Closed-Form Identification of Dynamic Discrete Choice Models with Proxies for Unobserved State Variables*

Yingyao Hu† and Yuya Sasaki

Department of Economics, Johns Hopkins University

March 14, 2015

Abstract

For dynamic discrete choice models of forward-looking agents where a continuous state variable is unobserved but its proxy is available, we derive closed-form identification of the structure by explicitly solving integral equations based on deconvolution techniques. For an empirical illustration, we analyze industry dynamics in Chile.

Keywords: closed-form identification, deconvolution, dynamic discrete choice models

*The paper was circulated as "Dynamic Discrete Choice Models with Proxies for Unobserved Technologies."

†The authors can be reached at yhu@jhu.edu and sasaki@jhu.edu. We benefited from useful comments by seminar participants at Cambridge, CREST, George Washington, LSE, Oxford, Rice, SHUFE, Texas A&M, Toulouse School of Economics, UCL, Wisconsin-Madison, 2013 Greater New York Metropolitan Area Econometrics Colloquium, and 2014 Shanghai Econometrics Workshop. The usual disclaimer applies.

1 Introduction

Forward-looking agents making dynamic decisions based on unobserved state variables are of interest in economic researches. While econometricians may not observe the true state variables, they often have access to or can construct proxy variables. To estimate the dynamic discrete choice models, would it make sense to substitute a proxy variable for the true state variable? Because of the nonlinearity of the forward-looking discrete choice structure, a naive substitution of the proxy generally biases the estimates of structural parameters, even if the proxy has only an independent error. In this paper, we develop closed-form identification of dynamic discrete choice models when a proxy for an unobserved continuous state variable is available.

Suppose that agent j at time t makes exit decisions $d_{j,t}$ based on its technology $x^*_{j,t}$. Suppose also that we obtain a proxy $x_{j,t} = x^*_{j,t} + \varepsilon_{j,t}$ for the unobserved technology $x^*_{j,t}$ with a classical error $\varepsilon_{j,t}$. If $x^*_{j,t}$ were observable, then identification of the structural parameters of forward-looking agents follows from identification of two auxiliary objects: (1) the conditional choice probability (CCP) denoted by $\Pr(d_t \mid x^*_t)$; and (2) the law of state transition denoted by $f(x^*_t \mid d_{t-1}, x^*_{t-1})$ (Hotz and Miller, 1993). We show that these two auxiliary objects, $\Pr(d_t \mid x^*_t)$ and $f(x^*_t \mid d_{t-1}, x^*_{t-1})$, are identified using the proxies $x_{j,t}$ without observing the true states $x^*_{j,t}$.

Indeed, dynamic discrete choice models with unobservables are extensively studied in the literature (e.g., Aguirregabiria and Mira, 2007; Kasahara and Shimotsu, 2009; Arcidiacono and Miller, 2011; Hu and Shum, 2012 – see also the survey by Aguirregabiria and Mira, 2013), but no preceding work handles continuous unobservables like technologies. Our methods allow for continuously distributed unobservables at the expense of the requirement of proxy variables for the unobservables. The use of proxy variables in dynamic structural models is related to Cunha, Heckman, and Schennach (2010) and Todd and Wolpin (2012). Since we estimate the parameters of forward-looking structural models, however, we follow a distinct approach outlined as follows.

In the first step, we identify the CCP and the law of state transition using a proxy variable. For this step, we use an approach related to the closed-estimator of Schennach (2004) and Hu and Sasaki (2015) for nonparametric regression models with measurement errors (cf. Li, 2002), as well as the deconvolution methods (Li and Vuong, 1998; Bonhomme and Robin, 2010). In the second step, the CCP-based method (Hotz, Miller, Sanders and Smith, 1994) is applied to the preliminary non-/semi-parametric estimates of the Markov components to obtain structural parameters of a current-time payoff in a simple closed-form expression. Because of its closed form, our estimator is practical and is free from common implementation problems of convergence and numerical global optimization.

First, an informal overview and a practical guideline of our methodology are presented in Section 2. Sections 3 and 4 present formal identification and estimation results. In Section 5, we present an empirical illustration. Section 6 summarizes the paper. A mathematical proof is included in the appendix. Large sample theories, Monte Carlo simulations, extended identification results, and other auxiliary materials are included in the supplementary note that is attached to this paper.

2 An Overview of the Methodology

In this section, we present a practical guideline of our methodology in the context of the problem of firms' exit decisions based on unobserved technologies. Formal identification and estimation results follow in Sections 3 and 4.

Let $d_{j,t} = 1$ indicate the decision of a firm to stay in the market, and let $d_{j,t} = 0$ indicate the decision to exit. The firm chooses $d_{j,t}$ given its technological level $x^*_{j,t}$, and based on its knowledge of the law of stochastic motion of $x^*_{j,t}$. Suppose that the technological state $x^*_{j,t}$ of a firm evolves according to the first-order process

$$x^*_{j,t} = \alpha_t + \gamma_t x^*_{j,t-1} + \eta_{j,t}. \tag{2.1}$$

A firm with its technological level $x^*_{j,t}$ is assumed to receive the current payoff of the affine form $\theta_0 + \theta_1 x^*_{j,t} + \omega^d_{j,t}$ if it is in the market, where $\omega^d_{j,t}$ is the choice-specific private shock.[1] On the other hand, the firm receives zero payoff if it is not in the market. Upon exit from the market, the firm may receive a one-time exit value θ_2, but they will not come back once exited. With this setting, the choice-specific value of the technological state $x^*_{j,t}$ can be written as

$$\begin{aligned} \text{With stay } (d_{j,t} = 1): \quad v_1(x^*_{j,t}) &= \theta_0 + \theta_1 x^*_{j,t} + \omega^1_{j,t} + \mathrm{E}\left[\rho V(x^*_{j,t+1}; \theta) \mid x^*_{j,t}\right] \\ \text{With exit } (d_{j,t} = 0): \quad v_0(x^*_{j,t}) &= \theta_0 + \theta_1 x^*_{j,t} + \theta_2 + \omega^0_{j,t} \end{aligned}$$

where $\rho \in (0,1)$ is the rate of time preference, $V(\,\cdot\,;\theta)$ is the value function, and the conditional expectation $\mathrm{E}[\,\cdot \mid x^*_{j,t}]$ is computed based on the the knowledge of the law (2.1) *including* the distribution of $\eta_{j,t}$.

The fist step toward estimation of the structural parameters is to find a proxy variable $x_{j,t}$ for the unobserved technology $x^*_{j,t}$ with a classical error $\varepsilon_{j,t}$, i.e., $x_{j,t} = x^*_{j,t} + \varepsilon_{j,t}$.

The second step is to estimate the parameters (α_t, γ_t) of the dynamic process (2.1) by the

[1] For continuous state variables, a more generic structure would be non-parametric, but we consider this parametric form in order to focus on the difficulty related to the unobservability of the state variable. Extending the model to a non-parametric one would entail the integral equation of second kind, the identification of which is developed by Srisuma and Linton (2012).

method-of-moment approach, e.g.,

$$\begin{bmatrix} \hat{\alpha}_t \\ \\ \hat{\gamma}_t \end{bmatrix} = \begin{bmatrix} 1 & \frac{\sum_{j=1}^N x_{j,t-1}\mathbb{1}\{d_{j,t-1}=1\}}{\sum_{j=1}^N \mathbb{1}\{d_{j,t-1}=1\}} \\ \frac{\sum_{j=1}^N w_{j,t-1}\mathbb{1}\{d_{j,t-1}=1\}}{\sum_{j=1}^N \mathbb{1}\{d_{j,t-1}=1\}} & \frac{\sum_{j=1}^N x_{j,t-1}w_{j,t-1}\mathbb{1}\{d_{j,t-1}=1\}}{\sum_{j=1}^N \mathbb{1}\{d_{j,t-1}=1\}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\sum_{j=1}^N x_{j,t}\mathbb{1}\{d_{j,t-1}=1\}}{\sum_{j=1}^N \mathbb{1}\{d_{j,t-1}=1\}} \\ \frac{\sum_{j=1}^N x_{j,t}w_{j,t-1}\mathbb{1}\{d_{j,t-1}=1\}}{\sum_{j=1}^N \mathbb{1}\{d_{j,t-1}=1\}} \end{bmatrix}$$

where $w_{j,t-1}$ is some observed variable that is correlated with $x^*_{j,t-1}$, but uncorrelated with the current technological shock $\eta_{j,t}$ and the idiosyncratic shocks $(\varepsilon_{j,t}, \varepsilon_{j,t-1})$. Examples include lags of the proxy, $x_{j,t-2}$. Note that the proxy $x_{j,t}$ as well as $w_{j,t}$ and the choice $d_{j,t}$ are observed, provided that the firm stays in the market. Because of the interaction with the indicator $\mathbb{1}\{d_{j,t-1} = 1\}$, all the sample moments in the above display are computable from observed data.

Having obtained $(\hat{\alpha}_t, \hat{\gamma}_t)$, the third step is to identify the distribution of the idiosyncratic shocks $\varepsilon_{j,t}$. Applying the deconvolution method presented by the references listed in the introduction, we can estimate its characteristic function by the formula

$$\hat{\phi}_{\varepsilon_t}(s) \;=\; \frac{\frac{\sum_{j=1}^N e^{isx_{j,t}}\cdot\mathbb{1}\{d_{j,t}=1\}}{\sum_{j=1}^N \mathbb{1}\{d_{j,t}=1\}}}{\exp\left[\int_0^s \frac{i\cdot\sum_{j=1}^N (x_{j,t+1}-\hat{\alpha}_t)\cdot e^{is'x_{j,t}}\cdot\mathbb{1}\{d_{j,t}=1\}}{\hat{\gamma}_t\cdot\sum_{j=1}^N e^{is'x_{j,t}}\cdot\mathbb{1}\{d_{j,t}=1\}} ds'\right]}.$$

All the moments in this formula involve only the observed variables $x_{j,t}$, $x_{j,t+1}$ and $d_{j,t}$, as opposed to the unobserved true state $x^*_{j,t}$. Thus, they are computable from observed data. Note also that $\hat{\alpha}_t$ and $\hat{\gamma}_t$ are already obtained in the previous step. Hence the right-hand side of this formula is directly computable.

The fourth step is to estimate the CCP, $\Pr(d_t \mid x^*_t)$, of stay given the current technological state x^*_t. Using the estimated characteristic function $\hat{\phi}_{\varepsilon_t}$ produced in the previous step and then applying Schennach (2004) or Hu and Sasaki (2015), we can estimate the CCP by the formula

$$p_t(\xi) \;:=\; \hat{\Pr}(d_{j,t} = 1 \mid x^*_{j,t} = \xi) \;=\; \frac{\int \left(\sum_{j=1}^N \mathbb{1}\{d_{j,t} = 1\} \cdot e^{is(x_{j,t}-\xi)}\right) \cdot \hat{\phi}_{\varepsilon_{j,t}}(s)^{-1} \cdot \phi_K(sh)ds}{\int \left(\sum_{j=1}^N e^{is(x_{j,t}-\xi)}\right) \cdot \hat{\phi}_{\varepsilon_{j,t}}(s)^{-1} \cdot \phi_K(sh)ds} \tag{2.2}$$

where ϕ_K is the Fourier transform of a kernel function K and h is a bandwidth parameter. A similar remark to the previous ones applies here: since $d_{j,t}$ and $x_{j,t}$ are observed, this CCP estimate is directly computable using observed data, even though the true state $x^*_{j,t}$ is unobserved.

The fifth step is to estimate the state transition law, $f(x^*_{j,t} \mid x^*_{j,t-1})$. Using the previously estimated characteristic function $\hat{\phi}_{\varepsilon_t}$, we can estimate the state transition law by the formula

$$\hat{f}(x^*_{j,t} = \xi_t \mid x^*_{j,t-1} = \xi_{t-1}) \;=\; \frac{1}{2\pi} \int \frac{\hat{\phi}_{\varepsilon_{j,t-1}}(s\gamma_t) \sum_{j=1}^N e^{is(x_{j,t}-\xi_t)} \cdot e^{is(\alpha_t+\gamma_t\xi_{t-1})}}{\hat{\phi}_{\varepsilon_{j,t}}(s) \sum_{j=1}^N e^{is(\alpha_t+\gamma_t x^*_{j,t-1})}} \cdot \phi_K(sh)ds. \quad (2.3)$$

Finally, by applying our estimated CCP (2.2) and our estimated state transition law (2.3) to the CCP-based method of Hotz and Miller (1993), we can now estimate the structural parameters $\theta = (\theta_0, \theta_1, \theta_2)$. Specifically, if we follow the standard assumption that the choice-specific private shocks independently follow the standard Gumbel (Type I Extreme Value) distribution, then we obtain the restriction

$$\ln p_t(x^*_t) - \ln(1 - p_t(x^*_t)) \;=\; \mathrm{E}[\rho V(x^*_{t+1}; \theta) \mid x^*_t] - \theta_2,$$

where the discounted future value can be written in terms of the parameters θ as

$$\begin{aligned}\mathrm{E}[\rho V(x^*_{t+1}; \theta) \mid x^*_t] \;=\; & \mathrm{E}\Bigg[\sum_{s=t+1}^{\infty} \rho^{s-t}\,(\theta_0 + \theta_1 x^*_s + \theta_2(1 - p_s(x^*_s)) + \bar{\omega} \\ & -(1 - p_s(x^*_s))\log(1 - p_s(x^*_s)) - p_s(x^*_s)\log p_s(x^*_s))\left(\prod_{s'=t+1}^{s-1} p_{s'}(x^*_{s'})\right)\Bigg|\, x^*_t\Bigg],\end{aligned}$$

where $\bar{\omega}$ denotes the Euler constant ≈ 0.5772. This conditional expectation can be computed by the state transition law estimated with (2.3), and the CCP $p_t(x^*_t)$ is estimated with (2.2). Hence, with our auxiliary estimates, (2.2) and (2.3), the estimator $\hat{\theta}$ solves the equation

$$\begin{aligned}\ln \hat{p}_t(x^*_t) - \ln(1 - \hat{p}_t(x^*_t)) \;=\; & \widehat{\mathrm{E}}\Bigg[\sum_{s=t+1}^{\infty} \rho^{s-t}\Big(\hat{\theta}_0 + \hat{\theta}_1 x^*_s + \hat{\theta}_2(1 - \hat{p}_s(x^*_s)) + \bar{\omega} \qquad (2.4) \\ -(1 - \hat{p}_s(x^*_s))\log(1 - \hat{p}_s(x^*_s)) - \hat{p}_s(x^*_s)\log \hat{p}_s(x^*_s)) & \left(\prod_{s'=t+1}^{s-1} \hat{p}_{s'}(x^*_{s'})\right)\Bigg|\, x^*_t\Bigg] - \hat{\theta}_2 \qquad \text{for all } x^*_t,\end{aligned}$$

which can be solved for $\hat{\theta}$ in an OLS-like closed form (cf. Motz, Miller, Sanders and Smith, 1994). The practical advantage of the above estimation procedure is that every single formula is provided with an explicit closed-form expression, and hence does not suffer from the common implementation problems of convergence and global optimization.

Given the structural parameters $\theta = (\theta_0, \theta_1, \theta_2)$ estimated through the above procedure, one can conduct counter-factual policy predictions in the usual manner. For example, consider the policy scenario where the exit value of the current period is reduced by rate r at time t, i.e., the exit value becomes $(1-r)\theta_2$. To predict the number of exits under this experimental setting, we can estimate the counter-factual CCP of stay by the formula

$$\hat{p}_t^c(x_t^*; r) \;=\; \frac{\exp\left(\ln \hat{p}_t(x_t^*) - \ln(1-\hat{p}_t(x_t^*)) + r\hat{\theta}_2\right)}{1+\exp\left(\ln \hat{p}_t(x_t^*) - \ln(1-\hat{p}_t(x_t^*)) + r\hat{\theta}_2\right)}.$$

Integrating $\hat{p}_t^c(\,\cdot\,; r)$ over the the unobserved distribution of $x_{j,t}^*$ yields the overall fraction of staying firms, where this unobserved distribution can be in turn estimated by the formula

$$\hat{f}(x_{j,t}^* = \xi_t) \;=\; \frac{1}{2\pi}\int \frac{\sum_{j=1}^{N} e^{is(x_{j,t} - xi_t)}}{N \cdot \hat{\phi}_{\varepsilon_{j,t}(s)}} \cdot \phi_K(sh)ds.$$

In this section, we proposed a practical step-by-step guideline of our proposed method. For ease of exposition, this informal overview of our methodology in the current section focused on a specific economic problem and skipped formal assumptions and formal justifications. Readers who are interested in more details of how we derive this methodology may want to go through Sections 3 and 4, where we provide formal identification and estimation results in a more general class of forward-looking structural models.

3 Markov Components: Identification and Estimation

Our basic notations are fixed as follows. A discrete control variable, taking values in $\{0, 1, \cdots, \bar{d}\}$, is denoted by d_t. For example, it indicates the discrete amounts of lumpy R&D investment, and can take the value of zero which is often observed in empirical panel data for firms. Another example is the binary choice of exit by firms that take into account the future fate of technological progress. An observed state variable is denoted by w_t. It is for example the stock of capital. An unobserved state variable is denoted by x_t^*. It is for example the stock of skills or technologies. Finally, a proxy for x_t^* is denoted by x_t. Throughout this paper, we consider the dynamics of this list of random variables.

3.1 Closed-Form Identification of the Markov Components

Our identification strategy is based on the assumptions listed below.

Assumption 1 (First-Order Markov Process)**.** *The quadruple* $\{d_t, w_t, x_t^*, x_t\}$ *jointly follows a first-order Markov process.*

This Markovian structure is decomposed into four independent modules as follows.

Assumption 2 (Independence)**.** *The Markov kernel can be decomposed as follows.*

$$\begin{aligned} & f\left(d_t, w_t, x_t^*, x_t | d_{t-1}, w_{t-1}, x_{t-1}^*, x_{t-1}\right) \\ = \; & f\left(d_t | w_t, x_t^*\right) f\left(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*\right) f\left(x_t^* | d_{t-1}, w_{t-1}, x_{t-1}^*\right) f\left(x_t | x_t^*\right) \end{aligned}$$

where the four components represent

$$\begin{aligned} f\left(d_t | w_t, x_t^*\right) \quad & \textit{conditional choice probability (CCP)}; \\ f\left(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*\right) \quad & \textit{transition rule for the observed state variable}; \\ f\left(x_t^* | d_{t-1}, w_{t-1}, x_{t-1}^*\right) \quad & \textit{transition rule for the unobserved state variable}; \text{and} \\ f\left(x_t | x_t^*\right) \quad & \textit{proxy model.} \end{aligned}$$

The CCP is the firm's investment or exit decision rule based on the observed capital stocks w_t and the unobserved productivity x_t^* for example. The two transition rules specify how the capital stock w_t and the technology x_t^* co-evolve endogenously with firm's forward-looking decision d_t. The proxy model is a stochastic relation between the true productivity x_t^* and a proxy x_t. Because the state variable x_t^* of interest is unit-less and unobserved, we require a restriction of location- and scale-scale normalization. To this goal, the transition rule for the unobserved state variable and the state-proxy relation are semi-parametrically specified as follows.

Assumption 3 (Semi-Parametric Restrictions on the Unobservables)**.** *The transition rule for the unobserved state variable and the state-proxy relation are semi-parametrically specified by*

$$f\left(x_t^* | d_{t-1}, w_{t-1}, x_{t-1}^*\right): \qquad x_t^* = \alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^* + \eta_t^d \qquad \text{if } d_{t-1} = d \tag{3.1}$$

$$f\left(x_t | x_t^*\right): \qquad x_t = x_t^* + \varepsilon_t \tag{3.2}$$

where ε_t and η_t^d have mean zero for each d, and satisfy

$$\begin{aligned} \varepsilon_t &\perp\!\!\!\perp (\{d_\tau\}_\tau, \{x_\tau^*\}_\tau, \{w_\tau\}_\tau, \{\varepsilon_\tau\}_{\tau \neq t}) && \text{for all } t \\ \eta_t^d &\perp\!\!\!\perp (d_\tau, x_\tau^*, w_\tau) && \text{for all } \tau < t \text{ for all } t. \end{aligned}$$

Remark 1. *The decomposition in Assumption 2 and the functional form for the evolution of x_t^* in addition imply that $\eta_t^d \perp\!\!\!\perp w_t$ for all d and t, which is also used to derive our result.*

In case where we consider the discrete choice d_t of investment decisions for example, it is important that the coefficients, $(\alpha^d, \beta^d, \gamma^d)$, are allowed to depend on the amount d of investments since how much a firm invests will likely affect the dynamics of technological evolution. As such, we allow these parameters to have the d superscripts in (3.1). The semi-parametric model (3.2) of the state-proxy relation specifies the proxy x_t as a measurement of the latent technology x_t^* with a classical error. Since it is often restrictive in applications, we also discuss how to relax this classical-error assumption in Section B.9 in the supplementary note.

By Assumption 3, closed-form identification of the transition rule for x_t^* and the proxy model for x_t^* follows from identification of the parameters $(\alpha^d, \beta^d, \gamma^d)$ for each d and from identification of the nonparametric distributions of the unobservables, ε_t, x_t^* , and η_t^d for each d. We show that identification of the parameters $(\alpha^d, \beta^d, \gamma^d)$ follows from the empirically testable rank condition stated as Assumption 4 below.[2] We also obtain identification of the nonparametric distributions of the unobservables, ε_t, x_t^*, and η_t^d, by deconvolution methods under the regularity condition stated as Assumption 5 below.

Assumption 4 (Testable Rank Condition)**.** *$Pr(d_{t-1} = d) > 0$ and the following matrix is nonsingular for each d.*

$$\begin{bmatrix} 1 & E[w_{t-1} \mid d_{t-1} = d] & E[x_{t-1} \mid d_{t-1} = d] \\ E[w_{t-1} \mid d_{t-1} = d] & E[w_{t-1}^2 \mid d_{t-1} = d] & E[x_{t-1} w_{t-1} \mid d_{t-1} = d] \\ E[w_t \mid d_{t-1} = d] & E[w_{t-1} w_t \mid d_{t-1} = d] & E[x_{t-1} w_t \mid d_{t-1} = d] \end{bmatrix}$$

Assumption 5 (Regularity)**.** *The random variables w_t and x_t^* have bounded conditional first moments given d_t. The conditional characteristic functions of w_t and x_t^* given $d_t = d$ do not vanish on the real line, and is absolutely integrable. The conditional characteristic function of (x_{t-1}^*, w_t) given (d_{t-1}, w_{t-1}) and the conditional characteristic function of x_t^* given w_t are absolutely integrable. Random variables ε_t and η_t^d have bounded first moments and absolutely integrable characteristic functions that do not vanish on the real line.*

The validity of Assumptions 1, 2, and 3 can be discussed with specific economic structures. Assumption 4 is empirically testable as is the common rank condition in generic econometric

[2] This matrix consists of moments estimable at the parametric rate of convergence, and hence the standard rank tests (e.g., Cragg and Donald, 1997; Robin and Smith, 2000; Kleibergen and Paap, 2006) can be used.

contexts. Assumption 5 consists of technical regularity conditions, but are automatically satisfied by common distribution families, such as the normal distributions among others. Under this list of five assumptions, we obtain the following closed-form identification result for the four components of the Markov kernel.

