# Sufficient Statistics for Unobserved Heterogeneity in Structural Dynamic Logit Models

Victor Aguirregabiria, Jiaying Gu, & Yao Luo

University of Toronto

**MICHIGAN STATE - ECONOMETRICS SEMINAR**

April 12, 2018

# Context & Motivation; UH in DPD

- Key empirical question in Dynamic Panel Data (DPD) models is distinguishing between **"true dynamics"** and **"spurious dynamics"** due to persistent **unobserved heterogeneity (UH).**

- The answer to this question should deal with two important econometric issues [Heckman, 1981]

  - *The* **incidental parameters problem** *[simple dummy-variables estimation is inconsistent, with fixed T]*
  - *The* **initial conditions problem** *[the joint distribution of UH and the initial values of endogeneous/predetermined explanatory variables is NOT nonparametrically identified].*

# Context & Motivation: CRE vs FE

- The standard methods to deal with UH in DPD models are **Fixed Effects (FE)** and **Correlated Random Effects (CRE)**.
- **CRE models** deal with the initial conditions problem by imposing restrictions (e.g., parametric, finite support) on the joint distribution of UH and explanatory variables.
- **FE approach**. The join distribution of the UH and the whole history of explanatory variables is NP specified.
- FE is more robust than CRE. [Several limitations: not all DPD can be estimated by FEs; Identification of Average Marginal Effects.]

# Context & Motivation: FE-Conditional MLE

- Within the FE methods, we can distinguish:

  *- Dummy-variables [incosistent in discrete choice];*
  *- Transformation-based / Maximum Score [inconsistent in Dynamic DC]*
  - **Sufficient statistics-Conditional ML (FE-CMLE)** *[Can identify parameters of interest only in some PD models]*

- In Dynamic DC models, **FE-CMLE** cannot point-identify parameters of interest if:

  *- Distribution of transitory unobs is not logit (EV) [Honore & Tamer (2006) extension to partial identification.]*
  *- UH and predetermined explan. vars are not additively separable.*

# Context & Motivation: Structural (Forward-Looking) DDC

- In **structural [forward-looking] dynamic logit models**, the common wisdom is that FE-CML cannot identify structural parameters.
- An agent's optimal decision depends on current utility but also on the **continuation value** of the current choice.
- Even if UH enters additively in one-period utility function, **UH appears non-additively in continuation value function**.
- All applications of structural DDC models: CRE approach. Typicaly finite mixture with strong restrictions on joint distribution with initial conditions.

# The contribution of this paper

- We revisit the **applicabiity of FE-CML** to a class of strutural dynamic logit models that includes many of the existing applications.
- This model includes two types of endogenous state variables: lagged decision; and duration in the last choice.
- We derive the **minimal sufficient statistic for the UH,** and show the identification of some structural parameters (switching costs, and returns of experience/duration).
- Based on this identification result, we propose a specification test (Hausman test) for the validity of a CRE model.
- We apply our results to a Machine Replacement model.

# Outline

1. Model
2. Identification
   (a) Models WITHOUT duration dependence
   (b) Models WITH duration dependence
3. Estimation and Inference
4. Monte Carlo Experiments
5. Empirical Application

# 1. Model

# FE Structural Dynamic Logit

- Decision variable: $y_{it} \in \mathcal{Y} = \{0, 1, ..., J\}$. Agent maximizes $\mathbb{E}_t \left[\sum_{s=0}^{\infty} \delta_i^s \; U_{i,t+s}\right]$. $U_{it}$ is the utility function:

$$U_{it}(y) = \alpha\left(y, \boldsymbol{\eta}_i, \mathbf{z}_{it}\right) + \beta\left(y, \mathbf{x}_{it}, \mathbf{z}_{it}\right) + \varepsilon_{it}(y).$$

  $\mathbf{z}_{it}$ = exogenous state vars. $\mathbf{x}_{it}$ = endogenous state vars. **Observable**.

