Direction des Études et Synthèses Économiques

G 2010 / 18

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

Romain AEBERHARDT et Laurent DAVEZIES

Document de travail

Institut National de la Statistique et des Études Économiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

Série des documents de travail
de la Direction des Études et Synthèses Économiques

G 2010 / 18

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

Romain AEBERHARDT et Laurent DAVEZIES*

DÉCEMBRE 2010

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.

* Département des Études Économiques d'Ensemble - 15, bd Gabriel Péri - BP 100 - 92244 MALAKOFF CEDEX

Département des Études Économiques d'Ensemble - Timbre G201 - 15, bd Gabriel Péri - BP 100 - 92244 MALAKOFF CEDEX - France - Tél. : 33 (1) 41 17 60 68 - Fax : 33 (1) 41 17 60 45 - CEDEX - E-mail : d3e-dg@insee.fr - Site Web Insee : http://www.insee.fr

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 régresseurs 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 régresseur 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 α_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:

$$y_{it} = \mathbb{1}_{\{\alpha_i + y_{it-1}\delta + \varepsilon_{it} > 0\}} \quad \text{where the} \quad \varepsilon_{it} \quad \text{have a logistic probability distribution}$$

This model can also be characterized by:

$$\mathbb{P}(y_{it}|y_{it-1}, \alpha_i) = \frac{e^{y_{it}(\alpha_i + y_{it-1}\delta)}}{1 + e^{\alpha_i + y_{it-1}\delta}} \tag{1}$$

In order to make the α_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:

$$\begin{aligned}\mathbb{P}(A_i|\alpha_i) &= \mathbb{P}(y_{i4}|y_{i3}=1,\alpha_i)\mathbb{P}(y_{i3}=1|y_{i2}=0,\alpha_i)\mathbb{P}(y_{i2}=0|y_{i1},\alpha_i)\mathbb{P}(y_{i1}|\alpha_i)\\ &= \frac{e^{\alpha_i}e^{y_{i4}(\delta+\alpha_i)}}{(1+e^{\alpha_i+y_{i1}\delta})(1+e^{\alpha_i})(1+e^{\alpha_i+\delta})}\mathbb{P}(y_{i1}|\alpha_i)\\ \mathbb{P}(B_i|\alpha_i) &= \frac{e^{y_{i1}\delta+\alpha_i}e^{y_{i4}\alpha_i}}{(1+e^{\alpha_i+y_{i1}\delta})(1+e^{\alpha_i+\delta})(1+e^{\alpha_i})}\mathbb{P}(y_{i1}|\alpha_i)\end{aligned}$$

hence,

$$\begin{aligned}\mathbb{P}(A_i|A_i\cup B_i,\alpha_i) &= \frac{e^{y_{i4}\delta}}{e^{y_{i1}\delta}+e^{y_{i4}\delta}}\\ \mathbb{P}(B_i|A_i\cup B_i,\alpha_i) &= \frac{e^{y_{i1}\delta}}{e^{y_{i1}\delta}+e^{y_{i4}\delta}}\end{aligned}$$

therefore, the conditional likelihood does not depend on α_i and can be written as:

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

Note that $\mathbb{P}(y_{i2},y_{i3}|y_{i1},y_{i4},y_{i2}+y_{i3}=0)$ and $\mathbb{P}(y_{i2},y_{i3}|y_{i1},y_{i4},y_{i2}+y_{i3}=2)$ do not depend on δ 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:

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

This model can also be characterized by:

$$\mathbb{P}(z_{it}=1|x_{it},\alpha_i) = \frac{e^{\alpha_i+x_{it}\delta}}{1+e^{\alpha_i+x_{it}\delta}} \tag{3}$$

Hence, with the same kind of calculations as before,

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

Note that, as in the dynamic case, $\mathbb{P}(z_{i1}, z_{i2}|z_{i1} + z_{i2} = 0)$ and $\mathbb{P}(z_{i1}, z_{i2}|z_{i1} + z_{i2} = 2)$ do not depend on δ 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

(a) Good one

t	z_i	x_i
1	y_{i2}	y_{i1}
2	y_{i3}	y_{i4}

(b) Bad one

t	z_i	x_i
1	y_{i2}	y_{i1}
2	y_{i3}	y_{i2}
3	y_{i4}	y_{i3}

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 $\sum_{t=2}^{T-1} y_{it} \notin \{0, T-2\}$ 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 $\sum_{t=1}^{T'} z_{it} \notin \{0, T'\}$ 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 $T \geq 4$:

$$\mathbb{P}(y_{i1}, ..., y_{iT} | y_{i1}, \sum_{t=2}^{T-1} y_{it}, y_{iT}) = \frac{\exp\left(\sum_{t=2}^{T} y_{it} y_{it-1} \delta\right)}{\sum_{\tilde{y} \in B(T)} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it} \tilde{y}_{it-1} \delta\right)} \tag{5}$$

where $B(T) = \left\{ (\tilde{y}_{i1}, \ldots, \tilde{y}_{iT}) | \tilde{y}_{i1} = y_{i1}, \tilde{y}_{iT} = y_{iT}, \sum_{t=2}^{T-1} \tilde{y}_{it} = \sum_{t=2}^{T-1} y_{it} \right\}$ 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 $\mathcal{O}(2^T \frac{(1-1/T)^T}{\sqrt{T-1}})$. 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 δ.

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 :

$$\mathbb{P}(y_{i1}, ..., y_{iT}|y_{i1}, \sum_{t=2}^{T-1} y_{it}, y_{iT}) =$$
$$\left[\sum_{j=\max(0, 2\sum_t y_{it} - T + 1 - y_1 - y_T)}^{\sum_t y_{it} - 1 - y_1 y_T} \binom{\sum_t y_{it} - 1}{j} \binom{T - \sum_t y_{it} - 1}{\sum_t y_{it} - j - y_1 - y_T} \exp\left\{ \left(j - \sum_{t=2}^{T} y_t y_{t-1} \right) \delta \right\} \right]^{-1}$$

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 $T \geq 2$:

$$\mathbb{P}(z_{i1}, \ldots, z_{iT'}|\sum_{t=1}^{T'} z_{it}, x_{i1}, \ldots, x_{iT'}) = \frac{\exp\left(\sum_{t=1}^{T'} z_{it} x_{it} \delta\right)}{\sum_{\tilde{z} \in B'(T')} \exp\left(\sum_{t=1}^{T'} \tilde{z}_{it} x_{it} \delta\right)} \tag{6}$$

