Panel effects in consumer research: Statistical models for underreporting
Kaper, E.

Citation for published version (APA):

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Estimation of underreporting using two indicators for reporting

3.1 Introduction

As shown in the previous chapter, panel effects can be so strong that one cannot use such data without correction. One way to correct for these effects is by defining consumption as a latent variable, of which the reported consumption is one indicator. This approach is possible if two indicators are available (for more on latent variable models, see e.g. Bollen, 1989). The objective of this chapter is to estimate underreporting using a model in which there are two indicators for consumption. In Section 3.2 the research design will be described, including a description of the model, and its identification. In Section 3.3, the estimation procedure will be described, with the parameter estimates of the model. Section 3.4 will contain the conclusions.

3.2 Research design

3.2.1 The model

A Structural Equations Model has been defined to describe expenditure and response behaviour, since Structural Equations Models allow the presence of latent variables. For each household i, period t and group g, true consumption \( C_{igt} \) is influenced by some household characteristics \( x_{ig} \), like family size, family income, age, and education level. Besides these variables, an intercept \( \alpha_g \) is present to represent a trend in consumption, and a disturbance-term \( \zeta_{igt} \) for random fluctuation. Consumption is measured using two
Chapter 3

indicators: each period for all products bought the amount and price are reported. The sum of the prices of these individual products is reported expenditure ($r_{ig}$). It is subject to grave response burden. Since this response burden leads to demotivation, which causes underreporting (Kaper and Saris, 1996; Kaper and Saris 1997), it has a slope ($\lambda_{ig}$) not necessarily equal to 1 and an intercept ($\tau_{ig}$) which is not restricted to being 0. Apart from these biases, random measurement errors are present for this indicator: $\varepsilon_{ig}$. Besides this reported expenditure, a second indicator is present, called the direct estimate ($d_{ig}$), which is measured retrospectively each quarter: at the end of the quarter the respondent is asked how much was spent on average per week during the last quarter. It is expected to be measured without systematic underreporting, since it is not subject to a large response burden like the reported expenditure, so a slope of 1 and an intercept of 0 are assumed. This assumption also defines a scale and origin for the latent variable true consumption: it is assumed that true consumption is measured with the same metric as this second indicator. There is nonetheless random measurement error, represented by $\delta_{ig}$.

The model can be described with the following equations for each respondent $i = 1,...,I$, period $t = 1,...,T$ and group $g = 1,...,G$:

$$c_{ig} = \alpha_{ig} + \gamma x_{ig} + \zeta_{ig}$$  
(3.1a)

$$r_{ig} = \tau_{ig} + \lambda_{ig} c_{ig} + \varepsilon_{ig}$$  
(3.1b)

$$d_{ig} = C_{ig} + \delta_{ig}$$  
(3.1c)

with $\mathbb{E}[\zeta_{ig}] = 0$, $\mathbb{E}[\varepsilon_{ig}] = 0$, and $\mathbb{E}[\delta_{ig}] = 0$; $\text{Var}[\zeta_{ig}] = \psi_i$ (variance of disturbances); $\text{Cov}[\zeta_{ig}, \delta_{ig}] = \psi_{it}$ (correlated disturbances); $\text{Var}[\delta_{ig}] = \theta_{it}$ (variances of measurement errors in $\delta_{ig}$); $\text{Cov}[\delta_{ug}, \delta_{it}] = \theta_{itu} \forall u \neq t$ (correlated measurement errors: same respondents); $\text{Var}[\varepsilon_{ig}] = \theta_{it}$ (variances of measurement errors in $\varepsilon_{ig}$), $\text{Cov}[\varepsilon_{ug}, \varepsilon_{it}] = 0 \forall u \neq t$ (correlations in measurement errors are accounted for in $\lambda_{ig}$).

In the remainder, equation (3.1a) is referred to as the Consumption model, and (3.1b) and (3.1c) together as the Reporting model.

3.2.2 Identification

Prior to the aspect of estimation, identification of the model has to be considered. In each period, the model is a MIMIC-model, a special case of a Structural Equations Model in which there are multiple indicators and multiple causes for one latent variable (Jöreskog and Goldberger, 1975). Although it is known that MIMIC-models are identified if there are at least two indicators and at least one causal variable (e.g. see Bollen, 1989), identification has
Two indicators

Two indicators to be studied more carefully, since correlated errors in the direct estimate and correlated disturbance terms for consumption are introduced.

