Conditional Logit with one Binary Covariate: Link between the Static and Dynamic Cases

Romain AEBERHARDT et Laurent DAVEZIES
We are grateful to Magali Beffy, Bo Honoré as well as participants at the 15th International Conference on Panel Data in Bonn (July 2009) for helpful discussions. We also thank Pauline Givord for thoroughly reading the proofs and providing us with suggestions to improve the paper. The views expressed herein are attributable to the authors and do not reflect those of INSEE.
Conditional Logit with one Binary Covariate:  
Link between the Static and Dynamic Cases

Abstract
Disentangling state dependence from unobserved heterogeneity is a common issue in economics. It arises for instance when studying transitions between different states on the labor market. When the outcome variable is binary, one of the usual strategies consists in using a conditional logit model with an appropriate conditioning suitable for a dynamic framework.

Although static conditional logit procedures are widely available, these procedures cannot be used directly in a dynamic framework. Indeed, it is inappropriate to use them with a lag dependent variable in the list of regressors. Moreover, reprogramming this kind of procedures in a dynamic framework can prove quite cumbersome because the likelihood can have a very high number of terms when the number of periods increases.

Here, we consider the case of a conditional logit model with one binary regressor which can be either exogenous or the lagged dependent variable itself. We provide closed forms for the conditional likelihoods in both cases and show the link between them.

These results show that in order to evaluate a conditional logit model with one lag of state dependence and no other covariate, it is possible to simply generate a two variable dataset and use standard procedures originally intended for models without state dependence. Moreover, the closed forms help reduce the computational burden even in the static case in which preimplemented procedures usually exist.

Keywords: conditional logit, state dependence, binary model, incidental parameter

Logit conditionnel avec une covariable binaire :  
lien entre les cas statiques et dynamiques

Résumé
Identifier les effets relevant de la dépendance d’état (c’est-à-dire de la trajectoire passée) tout en contrôlant des effets de variables inobservées est un problème classique en économétrie. C’est par exemple le cas lorsqu’on étudie les transitions d’individus entre différents états sur le marché du travail. Quand la variable d’intérêt est binaire, une des stratégies usuelles consiste à utiliser un modèle logit conditionnel.

Bien que les procédures informatiques permettant d’estimer des modèles de type logit conditionnel dans un cadre statique soient largement disponibles dans les logiciels statistiques, ces procédures ne doivent pas être utilisées directement dans le cas d’un modèle dynamique. Les utiliser en incluant naïvement la variable dépendante retardée dans la liste des régisseurs conduit à des estimations non convergentes. De plus, programmer la vraisemblance de ce type de modèles dans un cadre dynamique peut se révéler délicat car le nombre de termes dans la vraisemblance conditionnelle augmente exponentiellement avec le nombre de périodes d’observation.

Dans ce document, nous considérons un modèle logit conditionnel avec comme explicatives soit un régisseur binaire exogène, soit la variable d’intérêt retardée. Dans les deux cas, nous explicitons des formes fermées de la vraisemblance conditionnelle et nous mettons en évidence la relation entre ces deux expressions. En particulier, pour estimer un logit conditionnel avec une dépendance d’état d’une période et sans autre covariable, nous démontrons qu’il est possible de générer simplement une base de données constituée de deux variables et d’appliquer les procédures standards développées pour le cadre statique. De plus au prix de la programmation d’une maximisation, les formes fermées permettent de tirer parti du caractère binaire de la variable explicative pour réduire drastiquement les temps de calculs, y compris dans le cas statique pour lequel les procédures disponibles sont parfois lentes.

Mots-clés : Logit conditionnel, dépendance d’état, modèle binaire, paramètres incidents

Classification JEL : C23, C25, J62
1 Introduction

Disentangling state dependence from unobserved heterogeneity is a common issue in economics. While the former corresponds to the lasting effect of the past trajectory on the state or behavior of an individual, the latter corresponds to the effect of all other relevant individual characteristics which can not be controlled by the researcher. This question arises for instance when studying transitions between different states on the labor market. An application of such a method concerning subsidized training and youth employment can be found for instance in Magnac (2000). Indeed, being able to distinguish between these two aspects can be extremely important in terms of policy targeting. For instance, one can imagine that if state dependence is very high, voluntary policies aiming at bringing unemployed people back to jobs could have long term positive effects, whereas if there is no state dependence, efforts should be made to improve individual characteristics for instance through more education or on-the-job training.

When the outcome variable is binary, one of the usual strategies consists in using a conditional logit model with an appropriate conditioning suitable for a dynamic framework as developed in Chamberlain (1985). The conditional logit is a common procedure when dealing with unobserved heterogeneity in panel data econometrics when the outcome is binary and the covariates are exogenous, in particular when there is no state dependence. Indeed, classical maximum likelihood estimators are not consistent when the number of estimated parameters increases at the same rate as the number of observations. The conditional logit approach rules out this incidental parameters issue which arises with panel data. Seminal papers on this topic include Rasch (1960) and Andersen (1973), while Magnac (2004) generalizes this approach to other distribution families.

In such a framework with exogenous covariates one needs only to condition on the number of occurrences of each state. When there is state dependence, for instance when the outcome at date $t$ depends on the outcome at date $t-1$, the exogeneity of the covariates is lost. However it is still possible to use a similar approach to make the individual fixed effect disappear but with another conditioning statistic. In that case, there must be at least four observations by individual for the model to be identified.

Although static conditional logit procedures are widely available, these procedures cannot
be used directly in a dynamic framework. Indeed, it is inappropriate to use them with a lag dependent variable in the list of regressors. Moreover, reprogramming this kind of procedures in a dynamic framework can prove quite cumbersome because the likelihood can have a very high number of terms when the number of periods increases.

Here, we consider the case of a conditional logit model with one binary regressor which can be either exogenous or the lagged dependent variable itself. We provide closed forms for the conditional likelihoods in both cases and show the link between them. These results show that in order to evaluate a conditional logit model with one lag of state dependence and no other covariate, it is possible to simply generate a two-variable dataset and use standard procedures originally devised for a framework without state dependence. Moreover, the closed forms help reduce the computational burden even in the static case in which preimplemented procedures usually exist.

Note however that the dynamic procedure studied here is only suitable when there are no other varying covariates. This may seem very restrictive but in fact, first, the model takes into account, in a very flexible way, all the individual characteristics which are constant over time, and second, in the case of varying but stationary covariates, it is always possible to use a stronger conditioning. This extension consists in restricting the sample to individuals whose covariates do not change between two consecutive periods (see for instance Honoré and Kyriazidou, 2000).

The second part will give an idea of the method on a simple case, the third part compares the conditional likelihoods in the static and dynamic frameworks, provides closed forms for the corresponding likelihoods and shows how to generate an appropriate dataset which can be used for the estimation of a conditional logit model with state dependence using standard procedures intended for static conditional logit models. The last part concludes. Examples of programs which generate such an appropriate dataset as well as full proofs of the results are provided in appendix.

2 Introductory example

The minimum number of periods needed to identify the state dependence parameter in a dynamic conditional logit model is four (Chamberlain, 1985). In this section we consider such a
model with four time periods, one lag of state dependence and no other covariates than the constant individual unobserved heterogeneity. Then we show that the corresponding likelihood can be interpreted as that of a conditional logit with two time periods and one exogenous binary covariate.

For this example as well as for the rest of the article, we will use the following notations: \( y_{it}, z_{it}, x_{it} \) will all be binary variables; \( i \) and \( t \) represent respectively individuals and time periods; \( y \) will be reserved to the dynamic model, while \( z \) and \( x \) will be respectively the dependent and explanatory variables for the static model. In both types of models \( \alpha_i \) will represent the individual unobserved heterogeneity which is constant over time.

2.1 Conditional Logit example in a dynamic framework with four time periods

We consider the following model:

\[
y_{it} = \mathbb{1}_{\{\alpha_i + y_{it-1} + \varepsilon_{it} > 0\}}
\]

where the \( \varepsilon_{it} \) have a logistic probability distribution

This model can also be characterized by:

\[
P(y_{it}|y_{it-1}, \alpha_i) = \frac{e^{y_{it}(\alpha_i + y_{it-1})}}{1 + e^{\alpha_i + y_{it-1}}}
\]

(1)

In order to make the \( \alpha_i \) vanish, one usually conditions on \( y_{i1}, y_{i4} \) and \( y_{i2} + y_{i3} = 1 \) (corresponding to individuals switching state between periods 2 and 3). With this particular conditioning the remaining sets of events to consider correspond to the following:

\[
A_i = \{y_{i1}, y_{i2} = 0, y_{i3} = 1, y_{i4}\} \quad \text{and} \quad B_i = \{y_{i1}, y_{i2} = 1, y_{i3} = 0, y_{i4}\}
\]
The probabilities of these two events are:

\[ P(A_{i} | \alpha_i) = P(y_{i4} | y_{i3} = 1) P(y_{i3} = 1 | y_{i2} = 0, \alpha_i) P(y_{i2} = 0 | y_{i1}, \alpha_i) P(y_{i1} | \alpha_i) \]
\[ = \frac{e^{\alpha_i + y_{i4}(\delta + \alpha_i)}}{(1 + e^{\alpha_i + y_{i1}\delta})(1 + e^{\alpha_i})} P(y_{i1} | \alpha_i) \]
\[ P(B_{i} | \alpha_i) = \frac{e^{y_{i1}\delta + \alpha_i}}{(1 + e^{\alpha_i + y_{i1}\delta})(1 + e^{\alpha_i})} P(y_{i1} | \alpha_i) \]

hence,

\[ P(A_{i} | A_{i} \cup B_{i}, \alpha_i) = \frac{e^{y_{i4}\delta}}{e^{y_{i1}\delta} + e^{y_{i4}\delta}} \]
\[ P(B_{i} | A_{i} \cup B_{i}, \alpha_i) = \frac{e^{y_{i1}\delta}}{e^{y_{i1}\delta} + e^{y_{i4}\delta}} \]

therefore, the conditional likelihood does not depend on \( \alpha_i \) and can be written as:

\[ P(y_{i2}, y_{i3} | y_{i1}, y_{i4}, y_{i2} + y_{i3} = 1) = \frac{e^{\delta y_{i2} y_{i1}} e^{\delta y_{i4} y_{i3}}}{e^{\delta y_{i1}} + e^{\delta y_{i4}}} \] (2)

Note that \( P(y_{i2}, y_{i3} | y_{i1}, y_{i4}, y_{i2} + y_{i3} = 0) \) and \( P(y_{i2}, y_{i3} | y_{i1}, y_{i4}, y_{i2} + y_{i3} = 2) \) do not depend on \( \delta \) and therefore the estimation strategy consisting in maximizing the conditional likelihood relies only on the individuals who change state between periods 2 and 3.

