

# Using Subjective Expectations Data to Allow for Unobserved Heterogeneity in Hotz-Miller Estimation Strategies

Juan Pantano  
University of Arizona  
University of Hong Kong

Yu Zheng  
Queen Mary University of London and CEPR

DSE 2025 Conference, UCL  
July 4, 2025

- Importance of structural estimation as the cornerstone of theory-based empirical research in applied microeconomics.
- Seminal work of Hotz & Miller (1993) shows how to estimate the structural parameters of a discrete choice dynamic programming model without solving the optimization problem even once.
- The nature of HM implies no straight-forward way to incorporate unobserved heterogeneity. Recent progress:
  - Aguirregabiria & Mira (2007)
  - Kasahara & Shimotsu (2009)
  - Arcidiacono & Miller (2011)
  - Hu & Shum (2012)
  - Bonhomme, Lamadon & Manresa (2022)
- We propose a different approach that leverages on expectations data to introduce unobserved heterogeneity into a HM estimation strategy.

# Our Approach

- Use **subjective expectations of future choice probabilities** to allow for estimable unobserved heterogeneity in two-step estimation strategies for dynamic structural models.
- In the first step, we identify the unobserved types using the expectations data and group observations into homogeneous subsamples.
- In the second step, we employ Hotz-Miller in each subsample.
- We consider the ideal case where subjective expectations are precisely elicited.
- Also show that when they are rounded off to some **focal** points, most of our results from the ideal case still hold.

- Toy Model of Machine Replacement
- Monotonicity Property and A Single Report of Subjective Expectation
  - Precise Subjective Expectations
  - Focal Subjective Expectations
- More General Unobserved Heterogeneity and Two Reports of Subjective Expectations
  - Precise Subjective Probabilities
  - Focal Subjective Probabilities
  - Montecarlo Experiments
- Conclusion

# A Toy Model of Machine Replacement I

- Capital replacement “renewal” problem similar to that in Rust (1987).
- Let  $x_t$  be the age of the machine at time  $t$
- Current period profits from using a machine of age  $x_t$ :

$$\Pi(x_t, d_t, \varepsilon_{0t}, \varepsilon_{1t}) = \begin{cases} \theta x_t + \varepsilon_{0t} & \text{if } d_t = 0 \\ R + \varepsilon_{1t} & \text{if } d_t = 1 \end{cases}$$

where  $d_t = 1$  if the firm decides to replace the machine at  $t$ .

- $\theta$  captures the economic depreciation of the machine or maintenance cost
- $R$  is the net cost of a new machine
- $\varepsilon_t$ s are time specific i.i.d. type I extreme value shocks to the utilities/profits from replacing and not replacing

# A Toy Model of Machine Replacement II

- A Renewal Model with Stochastic Machine Aging

$$x_{t+1} = \begin{cases} \min \{5, x_t + 1\} & \text{with probability } \pi_f & \text{if } d_t = 0 \\ x_t & \text{with probability } 1 - \pi_f & \text{if } d_t = 0 \\ 1 & \text{with probability } 1 & \text{if } d_t = 1 \end{cases}$$

- Estimation is standard, and proceeds using either
  - Rust (1987) nested fixed point algorithm, or
  - Hotz-Miller (1993) two-step estimator
- Suppose now in at least one period, each firm, after making the replacement decision, is asked the question:  
*“What is the probability that you will replace your machine next year?”*
- Suppose firms are rational and give their answers using the solutions to the DP problem (subject to some form of measurement error)
- Leverage on this type of subjective expectations to help identify unobserved heterogeneity (UH)

## Adding Unobserved Heterogeneity: Monotonicity Property

- Assume time-invariant finite discrete types indexed by  $k = 1, \dots, K$ .
- Allow for heterogeneity in the structural parameter  $R$ :  $R_k$ ,  $k = 1, \dots, K$  such that  $R_1 > R_2 > \dots > R_K$ .
- Then the model implies

$$Pr(d = 1|x, k = 1) < Pr(d = 1|x, k = 2) < \dots < Pr(d = 1|x, k = K), \forall x$$

- A lower type has a lower CCP at any observable state
- We can estimate the model with the help of a single self reported (SR) subjective probability per firm.