Initially, identification of the slopes and variances is verified, assuming the variables are all measured in deviation from the mean. Once identification of these parameters is established, the means are identified as well, under the condition that apart from the scale of the latent variables also the origin of the latent variables is defined.

If first the two MIMIC-models for each period are considered separately, it is found that $\lambda_c$, $\gamma$, $\theta_{\xi}$, $\theta_{\varepsilon}$, and $\psi_t$ are identified. Cross-time covariances among the observed variables establish identification of both $\theta_{\xi u}$ and $\psi_{\varepsilon u}$ for $t \neq u$. A full proof of identification of the model is given in Appendix 3.B.

3.2.3 The data

The data are from the survey and questionnaire as described in Chapter 1. Since underreporting was expected in these data, at the end of the quarter all respondents were asked how many guilders they spent during the quarter, on average per week. In this manner, a second indicator for true expenditure on meat, poultry, and eggs is available. For this purpose (and other purposes), the purchases in meat are subdivided into four groups: meat, meat products, poultry and eggs. For each of these groups, this second estimator is present. Because the second indicator is measured on a quarterly basis, the first indicator, reported consumption, is averaged to get a quarterly measure as well.

Before the final analysis, some evident outliers are removed from the data. These outliers were defined as observations in which the average amount reported in the second quarter was more than five times larger or more than five times smaller than the amount reported in the third quarter (both direct estimate and reported expenditure). Also households with a zero reporting for the direct estimate and reporting expenditure larger than zero are not taken into consideration, since these responses are definitely erroneous. After these individual exclusions of observations, 467 households remain for which all relevant variables are available for all periods.

Since it is hypothesized that response behaviour, and thus underreporting, differs for households that report or consume at different levels, the sample is split into three more or less equal-sized groups based on the average number of reportings per week in the first quarter. (This variable is denoted $p_i$.) Following this procedure, there is a group of 132 households that reported over 0 up to and including 1 product, a group of 161 households that reported over 1 up to and including 2 products and a group of 174 households that
Chapter 3

reported more than 2 products on average per week in the first quarter and that reported both in the second and the third quarters.

### 3.3 Estimation

Estimation of the model is done using LISREL (Jöreskog and Sörbom, 1993), which is a standard program to estimate Structural Equations Models. The method of estimation used is Pseudo Maximum Likelihood, which is the use of Maximum Likelihood when the normality assumption does not necessarily hold. The fit function to be minimised for the multi-sample model with means and intercepts is

\[ F = \sum_i G_N g / N F_g \]

where

- \( F_g = 1/2(s^{(g)} - \sigma^{(g)}) W^{-1}(s^{(g)} - \sigma^{(g)}) + 1/2(m^{(g)} - \mu^{(g)}) \Sigma^{-1}(m^{(g)} - \mu^{(g)}) \)
- \( W(g) = (w^{h,y}) = N(2 - \delta_{ph})(2 - \delta_{py})(\sigma^h \sigma^y + \sigma^h \sigma^y) \) for group \( g \)
- \( s^{(g)} \) is the vector of non-redundant sample (co-)variances for group \( g \)
- \( \sigma^{(g)} \) is the vector of non-redundant population (co-)variances for group \( g \)
- \( m^{(g)} \) is the vector of sample means for group \( g \)
- \( \mu^{(g)} \) is the vector of population means for group \( g \)
- \( \Sigma^{(g)} \) is the population variance-covariance matrix for group \( g \)
- \( \delta \) is Kronecker delta
- \( N \) is the total number of persons in the sample
- \( N_g \) is the total number of persons in group \( g \)

Maximum Likelihood estimators have a number of desirable properties, among which the property that the estimates are consistent, and under normality also efficient. Robustness of Maximum Likelihood estimators has been studied for non-normal data (Satorra, 1993).

