Before describing the value functions of a household, it is important to have a complete accounting of the state space. The state variables for a household can be divided into objects that are permanent, transitory, endogenous and aggregate.

- **Permanent productivity state.** Each household is endowed with z efficiency units in the urban area and one efficiency unit in the rural area. This is the "static Roy model" aspect of the model.
- **Transitory productivity state.** Each household is subject to transitory productivity shocks, s.
- **Transitory moving shock.** Each household is subject to i.i.d. moving shocks ν.
- **Endogenous state variables.** There are three endogenous (individual) state variables. The first is the household's asset holdings, a. The second is a composite variable that describes the household's location and migration status. The possible states are: rural, seasonal-migrant (living in the rural area but working in the urban area for one period), and urban. The third is whether or not the household is an inexperienced migrant, x, and, thus, whether or not it suffers disutility $\bar{u}$ from locating in the urban area.
- **Aggregate state variables.** There are two aggregate state variables: the season, $i \in \{g, \ell\}$, and the number of workers in the rural area, N_r. The season determines the current and future productivity in the rural area, and jointly, the two aggregate states determine the current wage per efficiency unit as in equation (**??**).

We begin with the problem of a rural household. Because z is time-invariant for each household, we omit it from the formulation of the household's problem below.

Rural Households. A rural household with productivity z solves the following problem:

$$v(a, r, s, \nu, x, i, N_r) = \max \left\{ v(a, r, s, x, i, N_r | \text{ stay}) + \nu^{\text{stay}},\ v(a, r, s, x, i, N_r | \text{ seas}) + \nu^{\text{seas}},\ v(a, r, s, x, i, N_r | \text{ perm}) + \nu^{\text{perm}} \right\}, \quad (1)$$

where a household chooses among staying in the rural area, seasonally moving, and permanently moving. Influencing this choice is the value function associated with each option and the household's taste shock associated with each choice. Here we will follow the literature and assume that these taste shocks are independently and identically distributed across time and is distributed Type 1 extreme value distribution with scale parameter σ_ν.

The distributional assumption on the taste shocks implies that the choice probability to stay in a location, for example, is:

$$P(a, r, s, x, i, N_r | \text{ stay}) = \frac{\exp\{\sigma_\nu^{-1} v(a, r, s, x, i, N_r | \text{ stay}\}}{\sum_{j_r} \exp\{\sigma_\nu^{-1} v(a, r, s, x, i, N_r, |\ j_r)\}}$$

where the sum across j's are the different choices. Here the scale parameter shows up and modulates the strength of the preference shock in determining the move. For example, if σ_ν goes to infinity, then only the shock matters for the moving choice, and the probability of each individual choice is simply one over the number of choices. Another important feature to note with this specification is that the expected value function (with respect to the preference shocks) is

$$E_\nu v(a, r, s, \nu, x, i, N_r) = \sigma_\nu \log \left(\sum_{j_r} \exp\{\sigma_\nu^{-1} v(a, r, s, x, i, N_r, |\ j_r)\} \right).$$

An important feature here is that this is not a simple probability weighted sum. The reason is that the taste shock is "felt" and the household is choosing over it. In other words, it's the expectation over the max operator, hence, this funky feature.

Below, the value functions conditional on a choice are described.

Conditional on staying in the rural area, the value function is:

$$v(a, r, s, x, i, N_r | \text{ stay}) = \max_{a' \in \mathcal{A}} \Big\{ u(Ra + w_r(s, i, N_r) - a') + \beta\, \mathbb{E}[v(a', r, s', \nu', x', i', N_r')] \Big\}, \tag{2}$$

which says that the household chooses future asset holdings to maximize the expected present discounted value of utility. The asset holdings must respect the borrowing constraint and, thus, must lie in the set $\mathcal{A}$. Given asset choices, a household's consumption equals the gross return on current asset holdings, Ra, plus labor income from working in the rural area, $w_r(z, s, i)$, minus future asset holdings. Next period's state variables are the new asset holdings, location in the rural area, the transitory productivity shock, the experience level, the subsequent season, and the aggregate rural efficiency units in the next period. The expectation operator is defined over two uncertain outcomes: the transitory shocks and the change in experience. Recall, that if the household is experienced, it stays that way with probability π and becomes inexperienced with probability $1 - \pi$; if the household is inexperienced, then it stays inexperienced.

The value function associated with a permanent move is:

$$v(a, r, s, x, i, N_r | \text{ perm}) = \max_{a' \in \mathcal{A}} \Big\{ u(Ra + w_r(z, s, i, N_r) - a' - m_p) + \beta\, \mathbb{E}[v(a', u, s', \nu', x', i', N_r')] \Big\}.$$

While similar to the staying value function, there are several points of difference. First, the agent must pay m_p to make the permanent move, and this costs resources. Second, the continuation value function denotes that the household's location changes from the rural to the urban area.

