Semiparametric Identification in Panel Data Discrete Response Models∗

Eleni Aristodemou†
UCL
March 8, 2016
Please click [here] for the latest version.‡

Abstract

This paper studies partial identification in fixed effects panel data discrete response models. In particular, semiparametric identification in linear index binary and ordered response panel data models with fixed effects is examined. It is shown that under unrestrictive distributional assumptions on the fixed effect and the time-varying unobservables and failure of point identification, informative bounds on the regression coefficients can still be derived under fairly weak conditions. Partial identification is achieved by eliminating the fixed effect and finding features of the distribution that do not depend on the unobserved heterogeneity. Numerical analysis illustrates how these sets shrink as the support of the explanatory variables increases and how higher variation in the unobservables results in wider identification bounds.

Keywords: Static and Dynamic Panel Data, Binary Response Models, Ordered Response Models

JEL classification Numbers: C01, C33, C35.

∗I would like to thank Katerina Aristodemou, Andrew Chesher, Aureo De Paula, Bo Honoré, Toru Kitagawa, Lars Nesheim, Adam Rosen, Martin Weidner and Frank Windmeijer for their useful detailed discussions and suggestions. Special thanks go to my colleagues at UCL as well as participants at the departmental seminar of the University of Cyprus in September 2015, the EEA conference in August 2015, the Bristol Econometrics Study Group in July 2015, the Spring Meeting of Young Economists in May 2015, the RES Annual Conference and the ENTER Jamboree Conference in March 2015 for their comments.

†Address: Department of Economics, UCL, Gordon Street, London WC1H0AX, United Kingdom.

‡For the latest version please see [http://www.homepages.ucl.ac.uk/~uctpear/]
1 Introduction

This paper studies semiparametric identification in fixed effects panel data discrete response models, and in particular linear index binary and ordered response panel data models with additively separable fixed effects. It is shown that in models with unrestrictive distributional assumptions on the fixed effect and the time-varying unobservables where point identification fails, informative identification bounds on the regression coefficients can still be derived under fairly weak conditions. The approach builds on the work on semiparametric identification in static binary panel data models of Rosen and Weidner (2013, WP) and extends it to dynamic binary and static and dynamic ordered response fixed effects panel data models. Partial identification is achieved by relying on features of the distribution of observable variables that do not depend on the fixed effect. Numerical analysis illustrates how these sets shrink as the support of the explanatory variables increases and how higher variation in the unobservables results in wider identification bounds.

It is often noted that individuals who were observed making a specific choice in the past are more likely to make the same choice in the future. For example, in labour economics it has been documented that the employment choices individuals make in different periods are functions of past employment decisions. In panel data demand settings, when consumers make a specific purchasing decision regularly, it is often evident that decisions are linked intertemporally. Such correlation might be a result of the formation of habits, the persistence in choice or the presence of switching costs.

As pointed out by Heckman (1981) in panel data models intertemporal correlation between choices in different periods comes in general through the presence of time-invariant unobservables as well as lagged dependent variables in the utility function. In the first case individuals differ by some unobserved time-invariant factor that correlates decisions over time, also referred to as spurious state dependence. In the latter, past choices enter the utility function directly and therefore affect future decisions. This dependence is referred to as the true state dependence. Ignoring this dynamic behaviour of individuals can result in inconsistent estimates of regression coefficients and other quantities of interest, such as price elasticities.
Linear panel data model settings where the dependent outcome is continuous, can be seen as solving an omitted variables problem, arising from the presence of this additively separable time-invariant unobserved component. Even when this fixed effect is not restricted to be independent of the explanatory variables, point identification of the regression parameters can be achieved by differencing out this additively separable time-invariant unobserved heterogeneity.

In non-linear panel data models with additively separable fixed effects, the differencing approach can not be directly implemented. Identification and estimation in these models therefore rely heavily on the assumptions econometricians place on the individual specific heterogeneity and the challenges these models face have been well documented in the literature. Choosing between random and fixed effects, how to deal with initial conditions and lagged dependent variables, as well as the incidental parameters problem and the calculation of marginal effects, have all been extensively studied by a number of authors. A detailed summary of developments can be found in Arellano and Honoré (2001) and more recently in Arellano and Bonhomme (2011).

This paper studies identification in fixed effects linear index discrete response models, without imposing distributional assumptions on the additively separable unobserved heterogeneity. It uses a similar approach as in linear panel data models, where identification relies on observable implications in which the additively separable fixed effect does not appear. As the model is non-linear, it is shown that eliminating the individual fixed effect leads to partial identification of the regression parameters.

This paper examines two different discrete response settings. The first setting corresponds to the case where individuals’ choice set consists of a binary outcome, for example whether to purchase a specific product or not in a given period, and examines semiparametric identification in linear index binary panel data models. Several papers, including Chamberlain (1984, 2010), Honoré (2002) and Honoré and Kyriaezidou (2000), have shown that in linear index panel data models with binary outcomes, parametric point-identification of the regression parameters when regressors have bounded support can only be achieved under the assumption of independently and identically logistically distributed time-varying unobservables. Manski (1987), using a conditional version of the Maximum Score Estimator, shows that the panel
data binary model can have a median regression interpretation and inference is still possible if the disturbances for each panel member are known to be time-stationary conditional on the strictly exogenous variables and the fixed effects, when the strictly exogenous explanatory variables vary enough over time with at least one component having unbounded support. \cite{Honoré and Lewbel, 2002} also show point-identification of binary panel data models with predetermined regressors, under the assumption that there exists a special regressor which conditional on the rest of the regressors and the instruments is independent of the fixed effect.

Point-identification of the regression parameters in discrete response panel data models relies on strong assumptions, such as logistically distributed unobservables and large support conditions, which are strong, untestable and difficult to satisfy in applications. In an attempt to overcome this limitation, several papers study semiparametric identification in a general class of panel data models that results to partial identification of parameters and quantities of interest. For example, \cite{Chernozhukov, Hahn, and Newey, 2005} focus on nonparametric bound identification, estimation and inference in multinomial panel data models with correlated random effects with only discrete explanatory variables. \cite{Honoré and Tamer, 2006} study bounds on parameters in dynamic discrete choice models, mainly focusing on the initial condition problem. In linear panel data settings \cite{Rosen, 2012} studies the identifying power of conditional quantile restrictions in short panels with fixed effects, while \cite{Chernozhukov, Fernández-Val, Hahn, and Newey, 2013} provide sharp identification bounds for average and quantile treatments effects in fully parametric and semiparametric nonseparable panel data models. In a very recent paper, \cite{Pakes and Porter, 2014} develop a new approach to semiparametric analysis of multinomial models with additively separable fixed effects. Set identification of the model parameters in this setting comes through the weak monotonicity assumption on the index function and the group homogeneity condition on the disturbances conditional on the contemporaneous explanatory variables and the fixed effects.

This paper focuses on semiparametric identification of the regression coefficients in linear index binary panel data models with additive fixed effects, where the main focus is on the elimination of the unobserved heterogeneity. No assumptions are imposed on the fixed effect, which is allowed to be correlated with the explanatory variables and the time-varying
unobservables. Identification is achieved by finding features of the distribution that do not depend on the fixed effect. This approach extends the work on semiparametric identification in static binary panel data models in [Rosen and Weidner (2013, WP)], thereafter RW2013, to dynamic binary panel data models, where last periods choice directly enters current period’s decision rule. Conditioning on the initial condition, identification relies on individuals who switch options from one period to the next. It is shown that the joint probability of the choices these individuals make in two consecutive periods is bounded by features of the distribution that do not depend on the unobserved fixed effect. Therefore, under an exogeneity condition for the time varying unobservables, this allows for the derivation of identification bounds on the coefficients of the contemporaneous explanatory variables as well as the lagged dependent variables, without parametric assumptions on the fixed effect or the time-varying unobservables.

Panel data binary response models are suitable for explaining individual choices when the choice set includes two alternatives, such as the choice between purchasing a product or not, the choice between two different brands and the choice between employment and unemployment. Most of the papers dealing with identification of discrete choice panel data models, use the binary case as an example. However, when the choice set includes more than two alternatives then the binary choice model may fail to take into account all the information.

Several papers recognizing that the choice set individuals face may include more than two options, have extended the binary response panel data model to a multinomial response panel data model where the choice set includes a variety of unordered alternatives. Although, these papers provide clear identification results, they usually require the comparison of each option against every other alternative, which might be intractable and computationally heavy in practice. This paper departs from these models and imposes some additional shape restrictions on the functional form, such that the binary model is extended to an ordered model. Such a method reduces the amount of between alternative comparisons needed and gives a clearer picture on which option combinations are needed for identification.

Models that impose ordering restrictions can describe situations where consumers choose between vertically differentiated products, such as the choice between flying first, business
or economy class and the choice between buying a two-blade or a three blade razor, as in Shaked and Sutton (1983), Berry (1994), Bresnahan (1987), and Aristodemou and Rosen (2015,WP). In market settings where consumers are observed purchasing a specific product regularly, consumers base their purchasing decisions on past choices or decisions are somehow correlated because consumers are locked-in. As a result, in these markets dynamics play a crucial role in individuals’ consumption decisions.

Identification of panel data ordered response models has not been extensively studied in the literature. Following the work by Honoré (1992) that shows how to consistently estimate the parameters in the truncated/censored panel data model, this paper focuses on the “in-between” case of ordered outcomes. Since every ordered response model can be expressed as a dichotomous/binary response model, parametric point-identification can be achieved under the assumptions of logistically distributed time-varying unobservables as discussed in Chamberlain (1984, 2010) and time-constant thresholds. As discussed in Baetschmann, Staub, and Winkelmann (2015) literature has estimated the fixed effects ordered logit model either with a single dichotomization with constant or individual-specific thresholds, or by combining all possible dichotomizations by various estimation methods, such as a two-step minimum distance method and generalized method of moments. The former approach is inefficient as it does not use all the information and in the case of individual specific thresholds shown to be inconsistent. The latter can be consistent and nearly efficient, under correct specification.

The approaches used in the literature rely on the distributional assumption of iid logistically distributed unobservables independent of the fixed effect and do not use all the information provided. Furthermore, they require time-invariant thresholds, which might be too restrictive in some settings. This paper departs from these approaches and examines semiparametric identification of fixed effects ordered response panel data models without imposing distributional assumptions on the fixed effect or the time-varying unobservables. The complete structure of the ordered choice model is used, thus utilizing information provided from the additional shape restrictions, while at the same time allowing for the possibility of time-varying thresholds.

As already discussed, allowing for the fixed effect to be correlated with the explanatory
variables, creates an endogeneity problem. Chesher (2010) and Chesher and Smolinski (2012) derive sharp identification bounds in nonparametric cross-sectional ordered response models in the presence of endogenous variables. They show that these bounds can shrink at a relatively fast rate as the relevance of the instruments increases and as the number of ordered outcomes becomes larger. Furthermore, they show that when the ordered model is expressed as a binary response model, the identified sets are not sharp. In a more general setting a series of papers Chesher, Rosen, and Smolinski (2013), Chesher and Rosen (2012, 2013) study identification of instrumental variables models with discrete dependent variables. This paper departs from the instrumental variables approach and tackles this problem by eliminating this unobserved heterogeneity and finding features of the distribution of the observed outcomes that do not depend on this time-invariant unobservable. In contrast to the binary response panel data model, identification of the regression parameters does not rely only on observations who switch choices in two consecutive periods. It is shown that individuals who choose the middle option in two consecutive periods are also a useful source for identification. The greater number of pairs that can be used in the ordered model in comparison to the binary model can help in achieving tighter bounds for the regression parameters, whilst keeping the number of inequalities to a trackable number. Furthermore, the use of information from individuals who stay with the same option might be useful in comparing the behaviour of switchers to non-switchers.

In this paper identification in the static and dynamic, binary and ordered panel data models is studied. The main result of this paper is that in general classes of panel data discrete response models where when point-identification in known to fail, informative identification sets of the regression coefficients can still be derived. Numerical examples examine how these sets behave under different conditions.

The rest of the paper is structured as follows. Section 2 reviews the work in RW2013 and examines identification in dynamic binary response models. Section 3 extends the binary response panel data model to an ordered response panel data model and examines identification under the static and dynamic settings. Section 4 gives some numerical results for the binary panel data model. Section 5 concludes with some final remarks and future steps.
2 The Binary Response Panel Data Model

This section examines identification in fixed effects binary response panel data models. Each individual in the population is observed for two time periods, \( t = 1 \) and \( t = 2 \), and in each time period he can choose one outcome from the set of binary options, \( Y_t \equiv \{0, 1\} \). Therefore, each individual is characterized by a set of observables \((Y, X)\) such that \( Y = (Y_1, Y_2) \), \( X = (X_1, X_2) \), and a set of unobservables \((V, \alpha)\), where \( V = (V_1, V_2) \) and \( \alpha \in \mathbb{R} \). The utility an individual, with covariates \( x_t, y_{t-1} \) and unobservables \( v_t, \alpha \), receives from choosing a specific outcome \( y \) in period \( t, y_t \), is then given by

\[
U_t = g(x_t, y_{t-1}, \alpha, v_t)
\] (1)

where, \( x_t \) are observed individual characteristics, \( y_{t-1} \) is the option chosen last period, \( \alpha \) is the unobserved to the econometrician time-invariant individual fixed effect, \( v_t \) is the unobserved to the econometrician time-varying component and \( g() \) is function not yet specified, which is an element of some parameter space \( \Theta \). Normalizing the utility of the outside option, \( Y_t = 0 \) in each period to be zero and under the assumption that in each period \( t \) individuals choose the outcome to maximize [1], the panel data binary response model can be shown to be equivalent to,

\[
y_t = 1[g(X_t, Y_{t-1}, \alpha, V_t) > 0]
\]

where

\[
1[g(X_t, Y_{t-1}, \alpha, V_t) > 0] = \begin{cases} 
1 & \text{if } g(X_t, Y_{t-1}, \alpha, V_t) > 0 \\
0 & \text{otherwise} 
\end{cases}
\] (2)

Such a setting could be used to describe situations where individuals in each periods decide whether to purchase a product or not, or whether to seek employment or stay unemployed. In all these situations persistence in decisions has been documented by a number of authors.

2.1 Static Binary Panel Data Model

This section examines the simplest version of the model in (2), where the lagged dependent variable does not enter directly the decision rule function and the correlation between choices
in different periods comes only through the unobservables, $\alpha$ and $V_t$, such that,

$$y_t = 1[g(X_t, \alpha, V_t) > 0]$$

where

$$1[g(X_t, \alpha, V_t) > 0] = \begin{cases} 1 & \text{if } g(X_t, \alpha, V_t) > 0 \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (3)$$

\subsection*{2.1.1 Model Assumptions}

\textbf{Assumption 1.} The observed data comprise a random sample of $N$ individuals from the population. For each individual $(Y, X, V, \alpha)$ are defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ contains the Borel Sets. The support of $(X, V, \alpha)$ is $(\mathcal{X} \times \mathcal{V} \times \mathcal{A})$ where $\mathcal{V} \subseteq \mathbb{R}^2$ and $\mathcal{A} \subseteq \mathbb{R}$.

\textbf{Assumption 2.} For each value of $x \in \mathcal{X}$ there is a proper conditional distribution of $(Y_1, Y_2)$ given $X = x$,

$$P^0(y_1, y_2|x) \equiv \mathbb{P}(Y_1 = y_1 \wedge Y_2 = y_2|X = x)$$

and the conditional probability of each pair $(y_1, y_2)$, is point identified over the support of $(Y_1, Y_2)$ for almost every $x \in \mathcal{X}$.

\textbf{Assumption 3.} The conditional distribution of $V$ given $(X = x)$, $F_{V|X}$ is absolutely continuous with respect to Lebesgue measure with everywhere positive density and the marginal distribution of $\Delta V|X = x$ is given by $F_{\Delta V|X}$, where $\Delta V = V_2 - V_1$.

\textbf{Assumption 4.} The conditional distribution of $\alpha$ given $(X = x)$ is absolutely continuous with respect to Lebesgue measure with everywhere positive density on $\mathbb{R}$ and marginal distribution $F_{\alpha|X}$.

\textbf{Assumption 5.} $X$ and $V$ are stochastically independent. $\alpha$ is allowed to be correlated with both $V$ and $X$ in an arbitrary way. The joint distribution of $(V, \alpha)$ conditional on $(X = x)$ is given by $F_{(V, \alpha)|X}$.
Assumptions 3-5 provide restrictions on the conditional distributions of the unobservables given the observed covariates. These conditional distributions are elements of the generic collection of conditional distributions defined as,

- \( \{F_{V|X}(\cdot|x) : x \in X\} \) is an element of \( F_{V|X}\)
- \( \{F_{\Delta V|X}(\cdot|x) : x \in X\} \) is an element of \( F_{\Delta V|X}\)
- \( \{F_{\alpha|X}(\cdot|x) : x \in X\} \) is an element of \( F_{\alpha|X}\)
- \( \{F_{(V,\alpha)V|X}(\cdot|x) : x \in X\} \) is an element of \( F_{(V,\alpha)V|X}\)

Assumption 1 defines the underlying probability space and notation for the support of random variables \((Y, X, V, \alpha)\). Assumption 2 stipulates that the conditional distribution of \((Y_1, Y_2)\) given covariates \(x\) is point identified for almost every \(x \in X\), as would be the case for example under random sampling. Assumption 3 and Assumption 4 requires that the time varying unobservable \(V\) and the unobserved heterogeneity \(\alpha\) are absolutely continuously distributed conditional on \(X\) with full support in Euclidean space. Assumption 5 imposes independence of \(X\) and \(V\), but allows \(\alpha\) and \(V\) to be arbitrary correlated with a joint distribution \(F_{(V,\alpha)|X}\).

Assumption 4 imposes no restrictions on the time-invariant unobservable and allows correlation with the explanatory variables. Therefore, the presence of \(\alpha\), creates an endogeneity problem, that needs to be addressed for identification and consistent estimation of the parameters of interest. In linear panel data models with continuous outcomes, differencing out the fixed effect is sufficient to guarantee point-identification of the regression parameters. This paper examines identification in linear index discrete panel data models and mimics the approach used in linear panel data models with continuous outcomes to solve the problem of the fixed effect. To be able to proceed the following additional assumption is imposed.

**Assumption 6.** The individual specific part, \(g(\cdot)\), of the model in (3) is linear in the variables such that,

\[
g(X_t, \alpha, V_t) = X_t \beta + \alpha + V_t
\]

with \(\beta \in \Theta\).
Finally the set of structures $S \equiv (\beta, F_{(V,\alpha)|X})$ admitted by model $[3]$ is defined through Assumption $[7]$

**Assumption 7.** $S \equiv (\beta, F_{(V,\alpha)|X}) \in \mathcal{S}$ is a specified collection of parameters $\beta$ and joint distributions of the time-varying unobservable and the unobserved heterogeneity, $F_{(V,\alpha)|X}$. Such a $S$ is called a structure.