# Adding Unobserved Heterogeneity: Monotonicity Property

- Assume time-invariant finite discrete types indexed by  $k = 1, \dots, K$ .
- Allow for heterogeneity in the structural parameter  $R$ :  $R_k$ ,  $k = 1, \dots, K$  such that  $R_1 > R_2 > \dots > R_K$ .
- Then the model implies

$$Pr(d = 1|x, k = 1) < Pr(d = 1|x, k = 2) < \dots < Pr(d = 1|x, k = K), \forall x$$

- A lower type has a lower CCP at any observable state
- We can estimate the model with the help of a single self reported (SR) subjective probability per firm.

More generally, this belongs to the class of models that satisfies a Monotonicity Property.

**Definition (Monotonicity Property):** The structure of the unobserved heterogeneity has the **monotonicity property** if for unobserved types  $k = 1, 2, \dots, K$ ,

$$Pr(d = 1|x, k = 1) < Pr(d = 1|x, k = 2) < \dots < Pr(d = 1|x, k = K), \forall x.$$



# Monotonicity: A Single Precise Self Report I

Let's start with an ideal (but unrealistic) case.

**Assumption SR-Precise:** Subjective probabilities are elicited with precision.

- Firm  $i$  uses the solution to the DP to construct:

$$\begin{aligned} & \tilde{p}_i^{SR}(d_{t+1} = 1 | x_t, d_t) \\ &= p_i^{SR}(d_{t+1} = 1 | x_t, d_t, k) \\ &= \sum_{x_{t+1}} Pr(d_{t+1} = 1 | x_{t+1}, k) f_x(x_{t+1} | x_t, d_t) \end{aligned}$$

## Monotonicity: A Single Precise Self Report II

- Basic intuition: if two firms  $A$  and  $B$  with  $x_{At} = x_{Bt} = x_t$ , and,  $d_{At} = d_{Bt} = d_t$ , but

$$\begin{aligned}\tilde{p}_A^{SR}(d_{A,t+1} = 1 | x_{At}, d_{At}) &\neq \tilde{p}_B^{SR}(d_{B,t+1} = 1 | x_{Bt}, d_{Bt}) \\ \Leftrightarrow p_A^{SR}(d_{A,t+1} = 1 | x_t, d_t, k_A) &\neq p_B^{SR}(d_{B,t+1} = 1 | x_t, d_t, k_B)\end{aligned}$$

$$\Rightarrow k_A \neq k_B.$$

- On the contrary, if two firms report the exact identical expected probability conditioning on the same state and choice, then they must have the same type
- Key Idea: differences in self-reports are informative about underlying unobserved heterogeneity.**
- In this case, in every state-choice cell, we will see exactly  $K$  different SRs and the ranking corresponds to their type

# Monotonicity: A Single Focal Self Report

- In reality, Assumption SR-Precise rarely holds.
- There is likely to be substantial “heaping” or “bunching” at common reference points like 0, 0.25, 0.50, 0.75 and 1. Call these points “focal” points.
- Let  $p_i^{SRF}(d_{t+1} = 1|x_{it}, d_{it}, k_i)$  be  $p_i^{SR}$  rounded off to the closest focal point.
- SRF, by itself, does not necessarily leads to identification problems: if at every  $(x, d)$ , we still have  $K$  different SRFs, then we are essentially back to the precise case
- It creates difficulty in the identification of types when different types bunch at focal points
  - Two obs of different types give reports at the same  $(x, d)$
  - Under Assumption SR-Precise, type 1 reports 22%, while type 2 reports 27%.
  - With bunching, both report 25%.
- Under bunching, we sketch an iterative procedure to estimate the model with unobserved types

# Monotonicity: A Single Focal Self Report

- Suppose we recognize  $B$  focal points (motivated by patterns in the actual SRs data).
- When we rank the SRFs at a given  $(x, d)$ , there can be at most  $B$  different SRFs.  $K \leq B$
- $K$  is pinned down by the max # of different SRFs in any state-choice cell.
- In state-choice cells with  $K$  different answers, the type is simply given by the rank of the firm in the ascending order of the SRs.
  - We say types are fully expressed in these state-choice cells
- In state-choice cells where we have  $< K$  different answers, though we can rank the SRs, we don't really know which answer corresponds to which type except that a smaller answer is associated with a lower type.
- This motivates us to consider an iterative procedure.