Theorem 1 (Closed-Form Identification)**.** *If Assumptions 1, 2, 3, 4, and 5 are satisfied, then the four components* $f\left(d_t|w_t, x_t^*\right)$*,* $f\left(w_t|d_{t-1}, w_{t-1}, x_{t-1}^*\right)$*,* $f\left(x_t^*|d_{t-1}, w_{t-1}, x_{t-1}^*\right)$*,* $f\left(x_t|x_t^*\right)$ *of the Markov kernel* $f\left(d_t, w_t, x_t^*, x_t|d_{t-1}, w_{t-1}, x_{t-1}^*, x_{t-1}\right)$ *are identified with closed-form formulas.*

A proof is given in Section A.1 in the appendix. While the full closed-form identifying formulas for practitioners' reference are provided in Section B.1 in the supplementary note, we also show them with short-hand notations below for convenience of readers. Let $i := \sqrt{-1}$ denote the unit imaginary number. We introduce the Fourier transform operators $\mathcal{F}$ and $\mathcal{F}_2$ defined by

$$\begin{aligned} \mathcal{F}\phi(\xi) &= \frac{1}{2\pi}\int e^{-is\xi}\phi(s)ds && \text{for all } \phi \in L^1(\mathbb{R}) \text{ and } \xi \in \mathbb{R} \\ \mathcal{F}_2\phi(\xi_1,\xi_2) &= \frac{1}{4\pi^2}\int e^{-is_1\xi_1 - is_2\xi_2}\phi(s_1,s_2)ds_1ds_2 && \text{for all } \phi \in L^1(\mathbb{R}^2) \text{ and } (\xi_1,\xi_2) \in \mathbb{R}^2. \end{aligned}$$

First, with these notations, the CCP (e.g., the conditional probability of choosing the amount d of investment given the capital stock w_t and the technological state x_t^*) is identified in closed form by

$$\Pr\left(d_t = d|w_t, x_t^*\right) = \frac{\mathcal{F}\phi_{(d)x_t^*|w_t}(x_t^*)}{\mathcal{F}\phi_{x_t^*|w_t}(x_t^*)}$$

for each choice $d \in \{0, 1, \cdots, \bar{d}\}$, where $\phi_{(d)x_t^*|w_t}(s)$ and $\phi_{x_t^*|w_t}(s)$ are identified in closed form by

$$\phi_{(d)x_t^*|w_t}(s) = \frac{\mathrm{E}[\mathbb{1}\{d_t = d\} \cdot e^{isx_t} \mid w_t]}{\phi_{\varepsilon_t}(s)} \qquad \text{and} \qquad \phi_{x_t^*|w_t}(s) = \frac{\mathrm{E}[e^{isx_t} \mid w_t]}{\phi_{\varepsilon_t}(s)},$$

respectively, where $\phi_{\varepsilon_t}(s)$ is identified in closed form by

$$\phi_{\varepsilon_t}(s) = \frac{\mathrm{E}[e^{isx_t} \mid d_t = d']}{\exp\left[\int_0^s \frac{\mathrm{E}[i(x_{t+1}-\alpha^{d'}-\beta^{d'}w_t)\cdot e^{is'x_t}|d_t=d']}{\gamma^{d'}\,\mathrm{E}[e^{is'x_t}|d_t=d']}ds'\right]} \tag{3.3}$$

with any choice d'. For this closed form identifying formula, the parameter vector $(\alpha^d, \beta^d, \gamma^d)^T$ is in turn explicitly identified for each d by the matrix composition

$$\begin{bmatrix} 1 & \mathrm{E}[w_{t-1} \mid d_{t-1} = d] & \mathrm{E}[x_{t-1} \mid d_{t-1} = d] \\ \mathrm{E}[w_{t-1} \mid d_{t-1} = d] & \mathrm{E}[w_{t-1}^2 \mid d_{t-1} = d] & \mathrm{E}[x_{t-1}w_{t-1} \mid d_{t-1} = d] \\ \mathrm{E}[w_t \mid d_{t-1} = d] & \mathrm{E}[w_{t-1}w_t \mid d_{t-1} = d] & \mathrm{E}[x_{t-1}w_t \mid d_{t-1} = d] \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{E}[x_t \mid d_{t-1} = d] \\ \mathrm{E}[x_t w_{t-1} \mid d_{t-1} = d] \\ \mathrm{E}[x_t w_t \mid d_{t-1} = d] \end{bmatrix}.$$

Second, the transition rule for the observed state variable w_t (e.g., the law of motion of capital) is identified in closed form by

$$f\left(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*\right) = \frac{\mathcal{F}_2 \phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(x_{t-1}^*, w_t)}{\int \mathcal{F}_2 \phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(x_{t-1}^*, w_t) dw_t},$$

where $\phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}$ is identified in closed form by

$$\phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(s_1, s_2) \;=\; \frac{\mathrm{E}[e^{is_1 x_{t-1} + is_2 w_t} \mid d_{t-1}, w_{t-1}]}{\phi_{\varepsilon_{t-1}}(s_1)}.$$

Third, the transition rule for the unobserved state variable x_t^* (e.g., the evolution of technology) is identified in closed form by

$$f(x_t^* \mid d_{t-1}, w_{t-1}, x_{t-1}^*) = \mathcal{F}\phi_{\eta_t^d}(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*),$$

where $d := d_{t-1}$ for short-hand notation, and $\phi_{\eta_t^d}$ is identified in closed form by

$$\phi_{\eta_t^d}(s) \;=\; \frac{\mathrm{E}[e^{isx_t} \mid d_{t-1} = d] \cdot \phi_{\varepsilon_{t-1}}(s\gamma^d)}{\mathrm{E}[e^{is(\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1})} \mid d_{t-1} = d] \cdot \phi_{\varepsilon_t}(s)}.$$

Lastly, the proxy model for x_t^* (e.g., the distribution of the idiosyncratic shock as the proxy error) is identified in closed form by

$$f(x_t \mid x_t^*) \quad = \quad \mathcal{F}\phi_{\varepsilon_t}(x_t - x_t^*),$$

where $\phi_{\varepsilon_t}(s)$ is identified in closed form by (3.3).

In summary, we obtained the four components of the Markov kernel identified with closed-form expressions written in terms of observed data even though we do not observe the true state variable x_t^*. These identified components can be in turn plugged in to the structural restrictions to estimate relevant parameters for the model of forward-looking agents. We present how this step works in Section 4 – also see Section B.4 in the supplementary note for concrete expressions. Before proceeding with structural estimation, we first show that these identified components of the Markov kernel can be easily estimated by their sample counterparts.

3.2 Closed-Form Estimation of the Markov Components

Using the sample counterparts of the closed-form identifying formulas presented in Section 3.1, we develop straightforward closed-form estimators of the four components of the Markov kernel. Throughout this section, we assume homogeneous dynamics, i.e., time-invariant Markov kernel, for simplicity. This assumption is not crucial, and can be easily removed with minor modifications. Let h_w and h_x denote bandwidth parameters and let ϕ_K denote the Fourier transform of a kernel function K used for the purpose of regularization.

First, the sample-counterpart closed-form estimator of the CCP $f(d_t \mid w_t, x_t^*)$ is given by

$$\hat{\Pr}\left(d_t = d | w_t, x_t^*\right) \quad = \quad \frac{\int e^{-isx_t^*} \cdot \hat{\phi}_{(d)x_t^*|w_t}(s) \cdot \phi_K(sh_x) ds}{\int e^{-isx_t^*} \cdot \hat{\phi}_{x_t^*|w_t}(s) \cdot \phi_K(sh_x) ds}$$

for each choice $d \in \{0, 1, \cdots, \bar{d}\}$, where $\hat{\phi}_{(d)x_t^*|w_t}(s)$ and $\hat{\phi}_{x_t^*|w_t}(s)$ are given by

$$\hat{\phi}_{(d)x_t^*|w_t}(s) = \frac{\sum_{j=1}^N \sum_{t=1}^T \mathbb{1}\{D_{j,t} = d\} \cdot e^{isX_{j,t}} \cdot K\left(\frac{W_{j,t}-w_t}{h_w}\right)}{\hat{\phi}_{\varepsilon_t}(s) \cdot \sum_{j=1}^N \sum_{t=1}^T K\left(\frac{W_{j,t}-w_t}{h_w}\right)} \quad \text{and}$$

$$\hat{\phi}_{x_t^*|w_t}(s) = \frac{\sum_{j=1}^N \sum_{t=1}^T e^{isX_{j,t}} \cdot K\left(\frac{W_{j,t}-w_t}{h_w}\right)}{\hat{\phi}_{\varepsilon_t}(s) \cdot \sum_{j=1}^N \sum_{t=1}^T K\left(\frac{W_{j,t}-w_t}{h_w}\right)},$$

respectively, where $\hat{\phi}_{\varepsilon_t}(s)$ is given with any d' by

$$\hat{\phi}_{\varepsilon_t}(s) = \frac{\sum_{j=1}^N \sum_{t=1}^T e^{isX_{j,t}} \cdot \mathbb{1}\{D_{j,t} = d'\} \Big/ \sum_{j=1}^N \sum_{t=1}^T \mathbb{1}\{D_{j,t} = d'\}}{\exp\left[\int_0^s \frac{i \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'} - \beta^{d'} W_{j,t}) \cdot e^{is' X_{j,t}} \cdot \mathbb{1}\{D_{j,t}=d'\}}{\gamma^{d'} \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is' X_{j,t}} \cdot \mathbb{1}\{D_{j,t}=d'\}} ds'\right]}. \tag{3.4}$$

While the notations may make things appear sophisticated, all these expressions are straightforward sample-counterparts of the corresponding closed-form identifying formulas provided in the previous section. This CCP estimator is derived in a similar manner to Schennach (2004) and Hu and Sasaki (2015). Large sample properties of this CCP estimator can be found in Section B.6 in the supplementary note.

Second, the sample-counterpart closed-form estimator of $f(w_t \mid d_{t-1}, w_{t-1}, x_{t-1}^*)$ is given by

$$\hat{f}\left(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*\right) = \frac{\int\int e^{-s_1 x_{t-1}^* - s_2 w_t} \cdot \hat{\phi}_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(s_1, s_2) \cdot \phi_K(s_1 h_x) \cdot \phi_K(s_2 h_w) ds_1 ds_2}{\int\int\int e^{-s_1 x_{t-1}^* - s_2 w_t} \cdot \hat{\phi}_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(s_1, s_2) \cdot \phi_K(s_1 h_x) \cdot \phi_K(s_2 h_w) ds_1 ds_2 dw_t},$$

where $\hat{\phi}_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}$ is given by

$$\hat{\phi}_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(s_1, s_2) = \frac{\sum_{j=1}^N \sum_{t=2}^T e^{is_1 X_{j,t-1} + is_2 W_{j,t}} \cdot \mathbb{1}\{D_{j,t-1} = d_{t-1}\} \cdot K\left(\frac{W_{j,t-1}-w_{t-1}}{h_w}\right)}{\hat{\phi}_{\varepsilon_{t-1}}(s_1) \cdot \sum_{j=1}^N \sum_{t=2}^T \mathbb{1}\{D_{j,t-1} = d_{t-1}\} \cdot K\left(\frac{W_{j,t-1}-w_{t-1}}{h_w}\right)}.$$

Third, the sample-counterpart closed-form estimator of $f(x_t^* \mid d_{t-1}, w_{t-1}, x_{t-1}^*)$ is given by

$$f(x_t^* \mid d_{t-1}, w_{t-1}, x_{t-1}^*) = \frac{1}{2\pi} \int e^{-is(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*)} \cdot \hat{\phi}_{\eta_t^d}(s) \cdot \phi_K(s h_x) ds,$$

where $d := d_{t-1}$ for short-hand notation, and $\hat{\phi}_{\eta_t^d}$ is given by

$$\hat{\phi}_{\eta_t^d}(s) \;=\; \frac{\hat{\phi}_{\varepsilon_{t-1}}(s\gamma^d) \cdot \sum_{j=1}^N \sum_{t=2}^T e^{isX_{j,t}} \cdot \mathbb{1}\{D_{j,t-1} = d\}}{\hat{\phi}_{\varepsilon_t}(s) \cdot \sum_{j=1}^N \sum_{t=2}^T e^{is(\alpha^d + \beta^d W_{j,t-1} + \gamma^d X_{j,t-1})} \cdot \mathbb{1}\{D_{j,t-1} = d\}}.$$

Lastly, the sample-counterpart closed-form estimator of $f(x_t \mid x_t^*)$ is given by

$$\hat{f}(x_t \mid x_t^*) \;=\; \frac{1}{2\pi} \int e^{-is(x_t - x_t^*)} \cdot \hat{\phi}_{\varepsilon_t}(s) \cdot \phi_K(sh_x) ds,$$

where $\hat{\phi}_{\varepsilon_t}(s)$ is given by (3.4).

In each of the above four closed-form estimators, the choice-dependent parameters $(\alpha^d, \beta^d, \gamma^d)$ are also explicitly estimated by the matrix composition:

$$\begin{bmatrix} 1 & \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} W_{jt} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} X_{jt} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} W_{jt} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} W_{jt}^2 \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} X_{jt} W_{jt} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} W_{j,t+1} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} W_{jt} W_{j,t+1} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} X_{jt} W_{j,t+1} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} \end{bmatrix}^{-1} \times \begin{bmatrix} \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} X_{j,t+1} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} X_{j,t+1} W_{jt} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} X_{j,t+1} W_{j,t+1} \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt}=d\}} \end{bmatrix}.$$

Each element of the above matrix and vector consists of sample moments of observed data. In fact, not only these matrix elements, but also all the expressions in the estimation formulas provided in this section consist of sample moments of observed data. Thus, despite their apparently sophisticated expressions, computation of these estimators is not that difficult.

4 Structural Dynamic Discrete Choice Models

In this section, we focus on a class of concrete structural models of forward-looking economic agents. We apply our earlier auxiliary identification results to obtain closed-form estimation of the structural parameters. Agents observe the current state (w_t, x_t^*), where x_t^* is not observed by econometricians. Recall that we deal with a continuous observed state variable w_t and a continuous unobserved state variable x_t^*, and it is not practically attractive to work with nonparametric current-time payoff functions with respect to these continuous state variables. As such, suppose that agents receive the the current payoff of the affine form

$$\theta_d^0 + \theta_d^w w_t + \theta_d^x x_t^* + \omega_{dt}$$

at time t if they make the choice $d_t = d$ under the state (w_t, x_t^*), where ω_{dt} is a private payoff shock at time t that is associated with the choice of $d_t = d$. We may of course extend this affine payoff function to higher-order polynomials at the cost of increased number of parameters. The closed-form identifiability continues to hold as far as the payoff linear with respect to the parameters. Forward-looking agents sequentially make decisions $\{d_t\}$ so as to maximize the expected discounted sum of payoffs

$$\mathrm{E}_t \left[\sum_{s=t}^{\infty} \rho^{s-t} \left(\theta_{d_s}^0 + \theta_{d_s}^w w_s + \theta_{d_s}^x x_s^* + \omega_{d_s s} \right) \right],$$

where ρ is the rate of time preference. To conduct counterfactual policy predictions, economists estimate these structural parameters, θ_d^0, θ_d^w, and θ_d^x.

For ease of exposition under many notations, let us focus on the case of binary decision, where d_t takes values in $\{0, 1\}$. Since the payoff structure is generally identifiable only up to differences, we normalize one of the intercept parameters to zero, say $\theta_1^0 = 0$. [3] Furthermore, we assume that ω_{dt} is independently distributed according to the Type I Extreme Value Distribution in order to obtain simple closed-form expressions, although this distributional assumption is not essential. Under this setting, an application of Hotz and Miller's (1993) inversion theorem

[3] We may alternatively impose a system of restrictions and augment the least-square estimator following Pesendorfer and Schmidt-Dengler (2007) – see also Sanches, Silva, and Srisuma (2013).

and some calculations yield the restriction

$$\xi(\rho; w_t, x_t^*) = \theta_0^0 \cdot \xi_0^0(\rho; w_t, x_t^*) + \theta_0^w \cdot \xi_0^w(\rho; w_t, x_t^*) + \theta_1^w \cdot \xi_1^w(\rho; w_t, x_t^*) + \theta_0^x \cdot \xi_0^x(\rho; w_t, x_t^*) + \theta_1^x \cdot \xi_1^x(\rho; w_t, x_t^*) \tag{4.1}$$

for all (w_t, x_t^*) for all t, where

$$\begin{aligned} \xi(\rho; w_t, x_t^*) = {} & \ln f(1 \mid w_t, x_t^*) - \ln f(0 \mid w_t, x_t^*) + \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(0 \mid w_s, x_s^*) \cdot \ln f(0 \mid w_s, x_s^*) \mid d_t = 1, w_t, x_t^*\right] + \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(1 \mid w_s, x_s^*) \cdot \ln f(1 \mid w_s, x_s^*) \mid d_t = 1, w_t, x_t^*\right] - \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(0 \mid w_s, x_s^*) \cdot \ln f(0 \mid w_s, x_s^*) \mid d_t = 0, w_t, x_t^*\right] - \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(1 \mid w_s, x_s^*) \cdot \ln f(1 \mid w_s, x_s^*) \mid d_t = 0, w_t, x_t^*\right] \end{aligned} \tag{4.2}$$

$$\begin{aligned} \xi_0^0(\rho; w_t, x_t^*) = {} & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(0 \mid w_s, x_s^*) \mid d_t = 1, w_t, x_t^*\right] - \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(0 \mid w_s, x_s^*) \mid d_t = 0, w_t, x_t^*\right] - 1 \end{aligned} \tag{4.3}$$

$$\begin{aligned} \xi_d^w(\rho; w_t, x_t^*) = {} & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(d \mid w_s, x_s^*) \cdot w_s \mid d_t = 1, w_t, x_t^*\right] - \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(d \mid w_s, x_s^*) \cdot w_s \mid d_t = 0, w_t, x_t^*\right] - (-1)^d \cdot w_t \end{aligned} \tag{4.4}$$

$$\begin{aligned} \xi_d^x(\rho; w_t, x_t^*) = {} & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(d \mid w_s, x_s^*) \cdot x_s^* \mid d_t = 1, w_t, x_t^*\right] - \\ & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f(d \mid w_s, x_s^*) \cdot x_s^* \mid d_t = 0, w_t, x_t^*\right] - (-1)^d \cdot x_t^* \end{aligned} \tag{4.5}$$

for each $d \in \{0, 1\}$. See Section B.2 in the supplementary note for derivation of (4.1)–(4.5).

In the context of their model, Hotz, Miller, Sanders, and Smith (1994) propose to use (4.1) to construct moment restrictions. We adapt this approach to our model with unobserved state variables. To this end, define the function Q by

$$Q(\rho,\theta;w_t,x_t^*) = \xi(\rho;w_t,x_t^*) - \theta_0^0 \cdot \xi_0^0(\rho;w_t,x_t^*) + \theta_0^w \cdot \xi_0^w(\rho;w_t,x_t^*)$$
$$-\theta_1^w \cdot \xi_1^w(\rho;w_t,x_t^*) - \theta_0^x \cdot \xi_0^x(\rho;w_t,x_t^*) - \theta_1^x \cdot \xi_1^x(\rho;w_t,x_t^*)$$

where $\theta = (\theta_0^0, \theta_0^w, \theta_1^w, \theta_0^x, \theta_1^x)'$. From (4.1), we obtain the moment restriction

$$\mathrm{E}[R(\rho,\theta;w_t,x_t^*)'\ Q(\rho,\theta;w_t,x_t^*)] = 0 \tag{4.6}$$

for any list (row vector) of bounded functions $R(\rho,\theta;\ \cdot,\ \cdot)$. This paves the way for GMM estimation of the structural parameters (ρ,θ). Furthermore, if the rate ρ of time preference is not to be estimated (which is indeed the case in many applications in the literature),[4] then the moment restriction (4.6) can even be written linearly with respect to the structural parameters θ by defining the function R by

$$R(\rho;w_t,x_t^*) = [\xi_0^0(\rho;w_t,x_t^*),\ \xi_0^w(\rho;w_t,x_t^*),\ \xi_1^w(\rho;w_t,x_t^*),\ \xi_0^x(\rho;w_t,x_t^*),\ \xi_1^x(\rho;w_t,x_t^*)].$$

(Note that we can drop the argument θ from this function since none of the right-hand-side components depends on θ.) In this case, the moment restriction (4.6) yields the structural parameters θ by the OLS-like closed-form expression

$$\theta \ = \ \mathrm{E}\left[R(\rho;w_t,x_t^*)'\ R(\rho;w_t,x_t^*)\right]^{-1} \mathrm{E}\left[R(\rho;w_t,x_t^*)'\ \xi(\rho;w_t,x_t^*)\right], \tag{4.7}$$

provided that the following condition is satisfied.

Assumption 6 (Rank Condition)**.** *$E[R(\rho;w_t,x_t^*)'\ R(\rho;w_t,x_t^*)]$ is nonsingular.*

[4]This rate is generally non-identifiable together with the payoffs (Rust, 1994; Magnac and Thesmar, 2002).

While this result is indeed encouraging, an important remark is in order. Since the generated random variables $R(\rho; w_t, x_t^*)$ and $\xi(\rho; w_t, x_t^*)$ depend on the unobserved state variables x_t^* and their unobserved dynamics by their definitional equations (4.2)–(4.5), they need to be constructed properly based on observed variables. This issue can be solved by using the components of the Markov kernel identified with closed-form formulas in Section 3.1. Note that the elements of all these generated random variables $R(\rho; w_t, x_t^*)$ and $\xi(\rho; w_t, x_t^*)$ take the form $\mathrm{E}[\zeta(w_s, x_s^*) \mid d_t, w_t, x_t^*]$ of the unobserved conditional expectations for various $s > t$, where $\zeta(w_s, x_s^*)$ consists of the explicitly identified CCP $f(d_s \mid w_s, x_s^*)$ and its interactions with w_s, x_s^*, and the log of itself in the formulas (4.2)–(4.5). We can recover these unobserved components in the following manner. If $s = t + 1$, then

$$\mathrm{E}[\zeta(w_s, x_s^*) \mid d_t, w_t, x_t^*] = \int\int \zeta(w_{t+1}, x_{t+1}^*) \cdot f(w_{t+1} \mid d_t, w_t, x_t^*) \times f(x_{t+1}^* \mid d_t, w_t, x_t^*)\, dw_{t+1} dx_{t+1}^* \tag{4.8}$$

where $f(w_{t+1} \mid d_t, w_t, x_t^*)$ and $f(x_{t+1}^* \mid d_t, w_t, x_t^*)$ are identified with closed-forms formulas in Theorem 1. On the other hand, if $s > t + 1$, then

$$\begin{aligned}\mathrm{E}[\zeta(w_s, x_s^*) \mid d_t, w_t, x_t^*] = & \sum_{d_{t+1}=0}^{1} \cdots \sum_{d_{s-1}=0}^{1} \int \cdots \int \zeta(w_s, x_s^*) \cdot f(w_s \mid d_{s-1}, w_{s-1}, x_{s-1}^*) \times \\ & f(x_s^* \mid d_{s-1}, w_{s-1}, x_{s-1}^*) \cdot \prod_{\tau=t}^{s-2} f(d_{\tau+1} \mid w_\tau, x_\tau^*) \cdot f(w_{\tau+1} \mid d_\tau, w_\tau, x_\tau^*) \times \\ & \cdot f(x_{\tau+1}^* \mid d_\tau, w_\tau, x_\tau^*)\, dw_{t+1} \cdots dw_s\, dx_{t+1}^* \cdots dx_s^*,\end{aligned} \tag{4.9}$$

where $f(d_t \mid w_t, x_t^*)$, $f(w_{t+1} \mid d_t, w_t, x_t^*)$, and $f(x_{t+1}^* \mid d_t, w_t, x_t^*)$ are identified with closed-form formulas in Theorem 1.

In light of the explicit decompositions (4.8) and (4.9), the generated random variables $\xi(\rho; w_t, x_t^*)$ and $R(\rho; w_t, x_t^*) = [\xi_0^0(\rho; w_t, x_t^*),\ \xi_0^w(\rho; w_t, x_t^*),\ \xi_1^w(\rho; w_t, x_t^*),\ \xi_0^x(\rho; w_t, x_t^*),\ \xi_1^x(\rho; w_t, x_t^*)]$ defined in (4.2)–(4.5) are identified with closed-form formulas. Therefore, the structural parameters θ are in turn identified in the closed form (4.7). We summarize this result as the following corollary.

Corollary 1 (Closed-Form Identification of Structural Parameters)**.** *Suppose that Assumptions 1, 2, 3, 4, 5, and 6 are satisfied. Given* ρ*, the structural parameters* θ *are identified in the closed form (4.7), where the generated random variables* $\xi(\rho; w_t, x_t^*)$ *and* $R(\rho; w_t, x_t^*) = [\xi_0^0(\rho; w_t, x_t^*),\ \xi_0^w(\rho; w_t, x_t^*),\ \xi_1^w(\rho; w_t, x_t^*),\ \xi_0^x(\rho; w_t, x_t^*),\ \xi_1^x(\rho; w_t, x_t^*)]$ *which appear in (4.7) are in turn identified with closed-form formulas through Theorem 1, (4.2)–(4.5), (4.8), and (4.9).*

Remark 2. *We have left unspecified the measure with respect to which the expectations in (4.6) and thus in (4.7) are taken. The choice is in fact flexible because the original restriction (4.1) holds point-wise for all* (w_t, x_t^*)*. A natural choice is the distribution of* (w_t, x_t^*)*, but it is unobserved. In Section B.3 in the supplementary note, we propose how to evaluate those expectations with respect to this unobserved distribution of* (w_t, x_t^*) *using observed distribution of* (w_t, x_t) *while, of course, keeping the closed form formulas. We emphasize that one can pick any distribution with which the testable rank condition of Assumption 6 is satisfied.*

The closed-form identifying formula for the structural parameters directly translates into a closed-form estimator by substituting the closed-form estimators of the Markov kernel developed in Section 3.2. In Section B.4 in the supplementary note, we provide a concrete expression for the closed-form estimator of the structural parameters. Due to the consistency of the Markov component estimators (see Section B.6 in the supplementary note), the consistency of the sample-counterpart estimator of the structural parameters also follows by the continuous mapping theorem. However, asymptotic normality does not hold under mild conditions, as it requires among others sufficiently fast convergence rates of the preliminary Markov component estimators, which do not hold in general.[5]

[5]Specifically, super-smooth distributions cause logarithmic rates of convergence – see Fan (1991), Fan and

5 Empirical Illustration

Survival selection of firms based on their unobserved dynamic attributes is a long-lasting interest in the economics literature. Theoretical and empirical studies include Jovanovic (1982), Hopenhayn (1992), Ericson and Pakes (1995), Abbring and Campbell (2004), Asplund and Nocke (2006), and Foster, Haltiwanger and Syverson (2008).