Under Assumptions $[1][7]$ the identified set of admissible structures, denoted by $S^0$, is characterized. Define by $u(\tilde{y}_1, \tilde{y}_2, x, \alpha, v_1, v_2; \beta)$, the utility an individual with covariates $x$, and unobservables $\alpha, v_1, v_2$ receives from choosing any pair $(\tilde{y}_1, \tilde{y}_2) \in (\mathcal{Y}_1, \mathcal{Y}_2)$ and by $\mathcal{R}_{(y_1,y_2)}(x; \beta)$ the region of unobservables $(V, \alpha)$ associated with $(y_1, y_2)$ maximizing $u(\tilde{y}_1, \tilde{y}_2, x, \alpha, v_1, v_2; \beta)$, such that

$$\mathcal{R}_{(y_1,y_2)}(x; \beta) \equiv \left\{ (V, \alpha) \in (\mathcal{V}, \mathcal{A}) : (y_1, y_2) = \arg \max_{(\tilde{y}_1, \tilde{y}_2) \in (\mathcal{Y}_1 \times \mathcal{Y}_2)} u(\tilde{y}_1, \tilde{y}_2, x, \alpha, v_1, v_2; \beta) \right\}$$

Then $S^0$ is characterized by,

$$S^0 = \left\{ (\beta, F_{(V,\alpha)|X}) \in \mathcal{S} : \forall (y_1, y_2) \in (\mathcal{Y}_1 \times \mathcal{Y}_2), F_{(V,\alpha)|X} \left( \mathcal{R}_{(y_1,y_2)}(x; \beta) \right) = P^0(y_1, y_2|x) \quad \text{a.e. } x \in \mathcal{X} \right\} \quad (4)$$

and the identified set for the model parameters $\beta$ is then characterized by,

$$\Theta^0 = \left\{ \beta \in \Theta : \forall (y_1, y_2) \in (\mathcal{Y}_1 \times \mathcal{Y}_2), \exists F_{(V,\alpha)|X} \in \mathcal{F}_{(V,\alpha)|X} \right\}$$

$$F_{(V,\alpha)|X} \left( \mathcal{R}_{(y_1,y_2)}(x; \beta) \right) = P^0(y_1, y_2|x) \quad \text{a.e. } x \in \mathcal{X} \right\} \quad (5)$$

**2.1.2 Identified Set: Binary Static Panel Data Model**

Point-identification of the regression coefficients under both the logistic distribution assumption, Chamberlain (1984), and the time-stationarity assumption, Manski (1987), comes by observing individuals who change behaviour from one period to the next. This behaviour gives rise to features of the distribution that do not depend on the unobserved heterogeneity.
This paper uses exactly the same approach, and finds features of the distribution that do not depend on $\alpha$, without imposing these strong assumptions on the time-varying unobservables. As it will be shown below removing the unobserved heterogeneity, even under Assumption 6, will lead to partial identification of the regression parameters. Under Assumption 6, the model in (3) in each period, $t = \{1, 2\}$, is given by,

$$y_t = 1(X_t\beta + \alpha + V_t > 0)$$

This model has been studied in RW2013. From the specification of the model in (6), the authors obtain the regions $R_{SB}^{(y_1, y_2)}(x; \beta)$ that partition the $\text{supp}(V, \alpha)$ for all $(Y_1, Y_2) = (y_1, y_2)$ and $(X_1, X_2) = (x_1, x_2)$ such that:

$$R_{SB}^{(0, 0)}(x; \beta) = \{(V, \alpha) \in (V, A) : x_1\beta + \alpha + V_1 \leq 0 \land x_2\beta + \alpha + V_2 \leq 0\}$$

$$R_{SB}^{(0, 1)}(x; \beta) = \{(V, \alpha) \in (V, A) : x_2\beta + \alpha + V_2 > 0 \land x_1\beta + \alpha + V_1 \leq 0\}$$

$$R_{SB}^{(1, 0)}(x; \beta) = \{(V, \alpha) \in (V, A) : x_2\beta + \alpha + V_2 \leq 0 \land x_1\beta + \alpha + V_1 > 0\}$$

$$R_{SB}^{(1, 1)}(x; \beta) = \{(V, \alpha) \in (V, A) : x_1\beta + \alpha + V_1 > 0 \land x_2\beta + \alpha + V_2 > 0\}$$

and the conditional joint distribution of $(V, \alpha)|X = x$ for any given pair $(y_1, y_2) \in Y_1 \times Y_2$ is,

$$P(y_1, y_2|x) = F_{(V, \alpha)|X}(R_{SB}^{(y_1, y_2)}(x; \beta)|X = x)$$

where $P(y_1, y_2|x) = P(Y_1 = y_1 \land Y_2 = y_2|X = x)$.

**Theorem 1.** Let $S^{SB} = (\beta^{SB}, F_{(V, \alpha)|X}^{SB})$ be a structure admitted by model (6). If $S^{SB}$ is an observationally equivalent structure to $S^0$, then under assumptions 1, 7, $\beta^{SB}$ satisfies a.e. $x \in X$

$$P(1, 0|x) \leq F_{\Delta V}[-\Delta x\beta^{SB}] \leq [1 - P(0, 1|x)]$$

where $\Delta X = X_2 - X_1$ with $(X_1, X_2) \in X \times X$ and $\Delta V = V_2 - V_1$ with $(V_1, V_2) \in V \times V$.

**Proof.** The proof follows the work in RW2013 and it is provided here for completeness. The crucial element in Theorem 1 is that by limiting the attention to individuals who change their choice from one period to the next, equation (8) becomes invariant to changes in the value
of the unobserved time-varying heterogeneity, $\alpha$. This implies that only realizations that fall under the following two types can be used,

\[ \{ Y_1 = 0 \land Y_2 = 1 \} \text{ and } \{ Y_1 = 1 \land Y_2 = 0 \}. \]

The choice probabilities of these events are given by:

\[
P(0, 1 | x) = P(V, \alpha | x) \left[ \{ 0 \geq X_1 \beta + \alpha + V_1 \} \land \{ X_2 \beta + \alpha + V_2 > 0 \} \right] | X = x
\]

\[
P(1, 0 | x) = P(V, \alpha | x) \left[ \{ X_1 \beta + \alpha + V_1 > 0 \} \land \{ 0 \geq X_2 \beta + \alpha + V_2 \} \right] | X = x
\]

(9)

From (9) it is clear:

\[
P(0, 1 | x) = P(V, \alpha | x) \left[ \{ 0 \geq X_1 \beta + \alpha + V_1 \} \land \{ X_2 \beta + \alpha + V_2 > 0 \} \right] | X = x
\]

\[
\leq P(V, \alpha | x) \left[ \{ 0 > (X_1 - X_2) \beta + (V_1 - V_2) \} \right] | X = x
\]

\[
P(1, 0 | x) = P(V, \alpha | x) \left[ \{ X_1 \beta + \alpha + V_1 > 0 \} \land \{ 0 \geq X_2 \beta + \alpha + V_2 \} \right] | X = x
\]

\[
\leq P(V, \alpha | x) \left[ \{ 0 > (X_2 - X_1) \beta + (V_2 - V_1) \} \right] | X = x
\]

(10)

Then

\[
P(0, 1 | x) \leq P_{\Delta V | X}[\Delta V > -\Delta X \beta | X = x]
\]

\[
P(1, 0 | x) \leq P_{\Delta V | X}[-\Delta X \beta > \Delta V | X = x]
\]

(11)

\[
1 - P(0, 1 | x) \geq P_{\Delta V | X}[\Delta V < -\Delta X \beta | X = x]
\]

\[
P(1, 0 | x) \leq P_{\Delta V | X}[-\Delta X \beta > \Delta V | X = x]
\]

(12)

\[
\Rightarrow P(1, 0 | x) \leq P_{\Delta V | X}[\Delta V < -\Delta X \beta | X = x] \leq [1 - P(0, 1 | x)]
\]

For any given $X = x$ and using Assumption 5,

\[
P(1, 0 | x) \leq F_{\Delta V}[-\Delta x \beta] \leq [1 - P(0, 1 | x)]
\]

(13)
The above relations provide restrictions on the distribution of $\Delta V|X$ for any realization of $x \in \mathcal{X}$ that does not depend on the fixed effect $\alpha$. The events $\{Y_1 = 0 \land Y_2 = 0\}$ and $\{Y_1 = 1 \land Y_2 = 1\}$ provide no restrictions on $\Delta V$ and can not be used to eliminate the fixed effect $\alpha$. Similarly to the binary logit fixed effects model and as discussed in Honoré (2002), the individuals who do not switch can not be used to identify the regression parameter, since for any value of $\beta$ the choices these individuals make can be rationalized by extremely large or extremely small values of the fixed effect. In other words, these events provide no restrictions on the values the fixed effect can take for a given value of the regression parameters. The distribution of $\Delta V|X \sim F_{\Delta V|X}$ is equivalent to $\Delta V \sim F_{\Delta V}$ by Assumption 5. Notice that in order for the bounds in Theorem 1 to be informative, there should exists $x \in \mathcal{X}$ such that $x_1 \neq x_2$ with positive probability.

It can be clearly seen that the bounds change according to different values of $x \in \mathcal{X}$. Suppose that $X_t \in \{0, 1\}$. There are two distinct cases where $x_t$ changes from $t = 1$ to $t = 2$, $x = (x_1, x_2) \in \{(0, 1), (1, 0)\}$ and the bounds in (13) are given by:

$$P(1, 0|0, 1) \leq F_{\Delta V}[-\beta] \leq [1 - P(0, 1|0, 1)]$$

$$P(1, 0|1, 0) \leq F_{\Delta V}[+\beta] \leq [1 - P(0, 1|1, 0)]$$

Now suppose that $X_t \in \{0, 1, 2\}$. There are 6 distinct cases where $x_t$ change from period $t = 1$ to period $t = 2$, $x = (x_1, x_2) \in \{(0, 1), (1, 0), (0, 2), (2, 0), (1, 2), (2, 1)\}$. Then the bounds in (13) are given by:

$$P(1, 0|0, 1) \leq F_{\Delta V}[-\beta] \leq [1 - P(0, 1|0, 1)]$$

$$P(1, 0|1, 0) \leq F_{\Delta V}[+\beta] \leq [1 - P(0, 1|1, 0)]$$

$$P(1, 0|0, 2) \leq F_{\Delta V}[-2\beta] \leq [1 - P(0, 1|0, 2)]$$

$$P(1, 0|2, 0) \leq F_{\Delta V}[2\beta] \leq [1 - P(0, 1|2, 0)]$$

$$P(1, 0|1, 2) \leq F_{\Delta V}[-\beta] \leq [1 - P(0, 1|1, 2)]$$

$$P(1, 0|2, 1) \leq F_{\Delta V}[\beta] \leq [1 - P(0, 1|2, 1)]$$

From the above it is clear that different values of $(x_1, x_2)$ lead to different identified distributions for $(Y_1, Y_2)$. Therefore, to characterize the identified set for $\beta$, the greatest lower bound
and the smallest upper bound of (13) over all possible pairs \((x_1, x_2)\) need to be used. Following RW2013 take an arbitrary constant \(\omega \in \mathbb{R}\) and notice that:

\[
(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega \Rightarrow \Delta V < \omega
\]

\[
(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega \Rightarrow \Delta V > \omega
\]  

(14)

Therefore, \(\forall \omega \in \mathbb{R}\)

\[
P[(Y_1 = 1 \land Y_2 = 0) \land -\Delta X \beta \leq \omega | X = x] \leq F_{\Delta V}(\omega) \leq 1 - P[(Y_1 = 0 \land Y_2 = 1) \land -\Delta X \beta \geq \omega | X = x]
\]

The upper and lower bounds will then be defined as:

\[
\sup_{x \in X} P[(Y_1 = 1 \land Y_2 = 0) \land -\Delta X \beta \leq \omega | X = x] \leq F_{\Delta V}(\omega) \leq \inf_{x \in X} 1 - P[(Y_1 = 0 \land Y_2 = 1) \land -\Delta X \beta \geq \omega | X = x]
\]  

(15)

Furthermore notice that depending on the value of \(\omega\) the lower and the upper bounds will change. When \(-\Delta X \beta \leq \omega\),

\[
1 - P[(Y_1 = 0 \land Y_2 = 1) \land -\Delta X \beta \geq \omega | X = x] = 1
\]

\[
P[(Y_1 = 1 \land Y_2 = 0) \land -\Delta X \beta \leq \omega | X = x] = P(1, 0|x)
\]

and when \(-\Delta X \beta \geq \omega\),

\[
1 - P[(Y_1 = 0 \land Y_2 = 1) \land -\Delta X \beta \geq \omega | X = x] = 1 - P(0, 1|x)
\]

\[
P[(Y_1 = 1 \land Y_2 = 0) \land -\Delta X \beta \leq \omega | X = x] = 0
\]

Therefore this implies that (15) is equivalent to

\[
\sup_{x: -\Delta x \beta \leq \omega} P(1, 0|x) \leq F_{\Delta V}(\omega) \leq \inf_{x: -\Delta x \beta \geq \omega} [1 - P(0, 1|x)]
\]  

(16)

**Theorem 2.** (RW2013) Let Assumptions \([7]\) hold. The identified set for \(\beta\) is given by:

\[
\Theta^{SB} = \left\{ \beta \in \Theta : \forall \omega \in \mathbb{R}, \sup_{x: -\Delta x \beta \leq \omega} P(1, 0|x) \leq \inf_{x: -\Delta x \beta \geq \omega} [1 - P(0, 1|x)] \text{ a.e. } x \in \mathcal{X} \right\}
\]  

(17)
Theorem 3. (RW2013) Let Assumptions 1-7 hold. Then if \( \beta \in \Theta^{SB} \), there exists a structure \( S^{SB} = (\beta^{SB}, F^{SB}_{(V,\alpha)|X}) \) that satisfies the restrictions of the model and is observationally equivalent to structure \( S^0 \) that generates \( P^0(y_1, y_2|x) \). The sharp identified set for \( \beta \) and \( F^{SB}_{(V,\alpha)|X} \) is given by:

\[
S^{SB} = \left\{ (\beta, F^{SB}_{(V,\alpha)|X}) \in (\Theta, F^{SB}_{(V,\alpha)|X}) : \forall \omega \in \mathbb{R}, \sup_{x: -\Delta x \beta \leq \omega} P(1, 0|x) \leq F^V(\omega) \leq \inf_{x: -\Delta x \beta \geq \omega} [1 - P(0, 1|x)], \right. \\
\left. \text{with } F^V(\omega) = \int_{-\infty}^{\infty} F_{V|\alpha,X}(V_2 - V_1 < \omega|\alpha, x) dF_{\alpha|X}(\alpha|x) \text{ a.e. } x \in X \right\}
\]

Proof. The proof of Theorem 2 and Theorem 3 can be found in RW2013.

Corollary 1. Let Assumptions 1-7 hold. Then if \( (\beta^{SB}, F^{SB}_{(V,\alpha)|X}) \in S^{SB} \), the identified set for \( \beta \) and \( F_{\Delta V} \) is defined as,

\[
S^{SB}\Delta = \left\{ (\beta, F_{\Delta V}) \in (\Theta, F_{\Delta V}) : \forall \omega \in \mathbb{R}, \sup_{x: -\Delta x \beta \leq \omega} [P(1, 0|x)] \leq F^V(\omega) \leq \inf_{x: -\Delta x \beta \geq \omega} [1 - P(0, 1|x)] \text{ a.e. } x \in X \right\}
\]

Proof. From Theorem 1 the existence of the conditional joint distribution of \( (V, \alpha)|X \) for every \( \beta \in \Theta^{SB} \) implies the existence of the marginal distribution \( F_{\alpha|X} \) for every \( \beta \in \Theta^{SB} \) and a marginal distribution \( F_V \) with an associated marginal distribution of the difference \( F_{\Delta V} \).

2.2 Dynamic Binary Panel Data Models

Section 2.1 studies static panel data models, where individuals’ choices are correlated across different periods only through the unobserved heterogeneity. In panel data settings with repeated observations it is evident and natural to assume that individuals’ past choices directly affect current and future decisions. For example, an individual’s decision to seek employment in the current period is likely to be affected on his employment status last period. The dynamic binary response panel data model that includes the lagged dependent variable as an additional explanatory variable is equivalent to model (2),

\[ y_t = 1[g(X_t, Y_{t-1}, \alpha, V_t) > 0] \]
where
\[
1[g(X_t, Y_{t-1}, \alpha, V_t) > 0] = \begin{cases} 
1 & \text{if } g(X_t, Y_{t-1}, \alpha, V_t) > 0 \\
0 & \text{otherwise}
\end{cases}
\]

Last period’s choice directly affects the decision so the choice in period \( t - 1 \) needs to be taken into account. This creates an initial condition problem in modeling the choice in period \( t = 1 \), since the choice in period \( t = 1 \) depends on the choice in period \( t = 0 \). In handling this problem an approach similar to [Wooldridge (2005)] is used. In particular, in this paper it is be assumed that the outcome in period \( t = 0 \), \( Y_0 = y_0 \), is observed, however no assumptions about its generation are imposed or about its relation with the fixed effect, such that the set of conditioning covariates consists of \( (x, y_0) \in \mathcal{X} \times \mathcal{Y}_0 \). Section 2.2.1 formalizes the assumptions.

### 2.2.1 Model Assumptions

**Assumption 8.** The observed data comprise a random sample of \( N \) individuals from the population. For each individual \((Y, Y_0, X, V, \alpha)\) are defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \( \mathcal{F} \) contains the Borel Sets. The support of \((X, V, \alpha)\) is \((\mathcal{X} \times \mathcal{V} \times \mathcal{A})\) where \( \mathcal{V} \subseteq \mathbb{R}^2 \) and \( \mathcal{A} \subseteq \mathbb{R} \).

**Assumption 9.** For each value of \( x \in \mathcal{X} \) and \( y_0 \in \mathcal{Y}_0 \) there is a proper conditional distribution of \((Y_1, Y_2)\) given \( X = x \) and \( Y_0 = y_0 \),

\[
P^0(y_1, y_2 | x, y_0) \equiv \mathbb{P}(Y_1 = y_1 \wedge Y_2 = y_2 | X = x, Y_0 = y_0)
\]

and the conditional probability of each pair \((y_1, y_2)\), is point identified over the support of \((Y_1, Y_2)\) for almost every \( x \in \mathcal{X} \) and \( y_0 \in \mathcal{Y}_0 \).

**Assumption 10.** The conditional distribution of \( V \) given \((X = x, Y_0 = y_0)\), \( F_{V|X,Y_0} \) is absolutely continuous with respect to Lebesgue measure with everywhere positive density and the marginal distribution of \( \Delta V | X = x, Y_0 = y_0 \) is given by \( F_{\Delta V|X,Y_0} \), where \( \Delta V = V_2 - V_1 \).

**Assumption 11.** The conditional distribution of \( \alpha \) given \( X = x, Y_0 = y_0 \) is absolutely continuous with respect to Lebesgue measure with everywhere positive density on \( \mathbb{R} \) and marginal distribution \( F_{\alpha|X,Y_0} \).
Assumption 12. \( X \) and \( V \) are stochastically independent conditional on \( Y_0 \), i.e. \( V \perp X|Y_0 \). \( \alpha \) is allowed to be correlated with both \( V \) and \( X \) in an arbitrary way. The joint distribution of \((V,\alpha)\) conditional on \( (X=x,Y_0=y_0) \) is given by \( F_{(V,\alpha)|X,Y_0} \).