# Monotonicity: Iterative Procedure I

- 1 Divide the sample into two subsamples:
  - the subsample of firms whose types are revealed, as their SRFs are elicited in state-choice cells where types are fully expressed;
  - the complementary sample
- 2 Compute type-specific CCPs at as many states as possible for the obs from the first subsample
- 3 Fill in the remaining CCPs with “incorrect” CCPs calculated from the whole sample ignoring UH
- 4 Estimate the model using HM by type and obtain  $(R_1^{(1)}, R_2^{(1)}, \dots, R_K^{(1)}, \Gamma^{(1)})$ , where  $\Gamma$  collects all other structural parameters not varying by type
- 5 Given  $(R_1^{(1)}, R_2^{(1)}, \dots, R_K^{(1)}, \Gamma^{(1)})$ , we can calculate what is the model-implied one-period ahead expected CCP  $Pr(d_{t+1} = 1 | x_t = x, d_t = d, k), \forall x, d, k$ .

## Monotonicity: Iterative Procedure II

- 6 Go back to the subsample with indeterminate types, assign firm's type to be the type whose model-implied expected CCP at the relevant  $(x, d)$  is the closest to the actual SRF at that  $(x, d)$ :

$$k_i = \underset{k \in \{1, \dots, K\}}{\operatorname{argmin}} |Pr(d_{t+1} = 1|x, d, k) - \tilde{p}_i^{SRF}(d_{t+1} = 1|x, d)|$$

- 7 Now we have a type assignment for all obs in our sample. Estimate the model using HM by type and obtain  $(R_1^{(2)}, R_2^{(2)}, \dots, R_K^{(2)}, \Gamma^{(2)})$
- 8 Repeat Steps 6 and 7 to reassign types for obs in the problematic subsample until there is no type reassignment from the last iteration

## Interim Summary: Monotonicity

- We illustrate the core of the idea using a class of models with UH that satisfy the monotonicity property
- Using one subjective expectation of future action per obs, we can identify the unobserved the types and reduce the problem to estimating several homogeneous models with HM
- In many applications, the monotonicity property does not hold
- Next we illustrate how, using two subjective expectations per obs, we can restore identification of unobserved types

# Adding More General Unobserved Heterogeneity

- Allow for heterogeneity in the structural parameters capturing. for  $k = 1, \dots, K$ 
  - Maintenance costs  $\theta_k$  and
  - Machine replacement costs  $R_k$ .
- Note that the Monotonicity Property fails here.
  - Imagine two machines: One depreciates faster *and* is more costly to replace.
  - It can happen that initially, when the machine is relatively new, the prob of replacement is lower (i.e. the replacement cost is more salient). But when it ages, the prob of replacement is higher (i.e. the depreciation is more salient)
- Suppose we have **two** subjective expectations **precisely** elicited in two points in time per firm.
- The two self-reports will allow us to **“link”** firms with the same type even if they have never been asked the question at the same  $(x, d)$



# Linking Technology under Precise SRs

- We can only know immediately whether A and B have the same type if at the time of expectation elicitation, they happen to have the same  $(x, d)$ .
- With two reports, it opens up the possibility of “linking” observations.
- If A and B are revealed to have the same type in a particular  $(x, d)$  and B and C are revealed to have the same type in another  $(x', d')$ , then A and C have the same type even if they have no  $(x, d)$  in common.
  - Note B gives SRs at  $(x, d)$  and  $(x', d')$ , while A gives SRs at  $(x, d)$  and  $(\tilde{x}, \tilde{d})$  and C gives SRs at  $(x', d')$  and  $(\hat{x}, \hat{d})$
- This is achieved by a “linking technology”, which mathematically is to define an equivalence relation between observations.