Our proposed method extends the approach of Hotz and Miller by allowing for the model to involve persistent unobserved state variables that are observed by the firms but are not observed by econometricians. In this section, we apply our closed-form identification methods to study the forward-looking structure of firm's decision of exit on unobserved technologies. We follow the model and the methodology presented in Section 2, except that we allow for time-varying levels θ_0 (i.e., time fixed effects) of the current-time payoff in order to reflect idiosyncratic macroeconomic shocks. Closely related to what we do is Foster, Haltiwanger and Syverson (2008), who use the total factor productivity of a production function as the measure of productivity. Our approach differs in that we explicitly distinguish between the persistent productivity component and the idiosyncratic component of the total factor productivity.

Levinsohn and Petrin (2003) estimate the production functions for Chilean firms using plant-level panel data.[6] We use the same data set of an 18-year panel from 1979 to 1996 recording real values in 1980 Pesos. Following Levinsohn and Petrin, we focus on the four largest industries, food products (311), textiles (321), wood products (331) and metals (381). We implement

Truong (1993). Also see Section B.6 in the supplementary note for some details. In previous versions of this draft, we used to propose the asymptotic normality under many strong restrictions. In the current draft, we now desist from doing that due to the potential conflicts among the restrictive assumptions that were hard to check.

[6]See also Olley and Pakes (1996), Ackerberg, Caves and Frazer (2006) and Wooldridge (2009) on related methods and discussions of them.

their method using energy and material as two proxies to estimate the production function. The residual $x_{j,t} := y_{j,t} - b_l l_{j,t} - b_k k_{j,t}$ of the estimated production function based on Levinsohn and Petrin is used as a proxy for the true technology $x^*_{j,t}$ in the sense that $x_{j,t} = x^*_{j,t} + \varepsilon_{j,t}$ holds by construction, where $\varepsilon_{j,t}$ denotes idiosyncratic component of Hicks-neutral shocks.[7]

Table 1 shows a summary of the data and construct proxy values for industry 311 (food products), the largest industry in the data. It shows the tendency that the number of firms decreases over time. The number of exiting firms is displayed for each year. Note that, since there are some entering firms, the difference in the number of firms across adjacent years does not necessarily correspond to the number of exits. However, since entry is much less frequent than exits, we exclusively focus on exit decisions in our study. The last three columns of the table list the mean values of the constructed proxy $x_{j,t}$. The third-to-last column displays mean levels for all the firms in this industry. We can see that the productivities steadily advanced since the late 1980s, a little while after the Chilean recession during the 1982-1983. The second-to-last column displays mean levels among the subsample of firms exiting at the end of the current year. The last column displays mean levels among the subsample of firms surviving into the next year. Comparing these two columns, it is clear that exiting firms overall have lower proxy levels for the production technology. Similar patterns result for the other three industries.

We follow the second and third steps in the practical guideline presented in Section 2 to estimated the parameters in the law of technological growth (2.1) as well as the distribution $f_{\varepsilon_{j,t}}$ of the idiosyncratic shocks. These two auxiliary steps are followed by the fifth step in which the

[7]Along the optimal path of investment, the input $(l_{j,t}, k_{j,t})$ is determined by the technological level $x^*_{j,t}$, and hence our assumption that the current-type payoff of firms are given by a function of $x^*_{j,t}$ is not unnatural, possibly except for the functional form assumption.

				Mean of the Proxy $x_{j,t}$		
Year	# Firms	# Exits	% Exits	All Firms	Exiting Firms	Staying Firms
1980	1322	74	0.056	2.90	2.85	2.90
1981	1253	57	0.046	2.93	2.80	2.93
1982	1191	56	0.047	2.85	2.74	2.85
1983	1157	60	0.052	2.84	2.61	2.85
1984	1152	51	0.044	2.86	2.77	2.86
1985	1157	56	0.048	2.86	2.71	2.87
1986	1105	69	0.062	2.87	2.69	2.89
1987	1110	36	0.032	2.83	2.69	2.83
1988	1120	54	0.048	2.84	2.67	2.85
1989	1086	38	0.035	2.87	2.78	2.87
1990	1082	30	0.028	2.90	2.66	2.91
1991	1097	45	0.041	2.93	2.87	2.93
1992	1122	36	0.032	2.98	2.85	2.99
1993	1118	50	0.045	3.02	3.04	3.02
1994	1106	65	0.059	3.06	3.02	3.06
1995	1098	80	0.073	3.05	2.93	3.06

Table 1: Summary statistics for industry 311 (food products). Since there are entries too, the difference in the number of firms across adjacent years does not correspond to the displayed number of exits. The proxy $x_{j,t}$ for the unobserved technologies is constructed as the residual of the estimated production function. Since the mean of the idiosyncratic shocks $\varepsilon_{j,t}$ is zero, the mean of the proxy $x_{j,t}$ equals the mean of the truth $x^*_{j,t}$, but their distributions differ.

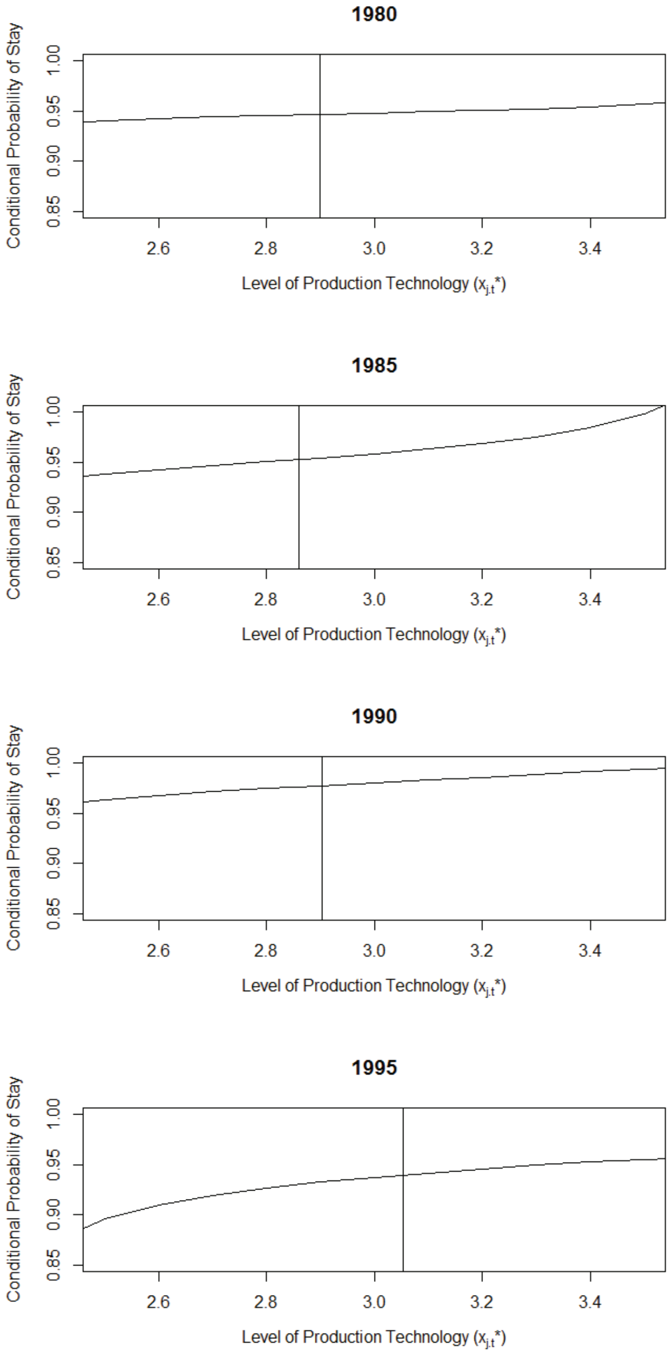

Figure 1: The estimated conditional choice probabilities of stay given the latent levels of production technology, $x^*_{j,t}$, for industry 311 in years 1980, 1985, 1990 and 1995. The vertical lines indicate the mean levels of the unobserved production technology, $x^*_{j,t}$.

conditional choice probability (CCP) of stay, $\Pr(D_{j,t} = 1 \mid x^*_{j,t})$ is estimated by (2.2). Figure 1 illustrates the estimated CCPs for years 1980, 1985, 1990 and 1995. The curves indicate our estimates of the CCPs on the unobserved technological state $x^*_{j,t}$. The probability of stay tends to be higher as the level of production technologies becomes higher. This is consistent with the presumption that firms with lower levels of technologies are more likely to exit. Note also that the levels of the estimated CCPs change across time. This evidence implies that there are some idiosyncratic macroeconomics shocks to the current-time payoffs. As such, it it natural to introduce time-varying intercepts θ_0 (i.e., time-fixed effects) for the payoff parameters when we take these preliminary CCPs estimates to structural estimation. Although the figure shows CCP estimates only for industry 311 (food products), similar remarks apply to the other three industries.

Along with the CCPs, we also estimate the transition kernel for the unobserved technology by (2.3). These two preliminary estimates are taken together to compute the elements in the restriction (2.4), and we estimate the structural parameters with this constructed restriction – see Section 4 for more details about this estimation procedure. The rate ρ of time preference is not to be estimated together with the payoffs given the general non-identification results (Rust, 1994; Magnac and Thesmar, 2002). We thus present estimates of the structural parameters that result under alternative values of $\rho \in \{0.80, 0.90\}$. Table 2 shows our estimates for each of the four industries. The marginal payoff of unit production technology is measured by θ_1. The one-time exit value is measured by θ_2. The magnitudes of these parameter estimates are only relative to the fixed scale of the logistic distribution followed by the difference in private shocks. For scale-normalized views of the structural estimates, we also show the ratio θ_2/θ_1, which measures the value of exit relative to the payoffs produced by each unit of technology. Not surprisingly, these relative exit values vary across alternative rates ρ of time preference.

Industry		Size	ρ	θ_1	θ_2	θ_2/θ_1
311	Food Products	18,276	0.80	1.047	16.491	15.749
321	Textiles	5,039	0.80	1.357	24.772	18.261
331	Wood Products	4,650	0.80	0.596	8.288	13.899
381	Metals	5,286	0.80	1.673	34.273	20.482
311	Food Products	18,276	0.90	0.998	34.553	34.633
321	Textiles	5,039	0.90	0.850	31.637	37.198
331	Wood Products	4,650	0.90	0.550	16.505	29.934
381	Metals	5,286	0.90	1.275	51.636	40.493

Table 2: Estimated structural parameters. The sample size is the number of non-missing entries in the unbalanced panel data used for estimation. The ratio θ_2/θ_1 measures how many units of production technologies are worth the exit value in terms of the current value, and thus indicates the value of exit relative to the payoffs produced by each unit of technology.

Subsample	Size	ρ	θ_1	θ_2	θ_2/θ_1
All firms	18,276	0.90	0.998	34.553	34.633
Real Estate below Average	15,652	0.90	0.905	29.897	33.027
Real Estate above Average	2,604	0.90	0.879	30.815	35.073
Machine & Furniture below Average	15,960	0.90	0.949	29.222	30.778
Machine & Furniture above Average	2,295	0.90	0.572	33.303	58.247

Table 3: Estimated structural parameters by sizes of capital stocks for food product industry (311). The sample size is the number of non-missing entries in the unbalanced panel data used for estimation. The ratio θ_2/θ_1 measures how many units of production technologies are worth the exit value in terms of the current value, and thus indicates the value of exit relative to the payoffs produced by each unit of technology.

However, the rankings of these relative exit values across the industries remain robust across the choice of ρ. Namely, industry 381 (metals) is associated with the largest relative value of exit, followed by industry 321 (textiles) and industry 311 (food products). Industry 331 (wood products) is associated with the smallest relative value of exit. Given that the relative exit value is determined partly by the value of sales and scarp of hard properties relative to the current-time contributory value of technologies, this ranking makes sense. For instance, it is reasonable to find that metals industry running intensively on physical capital exhibit the largest relative value of exit.

The values of exit are supposed to reflect one-time payoffs that firms receive by selling and scrapping capital stocks. In order to understand this feature of relative exit values in more details, we run our structural estimation for various subsets of firms grouped by the sizes of physical capital stocks, focusing on the largest industry (food products, 311). First, we consider the subsamples of firms below and above the average in terms of the amounts of real estate capital stocks. The middle rows of table 3 show structural estimates. Firms with larger stocks of real estate properties exhibit a slightly higher relative value of exit than those with smaller stocks. Second, we consider the subsamples of firms below and above the average in terms of the amounts of stocks of machine, furniture and tools. The bottom rows of table 3 show structural estimates. Firms with larger stocks of machine, furniture and tools exhibit a significantly higher relative value of exit than those with smaller stocks. These observations are consistent with the presumption that forward-looking firms make exit decisions by comparing between the future stream of payoffs attributed to its dynamic production technology and the one-time payoff that results from selling and scrapping physical capital stocks.

6 Summary

In this paper, we show that the structure of forward-looking firms can be identified provided that a proxy for the unobserved state variable is available in data. Our approach works in the following manner.

First, we identify the CCP and the law of state transition using a proxy variable. For this step, we use an approach related to the closed-estimator of Schennach (2004) and Hu and Sasaki (2015) for nonparametric regression models with measurement errors (cf. Li, 2002), as well as the deconvolution methods (Li and Vuong, 1998; Bonhomme and Robin, 2010). Second, the CCP-based method (Hotz, Miller, Sanders and Smith, 1994) is applied to the preliminary non-/semi-parametric estimates of the Markov components to obtain structural parameters of a current-time payoff in a simple closed-form expression.

Applying our methods to firm-level panel data, we analyze the structure of firms making exit decisions by comparing the expected future stream of payoffs attributed to the latent technologies and the exit value that they receive by selling or scrapping physical properties. We find that industries and firms that run intensively on physical capital exhibit greater relative values of exit. In addition, our CCP estimates show that the natural presumption that firms with lower levels of production technologies exit with higher probabilities is true.

References

Abbring, Jaap H. and Jeffrey R. Campbell (2004) "A Firm's First Year," Working Paper.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer (2006) "Structural Identification of Production Functions," Working Paper, UCLA.

Aguirregabiria, Victor, and Pedro Mira (2007) "Sequential Estimation of Dynamic Discrete Games," *Econometrica,* Vol. 75, No. 1, pp. 1-53.

Aguirregabiria, Victor, and Pedro Mira (2013) "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics,* Vol. 156, No. 1, pp. 38–67

Arcidiacono, Peter, and Robert A. Miller (2011) "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica,* Vol. 79, No. 6, pp. 1823–1867.

Asplund, Marcus and Volker Nocke (2006) "Firm Turnover in Imperfectly Competitive Markets," *Review of Economic Studies,* Vol. 73, No. 2, pp. 295–327.

Bonhomme, Stéphane and Jean-Marc Robin (2010) "Generalized Non-Parametric Deconvolution with an Application to Earnings Dynamics," *Review of Economic Studies,* Vol. 77, pp. 491–533.

Cragg, John G. and Stephen G.Donald (1997) "Inferring the Rank of a Matrix," *Journal of Econometrics,* Vol. 76, pp. 223-250.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach (2010) "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica,* Vol. 78, No. 3, pp. 883-931.

Ericson, Richard and Ariel Pakes (1995) "Markov-Perfect Industry Dynamics: a Framework for Empirical Work" *Review of Economic Studies,* Vol. 62, No. 1, pp.53–82.

Fan, Jianqing (1991) "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *Annals of Statistics,* Vol.19, No. 3, pp. 1257–1272.

Fan, Jianqing and Young K. Truong (1993) "Nonparametric Regression with Erros in Variables," *Annals of Statistics,* Vol.21, No. 4, pp. 1900–1925.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008) "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability," *American Economic Review,* Vol. 98, pp. 394–425.

Hopenhayn, Hugo A. (1992) "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica,* Vol. 60, No. 5, pp. 1127–1150.

Hotz, V. Joseph, and Robert A. Miller (1993) "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies,* Vol. 60, No. 3, pp. 497–529.

Hotz, V. Joseph, Robert A. Miller, Seth Sanders, and Jeffrey Smith (1994) “A Simulation Estimator for Dynamic Models of Discrete Choice,” *Review of Economic Studies,* Vol. 61, No. 2, pp. 265–289.

Hu, Yingyao and Yuya Sasaki (2015) “Closed-Form Estimation of Nonparametric Models with Non-Classical Measurement Errors,” *Journal of Econometrics,* Vol. 185, No. 2, pp. 392–408.

Hu, Yingyao, and Mathew Shum (2012) “Nonparametric identification of dynamic models with unobserved state variables,” *Journal of Econometrics,* Vol. 171, No. 1, pp. 32–44.

Jovanovic, Boyan (1982) “Selection and the Evolution of Industry,” *Econometrica,* Vol. 50, No. 3, pp. 649–670.

Kasahara, Hiroyuki, and Katsumi Shimotsu (2009) “Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices,” *Econometrica,* Vol. 77, No. 1, pp. 135-175.

Kleibergen, Frank and Richard Paap (2006) “Generalized Reduced Rank Tests Using the Singular Value Decomposition,” *Journal of Econometrics,* Vol. 133, No. 1, pp. 97–126.

Levinsohn, James, and Amil Petrin (2003) “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies,* Vol. 70, No. 2, pp. 317–342.

Li, Tong (2002) “Robust and Consistent Estimation of Nonlinear Errors-in-Variables Models,” *Journal of Econometrics,* Vol. 110, pp. 1–26.

Li, Tong and Quang Vuong (1998) “Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators,” *Journal of Multivariate Analysis,* Vol. 65, pp. 139–165.

Magnac, Thierry and David Thesmar (2002) “Identifying Dynamic Discrete Decision Processes,” *Econometrica,* Vol. 70, No. 2, pp. 801–816.

Olley, G. Steven., and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica,* Vol. 64, No. 6, pp. 1263–1297.

Pesendorfer, Martin and Philipp Schmidt-Dengler (2008) "Asymptotic Least Squares Estimators for Dynamic Games," *Review of Economic Studies,* Vol. 75, pp. 901–928.

Robin, Jean-Marc and Richard J. Smith (2000) "Tests of Rank," *Econometric Theory,* Vol. 16, No. 2, pp. 151–175.

Rust, John (1994) "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics,* Vol. 4, eds. R. Engle and D. McFadden. Amsterdam: North Holland, pp. 3081-3143.

Sanches, Fabio A.M., Daniel Silva Junior, and Sorawoot Srisuma (2013) "An Alternative Asymptotic Least Squares Estimator for Dynamic Games," Working Paper.

Schennach, Susanne (2004) "Nonparametric Regression in the Presence of Measurement Errors," *Econometric Theory,* Vol. 20, No. 6, pp. 1046–1093.

Srisuma, Sorawoot and Oliver Linton (2012) "Semiparametric Estimation of Markov Decision Processes with Continuous State Space," *Journal of Econometrics,* Vol. 166, No. 2, pp. 320–341.

Todd, Petra E. and Kenneth I. Wolpin (2012) "Estimating a Coordination Game in the Classroom," Working Paper, University of Pennsylvania.

Wooldridge, Jeffrey M. (2009) "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters,* Vol. 104, No. 3, pp. 112–114.

A Appendix

A.1 Proof of Theorem 1

Proof. Our closed-form identification includes four steps.

Step 1: Closed-form identification of the transition rule $f\left(x_t^*|d_{t-1}, w_{t-1}, x_{t-1}^*\right)$**:** First, we show the identification of the parameters and the distributions in transition of x_t^*. Since

$$\begin{aligned} x_t &= x_t^* + \varepsilon_t = \sum_d \mathbb{1}\{d_{t-1} = d\}[\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^* + \eta_t^d] + \varepsilon_t \\ &= \sum_d \mathbb{1}\{d_{t-1} = d\}[\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1} + \eta_t^d - \gamma^d \varepsilon_{t-1}] + \varepsilon_t \end{aligned}$$

we obtain the following equalities for each d:

$$\begin{aligned} \mathrm{E}[x_t \mid d_{t-1} = d] &= \alpha^d + \beta^d \,\mathrm{E}[w_{t-1} \mid d_{t-1} = d] + \gamma^d \,\mathrm{E}[x_{t-1} \mid d_{t-1} = d] \\ &\quad - \mathrm{E}[\gamma^d \varepsilon_{t-1} \mid d_{t-1} = d] + \mathrm{E}[\eta_t^d \mid d_{t-1} = d] + \mathrm{E}[\varepsilon_t \mid d_{t-1} = d] \\ &= \alpha^d + \beta^d \,\mathrm{E}[w_{t-1} \mid d_{t-1} = d] + \gamma^d \,\mathrm{E}[x_{t-1} \mid d_{t-1} = d] \\ \mathrm{E}[x_t w_{t-1} \mid d_{t-1} = d] &= \alpha^d \,\mathrm{E}[w_{t-1} \mid d_{t-1} = d] + \beta^d \,\mathrm{E}[w_{t-1}^2 \mid d_{t-1} = d] + \gamma^d \,\mathrm{E}[x_{t-1} w_{t-1} \mid d_{t-1} = d] \\ &\quad - \mathrm{E}[\gamma^d \varepsilon_{t-1} w_{t-1} \mid d_{t-1} = d] + \mathrm{E}[\eta_t^d w_{t-1} \mid d_{t-1} = d] + \mathrm{E}[\varepsilon_t w_{t-1} \mid d_{t-1} = d] \\ &= \alpha^d \,\mathrm{E}[w_{t-1} \mid d_{t-1} = d] + \beta^d \,\mathrm{E}[w_{t-1}^2 \mid d_{t-1} = d] + \gamma^d \,\mathrm{E}[x_{t-1} w_{t-1} \mid d_{t-1} = d] \\ \mathrm{E}[x_t w_t \mid d_{t-1} = d] &= \alpha^d \,\mathrm{E}[w_t \mid d_{t-1} = d] + \beta^d \,\mathrm{E}[w_{t-1} w_t \mid d_{t-1} = d] + \gamma^d \,\mathrm{E}[x_{t-1} w_t \mid d_{t-1} = d] \\ &\quad - \mathrm{E}[\gamma^d \varepsilon_{t-1} w_t \mid d_{t-1} = d] + \mathrm{E}[\eta_t^d w_t \mid d_{t-1} = d] + \mathrm{E}[\varepsilon_t w_t \mid d_{t-1} = d] \\ &= \alpha^d \,\mathrm{E}[w_t \mid d_{t-1} = d] + \beta^d \,\mathrm{E}[w_{t-1} w_t \mid d_{t-1} = d] + \gamma^d \,\mathrm{E}[x_{t-1} w_t \mid d_{t-1} = d] \end{aligned}$$

by the independence and zero mean assumptions for η_t^d and ε_t. From these, we have the linear equation

$$\begin{bmatrix} \mathrm{E}[x_t \mid d_{t-1}=d] \\ \mathrm{E}[x_t w_{t-1} \mid d_{t-1}=d] \\ \mathrm{E}[x_t w_t \mid d_{t-1}=d] \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{E}[w_{t-1} \mid d_{t-1}=d] & \mathrm{E}[x_{t-1} \mid d_{t-1}=d] \\ \mathrm{E}[w_{t-1} \mid d_{t-1}=d] & \mathrm{E}[w_{t-1}^2 \mid d_{t-1}=d] & \mathrm{E}[x_{t-1} w_{t-1} \mid d_{t-1}=d] \\ \mathrm{E}[w_t \mid d_{t-1}=d] & \mathrm{E}[w_{t-1} w_t \mid d_{t-1}=d] & \mathrm{E}[x_{t-1} w_t \mid d_{t-1}=d] \end{bmatrix} \begin{bmatrix} \alpha^d \\ \beta^d \\ \gamma^d \end{bmatrix}$$

Provided that the matrix on the right-hand side is non-singular, we can identify the parameters $(\alpha^d, \beta^d, \gamma^d)$ by

$$\begin{bmatrix} \alpha^d \\ \beta^d \\ \gamma^d \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{E}[w_{t-1} \mid d_{t-1}=d] & \mathrm{E}[x_{t-1} \mid d_{t-1}=d] \\ \mathrm{E}[w_{t-1} \mid d_{t-1}=d] & \mathrm{E}[w_{t-1}^2 \mid d_{t-1}=d] & \mathrm{E}[x_{t-1} w_{t-1} \mid d_{t-1}=d] \\ \mathrm{E}[w_t \mid d_{t-1}=d] & \mathrm{E}[w_{t-1} w_t \mid d_{t-1}=d] & \mathrm{E}[x_{t-1} w_t \mid d_{t-1}=d] \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{E}[x_t \mid d_{t-1}=d] \\ \mathrm{E}[x_t w_{t-1} \mid d_{t-1}=d] \\ \mathrm{E}[x_t w_t \mid d_{t-1}=d] \end{bmatrix}$$

