

Dynamic Discrete Choice Models with Proxies for Unobserved Technologies

Yingyao Hu* and Yuya Sasaki

Department of Economics, Johns Hopkins University

August 5, 2014

Abstract

Firms make forward-looking decisions based on technologies. The true technological states are not observed by econometricians, but its proxies are often available. For dynamic discrete choice models of forward-looking firms where a continuous state variable is unobserved but its proxy is available, we derive closed-form identification of the dynamic structure by explicitly solving integral equations without relying on the completeness assumption. We use this method to estimate the structure of forward-looking firms making exit decisions and to deduce the values of exit. The exit values are higher for capital-intensive industries and for firms that own more physical properties.

Keywords: dynamic discrete choice, unobserved state variable, exit, exit value, proxy, technology

*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

Behaviors of firms making forward-looking decisions based on unobserved state variables are of interest in economic researches. Even if the true states are unobserved by econometricians, we often have access to or can structurally construct proxy variables. To estimate the dynamic discrete choice models of forward-looking firms, can we 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 biases the estimates of structural parameters, even if the proxy incurs only a classical error. In this light, we develop methods to identify dynamic discrete choice structural models when a proxy for an unobserved continuous state variable is available.

Specifically, suppose that firm j at time t makes exit decisions $d_{j,t}$ based on its technology $x_{j,t}^*$. While econometricians do not observe the technology $x_{j,t}^*$, structures and/or data often allow us to construct a proxy for the unobserved technology. Suppose that we obtain a proxy $x_{j,t} = x_{j,t}^* + \varepsilon_{j,t}$ for the unobserved technology $x_{j,t}^*$ up to idiosyncratic shocks. It is known that identification of the structural parameters of forward-looking firms 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 constructed proxies $x_{j,t}$ in stead of the unobserved true states $x_{j,t}^*$. Indeed, dynamic discrete choice models with unobservables are studied in the literature,¹ but no preceding work handles continuous unobservables like production technologies. Our methods allow for continuously distributed unobservables at the expense of the requirement of proxy variables for the unobservables. Our study of the agents making exit/survival decisions based on dynamically

¹See for example Aguirregabiria and Mira (2007), Kasahara and Shimotsu (2009), Arcidiacono and Miller (2011), and Hu and Shum (2012).

evolving state variables is related to a number of papers (Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995; Abbring and Campbell, 2004; Asplund and Nocke, 2006; Abbring and Heckman, 2007; Heckman and Navarro, 2007; Foster, Haltiwanger and Syverson, 2008). To our best knowledge, none of the existing papers proposes identification of structural forward-looking exit decisions based on unobserved dynamic state variables with or without proxies.

The use of proxy variables in dynamic structural models is related to Cunha and Heckman (2008), Cunha, Heckman, and Schennach (2010), and Todd and Wolpin (2012). Since we analyze the structure of forward-looking firms, however, we follow a distinct approach. In the first step, we identify of 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 (2013) 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 stable 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 apply our methods and study the forward-looking structure of firms that make exit decisions based on unobserved production technologies, and estimate the option value of exit for each industry. Section 6 summarizes the paper. Technical proofs are included in the Appendix.

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' decisions based on unobserved technologies. Formal identification and estimation results behind this informal overview follow in Sections 3 and 4.

Firms with lower levels of production technologies produce less values added even at the optimal choice of inputs and even on the optimal investment paths, and may exit with a higher probability than firms with higher levels of production technologies. Let $d_{j,t} = 1$ indicate the decision of a firm to stay, and let $d_{j,t} = 0$ indicate the decision to exit. Firms choose $d_{j,t}$ given its technological level $x_{j,t}^*$, and based on their knowledge of the stochastic law of motion of $x_{j,t}^*$. Suppose that the technological state $x_{j,t}^*$ of a firm evolves according to the first-order process

$$x_{j,t}^* = \alpha_t + \gamma_t x_{j,t-1}^* + \eta_{j,t}. \quad (2.1)$$

As a reduced form of the underlying structural production process, a firm with its technological level $x_{j,t}^*$ is assumed to receive the current payoff of the affine form $\theta_0 + \theta_1 x_{j,t}^* + \omega_{j,t}^d$ if it is in the market, where $\omega_{j,t}^d$ is the choice-specific private shock. 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

$$\text{With stay } (d_{j,t} = 1) : \quad v_1(x_{j,t}^*) = \theta_0 + \theta_1 x_{j,t}^* + \omega_{j,t}^1 + \text{E} [\rho V(x_{j,t+1}^*; \theta) \mid x_{j,t}^*]$$

$$\text{With exit } (d_{j,t} = 0) : \quad v_0(x_{j,t}^*) = \theta_0 + \theta_1 x_{j,t}^* + \theta_2 + \omega_{j,t}^1$$

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

The first step toward estimation of the structural parameters is to construct a proxy variable $x_{j,t}$ for the unobserved technology $x_{j,t}^*$. This stage can take one of various routes. By using observed data and/or structural restrictions, construct an observed proxy variable $x_{j,t}$ for the true state variable $x_{j,t}^*$ up to 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 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 $\varepsilon_{j,t}$. Its characteristic function can be estimated by the formula

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

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, we can

estimate the CCP by the formula

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

where ϕ_K is the Fourier transform of a kernel function K and h is a bandwidth parameter. For example, $\phi_K(sh) = e^{-\frac{1}{2}s^2h^2}$ if the normal kernel is used. 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

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

As before, ϕ_K is the Fourier transform of a kernel function K and h is a bandwidth parameter.

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 $\theta = (\theta_0, \theta_1, \theta_2)$. If we follow that standard assumption that the choice-specific private shocks independently follow the standard Gumbel (Type I Extreme Value) distribution, then we obtain the restriction

$$\ln p_t(x_t^*) - \ln(1 - p_t(x_t^*)) = v_1(x_t^*) - v_0(x_t^*) = \text{E}[\rho V(x_{t+1}^*; \theta) \mid x_t^*] - \theta_2,$$

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

$$\begin{aligned} \text{E}[\rho V(x_{t+1}^*; \theta) \mid x_t^*] &= \text{E} \left[\sum_{s=t+1}^{\infty} \rho^{s-t} (\theta_0 + \theta_1 x_s^* + \theta_2 (1 - p_s(x_s^*))) + \bar{\omega} \right. \\ &\quad \left. - (1 - p_s(x_s^*)) \log(1 - p_s(x_s^*)) - p_s(x_s^*) \log p_s(x_s^*) \right] \left(\prod_{s'=t+1}^{s-1} p_{s'}(x_{s'}^*) \right) \Bigg| x_t^*, \end{aligned}$$

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

$$\begin{aligned} \ln \hat{p}_t(x_t^*) - \ln(1 - \hat{p}_t(x_t^*)) &= \widehat{\mathbb{E}} \left[\sum_{s=t+1}^{\infty} \rho^{s-t} \left(\hat{\theta}_0 + \hat{\theta}_1 x_s^* + \hat{\theta}_2 (1 - \hat{p}_s(x_s^*)) + \bar{\omega} \right. \right. \\ &\quad \left. \left. - (1 - \hat{p}_s(x_s^*)) \log(1 - \hat{p}_s(x_s^*)) - \hat{p}_s(x_s^*) \log \hat{p}_s(x_s^*) \right) \left(\prod_{s'=t+1}^{s-1} \hat{p}_{s'}(x_{s'}^*) \right) \middle| x_t^* \right] - \hat{\theta}_2 \quad \text{for all } x_t^*, \end{aligned} \quad (2.4)$$

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, one can conduct counter-factual 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

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

Again, ϕ_K is the Fourier transform of a kernel function K and h is a bandwidth parameter.

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. Others may skip the following two theoretical sections and move directly on to Section 5, where we present an empirical application of these methods to analyze the structure of forward-looking firms making exit decisions based on production technologies.