- Three **unobservables**: $\varepsilon_{it}(y)$ i.i.d. type I extreme value distributed; $\boldsymbol{\eta}_i$ in the payoff; and the discount factor $\delta_i$.

- $p(\boldsymbol{\eta}_i, \delta_i \mid \mathbf{x}_{it}, \mathbf{z}_{it} : t = 1, 2, ...\}$ is unrestricted, i.e., FE model.

- **Key condition:** $\boldsymbol{\eta}_i$ and $\mathbf{x}_{it}$ additively separable in payoff function.

# Endogenous State Variables

- Two types of endogenous state variables: $\mathbf{x}_{it} = (y_{i,t-1}, d_{it})$.

  (a) **lagged decision variable**, $y_{i,t-1} \in \mathcal{Y}$.

  (b) **duration** *in the last choice*, $d_{it}$

- The component of utility function that captures state dependence:

$$\beta\left(y, \mathbf{x}_{it}\right) = 1\{y = y_{i,t-1}\}\ \beta_d\left(y, d_{it}\right) + 1\{y \neq y_{i,t-1}\}\ \beta_y\left(y, y_{i,t-1}\right)$$

- $\beta_y\left(y, y_{i,t-1}\right)$ represents switching costs, e.g., the cost of switching from occupation $y_{i,t-1}$ to occupation $y$.

- $\beta_d\left(y, d_{it}\right)$ captures duration dependence, e.g., the effect of experience in occupation.

# Restrictions on beta_d function

- For the model with duration dependence (i.e., $\beta_d(y, d) \neq 0$), some of our identification results exploit the following assumptions. identificatWe impose the following restrictions on function $\beta_d$.

  **(A.1)** No duration dependence in the "outside alternative" $y = 0$:

$$\beta_d(y, d) = 0$$

- For the forward-looking model

  **(A.2)** For every $y > 0$, there is finite duration $d_y^*$ such that:

$$\beta_d(y, d) = \beta_d\left(y, d_y^*\right) \quad \text{for any } d \geq d_y^*$$

# Examples

- (1) *Market entry-exit models.*
- (2) *Machine replacement models.*
- (3) *Occupational choice models.*
- (4) *Dynamic demand of differentiated storable products.*

# Optimal decision rule

- Let $\boldsymbol{\theta}_i \equiv (\boldsymbol{\eta}_i, \delta_i)$ represent the UH. The model implies:

$$y_{it} = \arg\max_{y \in \mathcal{Y}} \left\{ \begin{array}{c} \alpha\left(y, \boldsymbol{\eta}_i, \mathbf{z}_{it}\right) + \beta\left(y, \mathbf{x}_{it}, \mathbf{z}_{it}\right) + \varepsilon_{it}(y) \\ + \delta_i \; \mathbb{E}\left[ V_{\boldsymbol{\theta}_i}\left(y, d_{i,t+1}(y), \mathbf{z}_{i,t+1}\right) \mid y, \mathbf{x}_{it}, \mathbf{z}_{it} \right] \end{array} \right\}$$

where $\delta_i \; \mathbb{E}\left[ V_{\boldsymbol{\theta}_i}\left(y, d_{i,t+1}(y), \mathbf{z}_{i,t+1}\right) \mid y, \mathbf{x}_{it}, \mathbf{z}_{it} \right]$ is the continuation value of choosing alternative $y$.

- Given the transition rule of duration:

$$v_{\boldsymbol{\theta}_i}\left(y, d_{it+1}(y), \mathbf{z}_{it}\right) = \left\{ \begin{array}{ll} v_{\boldsymbol{\theta}_i}\left(0, 0, \mathbf{z}_{it}\right) & \text{if} \quad y = 0 \\ v_{\boldsymbol{\theta}_i}\left(y, 1\{y = y_{it-1}\} d_{it} + 1, \mathbf{z}_{it}\right) & \text{if} \quad y > 0 \end{array} \right.$$

# Two important properties of continuation values

- **Property 1.** In the model **WITHOUT duration dependence**:

$$v_{\boldsymbol{\theta}_i}(y, \mathbf{z}_{it}) \text{ does not depend on } y_{it-1}$$