Next, we show identification of $f(\varepsilon_t)$ and $f\left(\eta_t^d\right)$ for each d. Observe that

$$\begin{aligned} & \mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 x_t\right) | d_{t-1}=d\right] \\ = \ & \mathrm{E}\left[\exp\left(is_1\left(x_{t-1}^* + \varepsilon_{t-1}\right) + is_2\left(\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^* + \eta_t^d + \varepsilon_t\right)\right) | d_{t-1}=d\right] \\ = \ & \mathrm{E}\left[\exp\left(i\left(s_1 x_{t-1}^* + s_2\alpha^d + s_2\beta^d w_{t-1} + s_2\gamma^d x_{t-1}^*\right)\right) | d_{t-1}=d\right] \\ & \times \mathrm{E}\left[\exp\left(is_1\varepsilon_{t-1}\right)\right] \mathrm{E}\left[\exp\left(is_2\left(\eta_t^d + \varepsilon_t\right)\right)\right] \end{aligned}$$

follows from the independence assumptions for η_t^d and ε_t. Taking the derivative with respect to s_2 yields

$$\begin{aligned} & \left[\frac{\partial}{\partial s_2} \ln \mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 x_t\right) | d_{t-1}=d\right]\right]_{s_2=0} \\ = \ & \frac{\mathrm{E}\left[i(\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^*) \exp\left(is_1 x_{t-1}^*\right) | d_{t-1}=d\right]}{\mathrm{E}\left[\exp\left(is_1 x_{t-1}^*\right) | d_{t-1}=d\right]} \\ = \ & i\alpha^d + \beta^d \frac{\mathrm{E}[iw_{t-1}\exp(is_1 x_{t-1}^*) \mid d_{t-1}=d]}{\mathrm{E}[\exp(is_1 x_{t-1}^*) \mid d_{t-1}=d]} + \gamma^d \frac{\partial}{\partial s_1} \ln \mathrm{E}\left[\exp\left(is_1 x_{t-1}^*\right) | d_{t-1}=d\right] \\ = \ & i\alpha^d + \beta^d \frac{\mathrm{E}[iw_{t-1}\exp(is_1 x_{t-1}) \mid d_{t-1}=d]}{\mathrm{E}[\exp(is_1 x_{t-1}) \mid d_{t-1}=d]} + \gamma^d \frac{\partial}{\partial s_1} \ln \mathrm{E}\left[\exp\left(is_1 x_{t-1}^*\right) | d_{t-1}=d\right] \end{aligned}$$

where the switch of the differential and integral operators is permissible provided that there exists $h \in L^1(F_{w_{t-1}x_{t-1}^*|d_{t-1}=d})$ such that $\left|i(\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^*) \exp\left(is_1 x_{t-1}^*\right)\right| < h(w_{t-1}, x_{t-1}^*)$

holds for all (w_{t-1}, x^*_{t-1}), which follows from the bounded conditional moment given in Assumption 5, and the denominators are nonzero as the conditional characteristic function of x^*_t given d_t does not vanish on the real line under Assumption 5. Therefore,

$$\begin{aligned} \mathrm{E}\left[\exp\left(isx^*_{t-1}\right)|d_{t-1}=d\right] &= \exp\left[\int_0^s \left[\frac{1}{\gamma^d}\frac{\partial}{\partial s_2}\ln \mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2x_t\right)|d_{t-1}=d\right]\right]_{s_2=0} ds_1 \right. \\ &\quad \left. -\int_0^s \frac{i\alpha^d}{\gamma^d}ds_1 - \int_0^s \frac{\beta^d}{\gamma^d}\frac{\mathrm{E}[iw_{t-1}\exp(is_1x_{t-1}) \mid d_{t-1}=d]}{\mathrm{E}[\exp(is_1x_{t-1}) \mid d_{t-1}=d]}ds_1\right] \\ &= \exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_t-\alpha^d-\beta^d w_{t-1})\exp\left(is_1x_{t-1}\right)|d_{t-1}=d\right]}{\gamma^d\,\mathrm{E}\left[\exp\left(is_1x_{t-1}\right)|d_{t-1}=d\right]}ds_1\right]. \end{aligned}$$

From the proxy model and the independence assumption for ε_t,

$$\mathrm{E}\left[\exp\left(isx_{t-1}\right)|d_{t-1}=d\right] = \mathrm{E}\left[\exp\left(isx^*_{t-1}\right)|d_{t-1}=d\right]\mathrm{E}\left[\exp\left(is\varepsilon_{t-1}\right)\right].$$

We then obtain the following result using any d.

$$\begin{aligned} \mathrm{E}\left[\exp\left(is\varepsilon_{t-1}\right)\right] &= \frac{\mathrm{E}\left[\exp\left(isx_{t-1}\right)|d_{t-1}=d\right]}{\mathrm{E}\left[\exp\left(isx^*_{t-1}\right)|d_{t-1}=d\right]} \\ &= \frac{\mathrm{E}\left[\exp\left(isx_{t-1}\right)|d_{t-1}=d\right]}{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_t-\alpha^d-\beta^d w_{t-1})\exp(is_1x_{t-1})|d_{t-1}=d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1x_{t-1})|d_{t-1}=d]}ds_1\right]}. \end{aligned}$$

This argument holds for all t so that we can identify $f\left(\varepsilon_t\right)$ with

$$\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right] = \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_t=d\right]}{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^d-\beta^d w_t)\exp(is_1x_t)|d_t=d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1x_t)|d_t=d]}ds_1\right]} \tag{A.1}$$

using any d.

In order to identify $f\left(\eta^d_t\right)$ for each d, consider

$$x_t + \gamma^d\varepsilon_{t-1} = \alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1} + \varepsilon_t + \eta^d,$$

and thus

$$\begin{aligned} \mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1}=d\right]\mathrm{E}\left[\exp\left(is\gamma^d\varepsilon_{t-1}\right)\right] &= \mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d x_{t-1})\right)|d_{t-1}=d\right] \\ &\quad \times \mathrm{E}\left[\exp\left(is\eta^d_t\right)\right]\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right] \end{aligned}$$

follows by the independence assumptions for η_t^d and ε_t. Therefore, by the formula (A.1), the characteristic function of η_t^d can be expressed by

$$\begin{aligned}
\mathrm{E}\left[\exp\left(is\eta_t^d\right)\right] &= \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1}=d\right]\cdot\mathrm{E}\left[\exp\left(is\gamma^d\varepsilon_{t-1}\right)\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d x_{t-1})\right)|d_{t-1}=d\right]\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right]} \\
&= \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1}=d\right]\cdot\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^d-\beta^d w_t)\exp(is_1 x_t)|d_t=d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1 x_t)|d_t=d]}ds_1\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d x_{t-1})\right)|d_{t-1}=d\right]\cdot\mathrm{E}\left[\exp\left(isx_t\right)|d_t=d\right]}\times \\
&\quad \frac{\mathrm{E}\left[\exp\left(is\gamma^d x_{t-1}\right)|d_{t-1}=d\right]}{\exp\left[\int_0^{s\gamma^d}\frac{\mathrm{E}\left[i(x_t-\alpha^d-\beta^d w_{t-1})\exp(is_1 x_{t-1})|d_{t-1}=d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1 x_{t-1})|d_{t-1}=d]}ds_1\right]}.
\end{aligned}$$

The denominator on the right-hand side is non-zero, as the conditional and unconditional characteristic functions do not vanish on the real line under Assumption 5. Letting $\mathcal{F}$ denote the operator defined by

$$\left(\mathcal{F}\phi\right)(\xi)=\frac{1}{2\pi}\int e^{-is\xi}\phi(s)ds \qquad \text{for all } \phi\in L^1(\mathbb{R}) \text{ and } \xi\in\mathbb{R},$$

we identify $f_{\eta_t^d}$ by

$$f_{\eta_t^d}(\eta)=\left(\mathcal{F}\phi_{\eta_t^d}\right)(\eta) \qquad \text{for all } \eta,$$

where the characteristic function $\phi_{\eta_t^d}$ is given by

$$\begin{aligned}
\phi_{\eta_t^d}(s) &= \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1}=d\right]\cdot\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^d-\beta^d w_t)\exp(is_1 x_t)|d_t=d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1 x_t)|d_t=d]}ds_1\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d x_{t-1})\right)|d_{t-1}=d\right]\cdot\mathrm{E}\left[\exp\left(isx_t\right)|d_t=d\right]}\times \\
&\quad \frac{\mathrm{E}\left[\exp\left(is\gamma^d x_{t-1}\right)|d_{t-1}=d\right]}{\exp\left[\int_0^{s\gamma^d}\frac{\mathrm{E}\left[i(x_t-\alpha^d-\beta^d w_{t-1})\exp(is_1 x_{t-1})|d_{t-1}=d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1 x_{t-1})|d_{t-1}=d]}ds_1\right]}.
\end{aligned}$$

We can use this identified density in turn to identify the transition rule $f\left(x_t^*|d_{t-1},w_{t-1},x_{t-1}^*\right)$ with

$$f\left(x_t^*|d_{t-1},x_{t-1},x_{t-1}^*\right)=\sum_d \mathbb{1}\{d_{t-1}=d\}f_{\eta_t^d}\left(x_t^*-\alpha^d-\beta^d w_{t-1}-\gamma^d x_{t-1}^*\right).$$

In summary, we obtain the closed-form expression

$$
\begin{aligned}
f(x_t^* \mid d_{t-1}, w_{t-1}, x_{t-1}^*) &= \sum_d \mathbb{1}\{d_{t-1} = d\} \left(\mathcal{F}\phi_{\eta_t^d}\right)(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*) \\
&= \sum_d \frac{\mathbb{1}\{d_{t-1} = d\}}{2\pi} \int \exp\left(-is(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*)\right) \times \\
&\quad \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1} = d\right] \cdot \exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1} - \alpha^{d'} - \beta^{d'} w_t)\exp(is_1 x_t)|d_t = d'\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1 x_t)|d_t = d']} ds_1\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1})\right)|d_{t-1} = d\right] \cdot \mathrm{E}\left[\exp\left(isx_t\right)|d_t = d\right]} \times \\
&\quad \frac{\mathrm{E}\left[\exp\left(is\gamma^d x_{t-1}\right)|d_{t-1} = d'\right]}{\exp\left[\int_0^{s\gamma^d} \frac{\mathrm{E}\left[i(x_t - \alpha^{d'} - \beta^{d'} w_{t-1})\exp(is_1 x_{t-1})|d_{t-1} = d'\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1 x_{t-1})|d_{t-1} = d']} ds_1\right]} ds.
\end{aligned}
$$

using any d'. This completes Step 1.

Step 2: Closed-form identification of the proxy model $f\left(x_t \mid x_t^*\right)$: Given (A.1), we can write the density of ε_t by

$$
f_{\varepsilon_t}(\varepsilon) = \left(\mathcal{F}\phi_{\varepsilon_t}\right)(\varepsilon) \qquad \text{for all } \varepsilon,
$$

where the characteristic function ϕ_{ε_t} is defined by (A.1) as

$$
\phi_{\varepsilon_t}(s) = \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_t = d\right]}{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1} - \alpha^d - \beta^d w_t)\exp(is' x_t)|d_t = d\right]}{\gamma^d\,\mathrm{E}[\exp(is' x_t)|d_t = d]} ds'\right]}.
$$

Provided this identified density of ε_t, we nonparametrically identify the proxy model

$$
f(x_t \mid x_t^*) = f_{\varepsilon_t}(x_t - x_t^*)
$$

In summary, we obtain the closed-form expression

$$
\begin{aligned}
f(x_t \mid x_t^*) &= \left(\mathcal{F}\phi_{\varepsilon_t}\right)(x_t - x_t^*) \\
&= \frac{1}{2\pi} \int \frac{\exp\left(-is(x_t - x_t^*)\right) \cdot \mathrm{E}\left[\exp\left(isx_t\right)|d_t = d\right]}{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1} - \alpha^d - \beta^d w_t)\exp(is_1 x_t)|d_t = d\right]}{\gamma^d\,\mathrm{E}[\exp(is_1 x_t)|d_t = d]} ds_1\right]} ds
\end{aligned}
$$

using any d. This completes Step 2.

Step 3: Closed-form identification of the transition rule $f\left(w_t|d_{t-1},w_{t-1},x^*_{t-1}\right)$: Consider the joint density expressed by the convolution integral

$$f\left(x_{t-1},w_t \mid d_{t-1},w_{t-1}\right)=\int f_{\varepsilon_{t-1}}\left(x_{t-1}-x^*_{t-1}\right)f\left(x^*_{t-1},w_t \mid d_{t-1},w_{t-1}\right)dx^*_{t-1}$$

We can thus obtain a closed-form expression of $f\left(x^*_{t-1},w_t \mid d_{t-1},w_{t-1}\right)$ by the deconvolution. To see this, observe

$$\begin{aligned}\mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right] &= \mathrm{E}\left[\exp\left(is_1x^*_{t-1}+is_1\varepsilon_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right]\\ &= \mathrm{E}\left[\exp\left(is_1x^*_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right]\mathrm{E}\left[\exp\left(is_1\varepsilon_{t-1}\right)\right]\end{aligned}$$

by the independence assumption for ε_t, and so

$$\begin{aligned}&\mathrm{E}\left[\exp\left(is_1x^*_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right]=\frac{\mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right]}{\mathrm{E}\left[\exp\left(is_1\varepsilon_{t-1}\right)\right]}\\ =\;&\frac{\mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right]\cdot\exp\left[\int_0^{s_1}\frac{\mathrm{E}\left[i(x_t-\alpha^d-\beta^dw_{t-1})\exp\left(is'_1x_{t-1}\right)|d_{t-1}=d\right]}{\gamma^d\,\mathrm{E}\left[\exp\left(is'_1x_{t-1}\right)|d_{t-1}=d\right]}ds'_1\right]}{\mathrm{E}\left[\exp\left(is_1x_{t-1}\right)|d_{t-1}=d\right]}\end{aligned}$$

follows. Letting $\mathcal{F}_2$ denote the operator defined by

$$\left(\mathcal{F}_2\phi\right)(\xi_1,\xi_2)=\frac{1}{4\pi^2}\int\int e^{-is_1\xi_1-is_2\xi_2}\phi(s_1,s_2)ds_1ds_2 \qquad \text{for all } \phi\in L^1(\mathbb{R}^2) \text{ and } (\xi_1,\xi_2)\in\mathbb{R}^2,$$

we can express the conditional density as

$$f\left(x^*_{t-1},w_t|d_{t-1},w_{t-1}\right)=\left(\mathcal{F}_2\phi_{x^*_{t-1},w_t|d_{t-1},w_{t-1}}\right)\left(w_t,x^*_{t-1}\right)$$

where the characteristic function is defined by

$$\begin{aligned}&\phi_{x^*_{t-1},w_t|d_{t-1},w_{t-1}}(s_1,s_2)\\ =\;&\frac{\mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2w_t\right)|d_{t-1},w_{t-1}\right]\cdot\exp\left[\int_0^{s_1}\frac{\mathrm{E}\left[i(x_t-\alpha^d-\beta^dw_{t-1})\exp\left(is'_1x_{t-1}\right)|d_{t-1}=d\right]}{\gamma^d\,\mathrm{E}\left[\exp\left(is'_1x_{t-1}\right)|d_{t-1}=d\right]}ds'_1\right]}{\mathrm{E}\left[\exp\left(is_1x_{t-1}\right)|d_{t-1}=d\right]}\end{aligned}$$

with any d. Using this conditional density, we can nonparametrically identify the transition rule $f(w_t \mid d_{t-1},w_{t-1},x^*_{t-1})$ with

$$f\left(w_t|d_{t-1},w_{t-1},x^*_{t-1}\right)=\frac{f\left(x^*_{t-1},w_t|d_{t-1},w_{t-1}\right)}{\int f\left(x^*_{t-1},w_t|d_{t-1},w_{t-1}\right)dw_t}.$$

In summary, we obtain the closed-form expression

$$f\left(w_t|d_{t-1},w_{t-1},x_{t-1}^*\right) = \frac{\left(\mathcal{F}_2\phi_{x_{t-1}^*,w_t|d_{t-1},w_{t-1}}\right)(x_{t-1}^*,w_t)}{\int\left(\mathcal{F}_2\phi_{x_{t-1}^*,w_t|d_{t-1},w_{t-1}}\right)(x_{t-1}^*,w_t)dw_t}$$

$$= \sum_d \mathbb{1}\{d_{t-1}=d\}\int\int \exp\left(-is_1w_t - is_2x_{t-1}^*\right)\cdot \mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2w_t\right)|d_{t-1}=d,w_{t-1}\right]\times$$
$$\frac{\exp\left[\int_0^{s_1}\frac{\mathrm{E}\left[i(x_t-\alpha^{d'}-\beta^{d'}w_{t-1})\exp\left(is_1'x_{t-1}\right)|d_{t-1}=d'\right]}{\gamma^{d'}\,\mathrm{E}\left[\exp\left(is_1'x_{t-1}\right)|d_{t-1}=d'\right]}ds_1'\right]}{\mathrm{E}\left[\exp\left(is_1x_{t-1}\right)|d_{t-1}=d'\right]}ds_1ds_2 \Bigg/$$
$$\int\int\int \exp\left(-is_1w_t - is_2x_{t-1}^*\right)\cdot \mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2w_t\right)|d_{t-1}=d,w_{t-1}\right]\times$$
$$\frac{\exp\left[\int_0^{s_1}\frac{\mathrm{E}\left[i(x_t-\alpha^{d'}-\beta^{d'}w_{t-1})\exp\left(is_1'x_{t-1}\right)|d_{t-1}=d'\right]}{\gamma^{d'}\,\mathrm{E}\left[\exp\left(is_1'x_{t-1}\right)|d_{t-1}=d'\right]}ds_1'\right]}{\mathrm{E}\left[\exp\left(is_1x_{t-1}\right)|d_{t-1}=d'\right]}ds_1ds_2dw_t$$

using any d'. This completes Step 3.

Step 4: Closed-form identification of the CCP $f\left(d_t|w_t,x_t^*\right)$: Note that we have

$$\begin{aligned}
\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(isx_t\right)|w_t\right] &= \mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(isx_t^*+is\varepsilon_t\right)|w_t\right] \\
&= \mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(isx_t^*\right)|w_t\right]\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*\right]\exp\left(isx_t^*\right)|w_t\right]\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right]
\end{aligned}$$

by the independence assumption for ε_t and the law of iterated expectations. Therefore

$$\begin{aligned}
\frac{\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(isx_t\right)|w_t\right]}{\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right]} &= \mathrm{E}\left[\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*\right]\exp\left(isx_t^*\right)|w_t\right] \\
&= \int \exp\left(isx_t^*\right)\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*\right]f\left(x_t^*|w_t\right)dx_t^*
\end{aligned}$$

This is the Fourier inversion of $\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*\right]f\left(x_t^*|w_t\right)$. On the other hand, the Fourier inversion of $f\left(x_t^*|w_t\right)$ can be found as

$$\mathrm{E}\left[\exp\left(isx_t^*\right)|w_t\right] = \frac{\mathrm{E}\left[\exp\left(isx_t\right)|w_t\right]}{\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right]}.$$

Therefore, we find the closed-form expression for CCP $f\left(d_t|w_t,x_t^*\right)$ as follows.

$$\Pr\left(d_t=d|w_t,x_t^*\right) = \mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*\right] = \frac{\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*\right]f\left(x_t^*|w_t\right)}{f\left(x_t^*|w_t\right)} = \frac{\left(\mathcal{F}\phi_{(d)x_t^*|w_t}\right)(x_t^*)}{\left(\mathcal{F}\phi_{x_t^*|w_t}\right)(x_t^*)}$$

where the characteristic functions are defined by

$$\begin{aligned}\phi_{(d)x_t^*|w_t}(s) &= \frac{\mathrm{E}\left[\mathbb{1}\{d_t = d\}\exp\left(isx_t\right)|w_t\right]}{\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right]} \\ &= \frac{\mathrm{E}\left[\mathbb{1}\{d_t = d\}\exp\left(isx_t\right)|w_t\right]\cdot\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^{d'}-\beta^{d'}w_t)\exp(is_1x_t)|d_t=d'\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1x_t)|d_t=d']}ds_1\right]}{\mathrm{E}\left[\exp\left(isx_t\right)|d_t = d'\right]}\end{aligned}$$

and

$$\begin{aligned}\phi_{x_t^*|w_t}(s) &= \frac{\mathrm{E}\left[\exp\left(isx_t\right)|w_t\right]}{\mathrm{E}\left[\exp\left(is\varepsilon_t\right)\right]} \\ &= \frac{\mathrm{E}\left[\exp\left(isx_t\right)|w_t\right]\cdot\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^{d'}-\beta^{d'}w_t)\exp(is_1x_t)|d_t=d'\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1x_t)|d_t=d']}ds_1\right]}{\mathrm{E}\left[\exp\left(isx_t\right)|d_t = d'\right]}\end{aligned}$$

by (A.1) using any d'. In summary, we obtain the closed-form expression

$$\begin{aligned}\Pr\left(d_t = d|w_t, x_t^*\right) &= \frac{\left(\mathcal{F}\phi_{(d)x_t^*|w_t}\right)(x_t^*)}{\left(\mathcal{F}\phi_{x_t^*|w_t}\right)(x_t^*)} \\ &= \int \exp\left(-isx_t^*\right)\cdot\mathrm{E}\left[\mathbb{1}\{d_t = d\}\exp\left(isx_t\right)|w_t\right]\times \\ &\quad \frac{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^{d'}-\beta^{d'}w_t)\exp(is_1x_t)|d_t=d'\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1x_t)|d_t=d']}ds_1\right]}{\mathrm{E}\left[\exp\left(isx_t\right)|d_t = d'\right]}ds \Bigg/ \\ &\quad \int \exp\left(-isx_t^*\right)\cdot\mathrm{E}\left[\exp\left(isx_t\right)|w_t\right]\times \\ &\quad \frac{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1}-\alpha^{d'}-\beta^{d'}w_t)\exp(is_1x_t)|d_t=d'\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1x_t)|d_t=d']}ds_1\right]}{\mathrm{E}\left[\exp\left(isx_t\right)|d_t = d'\right]}ds\end{aligned}$$

using any d'. This completes Step 4. □

Supplementary Note for "Closed-Form Identification of Dynamic Discrete Choice Models with Proxies for Unobserved State Variables"

Yingyao Hu and Yuya Sasaki

B Supplementary Appendix

B.1 The Full Closed-Form Estimator

Let $\widehat{\phi}_{x_t^*|d_t=d}$ denote the sample-counterpart estimator of the conditional characteristic function $\phi_{x_t^*|d_t=d}$, defined by

$$\widehat{\phi}_{x_t^*|d_t=d}(s) = \exp\left[\int_0^s \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} i(X_{j,t+1} - \alpha^d - \beta^d W_{jt}) \cdot \exp(is_1 X_{jt}) \cdot \mathbb{1}\{D_{jt} = d\}}{\gamma^d \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{jt}) \cdot \mathbb{1}\{D_{jt} = d\}} ds_1\right].$$

The closed-form estimator of the CCP, $f(d_t \mid w_t, x_t^*)$, is given by

$$\begin{aligned}\widehat{f}(d|w,x^*) \ = \ & \int \exp(-isx^*) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} K\left(\frac{W_{jt}-w}{h_w}\right)} \times \\ & \widehat{\phi}_{x_t^*|d_t=d'}(s) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x)\ ds \Bigg/ \\ & \int \exp(-isx^*) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} K\left(\frac{W_{jt}-w}{h_w}\right)} \times \\ & \widehat{\phi}_{x_t^*|d_t=d'}(s) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x)\ ds \qquad \text{(B.1)}\end{aligned}$$

with any d', where h_w denotes a bandwidth parameter and ϕ_K denotes the Fourier transform of a kernel function K used for the purpose of regularization. The closed-form estimator of the transition rule, $f(w_t \mid d_{t-1}, w_{t-1}, x^*_{t-1})$, for the observed state variable w_t is given by

$$\widehat{f}\,(w'^*) = \int\int \exp\left(-is_1 w' - is_2 x^*\right) \times$$
$$\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is_1 X_{jt} + is_2 W_{j,t+1}\right) \cdot \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)} \cdot \widehat{\phi}_{x^*_t|d_t=d'}(s_1) \times$$
$$\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is_1 X_{jt}\right) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(s_1 h_w) \cdot \phi_K(s_2 h_x)\ ds_1 ds_2 \Bigg/$$
$$\int\int\int \exp\left(-is_1 w'' - is_2 x^*\right) \times$$
$$\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is_1 X_{jt} + is_2 W_{j,t+1}\right) \cdot \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)} \cdot \widehat{\phi}_{x^*_t|d_t=d'}(s_1) \times$$
$$\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is_1 X_{jt}\right) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(s_1 h_w) \cdot \phi_K(s_2 h_x)\ ds_1 ds_2 dw'' \quad \text{(B.2)}$$

with any d'. The closed-form estimator of the transition rule, $f(x^*_t \mid d_{t-1}, w_{t-1}, x^*_{t-1})$, for the unobserved state variable x^*_t is given by

$$\widehat{f}(x^{*\prime *}) = \frac{1}{2\pi}\int \exp\left(-is(x^{*\prime d} - \beta^d w - \gamma^d x^*)\right) \times$$
$$\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is X_{j,t+1}\right) \cdot \mathbb{1}\{D_{jt} = d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is(\alpha^d + \beta^d W_{jt} + \gamma^d X_{jt})\right) \cdot \mathbb{1}\{D_{jt} = d\}} \times$$
$$\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is\gamma^d X_{jt}\right) \cdot \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is X_{jt}\right) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \frac{\widehat{\phi}_{x^*_t|d_t=d'}(s)}{\widehat{\phi}_{x^*_t|d_t=d'}(s\gamma^d)}\ \phi_K(s h_x)\ ds \quad \text{(B.3)}$$

with any d'. Finally, the the closed-form estimator of the proxy model, $f(x_t \mid x^*_t)$, is given by

$$\widehat{f}(x \mid x^*) = \frac{1}{2\pi}\int \exp\left(-is(x - x^*)\right) \cdot \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \exp\left(is X_{jt}\right) \cdot \mathbb{1}\{D_{jt} = d'\}}{\widehat{\phi}_{x^*_t|d_t=d'}(s) \cdot \sum_{j=1}^{N}\sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(s h_x)\ ds \quad \text{(B.4)}$$

using any d'.

