Chapter 5

Is the Choice of Mode of Payment Rational?

5.1 Introduction

When someone pays for an over-the-counter purchase, he has the choice between different modes of payment, e.g. payment in cash, by means of a cheque, a credit card, etc. Insight in the choice of mode of payment is important, because of its implications for the demand of token money, the development of new monetary products, etc. If we can model the choice of mode of payment in a utility maximizing framework, this model could be used to perform simulations. Hence, it is of interest to determine whether a discrete choice model of the mode of payment is compatible with utility maximization.

In the previous chapter we have examined when discrete choice models are compatible with stochastic utility maximization. In the present chapter, we apply this theory to models of the mode of payment. We try to answer the question whether the estimated choice probabilities are compatible with stochastic utility maximization. Even though a complete analysis of the choice of mode of payment is interesting in its own right, it is beyond the scope of this chapter. A detailed analysis of the choice of mode of payment based on the dataset that we use has been published by Mot, Cramer and Gulik (1989).

The data that we use are taken from the Intomart Bestedingen Index (which will be translated as the Intomart Expenditure Index). In that survey various kinds of expenditures are recorded, together with information on the household. We distinguish between three possible modes of payment: payment by a giro cheque, payment by a bank
cheque\textsuperscript{1} and cash payment. The alternatives are labelled in this order 1, 2 and 3. Since our data were collected in 1987, hardly any transactions with modern forms of payment like a credit card or a PIN-card are recorded, so these observations are deleted from the analysis.

The remainder of this chapter is organized as follows. In section 5.2 we estimate two models for the choice of payment mode. We examine whether the estimated choice probabilities are compatible with stochastic utility maximization in section 5.3. We summarize and conclude in section 5.4.

5.2 Model Specification and Estimation Results

Mode of payment is recorded in the Dutch Intomarkt Expenditure Index. This is a panel of slightly more than 1000 households. These households have on average slightly more than 2 members older than 12 years. Each household records its expenditures daily. We have used the 1987 wave, in which 2161 over-the-counter purchases were recorded. We restrict ourselves to households who both have a bank and a giro account, and to one transaction per household. This leaves us with 225 observations. We could have used all observations but that would complicate the statistical model because the number of observed purchases per household varies between 1 and 48, and the independence between observations which we make is hard to maintain if multiple transactions per household are recorded. We have selected randomly one transaction if more than one transaction is observed for a particular household. For each transaction the mode of payment is recorded, as well as the amount paid (\textit{AMT}), size of the household (\textit{SIZE}), age of the head of the household (\textit{AGE}), sex of the person making the transaction (\textit{SEX}) (this dummy is 1 for females and 0 for males) and household income (\textit{INC}).

The distribution of the observations over the mode of payment is given in table 5.1. It is expected that cheques are used to pay larger amounts, which is confirmed by the data. The average amount per transaction paid in cash is Dfl. 38.34 and the average amount paid with giro cheques is Dfl. 104.06 and with bank cheques Dfl. 88.93.

We model the choice of mode of payment using two different functional forms for the choice probabilities. The first model we estimate

\textsuperscript{1} This category consists of two subcategories, viz. Eurocheques and green bank cheques. All types of cheques are guaranteed.
Model Specification and Estimation Results

<table>
<thead>
<tr>
<th>Payment by cheque</th>
<th>39</th>
</tr>
</thead>
<tbody>
<tr>
<td>giro cheques</td>
<td>20</td>
</tr>
<tr>
<td>bank cheques</td>
<td>19</td>
</tr>
<tr>
<td>Payment in cash</td>
<td>186</td>
</tr>
</tbody>
</table>

Table 5.1: Distribution of modes of payment

is the multinomial logit model (MNL) (see Maddala (1983)), where the choice probabilities are given by

$$P_i(v_t) = \frac{\exp(-v_{i1})}{\exp(-v_{i1}) + \exp(-v_{i2}) + \exp(-v_{i3})}, \quad i = 1, 2, 3. \quad (5.1)$$

We use the same notation as in chapter 4. It is easily established that the MNL choice probabilities satisfy the Daly–Zachary–Williams conditions ((C.1)-(C.5) on page 70) for all $v \in \mathbb{R}^3$, and hence, that they are globally compatible with stochastic utility maximization.

