The starred quantities in the b equations are related to the unstarred ones in the a equations in an obvious way. For example, the parameters and error terms of (7.08b) are related to those of (7.08a) as follows:

$$\alpha^* = \alpha^{-1}; \quad \boldsymbol{\beta}^* = -\alpha^{-1}\boldsymbol{\beta}; \quad u_t^{d*} = -\alpha^{-1}u_t^d.$$

We can combine either (7.08a) or (7.08b) with either (7.09a) or (7.09b) when writing the entire model. There are thus *four* different ways that we could write this system of equations, each of them just as valid as any of the others. It is conventional to write simultaneous equations models so that each endogenous variable appears on the left-hand side of one and only one equation, but there is nothing sacrosanct about this convention. Indeed, from the point of view of economic theory, it is probably most natural to combine (7.08a) with (7.09a), putting quantity on the left-hand side of both the demand and supply equations.

We have just seen that **normalization** (i.e., determining which endogenous variable should be given a coefficient of unity and put on the left-hand side of each equation) is necessary whenever we deal with a system of simultaneous equations. Because there are two or more endogenous variables, there is no unique way to write the system. Thus, contrary to what some treatments of the subject may seem to imply, there is no such thing as a single structural form for a linear simultaneous equations model. There are as many structural forms as there are ways in which the equation system can be normalized.

The structural form(s) of a simultaneous equations model are to be contrasted with the **reduced forms**, of which there are two varieties. The **restricted reduced form**, or **RRF**, involves rewriting the model so that each endogenous variable appears once and only once. To derive it in this case, we begin by writing the structural form consisting of (7.08a) and (7.09a):

$$\begin{aligned} Q_t - \alpha P_t &= \boldsymbol{Z}_t^d\boldsymbol{\beta} + u_t^d \\ Q_t - \gamma P_t &= \boldsymbol{Z}_t^s\boldsymbol{\delta} + u_t^s. \end{aligned}$$

These two equations can be rewritten using matrix notation as

$$\begin{bmatrix} 1 & -\alpha \\ 1 & -\gamma \end{bmatrix} \begin{bmatrix} Q_t \\ P_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_t^d\boldsymbol{\beta} \\ \boldsymbol{Z}_t^s\boldsymbol{\delta} \end{bmatrix} + \begin{bmatrix} u_t^d \\ u_t^s \end{bmatrix}.$$

Solving this system for Q_t and P_t, we obtain the restricted reduced form:

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = \begin{bmatrix} 1 & -\alpha \\ 1 & -\gamma \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{Z}_t^d\boldsymbol{\beta} \\ \boldsymbol{Z}_t^s\boldsymbol{\delta} \end{bmatrix} + \begin{bmatrix} 1 & -\alpha \\ 1 & -\gamma \end{bmatrix}^{-1} \begin{bmatrix} u_t^d \\ u_t^s \end{bmatrix},$$

which can be written more explicitly as

$$Q_t = \frac{1}{\alpha - \gamma}\big(\alpha\boldsymbol{Z}_t^s\boldsymbol{\delta} - \gamma\boldsymbol{Z}_t^d\boldsymbol{\beta}\big) + v_t^1 \tag{7.10}$$

$$P_t = \frac{1}{\alpha - \gamma}\big(\boldsymbol{Z}_t^s\boldsymbol{\delta} - \boldsymbol{Z}_t^d\boldsymbol{\beta}\big) + v_t^2, \tag{7.11}$$

where the error terms v_t^1 and v_t^2 are linear combinations of the original error terms u_t^d and u_t^s.

Observe that the equations of the RRF, (7.10) and (7.11), are nonlinear in the parameters but linear in the variables $\boldsymbol{Z}_t^d$ and $\boldsymbol{Z}_t^s$. In fact, they are simply restricted versions of the **unrestricted reduced form**, or **URF**,

$$Q_t = \boldsymbol{Z}_t \boldsymbol{\pi}_1 + v_t^1 \tag{7.12}$$

$$P_t = \boldsymbol{Z}_t \boldsymbol{\pi}_2 + v_t^2, \tag{7.13}$$

where $\boldsymbol{Z}_t$ is a vector consisting of all variables that appear in either $\boldsymbol{Z}_t^d$ or $\boldsymbol{Z}_t^s$, and $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ are parameter vectors. The two equations of the URF can evidently be estimated consistently by OLS, since only exogenous or predetermined variables appear on the right-hand side. The RRF would be harder to estimate, however, since it involves nonlinear cross-equation restrictions. In fact, estimating the RRF is equivalent to estimating the structural form on which it is based, as we will see in Chapter 18.

If we were content simply to estimate the URF, we could stop at this point, since OLS estimates of (7.12) and (7.13) will clearly be consistent.[1] However, economists often want to estimate a structural form of a simultaneous equations model, either because the parameters of that structural form are of interest or because imposing the cross-equation restrictions implicit in the structural form may lead to substantially increased efficiency. Thus it is of interest to ask what happens if we apply OLS to any one of the equations of one of the structural forms. Consider equation (7.08a). The OLS estimates of α and $\boldsymbol{\beta}$ are

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}^\top \boldsymbol{P} & \boldsymbol{P}^\top \boldsymbol{Z}_d \\ \boldsymbol{Z}_d^\top \boldsymbol{P} & \boldsymbol{Z}_d^\top \boldsymbol{Z}_d \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{P}^\top \boldsymbol{Q} \\ \boldsymbol{Z}_d^\top \boldsymbol{Q} \end{bmatrix},$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ denote the vectors of observations on P_t and Q_t, and $\boldsymbol{Z}_d$ denotes the matrix of observations on $\boldsymbol{Z}_t^d$. If we assume that the model is correctly specified and replace $\boldsymbol{Q}$ by $\alpha_0 \boldsymbol{P} + \boldsymbol{Z}_d \boldsymbol{\beta}_0 + \boldsymbol{u}_d$, we find that

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \boldsymbol{\beta}_0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{P}^\top \boldsymbol{P} & \boldsymbol{P}^\top \boldsymbol{Z}_d \\ \boldsymbol{Z}_d^\top \boldsymbol{P} & \boldsymbol{Z}_d^\top \boldsymbol{Z}_d \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{P}^\top \boldsymbol{u}_d \\ \boldsymbol{Z}_d^\top \boldsymbol{u}_d \end{bmatrix}. \tag{7.14}$$

It is obvious that these estimates will be biased and inconsistent. They cannot possibly be unbiased, since the endogenous variable P_t appears on the right-hand side of the equation. They will be inconsistent because

$$\operatorname*{plim}_{n\to\infty} \left(n^{-1} \boldsymbol{P}^\top \boldsymbol{u}_d\right) \neq 0,$$

[1] It may seem that OLS estimation of the URF would be inefficient, because the error terms of (7.12) and (7.13) will clearly be correlated. However, as we will see in Chapter 9, this correlation cannot be exploited to yield more efficient estimates, because the regressors in the two equations are the same.