Assumptions (8)-(11) have the same interpretation as in Section 2.1.1. Assumption 12 imposes only conditional independence between \( X \) and \( V \) conditional on \( Y_0 \), which is less restrictive than the assumptions imposed in the literature, such as \( V \perp (X,Y_0) \) or specifying the conditional distribution \( V|X,Y_0 \). This assumption allows for example correlation between \( V_0 \) and \((V_1,V_2)\). Assumptions 10-12 provide restrictions on the conditional distributions of the unobservables given the observed covariates. These conditional distributions are elements of the generic collection of conditional distributions defined as,

- \( \{F_{V|X,Y_0}(\cdot|x,y_0) : x \in \mathcal{X}, y_0 \in \mathcal{Y}_0\} \) an element of \( \mathcal{F}_{V|X,Y_0} \)
- \( \{F_{\Delta V|X,Y_0}(\cdot|x,y_0) : x \in \mathcal{X}, y_0 \in \mathcal{Y}_0\} \) an element of \( \mathcal{F}_{\Delta V|X,Y_0} \)
- \( \{F_{\alpha|X,Y_0}(\cdot|x,y_0) : x \in \mathcal{X}, y_0 \in \mathcal{Y}_0\} \) an element of \( \mathcal{F}_{\alpha|X,Y_0} \)
- \( \{F_{(V,\alpha)|X,Y_0}(\cdot|x,y_0) : x \in \mathcal{X}, y_0 \in \mathcal{Y}_0\} \) an element of \( \mathcal{F}_{(V,\alpha)|X,Y_0} \)

As this paper examines identification in discrete linear-index panel data models the following additional assumption, similar to Assumption 6 is also imposed.

Assumption 13. The individual specific part, \( g(\cdot) \), of model (2) is linear in the variables such that,

\[
g(X_t,Y_{t-1},\alpha,V_t) = X_t\beta + l(Y_{t-1})\gamma + \alpha + V_t
\]

with \( \beta,\gamma \in \Theta \) and where \( l(Y_{t-1}) \) is a prespecified function of \( Y_{t-1} \).

Under Assumption 13 the model in (2) is given by:

\[
y_t = 1(X_t\beta + l(Y_{t-1})\gamma + \alpha + V_t > 0)
\] (18)

To account for last period’s choice, this paper examines two different settings. In the first setting described in this paper, \( l(Y_{t-1}) = 1(Y_{t-1} = 1) \) such that,

\[
1(Y_{t-1} = 1) = \begin{cases} 
1 & \text{if } Y_{t-1} = 1 \\
0 & \text{otherwise}
\end{cases}
\] (19)
and in the second, \( l(Y_{t-1}) = 1(Y_t = Y_{t-1}) \) such that,

\[
1(Y_t = Y_{t-1}) = \begin{cases} 
1 & \text{if } Y_t = Y_{t-1} \\
0 & \text{otherwise} 
\end{cases}
\]  

(20)

The rationale behind these settings is that the options in the models considered in this paper have a qualitative rather than quantitative meaning. The first setting is equivalent to the standard specification of dynamic binary panel data models, where the lagged dependent variable enters as an additional regressor in the form of \( Y_{t-1} \) and the parameter \( \gamma \) measures the “impact” of choosing option \( Y_{t-1} = 1 \). In the second one the parameter \( \gamma \) captures the effect of choosing the same option from one period to the next, rather than “a unit change” effect, and can be interpreted as a state dependence variable. The main focus of this paper will be in the first case. Finally, the set of admissible structures \( S \equiv (\beta, \gamma, F_{(V,\alpha)|X,Y_0}) \) is defined in Assumption 14.

**Assumption 14.** \( S \equiv (\beta, \gamma, F_{(V,\alpha)|X,Y_0}) \in S \) is a specified collection of parameters \( \beta \) and \( \gamma \), and joint distributions of the time-varying unobservable and the unobserved heterogeneity. Such a \( S \) is called a structure.

Under Assumptions 8-14 the identified set of admissible structures, denoted by \( S^0 \), is characterized. Define by \( u(\tilde{y}_1, \tilde{y}_2, x, y_0, \alpha, v_1, v_2; \beta, \gamma) \), the utility an individual with covariates \( x, y_0 \) and unobservables \( \alpha, v_1, v_2 \) receives from choosing any pair \((\tilde{y}_1, \tilde{y}_2) \in (Y_1, Y_2)\) and by \( R_{(y_1,y_2)}(x, y_0; \beta, \gamma) \) the region of unobservables \((V, \alpha)\) associated with \((y_1, y_2)\) maximizing \( u(\tilde{y}_1, \tilde{y}_2, x, y_0, \alpha, v_1, v_2; \beta, \gamma) \), such that

\[
R_{(y_1,y_2)}(x, y_0; \beta, \gamma) \equiv \left\{ (V, \alpha) \in (Y_1 \times Y_2) : (y_1, y_2) = \arg \max_{(\tilde{y}_1, \tilde{y}_2) \in (Y_1 \times Y_2)} u(\tilde{y}_1, \tilde{y}_2, x, y_0, \alpha, v_1, v_2; \beta, \gamma) \right\}
\]

Then \( S^0 \) is characterized by,

\[
S^0 = \left\{ (\beta, \gamma, F_{(V,\alpha)|X,Y_0}) \in S : \forall (y_1, y_2) \in (Y_1 \times Y_2), F_{(V,\alpha)|X,Y_0}(R_{(y_1,y_2)}(x, y_0; \beta, \gamma)) = P^0(y_1, y_2|x, y_0) \quad \text{a.e. } x \in X \text{ and } y_0 \in Y_0 \right\}
\]  

(21)
and the identified set for the model parameters \((\beta, \gamma)\) is then characterized by,

\[ \Theta^0 = \left\{ \begin{array}{l}
(\beta, \gamma) \in \Theta : \forall (y_1, y_2) \in (Y_1 \times Y_2), \exists F_{(V,\alpha)|X,Y_0} \in F_{(V,\alpha)|X,Y_0} \\
F_{(V,\alpha)|X,Y_0}(R_{(y_1,y_2)}(x,y_0;\beta,\gamma)) = P^0(y_1, y_2|x, y_0) \\
a.e. \ x \in X \text{ and } y_0 \in Y_0
\end{array} \right\} \] (22)

Point-identification of the regression coefficients in a dynamic binary response model under the logistic distribution assumption comes by observing individuals for (at least) four time periods who change behaviour from period \(t = 2\) and \(t = 3\), as shown in Honéré and Kyriazidou (2000). This behaviour gives rise to features of the distribution that do not depend on the unobserved heterogeneity. This paper uses exactly the same approach, and finds features of the distribution that do not depend on \(\alpha\), without imposing these strong assumptions on the time-varying unobservables. As it will be shown below removing the unobserved heterogeneity, even under Assumption 13, will lead to partial identification of the regression parameters.

2.2.2 Identified Set: Binary Response Dynamic Panel Data Model

This section considers an extension of model (6) that includes last period’s choice and uses the specification in (19). The model in (18) can then be expressed as:

\[ y_t = 1(X_t\beta + 1(Y_{t-1} = 1)\gamma + \alpha + V_t > 0) \] (23)

The model in (23) is equivalent to standard linear-index dynamic binary response panel data models, like the one discussed in Honéré and Kyriazidou (2000), where the lagged binary dependent variable enters as an additional regressor. As in Section 2.1 identification of the model parameters \((\beta, \gamma)\) comes through features of the distribution that are invariant to changes in \(\alpha\), by considering the joint probability of the choices individuals make in periods \(t = 1\) and \(t = 2\), conditional on \(Y_0 = y_0\). The regions \(R_{(y_1,y_2)}^{DB}(x,y_0;\beta,\gamma)\) that partition the \(\text{supp}(V,\alpha)\)
such that for all \((V, \alpha) \in (V, A)\), \((Y_1, Y_2) = (y_1, y_2)\) when \(X = x\) and \(Y_0 = y_0\) are given by:

\[
R_{(0,0)}^{DB}(x, 0; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_1 \beta + \alpha + V_1 \leq 0 \& x_2 \beta + \alpha + V_2 \leq 0\}
\]

\[
R_{(0,1)}^{DB}(x, 0; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 \geq x_1 \beta + \alpha + V_1\}
\]

\[
R_{(1,0)}^{DB}(x, 0; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_2 \beta + \gamma + \alpha + V_2 \leq 0 < x_1 \beta + \alpha + V_1\}
\]

\[
R_{(1,1)}^{DB}(x, 0; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_1 \beta + \alpha + V_1 > 0 \& x_2 \beta + \gamma + \alpha + V_2 > 0\}
\]

\[
R_{(0,0)}^{DB}(x, 1; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_1 \beta + \alpha + V_1 \leq 0 \& x_2 \beta + \alpha + V_2 \leq 0\}
\]

\[
R_{(0,1)}^{DB}(x, 1; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 \geq x_1 \beta + \gamma + \alpha + V_1\}
\]

\[
R_{(1,0)}^{DB}(x, 1; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_2 \beta + \gamma + \alpha + V_2 \leq 0 < x_1 \beta + \gamma + \alpha + V_1\}
\]

\[
R_{(1,1)}^{DB}(x, 1; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_1 \beta + \gamma + \alpha + V_1 > 0 \& x_2 \beta + \gamma + \alpha + V_2 > 0\}
\]

(24)

The conditional joint probabilities of \(F_{(V, \alpha)|X, Y_0}\) for any given pair \((y_1, y_2) \in (Y_1 \times Y_2)\) for almost every \(x \in X\) and \(y_0 \in Y_0\) are thus defined by,

\[
P(y_1, y_2|x, y_0) = F_{(V, \alpha)|X, Y_0}(R_{(y_1, y_2)}^{DB}(x, y_0; \beta, \gamma)|X = x, Y_0 = y_0)
\]

(25)

where \(P(y_1, y_2|x, y_0) = P(Y_1 = y_1 \land Y_2 = y_2|X = x, Y_0 = y_0)\).

From the regions in (24) and Figure 1(a), it can be seen that the model in (23) is complete in the sense that conditioning on any value of the exogenous variables and the initial condition, there is a unique solution to the individual’s problem with probability one. Therefore, applying the same approach as in Section 2.1 that removes \(\alpha\), identification sets of the form of (21) and (22) can be derived. Since the outcome in period \(t = 0\) appears in the generation of the outcome in period \(t = 1\), the outcomes in all three periods \(t = \{0, 1, 2\}\) are used. Assumption 11 imposes no restrictions on the fixed effect and identification of the parameters \(\beta\) and \(\gamma\) through the elimination of \(\alpha\) comes only by observing individuals who switch in periods \(t = 1\) and \(t = 2\) for each of the values of \(y_0\). This implies that consideration of observations of the
following types are required,

\[
A = \{ Y_0 = 0 \land Y_1 = 0 \land Y_2 = 1 \}
\]

\[
B = \{ Y_0 = 0 \land Y_1 = 1 \land Y_2 = 0 \}
\]

\[
C = \{ Y_0 = 1 \land Y_1 = 0 \land Y_2 = 1 \}
\]

\[
D = \{ Y_0 = 1 \land Y_1 = 1 \land Y_2 = 0 \}
\]

(26)

**Theorem 4.** Let \( S^{DB} = (\beta^{DB}, \gamma^{DB}, F^{DB}_{(V, \alpha)|X,Y_0}) \) be a structure admitted by model (23) that is observationally equivalent to \( S^0 \), then under Assumptions \[8\text{–}14\] \((\beta^{DB}, \gamma^{DB})\) satisfy

\[
1 - P(0, 1|x, 0) \geq P_{\Delta V|Y_0}[\Delta V < -\Delta x \beta^{DB}|Y_0 = 0]
\]

\[
P(1, 0|x, 0) \leq P_{\Delta V|Y_0}[\Delta V < -\Delta x \beta^{DB} - \gamma^{DB}|Y_0 = 0]
\]

\[
1 - P(0, 1|x, 1) \geq P_{\Delta V|Y_0}[\Delta V < -\Delta x \beta^{DB} + \gamma^{DB}|Y_0 = 1]
\]

\[
P(1, 0|x, 1) \leq P_{\Delta V|Y_0}[\Delta V < -\Delta x \beta^{DB}|Y_0 = 1]
\]

where \( P(y_1, y_2|x, y_0) = P(Y_1 = y_1 \land Y_2 = y_2|X = x, Y_0 = y_0), \Delta X = X_2 - X_1 \) with \((X_1, X_2) \in \mathcal{X} \times \mathcal{X}\) and \( \Delta V = V_2 - V_1 \) with \((V_1, V_2) \in \mathcal{V} \times \mathcal{V}\).

**Proof.** The proof of Theorem 4 follows similar arguments as in Theorem 1 and is provided in Appendix A.1. Figures 1(a) and 1(b) plot the regions of unobservables conditional on \( Y_0 = 0 \) and \( \gamma < 0 \) given in equations (24) and provide an outline of the main idea. The model in (23) is complete as it can be seen from Figure 1(a), since conditional on any value of the explanatory variable and the initial condition there is a unique solution to the individual decision problem with probability one. Using this fact it can be shown that the probability of any switching event is bounded by the probability of an event that is independent of the fixed effect. Figure 1(b) shows that for the event \((Y_1, Y_2) = (1, 0)\) conditional on \( Y_0 = 0 \) and \( \gamma < 0 \). Changing \( \alpha \) moves the region of \((1, 0)|Y_0 = 0\) up and down the line \( \Delta V = -\Delta X \beta - \gamma \). Therefore, it is clear that \((V^*, \alpha^*) \in R^{DB}_{(1,0)}(x, 0; \beta, \gamma)\) implies \((V^*, \alpha^*) \in \{(V, \alpha): \Delta V < -\Delta x \beta - \gamma\}\) and \( P(1, 0|x, 0) \leq P_{\Delta V|Y_0}[\Delta V < -\Delta x \beta - \gamma|Y_0 = 0] \). \( \square \)
Theorem 5. Following Theorem 4 under Assumptions 8-14 the bounds for $\beta, \gamma$ are given by the set:

$$
\Theta^{DB} = \left\{ (\beta, \gamma) \in \Theta : \forall \omega \in \mathbb{R}, \begin{array}{c}
\sup_{x: -\Delta x \beta - \gamma \leq \omega} P(1, 0 | x, 0) \leq \inf_{x: -\Delta x \beta + \gamma \geq \omega} 1 - P(0, 1 | x, 0) \\
\sup_{x: -\Delta x \beta \leq \omega} P(1, 0 | x, 1) \leq \inf_{x: -\Delta x \beta + \gamma \geq \omega} 1 - P(0, 1 | x, 1)
\end{array} \right\}
$$

and a.e. $x \in \mathcal{X}$

Proof. The proof is provided in Appendix A.2 and follows similar arguments as in Section 2.1 and in Chesher (2013) and RW2013.

In dynamic models this period’s choice depends on last periods’ choice. This implies that the choice in period $t = 1$ depends on the choice in period $t = 0$, which is the first period observed in the sample. Unless this period coincides with the first period of the process, it will depend on previous (not observed) periods, the exogenous variables in period $t = 0$ and the joint distribution of the outcome in the first period and the unobserved heterogeneity. This joint distribution is (in general) different from the joint distribution of future outcomes and
the unobserved heterogeneity. Therefore, since $V \not\perp \alpha$, Assumption 12 of $V \perp X|Y_0$ does not imply (in general) $V \perp (X,Y_0)$. Notice that unlike the Honor´e and Kyriazidou (2000) result, where point-identification in the dynamic binary panel data model is achieved iff the errors are logistically distributed and 4 time periods are observed, the identified set in Theorem 5 only requires 3 time periods. The identified set in Theorem 5 is the projection of the joint identified set of $(\beta, \gamma, F(V|\alpha)|X,Y_0)$.

Theorem 6. Let Assumptions 8-14 hold. Then if $(\beta, \gamma) \in \Theta^{DB}$, there exists a structure $S^{DB} = (\beta^{DB}, \gamma^{DB}, F^{DB}(V|\alpha)|X,Y_0)$ that satisfies the restrictions of the model and is observationally equivalent to structure $S^0$ that generates $P^0(y_1, y_2|x, y_0)$. Bounds for $\beta, \gamma$ and $F(V|\alpha)|X,Y_0$ are given by the set:

$$S^{DB} = \left\{ (\beta, \gamma, F(V|\alpha)|X,Y_0) \in (\Theta, F(V|\alpha)|X,Y_0) : \forall \omega \in \mathbb{R}, \sup_{x: -\Delta x \beta - \gamma \leq \omega} P(1,0|x,0) \leq F_{\Delta V|Y_0}(\omega|0) \leq \inf_{x: -\Delta x \beta \geq \omega} 1 - P(0,1|x,0) \right\}$$

and

$$\sup_{x: -\Delta x \beta \leq \omega} P(1,0|x,1) \leq F_{\Delta V|Y_0}(\omega|1) \leq \inf_{x: -\Delta x \beta + \gamma \geq \omega} 1 - P(0,1|x,1)$$

with $F_{\Delta V|Y_0}(\omega|y_0) = \int_{-\infty}^{\omega} F_{V|\alpha,X,Y_0}(V_2 - V_1 < \omega|\alpha,x,y_0)dF_{\alpha|X,Y_0}(\alpha|x,y_0)$ a.e. $x \in X, y_0 \in Y_0$

The identified set in Theorem 6 does not specify the sharp identified sets for $(\beta, \gamma, F(V|\alpha)|X,Y_0)$. The presence of the lagged dependent variable creates additional difficulties and defining the sharp identified set of structures is left for future work.

Suppose that instead of Assumption 12 of $X \perp V|Y_0$, Assumption 5 of $X \perp V$ as in Section 2.1.1 is imposed. Then the identified set for $\beta$ and $\gamma$ can be expressed in terms of the unconditional probabilities such that,

Theorem 7. Let Assumptions 8-11, 5, 13-14 hold. Then the (unconditional) identified set for $(\beta, \gamma)$ is given by,

$$\Theta^{DB}_U = \left\{ (\beta, \gamma) \in \Theta : \forall \omega \in \mathbb{R}, \sup_{x \in X} \{G(\omega|x,0)P_0(x) + G(\omega|x,1)P_1(x)\} \leq \inf_{x \in X} \{G(\omega|x,0)P_0(x) + G(\omega|x,1)P_1(x)\} \right\}$$
where

\[ G(\omega|x,0) = P[(Y_1,Y_2) = (1,0) \land -\Delta X\beta - \gamma \leq \omega|X = x,Y_0 = 0] \]

\[ G(\omega|x,1) = P[(Y_1,Y_2) = (1,0) \land -\Delta X\beta \leq \omega|X = x,Y_0 = 1] \]

\[ \overline{G}(\omega|x,0) = 1 - P[(Y_1,Y_2) = (0,1) \land -\Delta X\beta \geq \omega|X = x,Y_0 = 0] \]

\[ \overline{G}(\omega|x,1) = 1 - P[(Y_1,Y_2) = (0,1) \land -\Delta X\beta + \gamma \geq \omega|X = x,Y_0 = 1] \]

and

\[ P(Y_0 = 0|X = x) = P_0(x) \]

\[ P(Y_0 = 1|X = x) = P_1(x) \]

Proof. The proof applies the Total Law of Probabilities and is provided in Section A.3.

2.2.3 Dynamic Binary Panel Data Model with Changing Behaviour

Consider now the extension of model (6) that includes last period’s choice, but allows for the lagged dependent variable to enter as specified by (20),

\[ 1(Y_{t-1} = Y_t) = \begin{cases} 
1 & \text{if } Y_{t-1} = Y_t \\
0 & \text{otherwise} 
\end{cases} \]

Such a specification would be suitable in modeling cases where individuals form habits or are somehow locked-in once they made a decision and therefore staying with the same option from one period to the next has an additional effect on their utility, captured by the state dependence parameter \( \gamma \). The individual decision rule in (18) can be written as,

\[ y_t = 1(X_t\beta + 1(Y_{t-1} = Y_t)\gamma + \alpha + V_t > 0) \]  

(27)

As in Section 2.2.2 the choice in period \( t-1 \) directly affects the utility, and hence the choice in period \( t \). Consider now the two cases where \( Y_{t-1} \neq Y_t \) and \( Y_{t-1} = Y_t \), when \( Y_t \neq Y_{t+1} \).