**Definition (Linking Technology):** Define a binary relation,  $R$ , in the following way:  $\forall i, j \in \{1, 2, \dots, N\}$ ,

$$i R j \Leftrightarrow$$

$$\{\tilde{p}_i^{SR}(x_{it'}, d_{it'}), \tilde{p}_i^{SR}(x_{it''}, d_{it''})\} \cap \{\tilde{p}_j^{SR}(x_{jt'}, d_{jt'}), \tilde{p}_j^{SR}(x_{jt''}, d_{jt''})\} \neq \emptyset$$

The **linking technology** is a relation  $\sim$  on  $\{1, 2, \dots, N\}$ :  $\forall i, j \in \{1, 2, \dots, N\} = I$ ,

$$i \sim j$$

iff  $\exists$  a (possibly empty) subset of observations  $\{i_1, i_2, \dots, i_n\} \subseteq I$ , such that

$$i R i_1 R i_2 R \dots R i_n R j.$$

- It can be easily checked that  $\sim$  satisfies reflexivity, symmetry, and transitivity, and hence defines an equivalence relation.

# Lemma

- By the Fundamental Theorem of the Equivalence Relations, the linking technology partitions the set of all observations.
- But does the resulting partition give the partition of the sample in terms of types?

**Assumption SR-No Islands:** Define  $\Sigma^k$  to be the set of all state-choice cells at which a type  $k$  observation makes a self-report in the data. Then,  
 $\forall (x, d), (x', d') \in \Sigma^k, \exists$  observations  $m$  and  $n$  of type  $k$ , with  $m$  reporting at  $(x, d)$ , and  $n$  reporting at  $(x', d')$ , and  $m \sim n$ .

**Lemma.** Under **Assumptions SR-Precise** and **SR-No Islands**, the linking technology recovers the true number of types and the type membership for each observation.

Proof of the Lemma

When we allow for Focal SRs, we inevitably encounter bunching.

**Definition (Bunching):** Two SRFs are said to be **bunched at**  $(x, d)$  **for observations**  $i$  **and**  $j$  of different types, if  $p_i^{SRF}(x, d, k_i) = p_j^{SRF}(x, d, k_j)$  and  $k_i \neq k_j$ . Two SRs are said to be **bunched at**  $(x, d)$  **for types**  $k$  **and**  $k'$ , if  $p^{SRF}(x, d, k) = p^{SRF}(x, d, k')$ .

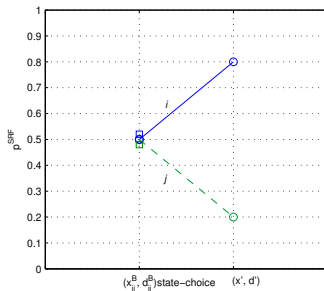
- Reporting different SRFs at a common  $(x, d) \Rightarrow$  different types
- Reporting the same SRFs at a common  $(x, d)$  may or may not imply the same type
- Therefore, we need some observations to reveal the bunching of types

## Focal SRs and Bunching II

**Assumption B1 (Detection of Bunching Types)** If two types bunch at the state-choice  $(x, d)$ , then  $\exists$  two observations  $i$  and  $j$  of different types and another state-choice  $(x', d')$  s.t.

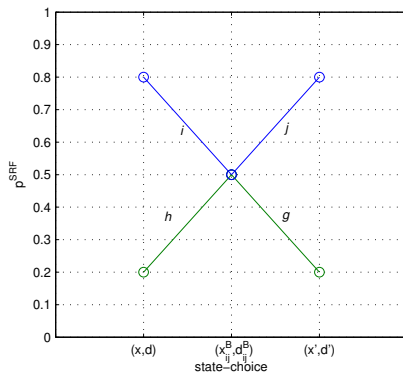
$$\tilde{p}_i^{SRF}(x, d) = \tilde{p}_j^{SRF} \text{ and } \tilde{p}_i^{SRF}(x', d') \neq \tilde{p}_j^{SRF}(x', d')$$

Denote  $(x, d) = (x_{ij}^B, d_{ij}^B) = (x_{k_i k_j}^B, d_{k_i k_j}^B)$ .