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 :

$$\mathbb{P}(z_{i1}, \ldots, z_{iT'} | \sum_{t=1}^{T'} z_{it}, x_{i1}, \ldots, x_{iT'}) =$$
$$\left[\sum_{j=max(0, \sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T')}^{min\left(\sum_{t=1}^{T'} z_{it}, \sum_{t=1}^{T'} x_{it}\right)} \binom{\sum_{t=1}^{T'} x_{it}}{j} \binom{T' - \sum_{t=1}^{T'} x_{it}}{\sum_{t=1}^{T'} z_{it} - j} \exp\left\{ \left(j - \sum_{t=1}^{T'} x_{it} z_{it} \right) \delta \right\} \right]^{-1}$$

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 ($\sum_{t=1}^{T-2} z_{it} = \sum_{t=2}^{T-1} y_{it}$ or $\sum_{t=1}^{T-2} z_{it} = T - 2 - \sum_{t=2}^{T-1} y_{it}$)*

2. • *for y such that $0 < \sum_{t=2}^{T-1} y_{it} < T-2$ and $1 < \sum_{t=1}^{T} y_{it} < T-1$*

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

or

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

• *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 $y_{i1} = y_{iT} = 1$)*

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

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 $\sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}$ 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 $\frac{\mathbb{1}_{\{y \in B(T)\}}}{T-2}$, 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 δ 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

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

The new dataset will look like the following:

$$\left.\begin{array}{cc}
z & x \\ \hline
\left.\begin{array}{cc} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{array}\right\} & \sum_{t=2}^{T} y_t y_{t-1} \\
\left.\begin{array}{cc} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{array}\right\} & \sum_{t=1}^{T} y_t - y_1 - y_T - \sum_{t=2}^{T} y_t y_{t-1} \\
\left.\begin{array}{cc} 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{array}\right\} & \sum_{t=1}^{T} y_t - 1 - \sum_{t=2}^{T} y_t y_{t-1} \\
\left.\begin{array}{cc} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{array}\right\} & T - 1 - 2\sum_{t=1}^{T} y_t + y_1 + y_T + \sum_{t=2}^{T} y_t y_{t-1}
\end{array}\right\} T - 2$$

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

$$\left\{\begin{array}{cc|l} z & x & \\ \hline 1 & 1 & \multirow{2}{*}{$\left.\right\} \sum_{t=2}^{T} y_t y_{t-1} = 2$} \\ 1 & 1 & \\ 1 & 0 & \left.\right\} \sum_{t=1}^{T} y_t - y_1 - y_T - \sum_{t=2}^{T} y_t y_{t-1} = 1 \\ 0 & 1 & \left.\right\} \sum_{t=1}^{T} y_t - 1 - \sum_{t=2}^{T} y_t y_{t-1} = 1 \end{array}\right.$$

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

$$\left\{\begin{array}{cc|l} z & x & \\ \hline 1 & 0 & \left.\right\} \sum_{t=1}^{T} y_t - y_1 - y_T - \sum_{t=2}^{T} y_t y_{t-1} = 1 \\ 0 & 1 & \multirow{2}{*}{$\left.\right\} \sum_{t=1}^{T} y_t - 1 - \sum_{t=2}^{T} y_t y_{t-1} = 2$} \\ 0 & 1 & \\ 0 & 0 & \left.\right\} T - 1 - 2\sum_{t=1}^{T} y_t + y_1 + y_T + \sum_{t=2}^{T} y_t y_{t-1} = 1 \end{array}\right.$$

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

Andersen, E. B. (1973): *Conditional inference and models for measuring.* Mentalhygiejnisk Forlag.

Chamberlain, G. (1985): "Heterogeneity, omitted variable bias, and duration dependence," in *Longitudinal Analysis of Labor Market Data*, ed. by J. J. Heckman, and B. Singer. Cambridge University Press.

Honoré, B. E., and E. Kyriazidou (2000): "Panel Data Discrete Choice with Lagged Dependent Variables," *Econometrica*, 68(4).

Magnac, T. (2000): "Subsidised training and youth employment: distinguishing unobserved heterogeneity from state dependence in labour market histories," *The Economic Journal*, 110.

——— (2004): "Panel Binary Variables and Sufficiency: Generalizing Conditional Logit," *Econometrica*, 72(6), 1859–1876.

Rasch, G. (1960): *Probabilistic Models for Some Intelligence and Attainment Tests.* Denmark Paedagogiske Institut.

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(delta) 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])*delta))[+]);
    end;
        return(value);
    finish;

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

    /*STANDARD DEVIATION*/
    start varas(delta) 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])*delta))[+])**2
                /((comb(sumx[i,1],j)#comb(t[i,1]-sumx[i,1],sumz[i,1]-j)#exp((j-
                    sumxz[i,1])*delta))[+])**2;
        end;
```

```
        value=1/nrow(t)*value;
        return(value);
    finish;

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

    /*PRINT 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.;
```

```
        __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;
    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 `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 {__sumYLY} into sumyly;
    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,sumyly,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-sumyly[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,sumyly,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+(
                    (
```

```
                    (j-sumyly[i,1])#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-sumyly[i
                        ,1])*delta)
                )[+]
            )**2/(
                (
                    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-sumyly[i,1])*delta)
                )[+]
            )**2;
    end;
        value=1/nrow(t)*value;
        return(value);
    finish;
    ec=1/sqrt(nrow(t)*varas(delta));
    print delta ec;
close;quit;
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*Y_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 < sumY - Y.1 - Y.T < T-2
### 1 < sumY < 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* ℓ *to state* ℓ',
- k_ℓ, $\ell \in \{0,1\}$ *the number of occurrences of state* ℓ,
- 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:

$$\begin{cases} j_{00} + j_{01} + (1 - y_T) = k_0 \\ j_{10} + j_{11} + y_T = k_1 \\ j_{00} + j_{10} + (1 - y_1) = k_0 \\ j_{01} + j_{11} + y_1 = k_1 \end{cases}$$

□

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]

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

[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:*

$$\binom{k_1 - 1}{j_{11}}\binom{k_0 - 1}{j_{00}}\mathbb{1}_{\{0\leq j_{11}\leq k_1-1-y_1y_T,\quad 0\leq j_{00}\leq k_0-1-(1-y_1)(1-y_T)\}}$$

Proof. To answer this question we shall introduce the following notations: let $y = (y_1, \ldots, y_T)$ be a trajectory, then define $Red(y)$ the "reduced trajectory" of y which is obtained by concatenation of the successive zeros and ones. For instance, $Red(1, 1, 1, 0, 0, 1, 1, 0) = (1, 0, 1, 0)$. 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: $Red(1, 1, 1, 0, 0, 1, 1, 0) = ((1, 0)_2)$.