3 Markov Components: Identification and Estimation

Our basic notations are fixed as follows. A discrete control variables, taking values in $\{0, 1, \dots, \bar{d}\}$, 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^* . In the context of the production literature, it is the technological term x_t^* in the production function $y_{j,t} = x_{j,t}^* + b_l l_{j,t} + b_w w_{j,t} + \varepsilon_{j,t}$. Finally, x_t denotes a proxy variable for x_t^* . For example, the residual $x_{j,t} := y_{j,t} - b_l l_{j,t} - b_w w_{j,t}$ can be used as a proxy in the sense that $x_{j,t} = x_{j,t}^* + \varepsilon_{j,t}$ automatically holds by the structural construction. Throughout this paper, we consider the dynamics of this list of random variables. The following subsection presents identification of the components of the dynamic law for these 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.*

In the production literature, the first-order Markov process for unobserved productivity as well as observed states and actions is commonly used as the core identifying assumption. This Markovian structure is decomposed into four modules as follows.

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

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

where the four components represent

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

Remark 1. *Depending on applications, we can alternatively specify the transition rule for the observed state variable as $f(w_t | d_{t-1}, w_{t-1}, x_t^*)$ which depends on the current unobserved state x_t^* instead of the lag x_{t-1}^* . A similar closed-form identification result follows in this case.*

In the context of the production models again, the four components of the Markov kernel can be economically interpreted as follows. 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^* . 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 . We provide a concrete example after the next assumption. Because the state variable x_t^* of interest is unit-less and unobserved, we require some restriction to tie hands of its location and scale. 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*

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

$$f(x_t|x_t^*) : \quad x_t = x_t^* + \varepsilon_t \quad (3.2)$$

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

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

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

Remark 2. *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.*

For the production models discussed earlier, these semi-parametric restrictions are interpreted as follows. In the special case of $\gamma^d = 1$, the semi-parametric model (3.1) of state transition yields super-/sub-Martingale process for the evolution of unobserved technology x_t^* depending on $\alpha^d + \beta^d w_t >$ or < 0 . In case where we consider the discrete choice d_t of investment decisions, 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 technological developments. The semi-parametric model (3.2) of the state-proxy relation is automatically valid as the proxy being the residual $x_t := y_t - b_l l_t - b_k k_t$ equals the productivity x_t^* plus the idiosyncratic shock ε_t .²

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

²While this classical error specification is valid for the specific example of production functions, it may be generally restrictive. We discuss how to relax this classical-error assumption in Section A.9 in the appendix.

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.³ 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 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 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 as we did using the production functions. Assumption 4 is empirically testable as is the common rank condition in generic econometric 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.

³This 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.

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

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

$$\begin{aligned}\mathcal{F}\phi(\xi) &= \frac{1}{2\pi} \int e^{-is\xi} \phi(s) ds && \text{for all } \phi \in L^1(\mathbb{R}) \text{ and } \xi \in \mathbb{R} \\ \mathcal{F}_2\phi(\xi_1, \xi_2) &= \frac{1}{4\pi^2} \int e^{-is_1\xi_1 - is_2\xi_2} \phi(s_1, s_2) ds_1 ds_2 && \text{for all } \phi \in L^1(\mathbb{R}^2) \text{ and } (\xi_1, \xi_2) \in \mathbb{R}^2.\end{aligned}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 firms. We present how this step works in Section 4. Before proceeding with structural estimation, we first show that these identified components of the Markov kernel can be easily estimated by their sample counterparts.

Since the CCP, $Pr(d_t = d \mid w_t, x_t^*)$, can be represented by the indicator mean regression, $E[\mathbb{1}\{d_t = d\} \mid w_t, x_t^*]$, the main identification strategy is based on the closed-form identification of mean regressions under measurement errors (Schennach, 2004; Hu and Sasaki, 2013). An alternative representation of the CCP is the CDF, $F(d_t \mid w_t, x_t^*)$, that may also be identified in a closed form using Fourier transforms in an analogous context (Wilhelm, 2013), but the CDF is a step function and closed-form estimators based on Fourier transforms are known to fail to converge at the points of discontinuity. As such, to the goal of closed-form consistent estimation, we take the route of the mean-regression representation of the CCP. Identification of the mean regression representation is also feasible in theory based on the generalized function approach (Schennach, 2007) indeed, but as Wilhelm (2013; Remark 3) remarks, this approach suffers from practical difficulties. For these reasons, we use the closed-form mean-regression estimator (Schennach, 2004; Hu and Sasaki, 2013) as our best possible identification strategy.

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.

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 denotes 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

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

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

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

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

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

While the notations may make them appear sophisticated, all these expressions are straightforward sample-counterparts of the corresponding closed-form identifying formulas provided in the previous section.

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

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

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

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

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

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

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. Firms 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 firms 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. Forward-looking firms sequentially make decisions $\{d_t\}$ so as to maximize the expected discounted sum of payoffs

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

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 . The following two subsections introduce closed-form identification and estimation of these structural parameters.

4.1 Closed-Form Identification of Structural Parameters

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 $\theta_1^0 = 0$.⁴ 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 and some calculations yield the restriction

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

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

$$\xi(\rho; w_t, x_t^*) = \ln f(1 | w_t, x_t^*) - \ln f(0 | w_t, x_t^*) + \quad (4.2)$$

$$\begin{aligned}&\sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathbb{E}[f(0 | w_s, x_s^*) \cdot \ln f(0 | w_s, x_s^*) | d_t = 1, w_t, x_t^*] + \\ &\sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathbb{E}[f(1 | w_s, x_s^*) \cdot \ln f(1 | w_s, x_s^*) | d_t = 1, w_t, x_t^*] - \\ &\sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathbb{E}[f(0 | w_s, x_s^*) \cdot \ln f(0 | w_s, x_s^*) | d_t = 0, w_t, x_t^*] - \\ &\sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathbb{E}[f(1 | w_s, x_s^*) \cdot \ln f(1 | w_s, x_s^*) | d_t = 0, w_t, x_t^*]\end{aligned}$$

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

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

⁴We 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).

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

for each $d \in \{0, 1\}$. See Section A.3 in the appendix for derivation of (4.1)–(4.5).

In the context of their models, 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

$$\begin{aligned}Q(\rho, \theta; w_t, x_t^*) &= \xi(\rho; w_t, x_t^*) - \theta_0^0 \cdot \xi_0^0(\rho; w_t, x_t^*) + \theta_0^w \cdot \xi_0^w(\rho; w_t, x_t^*) \\ &\quad - \theta_1^w \cdot \xi_1^w(\rho; w_t, x_t^*) - \theta_0^x \cdot \xi_0^x(\rho; w_t, x_t^*) - \theta_1^x \cdot \xi_1^x(\rho; w_t, x_t^*)\end{aligned}$$

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

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

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),⁵ 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

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

⁵This rate is generally non-identifiable together with the payoffs (Rust, 1994; Magnac and Thesmar, 2002).

provided that the following condition is satisfied.

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

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 $E[\zeta(w_s, x_s^*) \mid d_t, w_t, x_t^*]$ 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

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

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

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

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 3. *We have left unspecified the measure 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 A.4 in the appendix, 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.*

4.2 Closed-Form Estimation of Structural Parameters

The closed-form identifying formulas obtained at the population level in Section 4.1 can be directly translated into sample counterparts to develop a closed-form estimator of structural

parameters. Given Corollary 1 and Remark 3, we propose the following estimator.

$$\hat{\theta} = \left[\sum_{j=1}^N \sum_{t=1}^{T-1} \frac{\int \hat{R}(\rho; W_{j,t}, x_t^*)' \hat{R}(\rho; W_{j,t}, x_t^*) \cdot \hat{f}(X_{j,t} | x_t^*) \cdot \hat{f}(x_t^* | W_{j,t}) dx_t^*}{\int \hat{f}(X_{j,t} | x_t^*) \cdot \hat{f}(x_t^* | W_{j,t}) dx_t^*} \right]^{-1} \left[\sum_{j=1}^N \sum_{t=1}^{T-1} \frac{\int \hat{R}(\rho; W_{j,t}, x_t^*)' \hat{\xi}(\rho; W_{j,t}, x_t^*) \cdot \hat{f}(X_{j,t} | x_t^*) \cdot \hat{f}(x_t^* | W_{j,t}) dx_t^*}{\int \hat{f}(X_{j,t} | x_t^*) \cdot \hat{f}(x_t^* | W_{j,t}) dx_t^*} \right] \quad (4.10)$$

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

First, $\hat{f}(x_t | x_t^*)$ is given by (A.5) in Section 3.2. For convenience of readers, we repeat it here:

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

Second, $\hat{f}(x_t^* | w_t)$ is given by (A.8) in Section A.4 in the appendix. We write it here too:

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

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

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

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

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

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

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 (A.2) in Section 3.2.

