

IZA DP No. 1015

**Analytical Prediction of Transitions
Probabilities in the Conditional Logit Model**

Holger Bonin
Hilmar Schneider

February 2004

Analytical Prediction of Transitions Probabilities in the Conditional Logit Model

Holger Bonin

IZA Bonn

Hilmar Schneider

IZA Bonn

Discussion Paper No. 1015
February 2004

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

ABSTRACT

Analytical Prediction of Transitions Probabilities in the Conditional Logit Model

The paper derives analytical transitions probabilities following an exogenous shock to the deterministic component in the conditional logit model. The solution draws on the post-estimation distribution of the model's stochastic component, identified on the basis of a direct utility maximization interpretation of agents' revealed choice. Computational experiments confirm that analytical prediction of transitions probabilities might perform substantially better than the established calibration method with few repetitions. However, results obtained in an empirical application studying labor supply responses to social insurance reform in Germany suggest that previous calibration-based results accurately indicate the direction of incentive effects, while underpredicting small transitions frequencies.

JEL Classification: C35, C15, J22

Keywords: conditional logit model, logistic distributions, behavioral microsimulation, social insurance reform

Corresponding author:

Holger Bonin
IZA
P.O. Box 7240
53072 Bonn
Germany
Email: bonin@iza.org

1 Introduction

This paper derives an analytical form for transitions probabilities following an exogenous shock to the deterministic component in the conditional logit model (McFadden 1974), a commonplace discrete choice framework in empirical economics. The premise behind discrete choice models is that agents in many circumstances cannot optimize to a continuous or unconstrained decision variable, but only choose from a discrete set of exclusive states. One example is the labor supply decision where agents typically face a limited range of hours contracts, like full-time and part-time. An advantage of the discrete choice framework is that the budget set of the sample units needs to be determined only at a relatively small number of discrete points. It is perhaps mainly for this reason that following van Soest (1995) and Duncan and Giles (1996) qualitative models are increasingly in use for behavioral microsimulation of transitions effected by economic policy reforms.

Despite the popularity of the approach, the methods to predict state transitions in discrete choice models are somewhat underdeveloped. Creedy and Duncan (2002) distinguish between three approaches, which might be called the aggregate probability method, the maximum probability method and the calibration method. The aggregate probability approach is the most common one. Here the choice probabilities of each state, predicted by agent, are summed up over the entire sample, in order to compare the aggregate probabilities of each state before and after the simulated shock. A serious shortcoming of this method is that it ignores the probabilistic nature of state transitions at the individual level. As a result, one can gauge neither the impact of the shock in terms of individual costs or benefits, nor the specific response of agents who initially occupy a particular state.

The maximum probability method improves upon the aggregate probability method by allocating each agent to a unique state, which is the state with the highest choice probability conditional on observable characteristics. Transitions probabilities are derived from comparing the most likely state for each individual before and after the shock, which leads to a matrix of transitions frequencies. A problem with this approach is that it wastes information. It ignores the relative distance of the predicted probability levels and moreover the impact of unobserved heterogeneity, as the observed state is not necessarily the state with the highest predicted choice probability. As a consequence, especially when some states are sparsely populated, even well-specified discrete choice models have a tendency

for systematic underprediction and overprediction of certain state frequencies, which in turn biases simulated transitions frequencies.

This was shown by Duncan and Weeks (1998) who propose to approximate transitions probabilities on the basis of calibration. The first step is to draw realizations of the latent stochastic elements in the discrete choice model such that the predicted choice probability is maximized at the observed state. The calibration guarantees that all transitions start from a situation where each sample unit occupies its chosen state. In a second step, the maximum probability rule is applied to derive the preferred choice after the simulated shock given the calibrated random terms. If these two steps are repeated several times, individual transitions probabilities can be approximated by taking the mean of the predicted transitions between states over the repetitions. Although Monte Carlo simulations performed by Duncan and Weeks (1998) indicate that the calibration method might generate satisfactory estimates of transitions frequencies, it is comparatively rare in behavioral microsimulation studies; see Gerfin and Leu (2003) and Bargain (2003) for examples. The main drawback of the method is that resampling of the error terms will be computationally time-consuming in many applications. However, if only a small number of replications is used, the approximation of transitions probabilities obtained by calibration might be unreliable.

The present paper solves this problem for the conditional logit framework. We start from an analytical representation of the distribution of the unobserved heterogeneity terms, which the calibration method approximates through resampling. Post-estimation, on the basis of a direct utility maximization interpretation of agents' revealed choice, the distribution of the random error terms conditional on the estimated model parameters is identified in relative terms— the possible values of the prediction error in each state are determined relative to the value of the prediction error in the utility maximizing state. Transitions probabilities in response to a shock to the deterministic part in the model can be derived from the joint conditional distribution of the relative conditional prediction errors, by integrating over the relevant range of relative utility shifts induced by the shock. While this strategy would apply to any discrete choice model, it yields analytical transitions probabilities only in the special case of the conditional logit model. Here the specific distributional assumption on the unobserved random disturbances keeps the solution feasible irrespective of the number of choice dimensions.

A comparison of analytical versus calibrated transitions probabilities using stylized examples indicates that the latter might only converge slowly to the former. In order to judge the reliability of results found in the previous literature, we compare analytical and calibrated transitions frequencies in a typical empirical application, which studies the labor supply impact of an income allowance for social security contributions using German data. The example suggests that results obtained by calibration probably accurately reflect the direction of the incentive effects, while slightly underestimating small transitions frequencies. Of course, the analytical solution is preferable, not the least as a way of circumventing the burdensome resampling procedure.

The structure of the paper is as follows. Section 2 reviews the conditional logit framework and derives an analytical form for transitions probabilities following an exogenous shock. Section 3 compares the predictive performance of the analytical and the calibration procedures. Section 4 concludes.