The three groups were estimated simultaneously, incorporating the across-group restriction that the consumption model was invariant. Studying the parameter estimates of this model together with the standard errors led to hypothesizing that the loadings of reported expenditure on true consumption are equal through time and that the intercepts for reporting are equal to zero. After this hypothesis was incorporated in the model, it was tested again. The parameter estimates of this restricted model with \( \tau_{ig} = 0, \lambda_{mg} = \lambda_g \) are shown in Table 3.1. The LISREL input for this model is given in Appendix 3.A.
The standard errors show that education level does not significantly contribute to consumption. Furthermore, the disturbances are very large. This is also shown in the rather low squared correlation coefficients in row $R^2_{cx}$ of about 0.13. Apparently only a small part of consumption is explained by the model. The reporting model reveals the underreporting: $1 - \lambda_g$ gives the proportion of underreporting for group $g$. 

$$\chi^2 = 87.7238 \ (P = .021)$$

### Table 3.1: Maximum Likelihood Parameter estimates (standard errors within parentheses)

<table>
<thead>
<tr>
<th>Consumption model</th>
<th>ALL GROUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect Parameter</td>
<td>(All parameters of the consumption model restricted to being equal across groups)</td>
</tr>
<tr>
<td>Family size $\gamma_{11}$</td>
<td>3.2528 (0.5256)</td>
</tr>
<tr>
<td>Income $\gamma_{12}$</td>
<td>1.1540 (0.4290)</td>
</tr>
<tr>
<td>Age $\gamma_{13}$</td>
<td>0.1347 (0.0392)</td>
</tr>
<tr>
<td>Education $\gamma_{14}$</td>
<td>0.0701 (0.4896)</td>
</tr>
<tr>
<td>Disturbance $q^2_1$ $\psi_2$</td>
<td>115.8529 (19.2436)</td>
</tr>
<tr>
<td>Disturbance $q^3_2$ $\psi_3$</td>
<td>114.9435 (18.5008)</td>
</tr>
<tr>
<td>Correlated $\psi_4$</td>
<td>103.2474 (15.1454)</td>
</tr>
<tr>
<td>Intercept $q^2_1$ $\alpha_2$</td>
<td>7.1196 (3.0377)</td>
</tr>
<tr>
<td>Intercept $q^3_2$ $\alpha_3$</td>
<td>7.0148 (3.0335)</td>
</tr>
<tr>
<td>$R^2_{cq}$</td>
<td>0.1335</td>
</tr>
<tr>
<td>$R^2_{cq}$</td>
<td>0.1344</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reporting model</th>
<th>Group $g = 1$</th>
<th>Group $g = 2$</th>
<th>Group $g = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loading Parameter</td>
<td>$0 &lt; p_1 \leq 1; l_1 = 132$</td>
<td>$1 &lt; p_1 \leq 2; l_2 = 161$</td>
<td>$2 &lt; p_1; l_3 = 174$</td>
</tr>
<tr>
<td>Reported exp. $q^2$ $\lambda_2$</td>
<td>0.3172 (0.0211)</td>
<td>0.4553 (0.0301)</td>
<td>0.6345 (0.0400)</td>
</tr>
<tr>
<td>Reported exp. $q^3$ $\lambda_3$</td>
<td>0.3172 (0.0211)</td>
<td>0.4553 (0.0301)</td>
<td>0.6345 (0.0400)</td>
</tr>
<tr>
<td>Error reported $q^2$ $\delta_{22}$</td>
<td>8.2692 (1.9844)</td>
<td>16.6202 (3.9441)</td>
<td>15.8795 (5.9091)</td>
</tr>
<tr>
<td>Error reported $q^3$ $\delta_{33}$</td>
<td>7.2271 (1.8240)</td>
<td>16.2044 (3.9621)</td>
<td>19.9553 (5.7495)</td>
</tr>
<tr>
<td>Error direct $q^2$ $\delta_{22}$</td>
<td>943.8797 (120.9986)</td>
<td>1550.7089 (177.3229)</td>
<td>1656.3027 (180.5340)</td>
</tr>
<tr>
<td>Error direct $q^3$ $\delta_{33}$</td>
<td>812.0203 (104.5654)</td>
<td>3393.8856 (383.3019)</td>
<td>119.6740 (131.6203)</td>
</tr>
<tr>
<td>Correlated $\delta_{32}$</td>
<td>803.9784 (107.5261)</td>
<td>1698.4088 (229.2007)</td>
<td>1181.8373 (141.9609)</td>
</tr>
<tr>
<td>Error reported $q^2$ $\delta_{22}$</td>
<td>0.6268</td>
<td>0.6313</td>
<td>0.7722</td>
</tr>
<tr>
<td>$R^2_{ac}$ $q^3$</td>
<td>0.6563</td>
<td>0.6356</td>
<td>0.7282</td>
</tr>
<tr>
<td>$R^2_{ac}$ $q^2$</td>
<td>0.1276</td>
<td>0.0813</td>
<td>0.0747</td>
</tr>
<tr>
<td>$R^2_{ac}$ $q^3$</td>
<td>0.1445</td>
<td>0.0386</td>
<td>0.0997</td>
</tr>
<tr>
<td>Contribution to</td>
<td>30.6183 (34.90%)</td>
<td>31.5728 (35.99%)</td>
<td>25.5327 (29.11%)</td>
</tr>
</tbody>
</table>

