grouping with estimates based on a sample of $N = 467$ cannot be as clear-cut. A nonlinear Wald test for the null hypothesis that the RF ratios are the same in the first and second groups gave the test statistic 4.19 (p-value 0.90 under 9 degrees of freedom). Thus, the null hypothesis is easily accepted.

4.2 Two-group estimation after assignment

At this stage, we can proceed two ways. One is to accept two groups in the preceding subsection and go on to assignment. The other is to accept only one group (the long group with the global mode) and present a linear combination of two RF's free of the variables in the group; as already mentioned, we call this '*a singular SF*' because we get only one SF which may be solved for y_1 and y_2 to result in two SF's that are singular. This subsection presents our two group analysis, whereas the one group analysis and further remarks are provided in the next subsection. Note that, in the one-group analysis, we may pick up a loner as in (8) and (9) to do assigning. But because there is no clue for which loner to pick, we settle for a singular SF. For estimation in the following, we used the generalized method of moment (GMM) estimator stacking the two vector moment conditions $E(u_1 w) = 0$ and $E(u_1 w) = 0$; the estimator was implemented by slightly modifying a two-period panel data GMM program in Lee (2002).

Given the two groups, we need to assign each group to a SF. There are a couple of clues. Firstly, (14) provides a clue: as the coefficient of lnLP is near one, it is more plausible not to have lnLP in the DISC SF1, which assigns the long group to SF2. Secondly, as already noted, if we want to assign BATH and BATH^2 together to a single SF, the right SF is SF1 on the account of BATH. Table 3 is the result from these assignment. Thirdly, when these assignments get reversed, implausible SF estimates are obtained in Table 4 as discussed in the following paragraph. In an earlier version of this paper, we used smaller groups: only $L2$, $L3$, and lnLP for the long group, and only BIGS and BIGSYR for the short group. The ensuing estimation results, which are available from the authors upon request, are not much different from those using the above two groups with far more variables.

Table 4 looks implausible in a number of aspects. Firstly, regarding the estimates of α_1 and α_2 which are of our main interest, the estimates in Table 3 are 'stable' ($|\alpha_1\alpha_2| = 0.42 < 1$) while the estimates in Table 4 are not ($|\alpha_1\alpha_2| = 1.30 > 1$). Secondly, the assigned regressors in Table 4 look all insignificant, while this is not the case in Table 3. Thirdly, $COR(u_1, u_2) \simeq -1$ in Table 4, which is a concern.

Making the wrong assignment decision can have two consequences. One is having ineffective instruments for the endogenous regressor, since irrelevant variables are used as instruments. The other is omitted variable bias because

Table 3 SF under right assignment

	SF1 for DISC Est. (t-value)	SF2 for ln T Est. (t-value)
ln T	12.4 (6.13)	
DISC		0.034 (1.60)
1	18.1 (1.20)	−1.07 (−0.86)
BATH	−1.81 (−1.08)	
BATH^2	0.53 (1.22)	
ELEC		0.120 (1.65)
ROOM		−0.017 (−1.15)
TAX		−0.244 (−1.09)
TAX^2		0.039 (0.90)
YR	−0.23 (−3.26)	0.0052 (1.57)
$L1$		−0.162 (−1.87)
$L2$		−0.190 (−1.81)
$L3$		−0.109 (−1.64)
lnLP		0.346 (1.46)
BIGS	−9.42 (−1.69)	
BIGSYR	0.11 (1.58)	
RATE	−5.85 (−3.57)	0.441 (4.01)
SUPPLY	0.019 (1.01)	0.0011 (0.51)
$SD(u_j)$	8.62	0.68
	$COR(u_1, u_2) = -0.89$	