- **Property 2.** In the model **WITH duration dependence** under Assumption A.2:

$$v_{\boldsymbol{\theta}_i}(y, d, \mathbf{z}_{it}) = v_{\boldsymbol{\theta}_i}(y, d_y^*, \mathbf{z}_{it}) \text{ for any } d \geq d^*$$

This implies that for $d_{it} \geq d_y^* - 1$, the continuation value $v_{\boldsymbol{\theta}_i}(y, d_{it+1}(y), \mathbf{z}_{it})$ does not depend on the state variables $(y_{i,t-1}, d_{it})$.

# Conditional Choice Probabilities

- Under the extreme value type 1 distribution of the unobservables $\varepsilon$, implies that the *conditional choice probability* (CCP) function has the following form:

$$P\left(y|\mathbf{x}_t, \mathbf{z}_t, \theta\right) = \frac{\exp\left\{\ \alpha_\theta\left(y, \mathbf{z}_t\right) + \beta\left(y, \mathbf{x}_t\right) + v_\theta(y, d_{t+1}(y), \mathbf{z}_t)\ \right\}}{\sum\limits_{j\in\mathcal{Y}} \exp\left\{\ \alpha_\theta\left(j, \mathbf{z}_t\right) + \beta\left(j, \mathbf{x}_t\right) + v_\theta(j, d_{t+1}(j), \mathbf{z}_t)\ \right\}}$$

- Note change in notation: omit $i$; use subindex $\theta$ in those functions that depend on the incidental parameters $\theta$.

# 2. Identification using FE-CML

# Data

- The researcher observes panel data of individuals over several periods of time:

$$\text{Data} = \{\ y_{it},\ \mathbf{x}_{it}\ ,\ \mathbf{z}_{it} : i = 1, 2, ..., N\ ;\ \ t = 1, 2, ..., T\}$$

  $N$ is large and $T$ is small.

- Given these data and the restrictions from the model, the researcher is interested in the estimation of the structural parameters that capture "true dynamics" or "true state dependence", i.e., $\beta_d\left(y, d_{it}\right)$ and $\beta_y\left(y, y_{i,t-1}\right)$.

- We denote these structural parameters using the vector $\boldsymbol{\beta}$.

# Identification: General ideas

- Let $\mathbf{y}^T = \{y_1, y_2, ..., y_T\}$ be an individual's observed history

$$\mathbb{P}\left(\mathbf{y}^T \mid \mathbf{x}_1, \boldsymbol{\theta}, \boldsymbol{\beta}\right) = \prod_{t=1}^{T} \frac{\exp\{\ \alpha_{\boldsymbol{\theta}}(y_t) + \beta(y_t, \mathbf{x}_t) + v_{\boldsymbol{\theta}}(y_t, d_{t+1}(y_t))\ \}}{\sum\limits_{j \in \mathcal{Y}} \exp\{\ \alpha_{\boldsymbol{\theta}}(j) + \beta(j, \mathbf{x}_t) + v_{\boldsymbol{\theta}}(j, d_{t+1}(j))\ \}}$$

- All our identification results follow the same type of argument. We show that

$$\ln \mathbb{P}\left(\mathbf{y}^T \mid \mathbf{x}_1, \boldsymbol{\theta}, \boldsymbol{\beta}\right) = U(\mathbf{y}^T, \mathbf{x}_1)' g_{\boldsymbol{\theta}} + S(\mathbf{y}^T, \mathbf{x}_1)' \boldsymbol{\beta}^*$$

  $U$ and $S$ are vector of statistics; $g_{\boldsymbol{\theta}}$ is a vector of functions of $\boldsymbol{\theta}$; $\boldsymbol{\beta}^*$ is a vector of linear combinations of structural parameters $\boldsymbol{\beta}$.

- Each vector, $U$, $g_{\boldsymbol{\theta}}$, $S$, and $\boldsymbol{\beta}^*$, has elements which are linearly independent.

