Partially observed dynamic model

Fu Zixuan

May 12, 2025

1 Introduction

A dynamic model of schooling and occupation choice has existed for a long time. Since its inception, a richer set of elements has been incorporated into the model, either from a data perspective or through the relaxation of assumptions. Specifically, in the early days, with panel data containing individual characteristics, schooling decisions, and labor market outcomes, researchers typically estimated a choice probability in the form of $f_{y_t|x_t,c}$. Later, realizing that x_t and c only capture observed differences between individuals, researchers became more ambitious in recognizing unobserved differences—modeled as a latent type z_t. The conditional choice probability then became $f_{y_t|x_t,c,z_t}$.

In the case of a time-invariant latent type $z_t = z$, specific estimation methods have been proposed (Aguirregabiria and Mira, 2007), and identification arguments have been developed in Kasahara and Shimotsu (2009). To further relax this assumption and allow z_t to be time-varying, estimation methods are introduced by Arcidiacono and Miller (2008), and identification arguments are discussed in Hu and Shum (2012). This kind of model is commonly applied in different settings, from Miller (1984) on occupation choice and Pakes (1986) on patent choice, to Arcidiacono (2005) on schooling decision.

Another direction of extension is to enrich the set of data variables that can affect individuals' choices. Beyond socioeconomic status, distance, and cognitive ability, what else can influence students' schooling and occupational decisions? The paper by Todd and Zhang (2020) explores the non-cognitive dimension—specifically, personality traits measured over an extended period in their dataset. One might treat these measurements as the true values, but a critique is that personality, unlike other characteristics, is often measured with noise. Therefore, the authors treat personality—one of the observed state variables, denoted X^M—as a noisy measurement of the true X, with an error term ζ that is independent of X and normally distributed.

However, what assumptions can be further relaxed in this modeling framework? For example, in many of the seminal papers on errors-in-variables models (Hu, 2008; Hu and Schennach, 2008; Carroll et al., 2010), some form of repeated measurements of the true underlying variable is a common approach to address the issue. The key question is to define what constitutes a valid repeated measurement of unobserved varaibles, and how we can make use of them, albeit noisy, to recover the model primitives, which would involve the unobserved variables.

Motivated by the assumptions made in Todd and Zhang (2020), it is not immediately clear whether one can recover $f_{Y_t|X_t,Z_t}$ when only a panel of (Y_t, X_t^M) is observed. Since both X_t and its measurement X_t^M are time-varying, my initial intuition is that the model primitives may not be identified. If that is not the case, what additional assumptions would be required?

In the following section, I will start from a simplified case and then move on to the one corresponding to Todd and Zhang (2020). Unfortunately, I have not yet succeeded in identifying even the simplest case. Some preliminary investigations and discussions will be provided in the final section.

2 Review and Overview

Approximately folloiwng the notation in Hu and Shum (2012), we let W_t to represent the observed variables, Z_t as the latent type. The model is characterized by the folloiwng decompostion

$$f_{W_t,Z_t|W_{t-1},Z_t} = f_{W_t|W_{t-1},Z_t} f_{Z_t|W_{t-1},Z_{t-1}} \tag{1}$$

has shown that without the extra monotonicity assumption, the $f_{W_t,Z_t|W_{t-1},Z_t}$ are identified up to a common permutation/ordering of the latent variable Z_t for all t.

In the context of Markovian dynamic optimzation model, where the observed variable $W_t = (Y_t, M_t)$, the above probability can be further decomposed. A caveat here is that state variable M_t is realized after the transition of latent type Z_t.

$$f_{Y_t,M_t|Y_{t-1},M_{t-1},Z_t} f_{Z_t|Y_{t-1},M_{t-1},Z_{t-1}} = f_{Y_t|M_t,Z_t} f_{M_t|Y_{t-1},M_{t-1},Z_t} f_{Z_t|Y_{t-1},M_{t-1},Z_{t-1}} \tag{2}$$

This is the form of CCP $\times$ state transition $\times$ type transition.

In Todd and Zhang (2020), 1, thus 2, are not satisfied because here the assumption is that latent type transition after the state variable M_t is realized. Therefore, the Markovian dynamic optimization becomes

$$f_{Y_t|M_t,Z_t} f_{M_t|Y_{t-1},M_{t-1},Z_{t-1}} f_{Z_t|Y_{t-1},M_{t-1},Z_{t-1},M_t} \tag{3}$$

The identification of the model primitives is not entirely nested within Hu and Shum (2012) but is shown in Hu et al. (2017).[1]

Though in both Markovian decision process, the conditional choice probability[2] is the same, the differences lies in the order of realization of the observed and the latent.

As far as I am concerned, there is some arbitrariness in deciding which one to follow. Although a natural question would be to ask whether there is a way to test for this type of assumption? Since this is not the main focus of the proposal, I will not dive into the question.