Table 4 SF under reverse assignment

	SF1 for DISC Est. (t-value)	SF2 for ln T Est. (t-value)
ln T	18.6 (1.49)	
DISC		0.070 (6.29)
1	10.0 (0.30)	−0.810 (−0.74)
BATH		0.078 (1.06)
BATH^2		−0.020 (−0.85)
ELEC	−0.79 (−0.30)	
ROOM	0.08 (0.16)	
TAX	0.59 (0.07)	
TAX^2	0.03 (0.02)	
YR	−0.15 (−2.89)	0.013 (2.93)
$L1$	0.71 (0.18)	
$L2$	1.33 (0.34)	
$L3$	0.87 (0.37)	
lnLP	−2.54 (−0.30)	
BIGS		0.312 (0.99)
BIGSYR		−0.003 (−0.99)
RATE	−7.97 (−1.28)	0.458 (4.42)
SUPPLY	0.01 (0.26)	−0.001 (−1.05)
$SD(u_j)$	12.40	0.68
	$COR(u_1, u_2) = -0.99$	

the relevant variables are wrongly excluded. If the wrongly excluded variables are not correlated with the other regressors in the SF, then the omitted variable bias will disappear but the error term variance will increase. Comparing the error term standard deviations in Tables 3 and 4, the latter problem seems to occur in Table 4.

Recall that we assigned BATH to SF1 while keeping YR in both SF's. In view of the second case in Sect. 3.1, assigning BATH to SF1 implies $\alpha_2 = 0$, which is not rejected in Table 3. Also, keeping YR in both SF's, we should have $\alpha_2 = -\delta_{2,YR}/\delta_{1,YR}$. Using the estimates in Table 3, the two sides are, respectively, 0.034 and $0.023 = -0.0052/(-0.23)$: the difference $0.034 - 0.023 = 0.011$ is about one half of 0.021 that is the standard deviation of the α_2-estimator. Overall, although we could not present a general analytic approach for assignment, there seem to be enough telltale signs to favor one assignment over the other.

4.3 One-group singular SF

If we do not want to impose assigning decisions, then we can adopt the single "strong" long group for the global mode, and take a linear combination of the two RF's to exclude all variables in the long group:

$$\text{DISC} - 12.4 \times \ln T = \ldots, \text{ no } (\text{TAX}^2, \text{ROOM, L1,..., } \ln \text{LP}), \ \ldots;$$

12.4 comes from the $\ln T$ coefficient in Table 3. Thus we have

$$\text{DISC} - 12.4 \times \ln T = \text{only BATH, BATH}^2\text{, YR, BIGS,} \\ \text{BIGSYR, RATE, SUPPLY, } \ldots \tag{15}$$

This means that, although there are many variables that affect DISC and T separately as shown in the RF's, only the variables in (15) (or a subset of the variables) are relevant for DISC-12.4$\cdot$ lnT, for the other variables affect DISC and lnT in "12.4-proportional way" . The correlation coefficients of u_1 and u_2 in Tables 3 and 4 are quite high: -0.89 and -0.99, respectively, suggesting that two SF's might be one too many.

Further interpreting (15), for instance, no matter how the seller chooses L1–L3 (when to list) and LP (at how much to list), he will sell the house either quickly with a small discount or slowly with a large discount. The seller may try to set the list price higher than usual and offer a big discount to sell the house quickly. This strategy will not work according to (15) since lnLP is excluded; the seller may influence either DISC or T, but not both. The seller may then consider setting lnLP high hoping to get away with a small DISC and small T, not a large DISC and large T, because (15) holds anyway regardless of ln LP. But (15) does not tell whether a small DISC and small T or a large DISC and large T will be the case. If the seller wants to get a definite answer, he needs another SF. He may then take SF2 in Table 3 where a large ln LP leads to a large T; that is, a large DISC and large T will be the fate.

Judged by the sign of the estimate for YR in SF1 of Table 3, which is our (15), a newer house has a negative influence on $DISC - 12.4 \times \ln T$: it is sold either more slowly or at a lower discount compared with old houses. Also judged by the sign of the estimate for RATE, when RATE goes up, houses are sold either more slowly or at a lower discount. In short, without imposing arbitrary assigning restrictions, it is still possible to make useful inferences regarding simultaneous structural relationship between the two endogenous variables.

The relationship between sale price and duration in the residential real estate market is analogous to that between wage and unemployment duration in labor market. The main difference between the two markets is the presence of list price in the real estate market; Horowitz (1992) analyzes the role of the list price in detail with a structural real estate market model under a search framework. Lancaster (1985) shows that, under certain parameterization, post-unemployment wage and duration are simultaneously related. Devine and Kiefer (1991) survey various models in the literature for simultaneous relationships between the two variables. Addison and Portugal (1989, p. 292) estimate the SF's to get various SF estimates depending on model specification and estimation methods. Their α_1 (the effect of duration on wage, both in log) ranges from -0.02 to -0.23. Their α_2 (the effect of wage on duration) ranges from -0.79 to -1.46. This means that a large (small) duration goes with a small (large) wage, contrary to the argument that those who wait longer will get higher wages. The labor market empirical finding more or less agrees with our finding in the housing market: those who wait will suffer.