Using the utility notation in (1), \( U_t = g(X_t,Y_{t-1},\alpha,V_t) \), in the first case the underlying utility function in model (27) becomes,

\[ U_t = X_t\beta + \alpha + V_t \]

\[ U_{t+1} = X_{t+1}\beta + \alpha + V_{t+1} \]  

(28)
and $\gamma$ does not appear in the determination of $U_t$. In the latter case the utility function becomes,

$$
U_t = X_t \beta + \gamma + \alpha + V_t \\
U_{t+1} = X_{t+1} \beta + \alpha + V_{t+1}
$$

and $\gamma$ appears in the equation determining $U_t$. Notice one important feature of the model in (27), the underlying within period decision rule for each $y \in Y_t$ changes given $Y_{t-1}$. For example suppose that $Y_{t-1} = 1$

$$
Y_t = \begin{cases} 
1 & \text{if } X_t \beta + \gamma + \alpha + V_t > 0 \\
0 & \text{if } X_t \beta + \alpha + V_t < 0.
\end{cases}
$$

This implies that, as it will be shown below, the model in (27) is both incomplete and incoherent\footnote{The within period decision rule will be different for different outcomes even in the case where the direction of change is explicitly modeled and allowed to have heterogeneous coefficients.}

The regions of unobservables $(V, \alpha)$, $R^{DB}_{(y_1, y_2)}(x, y_0; \beta, \gamma)$ associated with each $(Y_1, Y_2) = (y_1, y_2)$ choice when $X = x$ and $Y_0 = y_0$ are given by,

$$
\begin{align*}
R^{DB}_{(0,0)}(x, 0; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_1 \beta + \gamma + \alpha + V_1 \leq 0 \& x_2 \beta + \gamma + \alpha + V_2 \leq 0\} \\
R^{DB}_{(0,1)}(x, 0; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 > 0 \geq x_1 \beta + \gamma + \alpha + V_1\} \\
R^{DB}_{(1,0)}(x, 0; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 \leq 0 < x_1 \beta + \alpha + V_1\} \\
R^{DB}_{(1,1)}(x, 0; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_1 \beta + \alpha + V_1 > 0 \& x_2 \beta + \gamma + \alpha + V_2 > 0\} \\
R^{DB}_{(0,0)}(x, 1; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_1 \beta + \alpha + V_1 \leq 0 \& x_2 \beta + \gamma + \alpha + V_2 \leq 0\} \\
R^{DB}_{(0,1)}(x, 1; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 > 0 \geq x_1 \beta + \gamma + \alpha + V_1\} \\
R^{DB}_{(1,0)}(x, 1; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 \leq 0 < x_1 \beta + \gamma + \alpha + V_1\} \\
R^{DB}_{(1,1)}(x, 1; \beta, \gamma) &= \{(V, \alpha) \in (V, A) : x_1 \beta + \gamma + \alpha + V_1 > 0 \& x_2 \beta + \gamma + \alpha + V_2 > 0\}
\end{align*}
$$

(31)
Proposition 1. From the regions defined in (31) and Figure 2, it can be seen that the dynamic model in (27) is incomplete, in the sense that \( \exists (x, y_0) \in X \times Y_0 \) conditional on which there is not a unique solution to model (27) with probability 1.

Proof. From (31) the regions of unobservables associated with the choice pairs \((Y_1, Y_2) = (1, 0)\) and \((Y_1, Y_2) = (0, 0)\) when \(Y_0 = 0\) are defined by,

\[
R^{DB}_{(1, 0)}(x, 0; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_2 \beta + \alpha + V_2 \leq 0 < x_1 \beta + \alpha + V_1\}
\]

\[
R^{DB}_{(0, 0)}(x, 0; \beta, \gamma) = \{(V, \alpha) \in (V, A) : x_1 \beta + \alpha + V_1 \leq 0 \text{ and } x_2 \beta + \gamma + \alpha + V_2 \leq 0\}
\]

Suppose also \(\gamma < 0\) and consider any \((V, \alpha) \in V^*\) such that

\[
V^* = \{(V, \alpha) \in (V, A) : -x_1 \beta < V_1 + \alpha \leq -x_1 \beta - \gamma \text{ and } V_2 + \alpha \leq -x_2 \beta\}
\]

Then it can be shown that:

\[
(V, \alpha) \in V^* \Rightarrow (V, \alpha) \in R^{DB}_{(1, 0)}(x, 0; \beta, \gamma)
\]

\[
(V, \alpha) \in V^* \Rightarrow (V, \alpha) \in R^{DB}_{(0, 0)}(x, 0; \beta, \gamma)
\]

This implies that the model in (27) is incomplete in the sense that there exists a pair of unobservables \((V, \alpha)\) such that conditional on \(X\) and \(Y_0\) there are multiple solutions to the individual’s problem.

Proposition 2. From the regions defined in (31) and Figure 2, it can be seen that the dynamic model in (27) is incoherent, in the sense that \(\exists (x, y_0) \in X \times Y_0 \) conditional on which there is no solution to model (27).

Proof. Suppose \(\gamma < 0\) and \(Y_0 = 0\). Consider any \((V, \alpha) \in V^{**}\), where

\[
V^{**} = \{(V, \alpha) \in (V, A) : -x_1 \beta - \gamma \leq V_1 + \alpha \text{ and } -x_2 \beta < V_2 + \alpha < -x_2 \beta - \gamma\}
\]
If \(-x_1\beta - \gamma \leq V_1 + \alpha\) then \(Y_1 = 1\) and \(Y_0 = 1\), which contradicts the conditioning on \(Y_0 = 0\). However, \(-x_1\beta - \gamma \leq V_1 + \alpha \Rightarrow -x_1\beta \leq V_1 + \alpha\). This implies that conditioning on \(Y_0 = 0\), \(-x_1\beta - \gamma \leq V_1 + \alpha \Rightarrow Y_1 = 1\).

For \((V,\alpha) \in V^{**}\) it also means that \(-x_2\beta < V_2 + \alpha < -x_2\beta - \gamma\). This can correspond to \(-x_2\beta < V_2 + \alpha\) and \(V_2 + \alpha < -x_2\beta - \gamma\). However, conditioning on \(Y_1 = 1\) and \(Y_0 = 0\) none of the two inequalities can hold. First take, \(-x_2\beta < V_2 + \alpha\). This corresponds to the event \(Y_2 = 1|Y_1 = 0\), which contradicts the \(Y_1 = 1\) from above. Then the event \(V_2 + \alpha < -x_2\beta - \gamma\) corresponds to \(Y_2 = 0|Y_1 = 0\) which again contradicts \(Y_1 = 1\). Therefore, \(V^{**} = \emptyset\) and the dynamic model in (27) is incoherent. Figure 2 shows the regions \(V^*\) and \(V^{**}\).

\[
\begin{array}{ccc}
V_2 + \alpha & (0,1)|Y_0 = 0 & (0,1)|Y_0 = 0 \\
\text{or} & (1,1)|Y_0 = 0 & \text{or} \\
\hline
-x_2\beta - \gamma & (0,0)|Y_0 = 0 & (1,0)|Y_0 = 0 \\
\text{or} & (0,1)|Y_0 = 0 & \text{or} \\
-x_2\beta & (0,0)|Y_0 = 0 & (0,0)|Y_0 = 0 \\
\hline
-x_1\beta & (0,0)|Y_0 = 0 & (0,0)|Y_0 = 0 \\
-x_1\beta - \gamma & (1,0)|Y_0 = 0 & (1,0)|Y_0 = 0 \\
V_1 + \alpha & (0,0)|Y_0 = 0 & (0,0)|Y_0 = 0 \\
\end{array}
\]

Figure 2: Regions of unobservables for each \((Y_1, Y_2)\) choice when \(\gamma < 0\) and \(Y_0 = 0\)

The important feature of model in (27) is the presence of current period’s outcome in the generation of current period’s utility. Unless \(\gamma = 0\), the model in (27) is logically inconsistent as discussed in Maddala (1986). Such a model would resemble models of simultaneous response.
games like Tamer (2003) and more general binomial response models with dummy endogenous regressors as in Lewbel (2007). The incompleteness and incoherency generally result in an inability to obtain point-identification of the regression parameters. Chesher and Rosen (2012) examine identification in incomplete and incoherent models and the application of these results to the model in (27) is left for future research.

3 The Ordered Choice Panel Data Model

Section 2 studies identification in binary panel data models. As discussed in Section 1 several papers, such as Chintagunta, Kyriazidou, and Perktold (2001) and Pakes and Porter (2014), have extended the binary response panel data model to multinomial response models, where individuals choose from a set of unordered alternatives. This paper extends the binary response panel data model to one where the choice set consists of alternatives that can be ordered, such as the choice between unemployment, part-time employment or full-time employment and the choice of flying first, business or economy class. This approach might be beneficial in reducing the dimension of search for the identified set, since the shape restriction imposed by the ordering, reduces the between alternatives comparisons needed to determine the option chosen by the individual.

This section extends the model (2) to a model of three ordered outcomes, where in every period \( t = 1 \) and \( t = 2 \) each individual chooses one option from the set \( Y_t = \{0, 1, 2\} \). Such a model could be used, for example, in describing consumers’ choices when faced with vertically differentiated alternatives such that if all the options were offered at the same price, \( (Y_t = 1) \preceq (Y_t = 2) \), where \( A \preceq B \) denotes that \( B \) is weakly preferred to \( A \). \( Y_t = 0 \) denotes the outside option and corresponds to not choosing any of the available alternatives. By extending the model in (2) to include options that are ordered the choice rule for each individual can be expressed as,

\[
Y_t = \begin{cases} 
0 & \text{if } g(X_t, Y_{t-1}, \alpha, V_t) < c_{1t} \\
1 & \text{if } c_{1t} < g(X_t, Y_{t-1}, \alpha, V_t) \leq c_{2t} \\
2 & \text{if } c_{2t} < g(X_t, Y_{t-1}, \alpha, V_t)
\end{cases}
\] (32)
where $x_t$ are observed individual characteristics, $y_{t-1}$ is last period’s choice, $\alpha$ is the unobserved to the econometrician time-invariant individual heterogeneity, $v_t$ is the time-varying component unobserved to the econometrician and $g()$ is a function not yet specified, which is an element of some parameter space $\Theta$. $c = \{c_{11}, c_{12}, c_{21}, c_{22}\} \in C$ are the threshold parameters in the ordered model, such that $C \subseteq \mathbb{R}^4$ and $c_{2t} > c_{1t}, \forall t \in \{1, 2\}$. They can be interpreted as the disutility parameters, exogenously determined and common to all individuals, incurred in period $t$ to obtain quality level $y$. For the rest of the paper the vector $c$ is assumed to be observed. In application these parameters are more likely to be unknown and therefore included in the identified sets. The simplification is done to reduce the dimension of the identified set for the model parameters. In future research and application the restriction that the threshold are observed will be relaxed.

As already discussed in Section 1 the ordered response panel data model has not been extensively studied in the literature and the work has mainly focused in redefining the ordered response model as a set of binary response models and imposing logistically distributed unobservables. This paper departs from this approach and uses directly the ordered structure of the model to characterize the identified set, without imposing distributional assumption on the unobserved time-varying components or the fixed effect. Such an approach uses more information than the binary response representation and provides informative identification bounds for the regression parameters.

### 3.1 Static Ordered Panel Data Model

This section examines identification of model (32) with no lagged dependent variables where correlation comes only through the presence of the unobserved heterogeneity. Applying Assumption 6 of linearity in the variables to (32), the static ordered response model for the periods $t = 1$ and $t = 2$ can be expressed as,
\[ Y_t = \begin{cases} 
0 & \text{if } X_t \beta + \alpha + V_t < c_1 \\
1 & \text{if } c_1 < X_t \beta + \alpha + V_t \leq c_2 \\
2 & \text{if } c_2 < X_t \beta + \alpha + V_t 
\end{cases} \quad (33) \]

Under Assumptions 1-7 the identified set of admissible structures, denoted by \( S^0 \), is characterized. Define by \( u(\tilde{y}_1, \tilde{y}_2, x, \alpha, v_1, v_2; \beta) \), the utility an individual receives from choosing any pair \((\tilde{y}_1, \tilde{y}_2) \) given the observables \( x, c \) and the unobservables \( \alpha, v_1, v_2 \) and by \( R_{(y_1, y_2)}(x, c; \beta) \) the region of unobservables \((V, \alpha)\) associated with \((y_1, y_2)\) maximizing \( u(\tilde{y}_1, \tilde{y}_2, x, c, \alpha, v_1, v_2; \beta), \) such that

\[ R_{(y_1, y_2)}(x, c; \beta) \equiv \{(V, \alpha) \in (\mathcal{V}, \mathcal{A}) : (y_1, y_2) = \arg \max_{(\tilde{y}_1, \tilde{y}_2) \in (\mathcal{Y}_1 \times \mathcal{Y}_2)} u(\tilde{y}_1, \tilde{y}_2, x, c, \alpha, v_1, v_2; \beta)\}. \]

Then \( S^0 \) is characterized by:

\[
S^0 = \left\{ (\beta, F_{(V, \alpha)|X,c}) \in S : \forall (y_1, y_2) \in (\mathcal{Y}_1 \times \mathcal{Y}_2) \right. \\
\left. F_{(V, \alpha)|X,c}(R_{(y_1, y_2)}(x, c; \beta)) = P^0(y_1, y_2| x, c) \right. \\
a.e. \ x \in \mathcal{X} \ and \ c \in \mathcal{C}
\]

and the identified set for the model parameters \( \beta \) is given by:

\[
\Theta^0 = \left\{ \beta \in \Theta : \forall (y_1, y_2) \in (\mathcal{Y}_1, \mathcal{Y}_2) \right. \\
\left. \exists F_{(V, \alpha)|X,c} \in F_{(V, \alpha)|X,c} 
\right.
\left. F_{(V, \alpha)|X,c}(R_{(y_1, y_2)}(x, c; \beta)) = P^0(y_1, y_2| x, c) \right. \\
a.e. \ x \in \mathcal{X} \ and \ c \in \mathcal{C}
\]

### 3.1.1 Identified Set: Ordered Response Static Panel Data Model

As in Section 2, no assumptions are imposed on the fixed effect and identification bounds for the regression parameters, \( \beta \), are derived by finding feature of the distribution that do not depend on \( \alpha \). The regions \( R^{SO}_{(y_1, y_2)}(x, c; \beta) \) that partition the support of \((V, \alpha)\) such that \((Y_1, Y_2) = (y_1, y_2)\) when \( X = x \) and for any fixed \( c \), are given by:
\( R_{(0,0)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& x_1 \beta + \alpha + V_1 - c_{11} \leq 0 \\
& x_2 \beta + \alpha + V_2 - c_{12} \leq 0 \end{cases} \)

\( R_{(0,1)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& x_1 \beta + \alpha + V_1 - c_{11} \leq 0 \\
& -x_2 \beta + c_{12} < \alpha + V_2 \leq -x_2 \beta + c_{22} \end{cases} \)

\( R_{(0,2)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& x_1 \beta + \alpha + V_1 - c_{11} \leq 0 \\
& 0 < x_2 \beta + \alpha + V_2 - c_{22} \end{cases} \)

\( R_{(1,0)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& -x_1 \beta + c_{11} < \alpha + V_1 \leq -x_1 \beta + c_{21} \\
& x_2 \beta + \alpha + V_2 - c_{12} \leq 0 \end{cases} \)

\( R_{(1,1)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& -x_1 \beta + c_{11} < \alpha + V_1 \leq -x_1 \beta + c_{21} \\
& -x_2 \beta + c_{12} < \alpha + V_2 \leq -x_2 \beta + c_{22} \end{cases} \)

\( R_{(1,2)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& -x_1 \beta + c_{11} < \alpha + V_1 \leq -x_1 \beta + c_{21} \\
& 0 < x_2 \beta + \alpha + V_2 - c_{22} \end{cases} \)

\( R_{(2,0)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& 0 < x_1 \beta + \alpha + V_1 - c_{21} \\
& x_2 \beta + \alpha + V_2 - c_{12} \leq 0 \end{cases} \)

\( R_{(2,1)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& 0 < x_1 \beta + \alpha + V_1 - c_{21} \\
& -x_2 \beta + c_{12} < \alpha + V_2 \leq -x_2 \beta + c_{22} \end{cases} \)

\( R_{(2,2)}^{SO}(x, c; \beta) = \begin{cases} (V, \alpha) \in (V, A) : \\
& 0 < x_1 \beta + \alpha + V_1 - c_{21} \\
& 0 < x_2 \beta + \alpha + V_2 - c_{22} \end{cases} \)
The conditional joint probabilities of \((V, \alpha) | X = x, c, F_{(V, \alpha) | X, c} \) for any given pair \((y_1, y_2) \in (Y_1 \times Y_2)\), conditional on \(x \in \mathcal{X}\) and for any fixed \(c \in \mathcal{C}\) are given by

\[
P(y_1, y_2 | x, c) = F_{(V, \alpha) | X, c} \left( R_{SO}^{(y_1, y_2)} (x, c; \beta) | X = x, c \right)
\]

where \(P(y_1, y_2 | x, c) = P(Y_1 = y_1 \land Y_2 = y_2 | X = x, c)\).

From the regions defined in (34) and plotted in Figure 3 it is easy to see that for any fixed vector \(c\) the model in (33) is complete in the sense that conditional on any value of exogenous variables \(x \in \mathcal{X}\), there is a unique solution to the individual decision problem with probability one.

Follow the same line of argumentation as in the case of the binary panel data model, individuals who switch their choice from period \(t = 1\) to \(t = 2\), can be used for identification of the parameters \(\beta\), without imposing any assumptions on the fixed effect. This implies that

Figure 3: Regions of unobservables for each \((Y_1, Y_2)\) choice for a fixed vector \(c\)
the transitions that are informative are,

\begin{align*}
\{Y_1 = 0 \land Y_2 = 1\} \\
\{Y_1 = 1 \land Y_2 = 0\} \\
\{Y_1 = 0 \land Y_2 = 2\} \\
\{Y_1 = 2 \land Y_2 = 0\} \\
\{Y_1 = 1 \land Y_2 = 2\} \\
\{Y_1 = 2 \land Y_2 = 1\}
\end{align*}

(35)

In addition to the individuals who change from period \( t = 1 \) to \( t = 2 \) it can be shown that information that is independent of \( \alpha \) is also provided by considering individuals who choose that same option, \( Y = 1 \), in periods \( t = 1 \) and \( t = 2 \) such that,

\begin{align*}
\{Y_1 = 1 \land Y_2 = 1\}
\end{align*}

(36)

This might prove helpful when comparing the behaviour of switchers to non-switchers.