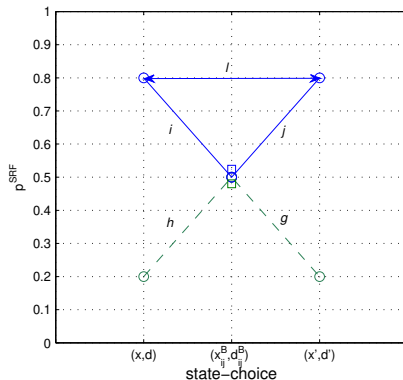
# Focal SRs and Bunching III

However, Assumption B1 is not enough to identify the types. Consider the following



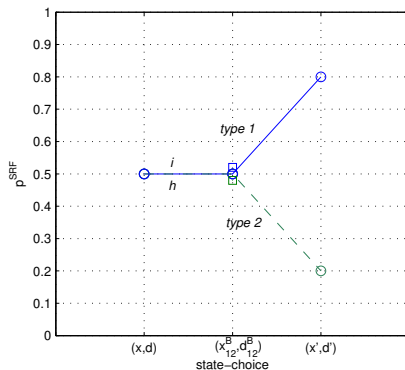
# Focal SRs and Bunching IV

**Assumption B2 (Bridging Bunchings)** For all observations  $i$  and  $j$  who belong to the same type, with singleton intersection of SRs at  $(x_{ih}^B, d_{ih}^B)$  for some  $h$ , there exists another observation  $l$  of the same type as  $i$  and  $j$ , who has SRs in the two non-bunching state-choice cells.



# Focal SRs and Bunching V

**Assumption B3 (No observations with two "bunched" self-reports)** Every observation  $i$  has at least one SR elicited at a state-choice in which there is no bunching.





# Linking Technology under Bunching I

**Definition (Linking Technology under Bunching):** Define a binary relation,  $R^B$ , in the following way:  $\forall i, j \in \{1, 2, \dots, N\}$ ,

$$i R^B j$$

iff the following conditions are met:

1. The pairs of self reports for  $i$  and  $j$  are such that

$$\{\tilde{p}_i^{SRF}(x_{it'}, d_{it'}), \tilde{p}_i^{SRF}(x_{it''}, d_{it''})\} \cap \{\tilde{p}_j^{SRF}(x_{jt'}, d_{jt'}), \tilde{p}_j^{SRF}(x_{jt''}, d_{jt''})\} \neq \emptyset$$

2. If  $\exists$  observation  $g$  and  $h$  of different types,

$$\begin{aligned} & \{\tilde{p}_i^{SRF}(x_{it'}, d_{it'}), \tilde{p}_i^{SRF}(x_{it''}, d_{it''})\} \cap \{\tilde{p}_j^{SRF}(x_{jt'}, d_{jt'}), \tilde{p}_j^{SRF}(x_{jt''}, d_{jt''})\} \\ &= \{\tilde{p}_i^{SRF}(x_{gh}^B, d_{gh}^B)\} \end{aligned}$$

then  $\exists l$  such that

$$\begin{aligned} & \{\tilde{p}_l^{SRF}(x_{lt'}, d_{lt'}), \tilde{p}_l^{SRF}(x_{lt''}, d_{lt''})\} \\ &= \{\tilde{p}_i^{SRF}(x_{it'}, d_{it'}), \tilde{p}_i^{SRF}(x_{it''}, d_{it''})\} \triangle \{\tilde{p}_j^{SRF}(x_{jt'}, d_{jt'}), \tilde{p}_j^{SRF}(x_{jt''}, d_{jt''})\} \end{aligned}$$

where  $\triangle$  denotes the set difference.