5 Conclusions and final remarks

In this paper, we decomposed exclusion restrictions in simultaneous equations into two: grouping and assigning. Instead of an arbitrary grouping, we showed how to derive coherent grouping restrictions for structural forms (SF) from the reduced form (RF) estimate ratios. Finding modes in the RF ratios was proposed as a main way to find group structure in the RF ratios. We also showed that there is less arbitrariness in over-identified SF exclusion restrictions than in just-identified SF exclusion restrictions. We proposed singular simultaneous systems to avoid the assigning arbitrariness. A singular system can be regarded as a reduction of the RF's and as a 'conditionally stable relationship' between the endogenous variables. An empirical example where the two dependent variables are price (discount) and duration in a real estate market was provided.

In this paper, we focused on two-SF systems. Although formally extending our approach to three or more SF's is left for future research, it is still possible to use our proposal when there are more than two SF's. Suppose y_1 has $y_2, y_3, \ldots, y_J$ as the endogenous regressors with coefficients $\alpha_{12}, \alpha_{13}, \ldots, \alpha_{1J}$. Then $(y_3, \ldots, y_J)$ can be substituted out with their RF's to leave only y_2 (and y_1), with which α_{12} can be estimated by our method. Analogously, $(y_2, y_4, \ldots, y_J)$ can be substituted out to leave only y_3 (and y_1), with which α_{13} can be estimated, and so forth. This is somewhat long-winded, but it shows that our approach is viable for more than two SF's as well.

Acknowledgements The research of the first author was supported by a Korea University grant.

References

Addison J, Portugal P (1989) Job displacement, relative wage changes and duration of unemployment. J Labor Econ 7:281–302

Bianchi M (1997) Testing for convergence: evidence from non-parametric multimodality tests. J Appl Econ 12:393–409

Blundell R, Smith RJ (1994) Coherency and estimation in simultaneous models with censored or qualitative dependent variables. J Econ 64:355–373

Cheng MY, Hall P (1998) Calibrating the excess mass and dip test of modality. J R Stat Soc (Series B) 60:579–589

Devine T, Kiefer NM (1991) Empirical labor economics: the search approach. Oxford University Press, Oxford

Eicher T, Osang T (2002) Protection for sale: an empirical investigation: comment. Am Econ Rev 92:1702–1710

Goldberg PK, Maggi G (1999) Protection for sale: an empirical investigation. Am Econ Rev 89:1135–1155

Grossman G, Helpman E (1994) Protection for sale. Am Econ Rev 84:833–50

Hall P, York M (2001) On the calibration of Silverman's test for multimodality. Stat Sin 11:515–536

Horowitz JL (1992) The role of the list price in housing markets: theory and an econometric model. J Appl Econ 7:115–129

Kang SJ, Lee MJ (2005) Q-convergence with interquartile ranges. J Econ Dyn Control 29:1785–1806

Lancaster T (1985) Simultaneous equations models in applied search theory. J Econ 28:113–126

Lee MJ (1995) Semiparametric estimation of simultaneous equations with limited dependent variables: a case study of female labor supply. J Appl Econ 10:187–200

Lee MJ (2002) Panel data econometrics: methods-of-moments and limited dependent variables. Academic, New York

Lee MJ, Kimhi A (2005) Simultaneous equations in ordered discrete responses with regressor-dependent thresholds. Econ J 8:176–196

Maasoumi E (1990) How to live with misspecification if you must. J Econ 44:67–86

Russo G, Rietveld P, Nijkamp P, Gorter C (2000) Recruitment channel use and applicant arrival: an empirical analysis. Emp Econ 25:673–697

Silverman BW (1981) Using kernel density estimates to investigate multimodality. J R Stat Soc (Series B) 43:97–99

Spanos A (1990) The simultaneous equations model revisited: statistical adequacy and identification. J Econ 44:87–105

Zhu F (2005) A nonparametric analysis of the shape dynamics of the US personal income distribution: 1962–2000. Bank for International Settlements, Working Papers 184