Novel Approaches to Coherency Conditions in Dynamic LDV Models:
Quantifying Financing Constraints and a Firm's Decision and Ability to Innovate

by

Vassilis Hajivassiliou[1,*], Department of Economics, LSE and FMG

and

Frédérique Savignac*, Banque de France

May 2019

Abstract

We develop two novel methods for establishing coherency conditions in Dynamic Limited Dependent Variables (LDV) Models, which have intuitive interpretations and are easy to implement and generalize. These methods lead to estimation strategies based on Conditional Maximum Likelihood Estimation (CMLE) for simultaneous LDV models without imposing recursivity. They also allow us to establish for the first time the coherency of several Dynamic LDV models. We develop a set of extensive Monte-Carlo experiments, which confirm substantive Mean-Squared-Error improvements of the CMLE approach over estimators that make overly restrictive coherency assumptions.

We apply our framework to analyse the existence and impact of financing constraints as a possibly serious obstacle to innovation by firms, with striking results: ceteris paribus, a firm facing binding finance constraints is approximately 30% less likely to undertake innovation, while the probability that a firm encounters a binding finance constraint more than doubles if the firm is classified as innovative. Finally, we establish a strong role for state dependence in dynamic versions of our models.

Keywords: Financing Constraints; Innovation; Dynamic Limited Dependent Variable Models; Joint Bivariate Probit Model; Econometric Coherency Conditions; State Dependence.

JEL Classifications: C51, C52, C15

*We have benefitted from constructive comments by participants at seminars at the London School of Economics, Cemmap/UCL, CERGE, Prague, Nuffield College, Oxford University, City University, University of Marseilles, University of Milan, and the University of Toulouse, as well as at the 2006 Conference on Panel Data at the University of Cambridge, the 2006 Corporate Finance Conference in London, and the IGIER 2008 Conference in Capri. We thank in particular Richard Blundell, Christian Bontemps, Andrew Chesher, Francesca Cornelli, Antoine Faure-Grimaud, Denis Gromb, Tatiana Komarova, Guy Laroque, Claire Lelarge, Thierry Magnac, Alex Michaelides, Peter Robinson, John Sutton, Elie Tamer, Alain Trognon, Dimitri Vayanos, Paolo Volpin, and Frank Windjmeier. We also thank Ryan Giordano for expert research assistance. The usual disclaimer applies for all remaining errors.

[1] Corresponding author.
eml: vassilis@econ.lse.ac.uk
Department of Economics and Financial Markets Group, London School of Economics
London WC2A 2AE, England

1 Introduction

In this paper, we investigate the fundamental identification issue of *coherency* of Limited Dependent Variable (LDV) models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. We develop two novel methods for establishing coherency, which have intuitive interpretations and are easy to implement and generalize. These methods lead to estimation strategies based on Conditional Maximum Likelihood Estimation (CMLE) for simultaneous LDV models without imposing recursivity. They also allow us to establish the coherency of several Dynamic LDV models that until now, it was imposible to determine whether they were coherent or incoherent using traditional methods. [Section 2]

SUMMARIZE: The first difficulty is that the existing derivations of formal conditions lack intuition, are difficult to generalize, and are sufficient but not necessary. The second one is that in practice, non-triangular or reverse triangular cases are the most interesting. To overcome the first difficulty, alternative ways for establishing coherency were developed in [Hajivassiliou, 2008], that are both intuitive and straightforward. In addition, this novel approach yields less strict conditions that in fact allow more interesting practical applications. Specifically, it is shown in the next Section how to establish coherency without recursiveness through the use of sign restrictions on model parameters. The fact that our novel approach for the first time eliminates the need to assume recursivity is quite important: recursivity corresponds to the key identifying assumption that innovation does not affect financial distress directly ($\delta = 0$). On a priori grounds, this assumption seems particularly dubious since innovation may lead to more profits and thus relax financial constraints (corresponding to $\delta > 0$). An alternative possibility is that innovation may lead to higher investment in intangible assets thus reinforcing binding financial constraints (corresponding to $\delta < 0$). Both possibilities violate the traditional coherency condition.

SUMMARIZE: The traditional approaches to model coherency suffer from several major difficulties. Firstly, derivations of formal conditions using the traditional approach lack intuition. Secondly, the derived conditions are impossible to generalize and verify in moderately more complicated LDV models, especially in cases where the models are allowed to contain intertemporal endogeneity of the type considered in [Falcetti and Tudela, 2008]. Similarly, in case the joint binary probit model (3)-(4) is extended intertemporally, as for example in the empirical dynamic application in Section 3.2, the coherency condition is impossible to generalize and verify using the traditional analysis of the previous subsection. Thirdly, in practice non-triangular or reverse triangular cases are the most interesting from an economic point of view.[2] Finally, the traditional approaches focus on establishing *sufficient* conditions for coherency, while our methods allow us to prove that they are *not necessary*. To overcome the first two difficulties, alternative ways for establishing coherency are developed here, that are both intuitive and straightforward, as well as

[2] Exceptions of this do exist: for example in industrial organization a two-agent discrete game may be employed to model the strategic interactions between firms in a duopoly setup. If one firm is a Stackelberg leader while the other is a follower, a recursive model may be applicable, even though the analogy is not precise.

much more generalizable. In addition, our methods allow us to resolve the last two difficulties leading to estimation based on CMLE for much more interesting practical applications. It is shown in the next Section how to establish coherency without model recursiveness through the use of (a) endogeneity in terms of latent variables and/or (b) sign restrictions on model parameters. The fact that our novel approach for the first time eliminates the need to assume model recursivity is quite important for the economic problem studied in the empirical application in Section 3.2. As we explain there, recursivity corresponds to the key identifying assumption that firm innovation does not affect financial distress directly ($\delta = 0$). On a priori grounds, this assumption seems particularly dubious since firm innovation may lead to more profits and thus relax financial constraints (corresponding to $\delta > 0$). An alternative possibility is that innovation may lead to higher investment in intangible assets thus reinforcing binding financial constraints (corresponding to $\delta < 0$). Both possibilities violate the traditional coherency condition. [Section 2.5]

EXPLAIN MISTAKE OF REFEREE SINCE SETUP IS NOT A DYNAMIC GAME

USING METHODS FOR COHERENCY OF DYNAMIC MODELS Panel LDV Models with Intertemporal Endogeneities using DGP Approach

Dynamic Model 1: Univariate Panel Data Probit with State Dependence

Dynamic Model 2: Bivariate Panel Data Probit with State Dependence

[Section 2.8]

MCs:

We also develop and summarize the results of a set of extensive Monte-Carlo experiments, which confirm very substantive Mean-Squared-Error estimation improvements of the CMLE approach over estimators that make overly restrictive coherency assumptions about the Data Generating Process (DGP).

Technical Appendix 2 describes how the CMLE methodology we develop is evaluated through an extensive set of Monte-Carlo experiments. The key results are summarized in Appendix 2:5. These experiments also allow us to study the consequences of employing estimators that make overly restrictive coherency assumptions about the DGP. The findings confirm very substantive improvements by employing the CMLE developed in this paper in terms of estimation Mean-Squared-Error.

We apply our framework to analyse the existence and impact of financing constraints as a possibly serious obstacle to innovation by firms. [Section 3]These results are quite striking: ceteris paribus, we estimate that a firm that faces a binding finance constraint is approximately 30% less likely to undertake innovation, while the probability that a firm encounters a binding finance constraint more than doubles if the firm is classified as innovative. Finally, we establish a strong role for state dependence in dynamic versions of our models. Applying our CML estimation methodology, we analyze the existence and impact of financing constraints as a possibly debilitating obstacle to innovation by firms. The approach here is innovative in several respects. Firstly, unlike in earlier literature, direct measures of binding constraints are employed, instead of using traditional indirect proxy variables like firm wealth, accumulated profits, etc. We construct a direct indicator based on firms' own assessments reported in a survey (FIT, Financement de l'Innovation Technologique) run by the French Ministry of Industry in order to know more about the financing of innovation at the firm level. This database includes three items used to define a firm as credit constrained: (i) unavailability of new financing; (ii) searching and waiting for new financing; and (iii) too high cost of new financing.

Since direct indicators may be viewed as too subjective, we start by comparing the balance sheet structure of constrained and unconstrained firms. Innovation surveys are merged with the "Centrale de Bilans" (CdB) of the Banque de France, which contains full balance sheet data on the firms in the sample thus giving us direct data on bank loans, tangible and intangible assets, as well as sources of finance. We find our direct indicator coherent with the financial and economic performances reflected by balance sheet ratios: firms without financial constraints exhibit a better profile than constrained firms in terms of financing structure, risk and economic performances.

Thirdly, a key question we pose in this paper is whether binding financing constraints have a seriously adverse impact on innovation by firms. We estimate the impact of financial constraints on the probability to be innovative with an econometric approach that allows explicitly the existence of binding financing constraints to be endogenously determined. Our main finding is that there is a very significantly negative effect on innovation due to the presence of financing constraints, ceteris paribus. We also show that ignoring the endogeneity of the constraint indicator together with the endogenous decision of whether a firm wishes to innovate or not, induces very serious upward biases in the estimated coefficient. A satisfactory resolution to an existing paradox is thus produced: not taking correct account of the endogeneity of financing constraints, leads one to incorrectly conclude that presence of financing constraints and innovation are positively correlated. An example of a study where this "paradox" is encountered is [Mohnen and Roller, 2005]. When allowing for endogeneity of a "Financing Constraints" indicator variable, it is found to have an extremely significant statistically *negative* effect on innovation, with the size of the effect almost tripled once endogeneity is introduced.

Fourthly, we use a new econometric approach to be able to estimate for the first time both direct and reverse interactions between innovation and financial constraints. Obviously, the propensity to innovate may be affected by financial constraints, and at the same time, innovative firms are likely to face specific financial constraints: innovation affects survival of firms (see [Audretsch, 1995] and [Klette and Kortum, 2004]), asset intangibility is higher for innovative firms which lowers their collateral value and due to their innovative nature informational asymmetries with external investors are more pronounced. But, the estimation of such of models is complicated by the problem of *coherency* specific to limited dependent variables models with qualitative endogenous regressors. Traditionally, the problem is tackled by avoiding recursive effects (for instance, by considering that financial constraints may affect the propensity to innovate while the probability of financing constraint is not affected by the innovative behavior).

COMPARATIVE TABLE OF RESULTS: [Section 3.3]

Here, we employ a novel econometric method to deal with such LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservable errors which is a restricted ML Estimation approach conditioned on prior sign restrictions on model parameters. Thus we are able to quantify, for the first time in the literature, the interaction between financing constraints and a firm's decision and ability to innovate without forcing the model to be recursive. The magnitudes of the direct and reverse effects that we find are quite striking: first, *financial constraints reduce by approximately 30% the likelihood of a firm to be innovative.* In addition, a firm *undertaking actively innovative activities more than doubles the probability of encountering a binding financing constraint*, possibly because potential lenders are particularly wary of granting loans to firms of such type because of the extra riskiness involved.

Finally, we merge our dataset with a previous wave of the innovation survey to study the dynamics of the interaction between innovation and financial constraints. We find evidence of state dependence for both variables: firms tend to innovate continuously rather than occasionally and past financial difficulties are correlated with the present ones even after conditioning on important firm characteristics. Moreover, it seems that firms with current but also past innovative experiences are more likely to find it difficult to finance their current projects.

CONCLUDING SECTION: [Section 4]

2 Econometric Coherency in LDV Models

In this section we present and study the fundamental identification issue of *coherency* of Limited Dependent Variable (LDV) models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. An LDV model was defined originally by [Gourieroux et al., 1980] to be "coherent" if it implies a valid function from the unobservables that drive the model to the observed limited dependent variables.

2.1 Defining Coherency in a Simultaneous LDV Model with Two Interactive Responses

We focus our discussion of the coherency problem in LDV models by using the Simultaneous LDV Model with Two Binary Responses. In this model, limited dependent variables y_1 and y_2 are jointly determined through filter functions $\tau_1(\cdot)$ and $\tau_2(\cdot)$ operating on latent variables y_1^* and y_2^* respectively:

$$y_{1it} = \tau_1\left(y_{1it}^* \equiv [h_1(x_{1it}'\beta_1, y_{2it}\gamma) + \epsilon_{1it}]\right) \tag{1}$$

$$y_{2it} = \tau_2\left(y_{2it}^* \equiv [h_2(x_{2it}'\beta_2, y_{1it}\delta) + \epsilon_{2it}]\right) \tag{2}$$

The (possibly non-linear) functions $h_1(\cdot)$ and $h_2(\cdot)$ are known up to parameter vectors β_1 and β_2 and the two interaction coefficients γ and δ. The interaction terms $y_{2it}\gamma$ and $y_{1it}\delta$ appear in the respective latent variables y_{1it}^* and y_{2it}^*. Let x_{1it} and x_{2it} denote the vectors of exogenous factors for each side of the model. The parameter vector to be estimated is $\theta \equiv (\beta_1', \beta_2', \gamma, \delta, \sigma_1^2, \sigma_2^2, \rho)'$ where $\rho \equiv correlation(\epsilon_{1it}, \epsilon_{2it})$. In the most general case, the sample is a panel data set indexed by $i = 1, \cdots, N$ and $t = 1, \cdots, T$.

The existing econometric literature has established as the typical coherency condition to be: $\gamma \cdot \delta = 0$, i.e., no reverse interaction terms are allowed among the two endogenous variables. This condition is sufficient for the joint distribution $(y_{1it}, y_{2it}|x_1, x_2, \theta)$ to be well-specified. [Gourieroux et al., 1980] explain the condition in terms of there being a valid function from $(\epsilon_{1it}, \epsilon_{2it})$ to the observable endogenous variables (y_{1it}, y_{2it}). [Lewbel, 2007] establishes necessary and sufficient conditions for coherency by approaching the problem as requiring a valid reduced form system for (y_{1it}, y_{2it}). For example, if $\delta = 0$ then the RF for y_{2it} is:

$$y_{2it} = \tau_2 \left(h_2(x_{2it}'\beta_2) + \epsilon_{2it} \right)$$

and hence the RF for y_{1it} is given by:

$$y_{1it} = \tau_1 \left(h_1 \left(x_{1it}'\beta_2, \gamma \cdot \left(\tau_2 h_2 \left(x_{2it}'\beta_2 \right) + \epsilon_{2it} \right) \right) + \epsilon_{1it} \right)$$

2.2 The Joint Bivariate Binary Probit Model

The leading case we focus on here is the binary threshold crossing response model in which:

$$\tau_j(z) \equiv \mathbf{1}(z > 0)$$

where $\mathbf{1}(z > 0)$ is the *indicator function* defined by: $\mathbf{1}(z > 0) \equiv \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z \leq 0 \end{cases}$. In terms of the two latent variables y_1^* and y_2^* and the observed binary indicators y_1 and y_2, and suppressing the observation indices:

$$y_1 = \begin{cases} 1 & if \quad y_1^* \equiv x_1'\beta_1 + \gamma y_2 + \epsilon_1 > 0 \\ 0 & if \quad y_1^* \equiv x_1'\beta_1 + \gamma y_2 + \epsilon_1 \leq 0 \end{cases} \tag{3}$$

$$y_2 = \begin{cases} 1 & if \quad y_2^* \equiv x_2'\beta_2 + \delta y_1 + \epsilon_2 > 0 \\ 0 & if \quad y_2^* \equiv x_2'\beta_2 + \delta y_1 + \epsilon_2 \leq 0 \end{cases} \tag{4}$$

In the Empirical application of Section 3.2, we employed this model to study the impact of financing constraints on a firm's decision and ability to innovate in a panel data context.[3]

The specific version of model becomes:

$$I_{it} = \begin{cases} 1 & if \quad I_{it}^* \equiv x_{it}^I\beta^I + \gamma F_{it} + \epsilon_{it}^I > 0 \\ 0 & if \quad I_{it}^* \equiv x_{it}^I\beta^I + \gamma F_{it} + \epsilon_{it}^I \leq 0 \end{cases} \tag{5}$$

