Some useful properties and formulas for random utility models with logit, nested logit, and ordered nested logit stochastic components

Victor Aguirregabiria*
University of Toronto

January 8, 2010

Abstract

Within the framework of discrete choice Random Utility Models (RUM) with additive stochastic components, this note reviews existing results on closed-form expressions for several key functions: the distribution of the maximum utility, the expected maximum utility, the choice probabilities, and the selection function. The analysis considers three different specifications for the distribution of the stochastic component: i.i.d. type I extreme value distribution, nested extreme value distribution, and ordered generalized extreme value distribution.

1 Random Utility Models

Consider a discrete choice Random Utility Model (RUM) with additive stochastic component. The optimal choice, a^*, is defined as:

$$a^* = \arg \max_{a \in \mathcal{A}} \ \{u_a + \varepsilon_a\} \tag{1}$$

where $\mathcal{A} = \{1, 2, ..., J\}$ is the set of feasible choice alternatives, $\mathbf{u} = (u_1, u_2, ..., u_J)$ is the vector with the deterministic component of the utility, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_J)$ is the vector with the stochastic component. The vector $\boldsymbol{\varepsilon}$ has a joint CDF $G(.)$ that is continuous and strictly increasing with respect to the Lebesgue measure in the Euclidean space.

*Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

This note derives closed-form expressions for the distribution of the maximum utility, $\max_{a\in\mathcal{A}}\{u_a+\varepsilon_a\}$, the expected maximum utility, $\mathbb{E}(\max_{a\in\mathcal{A}}\{u_a+\varepsilon_a\}|u)$, and the choice probabilities, $\Pr(a^*=a|\mathbf{u})$, under three different specifications for the distribution of the vector $\boldsymbol{\varepsilon}$: (1) i.i.d. Type I Extreme Value distribution (MNL model); (2) nested Extreme Value distribution (NL model); and (3) Ordered Generalized Extreme Value distribution (OGEV model).

The following definitions and properties are used in the note.

Definition: *A random variable X has a Double Exponential or Type I Extreme Value distribution with location parameter μ and dispersion parameter σ if its CDF is:*

$$G(x)=\exp\left\{-\exp\left(-\left[\frac{x-\mu}{\sigma}\right]\right)\right\} \tag{2}$$

for any $x\in(-\infty,+\infty)$. ■

Definition: *Maximum utility. Let v^* be the random variable that represents the maximum utility: $v^*\equiv\max_{a\in\mathcal{A}}\{u_a+\varepsilon_a\}$. This maximum utility is a random variable because it depends on the vector of random variables ε.* ■

Definition: *McFadden's Social Surplus function. The social surplus function $S(\mathbf{u})$ is the expected value of the maximum utility conditional on the vector of constants $\mathbf{u}$: $S(\mathbf{u})\equiv\mathbb{E}(\max_{a\in\mathcal{A}}\{u_a+\varepsilon_a\}|u)$.* ■

Definition: *Conditional choice probabilities (CCPs). The conditional choice probability $P(a|\mathbf{u})$ is the probability that alternative a is the optimal choice: $P(a|\mathbf{u})\equiv Pr(a^*=a|\mathbf{u})$.* ■

Definition: *Conditional choice expected utilities (CCEU). The conditional choice expected utility $e(a,\mathbf{u})$ is the expected value of utility $u_a+\varepsilon_a$ conditional on the vector $\mathbf{u}$ and on the event that alternative a is the optimal choice: $e(a,\mathbf{u})\equiv\mathbb{E}(u_a+\varepsilon_a|\mathbf{u},a^*=a)$.* ■

Definition: *Selection function. The selection function $\lambda(a,\mathbf{u})$ is the expected value of the stochastic component of the utility, ε_a, conditional on the vector $\mathbf{u}$ and on the event that alternative a is the optimal choice: $\lambda(a,\mathbf{u})\equiv\mathbb{E}(\varepsilon_a|\mathbf{u},a^*=a)$.* ■

Williams-Daly-Zachary (WDZ) Theorem is an important property of discrete choice RUM with additive stochastic component. It is the discrete-choice version of Roy's Identity in consumer theory. I use this property in several parts of this note. I include here an enunciation of the Theorem and a simple proof.