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

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

where the closed-form estimator $\widehat{f}(w_{t+1} \mid d_t, w_t, x_t^*)$ is given by (A.3), and the closed-form estimator $\widehat{f}(x_{t+1}^* \mid d_t, w_t, x_t^*)$ is given by (A.4). On the other hand, if $s > t + 1$, then

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

where the closed-form estimator $\hat{f}(d_t \mid w_t, x_t^*)$ is given by (A.2), the closed-form estimator $\hat{f}(w_{t+1} \mid d_t, w_t, x_t^*)$ is given by (A.3), and the closed-form estimator $\hat{f}(x_{t+1}^* \mid d_t, w_t, x_t^*)$ is given by (A.4). In summary, every component in (4.10) can be expressed explicitly by the previously obtained closed-form estimators, and hence the estimator $\hat{\theta}$ of the structural parameters is given in a closed form as well. Large sample properties for the estimator (4.10) is discussed in Section A.6 in the appendix. Monte Carlo simulations of the estimator are presented in Section A.8 in the appendix.

5 Exit on Production Technologies

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). Abbring and Campbell remark that their model violates Rust’s (1987) assumption of independent unobservables, and hence they cannot rely on the identification strategies of Hotz and Miller (1993) or other related approaches to dynamic discrete choice problems.

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 production 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 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. Following Levinsohn and Petrin, we focus on the four largest industries, food products (311), textiles (321), wood products (331) and metals (381). We implement 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 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.

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

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

Year	# Firms	# Exits	% Exits	Mean of the Proxy $x_{j,t}$		
				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.

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 estimate 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 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 is 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 .

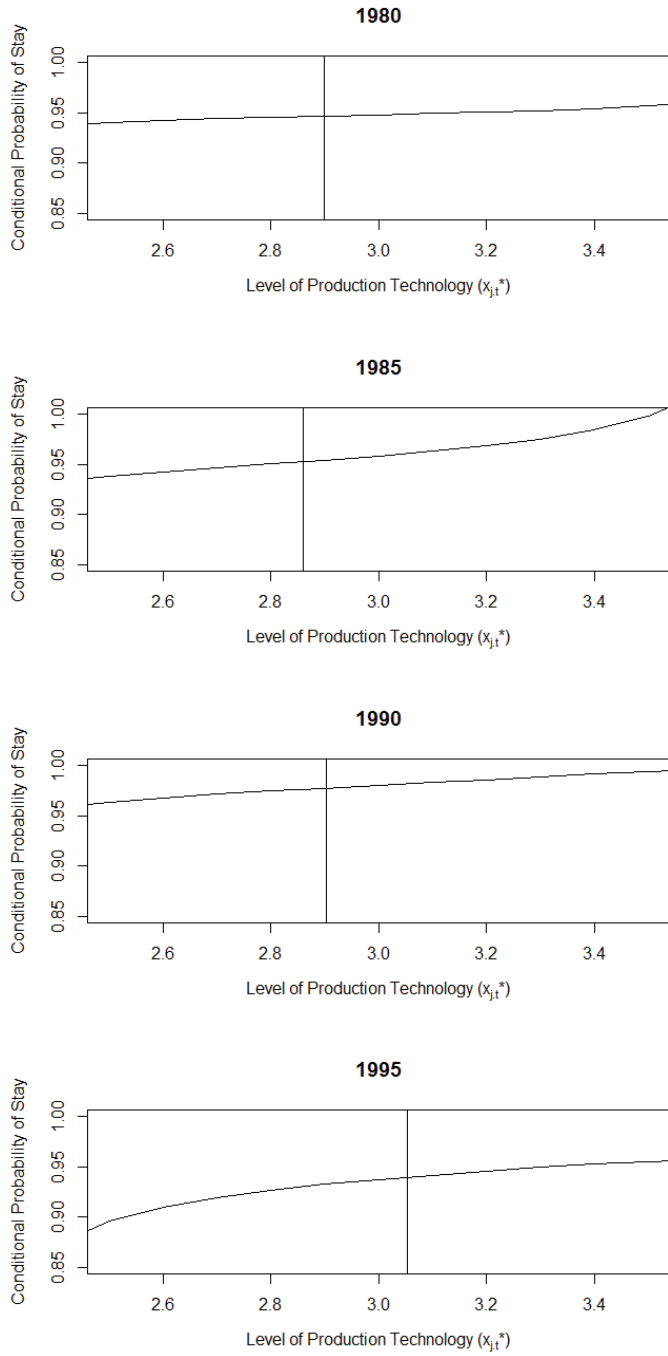


Figure 1: The estimated conditional choice probabilities of stay given the latent levels of production technology, $x_{j,t}^*$, for industry 311 in years 1980, 1985, 1990 and 1995. The vertical lines indicate the mean levels of the unobserved production technology, $x_{j,t}^*$.

Industry			Size	ρ	θ_1	θ_2
311	Food Products	18,276	0.80	1.047 (0.007)	16.491 (0.105)	15.749
321	Textiles	5,039	0.80	1.357 (0.024)	24.772 (0.434)	18.261
331	Wood Products	4,650	0.80	0.596 (0.010)	8.288 (0.126)	13.899
381	Metals	5,286	0.80	1.673 (0.026)	34.273 (0.532)	20.482
311	Food Products	18,276	0.90	0.998 (0.006)	34.553 (0.180)	34.633
321	Textiles	5,039	0.90	0.850 (0.031)	31.637 (1.083)	37.198
331	Wood Products	4,650	0.90	0.550 (0.009)	16.505 (0.225)	29.934
381	Metals	5,286	0.90	1.275 (0.030)	51.636 (1.140)	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. The numbers in parentheses show standard errors based on the calculations presented in Section A.6 in the appendix.

Subsample	Size	ρ	θ_1	θ_2	θ_2/θ_1
All firms	18,276	0.90	0.998	34.553	34.633
			(0.006)	(0.180)	
Real Estate below Average	15,652	0.90	0.905	29.897	33.027
			(0.008)	(0.233)	
Real Estate above Average	2,604	0.90	0.879	30.815	35.073
			(0.060)	(1.740)	
Machine & Furniture below Average	15,960	0.90	0.949	29.222	30.778
			(0.008)	(0.232)	
Machine & Furniture above Average	2,295	0.90	0.572	33.303	58.247
			(0.060)	(2.590)	

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. The numbers in parentheses show standard errors based on the calculations presented in Section A.6 in the appendix.

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. 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

Survival selection of firms based on their unobserved state variables has interested economic researchers. In this paper, we show that the structure of forward-looking firms can be identified and consistently estimated provided that a proxy for the unobserved state variable is available in data. Production technologies as unobservables fit in this framework due to the proxy methods developed in the production literature. 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.

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 $f(x_t^*|d_{t-1}, w_{t-1}, x_{t-1}^*)$: First,

we show the identification of the parameters and the distributions in transition of x_t^* . Since

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

we obtain the following equalities for each d :

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

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

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

Provided that the matrix on the right-hand side is non-singular, we can identify the parameters

$(\alpha^d, \beta^d, \gamma^d)$ by

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

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

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

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

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

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 $|i(\alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^*) \exp (i s_1 x_{t-1}^*)| < h(w_{t-1}, x_{t-1}^*)$ 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,

$$\begin{aligned}
\mathbb{E} [\exp (i s x_{t-1}^*) | d_{t-1} = d] &= \exp \left[\int_0^s \left[\frac{1}{\gamma^d} \frac{\partial}{\partial s_2} \ln \mathbb{E} [\exp (i s_1 x_{t-1} + i s_2 x_t) | d_{t-1} = d] \right]_{s_2=0} ds_1 \right. \\
&\quad \left. - \int_0^s \frac{i\alpha^d}{\gamma^d} ds_1 - \int_0^s \frac{\beta^d \mathbb{E} [i w_{t-1} \exp (i s_1 x_{t-1}) | d_{t-1} = d]}{\gamma^d \mathbb{E} [\exp (i s_1 x_{t-1}) | d_{t-1} = d]} ds_1 \right] \\
&= \exp \left[\int_0^s \frac{\mathbb{E} [i(x_t - \alpha^d - \beta^d w_{t-1}) \exp (i s_1 x_{t-1}) | d_{t-1} = d]}{\gamma^d \mathbb{E} [\exp (i s_1 x_{t-1}) | d_{t-1} = d]} ds_1 \right].
\end{aligned}$$