2 Transitions Probabilities in the Conditional Logit Model

Consider a framework where decision-making is described as the utility maximizing choice among J exclusive states. Assume that the level of utility U_{ij} enjoyed by agent i in state j can be expressed in terms of a vector of observable characteristics X_{ij} , which might be state variant, a vector of parameters β , and an unobservable component ε_{ij} , such that

$$U_{ij} = U(X_{ij}|\beta) + \varepsilon_{ij}, \quad j = 1, \dots, J. \quad (1)$$

Suppose that the random disturbance terms ε_{ij} are independent across states and agents, and that they are identically Type I Extreme Value distributed. This means that the cumulative density of ε_{ij} is given by $\Pr[\varepsilon_{ij} < \varepsilon] = \exp[-\exp(-\varepsilon)]$, for any $\varepsilon \in \mathbb{R}$. Furthermore suppose that agents choose the state with the largest utility across the possible states. From these assumptions, it follows that the probability $P_i(m)$ that agent i is observed to choose state m is given by (Maddala 1986):

$$P_i(m) = P[U_{im} > U_{ij} \forall j \neq m] = \frac{\exp[U(X_{im}|\beta)]}{\sum_{ij} \exp[U(X_{ij}|\beta)]}, \quad (2)$$

which summarizes the so-called conditional logit model, a variant of the multinomial logit model. The advantage of the conditional logit specification is that the parameters of the

distribution of the random disturbance terms are known *a priori* and do not need to be estimated.

Given observations on the choices made by a set of N agents and a set of characteristics in each possible state, the behavioral parameters entering the deterministic component in the utility function (1) can be estimated by maximizing the likelihood

$$\ell(\beta|X_{ij}) = \prod_{i=1}^N P_{im}(\beta|X_{ij}) \quad (3)$$

over β . Denote the maximum likelihood estimate of the parameter vector by $\hat{\beta}$, and the observable utility component in state j , predicted for an agent with characteristics X_{ij} , by $\hat{U}_{ij} = U(X_{ij}|\hat{\beta})$.

For the exact calculation of transitions frequencies following an exogenous shock on an agent's observable characteristics, it is necessary to account for the fact that the predicted value of utility \hat{U}_{ij} deviates from the latent value by the random component ε_{ij} . Post-estimation, distributions of the unobserved random disturbances can be recovered by exploiting the condition that the observed choice of the agent maximizes utility. Hence, it must be true that $U_{im} > U_{ik}$, or, inserting definition (1), that $\varepsilon_{ik} - \varepsilon_{im} < \hat{U}_{im} - \hat{U}_{ik}$, for all $k \neq m$. One approach to calibrate the model would be to rely on the so-called generalized residual $E[\varepsilon_{ik} - \varepsilon_{im} | \varepsilon_{ik} - \varepsilon_{im} < \hat{U}_{im} - \hat{U}_{ik}]$, i.e. the conditional expectation of the unobserved utility differential (Gourieroux et al. 1987). The calibration method instead obtains an approximation of the conditional distribution of the latent difference in the random model components, by repeatedly drawing at random vectors of pseudo-residuals $(\varepsilon_{i1}, \dots, \varepsilon_{iJ})$ satisfying $\hat{U}_{im} + \varepsilon_{im} = \max_j \hat{U}_{ij} + \varepsilon_{ij}$.

However, given the specific distribution of the random components in the conditional logit model, it is possible to derive an explicit form for the conditional probability distribution $f(\varepsilon_{ik} - \varepsilon_{im} | \varepsilon_{ik} - \varepsilon_{im} < \hat{U}_{im} - \hat{U}_{ik})$, irrespective of the dimension of the analyzed choice problem. In doing so it is important to note that the support of the conditional difference in prediction errors depends on the random disturbances in the possible alternative states as well. The condition $\varepsilon_{il} - \varepsilon_{im} < \hat{U}_{im} - \hat{U}_{il}$, for all $l \neq k, m$, imposes $J - 2$ additional constraints on the unobserved heterogeneity term ε_{im} .

To clarify the exposition, we first analyze the special case of a trinomial choice problem with states $j = k, l, m$. Results can then easily be extended to higher dimensions. In the trinomial case, it suffices to consider the conditional bivariate distribution $f(\varepsilon_{km}, \varepsilon_{lm} | \varepsilon_{km} <$

$\hat{U}_{mk}, \varepsilon_{lm} < \hat{U}_{ml}$), where we simplify notation by eliminating the agent index i , defining $\varepsilon_{km} = \varepsilon_{ik} - \varepsilon_{im}$, $\hat{U}_{mk} = \hat{U}_{im} - \hat{U}_{ik}$, etc. Using the results that the distribution of the difference of two Type I Extreme Value distributed variables is a logistic distribution, and that the joint distribution of two variables with logistic distribution is a bivariate logistic distribution (Kotz, Balakrishnan, and Johnson 2000), the joint distribution of the conditional prediction errors is given by

$$\begin{aligned} f(\varepsilon_{km}, \varepsilon_{lm} | \varepsilon_{km} < \hat{U}_{mk}, \varepsilon_{lm} < \hat{U}_{ml}) &= \frac{f(\varepsilon_{km}, \varepsilon_{lm})}{F(\hat{U}_{mk}, \hat{U}_{ml})} \\ &= \frac{2 \exp(-\varepsilon_{km}) \exp(-\varepsilon_{lm}) [1 + \exp(-\hat{U}_{mk}) + \exp(-\hat{U}_{ml})]}{[1 + \exp(-\varepsilon_{km}) + \exp(-\varepsilon_{lm})]^3} \end{aligned} \quad , \quad (4)$$

or, equivalently,

$$f(U_{km}, U_{lm} | U_{km} < 0, U_{lm} < 0) = \frac{2 \exp(\hat{U}_{km} - U_{km}) \exp(\hat{U}_{lm} - U_{lm}) [1 + \exp(\hat{U}_{km}) + \exp(\hat{U}_{lm})]}{[1 + \exp(\hat{U}_{km} - U_{km}) + \exp(\hat{U}_{lm} - U_{lm})]^3} \quad . \quad (5)$$