Williams-Daly-Zachary (WDZ) Theorem. *For any choice alternative $a \in \mathcal{A}$, the CCP $P(a|\mathbf{u})$ can be obtained as the partial derivative of the surplus function $S(\mathbf{u})$ with respect to utility $u(a)$:*

$$P(a|\mathbf{u}) = \frac{\partial S(\mathbf{u})}{\partial u_a} \qquad \blacksquare \tag{3}$$

Proof: By definition of $S(\mathbf{u})$, we have that:

$$\frac{\partial S(\mathbf{u})}{\partial u_a} = \frac{\partial}{\partial u_a} \int \max_{j \in \mathcal{A}} \{u_j + \varepsilon_j\} \; dG(\boldsymbol{\varepsilon}) \tag{4}$$

Given the conditions on the CDF of $\boldsymbol{\varepsilon}$, we can move the partial derivative inside the integral such that:

$$\begin{aligned} \frac{\partial S(\mathbf{u})}{\partial u_a} &= \int \frac{\partial \max_{j \in \mathcal{A}} \{u_j + \varepsilon_j\}}{\partial u_a} \; dG(\boldsymbol{\varepsilon}) \\ &= \int 1\{u_a + \varepsilon_a \geq u_j + \varepsilon_j, \; \forall j \in \mathcal{A}\} \; dG(\boldsymbol{\varepsilon}) \\ &= P(a|\mathbf{u}) \end{aligned} \tag{5}$$

where $1\{.\}$ is the indicator function. $\blacksquare$

I also use the following Theorem to derive several results in this note.

Theorem CEV. *For any distribution of $\boldsymbol{\varepsilon}$, any value of the vector $\mathbf{u}$, and any choice alternative $a \in \mathcal{A}$, we have that:*

$$\mathbb{E}(\varepsilon_a \mid \mathbf{u}, a^* = a) \; = \; S(\mathbf{u}) - u_a \qquad \blacksquare \tag{6}$$

Proof: First, I show that $e(a, \mathbf{u}) = S(\mathbf{u})$. Given that the random variable v^* represents maximum utility, we have that the event $\{a^* = a\}$ implies the event $\{v^* = u_a + \varepsilon_a\}$. Thus,

$$\begin{aligned} e(a, \mathbf{u}) &= \mathbb{E}(u_a + \varepsilon_a \mid \mathbf{u}, a^* = a) \\ &= \mathbb{E}(u_a + \varepsilon_a + v^* - v^* \mid \mathbf{u}, a^* = a) \\ &= \mathbb{E}(u_a + \varepsilon_a + v^* - v^* \mid \mathbf{u}, v^* = u_a + \varepsilon_a) \\ &= \mathbb{E}(v^* \mid \mathbf{u}) \;=\; S(\mathbf{u}) \end{aligned} \tag{7}$$

By definition, $e(a, \mathbf{u}) = u_a + \mathbb{E}(\varepsilon_a \mid u, a^* = a)$. Therefore, equation (7) implies that $\mathbb{E}(\varepsilon_a \mid \mathbf{u}, a^* = a) = S(\mathbf{u}) - u_a$. ■

2 Multinomial logit (MNL)

Suppose that the random variables in the vector $\boldsymbol{\varepsilon}$ are i.i.d. with Type I Extreme Value distribution with a location parameter $\mu = 0$ and unrestricted dispersion parameter σ. That is, for every alternative $a \in \mathcal{A}$, the CDF of ε_a is $G(\varepsilon_a) = \exp\left\{-\exp\left(-\frac{\varepsilon_a}{\sigma}\right)\right\}$.