[3] A related application of this setup in International Finance is the Banking and Currency Crises Model of [Falcetti and Tudela, 2008] where (C_{it}, B_{it}) refer to Currency and Banking Crises respectively. Their model is recursive, in that Currency crises are allowed to depend on Banking crises but not vice-versa.

$$F_{it} = \begin{cases} 1 & if \quad F^*_{it} \equiv x^F_{it}\beta^F + \delta I_{it} + \epsilon^F_{it} > 0 \\ 0 & if \quad F^*_{it} \equiv x^F_{it}\beta^F + \delta I_{it} + \epsilon^F_{it} \leq 0 \end{cases} \tag{6}$$

Analytically, $(I_{it}, F_{it}) \in \{(1,1),(1,0),(0,1),(0,0)\}$ such that:

(I_{it}, F_{it})	I^*_{it}		F^*_{it}
$(1,1)$	$x'_{1it}\beta_1 + \gamma + \epsilon_{1it} > 0$	,	$x'_{2it}\beta_2 + \delta + \epsilon_{2it} > 0$
$(1,0)$	$x'_{1it}\beta_1 + \epsilon_{1it} > 0$	,	$x'_{2it}\beta_2 + \delta + \epsilon_{2it} < 0$
$(0,1)$	$x'_{1it}\beta_1 + \gamma + \epsilon_{1it} < 0$	,	$x'_{2it}\beta_2 + \epsilon_{2it} > 0$
$(0,0)$	$x'_{1it}\beta_1 + \epsilon_{1it} < 0$	,	$x'_{2it}\beta_2 + \epsilon_{2it} < 0$

For a typical *it* observation, the probability $Prob(I_{it}, F_{it}|X, \theta)$ is thus characterized by the constraints on the unobservables:

$$(a^I, a^F)' < (\epsilon^I, \epsilon^F)' < (b^I, b^F)'$$

through the configuration:

I_{it}	F_{it}	a^I	b^I	a^F	b^F
1	1	$-x^I_{it}\beta^I - \gamma$	∞	$-x^F_{it}\beta^F - \delta$	∞
1	0	$-x^I_{it}\beta^I$	∞	$-\infty$	$-x^F_{it}\beta^F - \delta$
0	1	$-\infty$	$-x^I_{it}\beta^I - \gamma$	$-x^F_{it}\beta^F$	∞
0	0	$-\infty$	$-x^I_{it}\beta^I$	$-\infty$	$-x^F_{it}\beta^F$

In general, in the absence of coherency conditions, there will be *overlaps* and/or *gaps* in the domain of $(\epsilon_{1it} + x'_{1it}\beta_1, \epsilon_{2it} + x'_{2it}\beta_2)$. These would be ruled out by the aforementioned sufficient coherency condition.[4]

2.3 The Traditional Approach to Coherency Conditions

Let us use a slightly more complicated simultaneous LDV model to analyze coherency, namely the *binary & trinomial ordered probit model* of [Hajivassiliou and Ioannides, 2007].

[4] A related LDV model that does **not** exhibit similar coherency difficulties is the bivariate probit model with **latent variable** interactions (as opposed to limited variable interactions). Specifically:

$$y_{1it} = \tau_1\left(y^*_{1it} \equiv [h_1(x'_{1it}\beta_1, y^*_{2it}\gamma) + \epsilon_{1it}]\right)$$
$$y_{2it} = \tau_2\left(y^*_{2it} \equiv [h_2(x'_{2it}\beta_2, y^*_{1it}\delta) + \epsilon_{2it}]\right)$$

Then:

$$y^*_1 = x_1\beta_1 + y^*_2\gamma + \epsilon_1$$
$$y^*_2 = x_1\beta_1 + y^*_1\delta + \epsilon_2$$

and

$$y^*_1 = x_1\beta_1 + \gamma \cdot [x_2\beta_2 + y^*_1\delta + \epsilon_2] + \epsilon_1$$
$$y^*_2 = x_2\beta_2 + \delta \cdot [x_1\beta_1 + y^*_2\gamma + \epsilon_1] + \epsilon_2$$

Hence $y^*_1 = RF_1$ and $y^*_2 = RF_2$, allowing us to obtain $y_1 = \tau(RF_1)$ and $y_2 = \tau(RF_2)$. We thus see that it is considerably more straightforward to establish coherency identification of LDV models with latent variable interactions as opposed to limited variable interactions.

This model studies interactions between liquidity and employment constraints on individual households indexed by i at a given point in time indexed by t. The reason we select this model is because it can exhibit simultaneously *both* types of incoherency (overlaps and gaps). This is critical because it will allow us to devise estimation strategies that overcome certain types of incoherency.

Define two latent dependent variables y_{1it}^* and y_{2it}^*. The first denotes the propensity of individual i in period t to be liquidity constrained and the second its propensity to face employment hour constraints. The corresponding limited dependent variables are denoted by y_{1it} and y_{2it}. Dropping the it subscripts for simplicity, the model is defined by:

$$y_1 = \begin{cases} 1 & if \quad y_1^* > 0 \ \text{(liquidity constraint binding)}, \\ 0 & if \quad y_1^* \leq 0 \ \text{(liquidity constraint not binding)}. \end{cases}$$

$$y_2 = \begin{cases} -1 & if \quad y_2^* \leq \lambda^- \ \text{(overemployed)} \\ 0 & if \quad \lambda^- \leq y_2^* < \lambda^+ \ \text{(voluntarily employed)} \\ +1 & if \quad \lambda^+ \leq y_2^* \ \text{(under-/unemployed)}. \end{cases}$$

where the latent variables are given by:

$$y_1^* = \mathbf{1}(y_2^* < \lambda^-)\gamma_{11} + \mathbf{1}(\lambda^- < y_2^* < \lambda^+)\gamma_{12} + x_1'\beta_1 + \epsilon_1$$

$$y_2^* = \mathbf{1}(y_1^* > 0)\delta + x_2\beta_2 + \epsilon_2$$

Since (S, E) lie in $\{0,1\} \times \{-1, 0, 1\}$, the 6 possible configurations may be enumerated as follows:

S	E	y_1^*	y_2^*
0	-1	$\gamma_{11} + x_1\beta_1 + \epsilon_1 < 0,$	$x_2\beta_2 + \epsilon_2 < \lambda^-$
0	0	$x_1\beta_1 + \epsilon_1 < 0,$	$\lambda^- < x_2\beta_2 + \epsilon_2 < \lambda^+$
0	+1	$\gamma_{12} + x_1\beta_1 + \epsilon_1 < 0,$	$\lambda^+ < x_2\beta_2 + \epsilon_2$
1	-1	$\gamma_{11} + x_1\beta_1 + \epsilon_1 > 0,$	$\delta + x_2\beta_2 + \epsilon_2 < \lambda^-$
1	0	$x_1\beta_1 + \epsilon_1 > 0,$	$\lambda^- < \delta + x_2\beta_2 + \epsilon_2 < \lambda^+$
1	+1	$\gamma_{12} + x_1\beta_1 + \epsilon_1 > 0,$	$\lambda^+ < \delta + x_2\beta_2 + \epsilon_2$

In terms of the unobservables, the probability of a (y_1, y_2) observed pair is equivalent to the probability:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} < \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} < \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

where $(\epsilon_1, \epsilon_2)' \sim N(0, \Sigma_\epsilon)$, and a and b are given by:

S	E	a_1	a_2	b_1	b_2
0	-1	$-\infty$	$-\infty$	$-(\gamma_{11} + x_1\beta_1)$	$\lambda^- - x_2\beta_2$
0	0	$-\infty$	$\lambda^- - x_2\beta_2$	$-x_1\beta_1$	$\lambda^+ - x_2\beta_2$
0	+1	$-\infty$	$\lambda^+ - x_2\beta_2$	$-(\gamma_{12} + x_1\beta_1)$	$+\infty$
1	-1	$-(\gamma_{11} + x_1\beta_1)$	$-\infty$	$+\infty$	$\lambda^- - \delta - x_2\beta_2$
1	0	$-x_1\beta_1$	$\lambda^- - \delta - x_2\beta_2$	$+\infty$	$\lambda^+ - \delta - x_2\beta_2$
1	+1	$-(\gamma_{12} + x_1\beta_1)$	$\lambda^+ - \delta - x_2\beta_2$	$+\infty$	$+\infty$

Using traditional arguments, we obtain that a sufficient condition for coherency of the model is:

$$(\gamma_{11} + \gamma_{12})\delta = 0 \text{ and } \gamma_{11}\gamma_{12}\delta = 0.$$

To verify this condition, suppose $(S, E) = (0, 0)$. This rules out $(S, E) = (0, -1)$ because $x_2\beta_2 + \epsilon_2 > \lambda^-$, and rules out $(S, E) = (1, 0)$ because $x_1\beta_1 + \epsilon_1 < 0$.

But $(1, -1)$ is not ruled out if the coherency conditions do not hold, since γ_{11} could be sufficiently negative and δ sufficiently positive to imply the $(1, -1)$ conditions.

Similarly, the $(1, 1)$ possibility cannot be ruled out in the absence of the coherency conditions, since γ_{12} and δ can be sufficiently positive.

Such logical inconsistencies are prevented if either (a) $\delta = 0$ or (b) γ_{11} and γ_{12} are simultaneously 0.

2.4 Difficulties with the traditional approaches:

The first difficulty is that the existing derivations of formal conditions lack intuition, are difficult to generalize, and are sufficient but not necessary. The second one is that in practice, non-triangular or reverse triangular cases are the most interesting.

To overcome the first difficulty, alternative ways for establishing coherency were developed in [Hajivassiliou, 2008], that are both intuitive and straightforward. In addition, this novel approach yields less strict conditions that in fact allow more interesting practical applications.

Specifically, it is shown in the next Section how to establish coherency without recursiveness through the use of sign restrictions on model parameters. The fact that our novel approach for the first time eliminates the need to assume recursivity is quite important: recursivity corresponds to the key identifying assumption that innovation does not affect financial distress directly ($\delta = 0$). On a priori grounds, this assumption seems particularly dubious since innovation may lead to more profits and thus relax financial constraints (corresponding to $\delta > 0$). An alternative possibility is that innovation may lead to higher investment in intangible assets thus reinforcing binding financial constraints (corresponding to $\delta < 0$). Both possibilities violate the traditional coherency condition.

2.5 Extending the Traditional Approach to Coherency

The traditional approaches to model coherency suffer from several major difficulties. Firstly, derivations of formal conditions using the traditional approach lack intuition. Secondly, the derived conditions are impossible to generalize and verify in moderately more complicated LDV models, especially in cases where the models are allowed to contain intertemporal endogeneity of the type considered in [Falcetti and Tudela, 2008]. Similarly, in case the joint binary probit model (3)-(4) is extended intertemporally, as for example in the empirical dynamic application in Section 3.2, the coherency condition is impossible to generalize and verify using the traditional analysis of the previous subsection. Thirdly, in practice non-triangular or reverse triangular cases are the most interesting from an economic point of view.[5] Finally, the traditional approaches focus on establishing *sufficient* conditions for coherency, while our methods allow us to prove that they are *not necessary*.

To overcome the first two difficulties, alternative ways for establishing coherency are developed here, that are both intuitive and straightforward, as well as much more generalizable. In addition, our methods allow us to resolve the last two difficulties leading to estimation based on CMLE for much more interesting practical applications. It is shown in the next Section how to establish coherency without model recursiveness through the use of (a) endogeneity in terms of latent variables and/or (b) sign restrictions on model parameters. The fact that our novel approach for the first time eliminates the need to assume model recursivity is quite important for the economic problem studied in the empirical application in Section 3.2. As we explain there, recursivity corresponds to the key identifying assumption that firm innovation does not affect financial distress directly ($\delta = 0$). On a priori grounds, this assumption seems particularly dubious since firm innovation may lead to more profits and thus relax financial constraints (corresponding to $\delta > 0$). An alternative possibility is that innovation may lead to higher investment in intangible assets thus reinforcing binding financial constraints (corresponding to $\delta < 0$). Both possibilities violate the traditional coherency condition.[6]

2.5.1 Novel Approach 1: Graphical

Let us illustrate the first approach using the Liquidity-Employment constraints application of [Hajivassiliou and Ioannides, 2007]. This graphical approach was first included in the LSE working paper [Hajivassiliou, 2002] and was presented at the CRETE Conference in Syros in 2003. It should also be noted that our graphical approach presented here is related to that of [Tamer, 2003] who studied the problem of coherency in bivariate discrete models for games with multiple equilibria.

[5] Exceptions of this do exist: for example in industrial organization a two-agent discrete game may be employed to model the strategic interactions between firms in a duopoly setup. If one firm is a Stackelberg leader while the other is a follower, a recursive model may be applicable, even though the analogy is not precise.

[6] Note that throughout we expect $\gamma < 0$, i.e., the higher the probability that a firm faces a binding financial constraint, the less likely it is that it is able to innovate. So the two possibilities translate to: (a) $\gamma < 0$, $\delta > 0$ and (b) $\gamma < 0$, $\delta < 0$.

Figure 1 gives the 6 possible regimes $(y_1 \times y_2) = \{1, 0\} \times \{-1, 0, 1\}$ in terms of the two latent variables y_1^* and y_2^* and the possible configurations in terms of parameters $\bar{\lambda}$, $\underline{\lambda}$, δ, γ_{11}, and γ_{12}. y_1^* is on the horizontal axis and y_2^* on the vertical. The figure makes clear the role of the coherency condition (a) $\delta = 0$ or (b) $\gamma_{11} = \gamma_{12} = 0$: in general, regions $R2$ and $R6$ exhibit double-counting (cross-hatched area), as well as a white rectangle remains which makes the six regions not mutually exhaustive. These two logical incoherencies disappear when either $\delta = 0$ and/or $\gamma_{11} = \gamma_{12} = 0$ hold.

[Figure 1 approximately here.]

We develop further our graphical approach here and use it to highlight the fundamental distinction between two types of incoherency, the first corresponding to overlap regions in latent variables space, while the second to empty regions. A critical fact that this model illustrates is that that a model may exhibit **simultaneously incoherencies of one or both kinds**: *empty region incoherency* vs. *overlapping region incoherency.*

We then show below that under prior sign restrictions on model parameters, incoherencies of the empty region type can be eliminated by relying on suitable parameter sign restrictions through the use of Conditional MLE.

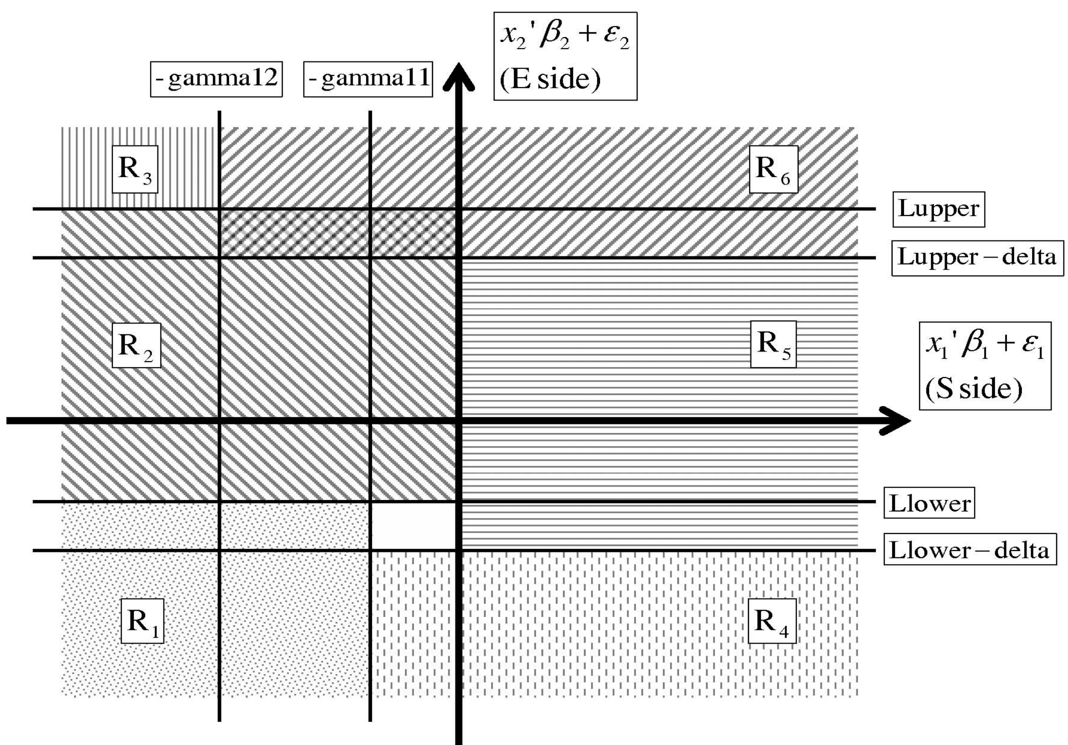

Figure 1: Regions in latent variables space that define the observed dependent qualitative variables Liquidity (binary) and Employment (trinomial) Constraints. Consequently there are six possible regimes. We see two regions of incoherency, one cross-hatched, the other empty. Both regions disappear under the coherency condition that Delta=0 and/or Gamma11=Gamma12=0.