**Theorem 8.** Let \( S^{SO} = (\beta^{SO}, F^{SO}_{(V,\alpha)\mid X, c}) \) be a structure admitted by model (33). Under Assumptions [4][7] for any fixed parameter vector \( c \), if \( S^{SO} \) is an observationally equivalent structure to \( S^0 \), then for any \( x \in X \) \( \beta^{SO} \) satisfies

\begin{align*}
P(1, 0 \mid x, c) & \leq F_{\Delta V}(-\Delta x \beta^{SO} - c_{11} + c_{12}) \leq 1 - P(0, 1 \mid x, c) \\
P(2, 0 \mid x, c) & \leq F_{\Delta V}(-\Delta x \beta^{SO} + c_{12} - c_{21}) \\
& \quad F_{\Delta V}(-\Delta x \beta^{SO} - c_{11} + c_{22}) \leq 1 - P(0, 2 \mid x, c) \\
P(2, 1 \mid x, c) & \leq F_{\Delta V}(-\Delta x \beta^{SO} + c_{22} - c_{21}) \leq 1 - P(1, 2 \mid x, c) \\
P(1, 1 \mid x, c) & \leq P_{\Delta V}(-\Delta x \beta^{SO} + c_{22} - c_{11} > \Delta V > -\Delta x \beta^{SO} + c_{12} - c_{21})
\end{align*}

where \( \Delta X = X_2 - X_1 \) and \( \Delta V = V_2 - V_1 \).

**Proof.** The proof that the conditional probabilities in \( P(y_1, y_2 \mid x, c) \) in Theorem 8 are bounded by features of the distribution that do not depend on the fixed effect \( \alpha \) is provided in Appendix.
Notice that in addition to the probabilities of the switching events, the conditional probability of the “in-between” event $(1, 1)$ is also bounded by a conditional probability invariant to the fixed effect. The rationale behind this result is that, as discussed in Section 2, in the binary model the events $(Y_1, Y_2) = (0, 0)$ and $(1, 1)$ give no information on $\beta$ since the behaviour for these “extreme” cases can be matched by extremely small or extremely large values of $\alpha$ regardless of the value of $\beta$. This is also true for the ordered model for the events $(0, 0)$ and $(2, 2)$. However, as in can also be seen from Figure 3 in the ordered response model where $\mathcal{Y}_t = \{0, 1, 2\}$, considered in this section, the middle event $(Y_1, Y_2) = (1, 1)$ provides restrictions on the possible values the fixed effect can take for each value of the parameter $\beta$, and hence it can be used to identify the regression coefficients.

\[\text{Theorem 9.}\] Let Assumptions 1-7 hold, for any fixed parameter vector $c$. Using the definitions in (37), an outer region for $\beta$ is given by the set:

$$\Theta^{SO} = \left\{ \beta \in \Theta : \forall \omega, \omega', \omega'' \in \mathbb{R}, \right. $$

$$\left. \max \{s(1,0)(\omega), s(2,0)(\omega), s(2,1)(\omega), s(1,1)(\omega)\} \leq \min \{i(0,1)(\omega), i(0,2)(\omega), i(1,2)(\omega), i(1,1)(\omega)\} \right\} a.e. \ x \in \mathcal{X} \ \text{and} \ c \in \mathcal{C}$$

where,

$$s(1,0)(\omega) = \sup_{x: -\Delta x \beta - c_{11} + c_{12} \leq \omega} P(1, 0|x, c)$$

$$s(2,0)(\omega) = \sup_{x: -\Delta x \beta + c_{12} - c_{21} \leq \omega} P(2, 0|x, c)$$

$$s(2,1)(\omega) = \sup_{x: -\Delta x \beta + c_{22} - c_{21} \leq \omega} P(2, 1|x, c)$$

$$s(1,1)(\omega) = \sup_{x \in \mathcal{X}^*} P(1, 1|x, c)$$

$$i(0,1)(\omega) = \inf_{x: -\Delta x \beta - c_{11} + c_{12} \geq \omega} [1 - P(0, 1|x, c)]$$

$$i(0,2)(\omega) = \inf_{x: -\Delta x \beta - c_{11} + c_{22} \geq \omega} [1 - P(0, 2|x, c)]$$

$$i(1,2)(\omega) = \inf_{x: -\Delta x \beta + c_{22} - c_{21} \geq \omega} [1 - P(1, 2|x, c)]$$

$$i(1,1)(\omega) = 1 - \inf_{x \in \mathcal{X}^*} P(1, 1|x, c)$$

(37)
and for any fixed \( \omega, \omega', \omega'' \in \mathbb{R} \)

\[
X^* = \{ x : \omega' \geq -\Delta x \beta + c_{22} - c_{11} - \Delta x \beta + c_{12} - c_{21} \geq \omega \}
\]

\[
X^{**} = \{ x : \omega \geq -\Delta x \beta + c_{22} - c_{11} - \Delta x \beta + c_{12} - c_{21} \geq \omega'' \}
\]

(38)

**Proof.** The proof is provided in Appendix A.5. Notice that following Theorem 8, information on the parameters \( \beta \) comes also from the non-switchers who choose option \((Y_1, Y_2) = (1, 1)\).

The identified set in Theorem 9 provides an outer region of the projection of the joint identified set for \((\beta, F_{(V, \alpha)|X,c})\), defined in the theorem below.

**Theorem 10.** Let Assumptions 1-7 hold for any fixed parameter vector \( c \). Then if \( \beta \in \Theta^S \), then there exists a structure \( S^S = (\beta^S, F^S_{(V, \alpha)|X,c}) \) that satisfies the restrictions of the model and is observationally equivalent to structure \( S_0 \) that generates \( P_0(y_1, y_2|x, c) \). The identified set for \( \beta \) and \( F_{(V, \alpha)|X,c} \) is given by:

\[
S^S = \left\{ (\beta, F_{(V, \alpha)|X,c}) \in (\Theta, F_{(V, \alpha)|X,c}) : \forall \omega, \omega', \omega'' \in \mathbb{R}, \text{s.t. the relations in (39) hold} \right\}
\]

where

\[
\begin{align*}
\sup_{x : -\Delta x \beta + c_{11} \geq \omega} P(1, 0|x, c) &\leq F_{\Delta V}(\omega) \\
F_{\Delta V}(\omega) &\leq \inf_{x : -\Delta x \beta + c_{12} \geq \omega} [1 - P(0, 1|x, c)] \\
\sup_{x : -\Delta x \beta + c_{22} \geq \omega} P(2, 0|x, c) &\leq F_{\Delta V}(\omega) \\
F_{\Delta V}(\omega) &\leq \inf_{x : -\Delta x \beta + c_{22} \geq \omega} [1 - P(0, 2|x, c)] \\
\sup_{x : -\Delta x \beta + c_{21} \geq \omega} P(2, 1|x, c) &\leq F_{\Delta V}(\omega) \\
F_{\Delta V}(\omega) &\leq \inf_{x : -\Delta x \beta + c_{22} \geq \omega} [1 - P(1, 2|x, c)] \\
\sup_{x \in X^*} P(1, 1|x, c) &\leq F_{\Delta V}(\omega') - F_{\Delta V}(\omega) \\
\sup_{x \in X^{**}} P(1, 1|x, c) &\leq F_{\Delta V}(\omega) - F_{\Delta V}(\omega'')
\end{align*}
\]

(39)

Notice that Theorem 10 does not characterize the sharp identified set. The question of sharpness remains an open question to be addressed in future work.
3.2 Dynamic Ordered Panel Data Model

Persistence in choice is evident when modeling individual behaviour in many different cases. For example, in modeling consumer demand in markets with differentiated products where consumers are observed purchasing products regularly it is often observed that consumers choices are linked intertemporally. As discussed in Section 1 in addition to the spurious state dependence, this correlation can also come through a true state dependence where past choices directly affect the utility function. This is also true in cases where consumers purchase a specific product regularly and decisions are over sets of differentiated alternatives. As discussed in Dubé, Hitsch, and Rossi (2009, 2010) the persistence in consumers brand and quality choices is evident even after controlling for the fixed effects. A common explanation is that consumers face monetary, psychological and search costs, that induce a lock-in effect, loyalty in specific alternatives and habit formation.

To capture these effects the model in (32) is examined where the lagged dependent variable enters the functional form of the individual choice rule. Applying Assumption (13) when $Y_t \equiv \{0, 1, 2\}$ the model in (32) can be expressed as,

\[
Y_t = \begin{cases} 
0 & \text{if } X_t \beta + 1(Y_{t-1} = 1)\gamma_1 + 1(Y_{t-1} = 2)\gamma_2 + \alpha + V_t < c_{1t} \\
1 & \text{if } c_{1t} < X_t \beta + 1(Y_{t-1} = 1)\gamma_1 + 1(Y_{t-1} = 2)\gamma_2 + \alpha + V_t \leq c_{2t} \\
2 & \text{if } c_{2t} < X_t \beta + 1(Y_{t-1} = 1)\gamma_1 + 1(Y_{t-1} = 2)\gamma_2 + \alpha + V_t 
\end{cases}
\]  

Similarly to the dynamic binary response panel data model, the presence of the lagged dependent variables results to an initial conditions problem. This paper does not examine the initial condition problem, but instead it is assumed that the outcome in the initial period $t = 0$, $y_0 \in Y_0 \equiv \{0, 1, 2\}$ is known, however the model is not specified and the explanatory variables do not need to be known. Notice that since the choice set now is $Y_0 \equiv \{0, 1, 2\}$ the parameters of interest are $(\beta, \gamma) \in \Theta$, with $\gamma = (\gamma_1, \gamma_2)^T$.

Under Assumptions 8-14 the identified set of admissible structures, denoted by $S^0$, is char-
acterized. Define by \( u(\tilde{y}_1, \tilde{y}_2, x, y_0, c, \alpha, v; \beta, \gamma) \), the utility an individual receives from choosing any pair \((\tilde{y}_1, \tilde{y}_2) \in (Y_1, Y_2)\) given the observables \(x, c, y_0\) and the unobservables \(\alpha, v\) and by \(R_{(y_1, y_2)}(x, y_0, c; \beta, \gamma)\) the region of unobservables \((V, \alpha)\) associated with \((y_1, y_2)\) maximizing \(u(\tilde{y}_1, \tilde{y}_2, x, y_0, c, \alpha, v_1, v_2; \beta, \gamma)\), such that

\[
R_{(y_1, y_2)}(x, y_0, c; \beta, \gamma) \equiv \{(V, \alpha) \in (V, A) : (y_1, y_2) = \arg \max_{(\tilde{y}_1, \tilde{y}_2) \in (Y_1 \times Y_2)} u(\tilde{y}_1, \tilde{y}_2, x, y_0, c, \alpha, v_1, v_2; \beta, \gamma)\}.
\]

Then \(S^0\) is characterized by:

\[
S^0 = \left\{ (\beta, \gamma, F_{(V, \alpha)|X,Y_0,c}) \in S : \forall (y_1, y_2) \in (Y_1 \times Y_2) \right. \\
F_{(V, \alpha)|X,Y_0,c}(R_{(y_1, y_2)}(x, y_0, c; \beta, \gamma)) = P^0(y_1, y_2|x, y_0, c) \\
a.e. x \in X, y_0 \in Y_0 and c \in C \right\}
\]

and the identified set for the model parameters \(\beta, \gamma\) is given by:

\[
\Theta^0 = \left\{ (\beta, \gamma) \in \Theta : \forall (y_1, y_2) \in (Y_1, Y_2), \exists F_{(V, \alpha)|X,Y_0,c} \in F_{(V, \alpha)|X,Y_0,c} \\
F_{(V, \alpha)|X,Y_0,c}(R_{(y_1, y_2)}(x, y_0, c; \beta, \gamma)) = P^0(y_1, y_2|x, y_0, c) \\
a.e. x \in X, y_0 \in Y_0 and c \in C \right\}
\]

3.2.1 Identified Set: Dynamic Ordered Response Model

As in Section 2.2, the initial condition \(Y_0\) is assumed to be known, however no information for its generation is assumed, and the joint probabilities for a sequence of events in periods \(t = 1\) and \(t = 2\) will be conditional on the choice in period \(t = 0\). As the model is complete, the regions \(R_{DO(y_1, y_2)}(x, y_0, c; \beta, \gamma)\) such that for all \((V, \alpha) \in (V, A)\), \((Y_1, Y_2) = (y_1, y_2)\) when \(X = x\), \(Y_0 = y_0\) and \(c\), partition the support of \((V, \alpha)\). The conditional joint probability of \((V, \alpha)|X, Y_0, c\), when \(X = x\), \(Y_0 = y_0\) and \(c\) is fixed, for any pair \((y_1, y_2) \in Y_1 \times Y_2\), is given by

\[
P(y_1, y_2|x, y_0, c) = F_{(V, \alpha)|X,Y_0,c}(R_{DO(y_1, y_2)}(x, y_0, c; \beta, \gamma)|X = x, Y_0 = y_0, c)
\]

where \(P(y_1, y_2|x, y_0, c) = P(Y_1 = y_1 \land Y_2 = y_2|X = x, Y_0 = y_0, c)\).

As in Section 3.1 for the dynamic ordered panel data model, for any value of \(Y_0 = y_0\), with no assumption on the fixed effect, restrictions on the realization of \(\Delta V|X, Y_0, c\) and
hence information for the values $\beta, \gamma$ that does not depend on the unobserved heterogeneity parameter are provided by the following observations,

\begin{align*}
\{Y_0 = y_0 & \land Y_1 = 0 & \land Y_2 = 1\} \\
\{Y_0 = y_0 & \land Y_1 = 1 & \land Y_2 = 0\} \\
\{Y_0 = y_0 & \land Y_1 = 0 & \land Y_2 = 2\} \\
\{Y_0 = y_0 & \land Y_1 = 2 & \land Y_2 = 0\} \\
\{Y_0 = y_0 & \land Y_1 = 2 & \land Y_2 = 2\} \\
\{Y_0 = y_0 & \land Y_1 = 2 & \land Y_2 = 1\} \\
\{Y_0 = y_0 & \land Y_1 = 1 & \land Y_2 = 1\}
\end{align*}

(41)

Appendix A.6 proves that the probabilities of the sequences of events described in (41) conditional on a specific value of $y_0 \in \{0, 1, 2\}$, $X = x$ and $c \in C$ are bounded by the features of the distribution that do not depend on the fixed effect. Theorem 11 formalizes the identified set for $(\beta, \gamma)$.

**Theorem 11.** Let Assumptions 8-14 hold for any fixed parameter vector $c$. The outer region for $\beta$ and $\gamma$ is given by the set:

\[
\Theta^{DO} = \left\{ (\beta, \gamma) \in \Theta : \forall \omega, \omega', \omega'' \in \mathbb{R}, \max \left[ s_{(1,0)}(\omega, y_0), s_{(2,0)}(\omega, y_0), s_{(2,1)}(\omega, y_0), s_{(1,1)}(\omega, y_0) \right] \leq \min \left[ i_{(0,1)}(\omega, y_0), i_{(0,2)}(\omega, y_0), i_{(1,2)}(\omega, y_0), i_{(1,1)}(\omega, y_0) \right] \right\}
\]

where

\[
i_{(0,1)}(\omega, 0) = \inf_{x: -\Delta x \beta - c_{11}, c_{12} \geq \omega} [1 - P(0, 1|x, 0, c)] \\
s_{(1,0)}(\omega, 0) = \sup_{x: -\Delta x \beta - \gamma_1 + c_{12} - c_{21} \leq \omega} P(1, 0|x, , 0, c)
\]
\[ i(0, 2)(\omega, 0) = \inf_{x : -\Delta x \beta - c_{11} + c_{22} \geq \omega} \left[ 1 - P(0, 2 | x, 0, c) \right] \]
\[ s_{(2, 0)}(\omega, 0) = \sup_{x : -\Delta x \beta - \gamma_2 + c_{12} - c_{21} \leq \omega} \left( P(2, 0 | x, 0, c) \right) \]
\[ i(1, 2)(\omega, 0) = \inf_{x : -\Delta x \beta - \gamma_1 + c_{12} - c_{21} \geq \omega} \left[ 1 - P(1, 2 | x, 0, c) \right] \]
\[ s_{(2, 1)}(\omega, 0) = \sup_{x : -\Delta x \beta - \gamma_2 + c_{12} - c_{21} \leq \omega} \left( P(2, 1 | x, 0, c) \right) \]
\[ i(1, 1)(\omega, 0) = \inf_{x \in X_0^*} \left[ 1 - P(1, 1 | x, 0, c) \right] \]
\[ s_{(1, 1)}(\omega, 0) = \sup_{x \in X_0^{**}} \left( P(1, 1 | x, 0, c) \right) \]
\[ i(0, 1)(\omega, 1) = \inf_{x : -\Delta x \beta + \gamma_1 - c_{11} + c_{12} \geq \omega} \left[ 1 - P(0, 1 | x, 1, c) \right] \]
\[ s_{(1, 0)}(\omega, 1) = \sup_{x : -\Delta x \beta + c_{12} - c_{11} \leq \omega} \left( P(1, 0 | x, 1, c) \right) \]
\[ i(0, 2)(\omega, 1) = \inf_{x : -\Delta x \beta + \gamma_1 - c_{11} + c_{22} \geq \omega} \left[ 1 - P(0, 2 | x, 1, c) \right] \]
\[ s_{(2, 0)}(\omega, 1) = \sup_{x : -\Delta x \beta + \gamma_2 + c_{12} - c_{21} \leq \omega} \left( P(2, 0 | x, 1, c) \right) \]
\[ i(1, 2)(\omega, 1) = \inf_{x : -\Delta x \beta + c_{12} - c_{21} \geq \omega} \left[ 1 - P(1, 2 | x, 1, c) \right] \]
\[ s_{(2, 1)}(\omega, 1) = \sup_{x \in X_1^*} \left( P(1, 1 | x, 1, c) \right) \]
\[ i(0, 1)(\omega, 2) = \inf_{x : -\Delta x \beta + \gamma_1 - c_{11} + c_{12} \geq \omega} \left[ 1 - P(0, 1 | x, 2, c) \right] \]
\[ s_{(1, 0)}(\omega, 2) = \sup_{x : -\Delta x \beta - \gamma_2 + c_{12} - c_{11} \leq \omega} \left( P(1, 0 | x, 2, c) \right) \]
\[ i(0, 2)(\omega, 2) = \inf_{x : -\Delta x \beta + \gamma_2 + c_{12} - c_{21} \geq \omega} \left[ 1 - P(0, 2 | x, 2, c) \right] \]
\[ s_{(2, 0)}(\omega, 2) = \sup_{x : -\Delta x \beta + c_{22} - c_{21} \leq \omega} \left( P(2, 0 | x, 2, c) \right) \]
\[ i(1, 2)(\omega, 2) = \inf_{x : -\Delta x \beta + c_{12} - c_{21} \leq \omega} \left[ 1 - P(1, 2 | x, 2, c) \right] \]
\[ s_{(2, 1)}(\omega, 2) = \sup_{x \in X_2^*} \left( P(1, 1 | x, 2, c) \right) \]
\[ i(1, 1)(\omega, 2) = \inf_{x \in X_2^{**}} \left[ 1 - P(1, 1 | x, 2, c) \right] \]
\[ s_{(1, 1)}(\omega, 2) = \sup_{x \in X_2^{**}} \left( P(1, 1 | x, 2, c) \right) \]
with

\[ X_0^* = \{ x : \omega' \geq -\Delta x\beta - \gamma_1 + c_{22} - c_{11} \land -\Delta X\beta - \gamma_1 + c_{12} - c_{21} \geq \omega \} \]

\[ X_0^{**} = \{ x : \omega \geq -\Delta x\beta - \gamma_1 + c_{22} - c_{11} \land -\Delta X\beta - \gamma_1 + c_{12} - c_{21} \geq \omega'' \} \]

\[ X_1^* = \{ x : \omega' \geq -\Delta x\beta + c_{22} - c_{11} \land -\Delta x\beta + c_{12} - c_{21} \geq \omega \} \]

\[ X_1^{**} = \{ x : \omega \geq -\Delta x\beta + c_{22} - c_{11} \land -\Delta x\beta + c_{12} - c_{21} \geq \omega'' \} \]

\[ X_2^* = \{ x : \omega' \geq -\Delta x\beta - \gamma_1 + \gamma_2 + c_{22} - c_{11} \land -\Delta x\beta - \gamma_1 + \gamma_2 + c_{12} - c_{21} \geq \omega \} \]

\[ X_2^{**} = \{ x : \omega \geq -\Delta x\beta - \gamma_1 + \gamma_2 + c_{22} - c_{11} \land -\Delta x\beta - \gamma_1 + \gamma_2 + c_{12} - c_{21} \geq \omega'' \} \]

**Proof.** The proof is similar to the proof of Theorem 9 and therefore omitted. 

4 Numerical Examples

This section provides numerical illustrations of the identified sets derived in the previous sections for the static and dynamic binary response models under different support conditions and Probability Generating Processes (PGP). All the models examined in this section have discrete support for the explanatory variables, \( X \). Even though point identification fails, the numerical examples illustrate that informative bounds can be achieved as the support of the discrete explanatory variables increases.