In each of the above four closed-form estimators, the parameters $(\alpha^d, \beta^d, \gamma^d)$ for each d are also explicitly estimated by the matrix composition:

$$\begin{bmatrix} 1 & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} W_{jt}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} X_{jt}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} W_{jt}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} W_{jt}^2\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} X_{jt}W_{jt}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} W_{j,t+1}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} W_{jt}W_{j,t+1}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} X_{jt}W_{j,t+1}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} \end{bmatrix}^{-1} \times \begin{bmatrix} \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} X_{j,t+1}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} X_{j,t+1}W_{jt}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} \\ \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} X_{j,t+1}W_{j,t+1}\mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d\}} \end{bmatrix}.$$

B.2 Derivation of Restriction (4.1)

Let $v(d, w, x^*)$ denote the policy value function defined by

$$v(d, w_t, x_t^*) = \theta_d^0 + \theta_d^w w_t + \theta_d^x x_t^* + \rho\,\mathrm{E}\left[V(w_{t+1}, x_{t+1}^*) \mid d_t = d, w_t, x_t^*\right]$$

where $V(w_t, x_t^*)$ denotes the value of state (w_t, x_t^*). With this notation, we can write the difference in the expected value functions as

$$\begin{aligned} & \rho\,\mathrm{E}\left[V(w_{t+1}, x_{t+1}^*) \mid d_t = 1, w_t, x_t^*\right] - \rho\,\mathrm{E}\left[V(w_{t+1}, x_{t+1}^*) \mid d_t = 0, w_t, x_t^*\right] \\ = \;& v(1, w_t, x_t^*) - v(0, w_t, x_t^*) - \theta_1^w w_t - \theta_1^x x_t^* + \theta_0^0 + \theta_0^w w_t + \theta_0^x x_t^* \\ = \;& \ln f_{D_t|W_tX_t^*}(1 \mid w_t, x_t^*) - \ln f_{D_t|W_tX_t^*}(0 \mid w_t, x_t^*) - \theta_1^w w_t - \theta_1^x x_t^* + \theta_0^0 + \theta_0^w w_t + \theta_0^x x_t^* \end{aligned}$$

where $f_{D_t|W_tX_t^*}(d_t \mid w_t, x_t^*)$ is the conditional choice probability CCP, which we show is identified in Section 3.1. On the other hand, this difference in the expected value functions can also be explicitly written as

$$
\begin{aligned}
& \rho\, \mathrm{E}\left[V(w_{t+1}, x^*_{t+1}) \mid d_t = 1, w_t, x^*_t\right] - \rho\, \mathrm{E}\left[V(w_{t+1}, x^*_{t+1}) \mid d_t = 0, w_t, x^*_t\right] \\
= & \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f_{D_s|W_s,X^*_s}(0 \mid w_s, x^*_s) \cdot (\theta^0_0 + \theta^w_0 w_s + \theta^x_0 x^*_s + \varpi - \ln f_{D_s|W_s,X^*_s}(0 \mid w_s, x^*_s)) + \right. \\
& \left. f_{D_s|W_s,X^*_s}(1 \mid w_s, x^*_s) \cdot (\theta^w_1 w_s + \theta^x_1 x^*_s + \varpi - \ln f_{D_s|W_s,X^*_s}(1 \mid w_s, x^*_s)) \mid d_t = 1, w_t, x^*_t\right] \;- \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}\left[f_{D_s|W_s,X^*_s}(0 \mid w_s, x^*_s) \cdot (\theta^0_0 + \theta^w_0 w_s + \theta^x_0 x^*_s + \varpi - \ln f_{D_s|W_s,X^*_s}(0 \mid w_s, x^*_s)) + \right. \\
& \left. f_{D_s|W_s,X^*_s}(1 \mid w_s, x^*_s) \cdot (\theta^w_1 w_s + \theta^x_1 x^*_s + \varpi - \ln f_{D_s|W_s,X^*_s}(1 \mid w_s, x^*_s)) \mid d_t = 0, w_t, x^*_t\right]
\end{aligned}
$$

by the law of iterated expectations, where $\varpi \approx 0.577$ is the Euler constant. Equating the above two equalities yields (4.1).

B.3 Feasible Computation of Moments – Remark 2

This section is referred to by Remark 2, where otherwise-infeasible computation of the expectation with respect to the unobserved distribution of (w_t, x^*_t) is warranted to be feasible. We show how to obtain a feasible computation of such moments. Suppose that we have a moment restriction

$$0 \;=\; \int\int \zeta(w_t, x^*_t)\; dF(w_t, x^*)$$

which is infeasible to evaluate because of the unobservability of x^*_t. By applying the Bayes' rule and our identifying assumptions, we can rewrite this moment equality as

$$
\begin{aligned}
0 \;&=\; \int\int \zeta(w_t, x^*_t)\; dF(w_t, x^*) \\
&=\; \int\int \frac{\int \zeta(w_t, x^*_t) \cdot f(x_t \mid x^*_t) \cdot f(x^*_t \mid w_t)\; dx^*_t}{\int f(x_t \mid x^*_t) \cdot f(x^*_t \mid w_t)\; dx^*_t}\; dF(w_t, x_t) \qquad \text{(B.5)}
\end{aligned}
$$

Now that the integrator $dF(w_t, x)$ is the observed distribution of (w_t, x_t), we can evaluate the last line provided that we know $f(x_t \mid x^*_t)$ and $f(x^*_t \mid w_t)$. By Theorem 1, we identify (w_t, x_t) in a closed form as the proxy model. Hence, in order to evaluate the last line of the transformed moment equality, it remains to identify $f(x_t^* \mid w_t)$. The next paragraph therefore is devoted to this identification problem.

By the same arguments as in Step 1 of the proof of Theorem 1 in Section A.1 in the appendix, we can deduce

$$\mathrm{E}\left[\exp\left(isx_t^*\right)|d_t = d, w_t\right] = \exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1} - \alpha^d - \beta^d w_t)\exp\left(is_1 x_t\right)|d_t = d, w_t\right]}{\gamma^d\,\mathrm{E}\left[\exp\left(is_1 x_t\right)|d_t = d, w_t\right]} ds_1\right].$$

Therefore, we can recover the density $f(x_t^* \mid d_t = d, w_t)$ by applying the the operator $\mathcal{F}$ to the right-hand side of the above equality as

$$f(x_t^* \mid d_t = d, w_t) = \frac{1}{2\pi}\int e^{-isx_t^*} \cdot \exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1} - \alpha^d - \beta^d w_t)\exp\left(is_1 x_t\right)|d_t = d, w_t\right]}{\gamma^d\,\mathrm{E}\left[\exp\left(is_1 x_t\right)|d_t = d, w_t\right]} ds_1\right] ds.$$

Since the conditional distribution of $d_t \mid w_t$ is observed in data, d_t can be integrated out from the above equality as

$$\begin{aligned} f(x_t^* \mid w_t) &= \frac{1}{2\pi}\sum_d \int e^{-isx_t^*} \cdot f(d_t = d \mid w_t) \times \\ &\quad \exp\left[\int_0^s \frac{\mathrm{E}\left[i(x_{t+1} - \alpha^d - \beta^d w_t) \cdot \exp\left(is_1 x_t\right)|d_t = d, w_t\right]}{\gamma^d \cdot \mathrm{E}\left[\exp\left(is_1 x_t\right)|d_t = d, w_t\right]} ds_1\right] ds. \qquad \text{(B.6)} \end{aligned}$$

Therefore, $f(x_t^* \mid w_t)$ is identified in a closed form. This shows that the expression in the last line of (B.5) can be evaluated in a closed-form.

Lastly, we propose a sample-counterpart estimation of (B.6). The conditional density $f(x_t^* \mid w_t)$ is estimated in a closed form by

$$\begin{aligned} \widehat{f}(x^* \mid w) &= \frac{1}{2\pi}\sum_d \int e^{-isx^*} \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} K\left(\frac{W_{j,t} - w}{h_w}\right)} \times \\ &\exp\left[\int_0^s \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} i(X_{j,t+1} - \alpha^d - \beta^d W_{j,t}) \cdot \exp\left(is_1 X_{j,t}\right) \cdot \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)}{\gamma^d \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} \exp\left(is_1 X_{j,t}\right) \cdot \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)} ds_1\right] ds. \end{aligned}$$
(B.7)

B.4 Closed-Form Estimation of Structural Parameters

The closed-form identifying formulas obtained at the population level in Section 4 can be directly translated into sample counterparts to develop a closed-form estimator of structural parameters. Given Corollary 1 and Remark 2, we propose the following estimator.

$$\hat{\theta} = \left[\sum_{j=1}^{N} \sum_{t=1}^{T-1} \frac{\int \widehat{R}(\rho; W_{j,t}, x_t^*)' \widehat{R}(\rho; W_{j,t}, x_t^*) \cdot \widehat{f}(X_{j,t} \mid x_t^*) \cdot \widehat{f}(x_t^* \mid W_{j,t}) \, dx_t^*}{\int \widehat{f}(X_{j,t} \mid x_t^*) \cdot \widehat{f}(x_t^* \mid W_{j,t}) \, dx_t^*} \right]^{-1}$$
$$\left[\sum_{j=1}^{N} \sum_{t=1}^{T-1} \frac{\int \widehat{R}(\rho; W_{j,t}, x_t^*)' \widehat{\xi}(\rho; W_{j,t}, x_t^*) \cdot \widehat{f}(X_{j,t} \mid x_t^*) \cdot \widehat{f}(x_t^* \mid W_{j,t}) \, dx_t^*}{\int \widehat{f}(X_{j,t} \mid x_t^*) \cdot \widehat{f}(x_t^* \mid W_{j,t}) \, dx_t^*} \right] \tag{B.8}$$

where closed-form formulas for $\widehat{f}(X_{j,t} \mid x_t^*)$, $\widehat{f}(x_t^* \mid W_{j,t})$, $\widehat{\xi}(\rho; W_{j,t}, x_t^*)$, and $\widehat{R}(\rho; W_{j,t}, x_t^*) = \left[\widehat{\xi}_0^0(\rho; w_t, x_t^*), \quad \widehat{\xi}_0^w(\rho; w_t, x_t^*), \ \widehat{\xi}_1^w(\rho; w_t, x_t^*), \ \widehat{\xi}_0^x(\rho; w_t, x_t^*), \ \widehat{\xi}_1^x(\rho; w_t, x_t^*) \right]$ are listed below.

First, $\widehat{f}(x_t \mid x_t^*)$ is given by (B.4) in Section B.1. For convenience of readers, we repeat it here:

$$\widehat{f}(x \mid x^*) \;=\; \frac{1}{2\pi} \int \exp\left(-is(x - x^*)\right) \cdot \frac{\sum_{j=1}^{N} \sum_{t=1}^{T-1} \exp\left(isX_{jt}\right) \cdot \mathbb{1}\{D_{jt} = d'\}}{\widehat{\phi}_{x_t^* \mid d_t = d'}(s) \cdot \sum_{j=1}^{N} \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x) \; ds.$$

Second, $\widehat{f}(x_t^* \mid w_t)$ is given by (B.7) in Section B.3 in the supplementary note. We write it here too:

$$\widehat{f}(x^* \mid w) \;=\; \frac{1}{2\pi} \sum_{d} \int e^{-isx^*} \cdot \frac{\sum_{j=1}^{N} \sum_{t=1}^{T-1} \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)}{\sum_{j=1}^{N} \sum_{t=1}^{T-1} K\left(\frac{W_{j,t} - w}{h_w}\right)} \times$$
$$\exp\left[\int_0^s \frac{\sum_{j=1}^{N} \sum_{t=1}^{T-1} i(X_{j,t+1} - \alpha^d - \beta^d W_{j,t}) \cdot \exp\left(is_1 X_{j,t}\right) \cdot \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)}{\gamma^d \cdot \sum_{j=1}^{N} \sum_{t=1}^{T-1} \exp\left(is_1 X_{j,t}\right) \cdot \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)} ds_1 \right] \; ds.$$

Third, $\widehat{\xi}(\rho; w_t, x_t^*)$ and the elements of $\widehat{R}(\rho; w_t, x_t^*)$ are given by

$$\begin{aligned}
\widehat{\xi}(\rho; w_t, x_t^*) \;=\;& \ln \widehat{f}(1 \mid w_t, x_t^*) - \ln \widehat{f}(0 \mid w_t, x_t^*) + \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(0 \mid w_s, x_s^*) \cdot \ln \widehat{f}(0 \mid w_s, x_s^*) \mid d_t = 1, w_t, x_t^*\right] + \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(1 \mid w_s, x_s^*) \cdot \ln \widehat{f}(1 \mid w_s, x_s^*) \mid d_t = 1, w_t, x_t^*\right] - \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(0 \mid w_s, x_s^*) \cdot \ln \widehat{f}(0 \mid w_s, x_s^*) \mid d_t = 0, w_t, x_t^*\right] - \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(1 \mid w_s, x_s^*) \cdot \ln \widehat{f}(1 \mid w_s, x_s^*) \mid d_t = 0, w_t, x_t^*\right]
\end{aligned}$$

$$\begin{aligned}
\widehat{\xi}_0^0(\rho; w_t, x_t^*) \;=\;& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(0 \mid w_s, x_s^*) \mid d_t = 1, w_t, x_t^*\right] - \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(0 \mid w_s, x_s^*) \mid d_t = 0, w_t, x_t^*\right] - 1
\end{aligned}$$

$$\begin{aligned}
\widehat{\xi}_d^w(\rho; w_t, x_t^*) \;=\;& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(d \mid w_s, x_s^*) \cdot w_s \mid d_t = 1, w_t, x_t^*\right] - \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(d \mid w_s, x_s^*) \cdot w_s \mid d_t = 0, w_t, x_t^*\right] - (-1)^d \cdot w_t
\end{aligned}$$

$$\begin{aligned}
\widehat{\xi}_d^x(\rho; w_t, x_t^*) \;=\;& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(d \mid w_s, x_s^*) \cdot x_s^* \mid d_t = 1, w_t, x_t^*\right] - \\
& \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \widehat{\mathrm{E}}\left[\widehat{f}(d \mid w_s, x_s^*) \cdot x_s^* \mid d_t = 0, w_t, x_t^*\right] - (-1)^d \cdot x_t^*
\end{aligned}$$

for each $d \in \{0, 1\}$, following the sample counterparts of (4.2)–(4.5). Of these four sets of expressions, the components of the form $\widehat{f}(d_t \mid w_t, x_t^*)$ are given by (B.1) in Section B.1.

Following the sample counterparts of (4.8) and (4.9), the estimated conditional expectations of the form $\widehat{\mathrm{E}}[\widehat{\zeta}(w_s, x_s^*) \mid d_t, w_t, x_t^*]$ in the above expressions are in turn given in the following manner. If $s = t + 1$, then

$$\widehat{\mathrm{E}}[\widehat{\zeta}(w_s, x_s^*) \mid d_t, w_t, x_t^*] \;=\; \int\int \widehat{\zeta}(w_{t+1}, x_{t+1}^*) \cdot \widehat{f}(w_{t+1} \mid d_t, w_t, x_t^*) \times \widehat{f}(x_{t+1}^* \mid d_t, w_t, x_t^*)\; dw_{t+1} dx_{t+1}^*$$

where the closed-form estimator $\widehat{f}(w_{t+1} \mid d_t, w_t, x_t^*)$ is given by (B.2), and the closed-form estimator $\widehat{f}(x_{t+1}^* \mid d_t, w_t, x_t^*)$ is given by (B.3). On the other hand, if $s > t+1$, then

$$\begin{aligned}\widehat{\mathrm{E}}[\zeta(w_s, x_s^*) \mid d_t, w_t, x_t^*] \quad = \quad & \sum_{d_{t+1}=0}^{1} \cdots \sum_{d_{s-1}=0}^{1} \int \cdots \int \widehat{\zeta}(w_s, x_s^*) \cdot \widehat{f}(w_s \mid d_{s-1}, w_{s-1}, x_{s-1}^*) \times \\ & \widehat{f}(x_s^* \mid d_{s-1}, w_{s-1}, x_{s-1}^*) \cdot \prod_{\tau=t}^{s-2} \widehat{f}(d_{\tau+1} \mid w_\tau, x_\tau^*) \cdot \widehat{f}(w_{\tau+1} \mid d_\tau, w_\tau, x_\tau^*) \times \\ & \cdot \widehat{f}(x_{\tau+1}^* \mid d_\tau, w_\tau, x_\tau^*) \; dw_{t+1} \cdots dw_s \; dx_{t+1}^* \cdots dx_s^*.\end{aligned}$$

where the closed-form estimator $\widehat{f}(d_t \mid w_t, x_t^*)$ is given by (B.1), the closed-form estimator $\widehat{f}(w_{t+1} \mid d_t, w_t, x_t^*)$ is given by (B.2), and the closed-form estimator $\widehat{f}(x_{t+1}^* \mid d_t, w_t, x_t^*)$ is given by (B.3). In summary, every component in (B.8) can be expressed explicitly by the previously obtained closed-form estimators, and hence the estimator $\widehat{\theta}$ of the structural parameters is given in a closed form as well. Monte Carlo simulations of the estimator are presented in Section B.8 in the supplementary note.

B.5 The Estimator without the Observed State Variable

With the observed state variable w_t dropped, the moment restriction with the additional notations we use for our analysis of large sample properties becomes

$$\mathrm{E}\left[R(\rho, f; x_t^*)_t^{\prime *})\theta - R(\rho, f; x_t^*)_t^{\prime *})\right] = 0$$

where

$$R(\rho, f; x_t^*) = [\xi_0^0(\rho, f; x_t^*), \; \xi_0^x(\rho, f; x_t^*), \; \xi_1^x(\rho, f; x_t^*)]$$

and

$$
\begin{aligned}
\xi(\rho, f; x_t^*) \;&=\; \ln f(1 \mid x_t^*) - \ln f(0 \mid x_t^*) \\
+ \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot &\left(\mathrm{E}_f\left[f(0 \mid x_s^*) \cdot \ln f(0 \mid x_s^*) \mid d_t = 1, x_t^*\right] + \mathrm{E}_f\left[f(1 \mid x_s^*) \cdot \ln f(1 \mid x_s^*) \mid d_t = 1, x_t^*\right]\right) \\
- \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot &\left(\mathrm{E}_f\left[f(0 \mid x_s^*) \cdot \ln f(0 \mid x_s^*) \mid d_t = 0, x_t^*\right] + \mathrm{E}_f\left[f(1 \mid x_s^*) \cdot \ln f(1 \mid x_s^*) \mid d_t = 0, x_t^*\right]\right)
\end{aligned}
$$

$$
\xi_0^0(\rho, f; x_t^*) \;=\; \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \left(\mathrm{E}_f\left[f(0 \mid x_s^*) \mid d_t = 1, x_t^*\right] - \mathrm{E}_f\left[f(0 \mid x_s^*) \mid d_t = 0, x_t^*\right]\right) - 1
$$

$$
\begin{aligned}
\xi_d^x(\rho, f; x_t^*) \;&=\; \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}_f\left[f(d \mid x_s^*) \cdot x_s^* \mid d_t = 1, x_t^*\right] - \\
&\quad \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathrm{E}_f\left[f(d \mid x_s^*) \cdot x_s^* \mid d_t = 0, x_t^*\right] - (-1)^d \cdot x_t^*
\end{aligned}
$$

for each $d \in \{0, 1\}$. The subscript f under the E symbol indicates that the conditional expectation is computed based on the components f of the Markov kernel.

The components of the Markov kernel are estimated as follows. Let $\widehat{\phi}_{x_t^*|d_t=d}$ denote the sample-counterpart estimator of the conditional characteristic function $\phi_{x_t^*|d_t=d}$, defined by

$$
\widehat{\phi}_{x_t^*|d_t=d}(s) = \exp\left[\int_0^s \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} i(X_{j,t+1} - \alpha^d) \cdot \exp(is_1 X_{jt}) \cdot \mathbb{1}\{D_{jt} = d\}}{\gamma^d \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{jt}) \cdot \mathbb{1}\{D_{jt} = d\}} ds_1\right]
$$

The CCP, $f(d_t \mid x_t^*)$, is estimated in a closed form by

$$
\begin{aligned}
\widehat{f}(d|x^*) \;=\; &\int \exp(-isx^*) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d\}}{N(T-1)} \times \\
&\widehat{\phi}_{x_t^*|d_t=d'}(s) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x)\; ds \Bigg/ \\
&\int \exp(-isx^*) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt})}{N(T-1)} \times \\
&\widehat{\phi}_{x_t^*|d_t=d'}(s) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x)\; ds
\end{aligned}
$$

with any d', where ϕ_K denotes the Fourier transform of a kernel function K used for the purpose of regularization. The transition rule, $f(w_t \mid d_{t-1}, w_{t-1}, x_{t-1}^*)$, for the observed state variable w_t

is no longer estimated given the absence of w_t. The transition rule, $f(x^*_t \mid d_{t-1}, x^*_{t-1})$, for the unobserved state variable x^*_t is estimated in a closed form by

$$\begin{aligned}\widehat{f}(x^{*\prime *}) \;=\; & \frac{1}{2\pi}\int \exp\left(-is(x^{*\prime d}-\gamma^d x^*)\right) \times \\ & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{j,t+1}\right)\cdot \mathbb{1}\{D_{jt}=d\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(is(\alpha^d+\gamma^d X_{jt})\right)\cdot \mathbb{1}\{D_{jt}=d\}} \times \\ & \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(is\gamma^d X_{jt}\right)\cdot \mathbb{1}\{D_{jt}=d'\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{jt}\right)\cdot \mathbb{1}\{D_{jt}=d'\}}\cdot \frac{\widehat{\phi}_{x^*_t\mid d_t=d'}(s)}{\widehat{\phi}_{x^*_t\mid d_t=d'}(s\gamma^d)}\cdot \phi_K(sh_x)\; ds\end{aligned}$$

with any d'. Finally, the proxy model, $f(x_t \mid x^*_t)$, is estimated in a closed form by

$$\widehat{f}(x \mid x^*) \;=\; \frac{1}{2\pi}\int \exp\left(-is(x-x^*)\right)\cdot \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{jt}\right)\cdot \mathbb{1}\{D_{jt}=d'\}}{\widehat{\phi}_{x^*_t\mid d_t=d'}(s)\cdot \sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}\cdot \phi_K(sh_x)\; ds$$

using any d'. When each of the above estimators is evaluated at the j-th data point, the j-th data point is removed from the sum for the leave-one-out estimation.

B.6 Large Sample Properties

In this section, we present theoretical large sample properties of our closed-form CCP estimator. To economize our writings, we focus on a simplified version of the baseline model and the estimator, where we omit the observed state variable w_t, because the unobserved state variable x^*_t is of the first-order importance in this paper. Accordingly, we modify the estimator by simply removing the w_t-relevant part. See B.5 in the supplementary note for the exact expressions of the estimator that we obtain under this setting.