2.1 Distribution of the maximum utility

The maximum utility v^* is a random variable because it depends on the vector of random variables ε. By definition, the cumulative probability distribution of v^* is:

$$\begin{aligned} F_{v^*}(v) \equiv \Pr(v^* \leq v) &= \textstyle\prod_{a \in \mathcal{A}} \Pr(u_a + \varepsilon_a \leq v) \\ &= \textstyle\prod_{a \in \mathcal{A}} \exp\left\{-\exp\left(-\dfrac{v - u_a}{\sigma}\right)\right\} \\ &= \exp\left\{-\exp\left(-\dfrac{v}{\sigma}\right)\; U\right\} \end{aligned} \tag{8}$$

where $U \equiv \sum_{a \in \mathcal{A}} \exp\left(\dfrac{u_a}{\sigma}\right)$. We can also write this expression as:

$$F_{v^*}(v) = \exp\left\{-\exp\left(-\frac{v - \sigma \, \ln U}{\sigma}\right)\right\} \tag{9}$$

This expression shows that the maximum utility v^* is a double exponential random variable with dispersion parameter σ and location parameter $\sigma \, \ln U$. Therefore, the maximum of a vector of i.i.d. double exponential random variables is also a double exponential random variable. This is the reason why this family of random variables is also called "extreme value". The density function of v^* is:

$$f_{v^*}(v) \equiv H'(v) \quad = \quad F_{v^*}(v) \; \frac{U}{\sigma} \; \exp\left(-\frac{v}{\sigma}\right) \tag{10}$$

2.2 Expected maximum utility

By definition, $S(\mathbf{u}) = \mathbb{E}(v^*|\mathbf{u})$. Therefore,

$$S(\mathbf{u}) \quad = \quad \int v^* \; h(v^*) \; dv^* \quad = \quad \int v^* \; \exp\left\{- \exp\left(-\frac{v^*}{\sigma}\right) U\right\} \; \frac{U}{\sigma} \; \exp\left(-\frac{v^*}{\sigma}\right) \; dv^* \tag{11}$$

Applying the change in variable $z = \exp(-v^*/\sigma)$, such that $v^* = -\sigma \ln(z)$, and $dv^* = -\sigma(dz/z)$, we have:

$$\begin{aligned} S(\mathbf{u}) \quad &= \quad \int_{+\infty}^{0} -\sigma \ln(z) \; \exp\{-z \; U\} \; \frac{U}{\sigma} \; z \; \left(-\sigma \frac{dz}{z}\right) \\ &= \quad -\sigma U \int_{0}^{+\infty} \ln(z) \; \exp\{-z \; U\} \; dz \end{aligned} \tag{12}$$

Using Laplace transformation we have that $\int_0^{+\infty} \ln(z) \; \exp\{-z \; U\} \; dz = \dfrac{\ln(U) + \gamma}{U}$, where γ is Euler's constant. Therefore, the expected maximum utility is:

$$S(\mathbf{u}) \quad = \quad \sigma U \left(\frac{\ln(U) + \gamma}{U}\right) \quad = \quad \sigma \; (\; \ln(U) + \gamma \;) \tag{13}$$

2.3 Choice probabilities

By Williams-Daly-Zachary (WDZ) theorem, the optimal choice probabilities can be obtained by differentiating the surplus function. Therefore, for the MNL model,

$$\begin{aligned} P(a|\mathbf{u}) \quad &= \quad \sigma \; \frac{\partial \ln(U)}{\partial u_a} \qquad = \sigma \; \frac{\partial U}{\partial u_a} \; \frac{1}{U} \\ &= \quad \exp\left(\frac{u_a}{\sigma}\right) \; \frac{1}{U} \quad = \frac{\exp\left(u_a/\sigma\right)}{\sum_{j \in \mathcal{A}} \exp\left(u_j/\sigma\right)} \end{aligned} \tag{14}$$

2.4 Conditional choice expected utilities

As shown in Theorem CEV, $\mathbb{E}(\varepsilon_a|\mathbf{u}, a^* = a) = S(\mathbf{u}) - u_a$. For the case of the i.i.d. double exponential ε we have that:

$$\mathbb{E}(\varepsilon_a|\mathbf{u}, a^* = a) = \sigma \;\left(\; \ln(U) + \gamma \;\right) - u_a \tag{15}$$

2.5 Relationship between selection function and CCPs

In some applications, we are interested in the selection function that relates the expected value $\mathbb{E}(\varepsilon_a|\mathbf{u}, a^* = a)$ with the conditional choice probabilities. From the expression for $P(a|\mathbf{u})$ in the MNL model, we have that $\ln P(a|\mathbf{u}) = u_a/\sigma - \ln U$, and therefore $\ln(U) = u_a/\sigma - \ln P_a$. Solving this expression in equation (15) we get:

$$\mathbb{E}(\varepsilon_a|\mathbf{u}, a^* = a) = \sigma \;\left(u_a/\sigma - \ln P(a|\mathbf{u}) + \gamma\right) - u_a = \sigma\left(\gamma - \ln P(a|\mathbf{u})\right) \tag{16}$$