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:

$$(\underbrace{1, \ldots, 1, 0, \ldots, 0}, \underbrace{1, \ldots, 1, 0, \ldots, 0}, \ldots, \underbrace{1, \ldots, 1, 0, \ldots, 0})\ ,$$

in each block of 1's, the number of 1's is equal to 1+ "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 $Red(y) = 0\,(10)_{k_1 - j_{11}}$
- $y_1 = 1$ and $y_T = 0$ then $Red(T) = (10)_{k_1 - j_{11}}$
- $y_1 = 0$ and $y_T = 1$ then $Red(T) = (01)_{k_1 - j_{11}}$
- $y_1 = 1$ and $y_T = 1$ then $Red(T) = 1\,(01)_{k_1 - j_{11} - 1}$

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 $Red^{-1}((0,1)_{k_1-j_{11}})$ 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 $\binom{k_1-1}{k_1-j_{11}-1}\binom{k_0-1}{k_1-j_{11}-1} = \binom{k_1-1}{j_{11}}\binom{k_0-1}{k_1-j_{11}-1}$,
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:

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

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

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

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

Finally the number of trajectories is $\binom{k_1-1}{j_{11}}\binom{T-k_1-1}{k_1-j_{11}-y_1-y_T}$
with $\max(0, 2k_1-T+1-y_1-y_T) \leq j_{11} \leq k_1-1-y_1y_T$

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:*

$$B(T, k_1, y_1, y_T) = \left\{ (\tilde{y}_1, \ldots, \tilde{y}_T) \in \{0,1\}^T \,|\tilde{y}_{i1} = y_{i1}, \tilde{y}_{iT} = y_{iT}, \sum_{t=1}^{T} \tilde{y}_t = k_1 \right\}$$

Then :

$$\sum_{\tilde{y}\in B(T,k_1,y_1,y_T)} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}\delta\right) = \sum_{j_{11}=\max(0,2k_1-T+1-y_1-y_T)}^{k_1-1-y_1y_T} \binom{k_1-1}{j_{11}}\binom{T-k_1-1}{k_1-j_{11}-y_1-y_T} e^{\delta j_{11}}$$

Proof. Let $A(T,k_1,y_1,y_T,j_{11}) = B(T,k_1,y_1,y_T) \bigcap \left\{(\tilde{y}_1,\ldots,\tilde{y}_T) \in \{0,1\}^T \mid \sum_{t=2}^{T} \tilde{y}_t\tilde{y}_{t-1} = j_{11}\right\}$

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

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

and then, from lemma B.2:

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

□

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:

$$P(y_1,\ldots,y_T|y_1,y_T,\sum_{t=1}^{T} y_t) = \frac{\exp\left(\sum_{t=2}^{T} y_t y_{t-1}\delta\right)}{\sum_{j_{11}=\max(0,2k_1-T+1-y_1-y_T)}^{k_1-1-y_1y_T} \binom{k_1-1}{j_{11}}\binom{T-k_1-1}{k_1-j_{11}-y_1-y_T} e^{\delta j_{11}}}$$

Proof. direct consequence of lemma B.3. □

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 $\sum_{t=1}^{T'} \tilde{z}_{it}x_{it} = k$.

n_k is equal to 0 if $k \notin \left[\max\left(\sum_{t=1}^{T'} z_{it} + \sum_{t=1}^{T'} x_{it} - T', 0\right), \min\left(\sum_{t=1}^{T'} z_{it}, \sum_{t=1}^{T'} x_{it}\right)\right]$.
Indeed, $\sum_{t=1}^{T'} \tilde{z}_{it} x_{it} \leq \min\left(\sum_{t=1}^{T'} \tilde{z}_{it}, \sum_{t=1}^{T'} x_{it}\right)$ and

$$\begin{aligned}
\sum_{t=1}^{T'} \tilde{z}_{it} x_{it} &= \sum_{t=1}^{T'} x_{it} - \sum_{t=1}^{T'} (1 - \tilde{z}_{it}) x_{it} \\
&\geq \sum_{t=1}^{T'} x_{it} - \min\left(T' - \sum_{t=1}^{T'} \tilde{z}_{it}, \sum_{t=1}^{T'} x_{it}\right) \\
&= \max\left(\sum_{t=1}^{T'} \tilde{z}_{it} + \sum_{t=1}^{T'} x_{it} - T', 0\right)
\end{aligned}$$

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

So, in this case, $n_k = \binom{\sum_{t=1}^{T'} x_{it}}{j} \binom{T' - \sum_{t=1}^{T'} x_{it}}{\sum_{t=1}^{T'} z_{it} - j}$ and the result follows.

B.3 Proof of 3.3

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

Using lemma 3.1 and lemma 3.2, and noting:

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

$$J_{dyn} = \left[\max\left(0, 2\sum_{t=1}^{T} y_{it} - T + 1 - y_1 - y_T\right), \sum_{t=1}^{T} y_{it} - 1 - y_1 y_T\right] \cap \mathbb{N},$$

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:

$$\left\{\left(\binom{\sum_{t=1}^{T'} x_{it}}{j}\binom{T'-\sum_{t=1}^{T'} x_{it}}{\sum_{t=1}^{T'} z_{it}-j}, j-\sum_{t=1}^{T'} x_{it}z_{it}\right)\right\}_{j\in J_{sta}}$$

$$\left\{\left(\binom{\sum_{t=1}^{T} y_{it}-1}{j}\binom{T-\sum_{t=1}^{T} y_{it}-1}{\sum_{t=1}^{T} y_{it}-j-y_1-y_T}, j-\sum_{t=2}^{T} y_{it}y_{it-1}\right)\right\}_{j\in J_{dyn}}$$