Effects labelled $q^2$ concern the second quarter, while labels $q^3$ denote the third quarter.
Almost 60% is underreported by the first group, more than 50% by the second group, and more than 30% by the third group. Hence, the group that reported least in the first quarter (the quarter on which the group definition is based) reports least in successive quarters. It appears that the first quarter was already subject to underreporting. By comparison of the measurement errors in reported expenditure and in the direct estimates and the squared multiple correlations ($R^2_{rc}$ and $R^2_{dC}$ in the table), it is seen that the direct estimate is indeed a less precise indicator than reported expenditure, though less biased. Finally, the $\chi^2$ statistic of 87.72, with 63 degrees of freedom, shows with a probability of 2.1% a moderate fit of the model.

The estimates for $\lambda_{gr}$, together with the standard errors, do not suggest that underreporting is equal across groups. This is also indicated by the corresponding $\chi^2$ statistic of 262.45 with 65 degrees of freedom. The difference-test with the less restrictive model gives a $\chi^2$ statistic of 174.73 with 2 degrees of freedom, which is highly significant, leading to a rejection of the hypothesis of equal underreporting across groups.

The parameter estimates can be used to predict average true consumption, by using the average of the prediction: average of $C_i^* = \alpha_i^* + \gamma^\prime K$. This gives us the predictions related to the indicators as shown in Table 3.2. The results show that although reported consumption decreases through time, predicted true consumption increases.

Table 3.2: Predictions for true consumption

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quarter 2 Quarter 3</td>
<td>Quarter 2 Quarter 3</td>
<td>Quarter 2 Quarter 3</td>
</tr>
<tr>
<td>Reported</td>
<td>$r$ 7.11 7.33</td>
<td>11.67 11.45</td>
<td>17.53 17.03</td>
</tr>
<tr>
<td>Direct</td>
<td>$d$ 17.98 18.68</td>
<td>28.42 31.92</td>
<td>30.28 30.82</td>
</tr>
<tr>
<td>True consumption</td>
<td>$C$ 15.64 16.42</td>
<td>17.59 18.09</td>
<td>22.10 22.63</td>
</tr>
</tbody>
</table>

3.4 Conclusions

Reconsidering the results, a quite simple model that incorporates latent variables is very well able to correct for underreporting in a panel survey. Essential for the estimations is the availability of a second indicator for consumption. Although the indicator present in the data is a very rough one, with very large error-variances, it can be assumed not to be subject to underreporting, enabling estimation of underreporting itself in the other indicator, reported consumption. Indeed underreporting is present in the data, varying from 30 to 60%, and this underreporting is not constant across groups: the group with a low reporting
in the first quarter underreports more than the group with high reporting in the first quarter. Apparently reportings in the first quarter already were subject to underreporting.

Furthermore, while reported consumption suggests that consumption decreases from the second to the third quarter, after correction for measurement errors, consumption actually increases.

No explanation is given for the underreporting itself. For that, a more complex model would be necessary. This is something that will be studied in the next chapter.