### 2.2 Conditional Logit example in a static framework with two independent time periods

We now turn to a conditional logit model with two time periods, no state dependence and one binary exogenous variable:

\[ z_{it} = 1_{\{\alpha_i + x_{it}\delta + \varepsilon_{it} > 0\}} \text{ where the } \varepsilon_{it} \text{ have a logistic probability distribution} \]

This model can also be characterized by:

\[ P(z_{it} = 1 | x_{it}, \alpha_i) = \frac{e^{\alpha_i + x_{it}\delta}}{1 + e^{\alpha_i + x_{it}\delta}} \] (3)
Hence, with the same kind of calculations as before,

\[ P(z_{i1}, z_{i2} | z_{i1} + z_{i2} = 1, x_{i1}, x_{i2}) = \frac{e^{\delta z_{i1} x_{i1}} e^{\delta z_{i2} x_{i2}}}{e^{\delta x_{i1}} + e^{\delta x_{i2}}} \]  (4)

Note that, as in the dynamic case, \( P(z_{i1}, z_{i2} | z_{i1} + z_{i2} = 0) \) and \( P(z_{i1}, z_{i2} | z_{i1} + z_{i2} = 2) \) do not depend on \( \delta \) and therefore the estimation strategy consisting in maximizing the conditional likelihood relies only on the individuals who change state between periods 1 and 2.

Note that this model is identified only on the individuals who change state between the two periods.

2.3 Link between the static and the dynamic frameworks

In this simple case, if you consider a four-period dynamic model \((y_{i1}, y_{i2}, y_{i3}, y_{i4})\) and a two-period static model \(((z_{i1}, x_{i1}), (z_{i2}, x_{i2}))\) such that \( z_{i1} = y_{i2}, x_{i1} = y_{i1}, z_{i2} = y_{i3} \) and \( x_{i2} = y_{i4} \), both conditional likelihoods are equal. That means, that with four time periods, if you feed a standard conditional logit procedure designed for a static framework with a two-period dataset such as in Table 1a, the software will maximize the “right” conditional likelihood of a dynamic framework. Note that reshaping the data as in Table 1b will not give a consistent estimator of the state dependence parameter with standard conditional logit procedures although it could have seemed intuitive to run a conditional logistic regression on the lagged dependent variable when looking at Equations (1) and (3).

In the following section we shall see that the link between the conditional likelihoods of static and dynamic models can be generalized to more than four time periods.

Table 1: Rearrangements with 4 time periods

<table>
<thead>
<tr>
<th></th>
<th>(a) Good one</th>
<th>(b) Bad one</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>z_i</td>
<td>x_i</td>
</tr>
<tr>
<td>1</td>
<td>y_{i2}</td>
<td>y_{i1}</td>
</tr>
<tr>
<td>2</td>
<td>y_{i3}</td>
<td>y_{i4}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3 Generalization for any $T \geq 4$

Here we consider a binary outcome with $T \geq 4$ periods in the dynamic framework: $y = (y_{i1}, \ldots, y_{iT})$ and $T' \geq 2$ in the static framework: $z = (z_{i1}, \ldots, z_{iT'})$. Because in the dynamic framework, the trajectories such that $\sum_{t=2}^{T-1} y_{it} = 0$ or $\sum_{t=2}^{T-1} y_{it} = T - 2$ are uninformative, estimation is made without these observations and in this section we assume without loss of generality but for convenience in notations that such observations are dropped from the data. In the following, individuals such that $\sum_{t=2}^{T-1} y_{it} / \in \{0, T - 2\}$ are called “movers” (in $y$). Similarly, trajectories such that $\sum_{t=1}^{T'} z_{it} = 0$ or $\sum_{t=1}^{T'} z_{it} = T'$ will not be considered, and individuals such that $\sum_{t=1}^{T'} z_{it} / \in \{0, T'\}$ are called “movers” (in $z$).

3.1 Likelihood of a conditional logit with one lag of state dependence but no other varying covariates

Equation (2) can easily be extended to $T \geq 4$:

$$P(y_{i1}, \ldots, y_{iT} | y_{i1}, \sum_{t=2}^{T-1} y_{it}, y_{iT}) = \frac{\exp \left( \sum_{t=2}^{T} y_{it}y_{it-1}\delta \right)}{\sum_{\tilde{y} \in B(T)} \exp \left( \sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}\delta \right)}$$

(5)

where $B(T) = \{(\tilde{y}_{i1}, \ldots, \tilde{y}_{iT}) | \tilde{y}_{i1} = y_{i1}, \tilde{y}_{iT} = y_{iT}, \sum_{t=2}^{T-1} \tilde{y}_{it} = \sum_{t=2}^{T-1} y_{it}\}$ is the set of all possible trajectories with the same first and last states and the same occurrences of each state as in the observed trajectory between $t = 2$ and $t = T - 1$. Intuitively, identification is achieved by analyzing if the visits in state 1 are rather concentrated or evenly distributed along the trajectory. Therefore, the conditional logit compares individual trajectories with the same number of visits in each state. Moreover, if there is state dependence, the states in the first and last periods will not be independent from the distribution of states in the rest of the trajectory, so it seems natural to also condition on the first and last periods.

However, Equation (5) depends on the set $B(T)$ which can become cumbersome to compute as $T$ increases: the number of terms for a trajectory such that $\sum_{t} y_{t} = T/2$ increases as $O(2^{T (1-1/T)^T})$. For instance, for $T = 5$ the number of terms is at most 10, but for $T = 10$ it is
252 and for $T = 20$ it is greater than 180,000. A simpler closed form can therefore be helpful to estimate $\delta$.

**Lemma 3.1** (Closed form of a conditional logit with state dependence).

For $T \geq 4$ and for any mover’s trajectory $y$, the likelihood of a conditional logit with one lag of state dependence but no other varying covariate can be written as:

$$
\mathbb{P}(y_1, \ldots, y_T | y_1, \sum_{t=2}^{T-1} y_{it}, y_{iT}) = \left[ \sum_{j=\max(0,2\sum_{t=1}^{T-1} y_{it} - y_T)}^{\sum_{t=1}^{T-1} y_{it} - 1} \left( \sum_{t=1}^{T-1} y_{it} - 1 \right) \left( \sum_{t=1}^{T-1} y_{it} - j - y_T \right) \exp \left\{ \left( j - \sum_{t=2}^{T-1} y_{it} \right) \delta \right\} \right]^{-1}
$$

The number of terms in the previous expression is bounded by $T/2 - 1$ and can be easily computed even when $T$ is larger than 20 or 30.

Proof : see appendix.

### 3.2 Likelihood of a conditional logit with one covariate but no state dependence

Equation (4) can easily be extended to $T \geq 2$:

$$
\mathbb{P}(z_{i1}, \ldots, z_{iT}, \sum_{t=1}^{T'} z_{it}, x_{i1}, \ldots, x_{iT}) = \frac{\exp \left( \sum_{t=1}^{T'} z_{it} x_{it} \delta \right)}{\sum_{\tilde{z} \in B'(T')} \exp \left( \sum_{t=1}^{T'} \tilde{z}_{it} x_{it} \delta \right)}
$$

(6)

where $B'(T') = \left\{ (\tilde{z}_{i1}, \ldots, \tilde{z}_{iT}) | \sum_{t=1}^{T'} \tilde{z}_{it} = \sum_{t=1}^{T} z_{it} \right\}$ is the set of all trajectories with the same occurrences of each state as in the observed trajectory between $t = 1$ and $t = T'$. Identification is intuitively achieved by the covariance of $x$ and $z$ for a known number of visits in each state.

Once again, the maximization of Likelihood (6) can be a computational burden because the number of elements in the set $B'(T')$ increases exponentially with $T$. But when $X$ is binary, a lot of trajectories $\tilde{z} \in B'(T')$ share the same value for $\sum_{t=1}^{T'} \tilde{z}_{it} x_{it} \in [0, T']$. If we define $B'(T', k) = B'(T') \cap \{(\tilde{z}_{i1}, \ldots, \tilde{z}_{iT}) | \tilde{z}_{it} x_{it} = k\}$, and $n_k = \#B'(T', k)$ the denominator of 6
becomes $\sum_{k=0}^{T'} n_k \exp(k\delta)$, and a simpler expression of the likelihood can be maximized. The following lemma gives the closed form obtained with this method.

**Lemma 3.2** (Closed form of a conditional logit with one binary covariate but without state dependence).

For $T' \geq 2$ and for any mover's trajectory $z$, the likelihood of conditional logit with one binary covariate but no state dependence can be written as:

$$P(z_{i1}, \ldots, z_{iT'} | \sum_{t=1}^{T'} z_{it}, x_{i1}, \ldots, x_{iT'}) =$$

$$\left[ \min_{j=\max(0, \sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T')} \frac{\left( \sum_{t=1}^{T'} x_{it} \right)}{\sum_{t=1}^{T'} z_{it} - j} \exp \left\{ \left( j - \sum_{t=1}^{T'} x_{it} z_{it} \right) \delta \right\} \right]^{-1}$$

Proof: see appendix.