Coherency of Binary+Trinomial Model

2.5.2 Novel approach 2: DGP From First Principles

Despite the usefulness of the graphical approach of the previous section to LDV problems with two latent variables, the method is very unwieldy or inapplicable to higher dimensional cases. To cover such problems, we develop a second approach to incoherency, which consists of designing a data-generating algorithm (hypothetical or implemented on a computer) to simulate random draws from an LDV model's structure. Again let us use the Liquidity-Employment Constraints application of [Hajivassiliou and Ioannides, 2007] to illustrate the method. We draw ϵ_1 and ϵ_2 under the joint bivariate normal distribution with zero mean vector and variance-covariance matrix Σ_ϵ, and given $x_1'\beta_1$ and $x_2'\beta_2$ we attempt to generate y_1^* and y_2^*. This is straightforward provided the coherency condition holds: If (a) $\delta = 0$, then latent y_2^* can be drawn, then LDV y_2, which together with ϵ_1 and $x_1'\beta_1$ determines the right hand side of y_1^*, thus allowing y_1 to be drawn. Similarly, if (b) $\gamma_{11} = \gamma_{12} = 0$, then y_1^* can be drawn from the first equation based on ϵ_1 and $x_1'\beta_1$, which determines y_1, thus giving y_2^* and hence y_2. In general it is impossible, however, to devise such a data generation mechanism in case the coherency condition does not hold.

This approach is related to the [Gourieroux et al., 1980] condition that a function exist from ϵ_1, ϵ_2 to y_1, y_2. It is also related to [Lewbel, 2005] in that coherency translates to there being a valid reduced form for the endogenous variables.

As we will show in section 2.8, the approach extends naturally to cases with intertemporal endogeneities in panel LDV models, and can be used to prove the coherency of the classic multiperiod panel probit with state dependence ([Heckman, 1981a]), as well as the intertemporal endogeneity versions of the models in Section 3.2 with explicit dynamic effects.

2.6 Identification Under Prior Sign Restrictions

The graphical approach we developed in the previous section highlights two distinct cases of incoherency: the first type of incoherency corresponds to regions of the observed endogenous variables of the model being *overlapping*, while the second to regions that are *empty*. We show that empty region incoherency can be overcome through conditional maximum likelihood (CMLE) of truncating the LDVs to lie outside the incoherency regions.[7] Our CMLE approach can also be motivated through the DGP approach for establishing coherency that we discussed in the previous subsection. In that case, we need to consider DGPs truncated to lie on a specific region of the latent variables space. A specific method for achieving this is given in Technical Appendix 2.

It is useful to highlight here the similarities and differences to the analysis in [Tamer, 2003], who also used a graphical approach to resolve an incomplete simultaneous discrete response model for a homogeneous two-agent discrete game of entry. Since the two rival firms in his setting were assumed identical, any incoherency arising was necessarily of the indeterminate type — see our two subcases 2.6.1 and 2.6.2,

[7] We also explain below that overlapping region incoherency *cannot* be transformed into empty region incoherency by redefining one of the observed binary LDVs to its complement.

where the interaction terms γ and δ are of the same sign. Consequently, the possibility of the interaction terms being of opposite sign was not under focus in his analysis and hence the applicability of CMLE to resolve those cases was not considered. It is also useful to note that our approach for establishing coherency through the use of prior sign restrictions developed here is related to the recent approach by [Uhlig, 2005] for Vector Autoregression identification under prior sign restrictions on impulse response functions.[8] [Dagenais, 1997] also makes a distinction between alternative types of incoherency regions.[9]

We illustrate the CMLE approach for establishing coherency through prior sign restrictions by using the joint binary probit model:[10]

$$I = \begin{cases} 1 & if \ I^* \equiv x_1'\beta_1 + \gamma F + \epsilon_1 > 0 \\ 0 & otherwise \end{cases} \tag{7}$$

$$F = \begin{cases} 1 & if \ F^* \equiv x_2'\beta_2 + \delta I + \epsilon_2 > 0 \\ 0 & otherwise \end{cases} \tag{8}$$

Obviously, there exist **four cases** based on the signs of γ and δ. These are presented in the four figures that follow.

2.6.1 Case 1: $\gamma > 0, \delta > 0$ — overlapping regions, incoherency

[Figure 2 approximately here.]

[8] We are indebted to Alain Trognon for pointing out the potential of parameter sign restrictions overcoming incoherency of the "empty region" type, and to Hashem Pesaran for bringing to our attention Uhlig's work on sign identification.

[9] Unfortunately his work remains incomplete and unpublished due to his untimely death.

[10] For the first equation, I^* is used for the latent and I for the observed LDV as a mnemonic to the *Innovation* side of the model of Section 3.2 below. Similarly, for the second equation we use F^* and F as a mnemonic to *Financing Constraints*.

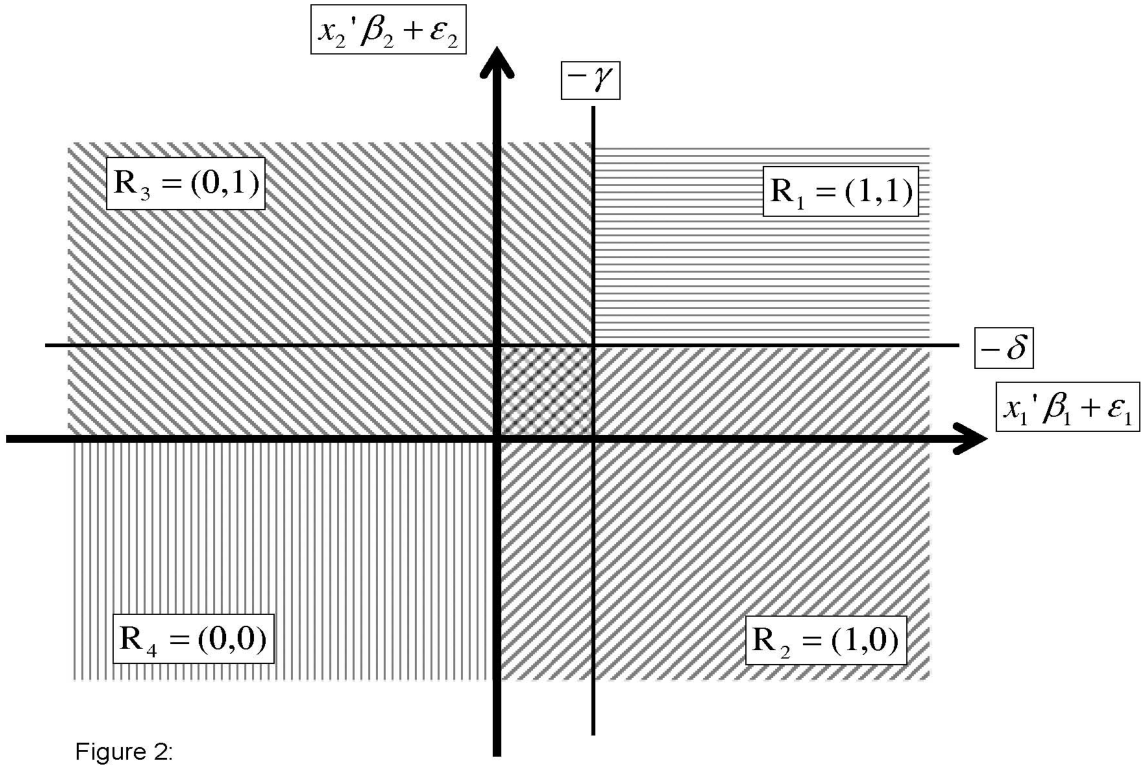

Figure 2:
Innovation and Finance Constraint bivariate binary model in latent variables space. Four implied regimes. Case 1: gamma>0, delta>0. Region of incoherency is of the overlap type (cross-hatched).

Case 1: $\gamma > 0, \delta > 0$

2.6.2 Case 2: $\gamma < 0, \delta < 0$ — overlapping regions, incoherency

[Figure 3 approximately here.]

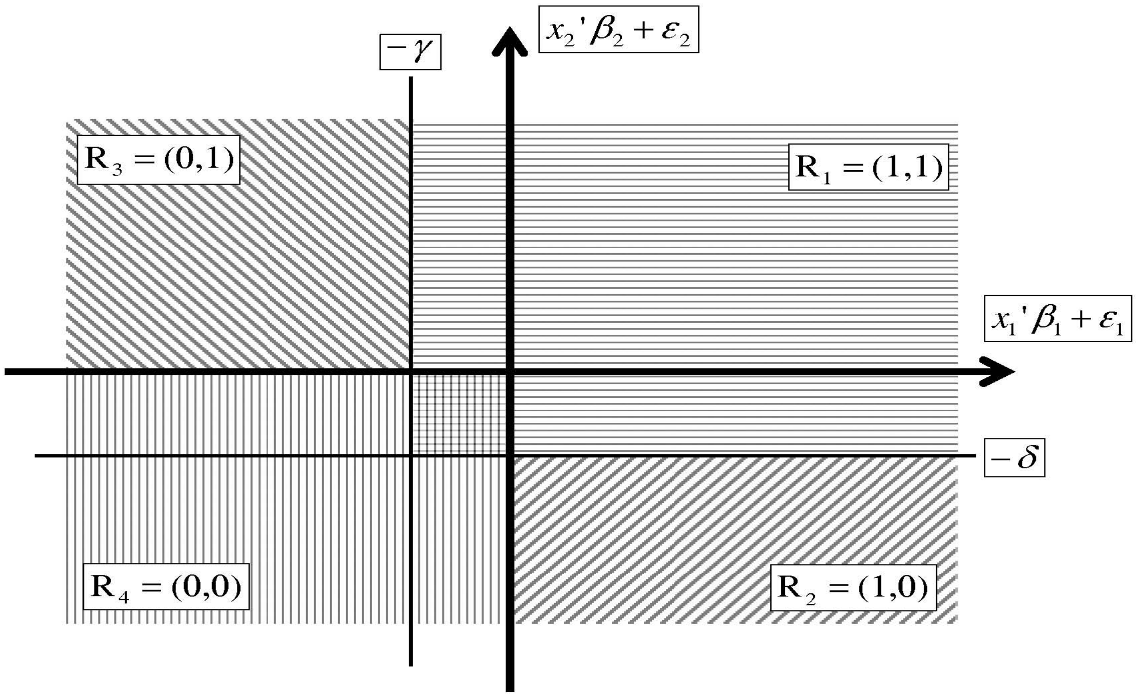

Figure 3:
Innovation and Finance Constraint bivariate binary model in latent variables space.
Four implied regimes. Case 2: gamma<0, delta<0. Region of incoherency is of the overlap type (cross-hatched).

Case 2: $\gamma < 0, \delta < 0$

2.6.3 Case 3: $\gamma > 0, \delta < 0$ — empty regions, coherency through conditioning

[Figure 4 approximately here.]

For this case, coherency can be achieved by conditioning to lie outside the "empty" region of figure 2.6.4, which has conditioning probability:

$$1 - Prob(-\gamma < \epsilon_1 + x_1'\beta_1 < 0, 0 < \epsilon_2 + x_2'\beta_2 < -\delta)$$

The estimation method that implements this is CMLE

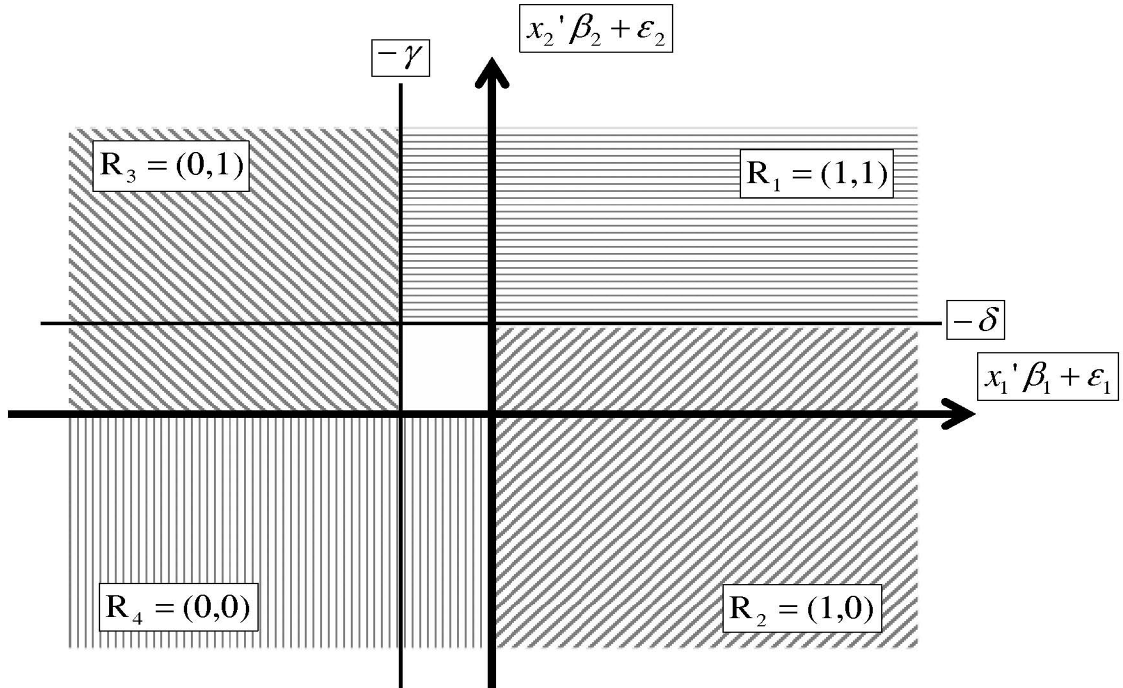

Figure 4:
Innovation and Finance Constraint bivariate binary model in latent variables space.
Four implied regimes. Case 3: gamma>0, delta<0. Region of incoherency is of the empty region type. Hence, Conditional MLE achieves coherency by ruling out the event:
[-gamma<eps1+x1b1<0,0<eps2+x2b2<-delta]

Case 3: $\gamma > 0, \delta < 0$

2.6.4 Case 4: $\gamma < 0, \delta > 0$ — empty regions, coherency through conditioning

[Figure 5 approximately here.]

For this case also, coherency is achieved by conditioning to lie outside the "empty" region of Figure 4. The conditioning probability is:

$$1 - Prob(0 < \epsilon_1 + x_1'\beta_1 < -\gamma, \delta < \epsilon_2 + x_2'\beta_2 < 0)$$

and the appropriate estimation method is CMLE.