From the proxy model and the independence assumption for ε_t ,

$$\mathbb{E} [\exp (i s x_{t-1}) | d_{t-1} = d] = \mathbb{E} [\exp (i s x_{t-1}^*) | d_{t-1} = d] \mathbb{E} [\exp (i s \varepsilon_{t-1})].$$

We then obtain the following result using any d .

$$\begin{aligned} \mathbb{E} [\exp (is\varepsilon_{t-1})] &= \frac{\mathbb{E} [\exp (isx_{t-1}) | d_{t-1} = d]}{\mathbb{E} [\exp (isx_{t-1}^*) | d_{t-1} = d]} \\ &= \frac{\mathbb{E} [\exp (isx_{t-1}) | d_{t-1} = d]}{\exp \left[\int_0^s \frac{\mathbb{E} [i(x_t - \alpha^d - \beta^d w_{t-1}) \exp(is_1 x_{t-1}) | d_{t-1} = d]}{\gamma^d \mathbb{E} [\exp(is_1 x_{t-1}) | d_{t-1} = d]} ds_1 \right]}. \end{aligned}$$

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

$$\mathbb{E} [\exp (is\varepsilon_t)] = \frac{\mathbb{E} [\exp (isx_t) | d_t = d]}{\exp \left[\int_0^s \frac{\mathbb{E} [i(x_{t+1} - \alpha^d - \beta^d w_t) \exp(is_1 x_t) | d_t = d]}{\gamma^d \mathbb{E} [\exp(is_1 x_t) | d_t = d]} ds_1 \right]} \quad (\text{A.1})$$

using any d .

In order to identify $f(\eta_t^d)$ for each d , consider

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

and thus

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

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

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

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

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

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

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

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

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

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

with

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

In summary, we obtain the closed-form expression

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

using any d' . This completes Step 1.

Step 2: Closed-form identification of the proxy model $f(x_t | x_t^*)$: Given (A.1), we can

write the density of ε_t by

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

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

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

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

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

In summary, we obtain the closed-form expression

$$\begin{aligned} f(x_t | x_t^*) &= (\mathcal{F}\phi_{\varepsilon_t})(x_t - x_t^*) \\ &= \frac{1}{2\pi} \int \frac{\exp(-is(x_t - x_t^*)) \cdot \mathbb{E}[\exp(isx_t) | d_t = d]}{\exp \left[\int_0^s \frac{\mathbb{E}[i(x_{t+1} - \alpha^d - \beta^d w_t) \exp(is_1x_t) | d_t = d]}{\gamma^d \mathbb{E}[\exp(is_1x_t) | d_t = d]} ds_1 \right]} ds \end{aligned}$$

using any d . This completes Step 2.

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

$$f(x_{t-1}, w_t | d_{t-1}, w_{t-1}) = \int f_{\varepsilon_{t-1}}(x_{t-1} - x_{t-1}^*) f(x_{t-1}^*, w_t | d_{t-1}, w_{t-1}) dx_{t-1}^*$$

We can thus obtain a closed-form expression of $f(x_{t-1}^*, w_t | d_{t-1}, w_{t-1})$ by the deconvolution.

To see this, observe

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

by the independence assumption for ε_t , and so

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

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

$$(\mathcal{F}_2\phi)(\xi_1, \xi_2) = \frac{1}{4\pi^2} \int \int e^{-is_1\xi_1 - is_2\xi_2} \phi(s_1, s_2) ds_1 ds_2 \quad \text{for all } \phi \in L^1(\mathbb{R}^2) \text{ and } (\xi_1, \xi_2) \in \mathbb{R}^2,$$

we can express the conditional density as

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

where the characteristic function is defined by

$$\begin{aligned} & \phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}}(s_1, s_2) \\ = & \frac{\mathbb{E}[\exp(is_1 x_{t-1} + is_2 w_t) | d_{t-1}, w_{t-1}] \cdot \exp \left[\int_0^{s_1} \frac{\mathbb{E}[i(x_t - \alpha^d - \beta^d w_{t-1}) \exp(is'_1 x_{t-1}) | d_{t-1}=d]}{\gamma^d \mathbb{E}[\exp(is'_1 x_{t-1}) | d_{t-1}=d]} ds'_1 \right]}{\mathbb{E}[\exp(is_1 x_{t-1}) | d_{t-1} = d]} \end{aligned}$$

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

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

In summary, we obtain the closed-form expression

$$\begin{aligned} f(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*) &= \frac{\left(\mathcal{F}_2 \phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}} \right) (x_{t-1}^*, w_t)}{\int \left(\mathcal{F}_2 \phi_{x_{t-1}^*, w_t | d_{t-1}, w_{t-1}} \right) (x_{t-1}^*, w_t) dw_t} \\ &= \sum_d \mathbb{1}\{d_{t-1} = d\} \int \int \exp(-is_1 w_t - is_2 x_{t-1}^*) \cdot \mathbb{E}[\exp(is_1 x_{t-1} + is_2 w_t) | d_{t-1} = d, w_{t-1}] \times \\ & \quad \frac{\exp \left[\int_0^{s_1} \frac{\mathbb{E}[i(x_t - \alpha^{d'} - \beta^{d'} w_{t-1}) \exp(is'_1 x_{t-1}) | d_{t-1}=d']}{\gamma^{d'} \mathbb{E}[\exp(is'_1 x_{t-1}) | d_{t-1}=d']} ds'_1 \right]}{\mathbb{E}[\exp(is_1 x_{t-1}) | d_{t-1} = d']} ds_1 ds_2 \Bigg/ \\ & \quad \int \int \int \exp(-is_1 w_t - is_2 x_{t-1}^*) \cdot \mathbb{E}[\exp(is_1 x_{t-1} + is_2 w_t) | d_{t-1} = d, w_{t-1}] \times \\ & \quad \frac{\exp \left[\int_0^{s_1} \frac{\mathbb{E}[i(x_t - \alpha^{d'} - \beta^{d'} w_{t-1}) \exp(is'_1 x_{t-1}) | d_{t-1}=d']}{\gamma^{d'} \mathbb{E}[\exp(is'_1 x_{t-1}) | d_{t-1}=d']} ds'_1 \right]}{\mathbb{E}[\exp(is_1 x_{t-1}) | d_{t-1} = d']} ds_1 ds_2 dw_t \end{aligned}$$

using any d' . This completes Step 3.

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

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

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

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

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

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

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

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

where the characteristic functions are defined by

$$\begin{aligned}
\phi_{(d)x_t^*|w_t}(s) &= \frac{\mathbb{E}[\mathbb{1}\{d_t = d\} \exp(isx_t) | w_t]}{\mathbb{E}[\exp(is\varepsilon_t)]} \\
&= \frac{\mathbb{E}[\mathbb{1}\{d_t = d\} \exp(isx_t) | w_t] \cdot \exp\left[\int_0^s \frac{\mathbb{E}[i(x_{t+1} - \alpha^{d'} - \beta^{d'} w_t) \exp(is_1 x_t) | d_t = d']}{\gamma^{d'} \mathbb{E}[\exp(is_1 x_t) | d_t = d']} ds_1\right]}{\mathbb{E}[\exp(isx_t) | d_t = d']}
\end{aligned}$$

and

$$\begin{aligned}
\phi_{x_t^*|w_t}(s) &= \frac{\mathbb{E}[\exp(isx_t) | w_t]}{\mathbb{E}[\exp(is\varepsilon_t)]} \\
&= \frac{\mathbb{E}[\exp(isx_t) | w_t] \cdot \exp\left[\int_0^s \frac{\mathbb{E}[i(x_{t+1} - \alpha^{d'} - \beta^{d'} w_t) \exp(is_1 x_t) | d_t = d']}{\gamma^{d'} \mathbb{E}[\exp(is_1 x_t) | d_t = d']} ds_1\right]}{\mathbb{E}[\exp(isx_t) | d_t = d']}
\end{aligned}$$