### 3.3 Link between the static and dynamic frameworks

The aim is to appropriately choose $z = (z_{it})_{1 \leq t \leq T'}$ and $x = (x_{it})_{1 \leq t \leq T'}$, such that the conditional likelihood in a static framework corresponds to the likelihood in a dynamic one.

Many statistical softwares provide procedures to estimate conditional logit models without state dependence, but not to estimate such models with state dependence. Despite the fact that 3.1 gives simple expression of the likelihood, econometricians who would like to estimate a conditional logit model with state dependence could be tempted to use standard procedures. The following theorem gives sufficient conditions for the two likelihoods to be the same.

**Theorem 3.3** (Conditions to link the conditional likelihoods of dynamic and static conditional logit models).

The likelihood of a dynamic model such as 3.1 and the likelihood of a static model with only one binary covariate such as 3.2 are the same if and only if

1. $T' = T - 2$ and $(\sum_{t=1}^{T-2} z_{it} = \sum_{t=2}^{T-1} y_{it} \text{ or } \sum_{t=1}^{T-2} z_{it} = T - 2 - \sum_{t=2}^{T-1} y_{it})$
for $y$ such that $0 < \sum_{t=1}^{T-1} y_{it} < T - 2$ and $1 < \sum_{t=1}^{T} y_{it} < T - 1$

- There are $\sum_{t=2}^{T} y_{it} y_{i(t-1)}$ periods such that $(z_{it}, x_{it}) = (1,1)$
- There are $\sum_{t=1}^{T} y_{it} - y_{i1} - y_{iT} - \sum_{t=2}^{T} y_{it} y_{i(t-1)}$ periods such that $(z_{it}, x_{it}) = (1,0)$
- There are $\sum_{t=1}^{T} y_{it} - 1 - \sum_{t=2}^{T} y_{it} y_{i(t-1)}$ periods such that $(z_{it}, x_{it}) = (0,1)$
- There are $T - 1 - 2 \sum_{t=1}^{T} y_{it} + \sum_{t=2}^{T} y_{it} y_{i(t-1)} + y_{i1} + y_{iT}$ periods such that $(z_{it}, x_{it}) = (0,0)$

or

- There are $\sum_{t=2}^{T} y_{it} y_{i(t-1)}$ periods such that $(z_{it}, x_{it}) = (0,0)$
- There are $\sum_{t=1}^{T} y_{it} - y_{i1} - y_{iT} - \sum_{t=2}^{T} y_{it} y_{i(t-1)}$ periods such that $(z_{it}, x_{it}) = (0,1)$
- There are $\sum_{t=1}^{T} y_{it} - 1 - \sum_{t=2}^{T} y_{it} y_{i(t-1)}$ periods such that $(z_{it}, x_{it}) = (1,0)$
- There are $T - 1 - 2 \sum_{t=1}^{T} y_{it} + \sum_{t=2}^{T} y_{it} y_{i(t-1)} + y_{i1} + y_{iT}$ periods such that $(z_{it}, x_{it}) = (1,1)$

• for $y$ such that $(\sum_{t=2}^{T-1} y_{it} = 1$ and $y_{i1} = y_{iT} = 0$) or such that $(\sum_{t=2}^{T-1} y_{it} = T - 3$ and $y_{i1} = y_{iT} = 1$)

$(\forall t \leq T - 2 \ x_{it} = 0$ or $\forall t \leq T - 2 \ x_{it} = 1)$ and $(\sum z \in \{1, T - 3\}$

Proof: see Appendix

Note that the trajectories such that $\sum_{t=2}^{T-1} y_{it} = \sum_{t=1}^{T} y_{it} = 1$ or $\sum_{t=2}^{T-1} y_{it} = \sum_{t=1}^{T} y_{it} - 2 = T - 3$ are the uninformative trajectories because $\sum_{t=2}^{T} \tilde{y}_{it} \tilde{y}_{i(t-1)}$ is constant when $\tilde{y}$ vary in $B(T)$. These special cases are degenerate ones. In these cases, $\#B(T) = T - 2$ and the conditional likelihood is equal to $\frac{1_{y \in B(T)}}{T - 2}$, to ensure equality of the likelihood $z$ must be in the same state at $T - 3$ periods and $x$ must be constant. In the other cases, if $y$ is such that $0 < \sum_{t=2}^{T-1} y_{it} < T - 2$ and $1 < \sum_{t=1}^{T} y_{it} < T - 1$, the likelihood of the dynamic model depends on $\delta$ and this dependence impose more restrictions on $z$ and $x$ to ensure the equality of the two likelihoods.
3.4 Simple algorithm to generate a new dataset allowing the estimation of a conditional logit model with state dependence using preimplemented procedures

**Corollary 3.4** (Corollary of 3.3).

It is equivalent to maximize the likelihood of a dynamic model such as 3.1 on the movers in $y$ and to maximize the likelihood of a static model with only one binary covariate such as 3.2 on the movers in $z$, such that $T' = T - 2$ and

\[
\begin{align*}
\text{There are } \sum_{t=2}^{T} y_{it} y_{i(t-1)} \text{ periods such that } (z_{it}, x_{it}) &= (1,1) \\
\text{There are } \sum_{t=1}^{T} y_{it} - y_{i1} - y_{iT} - \sum_{t=2}^{T} y_{it} y_{i(t-1)} \text{ periods such that } (z_{it}, x_{it}) &= (1,0) \\
\text{There are } \sum_{t=1}^{T} y_{it} - 1 - \sum_{t=2}^{T} y_{it} y_{i(t-1)} \text{ periods such that } (z_{it}, x_{it}) &= (0,1) \\
\text{There are } T - 1 - 2 \sum_{t=1}^{T} y_{it} + \sum_{t=2}^{T} y_{it} y_{i(t-1)} + y_{i1} + y_{iT} \text{ periods such that } (z_{it}, x_{it}) &= (0,0)
\end{align*}
\]

The new dataset will look like the following:

\[
\begin{array}{c|cc|c}
z & x & \sum_{t=1}^{T} y_{it} y_{i(t-1)} & \sum_{t=2}^{T} y_{it} y_{i(t-1)} \\
1 & 1 & \sum_{t=2}^{T} y_{it} y_{i(t-1)} & T - 2 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 1 & \sum_{t=1}^{T} y_{it} - y_{i1} - y_{iT} - \sum_{t=2}^{T} y_{it} y_{i(t-1)} & \sum_{t=2}^{T} y_{it} y_{i(t-1)} \\
1 & 0 & \vdots & \vdots \\
1 & 0 & \sum_{t=1}^{T} y_{it} - 1 - \sum_{t=2}^{T} y_{it} y_{i(t-1)} & \sum_{t=2}^{T} y_{it} y_{i(t-1)} \\
0 & 1 & \vdots & \vdots \\
0 & 1 & T - 1 - 2 \sum_{t=1}^{T} y_{it} + \sum_{t=2}^{T} y_{it} y_{i(t-1)} & \sum_{t=2}^{T} y_{it} y_{i(t-1)} \\
0 & 0 & \vdots & \vdots \\
0 & 0 & T - 1 - 2 \sum_{t=1}^{T} y_{it} + y_{i1} + y_{iT} + \sum_{t=2}^{T} y_{it} y_{i(t-1)} & \sum_{t=2}^{T} y_{it} y_{i(t-1)} \\
0 & 0 & \sum_{t=2}^{T} y_{it} y_{i(t-1)} & T - 2
\end{array}
\]

The result gives a very simple strategy to estimate a conditional logit model with state
dependence with a program which estimates conditional logit models without state dependence. After removing the observations such that \( \sum_{t=2}^{T-1} y_{it} \in \{0, T-2\} \), it suffices to generate a new dataset with two binary covariates as described above, and then to apply the maximization of the conditional likelihood in the static framework.

Note that, the uninformative trajectories such that \( \sum_{t=2}^{T-1} y_{it} = \sum_{t=1}^{T} y_{it} = 1 \) or \( \sum_{t=2}^{T-1} y_{it} = \sum_{t=1}^{T} y_{it} - 2 = T-3 \) do not require any special treatment. Indeed, it is equivalent to remove them from the analysis or to keep them and use the above transformation which will lead to uninformative static trajectories anyway.