Assumption 7 (Large Sample)**.** *(a) The data $\{D_{j,t}, X^*_{j,t}\}_{t=1}^{T}$ is i.i.d. across j. respect to some metric. (b) $\mathcal{X}^* = supp(X^*_{j,t})$ is compact and convex. (c) $f(x^*) > 0$. (d) The density function of x^*_t is k_1-time continuously differentiable and the k_1-th derivative is Hölder continuous with exponent k_2, i.e., there exists k_0 such that $\left|f(k_1)(x^*) - f^{(k_1)}(x^*+\delta)\right| \leqslant k_0\,|\delta|^{k_2}$ for all $x^* \in \mathcal{X}^*$ and $\delta \in \mathbb{R}$. Let $k = k_1 + k_2$ be the largest number satisfying this property. (e) $f(d \mid x^*)$ is l_1-time* *continuously differentiable with respect to x^* and the l_1-th derivative is Hölder continuous with exponent l_2. Let $l = l_1 + l_2$ be the largest number satisfying this property. (f) ϕ_K is symmetric, satisfies $\phi_K(0) = 1$, and has integrable second derivatives.*

Assumption 8 (Smoothness). *(i) The conditional distribution of X_t given $D_t = d$ is ordinary-smooth of order $q > 0$ for some choice d, i.e., $\left|\phi_{x_t|d_t=d}(s)\right| = \mathcal{O}\left(|s|^{-q}\right)$ as $s \to \pm\infty$. (ii) The conditional distribution of X_t given $D_t = d$ is super-smooth of order $q > 0$ for some choice d, i.e., there exists $r > 0$ such that $\left|\phi_{x_t|d_t=d}(s)\right| = \mathcal{O}\left(e^{-|s|^q/r}\right)$ as $s \to \pm\infty$.*

The convergence rate depends on these two categories of the smoothness, the concept developed by Fan (1991).

Proposition 1 (Convergence Rate). *Suppose that Assumptions 1, 2, 3, 4, 5, and 7 are satisfied. If Assumption 8 (i) holds, then we have the asymptotic convergence rate*

$$\left(E\left[\widehat{f}(d \mid x^*) - f(d \mid x^*)\right]^2\right)^{1/2} \;=\; \mathcal{O}\left(N^{\frac{-\min\{k,l\}}{2(2+2q+\min\{k,l\})}}\right)$$

by the choice of the bandwidth parameter so that $h_x = O\left(n^{-1/(4+4q+2\min\{k,l\})}\right)$ as $N \to \infty$. If Assumption 8 (ii) holds, then we have the asymptotic convergence rate

$$\left(E\left[\widehat{f}(d \mid x^*) - f(d \mid x^*)\right]^2\right)^{1/2} \;=\; \mathcal{O}\left((\log N)^{-\min\{k,l\}/q}\right)$$

by the choice of the bandwidth parameter so that $h_x = O\left((\log n)^{-\min\{k,l\}/q}\right)$ as $N \to \infty$

A proof is given in Section B.7.

B.7 Proof of Proposition 1

Proof. First, note that Theorem 1 guarantees the identification under Assumptions 1, 2, 3, 4, and 5. Since estimation of α^d and γ^d does not affect the nonparametric convergence rates of the component estimators, we take these parameters as given henceforth. For a short-hand notation we denote the CCP by $g_d(x_t^*) := \mathrm{E}[\mathbb{1}\{d_t = d\} \mid x_t^*]$. Our CCP estimator is written as $\widehat{g_d(x^*)f(x^*)}/\widehat{f(x^*)}$ where

$$\begin{aligned}\widehat{g_d(x^*)f(x^*)} &= \frac{1}{2\pi}\int \exp\left(-isx^*\right)\cdot\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{jt}\right)\cdot\mathbb{1}\{D_{jt}=d\}}{N(T-1)}\times \\ &\qquad \widehat{\phi}_{x_t^*|d_t=d'}(s)\cdot\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{jt}\right)\cdot\mathbb{1}\{D_{jt}=d'\}}\cdot\phi_K(sh)\;ds\end{aligned}$$

and

$$\begin{aligned}\widehat{f(x^*)} &= \frac{1}{2\pi}\int \exp\left(-isx^*\right)\cdot\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{jt}\right)}{N(T-1)}\times \\ &\qquad \widehat{\phi}_{x_t^*|d_t=d'}(s)\cdot\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(isX_{jt}\right)\cdot\mathbb{1}\{D_{jt}=d'\}}\cdot\phi_K(sh)\;ds\end{aligned}$$

where $\widehat{\phi}_{x_t^*|d_t=d}$ is given by

$$\widehat{\phi}_{x_t^*|d_t=d}(s) = \exp\left[\int_0^s \frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1} i(X_{j,t+1}-\alpha^d)\cdot\exp\left(is_1X_{jt}\right)\cdot\mathbb{1}\{D_{jt}=d\}}{\gamma^d\cdot\sum_{j=1}^{N}\sum_{t=1}^{T-1}\exp\left(is_1X_{jt}\right)\cdot\mathbb{1}\{D_{jt}=d\}}ds_1\right].$$

The absolute bias of $\widehat{f(x^*)}$ is bounded by the following terms.

$$\begin{aligned}\left|\mathrm{E}\,\widehat{f(x^*)} - f(x^*)\right| &\leqslant \left|\mathrm{E}\,\widehat{f(x^*)} - \frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds\right| + \\ &\quad \left|\frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds - f(x^*)\right|\end{aligned}$$

The first term on the right-hand side has the following asymptotic order.

$$\begin{aligned}&\left|\mathrm{E}\,\widehat{f(x^*)} - \frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds\right| \\ =\; &\left|\frac{1}{2\pi}\int e^{-isx^*}\phi_K(sh)\left\{\mathrm{E}\left[\exp\left(i\int_0^s\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}ds_1\right)\times\right.\right. \\ &\left.\left.\frac{\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{isX_{jt}}\sum_{j=1}^{N}\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{N(T-1)\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}}\right] - \phi_{x_t^*|d_t=d'}(s)\frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)}\right\}ds\end{aligned}$$

$$
\begin{aligned}
\leqslant\; & \frac{\|\phi_K\|_\infty \left\|\phi_{x_t^*|d_t=d'}\right\|_\infty}{2\pi h} \int_{-1}^{1}\int_0^{s/h} \\
& \left(\frac{\|\phi_{x_t}\|_\infty \operatorname{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \operatorname{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right|}{\left|\phi_{x_t|d_t=d'}(s/h)\right|\left|\phi_{x_t|d_t=d'}(s_1)\right|\left|\gamma^{d'}\right| f(d')} \right. \\
& + \frac{\|\phi_{x_t}\|_\infty \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty \operatorname{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \operatorname{E} e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right|}{\left|\phi_{x_t|d_t=d'}(s/h)\right|\left|\phi_{x_t|d_t=d'}(s_1)\right| f(d')} \\
& + \frac{\operatorname{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{isX_{jt}} - \operatorname{E} e^{isX_{jt}}\right|}{\left|\phi_{x_t|d_t=d'}(s/h)\right|} \\
& \left. + \frac{\|\phi_{x_t}\|_\infty \operatorname{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\} - \operatorname{E} e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right|}{\left|\phi_{x_t|d_t=d'}(s/h)\right|^2 f(d')} + hot(s_1) + hot(s/h)\right) ds_1 ds \\
=\; & \mathcal{O}\left(\frac{1}{n^{1/2}h^2\left|\phi_{x_t|d_t=d'}(1/h)\right|^2}\right)
\end{aligned}
$$

where the higher-order terms hot vanish faster than the leading terms uniformly as $N \to \infty$ under Assumption 7 (b), since the empirical process

$$
\mathbb{G}_N(s) := \sqrt{N}\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\} - \operatorname{E}(X_{j,t+1}-\alpha^{d'})e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)
$$

for example converges uniformly as $\operatorname{E}\left((X_{j,t+1}-\alpha^{d'})e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)^2 \leqslant \operatorname{E}(X_{j,t+1}-\alpha^{d'})^2$ is invariant from s. On the other hand, the second term has the following asymptotic order.

$$
\begin{aligned}
& \left|\frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds - f(x^*)\right| \\
\leqslant\; & \left|\int f(x)h^{-1}K\left(\frac{x-x^*}{h}\right)dx - f(x^*)\right| \;=\; \mathcal{O}\left(h^k\right)
\end{aligned}
$$

where k is the Hölder exponent provided in Assumption 7 (d). Consequently, we obtain the following asymptotic order for the absolute bias of $\widehat{f(x^*)}$.

$$
\left|\operatorname{E}\widehat{f(x^*)} - f(x^*)\right| \;=\; \mathcal{O}\left(\frac{1}{n^{1/2}h^2\left|\phi_{x_t|d_t=d'}(1/h)\right|^2}\right) + \mathcal{O}\left(h^k\right).
$$

Similarly, the absolute bias of $\widehat{g_d(x^*)f(x^*)}$ is bounded by the following terms.

$$
\begin{aligned}
& \left|\operatorname{E}\widehat{g_d(x^*)f(x^*)} - g_d(x^*)f(x^*)\right| \\
\leqslant\; & \left|\operatorname{E}\widehat{g_d(x^*)f(x^*)} - \frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\operatorname{E}[e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}]}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds\right| + \\
& \left|\frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\operatorname{E}[e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}]}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds - g_d(x^*)f(x^*)\right|
\end{aligned}
$$

The first term on the right-hand side has the following asymptotic order.

$$\begin{aligned}
&\left|\mathrm{E}\,\widehat{f(x^*)} - \frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds\right| \\
= \quad & \left|\frac{1}{2\pi}\int e^{-isx^*}\phi_K(sh)\left\{\mathrm{E}\left[\exp\left(i\int_0^s\frac{\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}ds_1\right)\times\right.\right.\right. \\
& \left.\left.\left.\frac{\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}\sum_{j=1}^N\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{N(T-1)\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}}\right] - \phi_{x_t^*|d_t=d'}(s)\frac{\mathrm{E}[e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}]}{\phi_{x_t|d_t=d'}(s)}\right\}ds\right| \\
\leqslant \quad & \frac{\|\phi_K\|_\infty\left\|\phi_{x_t^*|d_t=d'}\right\|_\infty}{2\pi h}\int_{-1}^1\int_0^{s/h} \\
& \left(\frac{\left\|\phi_{x_t|d_t=d}\right\|_\infty\mathrm{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right|f(d)}{\left|\phi_{x_t|d_t=d'}(s/h)\right|\left|\phi_{x_t|d_t=d'}(s_1)\right||\gamma^{d'}|f(d')}\right. \\
& + \frac{\left\|\phi_{x_t|d_t=d}\right\|_\infty\left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty\mathrm{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right|f(d)}{\left|\phi_{x_t|d_t=d'}(s/h)\right|\left|\phi_{x_t|d_t=d'}(s_1)\right|f(d')} \\
& + \frac{\mathrm{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}-\mathrm{E}\,e^{isX_{jt}}\right|}{\left|\phi_{x_t|d_t=d'}(s/h)\right|} \\
& \left.+ \frac{\left\|\phi_{x_t|d_t=d}\right\|_\infty\mathrm{E}\left|\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right|f(d)}{\left|\phi_{x_t|d_t=d'}(s/h)\right|^2f(d')} + hot(s_1) + hot(s/h)\right)ds_1ds \\
= \quad & \mathcal{O}\left(\frac{1}{n^{1/2}h^2\left|\phi_{x_t|d_t=d'}(1/h)\right|^2}\right)
\end{aligned}$$

where the higher-order terms *hot* vanish faster than the leading terms uniformly as $N\to\infty$ under Assumption 7 (b). On the other hand, the second term has the following asymptotic order.

$$\begin{aligned}
&\left|\frac{1}{2\pi}\int e^{-isx^*}\phi_{x_t^*|d_t=d'}(s)\frac{\mathrm{E}[e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}]}{\phi_{x_t|d_t=d'}(s)}\phi_K(sh)ds - g_d(x^*)f(x^*)\right| \\
\leqslant \quad & \left|\int g_d(x)f(x)h^{-1}K\left(\frac{x-x^*}{h}\right)dx - g_d(x^*)f(x^*)\right| \quad = \quad \mathcal{O}\left(h^{\min\{k,l\}}\right)
\end{aligned}$$

where k and l are the Hölder exponents provided in Assumption 7 (d) and (e), respectively. Consequently, we obtain the following asymptotic order for the absolute bias of $\widehat{g_d(x^*)f(x^*)}$.

$$\left|\mathrm{E}\,\widehat{g_d(x^*)f(x^*)} - g_d(x^*)f(x^*)\right| \quad = \quad \mathcal{O}\left(\frac{1}{n^{1/2}h^2\left|\phi_{x_t|d_t=d'}(1/h)\right|^2}\right) + \mathcal{O}\left(h^{\min\{k,l\}}\right).$$

Next, the variance of $\widehat{f(x^*)}$ has the following asymptotic order.

$$\frac{1}{4\pi^2}\,\mathrm{E}\left(\int e^{-isx^*}\phi_K(sh)\left[\exp\left(i\int_0^s \frac{\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}ds_1\right)\times\right.\right.$$
$$\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\right)\left(\frac{\sum_{j=1}^N\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}}\right)-$$
$$\mathrm{E}\exp\left(i\int_0^s \frac{\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}ds_1\right)\times$$
$$\left.\left.\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\right)\left(\frac{\sum_{j=1}^N\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}}\right)\right]ds\right)^2$$

$$= \frac{1}{4\pi^2}\,\mathrm{E}\left(\int\int e^{-i(s+r)x^*}\phi_K(sh)\phi_K(rh)\phi_{x_t^*|d_t=d'}(s)\phi_{x_t^*|d_t=d'}(r)\int_0^s\int_0^r\right.$$
$$\left[\frac{\phi_{x_t}(s)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(s)\phi_{x_t|d_t=d'}(s_1)\gamma^{d'}f(d')}\right.$$
$$-\frac{\phi_{x_t}(s)\phi'_{x_t^*|d_t=d'}(s_1)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(s)\phi_{x_t|d_t=d'}(s_1)f(d')}$$
$$+\frac{\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}-\mathrm{E}\,e^{isX_{jt}}}{\phi_{x_t|d_t=d'}(s)}$$
$$\left.-\frac{\phi_{x_t}(s)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(s)^2f(d')}+hot(s)+hot(s_1)\right]\times$$
$$\left[\frac{\phi_{x_t}(r)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(r)\phi_{x_t|d_t=d'}(r_1)\gamma^{d'}f(d')}\right.$$
$$-\frac{\phi_{x_t}(r)\phi'_{x_t^*|d_t=d'}(r_1)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(r)\phi_{x_t|d_t=d'}(r_1)f(d')}$$
$$+\frac{\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{irX_{jt}}-\mathrm{E}\,e^{irX_{jt}}}{\phi_{x_t|d_t=d'}(r)}$$
$$\left.-\frac{\phi_{x_t}(r)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{irX_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{irX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(r)^2f(d')}+hot(s)+hot(s_1)\right]dr_1ds_1drds$$

$$\leqslant \frac{\|\phi_K\|_\infty^2\left\|\phi_{x_t^*|d_t=d'}\right\|_\infty^2}{4\pi^2}\int_{-1}^1\int_{-1}^1\int_0^{s/h}\int_0^{r/h}I(s,r,s_1,r_1,h)dr_1ds_1drds \;=\; \mathcal{O}\left(\frac{1}{Nh^4\left|\phi_{x_t|d_t=d'}(1/h)\right|^4}\right)$$

where $I(s,r,s_1,r_1,h)$ consists of the following ten terms and higher-order terms that vanish faster uniformly.

$$I_1 = \frac{\|\phi_{x_t}\|_\infty^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r_1)\right| \cdot f(d')^2 \cdot (\gamma^{d'})^2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}(X_{j,t+1} - \alpha^{d'}) e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}$$

$$I_2 = \frac{\|\phi_{x_t}\|_\infty^2 \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r_1)\right| \cdot f(d')^2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}$$

$$I_3 = \frac{1}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|} \cdot \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{isX_{jt}/h} - \mathrm{E}\, e^{isX_{jt}/h} \right]^2 \right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{irX_{jt}/h} - \mathrm{E}\, e^{irX_{jt}/h} \right]^2 \right)^{1/2}$$

$$I_4 = \frac{\|\phi_{x_t}\|_\infty^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right|^2 \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')^2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{isX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{isX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{irX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{irX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}$$

$$I_5 = \frac{2\,\|\phi_{x_t}\|_\infty^2 \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r_1)\right| \cdot f(d')^2 \cdot |\gamma^{d'}|} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}$$

$$I_6 = \frac{2\left\|\phi_{x_t}\right\|_\infty}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot f(d') \cdot |\gamma^{d'}|} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h} - \mathrm{E}\, e^{irX_{jt}/h}\right]^2\right)^{1/2}$$

$$I_7 = \frac{2\left\|\phi_{x_t}\right\|_\infty^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')^2 \cdot |\gamma^{d'}|} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\, e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}$$

$$I_8 = \frac{2\left\|\phi_{x_t}\right\|_\infty \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot f(d')} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\, e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h} - \mathrm{E}\, e^{irX_{jt}/h}\right]^2\right)^{1/2}$$

$$I_9 = \frac{2\left\|\phi_{x_t}\right\|_\infty^2 \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')^2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\, e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\, e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}$$

$$I_{10} = \frac{2\left\|\phi_{x_t}\right\|_\infty}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')} \cdot \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{isX_{jt}/h} - \mathrm{E}\, e^{isX_{jt}/h}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\, e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}$$

Similarly, the variance of $\widehat{g_d(x^*)f(x^*)}$ has the following asymptotic order.

$$\frac{1}{4\pi^2}\,\mathrm{E}\left(\int e^{-isx^*}\phi_K(sh)\left[\exp\left(i\int_0^s \frac{\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}ds_1\right)\times\right.\right.$$
$$\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}\right)\left(\frac{\sum_{j=1}^N\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}}\right)-$$
$$\mathrm{E}\exp\left(i\int_0^s \frac{\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}}ds_1\right)\times$$
$$\left.\left.\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}\right)\left(\frac{\sum_{j=1}^N\sum_{t=1}^{T-1}\mathbb{1}\{D_{jt}=d'\}}{\gamma^{d'}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}}\right)\right]ds\right)^2$$

$$= \frac{1}{4\pi^2}\,\mathrm{E}\left(\int\int e^{-i(s+r)x^*}\phi_K(sh)\phi_K(rh)\phi_{x_t^*|d_t=d'}(s)\phi_{x_t^*|d_t=d'}(r)\int_0^s\int_0^r\right.$$
$$\left[\frac{\phi_{x_t|d_t=d}(s)f(d)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(s)\phi_{x_t|d_t=d'}(s_1)\gamma^{d'}f(d')}\right.$$
$$-\frac{\phi_{x_t|d_t=d}(s)\phi'_{x_t^*|d_t=d'}(s_1)f(d)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(s)\phi_{x_t|d_t=d'}(s_1)f(d')}$$
$$+\frac{\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}-\mathrm{E}\,e^{isX_{jt}}\mathbb{1}\{D_{jt}=d\}}{\phi_{x_t|d_t=d'}(s)}$$
$$\left.-\frac{\phi_{x_t|d_t=d}(s)f(d)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{isX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(s)^2f(d')}+hot(s)+hot(s_1)\right]\times$$
$$\left[\frac{\phi_{x_t|d_t=d}(r)f(d)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(r)\phi_{x_t|d_t=d'}(r_1)\gamma^{d'}f(d')}\right.$$
$$-\frac{\phi_{x_t|d_t=d}(r)\phi'_{x_t^*|d_t=d'}(r_1)f(d)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(r)\phi_{x_t|d_t=d'}(r_1)f(d')}$$
$$+\frac{\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{irX_{jt}}\mathbb{1}\{D_{jt}=d\}-\mathrm{E}\,e^{irX_{jt}}\mathbb{1}\{D_{jt}=d\}}{\phi_{x_t|d_t=d'}(r)}$$
$$\left.-\frac{\phi_{x_t|d_t=d}(r)f(d)\left(\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1}e^{irX_{jt}}\mathbb{1}\{D_{jt}=d'\}-\mathrm{E}\,e^{irX_{jt}}\mathbb{1}\{D_{jt}=d'\}\right)}{\phi_{x_t|d_t=d'}(r)^2f(d')}+hot(s)+hot(s_1)\right]dr_1ds_1drds$$

$$\leqslant \frac{\|\phi_K\|_\infty^2\left\|\phi_{x_t^*|d_t=d'}\right\|_\infty^2}{4\pi^2}\int_{-1}^1\int_{-1}^1\int_0^{s/h}\int_0^{r/h}J(s,r,s_1,r_1,h)dr_1ds_1drds = \mathcal{O}\left(\frac{1}{Nh^4\left|\phi_{x_t|d_t=d'}(1/h)\right|^4}\right)$$

where $J(s,r,s_1,r_1,h)$ consists of the following ten terms and higher-order terms that vanish faster uniformly.

$$
\begin{aligned}
J_1 \quad = \quad & \frac{\left\|\phi_{x_t|d_t=d}\right\|_\infty^2 f(d)^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r_1)\right| \cdot f(d')^2 \cdot (\gamma^{d'})^2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}(X_{j,t+1} - \alpha^{d'}) e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_2 \quad = \quad & \frac{\left\|\phi_{x_t|d_t=d}\right\|_\infty^2 \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty^2 f(d)^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r_1)\right| \cdot f(d')^2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_3 \quad = \quad & \frac{1}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{isX_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathrm{E}\, e^{isX_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{irX_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathrm{E}\, e^{irX_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_4 \quad = \quad & \frac{\left\|\phi_{x_t|d_t=d}\right\|_\infty^2 f(d)^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right|^2 \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')^2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{isX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{isX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)} \sum_{j=1}^{N} \sum_{t=1}^{T-1} e^{irX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathrm{E}\, e^{irX_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_5 \quad = \quad & \frac{2\left\|\phi_{x_t|d_t=d}\right\|_\infty^2 \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty f(d)^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r_1)\right| \cdot f(d')^2 \cdot |\gamma^{d'}|} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\,e^{ir_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_6 \quad = \quad & \frac{2\left\|\phi_{x_t|d_t=d}\right\|_\infty f(d)}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot f(d') \cdot |\gamma^{d'}|} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d\} - \mathrm{E}\,e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d\}\right]^2\right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_7 \quad = \quad & \frac{2\left\|\phi_{x_t|d_t=d}\right\|_\infty^2 f(d)^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')^2 \cdot |\gamma^{d'}|} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}(X_{j,t+1}-\alpha^{d'})e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\,e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
J_8 \quad = \quad & \frac{2\left\|\phi_{x_t|d_t=d}\right\|_\infty \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty f(d)}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right| \cdot f(d')} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\,e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times \\
& \left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^{N}\sum_{t=1}^{T-1}e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d\} - \mathrm{E}\,e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d\}\right]^2\right)^{1/2}
\end{aligned}
$$

$$J_9 = \frac{2\left\|\phi_{x_t|d_t=d}\right\|_\infty^2 \left\|\phi'_{x_t^*|d_t=d'}\right\|_\infty f(d)^2}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(s_1)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')^2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\,e^{is_1X_{jt}}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\,e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}$$

$$J_{10} = \frac{2\left\|\phi_{x_t|d_t=d}\right\|_\infty f(d)}{\left|\phi_{x_t|d_t=d'}(s/h)\right| \cdot \left|\phi_{x_t|d_t=d'}(r/h)\right|^2 \cdot f(d')} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{isX_{jt}/h}\mathbb{1}\{D_{jt}=d\} - \mathrm{E}\,e^{isX_{jt}/h}\mathbb{1}\{D_{jt}=d\}\right]^2\right)^{1/2} \times$$
$$\left(\mathrm{E}\left[\frac{1}{N(T-1)}\sum_{j=1}^N\sum_{t=1}^{T-1} e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\} - \mathrm{E}\,e^{irX_{jt}/h}\mathbb{1}\{D_{jt}=d'\}\right]^2\right)^{1/2}$$

Consequently, under Assumption 8 (i), the bandwidth parameter choice prescribed in part (i) of the proposition yields

$$\left(\mathrm{E}\left[\widehat{g_d(x^*)f(x^*)} - g_d(x^*)f(x^*)\right]^2\right)^{1/2} = \mathcal{O}\left(N^{\frac{-\min\{k,l\}}{2(2+2q+\min\{k,l\})}}\right)$$
$$\left(\mathrm{E}\left[\widehat{f(x^*)} - f(x^*)\right]^2\right)^{1/2} = \mathcal{O}\left(N^{\frac{-k}{2(2+2q+\min\{k,l\})}}\right).$$

Since the MSE of the CCP estimator is given by

$$\frac{1}{f(x^*)^2}\mathrm{MSE}\left(\widehat{g_d(x^*)f(x^*)}\right) + \frac{g_d(x^*)^2}{f(x^*)^2}\mathrm{MSE}\left(\widehat{f(x^*)}\right),$$

it follows that

$$\left(\mathrm{E}\left[\widehat{g_d(x^*)} - g_d(x^*)\right]^2\right)^{1/2} = \mathcal{O}\left(N^{\frac{-\min\{k,l\}}{2(2+2q+\min\{k,l\})}}\right)$$

and thus part (i) of the proposition holds.