# Identification: General ideas [2]

- Given the representation $\ln \mathbb{P}\left(\mathbf{y}^T \mid \mathbf{x}_1, \boldsymbol{\theta}, \boldsymbol{\beta}\right) = U' g_{\boldsymbol{\theta}} + S' \boldsymbol{\beta}^*$, and the lineal independence of the elements in each vector $U$, $g_{\boldsymbol{\theta}}$, $S$, and $\boldsymbol{\beta}^*$, we establish the following results.
- *(i) Sufficiency.* $U$ is a sufficient statistic for $\boldsymbol{\theta}$.
- *(ii) Minimal sufficiency.* $U$ is a *minimal sufficient statistic* for $\boldsymbol{\theta}$.
- *(iii) Identification.* We show that condition on $U$, the elements in the vector of $S$ are linearly independent, and that this implies identification of $\boldsymbol{\beta}^*$.

# Model without duration dependence

- We have that:

$$\ln \mathbb{P}\left(\mathbf{y}^{T} | y_0, \boldsymbol{\theta}\right) = \sum_{y=1}^{J} T^{(y)}\; g_{\boldsymbol{\theta},1}(y) + \sum_{y=1}^{J} \Delta^{(y)}\; g_{\boldsymbol{\theta},2}(y)$$

$$+ \sum_{y_{-1}=1}^{J} \sum_{y=1}^{J} D^{(y_{-1},y)}\; \widetilde{\beta}_y(y, y_{-1})$$

$T^{(y)} = \#$ of times alternative $y$ is observed in history $\mathbf{y}^{T}$;
$\Delta^{(y)} = 1\{y_T = y\} - 1\{y_0 = y\}$
$D^{(y_{-1},y)} = \#$ of times the sequence $(y_{-1}, y)$ is observed in history $\mathbf{y}^{T}$
$\widetilde{\beta}_y(y, y_{-1}) = \beta_y(y, y_{-1}) - \beta_y(0, y_{-1}) - \beta_y(y, 0)$: Difference between the cost of direct switching and the cost of switching via $y = 0$.

# Model without duration dependence [2]

- (i) $U = \{T^{(y)} : y \geq 1, \Delta^{(y)} : y \geq 1\}$ is a sufficient statistic for $\theta$.
- (ii) The elements in the vector $U$ are linearly independent such that $U$ is a minimal sufficient statistic.
- (iii) Conditional on $U$, the vector of statistics $\{D^{(y_{-1},y)} : y_{-1}, y \in Y - \{0\}\}$ are linearly independent such that they can identify the vector of parameters $\{\widetilde{\beta}_y(y, y_{-1}) : y_{-1}, y \in Y - \{0\}\}$. ■

- *EXAMPLE 1.* Suppose that $T = 3$. Two realizations of the history $(y_0|\mathbf{y}^T)$:

$$A = \{0 \mid 0, j, k\} \quad \text{and} \quad B = \{0 \mid j, 0, k\} \quad \text{with} \quad j, k \neq 0$$

$U(A) = U(B)$ and $\ln \mathbb{P}(A|U) - \ln \mathbb{P}(B|U) = \widetilde{\beta}_y(k, j)$.

# Model WITH duration dependence: Myopic

$$\begin{aligned}\ln \mathbb{P}\left(\mathbf{y}^T | y_0, d_1\right) &= \sum_{y=1}^{J}\sum_{d\geq 1} H^{(y)}(d)\ g_{\boldsymbol{\theta},1}(y,d) + \sum_{y=1}^{J} \Delta^{(y)}\ g_{\boldsymbol{\theta},2}(y) \\ &+ \sum_{y_{-1}=1}^{J}\ \sum_{y=1, y\neq y_{-1}}^{J} D^{(y_{-1},y)}\ \widetilde{\beta}_y(y, y_{-1}) \\ &+ \sum_{y=1}^{J}\sum_{d\geq 1} \Delta^{(y)}(d)\ \gamma(y, d-1)\end{aligned}$$

$H^{(y)}(d) =$ Histogram of duration.
$\Delta^{(y)}(d) = 1\{y_T = y, d_{T+1} = d\} - 1\{y_0 = y, d_1 = d\}$
$\widetilde{\beta}_y(y, y_{-1}) \equiv \beta_y(y, y_{-1}) - \beta_y(0, y_{-1}) - \beta_y(y, 0)$
$\gamma(y, d) \equiv \beta_d(y, d) - \beta_y(y, 0) - \beta_y(0, y)$.