### 3.5 Examples of data generation with \( T = 6 \)

The following two numerical examples are derived using corollary 3.4:

1. \( y = (0, 1, 1, 1, 0, 1) \) becomes

\[
\begin{align*}
\begin{cases}
  z & x \\
  1 & 1 \\
  1 & 1 \\
  1 & 0 \\
  0 & 1 \\
\end{cases}
\begin{array}{c}
\sum_{t=2}^{T} y_{it} y_{i,t-1} = 2 \\
\sum_{t=1}^{T} y_{i} - y_{1} - y_{T} - \sum_{t=2}^{T} y_{it} y_{i,t-1} = 1 \\
\sum_{t=1}^{T} y_{it} - 1 - \sum_{t=2}^{T} y_{it} y_{i,t-1} = 1 \\
\end{array}
\end{align*}
\]

2. \( y = (1, 0, 1, 0, 0, 1) \) becomes

\[
\begin{align*}
\begin{cases}
  z & x \\
  1 & 0 \\
  0 & 1 \\
  0 & 1 \\
\end{cases}
\begin{array}{c}
\sum_{t=1}^{T} y_{i} - y_{1} - y_{T} - \sum_{t=2}^{T} y_{it} y_{i,t-1} = 1 \\
\sum_{t=1}^{T} y_{it} - 1 - \sum_{t=2}^{T} y_{it} y_{i,t-1} = 2 \\
T - 1 - 2 \sum_{t=1}^{T} y_{i} + y_{1} + y_{T} + \sum_{t=2}^{T} y_{it} y_{i,t-1} = 1 \\
\end{array}
\end{align*}
\]

### 4 Conclusion

In the case of a binary outcome with one lag of state dependence and individual unobserved heterogeneity, we provide a simple closed form for the conditional likelihood. Moreover, we show that such a conditional likelihood is the same as that of a conditional logit model with one
binary exogenous variable and no state dependence. Therefore these results show that in order to evaluate a conditional logit model with one lag of state dependence, it is possible to simply generate a two-variable dataset and use standard procedures originally intended for a framework without state dependence. Examples of programs which can generate such a dataset are provided in appendix A. Proofs of lemmas and theorems are provided in appendix B. Extensions to more than two states and more than one lag of state dependence are left for further research.
References


Appendix

Appendix A provides a SAS macro and a piece of R code which help reshape datasets in a way that makes them directly usable by traditional conditional logit procedures. Appendix B gives the proofs concerning the closed forms of the conditional likelihoods in the static and in the dynamic case as well as the proof of the link between them.

A Examples of programs

A.1 SAS

A.1.1 Preimplemented procedure for the static framework

Estimation of a conditional logit in a static framework using a preimplemented procedure:

/*Example assuming dataset OBS contains three variables:
   - I for the individuals
   - Z for the LHS variable
   - X for the RHS variable
*/

proc logistic data=obs desc;
   strata i;
   model z=x;
run;

A.1.2 Faster estimation for the static framework

A faster computation can be obtained using lemma (3.2)

data obs2;
   set obs;
   zx=z*x;
run;

proc summary data=obs2 nway;
   class i;

var z x zx;
output out=tab(rename=(_freq_=t)) sum=sumz sumx sumxz;
run;

proc iml;
    use tab;
    read all var {sumxz} into sumxz;
    read all var {sumx} into sumx;
    read all var {sumz} into sumz;
    read all var {t} into t;

    /* LOG-LIKELIHOOD */
    start logvrais(d) global(sumxz,sumx,sumz,t);
        value=0;
        do i=1 to nrow(t);
            j=do(max(0,sumz[i,1]+sumx[i,1]-t[i,1]),min(sumz[i,1],sumx[i,1]),1);
            value=value-log(comb(sumx[i,1],j)*comb(t[i,1]-sumx[i,1],sumz[i,1]-j)*exp
                          (((j-sumxz[i,1])*d)[+]));
        end;
        return(value);
    finish;

    optn={1 0};
    delta0={0};
    call nlpnra(rc,delta,'logvrais',delta0,optn);

    /* STANDARD DEVIATION */
    start varas(d) global(sumxz,sumx,sumz,t);
        value=0;
        do i=1 to nrow(t);
            j=do(max(0,sumz[i,1]+sumx[i,1]-t[i,1]),min(sumz[i,1],sumx[i,1]),1);
            value=value+(((j-sumxz[i,1])*comb(sumx[i,1],j)*comb(t[i,1]-sumx[i,1],sumz[i,1]-j)*exp
                          (((j-sumxz[i,1])*d)[+])*2
                        /((comb(sumx[i,1],j)*comb(t[i,1]-sumx[i,1],sumz[i,1]-j)*exp(((j-
                          sumxz[i,1])*d)[+])*2;
        end;
value=1/nrow(t)*value;
return(value);
finish;

ec=1/sqrt(nrow(t)*varas(delta));

*NPRINT THE RESULTS (ESTIMATION AND STANDARD DEVIATION)*
print delta ec;
close;quit;run;

A.1.3 Reshaping the data to estimate the dynamic model using the preimplemented procedure intended for static models

%macro dyn2sta( table_in=, /* entry table (must be sorted by ident and time) */
    table_out=, /* output table */
    ident=, /* individual identifier */
    time=, /* period identifier */
    state=, /* state variable (0 or 1) in the original dataset */
    static=, /* LHS variable in the new dataset */
    static_1= /* RHS variable in the new dataset */
);
data &table_out.(drop= lagstate sumY sumYLY Y1 T k
    n11 n10 n01 n00);
set &table_in.;
by &ident. &time.;
retain lagstate sumY sumYLY Y1 T;
if first.&ident. then do;
    lagstate=&state.;
    sumY=&state.;
    sumYLY=0;
    Y1=&state.;
    T=1;
end;
else do;
    T=T+1;
    sumY=sumY+&state.;
    sumYLY=sumYLY+lagstate*&state.;

18
lagstate=&state;
if last.&ident. and 0<sumY - Y1 - &state. <T-2 and 1<sumY<T-1
   and T>=4 then do;
   _n11= sumYLY;
   _n10= sumY - Y1 - &state. - sumYLY;
   _n01= sumY - 1 - sumYLY;
   _n00= T - 2 - _n11 - _n10 - _n01;
   if _n11>0 then do _k=1 to _n11;
       &static.=1;
       &static_1.=1;
       output;
   end;
   if _n10>0 then do _k=1 to _n10;
       &static.=1;
       &static_1.=0;
       output;
   end;
   if _n01>0 then do _k=1 to _n01;
       &static.=0;
       &static_1.=1;
       output;
   end;
   if _n00>0 then do _k=1 to _n00;
       &static.=0;
       &static_1.=0;
       output;
   end;
end;
run;
%mend;

A.1.4 Estimation of the dynamic model using lemma (3.1)

For fast computation (necessary when \( n > 10000 \)), instead of using \texttt{proc logistic} reshape the
data and use lemma (3.2) or equivalently use lemma (3.1) as follows.
%macro dyn2iml(table_in=, /* entry table (must be sorted by ident and time) */
    table_out=, /* output table with variables : */
    __Y1, __YT, __sumY, __sumYLY */
    ident=, /* individual identifier */
    time=, /* period identifier */
    state=, /* state variable (0 or 1) in the original dataset */
  );
  data &table_out.(drop=lagstate);
    set &table_in ;
    by &ident &time ;
    retain lagstate sumY sumYLY Y1 T ;
    if first &ident then do;
      lagstate=&state ;
      sumY=&state ;
      sumYLY=0 ;
      Y1=&state ;
      T=1 ;
    end ;
    else do;
      T=T+1 ;
      sumY=sumY+&state ;
      sumYLY=sumYLY+lagstate &state ;
      lagstate=&state ;
      if last &ident and 0<sumY—Y1—&state.<T—2 and 1<sumY<T—1
            and T>4 then do;
        YT=&state ;
        output ;
      end ;
    end ;
  run ;
%mend ;
/*Example assuming dataset OBS contains three variables :
   -I for the individuals
   -T for the periods
   -Y for the dependant variable
*/
%dyn2iml(table_in=OBS,
table_out=tab,
ident=I,
time=T,
state=Y)
proc iml;
use tab;
read all var {__sumY} into sumy;
read all var {__sumLY} into sumly;
read all var {__y1} into y1;
read all var {__YT} into yT;
read all var {__T} into T;

/*/LOG-LIKELIHOOD*/
start logvrais(delta) global(sumy,sumly,y1,yT,t);
value=0;
do i=1 to nrow(t);
   j=do(max(0,2*sumy[i,1]-t[i,1]+1-y1[i,1]-yT[i,1]),sumy[i,1]-1-y1[i,1]*yT[i,1],1);
   value=value-log((comb(sumy[i,1]-1,j)#comb(t[i,1]-sumy[i,1]-1,sumy[i,1]-j-y1[i,1]-yT[i,1])-exp((j-sumly[i,1])*delta))]+)
end;
return(value);
finish;

optn={1 0};
delta0={0};
call nlpnra(rc,delta,'logvrais',delta0,optn);

/*/STANDARD DEVIATION*/
start varas(delta) global(sumy,sumly,y1,yT,t);
value=0;
do i=1 to nrow(t);
   j=do(max(0,2*sumy[i,1]-t[i,1]+1-y1[i,1]-yT[i,1]),sumy[i,1]-1-y1[i,1]*yT[i,1],1);
   value=value+(1/2*(sumy[i,1]-1-y1[i,1]*yT[i,1]-t[i,1]+1-y1[i,1]-yT[i,1]))
end;
return(value);
finish;
\[(j - \text{sumyly}[i,1]) \cdot \text{comb}(\text{sumy}[i,1] - 1, j) \cdot \text{comb}(t[i,1] - \text{sumy}[i,1] - j - y_1[i,1] - yT[i,1]) \cdot \exp \left( (j - \text{sumyly}[i,1]) \cdot \delta \right) + 1 \cdot \right] \right)^2 / \left( \left[ \text{comb}(\text{sumy}[i,1] - 1, j) \cdot \text{comb}(t[i,1] - \text{sumy}[i,1] - 1, \text{sumy}[i,1] - j - y_1[i,1] - yT[i,1]) \cdot \exp \left( (j - \text{sumyly}[i,1]) \cdot \delta \right) + 1 \right] \right)^2 \end{equation}

\[\text{value} = 1 / \text{nrow}(t) \cdot \text{value} ; \]
\[\text{return}(\text{value}) ; \]
\[\text{finish} ; \]
\[\text{ec} = 1 / \sqrt{\text{nrow}(t) \cdot \text{varas}(\delta)} ; \]
\[\text{print}\ delta ec ; \]
\[\text{close} ; \text{quit} ; \]
\[\text{run} ; \]

A.2 R

A.2.1 Preimplemented procedure for the static framework

###

# Example assuming dataset base.sta contains three variables:
# - I for the individuals
# - Z for the LHS variable
# - X for the RHS variable