First we show (1.) by summing all the terms in both sets and using Vandermonde's identity:

$$\sum_{j=\max(0,2\sum_t y_{it}-T+1-y_1-y_T)}^{\sum_{t=1}^{T} y_{it}-1-y_1y_T} \binom{\sum_{t=1}^{T} y_{it}-1}{j}\binom{T-\sum_t y_{it}-1}{\sum_{t=1}^{T} y_{it}-j-y_1-y_T} = \binom{T-2}{\sum_{t=2}^{T-1} y_{it}}$$

and

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

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 δ 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 δ 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 γ_i,

$$\mathbb{P}(U=u|V=v,\gamma_i)=\frac{e^{u(\delta v+\gamma_i)}}{1+e^{\delta v+\gamma_i}}=\frac{e^{(1-u)(\delta(1-v)-\delta-\gamma_i)}}{1+e^{\delta(1-v)-\delta-\gamma_i}}.$$

Therefore, since conditioning makes the unobserved heterogeneity disappear, for a dynamic model, the conditional likelihoods of the trajectories $y=(y_1,y_2,\dots,y_T)$ and $(1-y)=(1-y_1,1-y_2,\dots,1-y_T)$ will be the same, and likewise, for a static model, the conditional likelihoods of the trajectories $(z,x)=((z_1,x_1),\dots,(z_{T'},x_{T'}))$ and $(1-z,1-x)=((1-z_1,1-x_1),\dots,(1-z_{T'},1-x_{T'}))$ 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:

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

   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: $\sum_{t=1}^{T-2}z_{it}-\sum_{t=1}^{T-2}x_{it}z_{it}=\sum_{t=1}^{T}y_{it}-1-\sum_{t=2}^{T}y_{it}y_{it-1}$ and therefore, $\sum_{t=1}^{T-2}x_{it}z_{it}=\sum_{t=2}^{T}y_{it}y_{it-1}$.

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

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

$$2\sum_{t=1}^{T}y_{it}-T+1-y_1-y_T<0,\ =0,\ \text{and}\ >0$$

   .

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

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

$$T - 1 - \sum_{t=2}^{T-1} y_{it} = \binom{T - 1 - \sum_{t=2}^{T-1} y_{it}}{\sum_{t=2}^{T-1} y_{it}}$$

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

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

  We have $\sum_{t=1}^{T-2} z_{it} = \frac{T-1}{2}$, and the number of terms in the likelihood of the dynamic model is equal to $\frac{T-1}{2}$. The equality of the number of terms implies $\sum_{t=1}^{T-2} x_{it} + 1 \geq \frac{T-1}{2}$ and $T - 1 - \sum x_{it} \geq \frac{T-1}{2}$, and we deduce that $\sum_{t=1}^{T-2} x_{it} \in \{\frac{T-1}{2}, \frac{T-3}{2}\}$.

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

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

  We have $\sum_{t=1}^{T-2} z_{it} > T - 1 - \sum_{t=1}^{T-2} z_{it}$, the number of terms in each likelihood is

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

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

and

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

or equivalently

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

and

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

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

The necessary conditions given above are trivially sufficient using the closed forms of the conditional likelihoods.

G 9001 J. FAYOLLE et M. FLEURBAEY
Accumulation, profitabilité et endettement des entreprises

G 9002 H. ROUSSE
Détection et effets de la multicolinéarité dans les modèles linéaires ordinaires - Un prolongement de la réflexion de BELSLEY, KUH et WELSCH

G 9003 P. RALLE et J. TOUJAS-BERNATE
Indexation des salaires : la rupture de 1983

G 9004 D. GUELLEC et P. RALLE
Compétitivité, croissance et innovation de produit

G 9005 P. RALLE et J. TOUJAS-BERNATE
Les conséquences de la désindexation. Analyse dans une maquette prix-salaires

G 9101 Equipe AMADEUS
Le modèle AMADEUS - Première partie - Présentation générale

G 9102 J.L. BRILLET
Le modèle AMADEUS - Deuxième partie - Propriétés variantielles

G 9103 D. GUELLEC et P. RALLE
Endogenous growth and product innovation

G 9104 H. ROUSSE
Le modèle AMADEUS - Troisième partie - Le commerce extérieur et l'environnement international

G 9105 H. ROUSSE
Effets de demande et d'offre dans les résultats du commerce extérieur manufacturé de la France au cours des deux dernières décennies

G 9106 B. CREPON
Innovation, taille et concentration : causalités et dynamiques

G 9107 B. AMABLE et D. GUELLEC
Un panorama des théories de la croissance endogène

G 9108 M. GLAUDE et M. MOUTARDIER
Une évaluation du coût direct de l'enfant de 1979 à 1989

G 9109 P. RALLE et alii
France - Allemagne : performances économiques comparées

G 9110 J.L. BRILLET
Micro-DMS ***NON PARU***

G 9111 A. MAGNIER
Effets accélérateur et multiplicateur en France depuis 1970 : quelques résultats empiriques

G 9112 B. CREPON et G. DUREAU
Investissement en recherche-développement : analyse de causalités dans un modèle d'accélérateur généralisé

G 9113 J.L. BRILLET, H. ERKEL-ROUSSE, J. TOUJAS-BERNATE
"France-Allemagne Couplées" - Deux économies vues par une maquette macro-économétrique

G 9201 W.J. ADAMS, B. CREPON, D. ENCAOUA
Choix technologiques et stratégies de dissuasion d'entrée

G 9202 J. OLIVEIRA-MARTINS, J. TOUJAS-BERNATE
Macro-economic import functions with imperfect competition - An application to the E.C. Trade

G 9203 I. STAPIC
Les échanges internationaux de services de la France dans le cadre des négociations multilatérales du GATT
Juin 1992 (1ère version)
Novembre 1992 (version finale)

G 9204 P. SEVESTRE
L'économétrie sur données individuelles-temporelles. Une note introductive

G 9205 H. ERKEL-ROUSSE
Le commerce extérieur et l'environnement international dans le modèle AMADEUS (réestimation 1992)

G 9206 N. GREENAN et D. GUELLEC
Coordination within the firm and endogenous growth

G 9207 A. MAGNIER et J. TOUJAS-BERNATE
Technology and trade : empirical evidences for the major five industrialized countries

G 9208 B. CREPON, E. DUGUET, D. ENCAOUA et P. MOHNEN
Cooperative, non cooperative R & D and optimal patent life

G 9209 B. CREPON et E. DUGUET
Research and development, competition and innovation : an application of pseudo maximum likelihood methods to Poisson models with heterogeneity

G 9301 J. TOUJAS-BERNATE
Commerce international et concurrence imparfaite : développements récents et implications pour la politique commerciale

G 9302 Ch. CASES
Durées de chômage et comportements d'offre de travail : une revue de la littérature