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

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

using any d' . This completes Step 4. \square

A.2 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, $f(d_t | w_t, x_t^*)$, is given by

$$\begin{aligned}
\widehat{f}(d | w, x^*) &= \int \exp(-isx^*) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} K\left(\frac{W_{jt}-w}{h_w}\right)} \times \\
&\quad \widehat{\phi}_{x_t^*|d_t=d'}(s) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x) ds \Bigg/ \\
&\quad \int \exp(-isx^*) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} K\left(\frac{W_{jt}-w}{h_w}\right)} \times \\
&\quad \widehat{\phi}_{x_t^*|d_t=d'}(s) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(isX_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(sh_x) ds \quad (\text{A.2})
\end{aligned}$$

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. We discuss appropriate properties of

K required for desired large sample properties in Section A.6 in the appendix. 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

$$\begin{aligned} \hat{f}(w^*) &= \int \int \exp(-is_1 w' - is_2 x^*) \times \\ &\quad \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{jt} + is_2 W_{j,t+1}) \cdot \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)} \cdot \hat{\phi}_{x_t^*|d_t=d'}(s_1) \times \\ &\quad \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(s_1 h_w) \cdot \phi_K(s_2 h_x) ds_1 ds_2 \Bigg/ \\ &\quad \int \int \int \exp(-is_1 w'' - is_2 x^*) \times \\ &\quad \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{jt} + is_2 W_{j,t+1}) \cdot \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d\} \cdot K\left(\frac{W_{jt}-w}{h_w}\right)} \cdot \hat{\phi}_{x_t^*|d_t=d'}(s_1) \times \\ &\quad \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(s_1 h_w) \cdot \phi_K(s_2 h_x) ds_1 ds_2 dw'' \quad (\text{A.3}) \end{aligned}$$

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

$$\begin{aligned} \hat{f}(x^{*'}) &= \frac{1}{2\pi} \int \exp(-is(x^{*'} - \beta^d w - \gamma^d x^*)) \times \\ &\quad \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is X_{j,t+1}) \cdot \mathbb{1}\{D_{jt} = d\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is(\alpha^d + \beta^d W_{jt} + \gamma^d X_{jt})) \cdot \mathbb{1}\{D_{jt} = d\}} \times \\ &\quad \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is \gamma^d X_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}}{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is X_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}} \cdot \frac{\hat{\phi}_{x_t^*|d_t=d'}(s)}{\hat{\phi}_{x_t^*|d_t=d'}(s \gamma^d)} \phi_K(s h_x) ds \quad (\text{A.4}) \end{aligned}$$

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

$$\hat{f}(x \mid x^*) = \frac{1}{2\pi} \int \exp(-is(x - x^*)) \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is X_{jt}) \cdot \mathbb{1}\{D_{jt} = d'\}}{\hat{\phi}_{x_t^*|d_t=d'}(s) \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}} \cdot \phi_K(s h_x) ds \quad (\text{A.5})$$

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:

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

A.3 Derivation of Restriction (4.1)

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

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

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

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

where $f_{D_t|W_t X_t^*}(d_t \mid w_t, x_t^*)$ 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

$$\begin{aligned}
& \rho \mathbb{E} [V(w_{t+1}, x_{t+1}^*) \mid d_t = 1, w_t, x_t^*] - \rho \mathbb{E} [V(w_{t+1}, x_{t+1}^*) \mid d_t = 0, w_t, x_t^*] \\
&= \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathbb{E} [f_{D_s|W_s, X_s^*}(0 \mid w_s, x_s^*) \cdot (\theta_0^0 + \theta_0^w w_s + \theta_0^x x_s^* + c - \ln f_{D_s|W_s, X_s^*}(0 \mid w_s, x_s^*)) + \\
&\quad f_{D_s|W_s, X_s^*}(1 \mid w_s, x_s^*) \cdot (\theta_1^w w_s + \theta_1^x x_s^* + c - \ln f_{D_s|W_s, X_s^*}(1 \mid w_s, x_s^*)) \mid d_t = 1, w_t, x_t^*] - \\
&\quad \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot \mathbb{E} [f_{D_s|W_s, X_s^*}(0 \mid w_s, x_s^*) \cdot (\theta_0^0 + \theta_0^w w_s + \theta_0^x x_s^* + c - \ln f_{D_s|W_s, X_s^*}(0 \mid w_s, x_s^*)) + \\
&\quad f_{D_s|W_s, X_s^*}(1 \mid w_s, x_s^*) \cdot (\theta_1^w w_s + \theta_1^x x_s^* + c - \ln f_{D_s|W_s, X_s^*}(1 \mid w_s, x_s^*)) \mid d_t = 0, w_t, x_t^*]
\end{aligned}$$

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

A.4 Feasible Computation of Moments – Remark 3

This section is referred to by Remark 3, 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

$$\begin{aligned}
0 &= \int \int \zeta(w_t, x_t^*) dF(w_t, x^*) \\
&= \int \int \frac{\zeta(w_t, x_t^*) \cdot f(x_t \mid x_t^*) \cdot f(x_t^* \mid w_t)}{\int f(x_t \mid x_t^*) \cdot f(x_t^* \mid w_t) dx_t^*} dF(w_t, x_t) \tag{A.6}
\end{aligned}$$

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 $f(x_t^* | w_t)$. 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

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

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

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

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

$$f(x_t^* | w_t) = \frac{1}{2\pi} \sum_d \int e^{-isx_t^*} \cdot f(d_t = d | w_t) \times \exp \left[\int_0^s \frac{\mathbb{E}[i(x_{t+1} - \alpha^d - \beta^d w_t) \cdot \exp(is_1 x_t) | d_t = d, w_t]}{\gamma^d \cdot \mathbb{E}[\exp(is_1 x_t) | d_t = d, w_t]} ds_1 \right] ds. \quad (\text{A.7})$$

Therefore, $f(x_t^* | w_t)$ is identified in a closed form. This shows that the expression in the last line of (A.6) can be evaluated in a closed-form.

Lastly, we propose a sample-counterpart estimation of (A.7). The conditional density $f(x_t^* | w_t)$ is estimated in a closed form by

$$\begin{aligned} \hat{f}(x^* | w) &= \frac{1}{2\pi} \sum_d \int e^{-isx^*} \cdot \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)}{\sum_{j=1}^N \sum_{t=1}^{T-1} K\left(\frac{W_{j,t} - w}{h_w}\right)} \times \\ &\exp \left[\int_0^s \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} i(X_{j,t+1} - \alpha^d - \beta^d W_{j,t}) \cdot \exp(is_1 X_{j,t}) \cdot \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)}{\gamma^d \cdot \sum_{j=1}^N \sum_{t=1}^{T-1} \exp(is_1 X_{j,t}) \cdot \mathbb{1}\{D_{j,t} = d\} \cdot K\left(\frac{W_{j,t} - w}{h_w}\right)} ds_1 \right] ds. \end{aligned} \quad (\text{A.8})$$

A.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

$$\mathbb{E} [R(\rho, f; x_t^*)' \theta - R(\rho, f; x_t^*)] = 0$$

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

$$\begin{aligned} \xi(\rho, f; x_t^*) &= \ln f(1 | x_t^*) - \ln f(0 | x_t^*) \\ + \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot (\mathbb{E}_f[f(0 | x_s^*) \cdot \ln f(0 | x_s^*) | d_t = 1, x_t^*] &+ \mathbb{E}_f[f(1 | x_s^*) \cdot \ln f(1 | x_s^*) | d_t = 1, x_t^*]) \\ - \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot (\mathbb{E}_f[f(0 | x_s^*) \cdot \ln f(0 | x_s^*) | d_t = 0, x_t^*] &+ \mathbb{E}_f[f(1 | x_s^*) \cdot \ln f(1 | x_s^*) | d_t = 0, x_t^*]) \\ \xi_0^0(\rho, f; x_t^*) &= \sum_{s=t+1}^{\infty} \rho^{s-t} \cdot (\mathbb{E}_f[f(0 | x_s^*) | d_t = 1, x_t^*] - \mathbb{E}_f[f(0 | x_s^*) | d_t = 0, x_t^*]) - 1 \end{aligned}$$

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

for each $d \in \{0, 1\}$. The subscript f under the \mathbb{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

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

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

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

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

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

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.