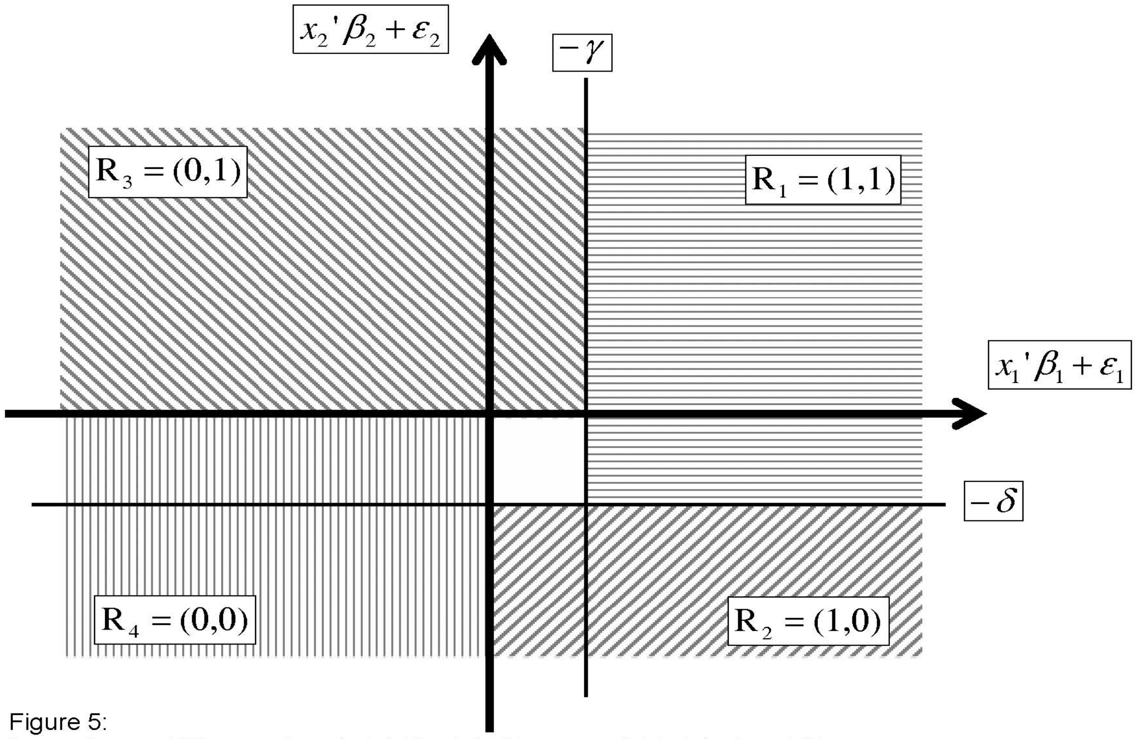

Figure 5:
Innovation and Finance Constraint bivariate binary model in latent variables space.
Four implied regimes. Case 4: gamma<0,delta>0. Region of incoherency is of the empty region type. Hence, Conditional MLE achieves coherency by ruling out the event:
[0<eps1+x1b1<-gamma,delta<eps2+x2b2<0]

Case 4: $\gamma < 0, \delta > 0$

2.6.5 Can Overlapping Regions Incoherency be Overcome through LDV Redefinition?

We have shown that in general, in the absence of coherency conditions, there will be *overlaps* and/or *gaps* in the domain of $(\epsilon_1 + x_1'\beta_1, \epsilon_2 + x_2'\beta_2)$. At this point, a researcher might be tempted to propose that the incoherency cases with overlapping regions (Cases 1 and 2 above) may be overcome by redefining one of the two limited dependent variables to their complement. According to this reasoning, since the incoherency is caused in these cases because γ and δ are of the same sign, and since changing y_2, say, to its complement $y_2^N \equiv (1 - y_2)$ would result in $\delta^N \equiv -\delta$, then coherency would be achieved since then $\gamma \cdot \delta^N < 0$.

Such reasoning would be incorrect, however. We analyze here this idea and show that such a redefinition would *maintain* the overlapping-region incoherency. This is because the $y_2^N \equiv (1 - y_2)$ redefinition would also switch the sign of γ and hence $\gamma^N \cdot \delta^N > 0$ just as $\gamma \cdot \delta > 0$.

Let us return to the bivariate binomial probit (7) and (8). Suppose we have incoherency because we believe $\gamma > 0$ (in our application below translating to binding finance constraints expected to raise the chance of innovation I) and that $\delta > 0$ (innovative firms face a higher chance that the banks will refuse them a loan). So

$\gamma \cdot \delta > 0$. This is Case 1 analyzed in subsection 2.6.1 as represented by Figure 2.6.1, and corresponding to the constraints on the unobservables:

$$(a^1, a^2)' < (\epsilon^1, \epsilon^2)' < (b^1, b^2)'$$

such that:

I	F	a^1	b^1	a^2	b^2	Shading	Region
1	1	$-x_1'\beta_1 - \gamma$	∞	$-x_2'\beta_2 - \delta$	∞	horizontal	R1
1	0	$-x_1'\beta_1$	∞	$-\infty$	$-x_2'\beta_2 - \delta$	///////	R2
0	1	$-\infty$	$-x_1'\beta_1 - \gamma$	$-x_2'\beta_2$	∞	\\\\\\\	R3
0	0	$-\infty$	$-x_1'\beta_1$	$-\infty$	$-x_2'\beta_2$	vertical	R4

Now consider the transformed model with NF **instead of** F. This transformation still gives an overlapping region in the transformed variables, and hence corresponds to an incoherent model. To see this, proceed as follows:

In terms of the two latent variables I^* and $NF^* = -F^*$ and the observed binary indicators I and $NF = 1 - F$, and suppressing the observation index:

$$I = \begin{cases} 1 & if \quad I^* \equiv x_1'\beta_1 + \gamma^N NF + \epsilon_1 > 0 \\ 0 & if \quad I^* \equiv x_1'\beta_1 + \gamma^N NF + \epsilon_1 \leq 0 \end{cases} \tag{9}$$

$$NF = \begin{cases} 1 & if \quad NF^* \equiv x_2'\beta_2^N + \delta^N I + \epsilon_2^N > 0 \\ 0 & if \quad NF^* \equiv x_2'\beta_2^N + \delta^N I + \epsilon_2^N \leq 0 \end{cases} \tag{10}$$

Given this transformation, we expect that $\gamma^N < 0$ (high NF means not very binding constraints so cause dampening of I) and that $\delta^N < 0$ (firms who have high I i.e., innovate raise the chance the banks will refuse them a loan so low NF). So $\gamma^N \cdot \delta^N > 0$. See Figure 2.6.5.

[Figure 6 approximately here.]

For a typical i observation, the probability $Prob(y_{1i}, y_{2i}|X, \theta)$ is characterized by the constraints on the unobservables:

$$(a^1, a^2)' < (\epsilon_1, \epsilon_2^N)' < (b^1, b^2)'$$

through the configuration:

I	NF	a^1	b^1	a^2	b^2	Shading	Region
1	0	$-x_1'\beta_1$	∞	$-\infty$	$-x_2'^N\beta_2 - \delta^N$	horizontal	R1
1	1	$-x_1'\beta_1 - \gamma^N$	∞	$-x_2'\beta_2^N - \delta^N$	∞	///////	R2
0	0	$-\infty$	$-x_1'\beta_1$	$-\infty$	$-x_2'^N\beta_2$	\\\\\\\	R3
0	1	$-\infty$	$-x_1'\beta_1 - \gamma^N$	$-x_2'^N\beta_2$	∞	vertical	R4

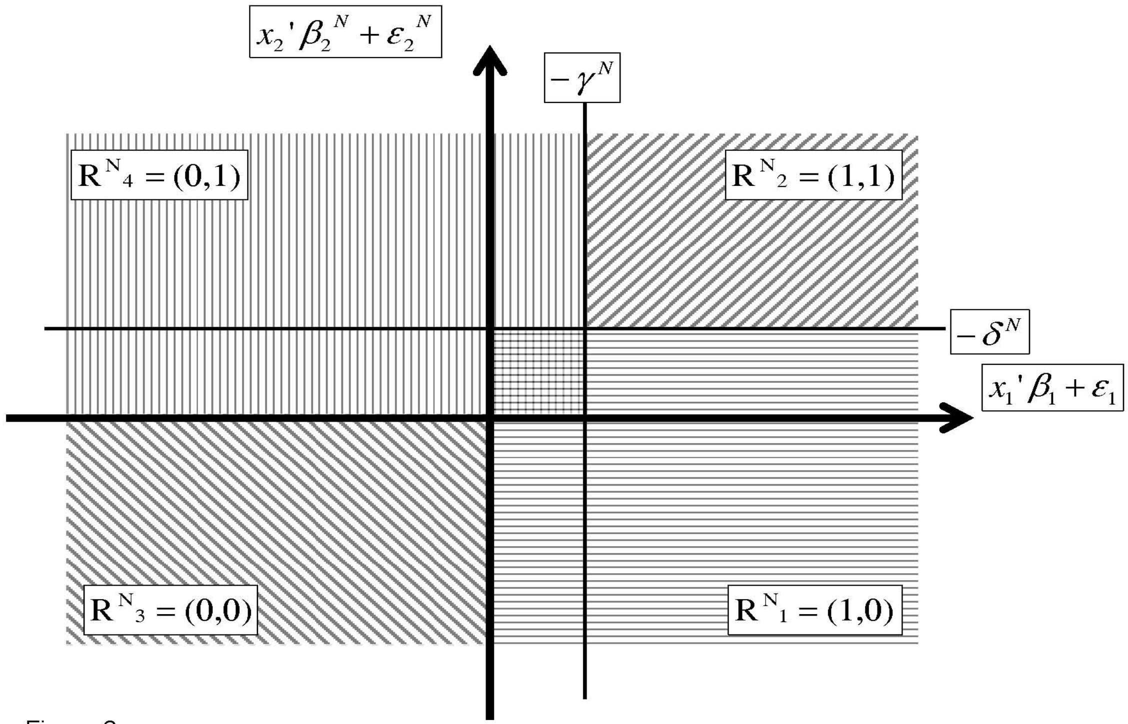

Figure 6:
Innovation and Finance Constraint bivariate binary model in latent variables space. Situation identical to Figure 1 (Case 3) but with Innovation negated to its complement (1-Innovation). Both gammaN and deltaN switch sign by this redefinition, and hence: gammaN<0,deltaN<0. Consequently, the model remains incoherent with an overlapping (cross-hatched) region.

$$\gamma^N < 0, \delta^N < 0$$

2.7 Efficient Estimation through Conditional Maximum Likelihood

The optimal parametric estimation approach for the models with empty region incoherency (Cases 3 and 4 above) will be *conditional maximum likelihood (CMLE)*, employing the appropriate likelihood contributions that characterize correctly the necessary conditioning through truncation that ensures that the LDVs stay out of the empty region of incoherency. For example, assuming independence across observations $\{it\}$, the likelihood contribution in Case 3 will be:

$$l_{it} = \frac{Prob\,(\epsilon_1, \epsilon_2 : I = 1(I^* > 0) \;\&\; F = 1(F^* > 0))}{(1 - Prob(-\gamma < \epsilon_1 + x_1'\beta_1 < 0, 0 < \epsilon_2 + x_2'\beta_2 < -\delta)}$$

while for Case 4:

$$l_{it} = \frac{Prob\,(\epsilon_1, \epsilon_2 : I = 1(I^* > 0) \;\&\; F = 1(F^* > 0))}{(1 - Prob(0 < \epsilon_1 + x_1'\beta_1 < -\gamma, \delta < \epsilon_2 + x_2'\beta_2 < 0)}$$

These likelihood contributions make it clear why approaches that ignore the coherency issue are inconsistent in general: the inconsistency would arise because the conditioning probability expressions in the denominator are functions of the underlying parameters and data, and hence affect critically the evaluation of the correct likelihood function.

It should be noted also that Cases 1 and 2 may be handled in an analogous fashion *provided it is assumed first that* the Data Generating Process (DGP) that overcomes the overlapping-regions incoherency is one where $(\epsilon_{1i}, \epsilon_{2i})$ are drawn from an unrestricted bivariate normal distribution and then any draws falling into the overlap region are rejected. To find the correct likelihood contributions in these two cases, first define:

$$\begin{aligned} p^*_{11} &\equiv Prob(I^* > 0, F^* > 0) \\ p^*_{10} &\equiv Prob(I^* > 0, F^* \leq 0) \\ p^*_{01} &\equiv Prob(I^* \leq 0, F^* > 0) \\ p^*_{00} &\equiv Prob(I^* \leq 0, F^* \leq 0) \end{aligned}$$

Then, note that:

$$p^*_{11} + p^*_{10} + p^*_{01} + p^*_{00} = S > 1$$

where $S - 1 \equiv d$, the probability of the overlap region. In Case 1, the overlap occurs between regions $(1, 1)$ and $(0, 0)$, while for Case 2 between regions $(1, 0)$ and $(0, 1)$. Consequently, assuming an Accept/Reject DGP out of the overlap region, the likelihood contribution for observation $\{it\}$ for Case 1 is:

$$l_{it} = \begin{cases} p_{11} \equiv Prob(I = 1 \& F = 1) = (p^*_{11} - d)/(2 - S) \\ p_{10} \equiv Prob(I = 1 \& F = 0) = p^*_{10}/(2 - S) \\ p_{01} \equiv Prob(I = 0 \& F = 1) = p^*_{01}/(2 - S) \\ p_{00} \equiv Prob(I = 0 \& F = 0) = (p^*_{00} - d)/(2 - S) \end{cases}$$

while for Case 2:

$$l_{it} = \begin{cases} p_{11} \equiv Prob(I = 1 \& F = 1) = p^*_{11}/(2 - S) \\ p_{10} \equiv Prob(I = 1 \& F = 0) = (p^*_{10} - d)/(2 - S) \\ p_{01} \equiv Prob(I = 0 \& F = 1) = (p^*_{01} - d)/(2 - S) \\ p_{00} \equiv Prob(I = 0 \& F = 0) = p^*_{00}/(2 - S) \end{cases}$$

This Accept/Reject DGP approach for overcoming *overlapping regions incoherency* is arguably less unambiguous and clear-cut, however, compared to the conditioning through truncation we proposed to handle *empty regions incoherency.* One may consider instead alternative schemes for the overlapping regions case which are particularly suitable for specific economic applications — see, for example, the game-theoretic models of entry analyzed by [Tamer, 2003].

2.8 Establishing the Coherency of Panel LDV Models with Intertemporal Endogeneities using DGP Approach

Extending the analysis to a panel data set, [Hajivassiliou, 2007] explains how the probability of a pair (S_{it}, E_{it}) in subsection 2.3 and a pair (y_{1it}, y_{2it}) in subsection 2.2, can be represented in terms of the linear inequality:

$$(a_1, a_2)' < (\epsilon_1, \epsilon_2)' < (b_1, b_2)'$$

where the error vector has a flexible autocorrelation structure. For example, one-factor random effect assumptions will imply an equicorrelated block structure on Σ_ϵ, while our most general assumption of one-factor random effects *combined with* an AR(1) process for each error implies that Σ_ϵ combines equicorrelated and Toeplitz-matrix features. Consequently, the approach incorporates fully (a) the contemporaneous correlations in ϵ_{it}, (b) the one-factor plus AR(1) serial correlations in ϵ_i, and (c) the dependency of S_{it} on E_{it} and vice versa. The coherency issue expands naturally to the panel sequence of data, by thinking of each (correlated) time-period for a given individual i as a **distinct probit equation** and then dealing with the independent cross-section of equations across individuals. Details of the analysis can be found in [Hajivassiliou, 2007].