Recovering the c.d.f. corresponding to (5) is straightforward:

$$F(U_{km}, U_{lm} | U_{km} < 0, U_{lm} < 0) = \frac{1 + \exp(\hat{U}_{km}) + \exp(\hat{U}_{lm})}{1 + \exp(\hat{U}_{km} - U_{km}) + \exp(\hat{U}_{lm} - U_{lm})} \quad . \quad (6)$$

Now assume an exogenous shock that changes the set of the agent's observable characteristics, from X_j to \tilde{X}_j , and denote the corresponding predicted utility level by $\tilde{U}_j = U(\tilde{X}_j | \hat{\beta})$. Given that the shock does not affect the unobserved random disturbance terms, as it is usually assumed in behavioral microsimulation studies, the actual utility changes, relevant for the transitions frequencies between states, are equal to the predicted utility changes. Denote the change of utility enjoyed in state j following the shock by $\Delta_j = \tilde{U}_j - \hat{U}_j$. To find the probability that the agent stays in the initial state m , $P(m|m)$, one has to evaluate the probability that $U_m + \Delta_m > U_j + \Delta_j$, for $j = k, l$, conditional on m being selected prior to the shock

$$P(m|m) = P(U_{km} < \Delta_{mk}, U_{lm} < \Delta_{ml} | U_{km} < 0, U_{lm} < 0) \quad , \quad (7)$$

where we introduce the notation $\Delta_{mk} = \Delta_m - \Delta_k$, etc. Provided that $\Delta_{mk} < 0$ and $\Delta_{ml} < 0$, one can immediately calculate this probability by evaluating the cumulative density (6) at the pair $(\Delta_{mk}, \Delta_{ml})$.

The initial conditions, $U_{km} < 0$, $U_{lm} < 0$, constrain the net utility changes between the original state and the alternative states that can enter into the transitions frequencies to non-positive values. In other words, the probability that the agent prefers state m over

state l is not larger if $\Delta_{ml} > 0$ than if $\Delta_{ml} = 0$. The intuition for this result is that if $\Delta_{ml} > 0$, the shock makes any initial utility gap between states m and l even larger. Therefore there is a zero probability that the agent moves from state m to state l . The bivariate probability (7) collapses to the marginal probability $P(U_{km} < \Delta_{mk}|U_{km} < 0)$. Integrating (6) over the range of U_{lm} yields the following marginal c.d.f.

$$F(U_{km}|U_{km} < 0) = \frac{1 + \exp(\hat{U}_{km}) + \exp(\hat{U}_{lm})}{1 + \exp(\hat{U}_{km} - U_{km}) + \exp(\hat{U}_{lm})}. \quad (8)$$

A parallel argument can be made if $\Delta_{mk} > 0$. As a corollary, if $\Delta_{mk} > 0$ and $\Delta_{ml} > 0$, the probability that the agent remains in the original state m is invariably identical to unity.

As a next step, consider the probability that the agent responds to the shock by making a transition from state m to state k . First assume that $\Delta_{mk} < 0$ and $\Delta_{ml} < 0$ so that the transitions probabilities from state m to state k and from state m to state l are positive. The probability that the agent post reform is in state k and not state l is equal to the probability that $U_k + \Delta_k > U_m + \Delta_m$ and $U_k + \Delta_k > U_l + \Delta_l$ conditional on $U_{km} < 0$ and $U_{lm} < 0$, which can be stated as

$$P(k|m) = P(U_{km} > \Delta_{mk}, U_{lm} < U_{km} + \Delta_{ml} - \Delta_{mk}|U_{km} < 0, U_{lm} < 0). \quad (9)$$

The value of this probability depends on the relative order of Δ_{ml} and Δ_{mk} . If $\Delta_{ml} \leq \Delta_{mk}$, the upper bound on U_{lm} , i.e. $U_{km} + \Delta_{ml} - \Delta_{mk}$, can never violate against the condition $U_{lm} < 0$ since U_{km} must satisfy the condition $U_{km} < 0$. Therefore the probability is obtained by solving the integral over the density function (5), as follows¹

$$\begin{aligned} P(k|m) &= \int_{\Delta_{mk}}^0 \int_{-\infty}^{U_{km} + \Delta_{ml} - \Delta_{mk}} f(U_{km}, U_{lm}|U_{km} < 0, U_{lm} < 0) dU_{lm} dU_{km} \\ &= \frac{[1 + \exp(\hat{U}_{km}) + \exp(\hat{U}_{lm})] \exp(\hat{U}_{km})}{[\exp(\hat{U}_{km}) + \exp(\hat{U}_{lm} - \Delta_{ml} + \Delta_{mk})][1 + \exp(\hat{U}_{km}) + \exp(\hat{U}_{lm} - \Delta_{ml} + \Delta_{mk})]} - \\ &\quad \frac{[1 + \exp(\hat{U}_{km}) + \exp(\hat{U}_{lm})] \exp(\hat{U}_{km})}{[\exp(\hat{U}_{km}) + \exp(\hat{U}_{lm} - \Delta_{ml} + \Delta_{mk})][1 + [\exp(\hat{U}_{km}) + \exp(\hat{U}_{lm} - \Delta_{ml} + \Delta_{mk})] \exp(-\Delta_{mk})]}. \end{aligned} \quad (10)$$

On the other hand, if $\Delta_{ml} > \Delta_{mk}$, the unconstrained upper bound on U_{lm} is positive for realizations of U_{km} larger than $\Delta_{mk} - \Delta_{ml}$. In these cases, the constraint $U_{lm} < 0$ limits the possible range of the variable. This leads to a different solution for the probability of a

¹ The procedures for solving the multi-variate integrals leading to the transitions frequencies discussed in this paper are straightforward. The algebra can be obtained from the authors upon request.