A.6 Large Sample Properties

In this section, we present theoretical large sample properties of our closed-form estimator of the structural parameters. 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 parts from the functions R and ξ as well as from the components of the Markov kernel. Furthermore, we use a slight variant of our baseline estimator of the Markov components for the sake of obtaining asymptotic normality for the closed-form estimator of the structural parameters. See A.5 for the exact expressions of the estimator that we obtain under this setting.

For convenience of our analyses of large sample properties, we make explicit the dependence of the functions R and ξ on the Markov components by writing

$$\begin{aligned} R(\rho, f; x_t^*) &= R(\rho; x_t^*) & R(\rho, \hat{f}; x_t^*) &= \hat{R}(\rho; x_t^*) & \text{and} \\ \xi(\rho, f; x_t^*) &= \xi(\rho; x_t^*) & \xi(\rho, \hat{f}; x_t^*) &= \hat{\xi}(\rho; x_t^*), \end{aligned}$$

where f denotes the vector of the components of the Markov kernel, i.e., $f(d_t, x_t^*, x^*; d_{t-1}, x_{t-1}^*) = (f(d_t | x_t^*), f(x_t^* | d_{t-1}, x_{t-1}^*), f(x_t | x_t^*))$, and \hat{f} denotes its estimate. The moment restriction is written as

$$\mathbb{E} [R(\rho, f; x_t^*)' R(\rho, f; x_t^*) \theta - R(\rho, f; x_t^*)' \xi(\rho, f; x_t^*)] = 0$$

and the sample-counterpart closed form estimator $\hat{\theta}$ is obtained by substituting \hat{f} for f in this expression. Furthermore, we simply take the above expectation with respect to the observed distribution of x_t , while Section 4 introduces a way to compute the expectation with respect to the unobserved distribution of x_t^* . Note that the moment restriction continues to hold even after this substitution of the integrators, because the population restriction (4.1) holds point-wise – also see Remark 3.

With these new setup and notations, it is clear that our estimator is essentially the semi-parametric two-step estimator, where f is an infinite-dimensional nuisance parameter. Reflect-

ing this characterization of the estimator, the score is denoted by

$$\bar{m}_{N,T}(\rho, \theta, f) = \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} m_{j,t}(\rho, \theta, f; X_{j,t}^*),$$

where $m_{j,t}$ is defined by

$$m_{j,t}(\rho, \theta, f; X_{j,t}^*) = R(\rho, f; X_{j,t}^*)' R(\rho, f; X_{j,t}^*) \theta - R(\rho, f; X_{j,t}^*)' \xi(\rho, f; X_{j,t}^*).$$

To derive asymptotic normality of our closed-form estimator $\hat{\theta}$ of the structural parameters, we make the following set of assumptions.

Assumption 7 (Large Sample). (a) The data $\{D_{j,t}, X_{j,t}^*\}_{t=1}^T$ is i.i.d. across j . (b) $\theta_0 \in \Theta$ where Θ is compact. (c) $f_0 \in \mathcal{F}$ where \mathcal{F} is compact with respect to some metric. (d) $\mathcal{X}^* = \text{supp}(X_{j,t}^*)$ is bounded and convex. (e) The CCP $f(d_t \mid \cdot)$ is uniformly bounded away from 0 and 1 over \mathcal{X}^* . (f) $\rho_0 \in (0, 1)$. (g) $\sup_{f \in \mathcal{F}} \|m_{j,t}(\rho_0, \theta_0, f; \cdot)\|_{2,1,\mathcal{X}^*} < \infty$, where $\|\cdot\|_{2,1,\mathcal{X}^*}$ is the first-order L^2 Sobolev norm on \mathcal{X}^* . (h) $m_{j,t}(\rho_0, \theta, f, \cdot)$ is continuous for all $(\theta, f) \in \Theta \times \mathcal{F}$. (i) $E[\sup_{(\theta,f) \in \Theta \times \mathcal{F}} |m_{j,t}(\rho_0, \theta, f, X^*)|] < \infty$. (j) $m_{j,t}(\rho_0, \cdot, f, x^*)$ is twice continuously differentiable for all $f \in \mathcal{F}$ and for all $x^* \in \mathcal{X}^*$. (k) $\sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta_0, f_0; X_{j,t}^*)$ has finite $(2+r)$ -th moment for some $r > 0$. (l) 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 $|f^{(k_1)}(x^*) - f^{(k_1)}(x^* + \delta)| \leq k_0 |\delta|^{k_2}$ for all $x^* \in \mathcal{X}^*$ and $\delta \in \mathbb{R}$. Let $k = k_1 + k_2$ be the largest number satisfying this property. (m) $f(d \mid x^*)$ 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. (n) The conditional distribution of X_t given $D_t = d$ is ordinary-smooth of order $q > 0$ for some choice d , i.e., $|\phi_{x_t|d_t=d}(s)| = \mathcal{O}(|s|^{-q})$ as $t \rightarrow \pm\infty$. (o) The bandwidth parameter is chosen so that $h_x \rightarrow 0$ and $nh_x^{4+4q+2\min\{k,l\}} \rightarrow c$ as $N \rightarrow \infty$ for some nonzero constant c . (p) $\min\{k, l\} > 2 + 2q$.

The major role of each part of Assumption 7 is as follows. The i.i.d. requirement (a) is useful to obtain the asymptotic independence of the nonparametric estimator \hat{f} , which in turn is important to derive the desired asymptotic normality result for $\hat{\theta}$. The compactness of the parameter space $\Theta \times \mathcal{F}$ in (b) and (c) are used in the common manner to apply the uniform weak law of large numbers among others. The boundedness of the state space \mathcal{X}^* in (d) is used primarily for two important objectives. First, together with the convexity requirement in (d) as well as what is discussed later about (g), it can be used to guarantee the stochastic equicontinuity of the empirical processes. Second, the bounded state space is necessary to uniformly bound the density function of X_t^* away from 0, which in turn is convenient for us to obtain a uniform convergence rate of the nonparametric estimator \hat{f} of the infinite dimensional nuisance parameters f_0 so as to prove the asymptotic independence. The assumption (f) that the true rate ρ_0 of time preference lies strictly between 0 and 1 is used to guarantee the existence and continuity of the score and its derivatives. The bounded first-order Sobolev norm in (g) is used to guarantee the stochastic equicontinuity of the empirical processes. Parts (h) and (i) are used derive consistency together with parts (a) and (b) as well the uniform law of large numbers. The twice-continuous differentiability in (j) and bounded $(2+r)$ -th moment in (k) are used for the asymptotic normality of the part of the empirical process evaluated at (ρ_0, θ_0, f_0) by the Lyapunov central limit theorem. The Hölder continuity assumptions in (l) and (m) admit use of higher-order kernels to mitigate the asymptotic bias of the nonparametric estimates of the components of the Markov kernel sufficiently enough to achieve asymptotic independence. The smoothness parameter in (n) determines the best convergence rate of the nonparametric estimates of the Markov components. The bandwidth choice in (o) is to assure that the squared bias and the variance of the nonparametric estimates of the Markov components converge at the same asymptotic rate so we can control their order. Lastly, part (p) requires that the

marginal density $f(x_t^*)$ and the CCP $f(d_t \mid x_t^*)$ are smooth enough with respect to x_t^* , and that characteristic function $\phi_{x_t \mid d_t=d}$ vanish relatively slowly toward the tails. On one hand, the smoothness of the marginal density $f(x_t^*)$ and the CCP $f(d_t \mid x_t^*)$ helps to reduce the asymptotic bias. On the other hand, the smoothness of the conditional distribution of x_t given d_t exacerbates the asymptotic variance. This relative rate restriction balances the subtle trade-offs, and is used to have the nonparametric nuisance parameters converge fast enough, specifically at least the rate faster than $n^{1/4}$. Under this set of assumptions, we obtain the following asymptotic normality result for the estimator $\hat{\theta}$ of the structural parameters.