Likewise, under Assumption 8 (ii), the bandwidth parameter choice prescribed in part (ii) of the proposition yields

$$\left(\mathrm{E}\left[\widehat{g_d(x^*)f(x^*)} - g_d(x^*)f(x^*)\right]^2\right)^{1/2} = \mathcal{O}\left((\log N)^{-\min\{k,l\}/q}\right).$$
$$\left(\mathrm{E}\left[\widehat{f(x^*)} - f(x^*)\right]^2\right)^{1/2} = \mathcal{O}\left((\log N)^{-k/q}\right).$$

It follows that

$$\left(\mathrm{E}\left[\widehat{g_d(x^*)} - g_d(x^*)\right]^2\right)^{1/2} \;=\; \mathcal{O}\left((\log N)^{-\min\{k,l\}/q}\right)$$

and thus part (ii) of the proposition holds. □

B.8 Monte Carlo Simulations

In this section, we evaluate finite sample performance of our estimator using artificial data based on a benchmark structural model of the literature, that is reminiscent of Rust (1987). We focus on a parsimonious version of the general model, where states consist only of the unobserved variable x_t^*, hence suppressing w_t from the baseline model.[8] The transition rule for the unobserved state variable x_t^* is defined by

$$x_t^* \;=\; 1.000 + 1.000 \cdot x_{t-1}^* + \eta_t^0 \qquad \text{if } d = 0$$

$$x_t^* \;=\; 0.000 + 0.000 \cdot x_{t-1}^* + \eta_t^1 \qquad \text{if } d = 1$$

where η_t^0 and η_t^1 are independently distributed according to the standard normal distribution, $N(0,1)$. In the context of Rust' model, the true state x_t^* of the capital (e.g., mileage of the engine) accumulates if continuation $d = 0$ is selected, while it is reset to zero if replacement $d = 1$ is selected. The parameters of the state transition are summarized by the vector

[8] In this case, there arises a minor modification in our closed-form identifying formulas and estimators. First, the random variable w_t and the parameter β^d become absent. Second, the transition rule for the observed state w_t becomes unnecessary. Third, most importantly, the estimator of (α^d, γ^d) will be based on two equations for two unknowns, where w_{t-1} is replaced by x_{t-2}. Other than these points, the main procedure continues to be the same. Also see Section B.6 where large sample properties of the estimator are studied in a similarly simplified setup.

$(\alpha^0, \gamma^0, \alpha^1, \gamma^1) = (1.000,\ 1.000,\ 0.000,\ 0.000)$. The current utility is given by

$$1.000 - 0.015 \cdot x_t^* + \omega_{0t} \qquad \text{if } d = 0$$

$$0.000 \cdot x_t^* + \omega_{1t} \qquad \text{if } d = 1$$

where ω_{0t} and ω_{1t} are independently distributed according to the Type I Extreme Value Distribution. In the context of Rust's model, continuation $d = 0$ incurs the marginal cost of 0.015 per the true state x_t^*, whereas replacement incurs the fixed cost 1.000. The structural parameters for the payoff are summarized by the vector $\theta = (\theta_0^0, \theta_0^x, \theta_1^x)' = (1.000,\ -0.015,\ 0.000)'$. The rate of time preference is set to $\rho = 0.9$. Lastly, the proxy model is defined by

$$x_t \;=\; x_t^* + \varepsilon_t$$

where $\varepsilon_t \sim N(0, 2)$. We thus let $\text{Var}(\varepsilon_t) = 2 > 1 = \text{Var}(\eta_t)$ to assess how our method performs under relatively large stochastic magnitudes of measurement errors.

With this setup, we present Monte Carlo simulation results for the closed-form estimator of the structural parameters θ developed in Section B.4. The shaded rows, (III) and (VI), in Table 4 provide a summary of MC-simulation results for our estimator. The other rows in Table 4 report MC-simulation results for alternative estimators for the purpose of comparison. The results in row (I) are based on the traditional estimator and the assumption that the observed proxy x_t is mistakenly treated as the unobserved state variable x_t^*. The results in row (II) are based on the traditional estimator and the metaphysical assumption that the unobserved state variable x_t^* were observed. On the other hand, the results in row (III) are based on the assumption that x_t^* is not observed, but the measurement error of x_t is accounted for by our closed-form estimator. All the results in these three rows are based on sampling with $N = 100$, $T = 10$. Rows (IV), (V) and (VI) are analogous to rows (I), (II) and (III), respectively, except that the sample size of $N = 500$ is used instead of $N = 100$. While the estimates θ_1^x are good

	Data		True	Interdecile Range	Mean	Bias	$\sqrt{\text{Var}}$	RMSE
(I)	$N = 100$ & $T = 10$	θ_0^x	1.000	$[0.772, 1.066]$	0.916	-0.084	0.104	0.133
	x_t treated wrongly as x_t^*	θ_1^x	-0.015	$[-0.049, 0.021]$	-0.014	0.001	0.024	0.024
(II)	$N = 100$ & $T = 10$	θ_0^x	1.000	$[0.837, 1.140]$	0.982	-0.018	0.101	0.103
	If x_t^* *were* observed	θ_1^x	-0.015	$[-0.041, 0.007]$	-0.016	-0.001	0.017	0.018
(III)	$N = 100$ & $T = 10$	θ_0^x	1.000	$[0.671, 1.414]$	0.988	-0.012	0.418	0.417
	Accounting for ME of x_t^*	θ_1^x	-0.015	$[-0.098, 0.032]$	-0.024	-0.009	0.050	0.052
(IV)	$N = 500$ & $T = 10$	θ_0^x	1.000	$[0.844, 0.981]$	0.911	-0.089	0.045	0.100
	x_t treated wrongly as x_t^*	θ_1^x	-0.015	$[-0.036, -0.006]$	-0.020	-0.005	0.010	0.014
(V)	$N = 500$ & $T = 10$	θ_0^x	1.000	$[0.898, 1.046]$	0.973	-0.027	0.050	0.057
	If x_t^* *were* observed	θ_1^x	-0.015	$[-0.030, -0.007]$	-0.019	-0.004	0.008	0.012
(VI)	$N = 500$ & $T = 10$	θ_0^x	1.000	$[0.835, 1.160]$	0.970	-0.030	0.220	0.222
	Accounting for ME of x_t^*	θ_1^x	-0.015	$[-0.042, 0.023]$	-0.014	0.001	0.027	0.027

Table 4: Summary statistics of the Monte Carlo simulated estimates of the structural parameters. The interdecile range shows the 10-th and 90-th percentiles of Monte Carlo distributions. All the other statistics are based on five-percent trimmed sample to suppress the effects of outliers. The results are based on (I) sampling with $N = 100$ and $T = 10$ where x_t is treated wrongly as x_t^*, (II) sampling with $N = 100$ and $T = 10$ where x_t^* is assumed to be observed, (III) sampling with $N = 100$ and $T = 10$ where the measurement error of unobserved x_t^* is accounted. Results shown in rows (IV), (V) and (VI) are analogous to those in rows (I), (II) and (III), respectively, except that the sample size of $N = 500$ is used. The shaded rows (III) and (VI) indicate use of our closed-form estimators which can handle unobserved state variables.

enough across all the six sets of experiments, the estimates of θ_0^x are substantially biased under rows (I) and (IV), which fail to account for the measurement error. Furthermore, the interdecile range of θ_0^x in row (IV) does not contain the true value. On the other hand, the MC-results of our estimator shown in rows (III) and (VI) of Table 4 are much less biased, like those in rows (II) and (V) of Table 4.

B.9 Extending the Proxy Model

The baseline model presented in Section 3.1 assumes classical measurement errors. To relax this assumption, we may allow the relationship between the proxy and the unobserved state variable to depend on the endogenous choice made in previous period. This generalization is useful if the past action can affect the measurement nature of the proxy variable. For example, when the choice d_t leads to entry and exit status of a firm, what proxy measure we may obtain for the unobserved productivity of the firm may differ depending whether the firm is in or out of the market.

To allow the proxy model to depend on edogeneous actions, we modify Assumptions 2, 3, 4 and 5 as follows.

Assumption 2′. The Markov kernel can be decomposed as follows.

$$\begin{aligned} & f\left(d_t, w_t, x_t^*, x_t | d_{t-1}, w_{t-1}, x_{t-1}^*, x_{t-1}\right) \\ = \; & f\left(d_t | w_t, x_t^*\right) f\left(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*\right) f\left(x_t^* | d_{t-1}, w_{t-1}, x_{t-1}^*\right) f\left(x_t | d_{t-1}, x_t^*\right) \end{aligned}$$

where the proxy model now depends on the endogenous choice d_{t-1} made in the last period.

Assumption 3′. The transition rule for the unobserved state variable and the state-proxy relation are semi-parametrically specified by

$$f\left(x_t^*|d_{t-1},w_{t-1},x_{t-1}^*\right): \qquad x_t^* = \alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^* + \eta_t^d \qquad \text{if } d_{t-1}=d$$

$$f\left(x_t|d_{t-1},x_t^*\right): \qquad x_t = \delta^d x_t^* + \varepsilon_t^d \qquad \text{if } d_{t-1}=d$$

where ε_t and η_t^d have mean zero for each d, and satisfy

$$\varepsilon_t^d \perp\!\!\!\perp (\{d_\tau\}_\tau, \{x_\tau^*\}_\tau, \{w_\tau\}_\tau, \{\varepsilon_\tau\}_{\tau\neq t}) \qquad \text{for all } t$$

$$\eta_t^d \perp\!\!\!\perp (d_\tau, x_\tau^*, w_\tau) \qquad \text{for all } \tau < t \text{ for all } t.$$

where $\varepsilon_t = (\varepsilon_t^0, \varepsilon_t^1, \cdots, \varepsilon_t^{\bar{d}})$.

Assumption 4′. For each d, $((d_{t-1}=d) > 0$ and the following matrix is nonsingular for each of $d'=d$ and $d'=0$.

$$\begin{bmatrix} 1 & \mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[w_{t-1}^2 \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1}w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[w_t \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[w_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] \end{bmatrix}$$

Assumption 5′. The random variables w_t and x_t^* have bounded conditional first moments given (d_t, d_{t-1}). The conditional characteristic functions of w_t and x_t^* given (d_t, d_{t-1}) do not vanish on the real line, and is absolutely integrable. The conditional characteristic function of (x_{t-1}^*, w_t) given $(d_{t-1}, d_{t-2}, w_{t-1})$ and the conditional characteristic function of x_t^* given (w_t, d_{t-1}) are absolutely integrable. Random variables ε_t and η_t^d have bounded first moments and absolutely integrable characteristic functions that do not vanish on the real line.

Because x_t^* is unit-less unobserved variable, there would be a continuum of observationally equivalent set of $(\delta^0, \cdots, \delta^{\bar{d}})$ and distributions of $(\varepsilon_t^0, \cdots, \varepsilon_t^{\bar{d}})$, unless we normalize δ^d for one of the choices d. We therefore make the following assumption in addition to the baseline assumptions.

Assumption 9. *WLOG, we normalize* $\delta^0 = 1$.

Under this set of assumptions that are analogous to those we assumed for the baseline model in Section 3.1, we obtain the following closed-form identification result analogous to Theorem 1.

Theorem 2 (Closed-Form Identification)**.** *If Assumptions 1, 2′, 3′, 4′, 5′, and 9 are satisfied, then the four components* $f(d_t|w_t, x_t^*)$, $f(w_t|d_{t-1}, w_{t-1}, x_{t-1}^*)$, $f(x_t^*|d_{t-1}, w_{t-1}, x_{t-1}^*)$, $f(x_t|d_{t-1}, x_t^*)$ *of the Markov kernel* $f(d_t, w_t, x_t^*, x_t|d_{t-1}, w_{t-1}, x_{t-1}^*, x_{t-1})$ *are identified by closed-form formulas.*

A proof and a set of full closed-form identifying formulas are given in Section B.10. This section demonstrated that, even if endogenous actions of firms, such as the decision of exit, can potentially affect the measurement nature of proxy variables through market participation status, we still obtain similar closed-form estimator with slight modifications.

B.10 Proof of Theorem 2

Proof. Similarly to the baseline case, our closed-form identification includes four steps.

Step 1: Closed-form identification of the transition rule $f(x_t^*|d_{t-1}, w_{t-1}, x_{t-1}^*)$**:** First, we show the identification of the parameters and the distributions in transition of x_t^*. Since

$$
\begin{aligned}
x_t &= \sum_d \mathbb{1}\{d_{t-1} = d\}[\delta^d x_t^* + \varepsilon_t^d] \\
&= \sum_d \mathbb{1}\{d_{t-1} = d\}[\alpha^d \delta^d + \beta^d \delta^d w_{t-1} + \gamma^d \delta^d x_{t-1}^* + \delta^d \eta_t^d + \varepsilon_t^d] \\
&= \sum_d \sum_{d'} \mathbb{1}\{d_{t-1} = d\}\mathbb{1}\{d_{t-2} = d'\} \left[\alpha^d \delta^d + \beta^d \delta^d w_{t-1} + \gamma^d \frac{\delta^d}{\delta^{d'}} x_{t-1} + \delta^d \eta_t^d + \varepsilon_t^d - \gamma^d \frac{\delta^d}{\delta^{d'}} \varepsilon_{t-1}^{d'}\right]
\end{aligned}
$$

we obtain the following equalities for each d and d':

$$\begin{aligned}
\mathrm{E}[x_t \mid d_{t-1}=d, d_{t-2}=d'] &= \alpha^d\delta^d + \beta^d\delta^d\,\mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\
&\quad +\gamma^d\frac{\delta^d}{\delta^{d'}}\,\mathrm{E}[x_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\
\mathrm{E}[x_t w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] &= \alpha^d\delta^d\,\mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\
&\quad +\beta^d\delta^d\,\mathrm{E}[w_{t-1}^2 \mid d_{t-1}=d, d_{t-2}=d'] \\
&\quad +\gamma^d\frac{\delta^d}{\delta^{d'}}\,\mathrm{E}[x_{t-1}w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\
\mathrm{E}[x_t w_t \mid d_{t-1}=d, d_{t-2}=d'] &= \alpha^d\delta^d\,\mathrm{E}[w_t \mid d_{t-1}=d, d_{t-2}=d'] \\
&\quad +\beta^d\delta^d\,\mathrm{E}[w_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] \\
&\quad +\gamma^d\frac{\delta^d}{\delta^{d'}}\,\mathrm{E}[x_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d']
\end{aligned}$$

by the independence and zero mean assumptions for η_t^d and ε_t^d. From these, we have the linear equation

$$\begin{bmatrix} \mathrm{E}[x_t \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[x_t w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[x_t w_t \mid d_{t-1}=d, d_{t-2}=d'] \end{bmatrix} =$$

$$\begin{bmatrix} 1 & \mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[w_{t-1}^2 \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1}w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[w_t \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[w_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] \end{bmatrix} \begin{bmatrix} \alpha^d\delta^d \\ \beta^d\delta^d \\ \gamma^d\frac{\delta^d}{\delta^{d'}} \end{bmatrix}$$

Provided that the matrix on the right-hand side is non-singular, we can identify the composite parameters $\left(\alpha^d\delta^d, \beta^d\delta^d, \gamma^d\frac{\delta^d}{\delta^{d'}}\right)$ by

$$\begin{bmatrix} \alpha^d\delta^d \\ \beta^d\delta^d \\ \gamma^d\frac{\delta^d}{\delta^{d'}} \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[w_{t-1}^2 \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1}w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[w_t \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[w_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] & \mathrm{E}[x_{t-1}w_t \mid d_{t-1}=d, d_{t-2}=d'] \end{bmatrix}^{-1} \times \begin{bmatrix} \mathrm{E}[x_t \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[x_t w_{t-1} \mid d_{t-1}=d, d_{t-2}=d'] \\ \mathrm{E}[x_t w_t \mid d_{t-1}=d, d_{t-2}=d'] \end{bmatrix}.$$

Once the composite parameters $\gamma^d \frac{\delta^d}{\delta^0}$ and $\gamma^d = \gamma^d \frac{\delta^d}{\delta^d}$ are identified by the above formula, we can in turn identify

$$\delta^d = \frac{\gamma^d \frac{\delta^d}{\delta^0}}{\gamma^d \frac{\delta^d}{\delta^d}}$$

for each d by the normalization assumption $\delta^0 = 1$. It in turn can be used to identify $(\alpha^d, \beta^d, \gamma^d)$ for each d from the identified composite parameters $\left(\alpha^d \delta^d, \beta^d \delta^d, \gamma^d \frac{\delta^d}{\delta^0}\right)$ by

$$(\alpha^d, \beta^d, \gamma^d) = \frac{1}{\delta^d}\left(\alpha^d \delta^d, \beta^d \delta^d, \gamma^d \frac{\delta^d}{\delta^0}\right).$$

Next, we show identification of $f\left(\varepsilon_t^d\right)$ and $f\left(\eta_t^d\right)$ for each d. Observe that

$$\begin{aligned}
& \mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 x_t\right)|d_{t-1} = d, d_{t-2} = d'\right] \\
= \; & \mathrm{E}\left[\exp\left(is_1\left(\delta^{d'} x_{t-1}^* + \varepsilon_{t-1}^{d'}\right) + is_2\left(\alpha^d\delta^d + \beta^d\delta^d w_{t-1} + \gamma^d\delta^d x_{t-1}^* + \delta^d\eta_t^d + \varepsilon_t^d\right)\right)|d_{t-1} = d, d_{t-2} = d'\right] \\
= \; & \mathrm{E}\left[\exp\left(i\left(s_1\delta^{d'} x_{t-1}^* + s_2\alpha^d\delta^d + s_2\beta^d\delta^d w_{t-1} + s_2\gamma^d\delta^d x_{t-1}^*\right)\right)|d_{t-1} = d, d_{t-2} = d'\right] \\
& \quad \times \mathrm{E}\left[\exp\left(is_1\varepsilon_{t-1}^{d'}\right)\right]\mathrm{E}\left[\exp\left(is_2\left(\delta^d\eta_t^d + \varepsilon_t^d\right)\right)\right]
\end{aligned}$$

follows for each pair (d, d') from the independence assumptions for η_t^d and ε_t^d for each d. We may then use the Kotlarski's identity

$$\begin{aligned}
& \left[\frac{\partial}{\partial s_2}\ln \mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 x_t\right)|d_{t-1} = d, d_{t-2} = d'\right]\right]_{s_2=0} \\
= \; & \frac{\mathrm{E}\left[i(\alpha^d\delta^d + \beta^d\delta^d w_{t-1} + \gamma^d\delta^d x_{t-1}^*)\exp\left(is_1\delta^{d'} x_{t-1}^*\right)|d_{t-1} = d, d_{t-2} = d'\right]}{\mathrm{E}\left[\exp\left(is_1\delta^{d'} x_{t-1}^*\right)|d_{t-1} = d, d_{t-2} = d'\right]} \\
= \; & i\alpha^d\delta^d + \beta^d\delta^d\frac{\mathrm{E}[iw_{t-1}\exp(is_1\delta^{d'} x_{t-1}^*) \mid d_{t-1} = d, d_{t-2} = d']}{\mathrm{E}[\exp(is_1\delta^{d'} x_{t-1}^*) \mid d_{t-1} = d, d_{t-2} = d']} \\
& + \gamma^d\frac{\delta^d}{\delta^{d'}}\frac{\partial}{\partial s_1}\ln \mathrm{E}\left[\exp\left(is_1\delta^{d'} x_{t-1}^*\right)|d_{t-1} = d, d_{t-2} = d'\right] \\
= \; & i\alpha^d\delta^d + \beta^d\delta^d\frac{\mathrm{E}[iw_{t-1}\exp(is_1 x_{t-1}) \mid d_{t-1} = d, d_{t-2} = d']}{\mathrm{E}[\exp(is_1 x_{t-1}) \mid d_{t-1} = d, d_{t-2} = d']} \\
& + \gamma^d\frac{\delta^d}{\delta^{d'}}\frac{\partial}{\partial s_1}\ln \mathrm{E}\left[\exp\left(is_1\delta^{d'} x_{t-1}^*\right)|d_{t-1} = d, d_{t-2} = d'\right]
\end{aligned}$$

Therefore,

$$
\begin{aligned}
& \mathrm{E}\left[\exp\left(is\delta^{d'}x^*_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right] \\
= & \exp\left[\int_0^s\left[\frac{\delta^{d'}}{\gamma^d\delta^d}\frac{\partial}{\partial s_2}\ln \mathrm{E}\left[\exp\left(is_1x_{t-1}+is_2x_t\right)|d_{t-1}=d,d_{t-2}=d'\right]\right]_{s_2=0}ds_1\right. \\
& \left.-\int_0^s\frac{i\alpha^d\delta^{d'}}{\gamma^d}ds_1-\int_0^s\frac{\beta^d\delta^{d'}}{\gamma^d}\frac{\mathrm{E}[iw_{t-1}\exp(is_1x_{t-1})\mid d_{t-1}=d,d_{t-2}=d']}{\mathrm{E}[\exp(is_1x_{t-1})\mid d_{t-1}=d,d_{t-2}=d']}ds_1\right] \\
= & \exp\left[\int_0^s\frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^d}x_t-\alpha^d\delta^{d'}-\beta^d\delta^{d'}w_{t-1})\exp\left(is_1x_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\gamma^d\,\mathrm{E}\left[\exp\left(is_1x_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]}ds_1\right].
\end{aligned}
$$

From the proxy model and the independence assumption for ε_t,

$$
\mathrm{E}\left[\exp\left(isx_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]=\mathrm{E}\left[\exp\left(is\delta^{d'}x^*_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]\mathrm{E}\left[\exp\left(is\varepsilon^{d'}_{t-1}\right)\right].
$$

We then obtain the following result using any d.

$$
\begin{aligned}
\mathrm{E}\left[\exp\left(is\varepsilon^{d'}_{t-1}\right)\right] &= \frac{\mathrm{E}\left[\exp\left(isx_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is\delta^{d'}x^*_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]} \\
&= \frac{\mathrm{E}\left[\exp\left(isx_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\exp\left[\int_0^s\frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^d}x_t-\alpha^d\delta^{d'}-\beta^d\delta^{d'}w_{t-1})\exp(is_1x_{t-1})|d_{t-1}=d,d_{t-2}=d'\right]}{\gamma^d\,\mathrm{E}[\exp(is_1x_{t-1})|d_{t-1}=d,d_{t-2}=d']}ds_1\right]}.
\end{aligned}
$$

This argument holds for all t so that we can identify $f\left(\varepsilon^d_t\right)$ for each d with

$$
\mathrm{E}\left[\exp\left(is\varepsilon^d_t\right)\right] = \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_t=d',d_{t-1}=d\right]}{\exp\left[\int_0^s\frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}}x_{t+1}-\alpha^{d'}\delta^d-\beta^{d'}\delta^dw_t)\exp(is_1x_t)|d_t=d',d_{t-1}=d\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1x_t)|d_t=d',d_{t-1}=d]}ds_1\right]}. \tag{B.9}
$$

using any d'.