4.1 Static Binary Panel Data Model

Consider the two time-period static binary panel data model,

\[ Y_t = 1(X_t\beta + \alpha + V_t > 0) \]

(42)

where \( X_t = (X_{1t}, X_{2t}) \), \( \beta = (\beta_1, \beta_2) \), \( \alpha | X \sim N(\bar{X}\delta, \sigma_\alpha^2) \) with \( \bar{X} = \frac{1}{2}(X_1 + X_2) \) and \( \delta = (1, -1) \) and \( V_t | X, \alpha \text{iid} f() \). The true parameter of \( \beta_2 = 1 \) after the normalization of \( \beta_1 = 1 \) and \( \sigma_\alpha^2 = 1 \).

For the purpose of comparison the baseline PGP is similar to the one used in RW2013\(^3\).

---

\(^3\) The Normal distribution for the fixed effect was approximated on a grid with 100 evenly spaced support points on \([\bar{X}\delta - 4, \bar{X}\delta + 4]\) as in RW2013, and the approximation error should be small.
Tables 1 and 2 give the identified sets for $\beta_2$ as described in Theorem 2 under different specification for the distribution of the time varying unobservables and as the support of the discrete explanatory variables $(X_{1t}, X_{2t})$ changes. The first Probit specification in Table 1 with $V_t | X, \alpha \overset{iid}{\sim} N(0, 1)$ is the same as in RW2013. In addition to the specification in RW2013 Table 1 also provides identified sets under the probit specification with $V_t \overset{iid}{\sim} N \left(0, \frac{\pi^2}{3}\right)$ and the standard logit specification $V_t \perp X, \alpha$ with iid logistic distribution.

Table 1: Identified sets for $\beta_2$ under Probit and Logit specifications with symmetric support for $(X_{1t}, X_{2t})$ around zero

<table>
<thead>
<tr>
<th>Support of $(X_{1t}, X_{2t})$</th>
<th>$V \sim N(0, 1)$</th>
<th>$V \sim N \left(0, \frac{\pi^2}{3}\right)$</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$(0, \infty)$</td>
<td>$(-0.5, \infty)$</td>
<td>$(-0.5, \infty)$</td>
</tr>
<tr>
<td>${0, 1}$</td>
<td>$\infty$</td>
<td>$2.38$</td>
<td>$2.38$</td>
</tr>
<tr>
<td>${0, 1, 2}$</td>
<td>$0.5$</td>
<td>$0.667$</td>
<td>$0.667$</td>
</tr>
<tr>
<td>${0, 1, 2, 3}$</td>
<td>$1.5$</td>
<td>$0.833$</td>
<td>$0.833$</td>
</tr>
<tr>
<td>${0, 1, 2, 3, 4}$</td>
<td>$(0.5, 2)$</td>
<td>$(0.29, 2.67)$</td>
<td>$(0.29, 2.67)$</td>
</tr>
<tr>
<td>${0, 1, 2, 3, 4, 5}$</td>
<td>$(0.667, 1.5)$</td>
<td>$(0.55, 1.71)$</td>
<td>$(0.55, 1.71)$</td>
</tr>
</tbody>
</table>

Table 2: Identified sets for $\beta_2$ under Probit and Logit specifications with asymmetric support for $(X_{1t}, X_{2t})$ around zero

<table>
<thead>
<tr>
<th>Support of $(X_{1t}, X_{2t})$</th>
<th>$V \sim N(0, 1)$</th>
<th>$V \sim N \left(0, \frac{\pi^2}{3}\right)$</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$(0.5, 6)$</td>
<td>$(0.286, \infty)$</td>
<td>$(0.286, \infty)$</td>
</tr>
<tr>
<td>${0, 1, 2}$</td>
<td>$5.5$</td>
<td>$1.333$</td>
<td>$1.333$</td>
</tr>
<tr>
<td>${0, 1, 2, 3}$</td>
<td>$1.333$</td>
<td>$1.955$</td>
<td>$1.955$</td>
</tr>
<tr>
<td>${0, 1, 2, 3, 4}$</td>
<td>$(0.667, 2)$</td>
<td>$(0.545, 2.5)$</td>
<td>$(0.545, 2.5)$</td>
</tr>
<tr>
<td>${0, 1, 2, 3, 4, 5}$</td>
<td>$1.333$</td>
<td>$1.955$</td>
<td>$1.955$</td>
</tr>
</tbody>
</table>

42
From Tables 1 and 2 two main conclusions can be drawn. The first one is that as the support of the explanatory variables increases the identified sets become narrower. This suggests that even thought the model only partially identify the regression parameters, those sets shrink around the true value as the support of the explanatory variables increases.

Secondly, it is evident that the model with $V_t \sim \mathcal{N}(0, \frac{\pi^2}{3})$ errors and the standard logit model with variance of the unobservables $\text{var}(V_t) = \frac{\pi^2}{3}$ give almost identical identified sets. Since the two distributions are very similar, they tend to behave in a same way. This might have implications on the identification and estimation of such models, not suggested by the theoretical results.

4.1.1 Changing PGP: Variance of unobservables

This section provides identified sets for $\beta_2$ under different PGP for the variance of the time varying unobservables and the fixed effect when the support of $(X_{1t}, X_{2t})$ is $\{-2, -1, 0, 1, 2\}$.

$$
\begin{pmatrix}
\alpha \\
V_t
\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}
\bar{X}\delta \\
0
\end{pmatrix}, \begin{pmatrix}
\sigma_\alpha^2 & 0 \\
0 & \sigma^2
\end{pmatrix}\right)
$$

Table 3: Identified sets for $\beta_2$ under different PGP-Variances

<table>
<thead>
<tr>
<th>Variance of $V_t$ when $\sigma_\alpha = 1$</th>
<th>$\sigma = 0.1$</th>
<th>$\sigma = 0.5$</th>
<th>$\sigma = 1.5$</th>
<th>$\sigma = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.57,1.6)</td>
<td>(0.57,1.6)</td>
<td>(0.38,2.67)</td>
<td>(0.25,4)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance of $\alpha$ when $\sigma = 1$</th>
<th>$\sigma_\alpha = 0$</th>
<th>$\sigma_\alpha = 0.5$</th>
<th>$\sigma_\alpha = 1.5$</th>
<th>$\sigma_\alpha = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.57,1.6)</td>
<td>(0.57,1.6)</td>
<td>(0.29,2.67)</td>
<td>(0.125,8)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 gives the identified set as expressed in Theorem 2 when the true variances of the time varying unobservables and the fixed effect change. As it can be concluded, the identified set for $\beta_2$ becomes wider for PGP with higher variance of the unobservables.
4.1.2 Binary Response Panel Data Model with Three Regressors

Consider the two time-period static binary response panel data model as in (42),

$$Y_t = 1(X_t\beta + \alpha + V_t > 0)$$

where $X_t = (X_{1t}, X_{2t}, X_{3t})$ with $X_{3t} \in \{0, 1\}$, $\beta = (\beta_1, \beta_2, \beta_3)$, $\alpha|X \sim N(\overline{X}\delta, 1)$ with $\overline{X} = \frac{1}{2}(X_1 + X_2)$ and $\delta = (1, -1, 0)$ and $V_t|X, \alpha \overset{iid}{\sim} N(0, 1)$. The true parameter of $\beta_2 = 1$ and $\beta_3 = 1$ after the normalization of $\beta_1 = 1$.

Figure 4 provides identified set for $(\beta_2, \beta_3)$ for different values of $\beta_3$ in the grid $[-1, 3]$. The sets shrink as the support of the discrete explanatory variables $(X_{1t}, X_{2t})$ increases and it can be seen the lower bound of the identified set for $\beta_2$ is achieved at the true value of $\beta_3$.

![Figure 4: Joint identified sets for $(\beta_2, \beta_3)$ when $X_{3t} \in \{0, 1\}$](image)

(a) $(X_{1t}, X_{2t}) \in \{-1, \ldots, 1\}$

(b) $(X_{1t}, X_{2t}) \in \{-2, \ldots, 2\}$

(c) $(X_{1t}, X_{2t}) \in \{-3, \ldots, 3\}$

(d) $(X_{1t}, X_{2t}) \in \{-1, \ldots, 3\}$

(e) $(X_{1t}, X_{2t}) \in \{-2, \ldots, 4\}$

(f) $(X_{1t}, X_{2t}) \in \{-4, \ldots, 2\}$

Figure 4: Joint identified sets for $(\beta_2, \beta_3)$ when $X_{3t} \in \{0, 1\}$
4.2 Dynamic Binary Response Panel Data Model

Consider the two time-period dynamic binary response panel data model as the one described in Section 2.2

\[ Y_t = 1(X_t\beta + 1(Y_{t-1} = 1)\gamma + \alpha + V_t > 0) \]

where \( X_t = (X_{1t}, X_{2t}) \), \( \beta = (\beta_1, \beta_2) \), \( \alpha|X \sim N(\overline{X}\delta, 1) \) with \( \overline{X} = \frac{1}{2}(X_1 + X_2) \) and \( \delta = (1, -1) \) and \( V_t|X, \alpha \sim N(0, 1) \). The true parameter of \( \beta_2 = 1 \) and \( \gamma = 0.5 \) after normalizing \( \beta_1 = 1 \).

Figure 5 provides identified sets for \((\beta_2, \gamma)\) as described in Theorem 5, for different values of \(\gamma\) in the grid \([-1, 6]\), as the support of \((X_{1t}, X_{2t})\) increases.

![Identified sets for \((\beta_2, \gamma)\) for different values of \(\gamma\).](image)

(a) \((X_{1t}, X_{2t}) \in \{-1, \ldots, 1\}\)
(b) \((X_{1t}, X_{2t}) \in \{-2, \ldots, 2\}\)
(c) \((X_{1t}, X_{2t}) \in \{-3, \ldots, 3\}\)
(d) \((X_{1t}, X_{2t}) \in \{-1, \ldots, 3\}\)
(e) \((X_{1t}, X_{2t}) \in \{-2, \ldots, 4\}\)
(f) \((X_{1t}, X_{2t}) \in \{-4, \ldots, 2\}\)

Figure 5: Joint identified sets for \((\beta_2, \gamma)\)

It is clear that for the specific range of values for the grid of \(\gamma\) chosen the identified set for \(\beta_2\) shrinks as the support of the discrete explanatory variables \((X_{1t}, X_{2t})\) increases, however with three periods the identified set for \(\gamma\) is only bounded by above. Recall that as shown in [Honoré and Kyriazidou (2000)](Honoré and Kyriazidou (2000)) the parameters in the dynamic binary response model with one lagged dependent variable and logistically distributed unobservables are point-identified with
four periods. Extending the model in Section 2.2 to include more than three time periods is currently in progress.

5 Concluding Remarks

This paper studies identification of discrete response panel data models with fixed effects. Under fairly weak conditions, identified sets in static and dynamic binary and ordered response models are derived, without assuming distributional assumptions on the time-varying unobservables and the fixed effect. The bounds are achieved by relying on observable implications in which the fixed effect does not appear.

Similarly to the static panel data model, the dynamic binary panel data model which specifies the option being chosen in the last period is complete and identification bounds for the parameters of interest are derived. When the definition of the dynamic binary panel data model changes to specify a switching behaviour, the model becomes both incomplete and incoherent and derivation of identification bounds is left for future research.

As discussed in Section 1 when the time-varying unobservables are independent and identically distributed with a logistic distribution, then the regression parameters in the linear index static and dynamic binary models can be point identified. The feature of the distribution these papers use that does not depend on the unobservable $\alpha$ is the conditional probability of the outcome variable in a specific period taking a specific value, conditional on the event that individuals change at some period in the past. The feature of the distribution that does not depend on the unobserved heterogeneity in this paper is the joint probability of the two outcome variables that take different values in two periods. If Assumptions 3 and 10 are strengthened such that the time-varying unobservables follow an independent and identically distributed logistic distribution, the bounds provided in this paper might still fail to be singletons. Chamberlain (1984, 2010) and Honoré and Kyriazidou (2000) prove point-identification of the regression parameters, when the regressors have bounded support, under the assumptions that the time-varying unobservables follow a logistic distribution and are independent of both the explanatory variables ($X_1, X_2$) and $\alpha$. Under Assumptions 5 and 12, $V$ is allowed
to be correlated with $\alpha$. Imposing the assumption that $V \perp \alpha$ may or may not lead to point-identification, but this would require an additional assumption which might not be credible or testable.

In conclusion, even though the identification bounds in this paper might not be singleton sets, they provide information on the regression parameters under fairly weak condition. Since the sets do not depend on any distributional assumption on the unobservables, they can provide information for a general class of linear index static and dynamic discrete response panel data models with fixed effects. Furthermore, they are relatively simple to construct and therefore might be easy to use for computation and inference. Extending the models to incorporate more time periods is currently in progress and in immediate next steps, the models studied in this paper will be applied to consumer demand for differentiated products.
A Proofs of Theorems

A.1 Proof of Theorem 4

This section provides the proof of Theorem 4. Following similar arguments as in Theorem 1, it can be shown that the event probabilities described in Theorem 4 do not depend on $\alpha$.

Consider the sequences of events in (26). Conditioning on $Y_0 = 0$ event $A$ implies the following conditional probability for the event $(Y_1, Y_2) = (0, 1)$,

\[
P(0, 1|x, 0) = P_{(V, \alpha)|X,Y_0} \{X_1 \beta + \alpha + V_1 \leq 0 \} \wedge \{X_2 \beta + \alpha + V_2 > 0\}|X = x, Y_0 = 0
\]

\[
\leq P_{V|X,Y_0}[(X_2 - X_1) \beta + V_2 - V_1 > 0|X = x, Y_0 = 0]
\]

\[
= P_{\Delta V|X,Y_0}[\Delta V > -\Delta X \beta |X = x, Y_0 = 0]
\]

Conditioning on $Y_0 = 0$ event $B$ implies the following conditional probability for the event $(Y_1, Y_2) = (1, 0)$,

\[
P(1, 0|x, 0) = P_{(V, \alpha)|X,Y_0} \{X_1 \beta + \gamma + \alpha + V_1 \leq 0 \} \wedge \{X_2 \beta + \alpha + V_2 < 0\}|X = x, Y_0 = 0
\]

\[
\leq P_{V|X,Y_0}[(X_2 - X_1) \beta + \gamma + V_2 - V_1 < 0|X = x, Y_0 = 0]
\]

\[
= P_{\Delta V|X,Y_0}[\Delta V < -\Delta X \beta - \gamma |X = x, Y_0 = 0]
\]

Conditioning on $Y_0 = 1$ event $C$ implies that the conditional probability of event $(Y_1, Y_2) = (0, 1)$ is equivalent to

\[
P(0, 1|x, 1) = P_{(V, \alpha)|X,Y_0} \{X_1 \beta + \gamma + \alpha + V_1 \leq 0 \} \wedge \{X_2 \beta + \alpha + V_2 > 0\}|X = x, Y_0 = 1
\]

\[
\leq P_{V|X,Y_0}[(X_2 - X_1) \beta - \gamma + V_2 - V_1 > 0|X = x, Y_0 = 1]
\]

\[
= P_{\Delta V|X,Y_0}[\Delta V > -\Delta X \beta + \gamma |X = x, Y_0 = 1]
\]

Conditioning on $Y_0 = 1$ event $D$ implies that the conditional probability of the event $(Y_1, Y_2) = (1, 0)$ is equivalent to

\[
P(1, 0|x, 1) = P_{(V, \alpha)|X,Y_0} \{X_1 \beta + \gamma + \alpha + V_1 \geq 0 \} \wedge \{X_2 \beta + \gamma + \alpha + V_2 < 0\}|X = x, Y_0 = 1
\]

\[
\leq P_{V|X,Y_0}[(X_2 - X_1) \beta + V_2 - V_1 < 0|X = x, Y_0 = 1]
\]

\[
= P_{\Delta V|X,Y_0}[\Delta V < -\Delta X \beta |X = x, Y_0 = 1]
\]
For any fixed \( X = x \) and by applying Assumption 12, then

\[
1 - P(0, 1|x, 0) \geq P_{\Delta V|X,Y_0}[\Delta V < -\Delta x \beta | Y_0 = 0] \\
P(1, 0|x, 0) \leq P_{\Delta V|X,Y_0}[\Delta V < -\Delta x \beta - \gamma | Y_0 = 0] \\
1 - P(0, 1|x, 1) \geq P_{\Delta V|X,Y_0}[\Delta V < -\Delta x \beta + \gamma | Y_0 = 1] \\
P(1, 0|x, 1) \leq P_{\Delta V|X,Y_0}[\Delta V < -\Delta x \beta | Y_0 = 1]
\]

which completes the proof.