transition from state m to state k :

$$\begin{aligned}
P(k|m) &= \int_{\Delta_{mk}}^{\Delta_{mk}-\Delta_{ml}} \int_{-\infty}^{U_{km}+\Delta_{ml}-\Delta_{mk}} f(U_{km}, U_{lm}|U_{km} < 0, U_{lm} < 0) dU_{lm} dU_{km} \\
&\quad + \int_{\Delta_{mk}-\Delta_{ml}}^0 \int_{-\infty}^0 f(U_{km}, U_{lm}|U_{km} < 0, U_{lm} < 0) dU_{lm} dU_{km} \\
&= 1 - \frac{1+\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})}{1+\exp(\hat{U}_{km}+\Delta_{ml}-\Delta_{mk})+\exp(\hat{U}_{lm})} + \\
&\quad \frac{[1+\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})] \exp(\hat{U}_{km})}{[\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})][1+[\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})] \exp(\Delta_{ml}-\Delta_{mk})]} - \\
&\quad \frac{[1+\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})] \exp(\hat{U}_{km})}{[\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})][1+[\exp(\hat{U}_{km})+\exp(\hat{U}_{lm})] \exp(-\Delta_{mk})]} .
\end{aligned} \tag{11}$$

Note that this solution contains the probability of a transition from state m to state k given that $\Delta_{ml} > 0$ as a special case. The probability is obtained by evaluating (11) at $\Delta_{ml} = 0$. It equals $1 - F(\Delta_{mk}|U_{km} < 0)$, where $F(\Delta_{mk}|U_{km} < 0)$ is the marginal conditional c.d.f. given by (8). Intuitively, as the transitions probabilities must sum up to unity over all alternatives, and considering that the probability of a transition from state m to state l is zero, the solution is the counter probability of the event that the agent stays in the initial state. Likewise, if $\Delta_{ml} > 0$ and $\Delta_{mk} > 0$, evaluating (10) respectively (11) at $\Delta_{ml} = \Delta_{mk} = 0$ yields zero, as required.

The probability of a transition from state m to state l can be obtained analogously. Before proceeding to the higher dimensional case, it is useful to re-state the two individual probabilities of leaving the original state m valid in the trinomial framework in more general terms. Define $\Delta = (\Delta_1, \Delta_2)$ as a vector containing the set of $\Delta_{mj}^* = \min(\Delta_{mj}, 0) \forall j \neq m$, in descending order. Using this notation, we can write the probability that the agent makes the transition from the state m to state k , for any $k \neq m$, as follows:

$$\begin{aligned}
P(k|m) &= P(U_{kj} > \Delta_{jk} \forall j \neq k | U_{jm} < 0 \forall j \neq m) \\
&= \int_{\Delta_{mk}}^{\Delta_{mk}-\Delta_1} \int_{-\infty}^{U_{km}+\delta_1-\Delta_{mk}} f(U_{km}, U_1 | U_{jm} < 0 \forall j \neq m) dU_1 dU_{km} \\
&\quad + I(\Delta_{mk} < \Delta_1) \int_{\Delta_{mk}-\Delta_1}^{\Delta_{mk}-\Delta_2} \int_{-\infty}^0 f(U_{km}, U_1 | U_{jm} < 0 \forall j \neq m) dU_1 dU_{km}
\end{aligned} \tag{12}$$

where $I(\cdot)$ is an indicator function taking the value of unity if the condition in the brackets is satisfied and zero otherwise. Furthermore, δ_1 refers to the first element in a vector δ identical to the vector Δ but omitting the element based on Δ_{mk} . In the trinomial case, δ is a scalar and hence $\delta_1 = \delta$. Finally, U_1 refers to the difference in utilities variable U_{jm} associated with the state j that is covered by δ_1 . To check that the expression is equivalent to the probability statements (10) and (11), evaluate (12) for the two states $j = k, l$ making an assumption on the ordering of Δ_{ml} and Δ_{mk} .

The analytical transition probabilities when there is a larger number of possible states, are obtained from a generalization of (12). For a framework with J exclusive states, the joint conditional distribution of differences in utility $U_{jm} \forall j \neq m$ is a $J - 1$ -variate logistic distribution. The set of $\Delta_{mj}^* = \min(\Delta_{mj}, 0) \forall j \neq m$ in descending order is represented by the vector $\Delta = (\Delta_1, \dots, \Delta_{J-1})$, whereas the set of $\Delta_{mj}^* = \min(\Delta_{mj}, 0) \forall j \neq m, k$ in descending order is denoted by the vector $\delta = (\delta_1, \dots, \delta_{J-2})$. The vector (U_1, \dots, U_{J-2}) contains the sequence of utility differential variables corresponding to the sequence of states in δ . The ordering enables us to represent the bounds of the relevant $J - 1$ post shock utility differences in a systematic manner. Using $f(U_{jm})$ as a shortcut for the conditional distribution $f(U_{km}, U_1, \dots, U_{J-2} | U_{jm} < 0 \forall j \neq m)$, the individual probability of a move from state m to a particular state $k \neq m$ can be written as

$$\begin{aligned}
P(k|m) &= P(U_{kj} > \Delta_{jk} \forall j \neq k | U_{jm} < 0 \forall j \neq m) = \\
&\int_{\Delta_{mk}}^{\Delta_{mk} - \Delta_1} \int_{-\infty}^{U_{km} + \delta_1 - \Delta_{mk}} \int_{-\infty}^{U_{km} + \delta_2 - \Delta_{mk}} \dots \int_{-\infty}^{U_{km} + \delta_{J-2} - \Delta_{mk}} f(U_{jm}) dU_{J-2} \dots dU_1 dU_{km} \\
&+ I(\Delta_{mk} < \Delta_1) \int_{\Delta_{mk} - \Delta_1}^{\Delta_{mk} - \Delta_2} \int_{-\infty}^0 \int_{-\infty}^{U_{km} + \delta_2 - \Delta_{mk}} \dots \int_{-\infty}^{U_{km} + \delta_{J-2} - \Delta_{mk}} f(U_{jm}) dU_{J-2} \dots dU_1 dU_{km} \\
&+ I(\Delta_{mk} < \Delta_2) \int_{\Delta_{mk} - \Delta_2}^{\Delta_{mk} - \Delta_3} \int_{-\infty}^0 \int_{-\infty}^0 \dots \int_{-\infty}^{U_{km} + \delta_{J-2} - \Delta_{mk}} f(U_{jm}) dU_{J-2} \dots dU_1 dU_{km} \\
&+ \dots \\
&+ I(\Delta_{mk} < \Delta_{J-2}) \int_{\Delta_{mk} - \Delta_{J-2}}^0 \int_{-\infty}^0 \dots \int_{-\infty}^0 f(U_{jm}) dU_{J-2} \dots dU_1 dU_{km}
\end{aligned} \tag{13}$$