2.8.1 Dynamic Model 1: Univariate Panel Data Probit with State Dependence

Our hypothetical DGP method presented in Subsection 2.5.2 for establishing coherency is now applied to the canonical panel data Probit model with state dependence, first analyzed by [Heckman, 1981a]. The model is defined by:

$$y_{iT} = \mathbf{1}(\lambda y_{i,T-1} + x_{iT}\beta + \epsilon_{iT} > 0)$$

$$y_{i,T-1} = \mathbf{1}(\lambda y_{i,T-2} + x_{i,T-1}\beta + \epsilon_{i,T-1} > 0)$$

$$\vdots$$

$$y_{i2} = \mathbf{1}(\lambda y_{i1} + x_{i2}\beta + \epsilon_{i2} > 0)$$

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + u_{i1} > 0)$$

The equation for $t = 1$ is a generalization of the [Barghava and Sargan, 1982] approach. Let $\Sigma \equiv VCov(\epsilon_{iT}, \cdots, \epsilon_{i1}, u_{i1})$. Imposing one-factor random effect assumptions will imply an equicorrelated block structure on the top left $T-1 \times T-1$ block of Σ, while more general assumptions of one-factor random effects *combined with* an AR(1) or ARMA(p,q) processes for each ϵ error implies that Σ combines equicorrelated and Toeplitz-matrix parts. The last row and last column of Σ giving the variance of u_{1i} and its covariances with all ϵ_{it} allow the flexibility stipulated by [Heckman, 1981b].

Define the Cholesky lower triangular times upper triangular factorization of $\Sigma =$

CC'. Given the assumed normality, the error vector can be written:

$$(\epsilon_i', u_{1i})' = C\nu_i \qquad \nu_i \sim N(0_T, I_T)$$

Theorem 1: *The Univariate PD Probit Model with State Dependence defined above is coherent.*

Proof: (using the DGP approach)
Let us begin with the simplified case of the initial condition being exogenous:

$$y_{iT} = \mathbf{1}(\lambda y_{i,T-1} + x_{iT}^{'}\beta + \epsilon_{iT} > 0) \tag{11}$$

$$y_{i,T-1} = \mathbf{1}(\lambda y_{i,T-2} + x_{i,T-1}^{'}\beta + \epsilon_{i,T-1} > 0) \tag{12}$$

$$\vdots \tag{13}$$

$$y_{i2} = \mathbf{1}(\lambda y_{i1} + x_{i2}^{'}\beta + \epsilon_{i2} > 0) \tag{14}$$

$$y_{i1} = \textbf{exogenous} \tag{15}$$

Suppose first the ϵ_{it} has the one-factor (equicorrelated) error components structure $\epsilon_{it} = \alpha_i + \nu_{it}$. Conditional on α_i, these $T-1$ equations are independent (since they only depend on the i.i.d. ν_{it}s). Hence draw an α_i and an independent ν_{i2}. Then use the exogenous y_{i1} outcome to generate y_{i2}. This completes equation 14 which allows to move sequentially to generating y_{i3}, then y_{i4}, etc. until y_{iT} is generated. This establishes the coherency of the model.

Now allow for a general $\Sigma \equiv VCov(\epsilon_{iT}, \cdots, \epsilon_{i2}) = CC'$. Given that we assume Gaussianity and dropping the i index, we obtain:

$$y_T = \mathbf{1}(\lambda y_{T-1} + x_T\beta + c_{T1}\nu_1 + c_{T2}\nu_2 + \cdots + c_{T,T-1}\nu_{T-1} + c_{TT}\nu_T > 0)$$

$$y_{T-1} = \mathbf{1}(\lambda y_{T-2} + x_{T-1}\beta + c_{T-1,1}\nu_1 + c_{T-1,2}\nu_2 + \cdots + c_{T,T-1}\nu_{T-1} > 0)$$

$$\vdots$$

$$y_2 = \mathbf{1}(\lambda y_1 + x_2\beta + c_{22}\nu_2 + c_{21}\nu_1 > 0)$$

$$y_1 = \textbf{exogenous}$$

Given a random draw of $\nu_{i1}, \cdots, \nu_{iT}$, an unambiguous rule gives sequentially $y_{i1} \rightarrow y_{i2} \rightarrow \cdots y_{i,T-1} \rightarrow y_{iT}$. Hence, the above defines a recursive DGP which establishes the coherency of the model.

Finally, consider the more general case when y_{i1} cannot be assumed as exogenous. We then supplement the system with an initial condition equation:

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + u_{i1}) > 0 \tag{16}$$

The following remarks are in order: First note that (16) is a generalization of the [Barghava and Sargan, 1982] approach. Second, one-factor random effect assumptions will imply an equicorrelated block structure on the top left $T-1 \times T-1$ block of Σ, while more general assumptions of one-factor random effects *combined with* an AR(1) or ARMA(p,q) processes for each ϵ error implies that Σ combines equicorrelated and Toeplitz-matrix parts. The last row and last column of Σ giving the variance of u_{1i} and its covariances with all ϵ_{it} allow the flexibility stipulated by [Heckman, 1981a]. The only modification now necessary is to change the initial condition equation to:

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + c_{11}\nu_{i1} > 0)$$

This recursive representation again establishes the coherency of the model: given a random draw of $\nu_{i1}, \cdots, \nu_{iT}$, an unambiguous DGP rule can be defined to establish sequentially $y_{i1} \to y_{i2} \to \cdots y_{i,T-1} \to y_{iT}$.

2.8.2 Dynamic Model 2: Bivariate Panel Data Probit with State Dependence

Parameter mnemonics:

- Exogenous variable coefficients: β, θ
- Simultaneous interaction terms: γ, δ
- Own state dependence: λ_y, λ_w
- Cross state dependence: ζ_w, ζ_y

$$\begin{aligned} y_{it} &= \mathbf{1}(x'_{i,t}\beta + \lambda_y y_{i,t-1} + \gamma w_{it} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0) \\ w_{it} &= \mathbf{1}(z'_{it}\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0) \end{aligned}$$

$$\begin{aligned} y_{iT} &= \mathbf{1}(x'_{iT}\beta + \lambda_y y_{i,T-1} + \gamma w_{iT} + \zeta_w w_{i,T-1} + \epsilon_{iT} > 0) \\ w_{iT} &= \mathbf{1}(z'_{iT}\theta + \lambda_w w_{i,T-1} + \delta y_{iT} + \zeta_y y_{i,T-1} + u_{iT} > 0) \end{aligned}$$

$$\begin{aligned} y_{i,T-1} &= \mathbf{1}(x'_{i,T-1}\beta + \lambda_y y_{i,T-2} + \gamma w_{i,T-1} + \zeta_w w_{i,T-2} + \epsilon_{i,T-1} > 0) \\ w_{i,T-1} &= \mathbf{1}(z'_{i,T-1}\theta + \lambda_w w_{i,T-2} + \delta y_{i,T-1} + \zeta_y y_{i,T-2} + u_{i,T-1} > 0) \end{aligned}$$

$$\vdots$$

$$\begin{aligned} y_{i2} &= \mathbf{1}(x'_{i2}\beta + \lambda_y y_{i1} + \gamma w_{i2} + \zeta_w w_{i1} + \epsilon_{i2} > 0) \\ w_{i2} &= \mathbf{1}(z'_{i2}\theta + \lambda_w w_{i1} + \delta y_{i2} + \zeta_y y_{i1} + u_{i2} > 0) \end{aligned}$$

$$y_{i1}$$
$$w_{i1}$$

Lemma 1: *Without any restrictions on the γ, δ parameters or the distribution of (ϵ, u), the General Bivariate PD Probit Model with State Dependence above is not coherent.*

Proof:

$$\begin{aligned} y_{it} &= \mathbf{1}(x_{it}'\beta + \lambda_y y_{i,t-1} + \gamma w_{it} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0) \\ w_{it} &= \mathbf{1}(z_{it}'\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0) \end{aligned}$$

Given y, w from period $t-1$, the λ and ζ terms are determined on the latent variable terms for period t (defining the event arguments of the indicator functions).

Together with unrestricted values of the random shocks and the exogenous variables of period t, everything in the event conditions is determined, except the simultaneous interaction terms γ, δ.

- But since the interaction terms appear both as conditioning variables on the RHS as well as dependent variable dummies on the LHS, they cannot be determined unambiguously. Hence, no complete DGP can be defined from ϵ, u to y, w.

Theorem 2: *The General Bivariate PD Probit Model with State Dependence above is coherent without any restrictions on the λ, ζ state dependence parameters or the distribution of (ϵ, u), if the simultaneous interaction terms satisfy $\gamma \cdot \delta = 0$, i.e., the model is triangular.*

Proof: Assume that $\gamma \cdot \delta = 0$ because $\gamma = 0$.

$$\begin{aligned} y_{it} &= \mathbf{1}(x_{it}'\beta + \lambda_y y_{i,t-1} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0) \\ w_{it} &= \mathbf{1}(z_{it}'\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0) \end{aligned}$$

Given y, w from period $t-1$, the λ and ζ terms are determined on the latent variable terms for period t (defining the event arguments of the indicator functions). Together with unrestricted values of the random shocks and the exogenous variables of period t, everything in the event condition of the y_t is determined, since there no simultaneous interaction term is present on the RHS (as $\gamma = 0$).

Entering the y_t value in the interaction term on the RHS of the w_t equation, everything in its event condition is now determined, which fixes w_t.

Hence, a complete DGP can be defined sequentially from the errors to the observables: $y_{i1}, w_{i1} \to y_{i2}, w_{i2} \to \cdots y_{i,T-1}, w_{i,T1} \to y_{iT}, w_{iT}$.

The proof for the $\delta = 0$ case is perfectly symmetric and will not be repeated.

Theorem 3: *The General Bivariate PD Probit Model with State Dependence above is coherent without any restrictions on the λ, ζ state dependence parameters, if:*

(i) the simultaneous interaction terms are of opposite signs, i.e., $\gamma \cdot \delta < 0$ *and*

(ii) the distribution of (ϵ, u) *satisfies* $F(\epsilon_t, u_t|\epsilon_{-t}, u_{-t}) = F(\epsilon_t, u_t|\epsilon_{<t}, u_{<t})$ *and the error r.v.s* (ϵ, u) *are restricted on rectangular regions that are determined recursively.*

- **Proof:** Assume that $\gamma \cdot \delta < 0$ because $\gamma < 0, \delta > 0$.

$$\begin{aligned} y_{it} &= \mathbf{1}(x_{it}'\beta + \lambda_y y_{i,t-1} + \gamma w_{it} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0) \\ w_{it} &= \mathbf{1}(z_{it}'\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0) \end{aligned}$$

Given y, w from period $t-1$, the λ and ζ terms are determined on the latent variable terms for period t (defining the event arguments of the indicator functions).

Given the exogenous variables of period t, the event conditions of y_t, w_t are determined except (a) the interaction terms γ, δ and (b) the error terms.

In the absence of condition (ii), the model would exhibit "empty region incoherency" as defined above. Employing the graphical approach of the Static Bivariate Probit above, defines the necessary rectangular exclusion region (drawn white) for the support of the truncated Gaussian:

$$\begin{aligned} 0 &< \epsilon_{it} + x_{it}'\beta_t + \lambda_y y_{i,t-1} + \zeta_w w_{i,t-1} < -\gamma \\ \delta &< u_{it} + z_{it}'\theta + \lambda_w w_{i,t-1} + \zeta_y y_{i,t-1} < 0 \end{aligned}$$

Based on the underlying uniform rv's drawn at the start of the DGP, the truncated Gaussian ϵ, u are drawn to satisfy the identifying rectangle restrictions using the probability integral transform method defined in the Monte Carlo section below.

Hence, the model under conditions (i) and (ii) is coherent, since a complete DGP could be defined sequentially from the errors to the observables: $y_{i1}, w_{i1} \rightarrow y_{i2}, w_{i2} \rightarrow \cdots y_{i,T-1}, w_{i,T1} \rightarrow y_{iT}, w_{iT}$.

The proof for $\gamma \cdot \delta < 0$ because $\gamma > 0, \delta < 0$ is exactly symmetric and will not be repeated.

2.8.3 Extensions to Bivariate Multinomial Ordered Probit

[Hajivassiliou, 2007] discusses how to extend the analysis to the case of two simultaneous (bivariate) *ordered probit equations with multiple regions*. We refer the interested reader to that study.

3 Empirical Application: Quantifying the Interactions between Financial Constraints and Firm Innovation

A large strand of the theoretical literature shows how investment is affected by informational asymmetries about the quality of the investment to be financed or relating to the behaviour of entrepreneurs. Such imperfections increase the cost of external finance and therefore, firms may be credit constrained. Due to their specificities inducing large informational asymmetries, high risk in terms of probability of failure, unpredictability in R&D returns and poor collateral, innovative firms are more likely to face agency issues and to be financially constrained ([Holmstrom, 1989]). Most of earlier papers study the link between firm level financial factors and R&D investment and obtain mixed evidences of such binding constraints on innovation (e.g., [Brown et al., 2009], [Brown et al., 2012], and see [Lerner and Hall, 2010] for a survey). Empirical evidence of the impact of financial constraints on the behaviour of firms is however not easy to obtain, essentially because the notional demand of firms for external finance is not observed directly (see [Hottenrott and Peters, 2012] for a test based on the use of a hypothetical payment received by the firm).

3.1 Direct Measures of Innovation and Financial Constraints

Instead, in this paper the existence of constraints is not deduced indirectly through the common arguments above nor identified through changes in financing supply conditions, but is directly measured by employing real data on the encountering of binding financing constraints as reported by firms in surveys by the European Union, as well as in a French survey about the financing of innovation. See Data Appendix 2 for details.

Due to the serious drawbacks of indirect approaches, direct measures of financial difficulties reported by firms can be useful, but very few surveys collect such information[11]. For instance, [Guiso, 1998] uses a direct qualitative measure given by a survey run by Banca d'Italia. In this paper, a firm is characterized as credit constrained "if at the rate of interest prevailing in the loan market, it would like to obtain a larger amount of loans but cannot". Such a precise definition of credit constrained firms is obtained thanks to the survey used where three questions are asked about access to credit (i) whether at the current market interest rate the firm wish a larger amount of credit, (ii) whether the firm would be willing to obtain more credit, (iii) whether the firm has applied for credit but has been turned down by the financial intermediary. Thanks to this information, the probability to be credit constrained is estimated which leads to the finding that low-tech firms are less likely to be financially constrained than high-tech firms. [Hottenrott and Peters, 2012] rely on hypothetical questions in a firm survey where firms are asked to imagine that they receive additional cash exogenously and indicate how they would spend it.

[11] We are currently investigating the availability of such direct measures of overall financing constraints from other sources, notably the Banque de France and the French National Institute of Statistical and Economic Studies (INSEE).

In the survey we use (FIT, Financement de l'Innovation Technologique) firms are asked whether some of their innovative projects were delayed, abandoned or non started because of (i) unavailability of new financing, (ii) searching and waiting for new financing, (iii) too high cost of finance. We define as financially constrained firms with hampered innovative projects because of one these three reasons so that our direct indicator of financial constraints takes into account both quantity rationing and higher cost of finance.[12]

3.2 Empirical Application

Using the econometric machinery developed in Section 3 that allows us to estimate joint binary probit equations with interaction terms on both sides, we apply those methods to the key issue of Being Innovative vs. Binding Financing Constraints interactions.

We take as our starting point the results obtained by [Savignac, 2008] who studied the impact of financial constraints on the decision to innovate by investigating the impact of financial constraints on innovation through a recursive model that did not allow for the probability of a binding finance constraint to depend on whether or not the firm is innovative. The propensity to innovate is explained by the traditional determinants of innovation exposed above (firm size and market power, technology push, latent consumer demand) and we account for financial constraints thanks to our qualitative indicator reflecting the financial difficulties encountered by firms to conduct their innovative projects.[13]

In sum, we model the probability that a firm decides to be innovative as:[14]

[12] For summary descriptive statistics, see Table 3 in the Data Appendix. See also Table 4 which confirms that the financial constraint indicators correlates strongly with the financial health of firms.

[13] For the importance of endogeneity in this setting, see [Mohnen and Roller, 2005]) for an example of another study that finds the “paradox” of a positive correlation between financial constraints and innovation .

[14] Main determinants of the propensity for a firm to innovate are known to be its size, its market power and its environment ([Cohen and Levin, 1989]).

The positive correlation between innovation and firm size is largely exposed in the literature (see [Cohen and Klepper, 1996]). Large firms can amortize sunk costs caused by their innovative activities by selling more units of output than smaller firms. In addition, they are able to diversify the risk incurred by innovation by running simultaneously several investment projects at the same time. And finally, large established firms are less likely to be financially constrained as they are able to generate cash-flow and to raise external funds.