###

### conditional logit

library(survival)

res <- clogit(Z ~ X + strata(I), data=base.sta)

summary(res)

A.2.2 Reshaping the data to estimate the dynamic model using the preimplemented procedure intended for static models
# example of a program which transforms a dataset with 3 variables
# − ID individual identifier
# − T time variable
# − binary state variable Y
# on which you want to perform a conditional logit analysis with state dependence,
# into a dataset with 4 variables:
# − ID
# − T
# − binary state variable Z
# − binary "modified lagged state variable" X
# which can be used directly with standard conditional logit procedures

###

# set working directory and load original data base
#setwd(""")
#base <- read.csv("", sep="", header=T)

### the dataset must be sorted by ID and T

### change names into ID, T and Y
# names(base)[names(base)=="Old.Y.Name"] <- "Y"

names(base)

### number of observations for each individual
base.agg <- aggregate(base$ID, by=base["ID"], length)
names(base.agg)[names(base.agg)=="x"] <- "nbobs"

### first and last observation for each individual
base.agg$Y.1 <- aggregate(base$Y, by=base["ID"], head, n=1)[,"x"]
base.agg$Y.T <- aggregate(base$Y, by=base["ID"], tail, n=1)[,"x"]

### sum of Y between 1 and T
base.agg$sum.Y <- aggregate(base$Y, by=base["ID"], sum)[,"x"]

### lag of Y: LY, put 0 for the first period of each individual
LY <- c(0, base$Y[-length(base$Y)])
LY[c(1, cumsum(base.agg$nbobs)[-nrow(base.agg)]+1)] <- 0

### sum of $Y_{i}Y_{i+1}$ between 2 and T
base.agg$sum.YLY <- aggregate(base$Y*LY, by=base[, "ID"], sum[, "x"])

### we keep only the individuals such that
### 0 < $\sum Y - Y_{i} - Y_{i+1} < T-2$
### 1 < $\sum Y < T-1$

base.agg.ident <- subset(base.agg, 
    sum.Y - Y.1 - Y.T < nbobs - 2 &
    sum.Y - Y.1 - Y.T > 0 &
    sum.Y < nbobs - 1 &
    sum.Y > 1)

names(base.agg.ident)[names(base.agg.ident)["sum.YLY"] <- "n.11"

base.agg.ident$n.10 <- base.agg.ident$sum.Y - base.agg.ident$Y.1 - 
    base.agg.ident$Y.T - base.agg.ident$n.11
base.agg.ident$n.01 <- base.agg.ident$sum.Y - 1 - base.agg.ident$n.11
base.agg.ident$n.00 <- base.agg.ident$nbobs - 2 - base.agg.ident$n.11 - 
    base.agg.ident$n.10 - base.agg.ident$n.01

### new data.frame with ID, LHS variable: ZZ, RHS variable: XX
repseq <- do.call(c, as.list(t(base.agg.ident[, c("n.11","n.10","n.01","n.00")])))
base.dyn <- data.frame(ID=rep(base.agg.ident$ID, base.agg.ident$nbobs-2), 
    ZZ=rep(rep(c(1,1,0,0), nrow(base.agg.ident)), repseq),
    XX=rep(rep(c(1,0,1,0), nrow(base.agg.ident)), repseq))

### conditional logit
library(survival)
res <- clogit(ZZ ~ XX + strata(ID), data=base.dyn)
summary(res)
B Proofs

B.1 Proof of 3.1

Lemma B.1 (Link between the number of occurrences of each state and the transitions from one state to another).

Let

- \(T \geq 4\) the number of time periods,
- \(j_{\ell,\ell'}, (\ell, \ell') \in \{0,1\}^2\) the number of transitions from state \(\ell\) to state \(\ell'\),
- \(k_{\ell}, \ell \in \{0,1\}\) the number of occurrences of state \(\ell\),
- \(y_1\) and \(y_T\) the first and last states,

the knowledge of \(k_1\) (and therefore also \(k_0 = T - k_1\), \(y_1, y_T\) and \(j_{11}\) determines \(j_{00}, j_{10}\) and \(j_{01}\) and conversely.

Proof. Immediate using:

\[
\begin{aligned}
    j_{00} + j_{01} + (1 - y_T) &= k_0 \\
    j_{10} + j_{11} + y_T &= k_1 \\
    j_{00} + j_{10} + (1 - y_1) &= k_0 \\
    j_{01} + j_{11} + y_1 &= k_1
\end{aligned}
\]

\[\square\]

Lemma B.2 (Counting trajectories as a function of the number of 1’s and the number of “11” transitions).

Let the two states be 0 and 1.

The number of trajectories of length \(T\), starting with \(y_1\), ending with \(y_T\), with \(k_1\) (1 \(\leq k_1 \leq T - 1\)) occurrences of 1 and \(j_{11}\) transitions from state 1 to state 1 is:\(^1\)

\[
\binom{k_1 - 1}{j_{11}} \binom{T - k_1 - 1}{k_1 - j_{11} - y_1 - y_T} \mathbb{I}_{\max(0, 2k_1 - T + 1 - y_1 - y_T) \leq j_{11} \leq k_1 - 1 - y_1 y_T}
\]

\(^1\)using the notation \(\binom{n}{p} = \frac{n!}{(n-p)p!}\)
Conversely, using the number of occurrences of state 0: \( k_0 = T - k_1 \), and the number of transitions from state 0 to state 0: \( j_{00} = k_0 - k_1 + j_{11} + y_1 + y_T - 1 \), the previous number of trajectories is also equal to:

\[
\left( k_1 - 1 \right) \left( k_0 - 1 \right) \mathbb{1}_{\{0 \leq j_{11} \leq k_1 - 1 - y_1, y_T, 0 \leq j_{00} \leq k_0 - 1 - (1 - y_1)(1 - y_T)\}}
\]

Proof. To answer this question we shall introduce the following notations: let \( y = (y_1, \ldots, y_T) \) be a trajectory, then define \( \text{Red}(y) \) the “reduced trajectory” of \( y \) which is obtained by concatenation of the successive zeros and ones. For instance, \( \text{Red}(1, 1, 1, 0, 1, 1, 0) = (1, 0, 1, 0) \). Then, let \((0, 1)_n \) (respectively \((1, 0)_n \)) be “\( n \) times” the sequence \((0, 1) \) (respectively the sequence \((1, 0) \)). This leads for instance to: \( \text{Red}(1, 1, 1, 0, 1, 1, 0) = ((1, 0)_2) \).

The interest of defining these reduced forms lies in the fact that two trajectories with the same reduced form also share the same likelihood.

The number of 1’s, \( k_1 \), in a trajectory of reduced form \((0, 1)_n \) or \((1, 0)_n \) is: \( k_1 = n + j_{11} \) where \( j_{11} \) is the number of transitions from state 1 to state 1. Indeed, noting that a trajectory of reduced form \((1, 0)_n \) will look like:

\[
(1, \ldots, 1, 0, \ldots, 0, 1, \ldots, 1, 0, \ldots, 0, \ldots, 1, \ldots, 1, 0, \ldots, 0),
\]

in each block of 1’s, the number of 1’s is equal to 1 plus “the number of transitions from 1 to 1 in that block”. Summing over the \( n \) blocks, the total number of 1’s will be \( j_{11} + n \).

Therefore,

- if \( y_1 = y_T = 0 \) then \( \text{Red}(y) = 0 (10)^{k_1 - j_{11}} \)
- \( y_1 = 1 \) and \( y_T = 0 \) then \( \text{Red}(T) = (10)^{k_1 - j_{11}} \)
- \( y_1 = 0 \) and \( y_T = 1 \) then \( \text{Red}(T) = (01)^{k_1 - j_{11}} \)
- \( y_1 = 1 \) and \( y_T = 1 \) then \( \text{Red}(T) = 1 (01)^{k_1 - j_{11} - 1} \)

Now we have to count the number of full trajectories which lead to each of these four reduced trajectories. We use the fact that the number of occurrences of state 0 is \( k_0 = T - k_1 \).
The number of elements in $\text{Red}^{-1}((0, 1)_{k_1-j_{11}})$ is the number of ways to make $k_1 - j_{11}$ groups out of $k_1$ identical elements (the 1’s), times the number of ways to make $k_1 - j_{11}$ groups out of $k_0$ identical elements (the 0’s). Using the fact that if you write down a list of $n$ identical elements on one line, making $p$ groups out of these $n$ elements is like setting $p - 1$ separators in the $n - 1$ spots between the elements, and there are $\binom{n-1}{p-1}$ ways to do so.

The answer is thus $(k_1-1)\binom{k_1-1}{j_{11}}(k_0-1)\binom{k_0-1}{j_{11}+1} = (k_1-1)\binom{k_0-1}{k_1-j_{11}-1}$, with $\max(0, 2k_1 - T) \leq j_{11} \leq k_1 - 1$.