G 9303 H. ERKEL-ROUSSE
Union économique et monétaire : le débat économique

G 9304 N. GREENAN - D. GUELLEC / G. BROUSSAUDIER - L. MIOTTI
Innovation organisationnelle, dynamisme technologique et performances des entreprises

G 9305 P. JAILLARD
Le traité de Maastricht : présentation juridique et historique

G 9306 J.L. BRILLET
Micro-DMS : présentation et propriétés

G 9307 J.L. BRILLET
Micro-DMS - variantes : les tableaux

G 9308 S. JACOBZONE
Les grands réseaux publics français dans une perspective européenne

G 9309 L. BLOCH - B. CŒURE
Profitabilité de l'investissement productif et transmission des chocs financiers

G 9310 J. BOURDIEU - B. COLIN-SEDILLOT
Les théories sur la structure optimal du capital : quelques points de repère

G 9311 J. BOURDIEU - B. COLIN-SEDILLOT
Les décisions de financement des entreprises françaises : une évaluation empirique des théories de la structure optimale du capital

G 9312 L. BLOCH - B. CŒURÉ
Q de Tobin marginal et transmission des chocs financiers

G 9313 Equipes Amadeus (INSEE), Banque de France, Métric (DP)
Présentation des propriétés des principaux modèles macroéconomiques du Service Public

G 9314 B. CREPON - E. DUGUET
Research & Development, competition and innovation

G 9315 B. DORMONT
Quelle est l'influence du coût du travail sur l'emploi ?

G 9316 D. BLANCHET - C. BROUSSE
Deux études sur l'âge de la retraite

G 9317 D. BLANCHET
Répartition du travail dans une population hétérogène : deux notes

G 9318 D. EYSSARTIER - N. PONTY
AMADEUS - an annual macro-economic model for the medium and long term

G 9319 G. CETTE - Ph. CUNÉO - D. EYSSARTIER - J. GAUTIÉ
Les effets sur l'emploi d'un abaissement du coût du travail des jeunes

G 9401 D. BLANCHET
Les structures par âge importent-elles ?

G 9402 J. GAUTIÉ
Le chômage des jeunes en France : problème de formation ou phénomène de file d'attente ? Quelques éléments du débat

G 9403 P. QUIRION
Les déchets en France : éléments statistiques et économiques

G 9404 D. LADIRAY - M. GRUN-REHOMME
Lissage par moyennes mobiles - Le problème des extrémités de série

G 9405 V. MAILLARD
Théorie et pratique de la correction des effets de jours ouvrables

G 9406 F. ROSENWALD
La décision d'investir

G 9407 S. JACOBZONE
Les apports de l'économie industrielle pour définir la stratégie économique de l'hôpital public

G 9408 L. BLOCH, J. BOURDIEU, B. COLIN-SEDILLOT, G. LONGUEVILLE
Du défaut de paiement au dépôt de bilan : les banquiers face aux PME en difficulté

G 9409 D. EYSSARTIER, P. MAIRE
Impacts macro-économiques de mesures d'aide au logement - quelques éléments d'évaluation

G 9410 F. ROSENWALD
Suivi conjoncturel de l'investissement

G 9411 C. DEFEUILLEY - Ph. QUIRION
Les déchets d'emballages ménagers : une analyse économique des politiques française et allemande

G 9412 J. BOURDIEU - B. CŒURÉ - B. COLIN-SEDILLOT
Investissement, incertitude et irréversibilité
Quelques développements récents de la théorie de l'investissement

G 9413 B. DORMONT - M. PAUCHET
L'évaluation de l'élasticité emploi-salaire dépend-elle des structures de qualification ?

G 9414 I. KABLA
Le Choix de breveter une invention

G 9501 J. BOURDIEU - B. CŒURÉ - B. SEDILLOT
Irreversible Investment and Uncertainty : When is there a Value of Waiting ?

G 9502 L. BLOCH - B. CŒURÉ
Imperfections du marché du crédit, investissement des entreprises et cycle économique

G 9503 D. GOUX - E. MAURIN
Les transformations de la demande de travail par qualification en France
Une étude sur la période 1970-1993

G 9504 N. GREENAN
Technologie, changement organisationnel, qualifications et emploi : une étude empirique sur l'industrie manufacturière

G 9505 D. GOUX - E. MAURIN
Persistance des hiérarchies sectorielles de salaires: un réexamen sur données françaises

G 9505 Bis D. GOUX - E. MAURIN
Persistence of inter-industry wages differentials: a reexamination on matched worker-firm panel data

G 9506 S. JACOBZONE
Les liens entre RMI et chômage, une mise en perspective
NON PARU - article sorti dans Economie et Prévision n°122 (1996) - pages 95 à 113

G 9507 G. CETTE - S. MAHFOUZ
Le partage primaire du revenu
Constat descriptif sur longue période

G 9601 Banque de France - CEPREMAP - Direction de la Prévision - Erasme - INSEE - OFCE
Structures et propriétés de cinq modèles macro-économiques français

G 9602 Rapport d'activité de la DESE de l'année 1995

G 9603 J. BOURDIEU - A. DRAZNIEKS
L'octroi de crédit aux PME : une analyse à partir d'informations bancaires

G 9604 A. TOPIOL-BENSAÏD
Les implantations japonaises en France

G 9605 P. GENIER - S. JACOBZONE
Comportements de prévention, consommation d'alcool et tabagie : peut-on parler d'une gestion globale du capital santé ?
Une modélisation microéconométrique empirique

G 9606 C. DOZ - F. LENGLART
Factor analysis and unobserved component models: an application to the study of French business surveys

G 9607 N. GREENAN - D. GUELLEC
La théorie coopérative de la firme

- Computerization in France: an evaluation based on individual company data

G2001/01 F. LEQUILLER
- La nouvelle économie et la mesure de la croissance du PIB
- The new economy and the measure ment of GDP growth

G2001/02 S. AUDRIC
La reprise de la croissance de l'emploi profite-t-elle aussi aux non-diplômés ?