The value function associated with a seasonal move is:

$$v(a, r, s, x, i, N_r | \text{ seas}) = \max_{a' \in \mathcal{A}} \Big\{ u(Ra + w_r(s, i, N_r) - a' - m_T) + \beta\, \mathbb{E}[v(a', seas, s', x', i', N_r')] \Big\}. \tag{3}$$

If a household decides to move seasonally, it pays the moving cost m_T, and works in the urban area in the next period. The key distinction between the permanent move and the seasonal move is that the seasonal move is for just one period. Hence, the location state variable is $seas$ and not u, as this indicates that the household is going to work in the urban area and return in the next period. The value function associated with a seasonal move while in the urban area is:

$$v(a', seas, s', x', i', N_r') = \max_{a'' \in \mathcal{A}} \Big[u(Ra' + w_u(z, s') - a'') \bar{u}^{x'} + \beta\, \mathbb{E}[v(a'', r, s'', \nu'', x'', i'', N_r'')] \Big]. \tag{4}$$

There are several important points to take note of in (4). First, this household has only one choice: how to adjust its asset holdings. By the definition of a seasonal move, the household works in the urban area for one period and then returns to the rural area. Second, note how the disutility from living in the urban area appears (i.e., the presence of $\bar{u}$). Moreover, the state variable of a household's experience x determines whether or not the disutility is experienced.

Equations (3) and (4) illustrate the forces that shape the decision to move seasonally and, in turn, our inferences from the experimental and survey results. Generally, the choice to move seasonally will relate to a household's comparative earnings advantage in the urban area relative to the rural area. However, several forces may lead a household with a permanent comparative advantage in the city not to move. First, the urban disutility may prevent the household from moving, even though its comparative advantage in the urban area is expected to be high. Second, there is risk associated with the move. A household does not know s', and, hence, there is a chance that the income realization in the urban area will not be favorable. Third, the household may have limited assets that simply make a move infeasible or not sufficient to insure against a bad outcome in the urban area.

Urban Households. Urban households face problems similar to those described above, though they choose between just two options: staying or making a permanent move. For a household with productivity level z, the problem is:

$$v(a, u, s, \nu, x, N_r, i) = \max \left\{ v(a, u, s, x, N_r, i | \text{ stay}) + \nu^{\text{stay}},\ v(a, u, s, x, N_r, i | \text{ perm}) + \nu^{\text{perm}} \right\}. \tag{5}$$

Again, influencing this choice is the value function associated with each option and the household's taste shock associated with each choice. These taste shocks are independently and identically distributed across time and distributed Type 1 extreme value distribution with the same scale parameter σ_ν. The distributional assumption on the taste shocks implies that the choice probability to stay in a location, for example, is:

$$P(a, u, s, x, i, N_r | \text{ stay}) = \frac{\exp\{\sigma_\nu^{-1} v(a, u, s, x, N_r, i | \text{ stay})\}}{\sum_{j_u} \exp\{\sigma_\nu^{-1} v(a, u, s, x, N_r, i | \ j_u)\}}$$

where the sum across j_u's are the different choices available to the urban household. As above the expected value function (with respect to the preference shocks) is

$$E_\nu v(a, u, s, \nu, x, i, N_r) = \sigma_\nu \log \left(\sum_{j_u} \exp\{\sigma_\nu^{-1} v(a, u, s, x, i, N_r, | \ j_u)\} \right).$$

NOTE Mechanically, this is lowering utility in the urban area because they have less options. In the urban area you get two draws, yet in the rural area you get three. From a preference standpoint, the rural area will naturally deliver higher expected utility, and will be a force to keep people in the urban area. This force has to be present in the old version of the model, i.e. the rural area will deliver more option value, but this will amplify the option value.

Conditional on staying in the urban area, the value is:

$$v(a, u, s, x, i, N_r, | \text{ stay}) = \max_{a' \in \mathcal{A}} \left\{ u(Ra + w_u(z, s) - a')\bar{u}^x + \beta\, \mathbb{E}[v(a', u, s', x', i', N_r')] \right\}. \tag{6}$$

Households staying in the urban area have several key differences from those staying in the rural area. First, their wage depends on their permanent productivity level, z, and not on the season or number of aggregate efficiency units in the rural areas. Moreover, the transitory productivity shocks may have more or less volatility relative to the rural area, as modulated by the γ parameter (see equation (**??**)). Third, the disutility from living in the urban area appears (i.e., the presence of $\bar{u}$), and the state variable of a household's experience x determines whether or not the disutility is experienced.

Finally, as with rural households, expectations are taken with respect to the transitory shock s and the change in experience. However, as these households are in the urban area, inexperienced households stay that way in the next period with probability λ and become experienced with probability $1-\lambda$. Experienced households retain their experience.

The value function associated with a permanent move back to the rural area is:

$$v(a, u, s, x, i, N_r | \text{ perm}) = \max_{a' \in \mathcal{A}} \left[u(Ra + w_u(z, s) - a' - m_p)\bar{u}^x + \beta\, \mathbb{E}[v(a', r, s', x', i', N_r')] \right]. \tag{7}$$

Here, the agent must pay m_p to make the permanent move. Furthermore, the continuation value function denotes the household's location changes from the urban to the rural area. After a permanent move to the rural area, experienced households keep their experience with probability π and lose it with probability $1 - \pi$.