At the present time, the link between innovation and competition is not well established both from the theoretical and empirical point of view. The Schumpeterian theory argues that market power and innovation are positively correlated whereas Arrow's theory shows that the gains to innovate are larger in an ex-ante competitive market. Recently, [Aghion et al., 2005] try to solve this puzzle and propose an inverted U shape relationship between innovation and competition : in a competitive environment, firms are incited to innovate to gain market power and increase their profits, but when competition becomes hard, the followers can be discouraged to innovate.

Other factors affecting innovative behaviour are driven by the firm environment. The technologic push that results from a lot of various factors such as the state of art, the process of diffusion of knowledge, the connections to academic research centers, etc., leads firms to develop or adopt innovative processes and products. The decision to innovate is also linked to the latent consumer demand perceived by the firm. (See among others [Crepon et al., 1998] or [Raymond et al., 2010] for

$$\text{Prob(Innovate?)} = \texttt{f}(\text{Financial Constraints, Size, Market Power, Technological Opportunities, Latent Consumer Demand for New Products, ...}) \quad (17)$$

We model the investment outcomes as depending on the discrete *outcome* from the other side (equation **??**), instead of on the continuous latent variable of the other side (equation **??**).[15] Nevertheless, with a simple probit, a surprising significant positive effect of the financial constraint is obtained ((***REFER TO FS PAPER AND OLD VERSIONS ON LINE***)). This positive effect is explained by two sources of bias: a selection bias due to firms not wishing to innovate, which we studied elsewhere and a problem of simultaneity between investment and financing decisions that we tackle below.[16]

To close the system, we now define also the probability of a binding financing constraint, which is assumed to have as an important determinant the (binary) decision of whether or not the firm chooses to be innovative:

$$\text{Prob(Binding Financing Constraint?)} = \text{f(Innovation?, Size, Guarantees or Collateral, Profit Margin, Banking Debt Structure, Internal Financing, ...)} \quad (18)$$

The key idea modelled by this equation is that prospective lenders will try to assess the creditworthiness of the applicant firm in the face of incomplete information. In particular, they do not know the precise riskiness of assets so they attempt to infer that using observable characteristics of firms. In the face of such uncertainty, it makes sense for lenders to be more cautious granting loans to innovative

empirical research on the firm level determinants of innovation).

[15] See the discussion of this issue in subsection ??? above.

[16] We conduct some robustness checks concerning the definition of the explanatory variables. Here, size is measured by the log of the number of employees to account for non-linear effects. We also test total assets as a measure of firm size both in terms of log and in level with a squared term. All these definitions lead to similar results. To distinguish between financial and *economic* distress it may be useful to add other liquidity measures such as available cash stock. However, as this variable is strongly correlated with profit margins it is redundant and does not provide significant results when it is introduced together with profit margins. Belonging to a holding group is likely to help innovation or to lead to less financial distress but we were not able to find such significant effects with our data, probably because they do not allow to account for *the size of the holding group* nor to identify the financial connections within the holding group. Finally, the likely endogeneity of the profitability and asset asset tangibility measures needs to be adressed.

firms since they present a higher inherent (but not directly observed) risk.

Such a system can be formulated as follows:

$$I_i = \begin{cases} 1 & if \quad I_i^* \equiv x_i^I \beta^I + \gamma F_i + \epsilon_i^I > 0 \\ 0 & if \quad I_i^* \equiv x_i^I \beta^I + \gamma F_i + \epsilon_i^I \leq 0 \end{cases} \tag{19}$$

$$F_i = \begin{cases} 1 & if \quad F_i^* \equiv x_i^F \beta^F + \delta I_i + \epsilon_i^F > 0 \\ 0 & if \quad F_i^* \equiv x_i^F \beta^F + \delta I_i + \epsilon_i^F \leq 0 \end{cases} \tag{20}$$

The econometric specifications we estimate below belong in three main groups. The first group contains *recursive specifications*, which ignore the possibility that the propensity to innovate may be affected by financial constraints. The second group allows for *reverse interactions*, whereby a firm undertaking actively innovative activities raises significantly the probability of it encountering a binding financing constraint, possibly because potential lenders are particularly wary of granting loans to firms of such type because of the extra riskiness involved. The third group of estimated specifications, investigates *state dependence* in financing and innovation experiences of firms:the nature of the available datasets can be exploited to study whether, ceteris paribus, past financial distress or innovation failures can affect a firm's current experiences in these two dimensions.

Though the surveys about innovation we use are not truly longitudinal "panel" sets, the information we use was collected in multiple biennial waves. Hence, we know whether a particular firm i has reported binding financing constraints in the past. Similarly, we can also tell whether a firm has failed in the past in its efforts to be innovative. Consequently, though our dataset is not a true panel, we can extend our econometric equations modelling the probabilities of being innovative and of encountering binding financing constraints at the end of the sample period to condition also on past experiences in these two dimensions. See the Data Appendix 3 for details about the waves employed here.

The sample design of the survey does not allow to observe every firm in the seven waves, therefore we restrict our "longitudinal" panel to two waves in order to limit the reduction of the sample size when merging the waves (see the transition tables in the Data Appendix).

The most general specification that we estimate below, accounting for both reverse and dynamic effects, is:

$$I_{it} = \begin{cases} 1 & if \quad I_{it}^* \equiv \alpha^I I_{it-1} + x_{it}^{I\prime} \beta^I + \gamma_0 F_{it} + \gamma_1 F_{it-1} + \epsilon_{it}^I > 0 \\ 0 & if \quad I_{it}^* \equiv \alpha^I I_{it-1} + x_{it}^{I\prime} \beta^I + \gamma_0 F_{it} + \gamma_1 F_{it-1} + \epsilon_{it}^I \leq 0 \end{cases} \tag{21}$$

$$F_{it} = \begin{cases} 1 & if \quad F_{it}^* \equiv \alpha^F F_{it-1} + x_{it}^{F\prime} \beta^F + \delta_0 I_{it} + \delta_1 I_{it-1} + \epsilon_{it}^F > 0 \\ 0 & if \quad F_{it}^* \equiv \alpha^F F_{it-1} + x_{it}^{F\prime} \beta^F + \delta_0 I_{it} + \delta_1 I_{it-1} + \epsilon_{it}^F \leq 0 \end{cases} \tag{22}$$

3.3 Comparative Analysis of Alternative Specifications

MUST REORGANIZE CAREFULLY

Table 1: Comparative Summary of Empirical Results

All Firms	Model 1 Triangular, Exogenous FC	Model 2 Triangular, Endogenous FC	Model 3 Full Joint Static	Model 4 Full Joint Dynamic
No.of Waves	One	One	One	Two
No.of Firms	1940	1940	1940	1512
FULL RESULTS:	Table ??.a	Table ??.a	Table 9.a	Table 10.a
INNOVATION EQUATION				
Size	0.33***	0.305	0.183***	0.256***
γ (FC dummy)	0.55***	-0.555***	-0.324**	-0.447***
Market Share	-0.01	-0.001***	0.020	0.027
Innov_{t-1}	—	—	—	0.829***
FinCon_{t-1}	—	—	—	0.301
avg $\widehat{P_I}$	0.418	0.418	0.418	0.543
$\widehat{P_I}$: $F = 0$	0.384	0.453	0.438	0.554
$\widehat{P_I}$: $F = 1$	0.601	0.250	0.316	0.377
$\%\widehat{\Delta P_I}$: $F = 0 \to 1$	56.42	-44.72	-27.93	-31.82
FINANCE EQUATION				
Size	-0.054	-0.002	-0.016	0.035
δ (Innov dummy)	—	—	0.647***	0.627***
Collateral	0.067	0.030	0.030	0.003
Banking Debt Ratio	0.010***	0.010***	0.015***	0.005
Own Financing Ratio	-0.003**	-0.003***	-0.001**	-0.008***
Profit Margin	-0.007**	-0.008***	-0.002***	-0.007***
Innov_{t-1}	—	—	—	0.236**
FinCon_{t-1}	—	—	—	0.135***
avg $\widehat{P_F}$	0.160	0.160	0.160	0.060
$\widehat{P_F}$: $I = 0$	0.160	0.160	0.103	0.029
$\widehat{P_F}$: $I = 1$	0.160	0.160	0.268	0.102
$\%\widehat{\Delta P_F}$: $I = 0 \to 1$	0.0	0.0	160.62	252.58
corr(Innov,FinCons)	—	0.572***	0.132**	0.500**
LogLikFunction	-1060-803=-1863	-1853	-1712	-1331 (-1706 imputed)

NOTES:

1. ***=significant at 1%; **=significant at 5%; *=significant at 10%.
2. Industry dummies (11) included in both Innovation and Financial Constraint equations.

Table 1 summarizes succinctly our key empirical results and presents the calculated direct and reverse effects that we obtained. To recapitulate, Models 1 and 2 adopt the existing approach of the past literature of forcing the econometric specification to be triangular with financial constraints allowed to affect the innovation decision, while the financial constraint outcome is assumed independent of being innovative or not. Models 3 and 4 allow the binary interactions on both the Innovation and the Finance sides, proved to be coherent through our Coherency analysis based on prior sign restrictions and estimated through the CMLE approach developed and analyzed in this paper. Model 3 is static with a single cross-section of firms, while Model 4 uses a two-wave, dynamic panel data set.

In summary, our findings in Tables 10.a and 10.b confirm the strong importance of such dynamic terms and establish very significant positive state dependency in our models.

First, our results show that firms tend to innovate persistently rather than occasionally.

Second, past financial difficulties are positively correlated with current binding financial constraints. As we take into account the experience of a firm concerning innovation, the state dependence of financial constraints seems particularly interesting. Indeed, firms currently implementing innovative projects as well as firms with innovative experience in the past are more likely to find it difficult to finance their current projects.[17]

Third, the probability for a firm to be currently conducting an innovative project is negatively impacted by the current financing difficulties as found in the static regressions but also positively correlated with financing constraints encountered in the past. One possible explanation for this positive correlation could be that financial difficulties mainly impact the beginning of the projects so that innovative projects that were initially hampered by financial difficulties are more likely to be continued when they become more mature. However, additional information on the stage of development of the innovative projects and on their duration would be necessary to further investigate this point. In particular, we are not able to identify whether the firms were continuing in Wave 2 (1997-1999) with projects that were already conducted in Wave 1 (1994-1996).

Apart from the coefficient estimates for the most important exogenous explanatory variables, Table 1 presents also: γ, the coefficient for the financial constraint dummy when entered in the Innovation side; δ, the coefficient for the Innovation dummy entered in the Finance side; and for the dynamic Model 4, the coefficients for the lags of the Finance and Innovation dummies. First, note how seriously misleading conclusions were reached by the early strands of the literature, that inappropriately

[17] An important issue discussed frequently in the econometrics literature is the possibility that state dependence may not be an important factor *per se*, but it might appear statistically significant if persistent heterogeneity among individual economic agents is ignored. As [Heckman, 1981a] shows, the two can be identified when a panel data set with more than two waves per individual is available. Since our dynamic sample consists only of two waves, we need to acknowledge the possibility that the strong state dependence we report here may be compounded by unobserved persistent heterogeneity that is not accounted for explicitly.

ignored the endogeneity of the Finance dummy: doing so yields a completely counter-intuitive γ estimate of $+0.55$, implying that finance constraints *raise* significantly the likelihood of innovation — while treating the Finance constraint dummy as endogenous gives a range of negative estimates from -0.32 to -0.56. Second, since Models 1 and 2 impose a triangular specification, they do not estimate δ coefficients for the Innovation dummy in the Finance equation. In contrast, our two CMLE models give statistically very significant δ estimates of over 0.6 — as we expected a priori, being innovative raises the probability of facing a binding finance constraint. In the dynamic Model 4, we find very significant state dependence over the two periods of the panel — note the statistical significance of three of the four lagged dummies entered as regressors.

In order to quantify the importance of our interaction findings about γ and δ, we present in Table 1 four estimated probability calculations for each of the two I and F sides: (a) $avg\ \widehat{P_I}$, the average probability of undertaking innovation; (b) $\widehat{P_I} : F = 0$, the estimated probability of Innovation given the Finance constraint is not binding; (c) $\widehat{P_I} : F = 1$, the estimated probability of Innovation given a binding Finance constraint; and (d) $\%\widehat{\Delta P_I} : F = 0 \rightarrow 1$, the percentage change in the estimated probability of changing from $F = 0$ to $F = 1$, while keeping everything else unchanged. For the finance side, the analogous four quantities are: $avg\ \widehat{P_F}$, $\widehat{P_F}$: $I = 0$, $\widehat{P_F}$: $I = 1$, and $\%\widehat{\Delta P_F}$: $I = 0 \rightarrow 1$.

Our estimated probability results are quite striking: for the Innovation equation, we find that *when a firm faces a binding finance constraint, the probability of being innovative falls ceteris paribus by* $30\% - 40\%$ depending on the version.[18] Moving to the Finance constraint side, the magnitudes of the results are even more impressive: *all other things equal, a firm being innovative more than doubles the probability of a binance Finance constraint.*[19]

4 Conclusions

In this paper we developed two novel methods for establishing coherency conditions in LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. The first is based on a graphical characterization and the second is based on a hypothetical Monte-Carlo Data Generating Process (DGP) approach. Our novel methods have intuitive interpretations and are easy to implement and generalize. A constructive consequence of the new approaches is that they indicate how to achieve coherency in models traditionally classified as incoherent through the use of prior sign restrictions on model parameters. This allowed us to develop estimation strategies based on CMLE for simultaneous LDV models without imposing recursivity. The proposed CMLE methodology was evaluated through an extensive set of Monte-Carlo experiments. The experiments allowed us also to study the conse-

[18] In the erroneous Model 1 that ignores the endogeneity of the Finance constraint, the probability of innovation is predicted to *rise by over* 50% as a result of a binding Finance constraint!

[19] Since Models 1 and 2 do not allow for reverse interactions by excluding the I dummy from the F side, they imply $\%\widehat{\Delta P_F}$: $I = 0 \rightarrow 1$ equal to zero.

quences of employing estimators that make overly restrictive coherency assumptions about the DGP. The findings confirmed very substantive improvements by employing the CMLE developed in this paper in terms of estimation Mean-Squared-Error.

Through the CMLE novel approach, we analyzed the existence and impact of financing constraints as a possibly serious obstacle to innovation by firms. We were able to quantify the interaction between financing constraints and a firm's decision and ability to innovate without forcing the econometric models to be recursive. Direct measures of financing constraints were employed using survey data, which helped us overcome the problems with the traditional approach in the past literature of trying to deduce the existence and impact of financing constraints through the significance of firm wealth variables. We thus obtained direct as well as reverse interaction effects, leading us to conclude that binding financing constraints discourage innovation and at the same time innovative firms are more likely to face binding financing constraints. The empirical results we obtained using CMLE were quite striking: ceteris paribus, we found that a firm facing a binding finance constraint is approximately *30% less likely* to undertake innovation, while the probability that a firm encounters a binding finance constraint *more than doubles* if the firm is classified as innovative. In addition, we investigated the importance of state-dependence in dynamic versions of our models and concluded that such issues are critical if direct and reverse interactions between innovation and financing constraints are to be quantified reliably.

References

[Aghion et al., 2005] Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation : An inverted u relationship. *The Quarterly Journal of Economics*, 120(2):701–728.

[Audretsch, 1995] Audretsch, D. (1995). Innovation, growth and survival. *International Journal of Industrial Organization*, 13(4):441–457.

[Barghava and Sargan, 1982] Barghava, A. and Sargan, J. (1982). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica*, 51:1635–1660.

[Brown et al., 2009] Brown, J., Fazzari, S., and Petersen, B. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s r&d boom. *The Journal of Finance*, 64:151–185.

[Brown et al., 2012] Brown, J., Martinsson, G., and Petersen, B. (2012). Do financing constraints matter for r&d? *European Economic Review*, 56(8):1512–1529.