G2001/03 I. BRAUN-LEMAIRE
Evolution et répartition du surplus de productivité

G2001/04 A. BEAUDU - Th. HECKEL
Le canal du crédit fonctionne-t-il en Europe ? Une étude de l'hétérogénéité des comportements d'investissement à partir de données de bilan agrégées

G2001/05 C. AUDENIS - P. BISCOURP - N. FOURCADE - O. LOISEL
Testing the augmented Solow growth model : An empirical reassessment using panel data

G2001/06 R. MAHIEU - B. SÉDILLOT
Départ à la retraite, irréversibilité et incertitude

G2001/07 Bilan des activités de la DESE - 2000

G2001/08 J. Ph. GAUDEMET
Les dispositifs d'acquisition à titre facultatif d'annuités viagères de retraite

G2001/09 B. CRÉPON - Ch. GIANELLA
Fiscalité, coût d'usage du capital et demande de facteurs : une analyse sur données individuelles

G2001/10 B. CRÉPON - R. DESPLATZ
Evaluation des effets des dispositifs d'allégements de charges sociales sur les bas salaires

G2001/11 J.-Y. FOURNIER
Comparaison des salaires des secteurs public et privé

G2001/12 J.-P. BERTHIER - C. JAULENT R. CONVENEVOLE - S. PISANI
Une méthodologie de comparaison entre consommations intermédiaires de source fiscale et de comptabilité nationale

G2001/13 P. BISCOURP - Ch. GIANELLA
Substitution and complementarity between capital, skilled and less skilled workers: an analysis at the firm level in the French manufacturing industry

G2001/14 I. ROBERT-BOBEE
Modelling demographic behaviours in the French microsimulation model Destinie: An analysis of future change in completed fertility

G2001/15 J.-P. ZOYEM
Diagnostic sur la pauvreté et calendrier de revenus : le cas du "Panel européen des ménages »

G2001/16 J.-Y. FOURNIER - P. GIVORD
La réduction des taux d'activité aux âges extrêmes, une spécificité française ?

G2001/17 C. AUDENIS - P. BISCOURP - N. RIEDINGER
Existe-t-il une asymétrie dans la transmission du prix du brut aux prix des carburants ?

G2002/01 F. MAGNIEN - J.-L. TAVERNIER - D. THESMAR
Les statistiques internationales de PIB par habitant en standard de pouvoir d'achat : une analyse des résultats

G2002/02 Bilan des activités de la DESE - 2001

G2002/03 B. SÉDILLOT - E. WALRAET
La cessation d'activité au sein des couples : y a-t-il interdépendance des choix ?

G2002/04 G. BRILHAULT
- Rétropolation des séries de FBCF et calcul du capital fixe en SEC-95 dans les comptes nationaux français
- Retropolation of the investment series (GFCF) and estimation of fixed capital stocks on the ESA-95 basis for the French balance sheets

G2002/05 P. BISCOURP - B. CRÉPON - T. HECKEL - N. RIEDINGER
How do firms respond to cheaper computers? Microeconometric evidence for France based on a production function approach

G2002/06 C. AUDENIS - J. DEROYON - N. FOURCADE
L'impact des nouvelles technologies de l'information et de la communication sur l'économie française - un bouclage macro-économique

G2002/07 J. BARDAJI - B. SÉDILLOT - E. WALRAET
Évaluation de trois réformes du Régime Général d'assurance vieillesse à l'aide du modèle de microsimulation DESTINIE

G2002/08 J.-P. BERTHIER
Réflexions sur les différentes notions de volume dans les comptes nationaux : comptes aux prix d'une année fixe ou aux prix de l'année précédente, séries chaînées

G2002/09 F. HILD
Les soldes d'opinion résument-ils au mieux les réponses des entreprises aux enquêtes de conjoncture ?

G2002/10 I. ROBERT-BOBÉE
Les comportements démographiques dans le modèle de microsimulation Destinie - Une comparaison des estimations issues des enquêtes Jeunes et Carrières 1997 et Histoire Familiale 1999

G2002/11 J.-P. ZOYEM
La dynamique des bas revenus : une analyse des entrées-sorties de pauvreté

G2002/12 F. HILD
Prévisions d'inflation pour la France

G2002/13 M. LECLAIR
Réduction du temps de travail et tensions sur les facteurs de production

G2002/14 E. WALRAET - A. VINCENT
- Analyse de la redistribution intragénérationnelle dans le système de retraite des salariés du privé - Une approche par microsimulation
- Intragenerational distributional analysis in the french private sector pension scheme - A microsimulation approach

G2002/15 P. CHONE - D. LE BLANC - I. ROBERT-BOBEE
Offre de travail féminine et garde des jeunes enfants

G2002/16 F. MAUREL - S. GREGOIR
Les indices de compétitivité des pays : interprétation et limites

G2003/01 N. RIEDINGER - E.HAUVY
Le coût de dépollution atmosphérique pour les entreprises françaises : Une estimation à partir de données individuelles

G2003/02 P. BISCOURP et F. KRAMARZ
Création d'emplois, destruction d'emplois et internationalisation des entreprises industrielles françaises : une analyse sur la période 1986-1992

G2003/03 Bilan des activités de la DESE - 2002

G2003/04 P.-O. BEFFY - J. DEROYON - N. FOURCADE - S. GREGOIR - N. LAÏB - B. MONFORT
Évolutions démographiques et croissance : une projection macro-économique à l'horizon 2020

G2003/05 P. AUBERT
La situation des salariés de plus de cinquante ans dans le secteur privé

G2003/06 P. AUBERT - B. CRÉPON
Age, salaire et productivité
La productivité des salariés décline-t-elle en fin de carrière ?

G2003/07 H. BARON - P.O. BEFFY - N. FOURCADE - R. MAHIEU
Le ralentissement de la productivité du travail au cours des années 1990

G2003/08 P.-O. BEFFY - B. MONFORT
Patrimoine des ménages, dynamique d'allocation et comportement de consommation

G2003/09 P. BISCOURP - N. FOURCADE
Peut-on mettre en évidence l'existence de rigidités à la baisse des salaires à partir de données individulles ? Le cas de la France à la fin des années 90

G2003/10 M. LECLAIR - P. PETIT
Présence syndicale dans les firmes : quel impact sur les inégalités salariales entre les hommes et les femmes ?

G2003/11 P.-O. BEFFY - X. BONNET - M. DARRACQ-PARIES - B. MONFORT
MZE: a small macro-model for the euro area

G2004/01 P. AUBERT - M. LECLAIR
La compétitivité exprimée dans les enquêtes trimestrielles sur la situation et les perspectives dans l'industrie

G2004/02 M. DUÉE - C. REBILLARD
La dépendance des personnes âgées : une projection à long terme

G2004/03 S. RASPILLER - N. RIEDINGER
Régulation environnementale et choix de localisation des groupes français

G2004/04 A. NABOULET - S. RASPILLER
Les déterminants de la décision d'investir : une approche par les perceptions subjectives des firmes

G2004/05 N. RAGACHE
La déclaration des enfants par les couples non mariés est-elle fiscalement optimale ?