A.2 Proof of Theorem 5

This section proves Theorem 5. Using arguments similar to the ones in Section 2.1 and in Chesher (2013) and RW2013, it can be shown that for any constant \( \omega \in \mathbb{R} \), conditioning on \( Y_0 = 0 \),

\[
(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega \Rightarrow \Delta V > \omega \\
(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega \Rightarrow \Delta V < \omega
\]

(43)

and conditioning on \( Y_0 = 1 \),

\[
(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega \Rightarrow \Delta V > \omega \\
(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega \Rightarrow \Delta V < \omega
\]

(44)

The relations in (43) and (44) in combination with Assumption 12 imply that, \( \forall \omega \in \mathbb{R} \):

\[
1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega | X = x, Y_0 = 0] \geq P[\Delta V < \omega | X = x, Y_0 = 0] \\
P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega | X = x, Y_0 = 0] \leq P[\Delta V < \omega | X = x, Y_0 = 0] \\
1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega | X = x, Y_0 = 1] \geq P[\Delta V < \omega | X = x, Y_0 = 1] \\
P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega | X = x, Y_0 = 1] \leq P[\Delta V < \omega | X = x, Y_0 = 1]
\]

\[\iff\]
\[
\inf_{x \in \mathcal{X}} 1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega | X = x, Y_0 = 0] \geq P[\Delta V < \omega | X = x, Y_0 = 0] \\
\sup_{x \in \mathcal{X}} P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta + \gamma \geq \omega | X = x, Y_0 = 0] \leq P[\Delta V < \omega | X = x, Y_0 = 0] \\
\inf_{x \in \mathcal{X}} 1 - P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega | X = x, Y_0 = 1] \geq P[\Delta V < \omega | X = x, Y_0 = 1] \\
\sup_{x \in \mathcal{X}} P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \leq \omega | X = x, Y_0 = 1] \leq P[\Delta V < \omega | X = x, Y_0 = 1]
\]

When \(-\Delta X \beta \geq \omega\)

\[P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega | X = x, Y_0 = 0] = P(0, 1|x, 0)\]

otherwise

\[P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \geq \omega | X = x, Y_0 = 0] = 0\]

When \(-\Delta X \beta - \gamma \leq \omega\)

\[P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega | X = x, Y_0 = 0] = P(1, 0|x, 0)\]

otherwise

\[P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta - \gamma \leq \omega | X = x, Y_0 = 0] = 0\]

When \(-\Delta X \beta + \gamma \geq \omega\)

\[P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega | X = x, Y_0 = 1] = P(0, 1|x, 1)\]

otherwise

\[P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega | X = x, Y_0 = 1] = 0\]

When \(-\Delta X \beta \leq \omega\)

\[P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega | X = x, Y_0 = 1] = P(1, 0|x, 1)\]

otherwise

\[P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega | X = x, Y_0 = 1] = 0\]

50
These equalities in combination with (43), (44) and by applying Assumption 12 imply that (45) is equivalent to:

\[
\inf_{x: -\Delta x \beta \geq \omega} 1 - P(0, 1|x, 0) \geq F_{\Delta V|Y_0}(\omega|0) \\
\sup_{x: -\Delta x \beta - \gamma \leq \omega} P(1, 0|x, 0) \leq F_{\Delta V|Y_0}(\omega|0) \\
\inf_{x: -\Delta x \beta + \gamma \geq \omega} 1 - P(0, 1|x, 1) \geq F_{\Delta V|Y_0}(\omega|1) \\
\sup_{x: -\Delta x \beta \leq \omega} P(1, 0|x, 1) \leq F_{\Delta V|Y_0}(\omega|1)
\]

\[
\iff 
\sup_{x: -\Delta x \beta - \gamma \leq \omega} P(1, 0|x, 0) \leq F_{\Delta V|Y_0}(\omega|0) \leq \inf_{x: -\Delta x \beta \geq \omega} 1 - P(0, 1|x, 0) \\
\sup_{x: -\Delta x \beta \leq \omega} P(1, 0|x, 1) \leq F_{\Delta V|Y_0}(\omega|1) \leq \inf_{x: -\Delta x \beta + \gamma \geq \omega} 1 - P(0, 1|x, 1)
\]

which completes the proof.

A.3 Proof of Theorem 7

To prove the unconditional identified set in Theorem 7, first notice that

\[
F_{\Delta V|X}(\omega|x) = F_{\Delta V|X,Y_0}(\omega|x,0)P(Y_0 = 0|X = x) + F_{\Delta V|X,Y_0}(\omega|x,1)P(Y_0 = 1|X = x).
\]

Define \( P(Y_0 = 0|X = x) = P_0(x) \) and \( P(Y_0 = 1|X = x) = P_1(x) \), which are fully observed. Following relations (43) and (44), \( \forall \omega \in \mathbb{R} \):

\[
1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega|X = x, Y_0 = 0] \geq P[\Delta V < \omega|X = x, Y_0 = 0] \\
P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega|X = x, Y_0 = 0] \leq P[\Delta V < \omega|X = x, Y_0 = 0] \\
1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega|X = x, Y_0 = 1] \geq P[\Delta V < \omega|X = x, Y_0 = 1] \\
P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega|X = x, Y_0 = 1] \leq P[\Delta V < \omega|X = x, Y_0 = 1].
\]
Define by

\[ G(\omega|x, 0) = P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega|X = x, Y_0 = 0] \]
\[ G(\omega|x, 1) = P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega|X = x, Y_0 = 1] \]
\[ \overline{G}(\omega|x, 0) = 1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega|X = x, Y_0 = 0] \]
\[ \overline{G}(\omega|x, 1) = 1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega|X = x, Y_0 = 1] . \]

Then multiplying by \( P_0(x) \) and \( P_1(x) \) such that,

\[ 1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta \geq \omega|X = x, Y_0 = 0]P_0(x) \geq P[\Delta V < \omega|X = x, Y_0 = 0]P_0(x) \]
\[ P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - \gamma \leq \omega|X = x, Y_0 = 0]P_0(x) \leq P[\Delta V < \omega|X = x, Y_0 = 0]P_0(x) \]
\[ 1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta + \gamma \geq \omega|X = x, Y_0 = 1]P_1(x) \geq P[\Delta V < \omega|X = x, Y_0 = 1]P_1(x) \]
\[ P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta \leq \omega|X = x, Y_0 = 1]P_1(x) \leq P[\Delta V < \omega|X = x, Y_0 = 1]P_1(x) , \]

implies that the relations in (47) can be expressed as

\[ G(\omega|x, 0)P_0(x) \leq F_{\Delta V|X,Y_0}(\omega|x, 0)P_0(x) \leq \overline{G}(\omega|x, 0)P_0(x) \]
\[ G(\omega|x, 1)P_1(x) \leq F_{\Delta V|X,Y_0}(\omega|x, 1)P_1(x) \leq \overline{G}(\omega|x, 1)P_1(x) \]

which implies,

\[ G(\omega|x, 0)P_0(x) + G(\omega|x, 1)P_1(x) \leq F_{\Delta V|X}(\omega|x) \leq \overline{G}(\omega|x, 0)P_0(x) + \overline{G}(\omega|x, 1)P_1(x) \]

\[ \iff \]

\[ G(\omega|x, 0)P_0(x) + G(\omega|x, 1)P_1(x) \leq F_{\Delta V}(\omega) \leq \overline{G}(\omega|x, 0)P_0(x) + \overline{G}(\omega|x, 1)P_1(x) \]

where the last result follows from Assumption 5 of \( V \perp X \). The last relation implies that,

\[ \sup_{x \in \mathcal{X}} \{ G(\omega|x, 0)P_0(x) + G(\omega|x, 1)P_1(x) \} \leq \inf_{x \in \mathcal{X}} \{ \overline{G}(\omega|x, 0)P_0(x) + \overline{G}(\omega|x, 1)P_1(x) \} \]

This completes the proof.
A.4 Proof of Theorem 8

The conditional probabilities \( P(Y_1 = y_1 \land Y_2 = y_2 | X = x, c) = P(y_1, y_2 | x, c) \) of the events given in (35) are then given by:

\[
P(0, 1 | x, c) = P_{V, \alpha}(X = (0 \geq X_1 \beta + \alpha + V_1 - c_{11}) \\
\wedge \{X_2 \beta + \alpha + V_2 - c_{12} > 0 \wedge 0 \geq X_2 \beta + \alpha + V_2 - c_{22}\}|X = x, c)
\]

\[
P(1, 0 | x, c) = P_{V, \alpha}(X = (X_1 \beta + \alpha + V_1 - c_{11} > 0 \wedge 0 \geq X_1 \beta + \alpha + V_1 - c_{21}) \\
\wedge \{0 \geq X_2 \beta + \alpha + V_2 - c_{12}\}|X = x, c)
\]

\[
P(0, 2 | x, c) = P_{V, \alpha}(X = (0 \geq X_1 \beta + \alpha + V_1 - c_{11}) \wedge \{X_2 \beta + \alpha + V_2 - c_{22} > 0\}|X = x, c)
\]

\[
P(2, 0 | x, c) = P_{V, \alpha}(X = (X_1 \beta + \alpha + V_1 - c_{21} > 0) \wedge \{0 \geq X_2 \beta + \alpha + V_2 - c_{12}\}|X = x, c)
\]

\[
P(1, 2 | x, c) = P_{V, \alpha}(X = (X_1 \beta + \alpha + V_1 - c_{11} > 0 \wedge 0 \geq X_1 \beta + \alpha + V_1 - c_{21}) \\
\wedge \{X_2 \beta + \alpha + V_2 - c_{12} > 0\}|X = x, c)
\]

\[
P(2, 1 | x, c) = P_{V, \alpha}(X = (X_1 \beta + \alpha + V_1 - c_{21} > 0) \wedge \{0 \geq X_2 \beta + \alpha + V_2 - c_{22}\}|X = x, c)
\]

\[
(48)
\]

From (48) it can be shown that:

\[
P(0, 1 | x, c) \leq P_{V, \alpha}(X = (0 \geq X_1 \beta + \alpha + V_1 - c_{11}) \wedge \{X_2 \beta + \alpha + V_2 - c_{12} > 0\}|X = x, c)
\]

\[
\leq P_{V, \alpha}(X = (0 > (X_1 - X_2) \beta + (V_1 - V_2) - c_{11} + c_{12})|X = x, c)
\]

\[
P(1, 0 | x, c) \leq P_{V, \alpha}(X = (X_1 \beta + \alpha + V_1 - c_{11} > 0) \wedge \{0 \geq X_2 \beta + \alpha + V_2 - c_{12}\}|X = x, c)
\]

\[
\leq P_{V, \alpha}(X = (0 > (X_2 - X_1) \beta + (V_2 - V_1) - c_{12} + c_{11})|X = x, c)
\]

\[
P(0, 2 | x, c) \leq P_{V, \alpha}(X = (0 > (X_1 - X_2) \beta + (V_1 - V_2) - c_{11} + c_{22})|X = x, c)
\]

\[
P(2, 0 | x, c) \leq P_{V, \alpha}(X = (0 > (X_2 - X_1) \beta + (V_2 - V_1) - c_{12} + c_{21})|X = x, c)
\]

\[
P(1, 2 | x, c) \leq P_{V, \alpha}(X = (0 \geq X_1 \beta + \alpha + V_1 - c_{21}) \wedge \{X_2 \beta + \alpha + V_2 - c_{22} > 0\}|X = x, c)
\]

\[
\leq P_{V, \alpha}(X = (0 > (X_1 - X_2) \beta + V_1 - V_2 + c_{22} - c_{21})|X = x, c)
\]

\[
P(2, 1 | x, c) \leq P_{V, \alpha}(X = (X_1 \beta + \alpha + V_1 - c_{21} > 0) \wedge \{0 \geq X_2 \beta + \alpha + V_2 - c_{22}\}|X = x, c)
\]

\[
\leq P_{V, \alpha}(X = (0 > (X_2 - X_1) \beta + V_2 - V_1 + c_{21} - c_{22})|X = x, c)
\]

\[
(49)
\]
Then using Assumptions 5 for any fixed \( c \in C \), the relations in (49) can be expressed as:

\[
P(0, 1|x, c) \leq P_{V|X,c}(V_2 - V_1 > -\Delta X\beta - c_{11} + c_{12}|X = x, c) = 1 - F_{\Delta V}[ -\Delta x\beta - c_{11} + c_{12}] \\
P(1, 0|x, c) \leq P_{V|X,c}(V_2 - V_1 < -\Delta X\beta + c_{12} - c_{11}|X = x, c) = F_{\Delta V}[ -\Delta x\beta - c_{11} + c_{12}] \\
P(0, 2|x, c) \leq P_{V|X,c}(-\Delta X\beta - c_{11} + c_{22} < V_2 - V_1|X = x, c) = 1 - F_{\Delta V}[ -\Delta x\beta - c_{11} + c_{22}] \\
P(2, 0|x, c) \leq P_{V|X,c}(V_2 - V_1 < -\Delta X\beta + c_{12} - c_{21}|X = x, c) = F_{\Delta V}[ -\Delta x\beta + c_{12} - c_{21}] \\
P(1, 2|x, c) \leq P_{V|X,c}(-\Delta X\beta + c_{22} - c_{21} < V_2 - V_1|X = x, c) = 1 - F_{\Delta V}[ -\Delta x\beta + c_{22} - c_{21}] \\
P(2, 1|x, c) \leq P_{V|X,c}(V_2 - V_1 < -\Delta X\beta + c_{22} - c_{21}|X = x, c) = F_{\Delta V}[ -\Delta x\beta + c_{22} - c_{21}] \\
\]

The inequalities in (50) lead to bounds for \( \beta \), such that for any given \( X = x \),

\[
P(1, 0|x, c) \leq F_{\Delta V}[ -\Delta x\beta - c_{11} + c_{12}] \leq 1 - P(0, 1|x, c) \\
P(2, 0|x, c) \leq F_{\Delta V}[ -\Delta x\beta + c_{12} - c_{21}] \\
F_{\Delta V}[ -\Delta x\beta - c_{11} + c_{22}] \leq 1 - P(0, 2|x, c) \\
P(2, 1|x, c) \leq F_{\Delta V}[ -\Delta x\beta + c_{22} - c_{21}] \leq 1 - P(1, 2|x, c) \\
\]

The above relations shows that changing choices from period \( t = 1 \) to \( t = 2 \) provide restrictions on the distributions of \( \Delta V \) that does not depend on the fixed effect, \( \alpha \). In the binary case discussed in Section 2 the events \((Y_1, Y_2) = (0, 0)\) and \((1, 1)\) gave no information on \( \beta \) since the behaviour for these “extreme” cases can be matched by extremely small or extremely large values of \( \alpha \). This is also true for the static ordered model for the events \((0, 0)\) and \((2, 2)\). However, in the ordered response model where \( Y_t = \{0, 1, 2\} \), considered in this section, there is
an “in-between” event, \((Y_1, Y_2) = (1, 1)\), that provides information on the \(\beta\) without involving the fixed effect. To see that consider the joint probability of choosing the event (36):

\[
P(1, 1|x, c) = P_{(V, \alpha)|X, c}[X_1 \beta + \alpha + V_1 - c_{11} > 0 \land 0 \geq X_1 \beta + \alpha + V_1 - c_{21} \\
\land \{X_2 \beta + \alpha + V_2 - c_{12} > 0 \land 0 \geq X_2 \beta + \alpha + V_2 - c_{22}\} | X = x, c]
\]

\[
P_{(V, \alpha)|X, c}[X_1 \beta + \alpha + V_1 - c_{11} > 0 \land 0 \geq X_1 \beta + \alpha + V_1 - c_{21} \\
\land X_2 \beta + \alpha + V_2 - c_{12} > 0 \land 0 \geq X_2 \beta + \alpha + V_2 - c_{22}|X = x, c]
\]

(52)

By combining

\[
X_1 \beta + \alpha + V_1 - c_{11} > 0 \text{ and } 0 \geq X_2 \beta + \alpha + V_2 - c_{22}
\]

and

\[
0 \geq X_1 \beta + \alpha + V_1 - c_{21} \text{ and } X_2 \beta + \alpha + V_2 - c_{12} > 0
\]

it can be shown that:

\[
P(1, 1|x, c) \leq P_V[0 > (X_2 - X_1) \beta + (V_2 - V_1) - c_{22} + c_{11} \\
\land (X_2 - X_1) \beta + (V_2 - V_1) - c_{12} + c_{21} > 0 | X = x, c]
\]

\[
P(1, 1|x, c) \leq P_{\Delta V}[-\Delta X \beta + c_{22} - c_{11} > \Delta V > -\Delta X \beta + c_{12} - c_{21}|X = x, c]
\]

(53)

which does not depend on \(\alpha\). This completes the proof.

### A.5 Proof of Theorem 9

The section proves the identified set of the static ordered model given in Theorem 9 using arguments as in Chesher (2013) and Rosen and Weidner (2013,WP). As discussed in Section 3.1 the events \{(0, 1), (1, 0), (0, 2), (2, 0), (1, 2), (2, 1), (1, 1)\} provide restrictions on the distribution of the \(\Delta V\) that do not depend on the unobserved heterogeneity, \(\alpha\). From the inequalities in (50) and (53) it can be concluded that,
For a given set of arbitrary constants $\omega, \omega', \omega'' \in \mathbb{R}$ the above relations imply that,

\[
\begin{align*}
(Y_1, Y_2) &= (0, 1) \quad \Rightarrow \quad \Delta V > -\Delta X \beta - c_{11} + c_{12} \\
(Y_1, Y_2) &= (1, 0) \quad \Rightarrow \quad \Delta V < -\Delta X \beta - c_{11} + c_{12} \\
(Y_1, Y_2) &= (0, 2) \quad \Rightarrow \quad \Delta V > -\Delta X \beta - c_{11} + c_{22} \\
(Y_1, Y_2) &= (2, 0) \quad \Rightarrow \quad \Delta V < -\Delta X \beta + c_{12} - c_{21} \\
(Y_1, Y_2) &= (1, 2) \quad \Rightarrow \quad \Delta V > -\Delta X \beta + c_{22} - c_{21} \\
(Y_1, Y_2) &= (2, 1) \quad \Rightarrow \quad \Delta V < -\Delta X \beta + c_{22} - c_{21} \\
(Y_1, Y_2) &= (1, 1) \quad \Rightarrow \quad -\Delta X \beta + c_{22} - c_{11} \geq \Delta V \geq -\Delta X \beta + c_{12} - c_{21}
\end{align*}
\]

For a given set of arbitrary constants $\omega, \omega', \omega'' \in \mathbb{R}$ the above relations imply that,

\[
\begin{align*}
(Y_1, Y_2) &= (0, 1) \quad \land \quad -\Delta X \beta - c_{11} + c_{12} \geq \omega \Rightarrow \Delta V > \omega \\
(Y_1, Y_2) &= (1, 0) \quad \land \quad -\Delta X \beta - c_{11} + c_{12} \leq \omega \Rightarrow \Delta V < \omega \\
(Y_1, Y_2) &= (0, 2) \quad \land \quad -\Delta X \beta - c_{11} + c_{22} \geq \omega \Rightarrow \Delta V > \omega \\
(Y_1, Y_2) &= (2, 0) \quad \land \quad -\Delta X \beta + c_{12} - c_{21} \leq \omega \Rightarrow \Delta V < \omega \\
(Y_1, Y_2) &= (1, 2) \quad \land \quad -\Delta X \beta + c_{22} - c_{21} \geq \omega \Rightarrow \Delta V > \omega \\
(Y_1, Y_2) &= (2, 1) \quad \land \quad -\Delta X \beta + c_{22} - c_{21} \leq \omega \Rightarrow \Delta V < \omega \\
\end{align*}
\]

\[
\begin{align*}
(Y_1, Y_2) &= (1, 1) \quad \land \quad \omega' \geq -\Delta X \beta + c_{22} - c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega \Rightarrow \omega' > \Delta V > \omega \\
(Y_1, Y_2) &= (1, 1) \quad \land \quad \omega \geq -\Delta X \beta + c_{22} - c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega'' \Rightarrow \omega > \Delta V > \omega''
\end{align*}
\]
The relations in (54) and Assumptions 5 imply that for each fixed $c \in C$, $\forall \omega, \omega', \omega'' \in \mathbb{R}$:

$$1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta - c_{11} + c_{12} \geq \omega | X = x, c] \geq P[\Delta V < \omega]$$
$$P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - c_{11} + c_{12} \leq \omega | X = x, c] \leq P[\Delta V < \omega]$$

$$1 - P[(Y_1, Y_2) = (0, 2) \land -\Delta X \beta - c_{11} + c_{22} \geq \omega | X = x, c] \geq P[\Delta V < \omega]$$
$$P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - c_{11} + c_{12} \leq \omega | X = x, c] \leq P[\Delta V < \omega]$$