# Linking Technology under Bunching II

The **linking technology under bunching** is a relation  $\sim^B$  on  $\{1, 2, \dots, N\}$ :  
 $\forall i, j \in \{1, 2, \dots, N\} = I$ ,

$$i \sim^B j$$

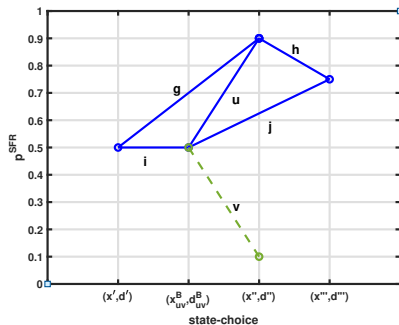
iff  $\exists$  a (possibly empty) subset of observations  $\{i_1, i_2, \dots, i_n\} \subseteq I$ , such that

$$i R^B i_1 R^B i_2 R^B \dots R^B i_n R^B j.$$

# Proposition

**Proposition** Under [Assumptions B1,B2,B3](#) and [SR-No Islands](#), the linking technology under bunching recovers the true number of types and the type membership of each observation.

The key mechanism is illustrated below. Consider firm  $i$  and  $j$ .



# Montecarlo: Model of Machine Replacement

We can write a computer algorithm to implement the linking technology for the toy model of machine replacement.

Suppose that there are  $K = 2$  types and the DGP is

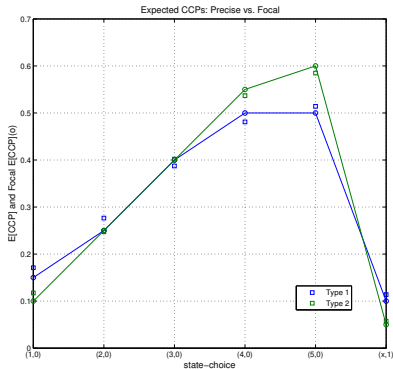
$$\text{Type 1: } (\theta_{11}, R_1) = (-0.27, -2.65)$$

$$\text{Type 2: } (\theta_{12}, R_2) = (-0.40, -3.75)$$

We simulate data on  $N = 10,000$  firms and  $T = 10$  periods from the model parametrized above. SRs are elicited in the 6th and 9th period for each firm.

DGP with SRF no Bunching

The simulated precise and focal self reports at each state-choice are



Focal self-reports lead to bunching in state choice-combinations  $(x, d) = (2, 0)$  and  $(x, d) = (3, 0)$ .

- Note that **Assumption B3 (No observations with two “bunched self-reports)** fail here as long as there are observations who give their two self-reports at (2,0) and (3,0). For these observations, there is no way of knowing which type they belong to.
- Approach 1: Drop these problematic obs and use only the remaining subsample to estimate the model: “Discarded”
- Approach 2: Assign the type of these problematic obs in the following way: “Full”
  1. Use the complementary subsample to form a system of equations linking the  $\tilde{p}^{SRF}$  to the CCPs for each type (ignore rounding) to solve for type-specific CCPs
  2. Use the type-specific CCPs and the transition prob to construct the conditional prob. of  $i$  being type  $k$  given  $i$ 's history of choices and states for every problematic obs

Table: Estimation Results, Linking Technology Under Bunching

	Truth	"Discarded"		"Full"	
		Mean	SD	Mean	SD
$\theta_{11}$	-0.27	-0.2735	0.0016	-0.2709	0.0015
$R_1$	-2.65	-2.6227	0.0088	-2.6561	0.0082
$\theta_{12}$	-0.40	-0.4006	0.0024	-0.3977	0.0022
$R_2$	-3.75	-3.7024	0.0138	-3.7329	0.0134
Avg. Time	-	24.1 seconds		76.6 seconds	

- Intuitive that "discarded" approach introduces an downward bias to cost of replacement (e.g.  $R_1$  ad  $R_2$ )
- The "Full" approach performs well

- New strategy to allow for unobserved heterogeneity in a two-step, CCP-based estimation strategies for discrete choice dynamic programming models
- Explicit use of expectations data, as subjective expectations about future choice probabilities are informative about unobserved types
- One can think of extensions to allow for continuous types or exogenous type transitions
- Complement methods that infer unobserved types from observed history of states and choices



# Extra Slides

# Proof of the Lemma

**Lemma.** Under [Assumptions SR-Precise](#) and [SR-No Islands](#), the linking technology recovers the true number of types and the type membership for each observation.