The integrals in (13) can be solved to yield:

$$\begin{aligned}
P(k|m) &= \left(\sum_j \exp(\hat{U}_{jm}) \right) \times \sum_{r=1}^{J-1} I(\Delta_{mk} < \Delta_{r-1}) \frac{\exp(\hat{U}_{km})}{\exp(\hat{U}_{km}) + \eta_r} \\
&\times \left(\frac{1}{1 + \exp(\Delta_r - \Delta_{mk})(\exp(\hat{U}_{km}) + \eta_r) + \theta_r} - \frac{1}{1 + \exp(\Delta_{r-1} - \Delta_{mk})(\exp(\hat{U}_{km}) + \eta_r) + \theta_r} \right) \tag{14}
\end{aligned}$$

with $\Delta_0 = 0$, $\eta_r = \sum_{s=r}^{J-2} \exp(\hat{U}_s - \delta_s + \Delta_{mk})$, and $\theta_r = \sum_{s=1}^{r-1} \exp(\hat{U}_s)$, where $(\hat{U}_1, \dots, \hat{U}_{J-2})$ represents a vector of observable utility differences corresponding to the vector of unobservable utility differences (U_1, \dots, U_{J-2}) .

Equation (14) provides an analytical prediction of the transition probability from the initial state m to a given alternative state k , for an individual member of the sample. To evaluate the vector of transition probabilities for the agent, one has to work through all possible states $j \neq m$, re-defining the vector δ accordingly. As the transition probabilities must add up to unity over all possible states including m , the probability that the agent

remains in the initial state after the shock can be calculated as a remainder. Alternatively, one might evaluate this probability by using the $J - 1$ -variate analogue of the c.d.f. (6). This is also useful for a validity check of the transitions frequencies evaluated on the basis of (14). Finally, to predict the transitions matrix for the sample, one has to sum up the transitions vectors evaluated at the individual level, over all sample units.

3 Analytical versus Calibration Results

In order to gauge the predictive performance of previous calibration-based results found in the behavioral microsimulation literature, we compare transitions frequencies evaluated by calibration of the unobserved utility components to those based on the analytical value of transitions probabilities in the conditional logit framework. We begin with some stylized examples of predicted transition probabilities at the individual level, which give us an element of control sufficient to highlight conditions under which calibrated results might differ more or less from analytical results. Recognizing that the observations based on such an approach may not directly relate to the sorts of problems normally encountered by analysts, a second part studies the reliability of calibrated transitions frequencies in the context of a typical empirical application.

3.1 Computational Experiments

For the computational analysis, we look at several examples of transitions probabilities obtained at the individual level. Both the calibration and the analytical procedure guarantee that transitions probabilities are evaluated starting from a perfect replication of the initial state. Therefore it seems plausible that calibrated probabilities converge to the analytical ones when the number of replications of the random terms in the utility model increases. However, this does not mean that the asymptotic value of the transitions probabilities is consistent with the data generating process. At an intuitive level, it appears to be likely that the predictions of marginal effects are unreliable, if the empirical model mapping observable characteristics into predictions of state-specific utility levels is mis-specified. But in order to provide an answer to the question how well calibrated results perform relative to our analytical solution, it is irrelevant whether the estimated model parameters are correct or not. For this reason, we can ignore the data generating process and refrain from performing a Monte-Carlo experiment with a larger sample.

In a first experiment, we simulate transitions in a design that provides the agent with a choice between three distinct states, $j = 1, 2, 3$. Assume that the agent initially occupies state 1. We examine two cases: a design where initial predicted utilities are given by a vector $\hat{U} = (0, 0, 0)$, and a design where this vector is given by $\hat{U} = (3.64, 0, 0)$. The first case represents a *balanced* design in which each state, ignoring the random component of the model, is predicted to be chosen with probability one third. Note that this misleading prediction is the starting point for the aggregate probability method of evaluating state transitions. The second case represents an *imbalanced* design in which, based on the mean utility equation, the observed choice with probability 0.95 is the utility maximizing one.

We examine transitions following an exogenous shock that changes the predictions of state-specific utilities according to the vector $\Delta_j = (0, 2, 1)$. In the balanced design, the vector of analytical transitions probabilities, evaluated on the basis of the results contained in the previous section, is $(0.27, 0.63, 0.10)$. An equivalent way to express this result is in terms of a transitions vector: $(-0.73, 0.63, 0.10)$. This means that within 100 identical agents, following the shock 73 would leave the initial state, 63 of whom would move into state 2, whereas 10 would move into state 3. Note that the transitions vector obtained by respecting the probabilistic nature of state transitions at the individual level, is substantially different from that obtained with the aggregate utility method $(-0.24, 0.33, -0.09)$, or that obtained with the maximum utility approach $(-1, 1, 0)$.² In the unbalanced design, the calculated vector of analytical transitions frequencies is $(-0.17, 0.13, 0.03)$.

For the calibration design, we take independent random draws from the extreme value distribution such that $\varepsilon_1 = \max_j(\varepsilon_j)$, which guarantees that prior to the shock, state 1 is the utility maximizing choice. Following the shock, we predict the new utility maximizing choice, on the basis of $n=(10, 20, 50, 100, 1000)$ repeats of the resampling procedure, and evaluate transitions probabilities taking the mean over the individual repeats. To generate confidence intervals for the calibrated transitions vectors, we replicate all steps 1000 times, for each resampling frequency n .