Proposition 1 (Asymptotic Normality). *If Assumptions 1, 2, 3, 4, 5, 6, and 7 are satisfied, then $\sqrt{N}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, V)$ as $N \rightarrow \infty$, where $V = M(\rho_0, f_0)^{-1}S(\rho_0, \theta_0, f_0)M(\rho_0, f_0)^{-1}$ with*

$$\begin{aligned} M(\rho_0, f_0) &= E[R(\rho_0, f_0; X_{j,t}^*)' R(\rho_0, f_0; X_{j,t}^*)] \quad \text{and} \\ S(\rho_0, \theta_0, f_0) &= \text{Var}\left(\frac{1}{T-1} \sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta_0, f_0; X_{j,t}^*)\right). \end{aligned}$$

In practice, it is also convenient to implement bootstrap re-sampling – see Chen, Linton and Keilegom (2003) for validity of bootstrapping.

A proof is given in Section A.7.

A.7 Proof of Proposition 1

Proof. First, note that identification of $f_0 \in \mathcal{F}$ and $\theta_0 \in \Theta$ is already obtained in the previous sections. We follow Andrews (1994) to prove the asymptotic normality of $\hat{\theta}$. First, we show that $m_{j,t}(\rho_0, \cdot, f; x_t^*)$ is continuously differentiable on Θ for all $f \in \mathcal{F}$, $x_t^* \in \mathcal{X}^*$, i and t . This is trivial from our definition of $m_{j,t}$ together with the boundedness of $R(\rho_0, f; x_j^*)$ that follows from Assumption 7 (e) and (f).

Next, we show that $\sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*)$ satisfies the uniform weak law of large numbers in the limit $N \rightarrow \infty$ over $\Theta \times \mathcal{F}$. To see this, note the compactness of the parameter space by Assumption 7 (b) and (c). Furthermore, $\sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*)$ is continuous with respect to (θ, f) due to Assumption 7 (e) and (f). The uniform boundedness $\mathbb{E} \sup_{(\theta, f) \in \Theta \times \mathcal{F}} \left| \sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*) \right| \leq \infty$ also follows from Assumption 7 (b), (c), (e), and (f). These suffice for $\sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*)$ to satisfy the conditions for the uniform weak law of large numbers in the limit $N \rightarrow \infty$ over $\Theta \times \mathcal{F}$. Furthermore, under the same set of assumptions, $m(\rho_0, \theta, f) = \frac{1}{T-1} \sum_{t=1}^{T-1} \mathbb{E} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*)$ exists and is continuous with respect to (θ, f) on $\Theta \times \mathcal{F}$. Similar lines of argument to show that the Hessian $\sum_{t=1}^{T-1} \frac{\partial}{\partial \theta} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*) = \sum_{t=1}^{T-1} R(\rho_0, \theta, f; X_{j,t}^*)' R(\rho_0, \theta, f; X_{j,t}^*)$ also satisfies the uniform weak law of large numbers in the limit $N \rightarrow \infty$ over $\Theta \times \mathcal{F}$, and that $M(\rho_0, f) = \mathbb{E} \frac{\partial}{\partial \theta} m_{j,t}(\rho_0, \theta, f; X_{j,t}^*) = \mathbb{E} R(\rho_0, f; X_{j,t}^*)' R(\rho_0, f; X_{j,t}^*)$ exists and is continuous with respect to f on \mathcal{F} .

To vanish the terms in the score that follow from estimating f by \hat{f} , we require that the empirical process $\nu_{NT}(\rho_0, \theta_0, f) := \sqrt{N} (\bar{m}_{NT}(\rho_0, \theta_0, f) - \mathbb{E} \bar{m}_{NT}(\rho_0, \theta_0, f))$ is stochastically equicontinuous at $f = f_0$. This can be shown to hold under Assumption 7 (a), (d), and (g) by applying the sufficient condition proposed by Andrews (1994).

To show that the empirical process under the true parameter values $\nu_{NT}(\rho_0, \theta_0, f_0)$ converge in distribution to a normal distribution as $N \rightarrow \infty$, it suffices to invoke the Lyapunov central limit theorem under Assumption 7 (a) and (k), where the N -asymptotic variance matrix is given by $S(\rho_0, \theta_0, f_0) = \text{Var} \left(\frac{1}{T-1} \sum_{t=1}^{T-1} m_{j,t}(\rho_0, \theta_0, f_0; X_{j,t}^*) \right)$.

Next, we show the asymptotic independence $\sqrt{N} \mathbb{E} \bar{m}_{NT}(\rho_0, \theta_0, \hat{f}) \xrightarrow{p} 0$. To this end, we show $\text{super-}n^{1/4}$ rate of uniform convergence of the leave-one-out nonparametric estimates of the components of the Markov kernel by the standard argument, but we need to perform several steps of calculations. 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 $g_d(x_t^*) := E[\mathbb{1}\{d_t = d\} \mid x_t^*]$. Our CCP estimator is written as $\widehat{g_d(x^*)f(x^*)}/\widehat{f(x^*)}$ where

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

and

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

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

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

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

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

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

$$\begin{aligned} &\left| E \widehat{f(x^*)} - \frac{1}{2\pi} \int e^{-isx^*} \widehat{\phi}_{x_t^*|d_t=d'}(s) \frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)} \phi_K(sh) ds \right| \\ &= \left| \frac{1}{2\pi} \int e^{-isx^*} \phi_K(sh) \left\{ E \left[\exp \left(i \int_0^s \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}}{\gamma^{d'} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}} ds_1 \right) \times \right. \right. \right. \\ &\quad \left. \left. \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} e^{isX_{jt}} \sum_{j=1}^N \sum_{t=1}^{T-1} \mathbb{1}\{D_{jt} = d'\}}{N(T-1) \sum_{j=1}^N \sum_{t=1}^{T-1} e^{isX_{jt}} \mathbb{1}\{D_{jt} = d'\}} \right] - \widehat{\phi}_{x_t^*|d_t=d'}(s) \frac{\phi_{x_t}(s)}{\phi_{x_t|d_t=d'}(s)} \right\} ds \right| \end{aligned}$$

$$\begin{aligned}
&\leq \frac{\|\phi_K\|_\infty \|\phi_{x_t^*|d_t=d'}\|_\infty}{2\pi h} \int_{-1}^1 \int_0^{s/h} \\
&\quad \left(\frac{\|\phi_{x_t}\|_\infty \mathbb{E} \left| \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right|}{|\phi_{x_t|d_t=d'}(s/h)| |\phi_{x_t|d_t=d'}(s_1)| |\gamma^{d'}| f(d')} \right. \\
&\quad + \frac{\|\phi_{x_t}\|_\infty \|\phi'_{x_t^*|d_t=d'}\|_\infty \mathbb{E} \left| \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right|}{|\phi_{x_t|d_t=d'}(s/h)| |\phi_{x_t|d_t=d'}(s_1)| f(d')} \\
&\quad + \frac{\mathbb{E} \left| \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}} - \mathbb{E} e^{is X_{jt}} \right|}{|\phi_{x_t|d_t=d'}(s/h)|} \\
&\quad \left. + \frac{\|\phi_{x_t}\|_\infty \mathbb{E} \left| \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right|}{|\phi_{x_t|d_t=d'}(s/h)|^2 f(d')} + \text{hot}(s_1) + \text{hot}(s/h) \right) ds_1 ds \\
&= \mathcal{O} \left(\frac{1}{n^{1/2} h^2 |\phi_{x_t|d_t=d'}(1/h)|^2} \right)
\end{aligned}$$