# Model WITH duration dependence: Myopic

*(i)* $U = \{H^{(y)}(d) : y \geq 1,\ d \geq 1,\ \Delta^{(y)} : y \geq 1\}$ *is a sufficient statistic of* $\boldsymbol{\theta}$.

*(ii) The elements in the vector $U$ are linearly independent such that $U$ is a minimal sufficient statistic.*

*(iii) Conditional on $U$, the vector of statistics* $\{D^{(y_{-1},y)} : y_{-1},\ y \geq 1;\ \Delta^{(y)}(d) : y \geq 1,\ d \geq 1\}$ *are linearly independent and they identify the vectors of structural parameters* $\{\widetilde{\beta}_y(y, y_{-1}) : y_{-1},\ y \geq 1,\ y \neq y_{-1};\ \gamma(y, d) : y \geq 1,\ d \geq 1\}$.

# Model WITH duration dependence: Myopic

*EXAMPLE 2.* Suppose that $T = 3$ and consider two realizations of the history $\left(y_0, d_1 | \mathbf{y}^T\right)$: for $j \neq k$,

$$A = \{0, 0 \mid 0, j, k\} \text{ and } B = \{0, 0 \mid j, 0, k\}$$

$U(A) = U(B)$ and $\ln \mathbb{P}\left(A|U\right) - \ln \mathbb{P}\left(B|U\right) = \widetilde{\beta}_y(k, j)$.

*EXAMPLE 3.* $T = n + 2$:

$$A = \{0, 0 \mid 0, \mathbf{y}_{n+1}\} \quad \text{and} \quad B = \{0, 0 \mid \mathbf{y}_n, 0, y\}$$

$U(A) = U(B)$ and $\ln \mathbb{P}\left(A|U\right) - \ln \mathbb{P}\left(B|U\right) = \gamma(y, n)$.

# Model WITH duration dependence: Forward

$\ln \mathbb{P}\left(\mathbf{y}^T | y_0, d_1\right)$ is

$$\sum_{y=1}^{J} \sum_{d \leq d_y^* - 1} H^{(y)}(d) \; g_{\boldsymbol{\theta},1}(y,d) + \left[\sum_{y=1}^{J} \sum_{d \geq d_y^*} H^{(y)}(d)\right] g_{\boldsymbol{\theta},1}(y, d_y^*)$$

$$+ \sum_{y=1}^{J} \sum_{d \leq d_y^* - 1} \Delta^{(y)}(d) \; g_{\boldsymbol{\theta},2}(y,d) + \left[\sum_{y=1}^{J} \sum_{d \geq d_y^*} \Delta^{(y)}(d)\right] g_{\boldsymbol{\theta},2}(y, d_y^*)$$

$$+ \sum_{y_{-1}=1}^{J} \sum_{y=1, y \neq y_{-1}}^{J} D^{(y_{-1},y)} \; \widetilde{\beta}_y(y, y_{-1}) - \sum_{y=1}^{J} \Delta^{(y)}(d_y^*) \; \Delta\beta_d(y, d_y^*)$$

# Model WITH duration dependence: Forward

*(i)* $U = \{H^{(y)}(d) : y \geq 1,\ d \leq d_y^* - 1,\ \sum_{d \geq d_y^*} H^{(y)}(d),\ \Delta^{(y)}(d) : y \geq 1,$ $d \leq d_y^* - 1,\ \sum_{d \geq d_y^*} \Delta^{(y)}(d)\}$ *is a sufficient statistic of* $\boldsymbol{\theta}$.