The upper panel of Table 1 displays the results for this experimental design. We get similar outcomes for the balanced and for the unbalanced design. Although the calibrated transitions probabilities are on average broadly correct, judged by the 95 percent confidence

² The first result is obtained by re-evaluating choice probabilities inserting the updated utility vector $\hat{U}_j + \Delta_j$ into (2). The second result is straightforward considering that state 2 is characterized by the largest utility gain following the shock.

band the obtained predictions only converge slowly to the analytical value. If the number of replications is smaller than 50, as it is generally the case in applications of the calibration method found in the literature, the predicted transitions frequencies appear to be quite unreliable.

The comparison with the analytical benchmark also confirms the suspicion that the calibration approach underpredicts small transition frequencies. In the balanced design, with less than 20 replications, there is a significant chance that the calibrated results indicate no transitions into state 3, although this least likely move still has a 0.10 probability of realization. In the unbalanced design, where the smallest transitions probability is only 3 percent, qualitatively misleading results are more frequent. It takes more than 50 repetitions of the calibration to exclude the zero predictions. Likewise, for comparatively large transitions frequencies, like $P(1 \rightarrow 1)$ in the unbalanced design, there is a significant chance that application of the calibration method with few replications obtains a corner solution at unity.

The lower panel of Table 1 shows the outcome of a second design, which uses the same parameters for states 1 to 3 as in the previous experiment. The difference is that we add a fourth state which is an exact copy of state 3, i.e. characterized by $\hat{U}_4 = 0$ and $\Delta_4 = 1$. This implies that the probabilities of a transition from state 1 to state 3 and from state 1 to state 4 are the same. The calibration method clearly fails to replicate this qualitative result. For example, in the balanced design with 10 replications, it might occur that $P(1 \rightarrow 3)$ is zero, whereas $P(1 \rightarrow 4)$ is as large as 0.30, whereas the analytical prediction is 0.09. Even the confidence bands of the predictions obtained by calibration are not always the same.

The observed patterns are robust for different experimental designs. Nevertheless this does not necessarily mean that variability of the predicted transitions probabilities at the individual level would challenge the results of previous empirical applications using calibration. In principle, aggregation over similar agents at the sample level might compensate for a small number of repeats at the individual level. To gauge the practical relevance of this effect, we check the performance of calibrated transitions frequencies against the analytical benchmark, in a typical empirical application.

3.2 An Empirical Application

We choose to follow Kaltenborn et al. (2003) in simulating the labor supply response to a general income allowance for payroll contributions to the social insurance system in Germany. For simplicity, our discrete choice model of labor supply assumes that agents decide on hours worked taking the labor supply decision of other household members as given. They can select themselves into one of six regimes: employment of $h_j = (10, 20, 30, 40, 50)$ hours per week, and non-participation ($h_j = 0$). Actual working hours observed in the data are rounded to fit the elements in the restricted choice set. We consider a translog specification of the direct utilities enjoyed by agent i in state j :

$$U_{ij} = x'_{ij}Ax_{ij} + b'x_{ij} + \varepsilon_{ij} \quad (15)$$

where $x = (y_{ij}, h_j)'$, A is a symmetric 2×2 matrix of parameters and b is a 1×2 vector of parameters. ε_{ij} represents the Type I Extreme Value distributed random disturbance term required to obtain the conditional logit model. Finally, y_{ij} stands for the (potential) net household income of agent i working hours h_j . To generate state-specific household net incomes, we calculate individual gross earnings assuming state-invariant gross wage rates, and derive the corresponding net income using a simplified but detailed microsimulation model of the German tax and transfer system. For agents whose gross wage rate is not observed, we base the calculation on a (selectivity-corrected) predicted wage rate.

The translog specification (15) implies that the interaction between hours worked and net household income is included in the estimation, as well as the elements of x squared. To introduce observed heterogeneity among agents, we specify parameters of the direct utility function as dependent on individual characteristics. Labor supply behavior is estimated on separate samples of 1729 women and 1675 men drawn from the 2000 wave of the German Socio-Economic Panel. The selected samples cover agents of age 25-55 living in West Germany who are not retired, self-employed, a civil servant, in education or in military (national) service, or on parental leave. Working hours are regular working hours including regular paid overtime. To ensure that the estimated model parameters are consistent with the monotonicity conditions implied by economic theory, we maximize the likelihood function of the conditional logit model (3) under inclusion of a penalty function, as suggested by Euwals and van Soest (1999).

Table 2 shows the estimated parameters for the female and male samples. A full discussion of the results is beyond the scope of this paper. We only note that the behavioral parameters obtained from our empirical specification are broadly in line with earlier work. For example, women who raise children or give care value leisure significantly more than a reference group who does not. The same is true for disabled men. Irrespective of gender, the parameters indicate a stigma effect of welfare recipients—agents appreciate a given amount of leisure more when they do not depend on welfare transfers.

In demonstrating the reliability of transitions frequencies evaluated on the basis of calibration, we simulate a hypothetical reform to the German social insurance system which replaces the existing rules for minor employment, exempting jobs that pay earnings below a certain threshold value (€325 per month, in 2002) from social insurance contributions, with a monthly allowance of €250 applying to all earnings. This proposal has in fact been put forward by the Federation of German Trade Unions. The behavioral microsimulation of the shock is generated by revising our detailed tax and transfer model, making the adjustments necessary to reproduce the legal state following the reform. By feeding the simulated post-reform net incomes into the estimated utility model, we can compute the state- and individual-specific changes in utility levels effected by the shock.

This information is sufficient to derive analytical transitions frequencies by aggregating individual transitions probabilities according to (14) over the sample. In applying the calibration method, we draw vectors of unobserved utility components ε_{ij} that maximize pre reform utility at the observed working hours, and recalculate the new utility maximizing choice given the simulated income change, for each agent in the sample. Tables 3 and 4 compare the analytical and the calibrated matrices of simulated transitions for the female and male samples. The calibrated transitions matrix is obtained taking the mean over 20 repetitions.