*Proof.* First, show that  $\forall i \sim j$ , we have  $k_i = k_j$ . By the transitivity of “=”, it is sufficient to show  $\forall m R n \Rightarrow k_m = k_n$ . True under [Assumption SR-Precise](#).

# Proof of the Lemma

**Lemma.** Under [Assumptions SR-Precise](#) and [SR-No Islands](#), the linking technology recovers the true number of types and the type membership for each observation.

*Proof.* First, show that  $\forall i \sim j$ , we have  $k_i = k_j$ . By the transitivity of “=”, it is sufficient to show  $\forall m R n \Rightarrow k_m = k_n$ . True under [Assumption SR-Precise](#).

Next, show that whenever  $k_i = k_j = k$ , we must have  $i \sim j$ . Suppose not:  $i \not\sim j$ . Let  $(x, d)$  (or  $(x', d')$ ) be the state-choice at which  $i$  (or  $j$ ) gives SR. Since  $(x, d) \in \Sigma^k$  and  $(x', d') \in \Sigma^k$ , by [Assumption SR-No Island](#),  $\exists m$  and  $n$  of type  $k$ , with  $m$  reporting at  $(x, d)$  and  $n$  reporting at  $(x', d')$ , and  $m \sim n$ . But note that  $i$  and  $m$  both have type  $k$  and report at the same  $(x, d)$ , then under [Assumption SR-Precise](#), their reports must be identical and  $i \sim m$ . Likewise for  $j \sim n$ . By transitivity,  $i \sim j$ . A contradiction.

# Proof of the Proposition

**Proposition** Under [Assumptions B1,B2,B3](#) and [SR-No Islands](#), the linking technology under bunching recovers the true number of types and the type membership of each observation. *Proof.* “ $\Rightarrow$ ”. Suffice to show that

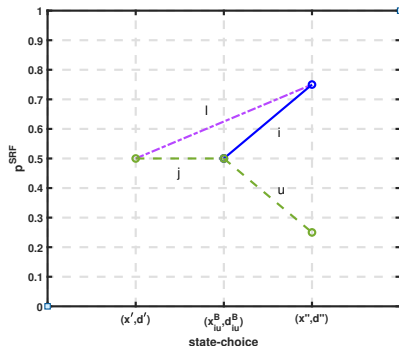
$i R^B j \Rightarrow k_i = k_j$ . Suppose  $k_i \neq k_j$ . By def of  $R^B$ ,  
 $\{\tilde{p}_i^{SRF}(x_{it'}, d_{it'}), \tilde{p}_i^{SRF}(x_{it''}, d_{it''})\} \cap \{\tilde{p}_j^{SRF}(x_{jt'}, d_{jt'}), \tilde{p}_j^{SRF}(x_{jt''}, d_{jt''})\} \neq \emptyset$ .  
By [Assumption B3](#), there is at most one common SR. Say the common SR is elicited at  $(x, d)$ . Then by [Assumption B1](#), the bunching of the two types is revealed by observations that do not necessarily involve  $i$  or  $j$ . There are thus three possibilities.

# Proof of the Proposition I

Case 1:  $i$  and  $j$  themselves reveal the bunching of their types. Then their another SRs are elicited at another common state-choice cell  $(x', d')$  and  $\tilde{p}_i^{SRF}(x', d') \neq \tilde{p}_j^{SRF}(x', d')$ . But then  $j$  satisfies the role of  $h$  in the def of  $R^B$ , it follows that  $\exists l$  that bridges the other SRs of  $i$  and  $j$ . This is impossible as  $l$  cannot give different SRs at the same state-choice cell. A contradiction.