Similar arguments can be made for the other trajectories and the results are summarized as follows:

- $\#\{\text{Red}^{-1}(0(1))_{k_1-j_{11}}\} = (k_1-1)\binom{k_1-1}{j_{11}}(k_0-1)\binom{k_0-1}{j_{11}+1}$ for $\max(0, 2k_1 - T + 1) \leq j_{11} \leq k_1 - 1$
- $\#\{\text{Red}^{-1}((10))_{k_1-j_{11}}\} = (k_1-1)\binom{k_1-1}{j_{11}}(k_0-1)\binom{k_0-1}{j_{11}+1} = (k_1-1)\binom{k_1-1}{j_{11}}(k_1-j_{11}-1)$ for $\max(0, 2k_1 - T) \leq j_{11} \leq k_1 - 1$
- $\#\{\text{Red}^{-1}((01))_{k_1-j_{11}}\} = (k_1-1)\binom{k_1-1}{j_{11}}(k_0-1)\binom{k_0-1}{j_{11}+1} = (k_1-1)\binom{k_1-1}{j_{11}}(k_1-j_{11}-1)$ for $\max(0, 2k_1 - T) \leq j_{11} \leq k_1 - 1$
- $\#\{\text{Red}^{-1}(1(10))_{k_1-j_{11}}\} = (k_1-1)\binom{k_1-1}{j_{11}}(k_0-1)\binom{k_0-1}{j_{11}+1} = (k_1-1)\binom{k_1-1}{j_{11}}(k_1-j_{11}-2)$ for $\max(0, 2k_1 - T - 1) \leq j_{11} \leq k_1 - 2$

Finally, the number of trajectories is $\binom{k_1-1}{j_{11}}(k_1-j_{11}-1)$ with $\max(0, 2k_1 - T + 1 - y_1 - y_T) \leq j_{11} \leq k_1 - 1 - y_T$.

The symmetrical expression using the 0’s follows directly.

**Lemma B.3** (Closed form for the denominator of a conditional logit with state dependence).

Let $B(T, k_1, y_1, y_T)$ be the number of trajectories of length $T$, with $k_1$ 1’s ($1 \leq k_1 \leq T - 1$), starting with $y_1$ and finishing with $y_T$:

$$B(T, k_1, y_1, y_T) = \left\{ (\tilde{y}_{1}, \ldots, \tilde{y}_{T}) \in \{0, 1\}^{T} | \tilde{y}_{11} = y_{11}, \tilde{y}_{T} = y_{T}, \sum_{t=1}^{T} \tilde{y}_t = k_1 \right\}$$
Then:

\[
\sum_{\tilde{y} \in B(T, k_1, y_T)} \exp \left( \sum_{t=2}^{T} \tilde{y}_{it} \delta_{it-1} \right) = \sum_{j_{11} = \max(0, 2k_1 - T + 1 - y_1 - y_T)}^{k_1 - 1 - y_1, y_T} \left( k_1 - 1 \right) \left( T - k_1 - 1 \right) \left( k_1 - j_{11} - y_1 - y_T \right) e^{\delta_{j_{11}}}
\]

Proof. Let \( A(T, k_1, y_1, y_T, j_{11}) = B(T, k_1, y_1, y_T) \cap \\{ (\tilde{y}_1, \ldots, \tilde{y}_T) \in \{0, 1\}^T | \sum_{t=2}^{T} \tilde{y}_t = j_{11} \} \)
then \( B(T, k_1, y_1, y_T) = \bigcup_{j_{11} \in \mathbb{N}} A(T, k_1, y_1, y_T, j_{11}) \)
and \( A(T, k_1, y_1, y_T, j_{11}) \cap A(T, k_1, y_1, y_T, j'_{11}) = \emptyset \)

It follows that

\[
\sum_{\tilde{y} \in B(T, k_1, y_1, y_T)} \exp \left( \sum_{t=2}^{T} \tilde{y}_{it} \delta_{it-1} \right) = \sum_{j_{11} \in \mathbb{N}} \# \{ A(T, k_1, y_1, y_T, j_{11}) \} e^{\delta_{j_{11}}}
\]

and then, from lemma B.2:

\[
\sum_{\tilde{y} \in B(T, k_1, y_1, y_T)} \exp \left( \sum_{t=2}^{T} \tilde{y}_{it} \delta_{it-1} \right) = \sum_{j_{11} = \max(0, 2k_1 - T + 1 - y_1 - y_T)}^{k_1 - 1} \left( k_1 - 1 \right) \left( T - k_1 - 1 \right) \left( k_1 - j_{11} - y_1 - y_T \right) e^{\delta_{j_{11}}}
\]

\( \square \)

**Theorem B.4** (Closed form for the conditional likelihood of a conditional logit with two states and one lag of state dependence).

The conditional likelihood of any trajectory of length \( T \), with \( k_1 \) 1's \((1 \leq k_1 \leq T - 1)\), starting with \( y_1 \) and finishing with \( y_T \) is equal to:

\[
P(y_1, \ldots, y_T | y_1, y_T, \sum_{t=1}^{T} y_t) = \frac{\exp \left( \sum_{t=2}^{T} y_t \delta_{it-1} \right)}{\sum_{j_{11} = \max(0, 2k_1 - T + 1 - y_1 - y_T)}^{k_1 - 1} \left( k_1 - 1 \right) \left( T - k_1 - 1 \right) \left( k_1 - j_{11} - y_1 - y_T \right) e^{\delta_{j_{11}}}}
\]

Proof. direct consequence of lemma B.3. \( \square \)

**B.2 Proof of 3.2**

Consider a trajectory \((z_{it}, x_{it})_{1 \leq t \leq T'} \in \{0, 1\}^{2T'} \) and let \( n_k \) be the number of trajectories \((\tilde{z}_{it})_{1 \leq t \leq T'} \in \{0, 1\}^{T'} \) such that \( \sum_{t=1}^{T'} \tilde{z}_{it} = \sum_{t=1}^{T'} z_{it} \) and \( \sum_{t=1}^{T'} \tilde{z}_{it} x_{it} = k \).
If $J$ is a nonempty subset of $\mathbb{R}$, $(\delta \mapsto \exp(k\delta))_{k \in J}$ is a free family of the mappings from $\mathbb{R}$ to $\mathbb{R}$, i.e. if $\forall \delta \in \mathbb{R}, \sum_{k \in J} \lambda_k e^{k\delta} = 0$ then, $\forall k \in J, \lambda_k = 0$.

Using lemma 3.1 and lemma 3.2, and noting:

$$J_{sta} = \left[ \max \left(0, \sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T', 0 \right), \min \left(\sum_{t=1}^{T'} z_{it}, \sum_{t=1}^{T'} x_{it} \right) \right] \cap \mathbb{N} \quad \text{and} \quad J_{dyn} = \left[ \max \left(0, 2 \sum_{t=1}^{T} y_{it} - T + 1 - y_1 - y_T \right), \sum_{t=1}^{T} y_{it} - 1 - y_1 y_T \right] \cap \mathbb{N},$$

the likelihood of a trajectory $(z, x)$ in the static model will be the same as the likelihood of a trajectory $y$ in the dynamic model if and only if the following sets are equal:

$n_k$ is equal to 0 if $k \not\in \max \left(\sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T', 0 \right), \min \left(\sum_{t=1}^{T'} z_{it}, \sum_{t=1}^{T'} x_{it} \right)] \cap \mathbb{N},$ choosing a trajectory $(\bar{z}_{it})_{1 \leq i \leq T'} \in \{0, 1\}^{T'}$, corresponds to choosing $k$ elements among $\sum_{t=1}^{T'} x_{it}$ (i.e. $k$ periods $t$ such that $\bar{z}_{it} = x_{it} = 1$), and $\sum_{t=1}^{T'} z_{it} - k$ elements among $T' - \sum_{t=1}^{T'} x_{it}$ (i.e. $\sum_{t=1}^{T'} z_{it} - k$ periods $t$ such that $\bar{z}_{it} = 1$ and $x_{it} = 0$).

So, in this case, $n_k = \left(\sum_{t=1}^{T'} x_{it}\right)\left(\sum_{t=1}^{T'} \bar{z}_{it} - y_{it} \right)$ and the result follows.

**B.3 Proof of 3.3**

If $J$ is a non-empty subset of $\mathbb{R}$, $(\delta \mapsto \exp(k\delta))_{k \in J}$ is a free family of the mappings from $\mathbb{R}$ to $\mathbb{R}$, i.e. if $\forall \delta \in \mathbb{R}, \sum_{k \in J} \lambda_k e^{k\delta} = 0$ then, $\forall k \in J, \lambda_k = 0$.

Using lemma 3.1 and lemma 3.2, and noting:

$$J_{sta} = \left[ \max \left(0, \sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T', 0 \right), \min \left(\sum_{t=1}^{T'} z_{it}, \sum_{t=1}^{T'} x_{it} \right) \right] \cap \mathbb{N} \quad \text{and} \quad J_{dyn} = \left[ \max \left(0, 2 \sum_{t=1}^{T} y_{it} - T + 1 - y_1 - y_T \right), \sum_{t=1}^{T} y_{it} - 1 - y_1 y_T \right] \cap \mathbb{N},$$

the likelihood of a trajectory $(z, x)$ in the static model will be the same as the likelihood of a trajectory $y$ in the dynamic model if and only if the following sets are equal:

$n_k$ is equal to 0 if $k \not\in \max \left(\sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T', 0 \right), \min \left(\sum_{t=1}^{T'} z_{it}, \sum_{t=1}^{T'} x_{it} \right)] \cap \mathbb{N},$ choosing a trajectory $(\bar{z}_{it})_{1 \leq i \leq T'} \in \{0, 1\}^{T'}$, corresponds to choosing $k$ elements among $\sum_{t=1}^{T'} x_{it}$ (i.e. $k$ periods $t$ such that $\bar{z}_{it} = x_{it} = 1$), and $\sum_{t=1}^{T'} z_{it} - k$ elements among $T' - \sum_{t=1}^{T'} x_{it}$ (i.e. $\sum_{t=1}^{T'} z_{it} - k$ periods $t$ such that $\bar{z}_{it} = 1$ and $x_{it} = 0$).