[Cohen and Klepper, 1996] Cohen, W. and Klepper, S. (1996). A reprise of size and r&d. *Economic Journal*, 106:925–951.

[Cohen and Levin, 1989] Cohen, W. and Levin, R. (1989). Empirical studies of innovation and market structure. In *Handbook of Industrial Organization*, volume chapter 18, pages 1060–1107, Boston. North-Holland.

[Crepon et al., 1998] Crepon, B., Duguet, E., and Mairesse, J. (1998). Research, innovation and productivity: an econometric analysis at the firm level. *Economics of Innovation and New Technology*, 7(2):115–158.

[Dagenais, 1997] Dagenais, M. (1997). A simultaneous probit model. GREQAM working paper.

[Falcetti and Tudela, 2008] Falcetti, E. and Tudela, M. (2008). What do twins share? a joint probit estimation of banking and currency crises. *Economica*, 75:199Ű221.

[Gourieroux et al., 1980] Gourieroux, C., Laffont, J., and Monfort, A. (1980). Coherency conditions in simultaneous linear equations models with endogenous switching regime. *Econometrica*, 48(3):75–96.

[Guiso, 1998] Guiso, L. (1998). High tech firms and credit rationing. *Journal of Economic Behaviour and Organization*, 35:39–59.

[Hajivassiliou, 2002] Hajivassiliou, V. (2002). Novel approaches to coherency conditions in ldv models. LSE Department of Economics working paper.

[Hajivassiliou, 2007] Hajivassiliou, V. (2007). Efficient and robust estimation of panel data ldv models with simultaneity, dynamics, and regressor-random effects correlation. LSE Department of Economics working paper.

[Hajivassiliou, 2008] Hajivassiliou, V. (2008). Novel approaches to coherency conditions in ldv models: Theory and monte-carlo experiments. LSE Department of Economics working paper, URL=http://econ.lse.ac.uk/staff/vassilis/papers/.

[Hajivassiliou and Ioannides, 2007] Hajivassiliou, V. and Ioannides, Y. (2007). Unemployment and liquidity constraints. *Journal of Applied Econometrics*, 22(3):479–510.

[Hajivassiliou and McFadden, 1998] Hajivassiliou, V. and McFadden, D. (1998). The method of simulated scores for the estimation of ldv models. *Econometrica*, 66(4):863–896.

[Heckman, 1981a] Heckman, J. (1981a). Dynamic discrete models. In *Structural Analysis of Discrete Data with Econometric Applications*, pages 179–195, Cambridge, Massachusetts.

[Heckman, 1981b] Heckman, J. (1981b). Statistical models for discrete panel data. In *Structural Analysis of Discrete Data with Econometric Applications*, pages 114–178, Cambridge, Massachusetts.

[Holmstrom, 1989] Holmstrom, B. (1989). Agency costs and innovation. *Journal of Economic Behavior and Organization*, 12(3):305–327 1989.

[Hottenrott and Peters, 2012] Hottenrott, H. and Peters, B. (2012). Innovative capability and financing constraints for innovation: More money, more innovation? *Review of Economics and Statistics*, 94(4):1126–1142.

[Klette and Kortum, 2004] Klette, T. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018.

[Lerner and Hall, 2010] Lerner, J. and Hall, B. (2010). The financing of r&d and innovation. In *Handbook of the Economics of Innovation*, volume chapter ?, vol.1, pages 1129–1155, Boston. North-Holland.

[Lewbel, 2005] Lewbel, A. (2005). Coherency of structural models containing a dummy endogenous variable. Boston College Working Papers in Economics, no.456.

[Lewbel, 2007] Lewbel, A. (2007). Coherency and completeness of structural models containing a dummy endogenous variable. *International Economic Review*, 48(4), November:1379–1392.

[Lhomme, 2002] Lhomme, Y. (2002). Le financement de l'innovation technologique dans l'industrie, références-chiffres clé no.217. Technical report, Ministère de l'économie, des finances et de l'industrie, Paris.

[Mairesse and Mohnen, 2010] Mairesse, J. and Mohnen, P. (2010). Using innovation surveys for econometric analysis. In *Handbook of the Economics of Innovation*, volume chapter 26, vol.2, pages 1129–1155, Boston. North-Holland.

[Mohnen and Roller, 2005] Mohnen, P. and Roller, L.-H. (2005). Complementarities in innovation policy. *European Economic Review*, 49:1431–1450.

[OECD and Eurostat, 1997] OECD and Eurostat (1997). Oslo-manual, proposed guidelines for collecting and interpreting technological innovation data. Technical report, Organization for Economic Cooperation and Development, Paris.

[Raymond et al., 2010] Raymond, W., Mohnen, P., Palm, F., and Schim van der Loeff, S. (2010). Persistence of innovation in dutch manufacturing:is it spurious? *The Review of Economics and Statistics*, 92(3):495–504.

[Savignac, 2008] Savignac, F. (2008). The impact of financial constraints on innovation: What can be learned from a direct measure? *Economics of Innovation and New Technology*, 17(6):553–569.

[Tamer, 2003] Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies*, 70:147–165.

[Uhlig, 2005] Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, vol.52(2):381–419.

5 Technical Appendix 1: Monte-Carlo Evidence on the Performance of CML

5.1 Overview

As we showed in Subsection 2.6, we obtain a coherent non-recursive model with interaction dummies included on both sides, provided we believe the feedback terms have opposite signs on the two sides. Note that it is sufficient to consider only the $\gamma \geq 0$, $\delta \leq 0$ case, since the reverse can always be subsumed by redefining both dependent binary variables to their complements $y'_{it} \equiv (1 - y_{it})$.

We performed extensive Monte Carlo experiments, designed to illustrate the consequences of adopting existing and novel estimation strategies for the problem of this paper. The experiments confirm that the CML approach under sign restrictions derived above provides reliable, consistent and efficient estimates of the underlying parameters including the two interaction terms. In contrast, the existing traditional approaches (unrestricted MLE ignoring possible incoherency and MLE that incorrectly assumes recursivity of the system) give seriously misleading and inconsistent results. The interested reader is referred to the online companion paper [Hajivassiliou, 2008] for an extensive presentation of the Monte Carlos summarized here and detailed analysis and findings.[20] Here we give a brief summary of the key findings:

- The Truncated CMLE proposed in this paper performs very satisfactorily, being the only consistent estimator for the reverse feedback cases, and with only small sacrifices in terms of efficiency in the recursive DGPs when it is not strictly necessary.
- The linear probability estimators perform very badly in all cases with endogenous interaction terms, thus suggesting that the inherent non-linearities of the

[20] The cited study considered *nine* estimation approaches:

(a) Incorrectly forcing the old coherency condition to hold, i.e., assuming recursivity when in fact both feedback terms are present (estimators `E-TRWN`=assuming $\delta = 0$ and `E-TRNW`=assuming $\gamma = 0$);

(b) unrestricted likelihood estimation, which ignores the resulting incoherency due to the empty or overlap region(s) (estimator `E-INCO`);

(c) restricted likelihood estimation conditioning on the data lying outside the empty region(s) of incoherency (estimators `E-SQPM`=assuming $(\gamma \geq 0, \delta \leq 0)$ and `E-SQMP`=assuming $(\gamma \leq 0, \delta \geq 0)$);

(d) restricted likelihood estimation conditioning on the data lying outside the overlap region(s) of incoherency (estimators `E-SQPM`=assuming $(\gamma \geq 0, \delta \geq 0)$ and `E-SQMP`=assuming $(\gamma \leq 0, \delta \leq 0)$); and

(e) LPOLS: (linear probability) ordinary least squares estimation of each binary probit equation ignoring the possible endogeneity of the interaction terms; and LP2SLS: applying two-stage least squares recognizing that the two interaction terms on the RHS of each probit equation can be endogenous.

In the cited study, six "true" models were generated, depending on whether interaction terms were allowed on one or both sides In each case, the nine estimators `E-TRWN`, `E-TRNW`, `E-INCO`, `E-SQPM`, `E-SQMP`, `E-SQPP`, `E-SQMM`, `LPOLS`, and `LP2SLS` were calculated.

Apart from confirming the excellent performance of the Truncated MLE approach adopted here, the study also confirmed that application of linear probability methods to the bivariate binary probit model typically leads to very unreliable findings, even if such methods attempt to take account of the endogeneity of the direct and reverse interaction effects.

bivariate probits cannot be safely ignored.

- Truncated CMLE also works well for the overlap region incoherency cases.
- Unrestricted likelihood estimation ignoring the resulting incoherency due to the empty or overlap region(s) is by far the worst performing estimator, dominated even by equation by equation univariate estimators which estimate the two equations separately while ignoring the other side of the model.[21]

5.2 Design of Monte-Carlo Experiments

The experiments were designed to illustrate the importance of coherency on the following *nine* estimation approaches:

(Est1&Est2) likelihood estimation that incorrectly forces the old coherency condition to hold, i.e., assuming recursivity when in fact both feedback terms are present (estimators `E-TRWN`=assuming $\delta = 0$ and `E-TRNW`=assuming $\gamma = 0$);

(Est3) unrestricted likelihood estimation, which ignores the resulting incoherency due to the empty or overlap region(s) (estimator `E-INCO`);

[21] Some further important features of [Hajivassiliou, 2008] are:

- The procedure for summarizing the Mean Squared Error (MSE) results was as follows: A given chart graphs the performance of each estimation algorithm in terms of a given estimation criterion (e.g., squared Root MSE (RMSE) etc.) with the best performing algorithm normalized to 100. The other methods are then given as a fraction of that best. For example, if method A is the best with RMSE=25, and methods B and C have RMSE equal to 75 and 125 respectively, method A will be reported as 100, B as 0.333 (one third as good since 3 times as high RMSE), and C as 0.20 (one fifth as good since 5 times as high RMSE).
- In a set of four-part figures, the *overall* RMSE results are presented with each method's performance averaged across all estimated parameters. The CMLE estimator dominates all other methods in impressive fashion when the true DGP possesses the opposite-signs restriction $\gamma \cdot \delta \leq 0$. It also performs very satisfactorily in case the true model is recursive, achieving almost as good a performance as the ideal recursive estimator for that case. Even in the case of no interaction terms being present in the true DGP ($\gamma \cdot \delta = 0$), the CMLE estimator loses out in terms of RMSE only because of the higher estimation variance in view of not imposing two true restrictions.
- Other sets of four-part figures give the relative RMSE performance for the δ interaction parameter, and of parameters β_{11}, β_{22}, and ρ respectively. CMLE also impresses in these sets of results in a similar ranking to the previous point.
- Similar figures summarize the *overall* results in terms of *absolute bias* instead of RMSE, as well as absolute *median* bias. The first set establishes that the CMLE estimator heads and shoulders above all the alternatives in terms of bias, and whenever it is less clearly the preferred estimator, this was only caused by higher estimation variance.

It may be noted that the dismal performance of the two estimators based on the Linear Probability (LP) approximation would have been alleviated had the average partial probability derivatives been calculated instead of the latent variable coefficients. This is because the LP estimators by construction have a constant probability derivative with respect to an explanatory variable, irrespective of the observation values. In our view, such calculations would not have been especially interesting since in most empirical LDV studies, investigators wish to allow for such probability derivatives to vary over the range of observations.

(Est4&Est5) restricted likelihood estimation conditioning on the data lying outside the empty region(s) of incoherency (estimators E-SQPM=assuming $(\gamma \geq 0, \delta \leq 0)$ and E-SQMP=assuming $(\gamma \leq 0, \delta \geq 0)$);

(Est6&Est7) restricted likelihood estimation conditioning on the data lying outside the overlap region(s) of incoherency (estimators E-SQPP=assuming $(\gamma \geq 0, \delta \geq 0)$ and E-SQMM=assuming $(\gamma \leq 0, \delta \leq 0)$).

(Est8&Est9) LPOLS: (linear probability) ordinary least squares estimation of each binary probit equation ignoring the possible endogeneity of the interaction terms; and LP2SLS: applying two-stage least squares recognizing that the two interaction terms on the RHS of each probit equation can be endogenous.

We generate six "true" models:

$$\begin{array}{llll} 1. & \mathtt{DGP}-\mathtt{TRWN}(\delta=0) & 4. & \mathtt{DGP}-\mathtt{SQMP}\ (\gamma \leq 0, \delta \geq 0) \\ 2. & \mathtt{DGP}-\mathtt{TRNW}(\gamma=0) & 5. & \mathtt{DGP}-\mathtt{SQPP}\ (\gamma \geq 0, \delta \geq 0) \\ 3. & \mathtt{DGP}-\mathtt{SQPM}\ (\gamma \geq 0, \delta \leq 0) & 6. & \mathtt{DGP}-\mathtt{SQMM}\ (\gamma \leq 0, \delta \leq 0) \end{array}$$

To simulate data from these six models, it is necessary to devise a methodology for generating standard Gaussian variates truncated to lie outside an interval $[\underline{\lambda}, \bar{\lambda}]$. The following algorithm achieves this: Let $z \sim N(0,1)$ and define $\tau \sim z|\{z \notin [\underline{\lambda}, \bar{\lambda}]\}$ Then $cdf(z) : F(z) = \Phi(z)$ and

$$cdf(\tau) : F(\tau) = \begin{cases} \frac{\Phi(z)}{1-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})} & \text{if } z < \underline{\lambda}, \\ \frac{\Phi(\underline{\lambda})}{1-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})} & \text{if } \underline{\lambda} < z \leq \bar{\lambda}, \\ \frac{\Phi(z)-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})}{1-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})} & \text{if } z > \bar{\lambda}. \end{cases}$$

The procedure is exact for a univariate z truncated on $\{z \notin [\underline{\lambda}, \bar{\lambda}]\}$, but it will not work for higher dimensions.[22]

Based on data generated from each of the six DGPs in turn, we calculate the nine estimators E-TRWN, E-TRNW, E-INCO, E-SQPM, E-SQMP, E-SQPP, E-SQMM, LPOLS, and LP2SLS.

The generating equations are:

$$\begin{aligned} ystar1 &= x1[nobs, kx1] * beta1 + gamma * y2 + eps1, \quad y1 = 1(ystar1 > 0) \\ ystar2 &= x2[nobs, kx2] * beta2 + delta * y1 + eps2, \quad y2 = 1(ystar2 > 0) \end{aligned}$$

where $x1$ is a $nobs \times kx1$ matrix and $x2$ is a $nobs \times kx2$ matrix.

5.2.1 Case DGP-TRWN: γ unrestricted, $\delta = 0$

$$\begin{aligned} ystar1 &= x1[nobs, kx1] * beta1 + gamma * y2 + eps1, \quad y1 = 1(ystar1 > 0) \\ ystar2 &= x2[nobs, kx2] * beta2 + eps2, \quad y2 = 1(ystar2 > 0) \end{aligned}$$

[22] For DGPs with higher dimensions, the leading alternative procedures are Acceptance-Rejection and Gibbs resampling — see [Hajivassiliou and McFadden, 1998] for discussion.

Given the recursivity of the $\gamma \cdot \delta = 0$ restriction in this case, $ystar2$ is generated first, which gives $y2$. This is then plugged into the RHS of the $ystar1$ equation thus allowing $ystar1$ and $y1$ to be obtained. The symmetric case `DGP-TRNW` is handled analogously.

5.2.2 Case `DGP-SQPM`: $\gamma \geq 0, \delta \leq 0$

$$0 \leq eps1 + x1 * b1 \leq gamma, -delta \leq eps2 + x2 * beta2 \leq 0 \tag{23}$$

Accept-reject methods are used to generate the data so that these restrictions are satisfied. The symmetric `DGP-SQMP` case is handled analogously. See also Appendix 1 for an exact algorithm for generating draws from truncated normal distributions restricted to lie on region (23).

5.2.3 Case `DGP-SQPP`: $\gamma \geq 0, \delta \geq 0$

Accept-reject methods are used to generate the data so that these restrictions are satisfied, as well as for the symmetric case `DGP-SQMM`.