The distribution of $(\varepsilon_1, \varepsilon_2, \varepsilon_3)'$ which generates the choice probabilities (5.1) is

$$F(\varepsilon) = \exp(-\exp(-\varepsilon_1) - \exp(-\varepsilon_2) - \exp(-\varepsilon_3)).$$

Note that $\varepsilon_1$, $\varepsilon_2$ and $\varepsilon_3$ are independently distributed.

We choose alternative 1 as the reference alternative, i.e. we set $v_{11} = 0$ for all individuals, and specify $v_{12}$ and $v_{13}$ as

$$v_{12} = \beta_2 x_t \quad (5.2)$$

$$v_{13} = \beta_3 x_t \quad (5.3)$$

The utility of a particular alternative depends on the log of the amount of the transaction ($\ln AMT$), the size of the household ($SIZE$), the age of the head of the household ($AGE$), the log of household income ($\ln INC/10000$) and sex of the person making the transaction ($SEX$).

The estimation results for this specification are given in table 5.2. One of the most important determinants of the choice between either a bank cheque or a post cheque and cash is the amount of the transaction: the higher the amount the more likely it is that either one of the cheques is chosen. From the second column it is clear that none of the included variables affects the choice between either type of cheque significantly.

The multinomial logit specification (5.1) is rather restrictive: it suffers from the ‘independence of irrelevant alternatives’ property: the
The odds of paying with a bank cheque over paying cash is independent of the presence and characteristics of the third alternative, viz. paying with a post cheque. Another drawback of the MNL model is that rationality of the choice is imposed a priori. A specification of the choice probabilities which recognizes the fact that the first two alternatives are rather similar is the nested multinomial logit (NMNL) specification. In the nested multinomial logit specification, this assumption can be tested. The functional form for the choice probabilities of the nested multinomial logit model are:

\[ P_i(v_t) = \frac{\exp(-v_{it}/\theta) \left( \exp(-v_{i1}/\theta) + \exp(-v_{i2}/\theta) \right)^{(\theta-1)}}{\left( \exp(-v_{i1}/\theta) + \exp(-v_{i2}/\theta) \right)^{\theta} + \exp(-v_{i3})} \]

\[ i = 1, 2 \]

\[ P_3(v_t) = \frac{\exp(-v_{i3})}{\left( \exp(-v_{i1}/\theta) + \exp(-v_{i2}/\theta) \right)^{\theta} + \exp(-v_{i3})} \]

Again, we choose the first alternative as the reference alternative. This model is estimated using the two-stage estimator for nested logit models (see, e.g., Maddala (1983)); we encountered convergence problems when we tried to estimate the model with the full maximum-likelihood estimator. The estimation results are presented in table 5.3. The column labelled ‘Bank cheques’ contains the estimate for \( \beta_2/\theta \), not for \( \beta_2 \) itself. This scaling makes the results of table 5.3 better comparable to those of table 5.2. As in the multinomial logit specification we see that the choice between the modes of payment is mainly governed by the amount of the

Table 5.2: Estimation results, multinomial logit model (standard errors in parentheses)
Compatibilty with Utility Maximization

Table 5.3: Estimation results, nested multinomial logit model (standard errors in parentheses)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Giro cheques</th>
<th>Bank cheques</th>
<th>Cash</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>-</td>
<td>-</td>
<td>-1.74 (2.60)</td>
</tr>
<tr>
<td>$\ln AMT$</td>
<td>-</td>
<td>-</td>
<td>0.61 (0.46)</td>
</tr>
<tr>
<td>$SIZE$</td>
<td>-</td>
<td>-</td>
<td>-0.33 (0.31)</td>
</tr>
<tr>
<td>$AGE/10$</td>
<td>-</td>
<td>-</td>
<td>0.35 (0.38)</td>
</tr>
<tr>
<td>$SEX$</td>
<td>-</td>
<td>-</td>
<td>0.79 (0.77)</td>
</tr>
<tr>
<td>$\ln INC/10000$</td>
<td>-</td>
<td>-</td>
<td>-1.40 (0.76)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-2.54 (1.50)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

transaction. The larger the amount of the transaction, the more likely it becomes that it is paid with a cheque. Moreover, households with a higher income have a lower probability of paying in cash.