So, in this case, $n_k = \left(\sum_{t=1}^{T'} x_{it}\right)\left(\sum_{t=1}^{T'} \bar{z}_{it} - y_{it} \right)$ and the result follows.
First we show (1.) by summing all the terms in both sets and using Vandermonde's identity:

\[
\sum_{t=1}^{T-1} y_{it} - 1 - y_1 y_T \sum_{t=1}^{T} y_{it} - 1 \sum_{t=1}^{T} y_{it} - j - y_1 - y_T = \left( \frac{T-2}{\sum_{t=2}^{T} y_{it}} \right)
\]

and

\[
\min \left( \sum_{t=1}^{T'} z_{it} \sum_{t=1}^{T'} x_{it} \right) \sum_{j=\max(0, \sum_{t=1}^{T'} z_{it} + \sum_{t'=1}^{T'} x_{it} - T')}^{\sum_{t=1}^{T'} z_{it} - j} \left( \sum_{t=1}^{T'} x_{it} \sum_{t=1}^{T'} z_{it} - j \right) = \left( \sum_{t=1}^{T'} z_{it} \right)
\]

Thus, assuming that we want a general relationship between \(T\) and \(T'\) which should be independent from the trajectories \(y\) and \((z, x)\), for one part \(T' = T - 2\) and for the other part,

\[
\sum_{t=1}^{T-2} z_{it} = \sum_{t=2}^{T-1} y_{it} \quad \text{or} \quad \sum_{t=1}^{T-2} z_{it} = T - 2 - \sum_{t=2}^{T-1} y_{it}
\]

Second, we consider the case of the uninformative trajectories (end of 2.):

- If \(\sum_{t=2}^{T-1} y_{it} = 1\) and \(y_{i1} = y_{iT} = 0\) then the equality of the likelihoods implies that \(\sum_{t=1}^{T-2} z_{it} \in \{1, T - 3\}\) and the likelihood of the static model does not depend on \(\delta\) if and only if \(x\) is constant.

- If \(\sum_{t=2}^{T-1} y_{it} = T - 3\) and \(y_{i1} = y_{iT} = 1\) then the equality of the likelihoods implies that \(\sum_{t=1}^{T-2} z_{it} \in \{1, T - 3\}\) and the likelihood of the static model does not depend on \(\delta\) if and only if \(x\) is constant.

Third, the show the rest of (2.). Note first, that we only need to prove the result for \(\sum_{t=1}^{T-2} z_{it} = \sum_{t=2}^{T-1} y_{it}\) and \((y_1, y_T) \in \{(0, 0), (1, 0)\}\). Indeed, considering a logistic model with two binary variables \(U\) and \(V\) and unobserved heterogeneity \(\gamma_i\),

30
\[ P(U = u \mid V = v, \gamma_i) = \frac{e^{u(\delta v + \gamma_i)}}{1 + e^{\delta v + \gamma_i}} = \frac{e^{(1-u)(\delta(1-v) - \gamma_i)}}{1 + e^{-(1-v) - \gamma_i}}. \]

Therefore, since conditioning makes the unobserved heterogeneity disappear, for a dynamic model, the conditional likelihoods of the trajectories \( y = (y_1, y_2, \ldots, y_T) \) and \( (1 - y) = (1 - y_1, 1 - y_2, \ldots, 1 - y_T) \) will be the same, and likewise, for a static model, the conditional likelihoods of the trajectories \( (z, x) = ((z_1, x_1), \ldots, (z_T, x_T)) \) and \( (1 - z, 1 - x) = ((1 - z_1, 1 - x_1), \ldots, (1 - z_T, 1 - x_T)) \) will also be the same.

Now, assume that \( T - 3 \geq \sum_{t=2}^{T-1} y_{it} \geq 1 \) and \( T - 2 \geq \sum_{t=1}^{T} y_{it} \geq 2 \) and \( \sum_{t=1}^{T-2} z_{it} = \sum_{t=2}^{T-1} y_{it} \).

The only conditions we need concern \( \sum_{t=1}^{T-2} x_{it} \) and \( \sum_{t=1}^{T-2} x_{it}z_{it} \).

1. Trajectories such that \((y_1, y_T) = (1, 0)\).

Considering the coefficient of the term of highest degree in each set gives:

\[
\left( \frac{\sum_{t=1}^{T-2} x_{it}}{\min(\sum_{t=1}^{T-2} z_{it}, \sum_{t=1}^{T-2} x_{it})} \right)^{T - 2 - \sum_{t=1}^{T-2} x_{it}} = 1
\]

If \( \sum_{t=1}^{T-2} z_{it} < \sum_{t=1}^{T-2} x_{it} \), then \( \sum_{t=1}^{T-2} z_{it} = 0 \) which is impossible, so we deduce that \( \sum_{t=1}^{T-2} x_{it} \leq \sum_{t=1}^{T-2} z_{it} \). Then we have either, \( T - 2 - \sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} z_{it} - \sum_{t=1}^{T-2} x_{it} \) or \( \sum_{t=1}^{T-2} z_{it} = \sum_{t=1}^{T-2} x_{it} \). The first choice is impossible because \( \sum_{t=1}^{T-2} z_{it} < T - 2 \) and then \( \sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} y_{it} \).

Now considering the highest degree in both sets, we have: \( \sum_{t=1}^{T-2} z_{it} = \sum_{t=1}^{T-2} x_{it}z_{it} = \sum_{t=1}^{T} y_{it} - 1 - \sum_{t=1}^{T} y_{it}y_{it-1} \) and therefore, \( \sum_{t=1}^{T} x_{it}z_{it} = \sum_{t=2}^{T} y_{it}y_{it-1} \).

2. Trajectories such that \((y_1, y_T) = (0, 0)\).

We consider the terms of lowest degree and therefore distinguish between:

\[ 2 \sum_{t=1}^{T} y_{it} - T + 1 - y_1 - y_T < 0, \quad = 0, \quad \text{and} \quad > 0. \]
We have $\sum_{t=1}^{T-1} z_{it} < T - 1 - \sum_{t=1}^{T-2} z_{it}$ and $\sum_{t=1}^{T} y_{it} - 1 = \sum_{t=1}^{T-2} z_{it} - 1$.

The equality of the number of terms in each likelihood gives:

$$\min \left( \sum_{t=1}^{T-2} z_{it} + 1, \sum_{t=1}^{T-2} x_{it} + 1, T - 1 - \sum_{t=1}^{T-2} x_{it}, T - 1 - \sum_{t=1}^{T-2} z_{it} \right) = \sum_{t=1}^{T-2} z_{it}.$$

Because $\sum_{t=1}^{T-2} z_{it} < \sum_{t=1}^{T-2} z_{it} + 1 \leq T - 1 - \sum_{t=1}^{T-2} z_{it}$, the equality of the likelihoods implies $\sum_{t=1}^{T-2} x_{it} + 1 = \sum_{t=1}^{T-2} z_{it}$ or $T - 1 - \sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} z_{it}$.

- If $T - 1 - \sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} z_{it}$, the equality of the terms with the smallest index in each likelihood implies that:

$$\sum_{t=1}^{T-2} x_{it} z_{it} = \sum_{t=2}^{T} y_{it} y_{it-1} + 1$$

and

$$T - 1 - \sum_{t=2}^{T-1} y_{it} = \left( T - 1 - \sum_{t=2}^{T-1} y_{it} \right) \sum_{t=2}^{T-1} y_{it}$$

Because we exclude $\sum_{t=2}^{T-1} y_{it} = 1$ (uninformative trajectory), we deduce $2 \sum_{t=2}^{T-1} y_{it} = T - 2$.

- If $\sum_{t=1}^{T-2} x_{it} + 1 = \sum_{t=1}^{T-2} z_{it}$, we deduce that $\sum_{t=1}^{T-2} z_{it} x_{it} = \sum_{t=2}^{T} y_{it} y_{it-1}$

$\cdot 2 \sum_{t=1}^{T} y_{it} - T + 1 < 0$

We have $\sum_{t=1}^{T} z_{it} < T - 1 - \sum_{t=1}^{T-2} z_{it}$ and $\sum_{t=1}^{T} y_{it} - 1 = \sum_{t=1}^{T-2} z_{it} - 1$.

The equality of the number of terms in each likelihood gives:

$$\min \left( \sum_{t=1}^{T-2} z_{it} + 1, \sum_{t=1}^{T-2} x_{it} + 1, T - 1 - \sum_{t=1}^{T-2} x_{it}, T - 1 - \sum_{t=1}^{T-2} z_{it} \right) = \sum_{t=1}^{T-2} z_{it}.$$

Because $\sum_{t=1}^{T-2} z_{it} < \sum_{t=1}^{T-2} z_{it} + 1 \leq T - 1 - \sum_{t=1}^{T-2} z_{it}$, the equality of the likelihoods implies $\sum_{t=1}^{T-2} x_{it} + 1 = \sum_{t=1}^{T-2} z_{it}$ or $T - 1 - \sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} z_{it}$.

- If $\sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} x_{it} z_{it}$, the equality of the terms with the smallest index in the two likelihoods implies that $\sum_{t=1}^{T-2} x_{it} = 1$, and next $T = 3$ which is absurd.

- If $\sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T-2} x_{it} z_{it}$, the likelihoods are the same if and only if $\sum_{t=1}^{T-2} z_{it} x_{it} = \sum_{t=2}^{T} y_{it} y_{it-1}$.

$\cdot 2 \sum_{t=1}^{T} y_{it} - T + 1 > 0$

We have $\sum_{t=1}^{T} z_{it} > T - 1 - \sum_{t=1}^{T-2} z_{it}$, the number of terms in each likelihood is
\[
\min(\sum_{t=1}^{T-2} z_{it} + 1, \sum_{t=1}^{T-2} x_{it} + 1, T-1 - \sum_{t=1}^{T-2} x_{it}, T-1 - \sum_{t=1}^{T-2} z_{it}) \quad \text{and} \quad T-1 - \sum_{t=1}^{T-2} z_{it}.
\]

So, \( \sum_{t=1}^{T-2} x_{it} + \sum_{t=1}^{T-2} z_{it} - T + 2 \geq 0 \) and \( \sum_{t=1}^{T-2} z_{it} \geq \sum_{t=1}^{T-2} x_{it} \).