3 Nested logit (NL)

Suppose that the random variables in the vector ε have the following joint CDF:

$$G(\boldsymbol{\varepsilon}) = \exp\left\{ -\sum_{r=1}^{R} \left[\sum\nolimits_{a \in \mathcal{A}_r} \exp\left(-\frac{\varepsilon_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \tag{17}$$

where $\{\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_R\}$ is a partition of $\mathcal{A}$, and δ, σ_1, σ_2, ..., σ_R are positive parameters, with $\delta \leq 1$.

3.1 Distribution of the Maximum Utility

Using the same approach as for the MNL model, we have:

$$
\begin{aligned}
F_{v^*}(v) \equiv \Pr(v^* \leq v) &= \textstyle\prod_{a\in\mathcal{A}} \Pr(u_a + \varepsilon_a \leq v, \forall\, a \in \mathcal{A}) \\
&= \textstyle\prod_{a\in\mathcal{A}} \exp\left\{ - \sum\limits_{r=1}^{R} \left[\sum_{a\in A_r} \exp\left(- \dfrac{v-u_a}{\sigma_r} \right) \right]^{\dfrac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ - \exp\left(- \dfrac{v}{\delta} \right) \sum\limits_{r=1}^{R} \left[\sum_{a\in\mathcal{A}_r} \exp\left(\dfrac{u_a}{\sigma_r} \right) \right]^{\dfrac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ - \exp\left(- \dfrac{v}{\delta} \right) U \right\}
\end{aligned}
\tag{18}
$$

where:

$$
U \equiv \sum_{r=1}^{R} \left[\sum_{a\in\mathcal{A}_r} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\dfrac{\sigma_r}{\delta}} = \sum_{r=1}^{R} U_r^{1/\delta} \tag{19}
$$

and

$$
U_r \equiv \left[\sum_{a\in\mathcal{A}_r} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\sigma_r} \tag{20}
$$

The density function of v^* is:

$$
f_{v^*}(v) \equiv H'(v) \;=\; F_{v^*}(v)\, \frac{U}{\delta}\, \exp\left(- \frac{v}{\delta} \right) \tag{21}
$$

3.2 Expected maximum utility

By definition, $S(\mathbf{u}) = \mathbb{E}(v^*)$. Therefore,

$$
S(\mathbf{u}) \;=\; \textstyle\int_{-\infty}^{+\infty} v^*\; h(v^*)\; dv^* \;=\; \textstyle\int_{-\infty}^{+\infty} v^*\; \exp\left\{ - \exp\left(- \dfrac{v^*}{\delta} \right) U \right\} \dfrac{U}{\delta}\; \exp\left(- \dfrac{v^*}{\delta} \right)\; dv^* \tag{22}
$$

Let's apply the following change in variable: $z = \exp(-v^*/\delta)$, such that $v^* = -\delta \ln(z)$, and $dv^* = -\delta(dz/z)$. Then,

$$
S(\mathbf{u}) \;=\; \textstyle\int_{+\infty}^{0} -\delta \ln(z)\; \exp\{-z\, U\}\; \dfrac{U}{\delta}\; z\; \left(-\delta \dfrac{dz}{z} \right) \;=\; -\delta U \textstyle\int_{+\infty}^{0} \ln(z)\; \exp\{-z\, U\}\; dz \tag{23}
$$

And using Laplace transformation:

$$S(\mathbf{u}) \;=\; \delta U \left(\frac{\ln(U) + \gamma}{U} \right) \;=\; \delta \;(\; \ln(U) + \gamma \;) \tag{24}$$

where γ is the Euler's constant.