G2004/06 M. DUÉE
L'impact du chômage des parents sur le devenir scolaire des enfants

G2004/07 P. AUBERT - E. CAROLI - M. ROGER
New Techlologies, Workplace Organisation and the Age Structure of the Workforce: Firm-Level Evidence

G2004/08 E. DUGUET - C. LELARGE
Les brevets accroissent-ils les incitations privées à innover ? Un examen microéconométrique

G2004/09 S. RASPILLER - P. SILLARD
Affiliating versus Subcontracting: the Case of Multinationals

G2004/10 J. BOISSINOT - C. L'ANGEVIN - B. MONFORT
Public Debt Sustainability: Some Results on the French Case

G2004/11 S. ANANIAN - P. AUBERT
Travailleurs âgés, nouvelles technologies et changements organisationnels : un réexamen à partir de l'enquête « REPONSE »

G2004/12 X. BONNET - H. PONCET
Structures de revenus et propensions différentes à consommer - Vers une équation de consommation des ménages plus robuste en prévision pour la France

G2004/13 C. PICART
Évaluer la rentabilité des sociétés non financières

G2004/14 J. BARDAJI - B. SÉDILLOT - E. WALRAET
Les retraites du secteur public : projections à l'horizon 2040 à l'aide du modèle de microsimulation DESTINIE

G2005/01 S. BUFFETEAU - P. GODEFROY
Conditions de départ en retraite selon l'âge de fin d'études : analyse prospective pour les générations 1945 à1974

G2005/02 C. AFSA - S. BUFFETEAU
L'évolution de l'activité féminine en France : une approche par pseudo-panel

G2005/03 P. AUBERT - P. SILLARD
Délocalisations et réductions d'effectifs dans l'industrie française

G2005/04 M. LECLAIR - S. ROUX
Mesure et utilisation des emplois instables dans les entreprises

G2005/05 C. L'ANGEVIN - S. SERRAVALLE
Performances à l'exportation de la France et de l'Allemagne - Une analyse par secteur et destination géographique

G2005/06 Bilan des activités de la Direction des Études et Synthèses Économiques - 2004

G2005/07 S. RASPILLER
La concurrence fiscale : principaux enseignements de l'analyse économique

G2005/08 C. L'ANGEVIN - N. LAÏB
Éducation et croissance en France et dans un panel de 21 pays de l'OCDE

G2005/09 N. FERRARI
Prévoir l'investissement des entreprises
Un indicateur des révisions dans l'enquête de conjoncture sur les investissements dans l'industrie.

G2005/10 P.-O. BEFFY - C. L'ANGEVIN
Chômage et boucle prix-salaires :
apport d'un modèle « qualiifés/peu qualifiés »

G2005/11 B. HEITZ
A two-states Markov-switching model of inflation in France and the USA: credible target VS inflation spiral

G2005/12 O. BIAU - H. ERKEL-ROUSSE - N. FERRARI
Réponses individuelles aux enquêtes de conjoncture et prévision macroéconomiques :
Exemple de la prévision de la production manufacturière

G2005/13 P. AUBERT - D. BLANCHET - D. BLAU
The labour market after age 50: some elements of a Franco-American comparison

G2005/14 D. BLANCHET - T. DEBRAND -
P. DOURGNON - P. POLLET
L'enquête SHARE : présentation et premiers résultats de l'édition française

G2005/15 M. DUÉE
La modélisation des comportements démographiques dans le modèle de microsimulation DESTINIE

G2005/16 H. RAOUI - S. ROUX
Étude de simulation sur la participation versée aux salariés par les entreprises

G2006/01 C. BONNET - S. BUFFETEAU - P. GODEFROY
Disparités de retraite de droit direct entre hommes et femmes : quelles évolutions ?

G2006/02 C. PICART
Les gazelles en France

G2006/03 P. AUBERT - B. CRÉPON -P. ZAMORA
Le rendement apparent de la formation continue dans les entreprises : effets sur la productivité et les salaires

G2006/04 J.-F. OUVRARD - R. RATHELOT
Demographic change and unemployment:
what do macroeconometric models predict?

G2006/05 D. BLANCHET - J.-F. OUVRARD
Indicateurs d'engagements implicites des systèmes de retraite : chiffrages, propriétés analytiques et réactions à des chocs démographiques types

G2006/06 G. BIAU - O. BIAU - L. ROUVIERE
Nonparametric Forecasting of the Manufacturing Output Growth with Firm-level Survey Data

G2006/07 C. AFSA - P. GIVORD
Le rôle des conditions de travail dans les absences pour maladie

G2006/08 P. SILLARD - C. L'ANGEVIN - S. SERRAVALLE
Performances comparées à l'exportation de la France et de ses principaux partenaires
Une analyse structurelle sur 12 ans

G2006/09 X. BOUTIN - S. QUANTIN
Une méthodologie d'évaluation comptable du coût du capital des entreprises françaises : 1984-2002

G2006/10 C. AFSA
L'estimation d'un coût implicite de la pénibilité du travail chez les travailleurs âgés

G2006/11 C. LELARGE
Les entreprises (industrielles) françaises sont-elles à la frontière technologique ?

G2006/12 O. BIAU - N. FERRARI
Théorie de l'opinion
Faut-il pondérer les réponses individuelles ?

G2006/13 A. KOUBI - S. ROUX
Une réinterprétation de la relation entre productivité et inégalités salariales dans les entreprises

G2006/14 R. RATHELOT - P. SILLARD
The impact of local taxes on plants location decision

G2006/15 L. GONZALEZ - C. PICART
Diversification, recentrage et poids des activités de support dans les groupes (1993-2000)

G2007/01 D. SRAER
Allègements de cotisations patronales et dynamique salariale

G2007/02 V. ALBOUY - L. LEQUIEN
Les rendements non monétaires de l'éducation : le cas de la santé

G2007/03 D. BLANCHET - T. DEBRAND
Aspiration à la retraite, santé et satisfaction au travail : une comparaison européenne

G2007/04 M. BARLET - L. CRUSSON
Quel impact des variations du prix du pétrole sur la croissance française ?