In order to identify $f\left(\eta^d_t\right)$ for each d, consider

$$
\begin{aligned}
& \mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1}=d,d_{t-2}=d'\right]\mathrm{E}\left[\exp\left(is\gamma^d\frac{\delta^d}{\delta^{d'}}\varepsilon^{d'}_{t-1}\right)\right] \\
= & \mathrm{E}\left[\exp\left(is(\alpha^d\delta^d+\beta^d\delta^dw_{t-1}+\gamma^d\frac{\delta^d}{\delta^{d'}}x_{t-1})\right)|d_{t-1}=d,d_{t-2}=d'\right] \\
& \times\mathrm{E}\left[\exp\left(is\delta^d\eta^d_t\right)\right]\mathrm{E}\left[\exp\left(is\varepsilon^d_t\right)\right]
\end{aligned}
$$

by the independence assumptions for η_t^d and ε_t^d. Therefore,

$$\mathrm{E}\left[\exp\left(is\delta^d\eta_t^d\right)\right] = \frac{\mathrm{E}\left[\exp\left(isx_t\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d\delta^d+\beta^d\delta^d w_{t-1}+\gamma^d\frac{\delta^d}{\delta^{d'}}x_{t-1})\right)|d_{t-1}=d,d_{t-2}=d'\right]} \times\frac{\mathrm{E}\left[\exp\left(is\gamma^d\frac{\delta^d}{\delta^{d'}}\varepsilon_{t-1}^{d'}\right)\right]}{\mathrm{E}\left[\exp\left(is\varepsilon_t^d\right)\right]}$$

and the characteristic function of η_t^d can be expressed by

$$\begin{aligned}\mathrm{E}\left[\exp\left(is\eta_t^d\right)\right] &= \frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d}x_t\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d\frac{1}{\delta^{d'}}x_{t-1})\right)|d_{t-1}=d,d_{t-2}=d'\right]} \times\frac{1}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d}\varepsilon_t^d\right)\right]\mathrm{E}\left[\exp\left(-is\gamma^d\frac{1}{\delta^{d'}}\varepsilon_{t-1}^{d'}\right)\right]}\\
&= \frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d}x_t\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d\frac{1}{\delta^{d'}}x_{t-1})\right)|d_{t-1}=d,d_{t-2}=d'\right]}\\
&\quad\times\frac{\exp\left[\int_0^{s/\delta^d}\frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}}x_{t+1}-\alpha^{d'}\delta^d-\beta^{d'}\delta^d w_t)\exp(is_1x_t)|d_t=d',d_{t-1}=d\right]}{\gamma^{d'}\mathrm{E}[\exp(is_1x_t)|d_t=d',d_{t-1}=d]}ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d}x_t\right)|d_t=d',d_{t-1}=d\right]}\\
&\quad\times\frac{\mathrm{E}\left[\exp\left(is\gamma^d\frac{1}{\delta^{d'}}x_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\exp\left[\int_0^{s\gamma^d/\delta^{d'}}\frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^d}x_t-\alpha^d\delta^{d'}-\beta^d\delta^{d'}w_{t-1})\exp(is_1x_{t-1})|d_{t-1}=d,d_{t-2}=d'\right]}{\gamma^d\mathrm{E}[\exp(is_1x_t)|d_{t-1}=d,d_{t-2}=d']}ds_1\right]}\end{aligned}$$

by the formula (B.9). We can then identify $f_{\eta_t^d}$ by

$$f_{\eta_t^d}(\eta)=\left(\mathcal{F}\phi_{\eta_t^d}\right)(\eta)\qquad\text{for all }\eta,$$

where the characteristic function $\phi_{\eta_t^d}$ is given by

$$\begin{aligned}\phi_{\eta_t^d}(s) &= \frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d}x_t\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d+\beta^d w_{t-1}+\gamma^d\frac{1}{\delta^{d'}}x_{t-1})\right)|d_{t-1}=d,d_{t-2}=d'\right]}\\
&\quad\times\frac{\exp\left[\int_0^{s/\delta^d}\frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}}x_{t+1}-\alpha^{d'}\delta^d-\beta^{d'}\delta^d w_t)\exp(is_1x_t)|d_t=d',d_{t-1}=d\right]}{\gamma^{d'}\mathrm{E}[\exp(is_1x_t)|d_t=d',d_{t-1}=d]}ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d}x_t\right)|d_t=d',d_{t-1}=d\right]}\\
&\quad\times\frac{\mathrm{E}\left[\exp\left(is\gamma^d\frac{1}{\delta^{d'}}x_{t-1}\right)|d_{t-1}=d,d_{t-2}=d'\right]}{\exp\left[\int_0^{s\gamma^d/\delta^{d'}}\frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^d}x_t-\alpha^d\delta^{d'}-\beta^d\delta^{d'}w_{t-1})\exp(is_1x_{t-1})|d_{t-1}=d,d_{t-2}=d'\right]}{\gamma^d\mathrm{E}[\exp(is_1x_t)|d_{t-1}=d,d_{t-2}=d']}ds_1\right]}.\end{aligned}$$

We can use this identified density in turn to identify the transition rule $f\left(x_t^*|d_{t-1}, w_{t-1}, x_{t-1}^*\right)$ with

$$f\left(x_t^*|d_{t-1}, x_{t-1}, x_{t-1}^*\right) = \sum_d \mathbb{1}\{d_{t-1} = d\} f_{\eta_t^d}\left(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*\right).$$

In summary, we obtain the closed-form expression

$$\begin{aligned}
f\left(x_t^*|d_{t-1}, x_{t-1}, x_{t-1}^*\right) &= \sum_d \mathbb{1}\{d_{t-1} = d\} \left(\mathcal{F}\phi_{\eta_t^d}\right)\left(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*\right) \\
&= \sum_d \frac{\mathbb{1}\{d_{t-1} = d\}}{2\pi} \int \exp\left(-is(x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*)\right) \times \\
&\quad \frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d} x_t\right)|d_{t-1} = d, d_{t-2} = d'\right]}{\mathrm{E}\left[\exp\left(is(\alpha^d + \beta^d w_{t-1} + \gamma^d \frac{1}{\delta^{d'}} x_{t-1})\right)|d_{t-1} = d, d_{t-2} = d'\right]} \times \\
&\quad \frac{\exp\left[\int_0^{s/\delta^d} \frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}} x_{t+1} - \alpha^{d'}\delta^d - \beta^{d'}\delta^d w_t)\exp(is_1 x_t)|d_t = d', d_{t-1} = d\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1 x_t)|d_t = d', d_{t-1} = d]} ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^d} x_t\right)|d_t = d', d_{t-1} = d\right]} \times \\
&\quad \frac{\mathrm{E}\left[\exp\left(is\gamma^d \frac{1}{\delta^{d'}} x_{t-1}\right)|d_{t-1} = d, d_{t-2} = d'\right]}{\exp\left[\int_0^{s\gamma^d/\delta^{d'}} \frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^d} x_t - \alpha^d\delta^{d'} - \beta^d\delta^{d'} w_{t-1})\exp(is_1 x_{t-1})|d_{t-1} = d, d_{t-2} = d'\right]}{\gamma^d\,\mathrm{E}[\exp(is_1 x_t)|d_{t-1} = d, d_{t-2} = d']} ds_1\right]}
\end{aligned}$$

using any d'. This completes Step 1.

Step 2: Closed-form identification of the proxy model $f(x_t \mid d_{t-1}, x_t^*)$: Given (B.9), we can write the density of ε_t^d by

$$f_{\varepsilon_t^d}(\varepsilon) = \left(\mathcal{F}\phi_{\varepsilon_t^d}\right)(\varepsilon) \qquad \text{for all } \varepsilon,$$

where the characteristic function $\phi_{\varepsilon_t^d}$ is defined by (B.9) as

$$\phi_{\varepsilon_t^d}(s) = \frac{\mathrm{E}\left[\exp(isx_t)|d_t = d', d_{t-1} = d\right]}{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}} x_{t+1} - \alpha^{d'}\delta^d - \beta^{d'}\delta^d w_t)\exp(is_1 x_t)|d_t = d', d_{t-1} = d\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1 x_t)|d_t = d', d_{t-1} = d]} ds_1\right]}.$$

Provided this identified density of ε_t^d, we nonparametrically identify the proxy model

$$f(x_t \mid d_{t-1} = d, x_t^*) = f_{\varepsilon_t^d|d_{t-1}=d}(x_t - \delta^d x_t^*) = f_{\varepsilon_t^d}(x_t - \delta^d x_t^*)$$

by the independence assumption for ε_t^d. In summary, we obtain the closed-form expression

$$
\begin{aligned}
f(x_t \mid d_{t-1}, x_t^*) &= \sum_d \mathbb{1}\{d_{t-1} = d\} \left(\mathcal{F}\phi_{\varepsilon_t^d}\right)(x_t - \delta^d x_t^*) \\
&= \sum_d \frac{\mathbb{1}\{d_{t-1} = d\}}{2\pi} \int \frac{\exp\left(-is(x_t - \delta^d x_t^*)\right) \cdot \mathrm{E}\left[\exp\left(isx_t\right)|d_t = d', d_{t-1} = d\right]}{\exp\left[\int_0^s \frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}} x_{t+1} - \alpha^{d'}\delta^d - \beta^{d'}\delta^d w_t) \exp(is_1 x_t)|d_t = d', d_{t-1} = d\right]}{\gamma^{d'}\,\mathrm{E}[\exp(is_1 x_t)|d_t = d', d_{t-1} = d]} ds_1\right]} ds
\end{aligned}
$$

using any d'. This completes Step 2.

Step 3: Closed-form identification of the transition rule $f\left(w_t|d_{t-1}, w_{t-1}, x_{t-1}^*\right)$: Consider the joint density expressed by the convolution integral

$$
f\left(x_{t-1}, w_t \mid d_{t-1}, w_{t-1}, d_{t-2} = d\right) = \int f_{\varepsilon_{t-1}^d}\left(x_{t-1} - \delta^d x_{t-1}^*\right) f\left(x_{t-1}^*, w_t \mid d_{t-1}, w_{t-1}, d_{t-2} = d\right) dx_{t-1}^*
$$

We can thus obtain a closed-form expression of $f\left(x_{t-1}^*, w_t \mid d_{t-1}, w_{t-1}, d_{t-2}\right)$ by the deconvolution. To see this, observe

$$
\begin{aligned}
&\mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 w_t\right)|d_{t-1}, w_{t-1}, d_{t-2} = d\right] \\
=\ &\mathrm{E}\left[\exp\left(is_1 \delta^d x_{t-1}^* + is_1 \varepsilon_{t-1}^d + is_2 w_t\right)|d_{t-1}, w_{t-1}, d_{t-2} = d\right] \\
=\ &\mathrm{E}\left[\exp\left(is_1 \delta^d x_{t-1}^* + is_2 w_t\right)|d_{t-1}, w_{t-1}, d_{t-2} = d\right] \mathrm{E}\left[\exp\left(is_1 \varepsilon_{t-1}^d\right)\right]
\end{aligned}
$$

by the independence assumption for ε_t^d, and so

$$
\begin{aligned}
\mathrm{E}\left[\exp\left(is_1 \delta^d x_{t-1}^* + is_2 w_t\right)|d_{t-1}, w_{t-1}, d_{t-2} = d\right] &= \frac{\mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 w_t\right)|d_{t-1}, w_{t-1}, d_{t-2} = d\right]}{\mathrm{E}\left[\exp\left(is_1 \varepsilon_{t-1}^d\right)\right]} \\
&= \mathrm{E}\left[\exp\left(is_1 x_{t-1} + is_2 w_t\right)|d_{t-1}, w_{t-1}, d_{t-2} = d\right] \\
&\quad \times \frac{\exp\left[\int_0^{s_1} \frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}} x_t - \alpha^{d'}\delta^d - \beta^{d'}\delta^d w_{t-1}) \exp\left(is_1' x_{t-1}\right)|d_{t-1} = d', d_{t-2} = d\right]}{\gamma^{d'}\,\mathrm{E}\left[\exp\left(is_1' x_{t-1}\right)|d_{t-1} = d', d_{t-2} = d\right]} ds_1'\right]}{\mathrm{E}\left[\exp\left(is_1 x_{t-1}\right)|d_{t-1} = d', d_{t-2} = d\right]}
\end{aligned}
$$

follows with any choice of d'. Rescaling s_1 yields

$$
\begin{aligned}
& \mathrm{E}\left[\exp\left(is_1 x^*_{t-1} + is_2 w_t\right) | d_{t-1}, w_{t-1}, d_{t-2} = d\right] \\
= \; & \mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^d} x_{t-1} + is_2 w_t\right) | d_{t-1}, w_{t-1}, d_{t-2} = d\right] \times \\
& \frac{\exp\left[\int_0^{s_1/\delta^d} \frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}} x_t - \alpha^{d'}\delta^d - \beta^{d'}\delta^d w_{t-1}) \exp\left(is_1' x_{t-1}\right) | d_{t-1}=d', d_{t-2}=d\right]}{\gamma^{d'} \mathrm{E}\left[\exp\left(is_1' x_{t-1}\right) | d_{t-1}=d', d_{t-2}=d\right]} ds_1'\right]}{\mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^d} x_{t-1}\right) | d_{t-1} = d', d_{t-2} = d\right]}.
\end{aligned}
$$

We can then express the conditional density as

$$
f\left(x^*_{t-1}, w_t | d_{t-1}, w_{t-1}, d_{t-2} = d\right) = \left(\mathcal{F}_2 \phi_{x^*_{t-1}, w_t | d_{t-1}, w_{t-1}, d_{t-2}=d}\right)\left(w_t, x^*_{t-1}\right)
$$

where the characteristic function is defined by

$$
\begin{aligned}
\phi_{w_t, x^*_{t-1} | d_{t-1}, w_{t-1}, d_{t-2}=d}(s_1, s_2) \; = \; & \mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^d} x_{t-1} + is_2 w_t\right) | d_{t-1}, w_{t-1}, d_{t-2} = d\right] \times \\
& \frac{\exp\left[\int_0^{s_1/\delta^d} \frac{\mathrm{E}\left[i(\frac{\delta^d}{\delta^{d'}} x_t - \alpha^{d'}\delta^d - \beta^{d'}\delta^d w_{t-1}) \exp\left(is_1' x_{t-1}\right) | d_{t-1}=d', d_{t-2}=d\right]}{\gamma^{d'} \mathrm{E}\left[\exp\left(is_1' x_{t-1}\right) | d_{t-1}=d', d_{t-2}=d\right]} ds_1'\right]}{\mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^d} x_{t-1}\right) | d_{t-1} = d', d_{t-2} = d\right]}.
\end{aligned}
$$

Using this conditional density, we nonparametrically identify the transition rule

$$
f\left(w_t | d_{t-1}, w_{t-1}, x^*_{t-1}\right) = \frac{\sum_d f\left(x^*_{t-1}, w_t | d_{t-1}, w_{t-1}, d_{t-2} = d\right) \Pr(d_{t-2} = d \mid d_{t-1}, w_{t-1})}{\int \sum_d f\left(x^*_{t-1}, w_t | d_{t-1}, w_{t-1}, d_{t-2} = d\right) \Pr(d_{t-2} = d \mid d_{t-1}, w_{t-1}) dw_t}.
$$

In summary, we obtain the closed-form expression

$$
\begin{aligned}
f\left(w_t|d_{t-1}, w_{t-1}, x^*_{t-1}\right) \;&=\; \sum_d \mathbb{1}\{d_{t-1}=d\} \times \\
&\frac{\sum_{d'} \left(\mathcal{F}_2 \phi_{x^*_{t-1}, w_t | d_{t-1}=d, w_{t-1}, d_{t-2}=d'}\right)(w_t, x^*_{t-1}) \cdot \Pr(d_{t-2}=d' \mid d_{t-1}=d, w_{t-1})}{\int \sum_{d'} \left(\mathcal{F}_2 \phi_{x^*_{t-1}, w_t | d_{t-1}=d, w_{t-1}, d_{t-2}=d'}\right)(w_t, x^*_{t-1}) \cdot \Pr(d_{t-2}=d' \mid d_{t-1}=d, w_{t-1}) dw_t} \\
= \sum_d \mathbb{1}\{d_{t-1}=d\} &\left\{ \sum_{d'} \Pr(d_{t-2}=d' \mid d_{t-1}=d, w_{t-1}) \int\int \exp\left(-is_1 w_t - is_2 x^*_{t-1}\right) \times \right. \\
&\frac{\mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^{d'}} x_{t-1} + is_2 w_t\right) | d_{t-1}=d, w_{t-1}, d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^{d'}} x_{t-1}\right) | d_{t-1}=d'', d_{t-2}=d'\right]} \times \\
&\left. \exp\left[\int_0^{s_1/\delta^{d'}} \frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^{d''}} x_t - \alpha^{d''}\delta^{d'} - \beta^{d''}\delta^{d'} w_{t-1}) \exp\left(is_1' x_{t-1}\right) | d_{t-1}=d'', d_{t-2}=d'\right]}{\gamma^{d''}\, \mathrm{E}\left[\exp\left(is_1' x_{t-1}\right) | d_{t-1}=d'', d_{t-2}=d'\right]} ds_1'\right] ds_1 ds_2 \right\} \Big/ \\
&\left\{ \sum_{d'} \int \Pr(d_{t-2}=d' \mid d_{t-1}=d, w_{t-1}) \int\int \exp\left(-is_1 w_t - is_2 x^*_{t-1}\right) \times \right. \\
&\frac{\mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^{d'}} x_{t-1} + is_2 w_t\right) | d_{t-1}=d, w_{t-1}, d_{t-2}=d'\right]}{\mathrm{E}\left[\exp\left(is_1 \frac{1}{\delta^{d'}} x_{t-1}\right) | d_{t-1}=d'', d_{t-2}=d'\right]} \times \\
&\left. \exp\left[\int_0^{s_1/\delta^{d'}} \frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^{d''}} x_t - \alpha^{d''}\delta^{d'} - \beta^{d''}\delta^{d'} w_{t-1}) \exp\left(is_1' x_{t-1}\right) | d_{t-1}=d'', d_{t-2}=d'\right]}{\gamma^{d''}\, \mathrm{E}\left[\exp\left(is_1' x_{t-1}\right) | d_{t-1}=d'', d_{t-2}=d'\right]} ds_1'\right] ds_1 ds_2 dw_t \right\}
\end{aligned}
$$

using any d' and d'' This completes Step 3.

Step 4: Closed-form identification of the CCP $f\left(d_t|w_t, x^*_t\right)$: Note that we have

$$
\begin{aligned}
\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(isx_t\right)|w_t, d_{t-1}=d'\right] =\; &\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(is\delta^{d'}x^*_t + is\varepsilon^{d'}_t\right)|w_t, d_{t-1}=d'\right] \\
=\; &\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(is\delta^{d'}x^*_t\right)|w_t, d_{t-1}=d'\right]\mathrm{E}\left[\exp\left(is\varepsilon^{d'}_t\right)\right] \\
=\; \mathrm{E}\left[\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t, x^*_t, d_{t-1}=d'\right]\exp\left(is\delta^{d'}x^*_t\right)|w_t, d_{t-1}=d'\right]&\mathrm{E}\left[\exp\left(is\varepsilon^{d'}_t\right)\right]
\end{aligned}
$$

by the independence assumption for $\varepsilon^{d'}_t$ and the law of iterated expectations. Therefore,

$$
\begin{aligned}
&\frac{\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(isx_t\right)|w_t, d_{t-1}=d'\right]}{\mathrm{E}\left[\exp\left(is\varepsilon^{d'}_t\right)\right]} \\
=\; &\mathrm{E}\left[\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t, x^*_t, d_{t-1}=d'\right]\exp\left(is\delta^{d'}x^*_t\right)|w_t, d_{t-1}=d'\right] \\
=\; &\int \exp\left(is\delta^{d'}x^*_t\right)\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t, x^*_t, d_{t-1}=d'\right] f\left(x^*_t|w_t, d_{t-1}=d'\right) dx^*_t
\end{aligned}
$$

and rescaling s yields

$$\begin{aligned}
&\frac{\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t,d_{t-1}=d'\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}\varepsilon_t^{d'}\right)\right]}\\
=&\int\exp\left(isx_t^*\right)\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*,d_{t-1}=d'\right]f\left(x_t^*|w_t,d_{t-1}=d'\right)dx_t^*
\end{aligned}$$

This is the Fourier inversion of $\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*,d_{t-1}=d'\right]f\left(x_t^*|w_t,d_{t-1}=d'\right)$. On the other hand, the Fourier inversion of $f\left(x_t^*|w_t,d_{t-1}\right)$ can be found as

$$\mathrm{E}\left[\exp\left(isx_t^*\right)|w_t,d_{t-1}=d'\right]=\frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t,d_{t-1}=d'\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}\varepsilon_t^{d'}\right)\right]}.$$

Therefore, we find the closed-form expression for CCP $f\left(d_t|w_t,x_t^*\right)$ as follows.

$$\begin{aligned}
\Pr\left(d_t=d|w_t,x_t^*\right) &= \sum_{d'}\Pr\left(d_t=d|w_t,x_t^*,d_{t-1}=d'\right)\Pr\left(d_{t-1}=d'\mid w_t,x_t^*\right)\\
&= \sum_{d'}\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*,d_{t-1}=d'\right]\Pr\left(d_{t-1}=d'\mid w_t,x_t^*\right)\\
&= \sum_{d'}\frac{\mathrm{E}\left[\mathbb{1}\{d_t=d\}|w_t,x_t^*,d_{t-1}=d'\right]f\left(x_t^*|w_t,d_{t-1}=d'\right)}{f\left(x_t^*|w_t,d_{t-1}=d'\right)}\Pr\left(d_{t-1}=d'\mid w_t,x_t^*\right)\\
&= \sum_{d'}\frac{\left(\mathcal{F}\phi_{(d)x_t^*|w_t(d')}\right)\left(x_t^*\right)}{\left(\mathcal{F}\phi_{x_t^*|w_t(d')}\right)\left(x_t^*\right)}\Pr\left(d_{t-1}=d'\mid w_t,x_t^*\right)
\end{aligned}$$

where the characteristic functions are defined by

$$\begin{aligned}
\phi_{(d)x_t^*|w_t(d')}(s) &= \frac{\mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t,d_{t-1}=d'\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}\varepsilon_t^{d'}\right)\right]}\\
&= \mathrm{E}\left[\mathbb{1}\{d_t=d\}\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t,d_{t-1}=d'\right]\\
&\quad\times\frac{\exp\left[\int_0^{s/\delta^{d'}}\frac{\mathrm{E}\left[i\left(\frac{\delta^{d'}}{\delta^{d''}}x_{t+1}-\alpha^{d''}\delta^{d'}-\beta^{d''}\delta^{d'}w_t\right)\exp(is_1x_t)|d_t=d'',d_{t-1}=d'\right]}{\gamma^{d''}\,\mathrm{E}[\exp(is_1x_t)|d_t=d'',d_{t-1}=d']}ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|d_t=d',d_{t-1}=d''\right]}
\end{aligned}$$

and

$$\begin{aligned}
\phi_{x_t^*|w_t(d')}(s) &= \frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}\varepsilon_t^{d'}\right)\right]}\\
=&\frac{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t\right]\cdot\exp\left[\int_0^{s/\delta^{d'}}\frac{\mathrm{E}\left[i\left(\frac{\delta^{d'}}{\delta^{d''}}x_{t+1}-\alpha^{d''}\delta^{d'}-\beta^{d''}\delta^{d'}w_t\right)\exp(is_1x_t)|d_t=d'',d_{t-1}=d'\right]}{\gamma^{d''}\,\mathrm{E}[\exp(is_1x_t)|d_t=d'',d_{t-1}=d']}ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|d_t=d',d_{t-1}=d''\right]}
\end{aligned}$$

by (B.9) using any d''. In summary, we obtain the closed-form expression

$$\begin{aligned}
\Pr\left(d_t = d | w_t, x_t^*\right) &= \sum_{d'} \frac{\left(\mathcal{F}\phi_{(d)x_t^*|w_t(d')}\right)(x_t^*)}{\left(\mathcal{F}\phi_{x_t^*|w_t(d')}\right)(x_t^*)} \Pr\left(d_{t-1} = d' \mid w_t, x_t^*\right) \\
&= \sum_{d'} \Pr\left(d_{t-1} = d' \mid w_t, x_t^*\right) \int \exp\left(-isx_t^*\right) \times \\
& \mathrm{E}\left[\mathbb{1}\{d_t = d\} \exp\left(is\frac{1}{\delta^{d'}}x_t\right) |w_t, d_{t-1} = d'\right] \times \\
& \frac{\exp\left[\int_0^{s/\delta^{d'}} \frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^{d''}}x_{t+1} - \alpha^{d''}\delta^{d'} - \beta^{d''}\delta^{d'}w_t)\exp(is_1 x_t)|d_t = d'', d_{t-1} = d'\right]}{\gamma^{d''}\mathrm{E}[\exp(is_1 x_t)|d_t = d'', d_{t-1} = d']} ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|d_t = d', d_{t-1} = d''\right]} ds \Bigg/ \\
& \int \exp\left(-isx_t^*\right) \cdot \mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|w_t\right] \times \\
& \frac{\exp\left[\int_0^{s/\delta^{d'}} \frac{\mathrm{E}\left[i(\frac{\delta^{d'}}{\delta^{d''}}x_{t+1} - \alpha^{d''}\delta^{d'} - \beta^{d''}\delta^{d'}w_t)\exp(is_1 x_t)|d_t = d'', d_{t-1} = d'\right]}{\gamma^{d''}\mathrm{E}[\exp(is_1 x_t)|d_t = d'', d_{t-1} = d']} ds_1\right]}{\mathrm{E}\left[\exp\left(is\frac{1}{\delta^{d'}}x_t\right)|d_t = d', d_{t-1} = d''\right]} ds.
\end{aligned}$$

This completes Step 4. □