It is apparent in Table 3 that for women the calibrated results are sufficient to correctly indicate the incentive effects of the reform—while the majority of simulated transitions are positive and mainly relate to agents not currently participating, there are also reductions in labor supply. Some agents currently working few hours reduce their working hours or even withdraw from the labor market, since the reform raises marginal taxes on earnings in minor employment. Furthermore, there is a pure income effect of the contribution rate allowance reducing labor supply of agents currently working higher hours.

Overall, the negative incentive effects of the reform do not compensate its positive incentive effects. Taking the average over the entire sample, working hours increase over the pre reform value of 20.59. However, although the calibration results indicate the correct direction of the incentive effects, the strength of off-diagonal transitions is generally underpredicted. As a result, post reform average working hours based on calibration (21.51) are about 0.17 percent smaller than average working hours based on the analytical solution (21.55).

The calibrated simulated transitions matrix for men is somewhat less reliable. Predicted average working hours obtained by calibration (33.96) underpredict the analytical value (34.05) by 0.29 percent. The intuition why the calibration method performs worse in this case is that men generally either work full-time or do not participate at all. Hence, several of the states permitted in our model specification initially are sparsely populated, so that individual transitions probabilities following the shock tend to be very small. As illustrated in section 3.1, in this situation there is high chance that the calibration method predicts zero transitions probabilities when they are in fact slightly positive, unless the number of repetitions in the re-sampling procedure is large. Nevertheless it is still possible to gauge the mixed incentive effects of the reform from Table 4.

4 Conclusions

In this paper, we have derived an explicit form for transitions probabilities following an exogenous shock to the deterministic component of a conditional logit framework with any number of dimensions. The proposed analytical solution significantly improves upon transitions estimators based on calibration of the unobserved model component, first of all because it helps to avoid computationally demanding re-sampling procedures. Moreover, computational results for stylized experimental designs confirm that calibrated transitions frequencies might perform substantially worse compared to analytical predictions provided that the number of repeats in an application cannot be made sufficiently large, and especially underpredict small transitions probabilities.

Nevertheless, comparison of simulated analytical and calibrated transitions frequencies in a typical empirical application focusing on the prediction of labour market transitions in response to economic policy reform suggests, that previous calibration-based behavioral microsimulations found in the literature probably indicate the strength and the direction

of the incentive effects, by and large, accurately. At the sample level, aggregation of calibrated transitions probabilities over comparable sample units might smooth out imprecise transitions probabilities at the individual level.

At the conclusion, it is important to point out that the suggested analytical prediction method does not immediately recover ‘true’ transitions probabilities. The reason is that the proposed solution ignores the probabilistic nature of the parameter estimates that yield the underlying predictions for the deterministic utility component in the model. An asymptotically exact solution can be obtained by averaging the calculated analytical transitions probabilities over repeated draws from the distribution of the estimated model parameters.

References

- BARGAIN, O. (2003): “Tax Reform Analysis Using Flexible Models of Labor Supply,” mimeo DELTA, Paris.
- CREEDY, J., AND A. DUNCAN (2002): “Behavioural Microsimulation with Labour Supply Responses,” *Journal of Economic Surveys*, 16, 1–39.
- DUNCAN, A., AND C. GILES (1996): “Labour Supply Incentives and Recent Family Credit Reforms,” *The Economic Journal*, 106, 142–155.
- DUNCAN, A., AND M. WEEKS (1998): “Simulating Transitions Using Discrete Choice Models,” *Proceedings of the American Statistical Association*, 106, 151–156.
- EUWALS, R., AND A. VAN SOEST (1999): “Desired and Actual Labour Supply of Unmarried Men and Women in the Netherlands,” *Labour Economics*, 6, 95–118.
- GERFIN, M., AND R. LEU (2003): “The Impact of In-Work Benefits on Poverty and Household Labour Supply: A Simulation Study for Switzerland,” IZA Discussion Paper, No. 762, Bonn.
- GOURIEROUX, C., A. MONFORT, E. RENAULT, AND A. TROGNON (1987): “Generalized Residuals,” *Journal of Econometrics*, 34, 5–32.
- KALTENBORN, B., S. KOCH, U. KRESS, U. WALWEI, AND G. ZIKA (2003): *Arbeitsmarkteffekte eines Freibetrags bei den Sozialabgaben*. Hampp Publishers, Munich.
- KOTZ, S., N. BALAKRISHNAN, AND N. JOHNSON (2000): *Continuous Multivariate Distributions*, vol. 1: Models and Applications. John Wiley & Sons, New York.
- MADDALA, G. (1986): *Limited-Dependent and Qualitative Variables in Econometrics*, Econometric Society Monographs No.3. Cambridge University Press, Cambridge.
- MCFADDEN, D. (1974): “Conditional Logit Analysis of Qualitative Choice Behavior,” in *Frontiers in Economics*, ed. by P. Zarembka, pp. 105–142. Academic Press, New York.
- VAN SOEST, A. (1995): “Structural Models of Family Labor Supply,” *Journal of Human Resources*, 30, 63–83.