The equality of the terms with the smallest index in each likelihood implies that

\[
\sum_{t=1}^{T-2} x_{it} + \sum_{t=1}^{T-2} z_{it} - T + 2 = 2 \sum_{t=1}^{T} y_{it} - T + 1 - \sum_{t=2}^{T} y_{it} y_{i,t-1}
\]

and

\[
\left( \frac{\sum_{t=1}^{T-2} x_{it}}{\sum_{t=1}^{T} x_{it} + \sum_{t=1}^{T-2} z_{it} - T + 2} \right) = \left( \frac{\sum_{t=1}^{T} y_{it} - 1}{2 \sum_{t=1}^{T} y_{it} - T + 1} \right)
\]

or equivalently

\[
\sum_{t=1}^{T-2} x_{it} + 1 - \sum_{t=1}^{T-2} x_{it} z_{it} = \sum_{t=2}^{T} y_{it} - \sum_{t=2}^{T} y_{it} y_{i,t-1}
\]

and

\[
\left( \frac{\sum_{t=1}^{T-2} x_{it}}{T - 2 - \sum_{t=1}^{T-2} z_{it}} \right) = \left( \frac{\sum_{t=1}^{T} y_{it} - 1}{T - 2 - \sum_{t=1}^{T-2} z_{it}} \right)
\]

Because we exclude \( \sum_{t=2}^{T-1} y_{it} = T - 2 \), we deduce \( \sum_{t=1}^{T-2} x_{it} = \sum_{t=1}^{T} y_{it} - 1 \) and

\( \sum_{t=1}^{T-2} z_{it} x_{it} = \sum_{t=2}^{T} y_{it} y_{i,t-1} \).

The necessary conditions given above are trivially sufficient using the closed forms of the conditional likelihoods.
<table>
<thead>
<tr>
<th>Numéro</th>
<th>Auteur(s)</th>
<th>Titre</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 9001</td>
<td>J. FAYOLLE et M. FLEURBAEY</td>
<td>Accumulation, profitabilité et endettement des entreprises</td>
</tr>
<tr>
<td>G 9002</td>
<td>H. ROUSSE</td>
<td>Détection et effets de la multicollinearité dans les modèles linéaires ordinaires - Un prolongement de la réflexion de BELSLEY, KUH et WATISCH</td>
</tr>
<tr>
<td>G 9003</td>
<td>P. RALLE et J. TOUJAS-BERNATE</td>
<td>Indexation des salaires : la rupture de 1983</td>
</tr>
<tr>
<td>G 9004</td>
<td>D. GUELLEC et P. RALLE</td>
<td>Compétitivité, croissance et innovation de produit</td>
</tr>
<tr>
<td>G 9005</td>
<td>P. RALLE et J. TOUJAS-BERNATE</td>
<td>Les conséquences de la désindexation. Analyse dans une maquette prix-salaires</td>
</tr>
<tr>
<td>G 9101</td>
<td>Equipe AMADEUS</td>
<td>Le modèle AMADEUS - 1ère partie - Présentation générale</td>
</tr>
<tr>
<td>G 9102</td>
<td>J.L. BRILLET</td>
<td>Le modèle AMADEUS - 2ème partie - Propriétés variantielles</td>
</tr>
<tr>
<td>G 9103</td>
<td>D. GUELLEC et P. RALLE</td>
<td>Endogène growth and product innovation</td>
</tr>
<tr>
<td>G 9104</td>
<td>H. ROUSSE</td>
<td>Le modèle AMADEUS - 3ème partie - Le commerce extérieur et l'environnement international</td>
</tr>
<tr>
<td>G 9105</td>
<td>H. ROUSSE</td>
<td>Effets de demande et offre dans les résultats du commerce extérieur manufacturé de la France au cours des deux dernières décennies</td>
</tr>
<tr>
<td>G 9106</td>
<td>B. CREPON</td>
<td>Innovation, taille et concentration : causalités et dynamiques</td>
</tr>
<tr>
<td>G 9107</td>
<td>B. AMABLE et D. GUELLEC</td>
<td>Un panorama des théories de la croissance endogène</td>
</tr>
<tr>
<td>G 9108</td>
<td>M. GLAUDE et M. MOUTARDIER</td>
<td>Une évaluation du coût direct de l'enfant de 1979 à 1989</td>
</tr>
<tr>
<td>G 9109</td>
<td>P. RALLE et all</td>
<td>France - Allemagne : performances économiques comparées</td>
</tr>
<tr>
<td>G 9110</td>
<td>J.L. BRILLET</td>
<td>Micro-DMS - NON PARU</td>
</tr>
<tr>
<td>G 9112</td>
<td>B. CREPON et G. DUREAU</td>
<td>Investissement en recherche-développement : analyse de causalités dans un modèle d'accélérateur généralisé</td>
</tr>
<tr>
<td>G 9113</td>
<td>J.L. BRILLET, H. EKREL-ROUSSE, J. TOUJAS-BERNATE</td>
<td>“France-Allemagne Couplées” - Deux économies vues par une maquette macro-économique</td>
</tr>
<tr>
<td>G 9201</td>
<td>W.J. ADAMS, B. CREPON et D. ENCAOUA</td>
<td>Choix technologiques et stratégies de dissuasion d'entrée</td>
</tr>
<tr>
<td>G 9202</td>
<td>J. OLIVEIRA-MARTINS, J. TOUJAS-BERNATE</td>
<td>Les théories de la structure optimal du capital : quelques points de repère</td>
</tr>
<tr>
<td>G 9203</td>
<td>I. STAPIC</td>
<td>Les échanges internationaux de services de la France dans le cadre des négociations multilatérales du GATT (1re version) Novembre 1992 (version finale)</td>
</tr>
<tr>
<td>G 9204</td>
<td>P. SEESTRE</td>
<td>L'économétrie sur données individuelles-temporelles. Une note introductive</td>
</tr>
<tr>
<td>G 9205</td>
<td>H. EKREL-ROUSSE</td>
<td>Le commerce extérieur et l'environnement international dans le modèle AMADEUS (réestimation 1992)</td>
</tr>
<tr>
<td>G 9206</td>
<td>N. GREENAN et D. GUELLEC</td>
<td>Coordination within the firm and endogenous growth</td>
</tr>
<tr>
<td>G 9207</td>
<td>A. MAGINIER et J. TOUJAS-BERNATE</td>
<td>Technology and trade : empirical evidences for the major five industrialized countries</td>
</tr>
<tr>
<td>G 9208</td>
<td>B. CREPON, E. DUGUET, D. ENCAOUA et P. MOHNER</td>
<td>Cooperative, non cooperative R &amp; D and optimal patent life</td>
</tr>
<tr>
<td>G 9209</td>
<td>B. CREPON et E. DUGUET</td>
<td>Research and development, competition and innovation : an application of pseudo maximum likelihood methods to Poisson models with heterogeneity</td>
</tr>
<tr>
<td>G 9301</td>
<td>J. TOUJAS-BERNATE</td>
<td>Commerce international et concurrence imparfaite : développements récents et implications pour la politique commerciale</td>
</tr>
<tr>
<td>G 9302</td>
<td>Ch. CASES</td>
<td>Durées de chômage et comportements d'offre de travail : une revue de la littérature</td>
</tr>
<tr>
<td>G 9303</td>
<td>H. EKREL-ROUSSE</td>
<td>Union économique et monétaire : le débat économique</td>
</tr>
<tr>
<td>G 9304</td>
<td>N. GREENAN, D. GUELLEC, G. BROUSAUD-LAURIER et L. MOITI</td>
<td>Innovation organisationnelle, dynamisme technologique et performances des entreprises</td>
</tr>
<tr>
<td>G 9305</td>
<td>P. JAILLARD</td>
<td>Le traité de Maastricht : présentation juridique et historique</td>
</tr>
<tr>
<td>G 9306</td>
<td>J.L. BRILLET</td>
<td>Micro-DMS : présantation et propriétés</td>
</tr>
<tr>
<td>G 9307</td>
<td>J.L. BRILLET</td>
<td>Micro-DMS - variantes : les tableaux</td>
</tr>
<tr>
<td>G 9308</td>
<td>S. JACOBZONE</td>
<td>Les grands réseaux publics français dans une perspective européenne</td>
</tr>
<tr>
<td>G 9309</td>
<td>L. BLOCH et B. CEURÉ</td>
<td>Problématique de l'investissement productif et transmission des chocs financiers</td>
</tr>
<tr>
<td>G 9310</td>
<td>J. BOURDIEU et B. COLIN-SEDILLOT</td>
<td>Les théories sur la structure optimal du capital : quelques points de repère</td>
</tr>
<tr>
<td>G 9311</td>
<td>J. BOURDIEU et B. COLIN-SEDILLOT</td>
<td>Les décisions de financement des entreprises françaises : une évaluation empirique des théories de la structure optimale du capital</td>
</tr>
</tbody>
</table>
À l'exemple de la prévision de la production manufacturière.


Le modèle MÉSANGE a été réestimé en base 2000. Ces travaux ont été réalisés par divers chercheurs, dont les travaux de M. DUÉE et P. AUBERT sont particulièrement importants.

Les chercheurs ont également étudié les économies d'agglomération et la productivité des entreprises. Par exemple, l'étude "Les rendements non monétaires de l'éducation: le cas de la santé" a été publiée en 2006.


La modélisation des comportements démographiques dans le modèle de microsimulation DÉTÉINE a été un autre sujet de recherche important. Les travaux de M. DUÉE sur "Les réactions des entreprises françaises à la baisse des tarifs douaniers étrangers" sont également significatifs.