*(ii) The elements in the vector* $U$ *are linearly independent such that* $U$ *is a minimal sufficient statistic.*

*(iii) Conditional on* $U$*, the vector of statistics* $\{D^{(y_{-1},y)} : y_{-1},\ y \geq 1\}$ *are linearly independent and they identify the vector of structural parameters* $\{\widetilde{\beta}_y(y, y_{-1}) : y_{-1},\ y, y \neq y_{-1} \geq 1\}$.

*Furthermore, the vector of statistics* $\{\Delta^{(y)}(d^*) : y \geq 1\}$ *are also linearly independent and they identify the vector of structural parameters* $\{\Delta^{(y)}(d^*) : y \geq 1\}$.

# Model WITH duration dependence: Myopic

*EXAMPLE 4*. Consider two realizations of the history $\left(y_0, d_1 | \mathbf{y}^T\right)$

$$A = \left\{0, 0 \mid \mathbf{y}_{d^*-1}, 0, \mathbf{y}_{d^*+1}\right\} \quad \text{and} \quad B = \left\{0, 0 \mid \mathbf{y}_{d^*}, 0, \mathbf{y}_{d^*}\right\}$$

$U(A) = U(B)$. and $\ln \mathbb{P}\left(A | U\right) - \ln \mathbb{P}\left(B | U\right) = \Delta\beta_d(y, d_y^*)$.

# Identification of d*

- For the forward-looking model, we can identify $\Delta\beta_d(y, d)$ for any $d$ in $\{d_y^*, d_y^* + 1, ..., T - 2\}$.

- By definition of $d_y^*$,

$$d_y^* = \arg\max_d \{d : \Delta\beta_d(y, d) \neq 0 \text{ and } \Delta\beta_d(y, d+1) = 0\}$$

- Suppose that $d_y^* \leq T - 3$, then the definition above identifies $d_y^*$ from the data.

# 3. Estimation & Inference

# Conditional MLE

- The structural parameters in $\beta_y$ and $\beta_d$ (or a parametric specification of these functions) can be identified using Conditional MLE.
- With exogenous state variables, $\mathbf{z}_{it}$, we can use the Kernel-weighted Conditional Likelihood in Honoré & Kiriadzidou (2000).

# 4. Monte Carlo Experiment

# Rust (1987) bus engine replacement

- Keeping bus engine ($y_t = 1$) or replacing it ($y_t = 0$).
- Payoff function:

  $-RC_i + \varepsilon_{it}(0)$, if replacement

  $-MC_i - c(d_{it}) + \varepsilon_{it}(1)$, if no replacement
- We allow or UH in $RC_i$ and $MC_i$.

# DGPs

Table 3

Description of DGPs in the Monte Carlo experiments

| Parameter | DGP 1 | DGP 2 | DGP 3 |
|---|---|---|---|
| $\alpha_i(0) = -RC_i$<br>Random draws from: | $N(\mu, \sigma^2)$<br>$\mu = 8, \sigma = 2$ | Two types<br>$RC_1 = 4.5$, $RC_2 = 9$<br>$\lambda_1 = \lambda_2 = 0.5$ | Two types<br>$RC_1 = 8$, $RC_2 = 9$<br>$\lambda_1 = \lambda_2 = 0.5$ |
| $\alpha_i(1) = -c_{0i}$ | 0 | 0 | 0 |
| $\beta_d(d) = \beta\ d$ | $\beta = 1$ | $\beta = 1$ | $\beta = 1$ |
| $d^*$ | 3 | 3 | 3 |
| $\delta$ | 0.95 | 0.95 | 0.95 |
| Initial $y_0, d_1$ | 0, 0 | 0, 0 | 0, 0 |
| Maximum $T$ | 25 | 25 | 25 |
| $N$ | 1000 | 1000 | 1000 |
| # simulations | 1000 | 1000 | 1000 |