Table 1: Convergence of Calibrated Transitions Probabilities

Transition	Number of Replications					
	10	20	50	100	1000	∞
balanced design						
$P(1 \rightarrow 1)$	[0.10, 0.50]	[0.10, 0.45]	[0.16, 0.38]	[0.20, 0.35]	[0.25, 0.29]	0.27
$P(1 \rightarrow 2)$	[0.40, 0.90]	[0.45, 0.80]	[0.52, 0.74]	[0.55, 0.71]	[0.61, 0.65]	0.63
$P(1 \rightarrow 3)$	[0.00, 0.30]	[0.00, 0.20]	[0.04, 0.18]	[0.05, 0.15]	[0.09, 0.11]	0.10
unbalanced design						
$P(1 \rightarrow 1)$	[0.60, 1.00]	[0.70, 0.95]	[0.74, 0.92]	[0.77, 0.89]	[0.81, 0.85]	0.83
$P(1 \rightarrow 2)$	[0.00, 0.30]	[0.00, 0.25]	[0.06, 0.22]	[0.08, 0.19]	[0.12, 0.15]	0.13
$P(1 \rightarrow 3)$	[0.00, 0.10]	[0.00, 0.10]	[0.00, 0.08]	[0.01, 0.07]	[0.03, 0.04]	0.03
balanced design						
$P(1 \rightarrow 1)$	[0.10, 0.50]	[0.15, 0.45]	[0.20, 0.40]	[0.22, 0.36]	[0.27, 0.31]	0.29
$P(1 \rightarrow 2)$	[0.30, 0.80]	[0.35, 0.70]	[0.42, 0.65]	[0.45, 0.62]	[0.51, 0.56]	0.54
$P(1 \rightarrow 3)$	[0.00, 0.20]	[0.00, 0.20]	[0.02, 0.16]	[0.04, 0.14]	[0.08, 0.10]	0.09
$P(1 \rightarrow 4)$	[0.00, 0.30]	[0.00, 0.20]	[0.04, 0.16]	[0.05, 0.14]	[0.08, 0.10]	0.09
unbalanced design						
$P(1 \rightarrow 1)$	[0.60, 1.00]	[0.65, 0.95]	[0.71, 0.90]	[0.74, 0.87]	[0.78, 0.83]	0.81
$P(1 \rightarrow 2)$	[0.00, 0.30]	[0.00, 0.25]	[0.06, 0.22]	[0.07, 0.18]	[0.11, 0.14]	0.13
$P(1 \rightarrow 3)$	[0.00, 0.10]	[0.00, 0.10]	[0.00, 0.08]	[0.01, 0.07]	[0.03, 0.04]	0.03
$P(1 \rightarrow 4)$	[0.00, 0.10]	[0.00, 0.10]	[0.00, 0.08]	[0.01, 0.07]	[0.02, 0.04]	0.03

Table 2: Parameter Estimates for Labor Supply Model

	Women		Men	
	Parameter Est.	Std. Dev.	Parameter Est.	Std. Dev.
log income	14.33**	3.90	18.38**	4.26
(log income) ²	0.13	0.18	-0.32*	0.19
log income × log leisure	-2.06**	0.43	-1.51**	0.44
log leisure	33.17**	5.56	48.01**	5.87
× potential experience	-0.10	0.06	-0.04	0.06
× (potential experience) ²	0.00	0.00	0.00	0.00
× unskilled	-0.63	0.42	-2.21**	0.39
× apprenticeship	-1.34**	0.34	-2.47**	0.31
× vocational training	-0.79*	0.41	-1.91**	0.39
× married	-0.66*	0.32	-1.22**	0.29
× foreigner	1.01**	0.36	0.68*	0.29
× disabled	0.86	0.53	2.72**	0.53
× giving care	2.59**	0.97	-0.04	0.72
× No. of children younger than 6	4.23**	0.41	0.41*	0.23
× No. of children younger than 16	1.29**	0.16	-0.33**	0.13
× welfare recipience	-0.42**	0.04	-0.14**	0.04
(log leisure) ²	-1.30**	0.38	-3.78**	0.41
No. of observations	1729		1675	
log likelihood	-2841.6		-2919.4	

Notes: Conditional logit for individual's choice of working hours. In couples, the choice of working hours made by the partner is treated as given. There are six possible choices of working hours per week: (0, 10, 20, 30, 40, 50). Maximum leisure time per week is set to 80 hours. Potential experience is calculated as age minus number of years at school minus 6. The omitted education category is university training. **, * are used to indicate that a parameter value is significant at the one respectively five percent level.

Table 3: Analytical and Calibrated Transitions Matrix – Women

Hours	0	10	20	30	40	50	Before
0	29.95 (30.57)	1.47 (1.08)	1.16 (1.20)	1.09 (0.95)	0.75 (0.60)	0.30 (0.30)	34.70
10	0.22 (0.14)	5.55 (5.66)	0.14 (0.12)	0.09 (0.11)	0.05 (0.05)	0.02 (0.01)	6.07
20	0.02 (0.01)	0.13 (0.05)	13.43 (13.52)	0.06 (0.05)	0.05 (0.06)	0.02 (0.02)	13.71
30	0.00 (0.00)	0.07 (0.03)	0.16 (0.10)	10.38 (10.47)	0.02 (0.03)	0.01 (0.00)	10.64
40	0.01 (0.01)	0.11 (0.13)	0.44 (0.27)	0.62 (0.30)	32.69 (33.18)	0.01 (0.00)	33.89
50	0.00 (0.00)	0.00 (0.00)	0.03 (0.03)	0.04 (0.02)	0.02 (0.00)	0.89 (0.93)	0.98
After	30.20 (30.73)	7.33 (6.95)	15.35 (15.23)	12.28 (11.91)	33.59 (33.92)	1.25 (1.26)	100.00

Table 4: Analytical and Calibrated Transitions Matrix – Men

Hours	0	10	20	30	40	50	Before
0	13.95 (14.52)	0.20 (0.11)	0.35 (0.35)	0.89 (0.73)	1.05 (0.87)	0.46 (0.31)	16.90
10	0.00 (0.00)	0.53 (0.57)	0.01 (0.00)	0.02 (0.02)	0.02 (0.01)	0.01 (0.00)	0.60
20	0.00 (0.00)	0.00 (0.00)	0.64 (0.65)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)	0.66
30	0.00 (0.00)	0.01 (0.00)	0.01 (0.00)	1.41 (1.43)	0.00 (0.00)	0.00 (0.00)	1.43
40	0.01 (0.00)	0.16 (0.15)	0.59 (0.36)	1.14 (0.55)	72.80 (73.73)	0.11 (0.01)	74.81
50	0.00 (0.00)	0.01 (0.00)	0.04 (0.04)	0.09 (0.08)	0.09 (0.05)	5.38 (5.44)	5.61
After	13.95 (14.53)	0.91 (0.84)	1.65 (1.40)	3.55 (2.81)	73.97 (74.67)	5.96 (5.76)	100.00

Notes: See Table 3.