3.A Appendix: LISREL-input

Panel-MIMIC model for two periods, red meat, group 1: 0 < RLVP1 <= 1
For a description of the model see section 3.2
Repl = r_tig, DIRT = d_tig, FAMSIZE...EDUC = x_tig, CONSt = C_tig
Data NG=3 NI=8 NO=132 MA=CM

KM
1.0000
.2801 1.0000
-.2229 -.0962 1.0000
-.0876 .1946 -.1621 1.0000
.3589 .1492 .0306 -.0440 1.0000
.3118 .2177 .0161 -.0839 .6182 1.0000
.2425 .0662 -.1066 -.0731 .1099 .2328 1.0000
.2182 .0786 -.1041 -.1195 .1090 .2713 .9089 1.0000

SD
1.2475 1.3497 15.0472 1.1817 4.7586 4.7167 31.0862 29.3582

ME
1.9697 2.9470 49.7197 2.5227 7.1120 7.3347 17.9848 18.6752

LA
FAMSIZE INCOME AGE EDUC REP2 REP3 DIR2 DIR3
SE
REP2 DIR2 REP3 DIR3 FAMSIZE INCOME AGE EDUC
Model NE=2 NY=4 NX=4 FI FS=SY, FR TE=SY, FI LY=FU, FI GA=FU, FR AL=FI TY=FI MA FR
LE
CONS2 CONS3
FR AL(1) AL(2)
FI TY(1) TY(3)
EQ GA(1,1) GA(2,1)
EQ GA(1,2) GA(2,2)
EQ GA(1,3) GA(2,3)
EQ GA(1,4) GA(2,4)
FR LY(1,1)
EQ LY(1,1) LY(3,2)
VA 1 LY(2,1) LY(4,2)
FR TE(2,1) TE(2,2) TE(3,3) TE(4,4)
FR TE(2,4)
MA LY
.3697 .0000
1.0000 .0000
.0000 .3783
.0000 1.0000
MA GA
2.3543 .9783 .1266 .2446
2.3543 .9783 .1266 .2446
MA FS
90.7120
79.1552 .82.6515
MA TE
7.8508
.0000 917.4685
Chapter 3

MIMIC MODEL FOR TWO PERIODS, RED MEAT, GROUP 2: 1 < RVLP1 <= 2
DATA NO=161

RM

\[
\begin{pmatrix}
1.0000 \\
0.2521 1.0000 \\
-0.3245 -0.0978 1.0000 \\
-0.0091 0.2571 -0.0931 1.0000 \\
-0.2631 0.3354 -0.0139 0.0742 1.0000 \\
-0.1293 0.2156 0.0871 -0.0101 0.5971 1.0000 \\
-0.0256 0.0729 0.0360 -0.1527 0.3150 0.2552 1.0000 \\
-0.0632 0.0205 -0.0608 -0.1386 0.2067 0.1425 0.7469 1.0000
\end{pmatrix}
\]

SD

\[
\begin{pmatrix}
1.2724 1.4361 15.4398 1.1940 6.8997 6.6898 41.7004 59.0476 \\
2.2609 3.4596 50.6211 2.6957 11.6720 11.4522 28.4161 31.9206
\end{pmatrix}
\]

LA

FAMILY SIZE INCOME AGE EDUC REP2 REP3 DIR2 DIR3

SE

REP2 DIR2 REP3 DIR3 FAMILY SIZE INCOME AGE EDUC

MODEL NE=2 NT=4 NX=4 FI PS=IN TE=SP LY=SP GA=IN AL=IN TY=SP KA=FR

LE

CONS2 CONS3

MA LI

-0.5572 .0000

1.0000 .0000

0.0000 .4887

0.0000 1.0000

Ma GA

2.9543 .9783 .1266 .2448

2.9543 .9783 .1266 .2448

Ma PS

90.7120 79.1552 82.6515

Ma TE

15.3422 .0000 1528.2568

0.0000 .0000 24.5548

0.0000 1499.3020 .0000 1699.7828

MA AL

-5.77

10.01

Ma TY

0.00

0.00

0.00

0.00

OU

PANEL-MIMIC MODEL FOR TWO PERIODS, RED MEAT, GROUP 3: 2 < RVLP1
DATA NO=174

RM

\[
\begin{pmatrix}
1.0000 \\
0.1921 1.0000 \\
-0.3492 0.0391 1.0000 \\
-0.0637 0.2846 0.0219 1.0000 \\
0.3911 0.0775 0.0229 0.0283 1.0000 \\
0.1496 0.1992 0.1357 0.0786 0.6955 1.0000 \\
0.0608 0.0615 0.0478 0.0262 0.2091 0.1193 1.0000 \\
0.0825 0.1240 0.0346 0.0757 0.2199 0.2086 0.8358 1.0000
\end{pmatrix}
\]