# Experiment. Estimate parameter

Table 4

Monte Carlo Experiments with DGP 1 (Normal RCs)

| Estimator of $\beta$ | Sample A (1 to 7) Estimate Mean | Sample A (1 to 7) Estimate St. dev. | Sample B (1 to 14) Estimate Mean | Sample B (1 to 14) Estimate St. dev. | Sample C (8 to 21) Estimate Mean | Sample C (8 to 21) Estimate St. dev. |
|---|---|---|---|---|---|---|
| FE-CMLE | 0.9996 | 0.1481 | 1.0021 | 0.0741 | 0.9958 | 0.0716 |
| MLE-2types | 0.9625 | 0.0478 | 0.8955 | 0.0265 | 0.8687 | 0.0289 |
| MLE-noUH | 0.6175 | 0.0277 | 0.5922 | 0.0210 | 0.5653 | 0.0213 |

# Test of the CRE

Table 4

Monte Carlo Experiments with DGP 1 (Normal RCs)

| Testing hypothesis | Freq. rejection significance level 1% | 5% | 10% | Freq. rejection significance level 1% | 5% | 10% | Freq. rejection significance level 1% | 5% | 10% |
|---|---|---|---|---|---|---|---|---|---|
| No UH | 0.504 | 0.753 | 0.854 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Two types | 0.007 | 0.046 | 0.095 | 0.138 | 0.348 | 0.466 | 0.244 | 0.481 | 0.610 |

# Using Rust data

| Empirical Distribution of Choice Histories | | | |
|---|---|---|---|
| | Frequency | | |
| Choice history | Absolute | % | % cumulative |
| 1101111111 | 3 | 5.17 | 5.17 |
| 1110111111 | 11 | 18.96 | 24.13 |
| 1111011111 | 9 | 15.51 | 39.64 |
| 1111101111 | 18 | 31.03 | 70.67 |
| 1111110111 | 7 | 12.07 | 82.74 |
| 1111111011 | 5 | 8.62 | 91.36 |
| 1111111101 | 3 | 5.17 | 96.53 |
| 1111111110 | 2 | 3.45 | 100.00 |

# Using Rust data

**Bus Engine Replacement (Rust, 1987)**

**Fixed-Effects-Conditional Maximum Likelihood**

| | $\beta$ | | *p-value* | |
|---|---|---|---|---|
| $d^*$ | $\widehat{\beta}$ | $se\left(\widehat{\beta}\right)$ | $H_0 : \beta = 0$ | log-*likelihood* |
| 4 | 0.2048 | 0.3576 | 0.5670 | -22.864 |
| **3** | **1.3218** | **0.5470** | **0.0160** | **-11.222*****

# Using Rust data

**Table 12**
**Bus Engine Replacement (Rust, 1987)**

**Hausman Test of Unobserved Heterogeneity**

| *Model* | $\widehat{\Delta\beta_d(d^*)}$ $(se)$ *MLE* | $\widehat{\Delta\beta_d(d^*)}$ $(se)$ *CMLE* | *Hausman* | *p-value* |
|---|---|---|---|---|
| Square root | 0.4294 (0.0700) | 1.3218 (0.5470) | 1.5479 | 0.2134 |
| Linear | 0.3601 (0.0552) | 1.3218 (0.5470) | 1.7181 | 0.1899 |
| Square | 0.3817 (0.0573) | 1.3218 (0.5470) | 1.6636 | 0.1971 |

# 5. Extensions / Conclusions

# Extensions

- Counterfactual experiments (and Marginal effects): Requires the identification of the distribution of $\theta$. Requires additional assumptions.
- **Stochastic evolution of $d_{it}$ (e.g., cumulative mileage)**
- **Dynamic games of incomplete information**.

# Summary and Conclusions

- We study identification in a fixed-effects structural dynamic logit, and obtain some positive identification results and a simple CMLE.
- Only parameters that capture structural state dependence can be identified. Counterfactual experiments?
- FE-estimator can be used also to test for the consistency of estimates using a CRE specification (Hausman type of test).