From the choice probabilities in equation (5.4) it is seen that the NMNL logit model reduces to the MNL logit model if $\theta = 1$. The $p$-value for the null hypothesis $H_0: \theta = 1$ against the two-sided alternative $H_1: \theta \neq 1$ is 0.005. Hence, the multinomial logit specification is rejected against the nested multinomial logit model.

The estimation results clearly indicate that the choice between either type of cheque is difficult to model. Other specifications did not lead to an improvement of fit or significance of variables. In fact, one can wonder whether both types of cheques should not be grouped together into a single alternative. This problem is examined in Cramer and Ridder (1991). In a different specification, they firmly reject the null hypothesis that both types of cheques are perceived as equivalent alternatives.

We now turn to the question whether the estimated NMNL choice probabilities are compatible with stochastic utility maximization.

5.3 Compatibility with Utility Maximization

In this section, we examine whether the estimated NMNL model of the previous section is compatible with stochastic utility maximization for three choices of the set of non-random utility components $\mathcal{V}$.

First, we examine whether the model is globally compatible with utility maximization. If the choice probabilities are compatible,
the density function of the random utility components \((\eta_1, \eta_2) = (\varepsilon_2 - \varepsilon_1, \varepsilon_3 - \varepsilon_1)\) which generates the choice probabilities (5.4) is given by equation (4.34) on page 89. This density is nonnegative for all values \(\eta \in \mathbb{R}^2\) if and only if \(0 < \theta \leq 1\). Our estimate for \(\theta\) is \(-2.54\), so we conclude that the NMNL model is not compatible with utility maximization for all values \(v \in \mathbb{R}^3\).

One possible reason for rejection of rationality of choice is that we require the Daly–Zachary–Williams conditions to hold for all values of \(v\), even if they are far away from the observed values. A less ambitious question is whether the model is compatible with utility maximization on a closed interval \(V_2 = [a, b]\). According to theorem 4.4 on page 79, the same Daly–Zachary–Williams conditions applied on an interval are necessary and sufficient conditions for compatibility with stochastic utility maximization on that interval. If all observed non-random utility components lie in that interval, we can conclude that the model is compatible with utility maximization on that interval.

In figure 5.1, we plot the estimated non-random utility components, and the curve \(h_1(\eta_1, \eta_2) = 0\). This curve is represented by the solid line, and given by

\[
\eta_2 = -\ln \frac{\theta + 1}{\theta - 1} - \theta \ln (1 + \exp(-\eta_1/\theta)),
\]

with \(\theta < -1\) or \(\theta > 1\). For the line in figure 5.1 we have taken \(\theta = -2.54\). \(h_1(\eta_1, \eta_2)\) is negative beneath the curve. A substantial number of observations are situated in the area where \(h_1(\eta_1, \eta_2)\) is negative and hence the model is not compatible with utility maximization on an interval.

Finally we check whether the estimated choice probabilities are compatible with utility maximization on the finite set \(V_3 = \{v_1, \ldots, v_T\}\) of observed utility components. Theorem 4.5 on page 84 gives a necessary and sufficient condition for this choice of \(V_3\), which is, however, rather difficult to implement. It is easier to check the necessary condition of corollary 4.4 on page 85. For all pairs \((t, t')\) there must hold:

\[
v_t - v_{it'} \geq v_t - v_{it} \Rightarrow P_i(v_t) \geq P_i(v_{it}) \quad i = 1, 2, 3. \tag{5.6}
\]

We have checked\(^3\) condition (5.6) for the estimated NMNL model. The condition was violated 2777 times, which is in 5.51% of

---

2. The function \(h_1(\eta_1, \eta_2)\) is negative for all values \(\eta \in \mathbb{R}^2\) if \(-1 \leq \theta < 0\).

3. Details are given in Appendix 5.A.
all comparisons. This means that the nested multinomial logit model is not locally compatible with stochastic utility maximization on the set $Y_3$, because the choice probabilities are not non-decreasing in $v$.