However, by recognizing the possible measurement issue in one of the observed state variable, e.g., personality, we can modify the above variables such that M_t is the measurement of the true state variable X_t. Therefore, the model is endowed with the following primitives

$$\begin{aligned} f_{Y_t,X_t,Z_t|Y_{t-1},X_{t-1},Z_{t-1}} &= \underbrace{f_{Y_t|X_t,Z_t}}_{\text{CCP}} \underbrace{f_{X_t|Y_{t-1},X_{t-1}}}_{\text{state transition}} \underbrace{f_{Z_t|Y_{t-1},X_{t-1},Z_{t-1},X_t}}_{\text{type transition}} \\ &= f_{Y_t|X_t,Z_t} f_{X_t|Y_{t-1},X_{t-1}} \underbrace{f_{Z_t|X_{t-1},Z_{t-1},X_t}}_{\text{not depend on } Y_{t-1}} \end{aligned} \tag{4}$$

As can be seen, the above decomposition does not involve the measurement M_t, which is undeirable for further identification arguments.

Now I move to the first simplified model where there is no latent type Z_t. The second would be to have a time-invariant latent type $Z_t = z$ before moving onto the full-fledged case as in Todd and Zhang (2020).

Toy Model 1 For a Markovian decision process without latent types, the model is characterized by the following primitives:

- $f_{Y_t|X_t}$: the conditional choice probability given the (unobserved) true state,
- $f_{X_t|X_{t-1},Y_{t-1}}$: the transition probability of the latent state variable,
- $f_{M_t|X_t}$: the measurement error distribution, where M_t is a noisy observation of the true state X_t.

[1]Previously, in my presentation, I thought this model fits exactly under the previous ones.
[2]Or the *output/emission* probability in the hidden Markov model setup.

In addition to the Markovian structure on (Y_t, X_t), we impose the following assumptions concerning the measurement process:

- $f_{Y_t|X_t,M_t} = f_{Y_t|X_t}$: the measurement M_t does not provide additional information about the decision outcome Y_t once the true state X_t is known,

- $f_{M_t|X_t,M_{t-1}} = f_{M_t|X_t}$: current measurements are conditionally independent of past measurements given the true state. This again mirrors the assumption in Hu (2008). A natural question is whether M_{t-1} can be used as an instrument for M_t. Translating the rank and invertibility conditions from Hu (2008) to this context, I need to verify whether the assumptions $\text{Rank}(f_{X_t|M_{t-1}}) = k$ and the invertibility of $f_{M_t|X_t}$ are consistent with the model's structure[3] and can thus be imposed,

- $f_{M_t|X_t} = f_{M_{t+1}|X_{t+1}}$: the measurement error distribution is time-invariant. This is a rather intuitive restriction that may or may not be necessary for identification.

We consider the joint distribution of observables under this measurement model. With X_t denoting the true latent state and M_t its noisy observation, the distribution of the panel $(Y_t, M_t, M_{t-1}, Y_{t-1})$ is given by:

$$f_{Y_t,M_t,M_{t-1},Y_{t-1}} = \sum_{x_t} f_{Y_t|x_t} \cdot f_{M_t|x_t} \cdot f_{x_t,M_{t-1},Y_{t-1}}$$

$$\begin{aligned} f_{X_t,M_{t-1},Y_{t-1}} &= \sum_{x_{t-1}} f_{X_t,M_{t-1},Y_{t-1}|X_{t-1}} \cdot f_{x_{t-1}} \\ &= \sum_{x_{t-1}} f_{X_t|X_{t-1},Y_{t-1}}\, f_{M_{t-1}|X_{t-1}}\, f_{Y_{t-1}|X_{t-1}} \cdot f_{x_{t-1}} \end{aligned}$$

Together, we have:

$$f_{Y_t,M_t,M_{t-1},Y_{t-1}} = \sum_{x_{t-1}} \sum_{x_t} f_{Y_t|x_t}\, f_{M_t|x_t}\, f_{x_t|x_{t-1},Y_{t-1}}\, f_{M_{t-1}|x_{t-1}}\, f_{Y_{t-1}|x_{t-1}}\, f_{x_{t-1}}.$$

In fact, if a stationary distribution is assumed for each t, and if M_{t-1} can serve as a valid instrument for X_t, then the only additional element beyond the setup in Hu (2008) is the state transition probability $f_{X_t|X_{t-1},Y_{t-1}}$, which I aim to identify using panel data. Once the (staionary) choice probability $f_{Y_t|X_t}$ and the measurement error distribution $f_{M_t|X_t}$ are identified, the state transition probability can be obtained by inverting the above equation.

[3] Not to be internally refuted by the model.

The only unknown is $f_{X_t|X_{t-1},Y_{t-1}}$. However, I am not sure whether the issue of ordering the latent state variable X_t is a concern here.

Toy Models 2 By adding back a constant latent type $Z_t = Z$, the model becomes:

- $f_{Y_t|X_t,Z}$: the conditional choice probability given the (unobserved) true state and fixed type,
- $f_{X_t|X_{t-1},Y_{t-1},Z}$: the transition probability of the latent state variable, conditional on the previous state, action, and type.