G2007/05 C. PICART
Flux d'emploi et de main-d'œuvre en France : un réexamen

G2007/06 V. ALBOUY - C. TAVAN
Massification et démocratisation de l'enseignement supérieur en France

G2007/07 T. LE BARBANCHON
The Changing response to oil price shocks in France : a DSGE type approach

G2007/08 T. CHANEY - D. SRAER - D. THESMAR
Collateral Value and Corporate Investment
Evidence from the French Real Estate Market

G2007/09 J. BOISSINOT
Consumption over the Life Cycle: Facts for France

G2007/10 C. AFSA
Interpréter les variables de satisfaction : l'exemple de la durée du travail

G2007/11 R. RATHELOT - P. SILLARD
Zones Franches Urbaines : quels effets sur l'emploi salarié et les créations d'établissements ?

G2007/12 V. ALBOUY - B. CRÉPON
Aléa moral en santé : une évaluation dans le cadre du modèle causal de Rubin

G2008/01 C. PICART
Les PME françaises : rentables mais peu dynamiques

G2008/02 P. BISCOURP - X. BOUTIN - T. VERGÉ
The Effects of Retail Regulations on Prices
Evidence form the Loi Galland

G2008/03 Y. BARBESOL - A. BRIANT
Économies d'agglomération et productivité des entreprises : estimation sur données individuelles françaises

G2008/04 D. BLANCHET - F. LE GALLO
Les projections démographiques : principaux mécanismes et retour sur l'expérience française

G2008/05 D. BLANCHET - F. TOUTLEMONDE
Évolutions démographiques et déformation du cycle de vie active : quelles relations ?

G2008/06 M. BARLET - D. BLANCHET - L. CRUSSON
Internationalisation et flux d'emplois : que dit une approche comptable ?

G2008/07 C. LELARGE - D. SRAER - D. THESMAR
Entrepreurship and Credit Constraints - Evidence from a French Loan Guarantee Program

G2008/08 X. BOUTIN - L. JANIN
Are Prices Really Affected by Mergers?

G2008/09 M. BARLET - A. BRIANT - L. CRUSSON
Concentration géographique dans l'industrie manufacturière et dans les services en France : une approche par un indicateur en continu

G2008/10 M. BEFFY - É. COUDIN - R. RATHELOT
Who is confronted to insecure labor market histories? Some evidence based on the French labor market transition

G2008/11 M. ROGER - E. WALRAET
Social Security and Well-Being of the Elderly: the Case of France

G2008/12 C. AFSA
Analyser les composantes du bien-être et de son évolution
Une approche empirique sur données individuelles

G2008/13 M. BARLET - D. BLANCHET -
T. LE BARBANCHON
Microsimuler le marché du travail : un prototype

G2009/01 P.-A. PIONNIER
Le partage de la valeur ajoutée en France, 1949-2007

G2009/02 Laurent CLAVEL - Christelle MINODIER
A Monthly Indicator of the French Business Climate

G2009/03 H. ERKEL-ROUSSE - C. MINODIER
Do Business Tendency Surveys in Industry and Services Help in Forecasting GDP Growth?
A Real-Time Analysis on French Data

G2009/04 P. GIVORD - L. WILNER
Les contrats temporaires : trappe ou marchepied vers l'emploi stable ?

G2009/05 G. LALANNE - P.-A. PIONNIER - O. SIMON
Le partage des fruits de la croissance de 1950 à 2008 : une approche par les comptes de surplus

G2009/06 L. DAVEZIES - X. D'HAULTFOEUILLE
Faut-il pondérer ?… Ou l'éternelle question de l'économètre confronté à des données d'enquête

G2009/07 S. QUANTIN - S. RASPILLER - S. SERRAVALLE
Commerce intragroupe, fiscalité et prix de transferts : une analyse sur données françaises

G2009/08 M. CLERC - V. MARCUS
Élasticités-prix des consommations énergétiques des ménages

G2009/09 G. LALANNE - E. POULIQUEN - O. SIMON
Prix du pétrole et croisssance potentielle à long terme

G2009/10 D. BLANCHET - J. LE CACHEUX - V. MARCUS
Adjusted net savings and other approaches to sustainability: some therotical backgroud

G2009/11 V. BELLAMY - G. CONSALES - M. FESSEAU - S. LE LAIDIER - É. RAYNAUD
Une décomposition du compte des ménages de la comptabilité nationale par catégorie de ménage en 2003

G2009/12 J. BARDAJI - F. TALLET
Detecting Economic Regimes in France: a Qualitative Markov-Switching Indicator Using Mixed Frequency Data

G2009/13 R. AEBERHARDT - D. FOUGÈRE - R. RATHELOT
Discrimination à l'embauche : comment exploiter les procédures de *testing* ?

G2009/14 Y. BARBESOL - P. GIVORD - S. QUANTIN
Partage de la valeur ajoutée, approche par données microéconomiques

G2009/15 I. BUONO - G. LALANNE
The Effect of the Uruguay round on the Intensive and Extensive Margins of Trade

G2010/01 C. MINODIER
Avantages comparés des séries des premières valeurs publiées et des séries des valeurs révisées - Un exercice de prévision en temps réel de la croissance trimestrielle du PIB en France

G2010/02 V. ALBOUY - L. DAVEZIES - T. DEBRAND
Health Expenditure Models: a Comparison of Five Specifications using Panel Data

G2010/03 C. KLEIN - O. SIMON
Le modèle MÉSANGE réestimé en base 2000
Tome 1 – Version avec volumes à prix constants

G2010/04 M.-É. CLERC - É. COUDIN
L'IPC, miroir de l'évolution du coût de la vie en France ? Ce qu'apporte l'analyse des courbes d'Engel

G2010/05 N. CECI-RENAUD - P.-A. CHEVALIER
Les seuils de 10, 20 et 50 salariés : impact sur la taille des entreprises françaises

G2010/06 R. AEBERHARDT - J. POUGET
National Origin Differences in Wages and Hierarchical Positions - Evidence on French Full-Time Male Workers from a matched Employer-Employee Dataset

G2010/07 S. BLASCO - P. GIVORD
Les trajectoires professionnelles en début de vie active : quel impact des contrats temporaires ?

G2010/08 P. GIVORD
Méthodes économétriques pour l'évaluation de politiques publiques

G2010/09 P.-Y. CABANNES - V. LAPÈGUE -
E. POULIQUEN - M. BEFFY - M. GAINI
Quelle croissance de moyen terme après la crise ?

G2010/10 I. BUONO - G. LALANNE
La réaction des entreprises françaises
à la baisse des tarifs douaniers étrangers