5.3 Implementation of the Monte-Carlo Experiments

We performed 32 Monte-Carlo experiments, indexed by MCxyz as follows:

	δ	γ
x=1	0	0
x=2	0.8	0
x=3	0.8	1
x=4	0.8	-1

	$\rho_{\epsilon_1,\epsilon_2}$
y=1	0.3
y=2	-0.3

	x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}
$z = 1$	const	$\chi^2(1)$	Bernoulli(0.7)	const	x_{12}	DoubleExponentialSS
$z = 2$	const	$\chi^2(1)$	Bernoulli(0.9)	const	x_{12}	DoubleExponentialSS
$z = 3$	const	$\chi^2(1)$	Bernoulli(0.7)	const	x_{12}	DoubleExponentialLS
$z = 4$	const	$\chi^2(1)$	Bernoulli(0.9)	const	x_{12}	DoubleExponentialLS

where *DoubleExponential* stands for a Double Exponential distribution with mean 0 with asymmetric two sides, SS for "small skewness" and LS with "large skewness." Each random data set had 2000 observations and 200 Monte-Carlo replications were generated. The true beta parameters were set at: $\beta_1 = (\ 0.8, \ \ -0.5, \ \ -0.3\ )'$ and $\beta_2 = (\ -0.3, \ \ 0.7, \ \ -0.4\ )'$.[23] Table 2 defines the four regime probabilities and their row and column sums across the 32 Monte-Carlo experiments we performed.

Table 2: Monte Carlo Regime Probabilities

	$Y_2 = 1$	$Y_2 = 0$	
$Y_1 = 1$	p_{11}	p_{10}	$p_{1\cdot}$
$Y_1 = 0$	p_{01}	p_{00}	$p_{0\cdot}$
	$p_{\cdot 1}$	$p_{\cdot 0}$	

Across Experiments	p_{11}	p_{10}	p_{01}	p_{00}	$p_{0\cdot}$	$p_{1\cdot}$	$p_{\cdot 1}$	$p_{\cdot 0}$
minimum	0.120	0.141	0.065	0.166	0.266	0.333	0.444	0.382
average	0.318	0.220	0.217	0.245	0.462	0.538	0.535	0.466
maximum	0.552	0.330	0.410	0.307	0.667	0.734	0.618	0.556

[23] The full tables presenting the detailed Monte-Carlo results in terms of various estimation criteria (root-mean-squared error, absolute bias, absolute median bias, variance, interquartile range, and nine-decile range) can be found in the online companion paper [Hajivassiliou, 2008].

6 Appendix 2: Data Sources and Constructions

We use data from two main sources:

6.1 Sources

6.1.1 The FIT survey

The survey "Financement de l'Innovation Technologique" (FIT) was conducted in 2000 by the French Ministry of Industry, in order to obtain statistical information about the financing conditions of innovative projects of manufacturing firms in France.[24] The survey identifies the firms which undertook innovative projects between 1997 and 1999 and gives qualitative information about the financial constraints experienced by firms. 5500 manufacturing companies with 20+ employees were surveyed (excluding agricultural-food and building sectors). The response rate was 85% overall, and 100% among firms with 500+ employees. It is important to note that start-ups and new established firms were not included, which implies a global response rate of 70%.

As the Community Innovation Surveys (CIS), the FIT survey is based upon the technological innovation definition in the Oslo manual (OECD 1997) and is less restrictive than R&D expenditures or patents data.[25]

- **Definition 1: Innovative firms**

A firm is "innovative" if it has introduced or developed a product or process innovation (or was in process of doing so) during the surveyed period. This identification is built on at least one positive answer to the three questions:

1) In 1997, 1998 or 1999, did Your enterprise introduce onto the market any new or significantly improved products for Your enterprise?

2) In 1997, 1998 or 1999, did Your enterprise introduce onto the market any new or significantly improved process for Your enterprise?

3) In 1997, 1998 or 1999, did Your enterprise have projects of new or significantly improved products or processes:

- Which are not yet completed or not yet introduce to the market?

[24] See [Lhomme, 2002] for details.

[25] The Oslo manual definition was set up to overcome some shortcomings associated with R&D and patents. For instance, innovative activities are not systematically associated with R&D investments and patents are also strategic tools that are not necessarily used by firms to protect innovation. Moreover, the set of innovative firms according to the OECD definition expands for practical reasons, as we need to observe both the innovative behaviour of the firm and its assessment about financial difficulties.

The Community Innovation Surveys (CIS) are conducted in each country by the national statistical entities in order to collect information about the innovative activities of firms.They are based on the same harmonised questionnaire that may be completed at the national level by additional questions. The survey used here (Financement de l'Innovation Technologique, FIT) is different because it is focused on the financing of innovation. However, its methodological framework is the same as the well-known CIS' one, in particular concerning the definition of innovation and the structure of the questionnaire.

- Which were failures?

- **Definition 2: Financing constraints**

All surveyed firms had to answer the following question:

In 1997, 1998 or 1999, what are the obstacles that have prevented your firm to conduct or to start innovative projects (multiple answers possible)?

- Excessive perceived economic risk
- Lack of qualified personnel
- Innovation costs too high
- Lack of sources of finance
- Slowness in the setting up of the financing
- Too high interest rates of the financing
- Excessive get out clause in the shareholder agreement
- Lack of knowledge about ad hoc financial networks
- No obstacle

The firm had to assess the severity of each negative factor (seriously delayed, abandoned, or prevented to be started).

A firm is defined as **financially constrained** when it reported seriously delayed, abandoned or non-started projects because of:

- Too high interest rates of the financing; Lack of sources of finance; or Slowness in the setting up of the financing

To understand better the nature of our data, we should think of our full sample as containing two subgroups:

The first subgroup consists of firms that (a) are classified as innovative according to Definition 1 (i.e., wished and were able to be innovative) **or** (b) reported that they faced financial obstacles to innovation according to Definition 2, which prevented them from being innovative. The common characteristic of the first subgroup is that a firm *wished* to be innovative.

- In contrast, the second subgroup consists of firms that did not wish to be innovative and did not encounter any financial obstacles to innovation.

6.1.2 The Banque de France Balance Sheet Data set

In order to obtain information about the size, economic performance, and financing structure of firms, we use the Banque de France Balance Sheet Data set, or Centrale de Bilans (CdB). This is a database of detailed accounting data of all French companies with 500+ employees, as well as of a fraction of smaller firms, giving a total of around 34,000 companies. It covers about 57% of all industries (by employment), and gives detailed information on financing sources (group financing, internal, etc) and financing expenditures (intangible goods, services, etc.)

We have verified that the direct indicator reported by the firms is in line with the balance sheet data: firms without financial constraints exhibit a better profile than constrained firms in terms of financing structure, risk and economic performances.

Table 3: Details of the financial obstacles and their consequences

	% of Constrained Firms	Consequences for their Innovative Project(s)		
		delayed	abandoned	non started
Unavailability of new financing	87.74	46.27	10.45	46.27
Searching and waiting for new financing	43.23	35.29	12.13	57.72
Too high cost of finance	22.90	28.17	15.49	57.75

Table 4: Direct indicator and balance sheet ratios

	Constrained				Unconstrained			
	Q1	Q2	Q3	Mean	Q1	Q2	Q3	Mean
Nber of employees	47	112	290	249.7	47	102	243	227.6
Debt/Equity	8.4	50.7	147.9	132.7	7.9	3.4	9.2	55.0
- Long term bank debt/Equity	0.6	21.9	62.3	43.7	1.7	18.6	51.9	33.4
- Short term bank debt/Equity	0.1	15.4	73.5	89.1	0.0	3.7	34.9	21.6
EBITDA/Sales	6.6	15.4	25.4	10.8	11.8	20.8	30.5	20.3
Cash-flow/Total assets	2.9	7.3	11.1	5.8	5.2	8.5	12.2	8.8
Immaterial Inv/Value added	0.4	1.5	3.9	4.6	0.3	1.1	3.0	3.3

6.2 Cross-section vs. Dynamic (Panel) Samples

The cross-section sample results from the matching of FIT and CdB in the 1997-1999 wave, allowing us to recover about 60% of the FIT companies.

6.2.1 Our Cross-Section Sample — Wave 1997-1999. (1940 firms)

After some necessary cleaning, our sample contains 1940 firms.[26]. The distribution of the firms in our sample according to their innovative behaviour and financing obstacles is given in the table below:

[26] Of these, 1082 were in the subgroup of potentially innovative firms explained above. The manufacture of coke, refined petroleum products and nuclear fuel has been deleted because only two firms were present in the merged dataset. In addition, two firms with negative value added or with abnormally high investment rates have been excluded.

Table 5: Number of firms in the cross-section sample

Potentially innovative firms (1082)				Others
with innovative activities		without innovative activities		
financially constrained	financially unconstrained	financially constrained	financially unconstrained	
198	613	112	159	858

Sources : Centrale de Bilans[27] (Banque de France), FIT (French Ministry of Industry)

Table 6: Definition of the variables

Name	Definition
Dependent variable : y_{1_i}	=1 if the firm was innovative, =0 otherwise
Explanatory : x_{1i}	
Size	log (number of employees)
Market share	$\frac{\text{sales of the firm}}{\text{sales of the sector}} \times 100$
TP1	=1 if the firm's market is technologically not innovative (reference)
TP2	=1 if the firm's market is weakly innovative,
TP3	=1 if the firm's market is moderately innovative
TP4	=1 if the firm's market is strongly innovative
Financial constraints	=1 if the firm faced financial constraints, =0 otherwise
Financial constraints equation	
Dependent variable : y_{2_i}	=1 if the firm faced financial constraints, =0 otherwise
Explanatory : x_{2i}	
Size	log (number of employees)
Collateral	log(tangible assets)
Banking debt ratio	$\frac{\text{Banking debt}}{\text{(Own financing+Market Financing+Financial debt)}} \times 100$
Own financing ratio	$\frac{\text{Own financing}}{\text{(Own financing+Market Financing+Financial debt)}} \times 100$
Gross operating profit margin	$\frac{\text{EBIDTA}}{\text{Value added}} \times 100$

Sources : Centrale de Bilans (Banque de France), FIT (French Ministry of Industry) and EAE (INSEE)

Table 7: Descriptive statistics

Full sample of 1940 firms

	Mean	Std	Min	Max
Innovation	0.418	0.493	0	1
Size	4.783	1.107	2.890	9.716
Market share	0.177	0.566	0.001	16.15
TP1	0.139	0.312	0	1
TP2	0.416	0.493	0	1
TP3	0.348	0.476	0	1
TP4	0.097	0.297	0	1
Financial constraints	0.160	0.366	0	1
Collateral	71.048	22.698	4.241	302.444
Banking debt ratio	17.678	15.758	0	92.307
Own financing ratio	31.827	24.195	-609.459	90.136
Gross operating profit margin	18.248	19.416	-197.600	76.850

Sources : Centrale de Bilans (Banque de France), FIT (French Ministry of Industry) and EAE (INSEE)

6.2.2 Our Dynamic (Panel) Sample — Waves 1994-1996 and 1997-1999. (1512 firms)

The panel sample is obtained by matching the survey FIT with two other sources: (i) the second French wave of the Community Innovation Survey (CIS2) run by the French Ministry of Industry for Eurostat; and (ii) the balance sheet data set of the Banque de France (Centrale de Bilans). The FIT and CIS2 surveys ask the same questions to identify innovative firms; and very similar questions about financial constraints.[28] The sample obtained by matching FIT, CIS2 and CdB contains 1512 firms. The transitions for innovation and financial constraints between the two surveyed periods are reported in the tables below:

Legend:	col % Cell Count row %

Table 8: **Transitions**

Part A: Innov Transitions 1994-6 −> 1997-9

		1997-1999 (FIT)		
		$I_{it} = 1$	$I_{it} = 0$	Total
1994-1996 (CIS2)	$I_{i,t-1} = 1$	84.45 543 60.54	40.74 354 39.46	42.53 897 100
	$I_{i,t-1} = 0$	15.55 100 16.26	59.26 515 83.74	57.47 615 100
	Total	100 643 42.53	100 869 57.47	100 1512 100

[28] Unfortunately, the following wave of the Community Innovation Survey (CIS3) covering 1998-2000 does not include questions about financial constraints and therefore we cannot use it here.

Table 8, Part B: FinCons Transitions 1994-6 –> 1997-9

		1997-1999 (FIT)		
		$F_{it} = 1$	$F_{it} = 0$	Total
1994-1996 (CIS2)	$F_{i,t-1} = 1$	41.32 100 33.00	15.98 203 67.00	20.04 303 100
	$F_{i,t-1} = 0$	58.68 142 67.00	84.02 1067 88.25	76.96 1209 100
	Total	100 242 16.01	100 1270 83.99	100 1512 100

Part C: 1994-6 –> 1997-9 Transitions

		1997-1999 (FIT)				
		$I_{it} = 1$ *and* $F_{it} = 1$	$I_{it} = 1$ *and* $F_{it} = 0$	$I_{it} = 0$ *and* $F_{it} = 1$	$I_{it} = 0$ *and* $F_{it} = 0$	Total
1994-1996 (CIS2)	$I_{i,t-1} = 1$ *and* $F_{i,t-1} = 1$	37.7 58 28.3	15.8 77 37.6	14.8 13 6.3	7.3 57 27.8	13.6 205 100
	$I_{i,t-1} = 1$ *and* $F_{i,t-1} = 0$	46.1 71 10.3	68.9 337 48.7	30.7 27 3.9	32.9 257 37.1	45.8 692 100
	$I_{i,t-1} = 0$ *and* $F_{i,t-1} = 1$	7.8 12 12.2	1.4 7 7.1	19.3 17 17.3	7.9 62 63.3	6.5 98 100
	$I_{i,t-1} = 0$ *and* $F_{i,t-1} = 0$	8.4 13 2.5	13.9 68 13.2	35.2 31 6.0	51.9 405 78.3	34.2 517 100
	Total	100 154 10.2	100 489 32.3	100 88 5.8	100 781 51.7	100 1512 100

7 Appendix 3: Detailed Estimation Results

Table 9: Innovation and Financing Constraints Joint Probit
With Reverse Interaction Effects
Full sample, nobs=1940

	Model 3	
	Coeff.	Std.
Innovation Equation		
Constant	-7.235***	0.118
Size	0.183***	0.020
Market share	0.020	0.045
TP4	1.822***	0.183
TP3	1.0110***	0.199
TP2	0.437***	0.176
Financial Constraints	-0.324**	0.255
11 Industry dummies	misc	
Financial Constraint Equation		
Constant	-1.221***	0.241
Firm Innovates	0.647***	0.032
Size	-0.016	0.073
Collateral amount	0.030	0.050
Banking debt ratio	0.015***	0.002
Own financing ratio	-0.001***	0.001
Profit margin	-0.002***	0.002
11 industry dummies	misc	
$corr_{12}$	-0.132***	0.013
Log lik Innovation		
Log lik Fin Constraint		
Log lik Bivariate	-1712	

Table 10: Innovation and Financing Constraints Joint Probit
With Reverse Interaction Effects and Dynamics

Full sample, nobs=1512		
	Model 4	
	Coeff.	Std.
Innovation Equation		
Constant	-2.441***	0.323
Innov_{t-1}	0.829***	0.094
Size	0.256***	0.037
Market share	0.027	0.071
TP4	1.461***	0.201
TP3	0.932***	0.156
TP2	0.621***	0.143
Financial Constraints	-0.447***	0.106
$\text{Financial Constraints}_{t-1}$	0.300	0.123
11 Industry dummies	misc	
Financial Constraint Equation		
Constant	-0.885***	0.311
$\text{Firm Innovates}_{t}$	0.627***	0.022
$\text{Firm Innovates}_{t-1}$	0.236**	0.133
$\text{Financial constraints}_{t-1}$	0.135***	0.093
Size	0.035	0.039
Collateral amount	0.003	0.002
Banking debt ratio	0.005	0.003
Own financing ratio	-0.008***	0.002
Profit margin	-0.007***	0.002
11 industry dummies	misc	
$corr_{12}$	0.500**	0.210
Log lik Innovation		
Log lik Fin Constraint		
Log lik Bivariate	-1331	