My attempt is to first identify the $f_{Y_t,M_t|Y_{t-1},M_{t-1},Z}$ by following the (constructive) arguement in Higgins and Jochmans (2023). Once we have the above element, can I apply the same argument as in Toy Model 1 where everything is conditional on the type $Z = z$? In this case, for each type z, we take the the distribution $f_{Y_t,M_t|Y_{t-1},M_{t-1},Z=z}$ (or the joint distribution $f_z(Y_t, M_t \mid Y_{t-1}, M_{t-1})$) and recover the choice probability $f_{Y_t|X_t,Z=z}$ and the state transition probability $f_{X_t|X_{t-1},Y_{t-1},Z=z}$. To the contrary of model 1, there is a need to ensure that for each type z, the ordering of the true state variable X_t is the same acrossing all types. Is this automatically ensured by some kind of joint decomposition or do we need to impose additional monoticity assumption?

Toy Model 3 For a time-varying latent type Z_t, the model admits two versions depending on the timing of type and state realization, as discussed earlier. I focus on the version studied in Hu et al. (2017), where the type is realized *after* the state. An additional assumption is imposed:

$$f_{Z_t|Y_{t-1},X_{t-1},Z_{t-1},X_t} = f_{Z_t|X_{t-1},Z_{t-1},X_t}$$

which excludes a direct dependence of Z_t on past action Y_{t-1}, conditional on past and current states and previous type.

First, in Hu et al. (2017), the ordering issue is resolved by an additional monotonicity assumption, as in Hu and Shum (2012). Therefore, I would need to adapt the argument in Higgins and Jochmans (2024) to show that this assumption is not necessary either.

Second, for the two models under consideration (2 and 3), new challenges arise. The idea is to first treat M as the true state variable and apply existing identification methods accordingly. Then, given the *mismeasured* version of the choice probability $f_{Y_t|M_t,Z_t}$, state transition probability, and type transition probability, we ask whether it is possible to recover the true underlying primitives—those in which M is replaced by the true latent state X.

Essentially, this is a two-step procedure. The first step is to identify the *mismeasured* model primitives. The second step is to recover the true model primitives. This paper focuses on the second step, as the first step has already been addressed in the existing literature. My conjecture is that The second step involves more severe issues related to the ordering of unobserved variables, which require further investigation.

3 Discussion

Starting from the simplest case, my goal is to identify the additional component of the model—namely, the state transition probability $f_{X_t|X_{t-1},Y_{t-1}}$—using the panel data (Y_t, M_t). This component is absent in Hu (2008), where there is no dynamic evolution.

Once this is achieved, introducing a constant latent type Z still appears to be a significant leap, as it introduces an additional unobserved variable. In terms of model primitives, this means that—on top of identifying the state transition—we must also identify the **type distribution**. It is not clear whether the data contain sufficient variation or structure to identify both components.

Going a step further and allowing the type to vary over time (i.e., letting Z_t evolve) makes the identification problem even more ambitious. An additional object to be identified is the **type transition probability**. I am even less optimistic about this case. I am not sure either whether introducing two state variables—say X^1 and X^2, where one is noisily measured while the other is well observed—would help.

Perhaps taking a step in the opposite direction, suppose that the measurement error is independent of the true state, i.e., $M_t = X_t + \zeta$, where ζ is independent of X_t. Under this assumption, is the model identified? If independence alone is not sufficient, is assuming normality of the measurement error—i.e., $\zeta \sim \mathcal{N}(0, \sigma^2)$—enough to ensure identification?

References

Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53.

Arcidiacono, P. (2005). Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica*, 73(5):1477–1524.

Arcidiacono, P. and Miller, R. A. (2008). Ccp estimation of dynamic discrete choice models with unobserved heterogeneity. *Manuscript. Duke University*.

Carroll, R. J., Chen, X., and Hu, Y. (2010). Identification and estimation of nonlinear models using two samples with nonclassical measurement errors. *Journal of nonparametric statistics*, 22(4):379–399.

Higgins, A. and Jochmans, K. (2023). Identification of mixtures of dynamic discrete choices. *Journal of Econometrics*, 237(1):105462.

Higgins, A. and Jochmans, K. (2024). Learning markov processes with latent variables. *Econometric Theory*, pages 1–13.

Hu, Y. (2008). Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *Journal of Econometrics*, 144(1):27–61.

Hu, Y. and Schennach, S. M. (2008). Instrumental variable treatment of nonclassical measurement error models. *Econometrica*, 76(1):195–216.

Hu, Y. and Shum, M. (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics*, 171(1):32–44.

Hu, Y., Shum, M., Tan, W., and Xiao, R. (2017). A simple estimator for dynamic models with serially correlated unobservables. *Journal of Econometric Methods*, 6(1):20150011.

Kasahara, H. and Shimotsu, K. (2009). Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica*, 77(1):135–175.

Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political economy*, 92(6):1086–1120.

Pakes, A. (1986). Patents as options: Some estimates of the value of holding european patent stocks. *Econometrica*, 54(4):755–784.

Todd, P. E. and Zhang, W. (2020). A dynamic model of personality, schooling, and occupational choice. *Quantitative Economics*, 11(1):231–275.