Following the same arguments as in proving Theorem 3, it can be shown that for any given $\omega, \omega', \omega'' \in \mathbb{R}$,

When $-\Delta X \beta - c_{11} + c_{12} \geq \omega$

$$1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta - c_{11} + c_{12} \geq \omega | X = x, c] = 1 - P(0, 1|x, c)$$

otherwise

$$1 - P[(Y_1, Y_2) = (0, 1) \land -\Delta X \beta - c_{11} + c_{12} \geq \omega | X = x, c] = 1$$

When $-\Delta X \beta - c_{11} + c_{12} \leq \omega$

$$P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - c_{11} + c_{12} \leq \omega | X = x, c] = P(1, 0|x, c)$$

otherwise

$$P[(Y_1, Y_2) = (1, 0) \land -\Delta X \beta - c_{11} + c_{12} \leq \omega | X = x, c] = 0$$

When $-\Delta X \beta - c_{11} + c_{22} \geq \omega$

$$1 - P[(Y_1, Y_2) = (0, 2) \land -\Delta X \beta - c_{11} + c_{22} \geq \omega | X = x, c] = 1 - P(0, 2|x, c)$$

otherwise

$$1 - P[(Y_1, Y_2) = (0, 2) \land -\Delta X \beta - c_{11} + c_{22} \geq \omega | X = x, c] = 1$$
When $-\Delta X \beta + c_{12} - c_{21} \leq \omega$

$$P[(Y_1, Y_2) = (2, 0) \land -\Delta X \beta + c_{12} - c_{21} \leq \omega | X = x, c] = P(2, 0 | x, c)$$

otherwise

$$P[(Y_1, Y_2) = (2, 0) \land -\Delta X \beta + c_{12} - c_{21} \leq \omega | X = x, c] = 0$$

When $-\Delta X \beta + c_{22} - c_{21} \geq \omega$

$$1 - P[(Y_1, Y_2) = (1, 2) \land -\Delta X \beta + c_{22} - c_{21} \geq \omega | X = x, c] = 1 - P(1, 2 | x, c)$$

otherwise

$$1 - P[(Y_1, Y_2) = (1, 2) \land -\Delta X \beta + c_{22} - c_{21} \geq \omega | X = x, c] = 1$$

When $-\Delta X \beta + c_{22} - c_{21} \leq \omega$

$$P[(Y_1, Y_2) = (2, 1) \land -\Delta X \beta + c_{22} - c_{21} \leq \omega | X = x, c] = P(2, 1 | x, c)$$

otherwise

$$P[(Y_1, Y_2) = (2, 1) \land -\Delta X \beta + c_{22} - c_{21} \leq \omega | X = x, c] = 0$$

When $\omega' \geq -\Delta X \beta + c_{22} - c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega,$

$$P[(Y_1, Y_2) = (1, 1) \land \{\omega' > -\Delta X \beta + c_{22} - c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega\} | X = x, c] = P(1, 1 | x, c)$$

otherwise

$$P[(Y_1, Y_2) = (1, 1) \land \{\omega' > -\Delta X \beta + c_{22} + c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega\} | X = x, c] = 0$$

When $\omega \geq c_{22} + c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega''$

$$P[(Y_1, Y_2) = (1, 1) \land \{\omega \geq c_{22} + c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega''\} | X = x, c] = P(1, 1 | x, c)$$

otherwise

$$P[(Y_1, Y_2) = (1, 1) \land \{\omega \geq -\Delta X \beta + p_{22} - c_{12} + c_{11} \land -\Delta X \beta + c_{12} - c_{21} \geq \omega''\} | X = x, c] = 0.$$
The above relations show that depending on the different values of \( x \in \mathcal{X} \) the lower and upper bounds of the inequalities in (55) change. Since for any value of \( \beta \) in the identified set the relations in (55) should hold simultaneously for all \( \omega, \omega', \omega'' \in \mathbb{R} \), a.e. \( x \in \mathcal{X} \), combining (51), (53), (55) and the above relations for the different values of \( x \in \mathcal{X} \), as in the case of the binary static model the distribution of \( \Delta V \) is shown to be bounded by,

\[
\begin{align*}
\sup_{x : -\Delta x \beta - c_{11} + c_{12} \leq \omega} P(1, 0|x, c) &\leq F_{\Delta V}(\omega) \\
F_{\Delta V}(\omega) &\leq \inf_{x : -\Delta x \beta - c_{11} + c_{12} \geq \omega} [1 - P(0, 1|x, c)] \\
\sup_{x : -\Delta x \beta + c_{12} - c_{21} \leq \omega} P(2, 0|x, c) &\leq F_{\Delta V}(\omega) \\
F_{\Delta V}(\omega) &\leq \inf_{x : -\Delta x \beta + c_{12} - c_{21} \geq \omega} [1 - P(0, 2|x, c)] \\
\sup_{x : -\Delta x \beta + c_{22} - c_{21} \leq \omega} P(2, 1|x, c) &\leq F_{\Delta V}(\omega) \\
F_{\Delta V}(\omega) &\leq \inf_{x : -\Delta x \beta + c_{22} - c_{21} \geq \omega} [1 - P(1, 2|x, c)] \\
\sup_{x \in \mathcal{X}^*} P(1, 1|x, c) &\leq F_{\Delta V}(\omega') - F_{\Delta V}(\omega) \\
\sup_{x \in \mathcal{X}^{**}} P(1, 1|x, c) &\leq F_{\Delta V}(\omega) - F_{\Delta V}(\omega'')
\end{align*}
\]

where

\[
\begin{align*}
\mathcal{X}^* &= \{ x : \omega' \geq -\Delta x \beta + c_{22} - c_{11} \land -\Delta x \beta + c_{12} - c_{21} \geq \omega \} \\
\mathcal{X}^{**} &= \{ x : \omega \geq -\Delta x \beta + c_{22} - c_{11} \land -\Delta x \beta + c_{12} - c_{21} \geq \omega'' \}
\end{align*}
\]

Furthermore, notice that at the limit \( \omega' \to \infty \) and \( \omega'' \to -\infty \):

\[
\begin{align*}
P(1, 1|x, c) &\leq 1 - F_{\Delta V}(\omega) \Leftrightarrow F_{\Delta V}(\omega) \leq 1 - P(1, 1|x, c), \text{ when } x \in \mathcal{X}^* \\
P(1, 1|x, c) &\leq F_{\Delta V}(\omega), \text{ when } x \in \mathcal{X}^{**}
\end{align*}
\]

and

\[
\begin{align*}
F_{\Delta V}(\omega) &\leq \inf_{x \in \mathcal{X}^*} 1 - P(1, 1|x, c) \\
\sup_{x \in \mathcal{X}^{**}} P(1, 1|x, c) &\leq F_{\Delta V}(\omega)
\end{align*}
\]

which completes the proof.
A.6 Dynamic Ordered Model Identification

The regions in (34) can be extended to include the lagged dependent variable, resulting to a total of 27 regions. The joint probabilities of the events described in Section 3.2 in relations (41) conditional on a specific value of \( y_0 \in \{0, 1, 2\} \) are then given by:

\[
P(0, 1|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{11} \leq 0 \}
\land \{X_2\beta + \alpha + V_2 - c_{12} > 0 \land X_2\beta + \alpha + V_2 - c_{22} \leq 0\}|x, 0, c]
\]

\[
P(1, 0|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{11} > 0 \land X_1\beta + \alpha + V_1 - c_{21} \leq 0\}
\land \{X_2\beta + \alpha + \gamma_1 + V_2 - c_{12} \leq 0\}|x, 0, c]
\]

\[
P(0, 2|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{11} \leq 0 \}
\land \{X_2\beta + \alpha + V_2 - c_{22} > 0\}|x, 0, c]
\]

\[
P(2, 0|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{21} > 0\}
\land \{X_2\beta + \gamma_2 + \alpha + V_2 - c_{12} \leq 0\}|x, 0, c]
\]

\[
P(1, 2|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{11} > 0 \land X_1\beta + \alpha + V_1 - c_{21} \leq 0\}
\land \{X_2\beta + \alpha + V_2 + \gamma_2 - c_{12} > 0 \land X_2\beta + \alpha + V_2 + \gamma_2 - c_{22} \leq 0\}|x, 0, c]
\]

\[
P(2, 1|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{21} > 0\}
\land \{X_2\beta + \alpha + V_2 + \gamma_1 - c_{12} > 0 \land X_2\beta + \alpha + V_2 + \gamma_1 - c_{22} \leq 0\}|x, 0, c]
\]

\[
P(1, 1|x, 0, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \alpha + V_1 - c_{11} > 0 \land X_1\beta + \alpha + V_1 - c_{21} \leq 0\}
\land \{X_2\beta + \alpha + V_2 + \gamma_1 - c_{12} > 0 \land X_2\beta + \alpha + V_2 + \gamma_1 - c_{22} \leq 0\}|x, 0, c]
\]

\[
P(0, 1|x, 1, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \gamma_1 + \alpha + V_1 - c_{11} \leq 0\}
\land \{X_2\beta + \alpha + V_2 - c_{12} > 0 \land X_2\beta + \alpha + V_2 - c_{22} \leq 0\}|x, 1, c]
\]

\[
P(1, 0|x, 1, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \gamma_1 + \alpha + V_1 - c_{11} > 0 \land X_1\beta + \gamma_1 + \alpha + V_1 - c_{21} \leq 0\}
\land \{X_2\beta + \gamma_1 + \alpha + V_2 - c_{12} \leq 0\}|x, 1, c]
\]

\[
P(0, 2|x, 1, c) = P_{(V, a)}|X, Y_0, c[\{X_1\beta + \gamma_1 + \alpha + V_1 - c_{11} \leq 0\}
\land \{X_2\beta + \alpha + V_2 - c_{22} > 0\}|x, 1, c]
\]

60
\[ P(2,0|x,1,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_1 + \alpha + V_1 - c_{21} > 0] \]
\[ \land \{X_2 \beta + \gamma_2 + \alpha + V_2 - c_{12} \leq 0\}|x,1,c] \]
\[ P(1,2|x,1,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_1 + \alpha + V_1 - c_{11} > 0 \land X_1 \beta + \gamma_1 + \alpha + V_1 - c_{21} \leq 0] \]
\[ \land \{X_2 \beta + \gamma_1 + \alpha + V_2 - c_{22} > 0\}|x,1,c] \]
\[ P(2,1|x,1,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_1 + \alpha + V_1 - c_{21} > 0] \]
\[ \land \{X_2 \beta + \gamma_2 + \alpha + V_2 - c_{12} > 0 \land X_2 \beta + \gamma_2 + \alpha + V_2 - c_{22} \leq 0\}|x,1,c] \]
\[ P(1,1|x,1,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_1 + \alpha + V_1 - c_{11} > 0 \land X_1 \beta + \gamma_1 + \alpha + V_1 - c_{21} \leq 0] \]
\[ \land \{X_2 \beta + \gamma_1 + \alpha + V_2 - c_{12} > 0 \land X_2 \beta + \gamma_1 + \alpha + V_2 - c_{22} \leq 0\}|x,1,c] \]
\[ P(0,1|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + \alpha + V_1 - c_{11} \leq 0] \]
\[ \land \{X_2 \beta + \alpha + V_2 - c_{12} > 0 \land X_2 \beta + \alpha + V_2 - c_{22} \leq 0\}|x,2,c] \]
\[ P(1,0|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + \alpha + V_1 - c_{11} > 0 \land X_1 \beta + \gamma_2 + \alpha + V_1 - c_{21} \leq 0] \]
\[ \land \{X_2 \beta + \alpha + \gamma_1 + V_2 - c_{12} \leq 0\}|x,2,c] \]
\[ P(0,2|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + \alpha + V_1 - c_{11} \leq 0] \]
\[ \land \{X_2 \beta + \gamma_2 + \alpha - c_{22} > 0\}|x,2,c] \]
\[ P(2,0|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + \alpha + V_1 - c_{21} > 0] \]
\[ \land \{X_2 \beta + \gamma_2 + \alpha + V_2 - c_{12} \leq 0\}|x,2,c] \]
\[ P(1,2|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + \alpha + V_1 - c_{11} > 0 \land X_1 \beta + \gamma_2 + \alpha + V_1 - c_{21} \leq 0] \]
\[ \land \{X_2 \beta + \gamma_1 + \alpha + V_2 - c_{22} > 0\}|x,2,c] \]
\[ P(2,1|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + \alpha + V_1 - c_{21} > 0] \]
\[ \land \{X_2 \beta + \gamma_2 + \alpha + V_2 - c_{12} > 0 \land X_2 \beta + \gamma_2 + \alpha + V_2 - c_{22} \leq 0\}|x,2,c] \]
\[ P(1,1|x,2,c) = P_{(V,o)|X,Y,c}[[X_1 \beta + \gamma_2 + V_1 + \alpha - c_{11} > 0 \land X_1 \beta + \gamma_2 + V_1 + \alpha - c_{21} \leq 0] \]
\[ \land \{X_2 \beta + \gamma_1 + V_2 + \alpha - c_{22} > 0 \land X_2 \beta + \gamma_1 + V_2 + \alpha - c_{22} \leq 0\}|x,2,c] \]
Then following the same arguments as in Section 3.1 it can be shown that the joint probabilities described in Section 3.2 in relations (41) conditional on a specific value of \( y_0 \in \{0, 1, 2\} \) are bounded by the following relations,

\[
\begin{align*}
P(0, 1|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_1 - V_2) - c_{11} + c_{12}|x, 0, c] \\
P(1, 0|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + (V_1 - V_1) - c_{12} + c_{11}|x, 0, c] \\
P(0, 2|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_1 - V_2) - c_{11} + c_{22}|x, 0, c] \\
P(2, 0|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + (V_2 - V_1) - c_{12} + c_{21}|x, 0, c] \\
P(1, 2|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_2 - V_2) - c_{12} + c_{21}|x, 0, c] \\
P(2, 1|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + V_1 - V_1 + c_{21}|x, 1, c] \\
P(1, 1|x, 0, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + (V_1 - V_1) - c_{12} + c_{21}|x, 0, c] \\
P(0, 1|x, 1, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_1 - V_2) - c_{11} + c_{12}|x, 1, c] \\
P(1, 0|x, 1, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + (V_2 - V_1) - c_{12} + c_{11}|x, 1, c] \\
P(0, 2|x, 1, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_2 - V_1) - c_{11} + c_{22}|x, 1, c] \\
P(2, 0|x, 1, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + (V_2 - V_1) - c_{12} + c_{21}|X = x, 1, c] \\
P(1, 2|x, 1, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + V_1 - V_2 + c_{22} - c_{21}|x, 1, c] \\
P(2, 1|x, 1, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + V_1 - V_2 + c_{22} - c_{21}|x, 1, c] \\
P(1, 1|x, 1, c) & \leq P_{V|X,Y_0,c}[-\Delta X \beta - \gamma_1 + c_{22} - c_{11} > \Delta V > -\Delta X \beta - \gamma_1 + c_{12} - c_{21}|x, 0, c] \\
P(0, 1|x, 2, c) & \leq P_{V|X,Y_0,c}[-\Delta X \beta + c_{22} - c_{11} > \Delta V > -\Delta X \beta + c_{12} - c_{21}|x, 1, c] \\
P(1, 0|x, 2, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_1 - V_2) - c_{11} + c_{12}|x, 1, c] \\
P(0, 2|x, 2, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + (V_1 - V_2) - c_{11} + c_{22}|x, 1, c] \\
P(2, 0|x, 2, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + (V_2 - V_1) - c_{12} + c_{21}|X = x, 1, c] \\
P(1, 2|x, 2, c) & \leq P_{V|X,Y_0,c}[0 > (X_1 - X_2)\beta + c_{12} - c_{21}|x, 1, c] \\
P(2, 1|x, 2, c) & \leq P_{V|X,Y_0,c}[0 > (X_2 - X_1)\beta + V_2 - V_1 + c_{21} - c_{22}|x, 1, c] \\
P(1, 1|x, 2, c) & \leq P_{V|X,Y_0,c}[\begin{array}{c}
-\Delta X \beta - \gamma_1 + c_{22} - c_{11} \\
\quad > \Delta V > \\
-\Delta X \beta - \gamma_1 + c_{12} - c_{21}
\end{array}|x, 1, c]
\end{align*}
\]
\[ \begin{align*}
1 - P(0, 1|x, 0, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - c_{11} + c_{12}|X = x, Y_0 = 0, c] \\
P(1, 0|x, 0, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_1 + c_{12} - c_{11}|X = x, Y_0 = 0, c] \\
1 - P(0, 2|x, 0, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - c_{11} + c_{22}|X = x, Y_0 = 0, c] \\
P(2, 0|x, 0, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + c_{12} - c_{21}|X = x, Y_0 = 0, c] \\
1 - P(1, 2|x, 0, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_1 + c_{22} - c_{21}|X = x, Y_0 = 0, c] \\
P(2, 1|x, 0, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + c_{22} - c_{21}|X = x, Y_0 = 0, c] \\
1 - P(1, 1|x, 0, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_1 + c_{22} - c_{11} > \Delta V > \Delta V > \Delta V > X = x, Y_0 = 0, c] \\
P(1, 1|x, 0, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_1 + c_{12} - c_{21} > \Delta V > \Delta V > \Delta V > X = x, Y_0 = 0, c] \\
1 - P(0, 0|x, 1, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_1 - c_{11} + c_{12}|X = x, Y_0 = 1, c] \\
P(1, 0|x, 1, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + c_{12} - c_{11}|X = x, Y_0 = 1, c] \\
1 - P(0, 1|x, 1, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_1 - c_{11} + c_{22}|X = x, Y_0 = 1, c] \\
P(2, 0|x, 1, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_1 - \gamma_2 + c_{12} - c_{21}|X = x, Y_0 = 1, c] \\
1 - P(1, 0|x, 1, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + c_{22} - c_{21}|X = x, Y_0 = 1, c] \\
P(2, 1|x, 1, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + \gamma_1 + c_{22} - c_{21}|X = x, Y_0 = 1, c] \\
1 - P(0, 2|x, 1, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + \gamma_1 + c_{22} - c_{21}|X = x, Y_0 = 1, c] \\
P(1, 2|x, 1, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + \gamma_1 + c_{12} - c_{11}|X = x, Y_0 = 1, c] \\
1 - P(0, 0|x, 2, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_2 - c_{11} + c_{12}|X = x, Y_0 = 2, c] \\
P(1, 0|x, 2, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + \gamma_1 + c_{22} - c_{11}|X = x, Y_0 = 2, c] \\
1 - P(0, 1|x, 2, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_2 + \gamma_1 + c_{12} - c_{21}|X = x, Y_0 = 2, c] \\
P(2, 0|x, 2, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_2 - c_{21}|X = x, Y_0 = 2, c] \\
1 - P(1, 2|x, 2, c) & \geq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + \gamma_2 - \gamma_1 + c_{22} - c_{21}|X = x, Y_0 = 2, c] \\
P(2, 1|x, 2, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta + c_{22} - c_{21}|X = x, Y_0 = 2, c] \\
P(1, 2|x, 2, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + \gamma_1 + c_{22} - c_{21}|X = x, Y_0 = 2, c] \\
P(1, 1|x, 2, c) & \leq P_{\Delta V|X,Y_0, c}[\Delta V < -\Delta X\beta - \gamma_2 + \gamma_1 + c_{12} - c_{21}|X = x, Y_0 = 2, c]
\end{align*}\]
References