This result raises the question whether one obtains a model compatible with utility maximization if pairs of observations which violate the necessary condition (5.6) are excluded. We have re-estimated the nested multinomial logit model twice. First, we have deleted observations for which the necessary condition was violated more than 25 times, which leaves us with 201 observations. The estimate for $\theta$ in this case is $-1.46$. This estimate is increased further if we select only those observations which violate condition (5.6) 15 times or less. In the latter case, we find $\hat{\theta} = -0.83$. A natural step would be to select only those observations which do not violate condition (5.6) at all. This, however, leaves us with too few observations to estimate the model.

The choice probability of the first alternative is drawn in figure 5.2 from two different angles. It is clear that $P_1(v)$ is not increasing in $v_3$ and $v_2$. To be specific, $P_1(v)$ decreases in $v_2$ for some values of $v_3$.

In figure 5.3 the same surface is shown for $\theta = -0.83$. $P_1(v)$ decreases even more clearly in $v_2$. For comparison, $P_1(v)$ is drawn in
Figure 5.2: $P_1(v)$ with $\theta = -2.54$

figure 5.4 for $\theta = 0.50$. The NMNL model is compatible with stochastic utility maximization for this value of $\theta$. 
Figure 5.3: $P_1(v)$ with $\theta = -0.83$

5.4 Conclusion

In this chapter we have used the theory developed in chapter 4 to examine whether the choice of mode of payment is rational. Rationality can be imposed by specifying a specific functional form for the choice probabilities that satisfies the Daly–Zachary–Williams conditions, as the MNL.
This model, however, is rejected against a more flexible functional form for the choice probabilities, viz. the NMNL model. The NMNL model is not guaranteed to be compatible with stochastic utility maximization; indeed we reject rationality for three choices of sets of non-random utility components.

Figure 5.4: $P_i(v)$ with $\theta = 0.50$
Of course, the nested multinomial logit model might be rejected against an even more flexible functional form for the choice probabilities. We are, however, unaware of convenient parametric alternatives to the nested logit specification. Moreover, even if the necessary condition (5.6) for compatibility with utility maximization is satisfied, one has to check the necessary and sufficient condition of theorem 4.5. This difficult problem remains a topic for future research.

5.A Check of the Necessary Condition

We have written a computer program\(^4\) which checks condition (5.6) for all different pairs of points. In the case with three alternatives, every two observations imply exactly two restrictions of the form (5.6) on the choice probabilities. This is illustrated in figure 5.A-1. Consider for example one observation \(t\). The other observation \(t'\) lies in either one of the areas \(A_1, A_2, \ldots, A_6\) and each case implies two restrictions on the choice probabilities, see table 5.A-4.

| \(t' \in A_1\) | \(P_1(v_{i_t}) \geq P_1(v_{i_{t'}})\) | \(P_3(v_{i_t}) \leq P_3(v_{i_{t'}})\) |
| \(t' \in A_2\) | \(P_1(v_{i_t}) \geq P_1(v_{i_{t'}})\) | \(P_2(v_{i_t}) \leq P_2(v_{i_{t'}})\) |
| \(t' \in A_3\) | \(P_3(v_{i_t}) \geq P_3(v_{i_{t'}})\) | \(P_2(v_{i_t}) \leq P_2(v_{i_{t'}})\) |
| \(t' \in A_4\) | \(P_3(v_{i_t}) \geq P_3(v_{i_{t'}})\) | \(P_1(v_{i_t}) \leq P_1(v_{i_{t'}})\) |
| \(t' \in A_5\) | \(P_1(v_{i_t}) \geq P_1(v_{i_{t'}})\) | \(P_3(v_{i_t}) \leq P_3(v_{i_{t'}})\) |
| \(t' \in A_6\) | \(P_2(v_{i_t}) \geq P_2(v_{i_{t'}})\) | \(P_3(v_{i_t}) \leq P_3(v_{i_{t'}})\) |

Table 5.A-4: Restrictions on choice probabilities

---

4. Both the C++-source code and the executable of this program are available on request.
Figure 5.A-1: Two restrictions per observation