SD

\[
\begin{pmatrix}
1.3245 1.4078 15.4398 1.1940 6.8997 6.6898 41.7004 59.0476
\end{pmatrix}
\]
Two indicators

3.B Appendix: Identification of the model

To establish identification, first the consumption model with latent variables is inserted to the measurement model. For each household i and period t (since we do not have any cross-group parameters, we treat all groups together, omitting the subscript g), the equations are:

\[ d_{it} = \alpha_t + \gamma'x_i + \zeta_{it} + \delta_i \]  
(3.B.1)

\[ r_{it} = \tau_t + \lambda_t(\alpha + \gamma'x_i + \zeta_{it}) + \epsilon_{it} \]  
(3.B.2)

The sample moments are compared with the moments implied by the model. First, the sample covariances (\( \Sigma_x \) and \( \sigma^t \)'s) are studied, resulting in the following equalities that need to be solved for the parameters of the model. The first equation can be solved for \( \Phi: \Phi = \Sigma_x \). In each consecutive equation only one new parameter appears, so the system of equation can be solved. Note that the following model assumptions are applied: \( \text{Cov}[x, \xi_t] = 0, \text{Cov}[x, \delta_i] = 0, \text{Cov}[x, \epsilon_i] = 0, \text{Cov}[\delta_i, \zeta_t] = 0, \text{Cov}[\delta_i, \xi_t] = 0, \text{Cov}[\delta_i, \epsilon_t] = 0, \text{Cov}[\delta_i, \delta_u] = 0 \) for all \( t \) and \( u \neq t \).
$$\Sigma_x = \text{Var}[x] = \Phi$$

$$\sigma_{x,t} \equiv \text{Cov}[x_t, d_t] = \text{Cov}[x_t, \alpha_t + \gamma x + \zeta_t + \delta_t] = \gamma' \Phi$$

$$\sigma_{x,r} \equiv \text{Cov}[x_r, r_r] = \text{Cov}[x_r, \lambda_r(\alpha_r + \gamma x + \zeta_r) + e_r] = \lambda_r' \gamma' \Phi$$

$$\sigma_{d,r} \equiv \text{Cov}[d_r, r_r] = \text{Cov}[\alpha_r + \gamma x + \zeta_r + \delta_r, \tau_r + \lambda_r(\alpha_r + \gamma x + \zeta_r) + e_r] = \lambda_r(\gamma' \Phi \gamma + \psi_r)$$

$$\sigma_{d,t}^2 \equiv \text{Var}[d_t] = \text{Var}[\alpha_t + \gamma x + \zeta_t + \delta_t] = \gamma' \Phi \gamma + \psi_t + \theta_{\delta_t}$$

$$\sigma_{r,t}^2 \equiv \text{Var}[r_t] = \text{Var}[\tau_t + \lambda_t(\alpha_t + \gamma x + \zeta_t) + e_t] = \lambda_t^2 \gamma' \Phi \gamma + \lambda_t^2 \psi_t + \theta_{\delta_t}$$

$$\sigma_{d_r, d_u} \equiv \text{Cov}[d_r, d_u] = \text{Cov}[\alpha_r + \gamma x + \zeta_r + \delta_r, \alpha_u + \gamma x + \zeta_u + \delta_u] = \gamma' \Phi \gamma + \psi_{d_u} + \theta_{\delta_{d_u}}$$

The same procedure can be used for establishing identification of the means ($\mu$'s), with the following assumptions from the model definition: $E[\delta] = 0, E[\epsilon] = 0, E[\zeta] = 0.$

$$\mu_x = E[x] = \kappa$$

$$\mu_{d_t} = E[d_t] = E[\alpha_t + \gamma x + \zeta_t + \delta_t] = \alpha_t + \gamma' \kappa$$

$$\mu_{r_t} = E[r_t] = E[\tau_t + \lambda_t(\alpha_t + \gamma x + \zeta_t) + e_t] = \tau_t + \lambda_t(\alpha_t + \gamma' \kappa)$$

which proves that the complete model is identified.