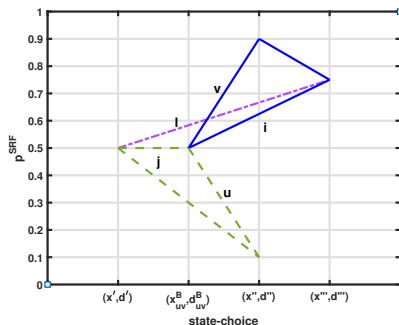
## Proof of the Proposition II

Case 2:  $\exists u$  with  $k_u = k_j \neq k_i$  which reveals the bunching,  $(x_{iu}^B, d_{iu}^B)$ . Because  $j R^B i$ , and  $u$  satisfies the role of  $h$  in the def of  $R^B$ ,  $\exists l$  that bridges  $j$  and  $i$ . Since  $l$  has a different SR at  $(x'', d'')$  than  $u$ , then  $k_l \neq k_u = k_j$ . Then it follows that  $(x', d')$  is also a bunching state-choice cell and  $j$  gives SRs at two bunching state-choice cells. A contradiction to [Assumption B3](#).



## Proof of the Proposition III

Case 3:  $\exists u$  and  $v$  with  $k_v = k_i \neq k_u = k_j$  who reveal the bunching of the types. Since  $i R^B j$ ,  $\exists l$  with  $\{\tilde{p}_j^{SRF}(x', d'), \tilde{p}_i^{SRF}(x''', d''')\}$ . If  $k_l = k_i$ , then  $(x', d')$  will be a bunching state-choice cell and  $j$ 's SRs are both at bunching state-choice cells. If  $k_l = k_j$ , then  $(x''', d''')$  will be a bunching state-choice cell and  $i$ 's SRs are both at bunching state-choice cells. If  $k_l \neq k_i$  and  $k_l \neq k_j$ , then  $l$ 's SRs are both at bunching state-choice cells. Contradiction to [Assumption B3](#).



## Proof of the Proposition IV

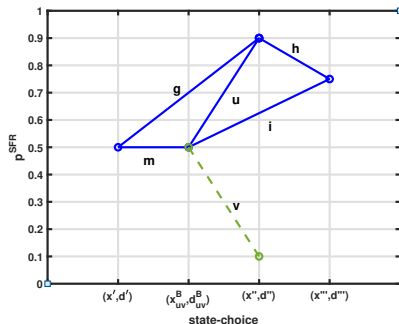
“ $\Leftarrow$ ”. Suppose  $k_i = k_j = k$  but  $i \not\sim^B j$ . Let  $(x, d)$  (or  $(x', d')$ ) be the state-choice at which  $i$  (or  $j$ ) gives SR. Since  $(x, d) \in \Sigma^k$  and  $(x', d') \in \Sigma^k$ , by [Assumption SR-No Island](#),  $\exists m$  and  $n$  of type  $k$ , with  $m$  reporting at  $(x, d)$  and  $n$  reporting at  $(x', d')$ , and  $m \sim^B n$ . Then  $i$  and  $m$  have the same type and make the same SR at  $(x, d)$ .

If  $(x, d)$  is not a bunching state-choice cell, then the linking technology immediately gives  $i \sim^B m$ .



# Proof of the Proposition V

If  $(x, d)$  is a bunching state-choice cell for type  $k$ , then [Assumption B1](#) implies that  $\exists$  possibly different obs  $u$  of type  $k$  and  $v$  of a different type to reveal  $(x_{uv}^B, d_{uv}^B)$ . [Assumption B2](#) makes sure that  $\exists h$  that bridges  $i$  and  $u$  and  $\exists g$  that bridges  $u$  and  $m$ . Under the linking technology, we will have  $i R^B u R^B m$  or  $i \sim^B m$



By the same token,  $n \sim^B j$  and hence  $i \sim^B m \sim^B n \sim^B j$ . Contradiction.

Back

# Montecarlo: DGP with SRFs but No Bunching

Suppose that there are  $K = 2$  types and the DGP is

Type 1:  $(\theta_{11}, R_1) = (-0.27, -2.65)$

Type 2:  $(\theta_{12}, R_2) = (-0.40, -3.75)$