3.3 Choice probabilities

By Williams-Daly-Zachary (WDZ) theorem, choice probabilities can be obtained differentiating the surplus function. For the NL model:

$$\begin{aligned} P(a|\mathbf{u}) &= \delta \frac{\partial \ln(U)}{\partial u_a} = \delta \frac{\partial U}{\partial u_a} \frac{1}{U} = \\ &= \delta \frac{\sigma_{ra}}{\delta} \left[\sum_{j \in A_{ra}} \exp\left(\frac{u_j}{\sigma_{ra}} \right) \right]^{\frac{\sigma_{ra}}{\delta} - 1} \frac{1}{\sigma_{ra}} \exp\left(\frac{u_a}{\sigma_{ra}} \right) \frac{1}{U} \\ &= \frac{\exp\left(u_a/\sigma_{ra}\right)}{\sum_{j \in \mathcal{A}_{ra}} \exp\left(u_j/\sigma_{ra}\right)} \; \frac{\left[\sum_{j \in A_{ra}} \exp\left(u_j/\sigma_{ra}\right) \right]^{\frac{\sigma_{ra}}{\delta}}}{\sum_{r=1}^{R} \left[\sum_{j \in \mathcal{A}_r} \exp\left(u_j/\sigma_r\right) \right]^{\frac{\sigma_r}{\delta}}} \end{aligned} \tag{25}$$

The first term is $q(a|r_a)$ (i.e., probability of choosing a given that we are in group A_{ra}), and the second term is $Q(r_a)$ (i.e., probability of selecting the group A_{ra}).

3.4 Conditional choice expected utilities

As shown in general, $e(a, \mathbf{u}) = S(\mathbf{u})$. This implies that $\mathbb{E}(\varepsilon_a \mid u,\, a^* = a) = S(\mathbf{u}) - u_a$. Given that for the NL model $S(\mathbf{u}) = \delta\left(\ln U + \gamma\right)$ we have that:

$$\mathbb{E}(\varepsilon_a | u, a^* = a) = \delta\gamma + \delta \ln U - u_a \tag{26}$$

3.5 Relationship between selection function and CCPs

To write $\mathbb{E}(\varepsilon_a | u, a^* = a)$ in terms of choice probabilities, note that from the definition of $q(a|r_a)$ and $Q(r_a)$, we have that:

$$\ln q(a|r_a) \;=\; \frac{u_a - \ln U_{ra}}{\sigma_{ra}} \;\Rightarrow\; \ln U_{ra} \;=\; u_a - \sigma_{ra} \ln q(a|r_a) \tag{27}$$

and

$$\ln Q(r_a) = \frac{\ln U_{ra}}{\delta} - \ln U \quad \Rightarrow \quad \ln U = \frac{\ln U_{ra}}{\delta} - \ln Q(r_a) \tag{28}$$

Combining these expressions, we have that:

$$\ln U = \frac{u_a - \sigma_{ra} \ln q(a|r_a)}{\delta} - \ln Q(r_a) \tag{29}$$

Therefore,

$$\begin{aligned} e_a &= \delta\gamma + \delta \left(\frac{u_a - \sigma_{ra} \ln q(a|r_a)}{\delta} - \ln Q(r_a) \right) - u_a \\ &= \delta\gamma - \sigma_{ra} \ \ln q(a|r_a) - \delta \ \ln Q(r_a) \end{aligned}$$

4 Ordered GEV (OGEV)

Suppose that the random variables in the vector $\boldsymbol{\varepsilon}$ have the following joint CDF:

$$G(\boldsymbol{\varepsilon}) = \exp\left\{ -\sum_{r=1}^{J+M} \left[\sum\nolimits_{a \in B_r} W_{r-a} \exp\left(-\frac{\varepsilon_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \tag{30}$$

where:

- M is a positive integer;
- $\{B_1, B_2, ..., B_{J+M}\}$ are $J + M$ subsets of A, with the following definition:

  $$B_r = \{a \in \mathcal{A} : r - M \leq a \leq r\} \tag{31}$$

  For instance, if $A = \{1, 2, 3, 4, 5\}$ and $M = 2$, then $B_1 = \{1\}$, $B_2 = \{1, 2\}$, $B_3 = \{1, 2, 3\}$, $B_4 = \{2, 3, 4\}$, $B_5 = \{3, 4, 5\}$, $B_6 = \{4, 5\}$, and $B_7 = \{5\}$.
- δ, and σ_1, σ_2, ..., σ_{J+M} are positive parameters, with $\delta \leq 1$;
- W_0, W_1, ..., W_M are constants (weights) such that: $W_m \geq 0$, and $\sum_{m=0}^{M} W_m = 1$.

4.1 Distribution of the Maximum Utility

$$
\begin{aligned}
F_{v^*}(v) \equiv \Pr(v^* \leq v) &= \Pr(\varepsilon_a \leq v - u_a : for\ any\ a \in \mathcal{A}) \\
&= \exp\left\{ - \sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(- \frac{v - u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ - \exp\left(- \frac{v}{\delta} \right) \sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ - \exp\left(- \frac{v}{\delta} \right) U \right\}
\end{aligned}
\tag{32}
$$

where:

$$
U \equiv \sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} = \sum_{r=1}^{J+M} U_r^{1/\delta} \tag{33}
$$

where $U_r \equiv \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\sigma_r}$. The density function of v^* is:

$$
f_{v^*}(v) \equiv H'(v) \;=\; F_{v^*}(v) \, \frac{U}{\delta} \, \exp\left(- \frac{v}{\delta} \right) \tag{34}
$$

4.2 Expected maximum utility

By definition, $S(\mathbf{u}) = \mathbb{E}(v^*|u)$. Therefore,

$$
S(\mathbf{u}) \;=\; \int_{-\infty}^{+\infty} v^* \; h(v^*) \; dv^* \;=\; \int_{-\infty}^{+\infty} v^* \; \exp\left\{ - \exp\left(- \frac{v^*}{\delta} \right) U \right\} \frac{U}{\delta} \; \exp\left(- \frac{v^*}{\delta} \right) \; dv^* \tag{35}
$$

Let's apply the following change in variable: $z = \exp(-v^*/\delta)$, such that $v^* = -\delta \ln(z)$, and $dv^* = -\delta(dz/z)$. Then,

$$
S \;=\; \int_{+\infty}^{0} -\delta \ln(z) \; \exp\{-z\,U\} \; \frac{U}{\delta} \; z \left(-\delta \frac{dz}{z} \right) \;=\; -\delta U \int_{0}^{+\infty} \ln(z) \; \exp\{-z\,U\} \; dz \tag{36}
$$

And using Laplace transformation:

$$
S \;=\; \delta U \left(\frac{\ln U + \gamma}{U} \right) \;=\; \delta \left(\ln U + \gamma \right) = \delta\gamma + \delta \; \ln \left[\sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right] \tag{37}
$$

where γ is the Euler's constant.

4.3 Choice probabilities

By Williams-Daly-Zachary (WDZ) theorem, choice probabilities can be obtained differentiating the surplus function.

$$P(a|u) = \frac{1}{U} \sum_{r=a}^{a+M} \left[\sum_{j \in B_r} W_{r-j} \exp\left(\frac{u_j}{\sigma_r}\right) \right]^{\frac{\sigma_r}{\delta}-1} W_{r-a} \exp\left(\frac{u_a}{\sigma_r}\right) = \sum_{r=a}^{a+M} q(a|r)\ Q(r) \tag{38}$$

where:

$$\begin{aligned} q(a|r) &= \frac{W_{r-a}\ \exp\left(u_a/\sigma_r\right)}{\sum_{j \in B_r} W_{r-j} \exp\left(u_j/\sigma_r\right)} = \frac{\exp\left(u_a/\sigma_r\right)}{\exp(\ln U_r/\sigma_r)} \\ Q(r) &= \frac{\exp\left(\ln U_r/\delta\right)}{\sum_{j=1}^{J+M} \exp\left(\ln U_j/\delta\right)} = \frac{\exp\left(\ln U_r/\delta\right)}{U} \end{aligned} \tag{39}$$

4.4 Conditional choice expected utilities

As shown in general, $e(a, \mathbf{u}) = S(\mathbf{u})$. This implies that $\mathbb{E}(\varepsilon_a \mid u,\, a^* = a) = S(\mathbf{u}) - u_a$. Given that for the OGEV model $S(\mathbf{u}) = \delta\left(\ln U + \gamma\right)$ we have that:

$$\mathbb{E}(\varepsilon_a | u, a^* = a) = \delta\gamma + \delta \ln U - u_a \tag{40}$$