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
& \left| \mathbb{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\{ \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}}{\gamma^{d'} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}} ds_1 \right) \times \right. \right. \right. \\
&\quad \left. \left. \frac{\sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}} \mathbb{1}\{D_{jt} = d\}}{N(T-1) \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}} \mathbb{1}\{D_{jt} = d'\}} \right] - \phi_{x_t^*|d_t=d'}(s) \frac{\mathbb{E}[e^{is X_{jt}} \mathbb{1}\{D_{jt} = d\}]}{\phi_{x_t|d_t=d'}(s)} \right\} ds \right| \\
&\leq \frac{\|\phi_K\|_\infty \|\phi_{x_t^*|d_t=d'}\|_\infty}{2\pi h} \int_{-1}^1 \int_0^{s/h} \\
&\quad \left(\frac{\|\phi_{x_t|d_t=d}\|_\infty \mathbb{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'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right| f(d)}{|\phi_{x_t|d_t=d'}(s/h)| |\phi_{x_t|d_t=d'}(s_1)| |\gamma^{d'}| f(d')} \right. \\
&\quad + \frac{\|\phi_{x_t|d_t=d}\|_\infty \|\phi_{x_t^*|d_t=d'}'\|_\infty \mathbb{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'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right| f(d)}{|\phi_{x_t|d_t=d'}(s/h)| |\phi_{x_t|d_t=d'}(s_1)| f(d')} \\
&\quad + \frac{\mathbb{E} \left| \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}} - \mathbb{E} e^{is X_{jt}} \right|}{|\phi_{x_t|d_t=d'}(s/h)|} \\
&\quad \left. + \frac{\|\phi_{x_t|d_t=d}\|_\infty \mathbb{E} \left| \frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right| f(d)}{|\phi_{x_t|d_t=d'}(s/h)|^2 f(d')} + \text{hot}(s_1) + \text{hot}(s/h) \right) ds_1 ds \\
&= \mathcal{O} \left(\frac{1}{n^{1/2} h^2 |\phi_{x_t|d_t=d'}(1/h)|^2} \right)
\end{aligned}$$

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

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

where k and l are the Hölder exponents provided in Assumption 7 (l) and (m), respectively.

Consequently, we obtain the following asymptotic order for the absolute bias of $g_d(\widehat{x^*}) f(x^*)$.

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

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

$$\begin{aligned}
& \frac{1}{4\pi^2} \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}}{\gamma^{d'} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}} ds_1 \right) \times \right. \right. \\
& \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. \\
& \left. \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}}{\gamma^{d'} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}} ds_1 \right) \times \right. \\
& \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} \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{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{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'\} - \mathbb{E} e^{isX_{jt}} \mathbb{1}\{D_{jt} = d'\} \right)}{\phi_{x_t|d_t=d'}(s)^2 f(d')} + \text{hot}(s) + \text{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{ir_1 X_{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir_1 X_{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{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'\} - \mathbb{E} e^{irX_{jt}} \mathbb{1}\{D_{jt} = d'\} \right)}{\phi_{x_t|d_t=d'}(r)^2 f(d')} + \text{hot}(s) + \text{hot}(s_1) \right] dr_1 ds_1 dr ds \\
& \leq \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_1 ds_1 dr ds = \mathcal{O} \left(\frac{1}{nh^4 |\phi_{x_t|d_t=d'}(1/h)|^4} \right)
\end{aligned}$$

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

faster uniformly.

$$\begin{aligned}
I_1 &= \frac{\|\phi_{x_t}\|_\infty^2}{|\phi_{x_t}|_{d_t=d'}(s/h)| \cdot |\phi_{x_t}|_{d_t=d'}(s_1)| \cdot |\phi_{x_t}|_{d_t=d'}(r/h)| \cdot |\phi_{x_t}|_{d_t=d'}(r_1)| \cdot f(d')^2 \cdot (\gamma^{d'})^2} \times \\
&\quad \left(\mathbb{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'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{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'\} - \mathbb{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}{|\phi_{x_t}|_{d_t=d'}(s/h)| \cdot |\phi_{x_t}|_{d_t=d'}(s_1)| \cdot |\phi_{x_t}|_{d_t=d'}(r/h)| \cdot |\phi_{x_t}|_{d_t=d'}(r_1)| \cdot f(d')^2} \times \\
&\quad \left(\mathbb{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'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{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'\} - \mathbb{E} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \\
I_3 &= \frac{1}{|\phi_{x_t}|_{d_t=d'}(s/h)| \cdot |\phi_{x_t}|_{d_t=d'}(r/h)|} \cdot \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}/h} - \mathbb{E} e^{is X_{jt}/h} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} - \mathbb{E} e^{ir X_{jt}/h} \right]^2 \right)^{1/2} \\
I_4 &= \frac{\|\phi_{x_t}\|_\infty^2}{|\phi_{x_t}|_{d_t=d'}(s/h)|^2 \cdot |\phi_{x_t}|_{d_t=d'}(r/h)|^2 \cdot f(d')^2} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{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}{|\phi_{x_t}|_{d_t=d'}(s/h)| \cdot |\phi_{x_t}|_{d_t=d'}(s_1)| \cdot |\phi_{x_t}|_{d_t=d'}(r/h)| \cdot |\phi_{x_t}|_{d_t=d'}(r_1)| \cdot f(d')^2 \cdot |\gamma^{d'}|} \times \\
&\quad \left(\mathbb{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'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{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'\} - \mathbb{E} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
I_6 &= \frac{2 \|\phi_{x_t}\|_\infty}{|\phi_{x_t}|_{d_t=d'}(s/h) \cdot |\phi_{x_t}|_{d_t=d'}(s_1) \cdot |\phi_{x_t}|_{d_t=d'}(r/h) \cdot f(d') \cdot |\gamma^{d'}|} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} - \mathbb{E} e^{ir X_{jt}/h} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
I_7 &= \frac{2 \|\phi_{x_t}\|_\infty^2}{|\phi_{x_t}|_{d_t=d'}(s/h) \cdot |\phi_{x_t}|_{d_t=d'}(s_1) \cdot |\phi_{x_t}|_{d_t=d'}(r/h)|^2 \cdot f(d')^2 \cdot |\gamma^{d'}|} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
I_8 &= \frac{2 \|\phi_{x_t}\|_\infty \left\| \phi'_{x_t^*}|_{d_t=d'} \right\|_\infty}{|\phi_{x_t}|_{d_t=d'}(s/h) \cdot |\phi_{x_t}|_{d_t=d'}(s_1) \cdot |\phi_{x_t}|_{d_t=d'}(r/h) \cdot f(d')} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} - \mathbb{E} e^{ir X_{jt}/h} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
I_9 &= \frac{2 \|\phi_{x_t}\|_\infty^2 \left\| \phi'_{x_t^*}|_{d_t=d'} \right\|_\infty}{|\phi_{x_t}|_{d_t=d'}(s/h) \cdot |\phi_{x_t}|_{d_t=d'}(s_1) \cdot |\phi_{x_t}|_{d_t=d'}(r/h)|^2 \cdot f(d')^2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
I_{10} &= \frac{2 \|\phi_{x_t}\|_\infty}{|\phi_{x_t}|_{d_t=d'}(s/h) \cdot |\phi_{x_t}|_{d_t=d'}(r/h)|^2 \cdot f(d')} \cdot \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}/h} - \mathbb{E} e^{is X_{jt}/h} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

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

$$\begin{aligned}
& \frac{1}{4\pi^2} \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}}{\gamma^{d'} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}} ds_1 \right) \times \right. \right. \\
& \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. \\
& \left. \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}}{\gamma^{d'} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\}} ds_1 \right) \times \right. \\
& \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 \Big)^2 \\
& = \frac{1}{4\pi^2} \mathbb{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{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\} - \mathbb{E} e^{isX_{jt}} \mathbb{1}\{D_{jt} = d\}}{\phi_{x_t|d_t=d'}(s)} \\
& - \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'\} - \mathbb{E} e^{isX_{jt}} \mathbb{1}\{D_{jt} = d'\} \right)}{\phi_{x_t|d_t=d'}(s)^2 f(d')} + \text{hot}(s) + \text{hot}(s_1) \Big] \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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{ir_1 X_{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_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir_1 X_{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\} - \mathbb{E} e^{irX_{jt}} \mathbb{1}\{D_{jt} = d\}}{\phi_{x_t|d_t=d'}(r)} \\
& - \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'\} - \mathbb{E} e^{irX_{jt}} \mathbb{1}\{D_{jt} = d'\} \right)}{\phi_{x_t|d_t=d'}(r)^2 f(d')} + \text{hot}(s) + \text{hot}(s_1) \Big] dr_1 ds_1 dr ds \\
& \leq \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_1 ds_1 dr ds = \mathcal{O} \left(\frac{1}{nh^4 |\phi_{x_t|d_t=d'}(1/h)|^4} \right)
\end{aligned}$$

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

faster uniformly.

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

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

$$J_3 = \frac{1}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(r/h)|} \times \\ \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathbb{E} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2} \times \\ \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2}$$

$$J_4 = \frac{\|\phi_{x_t|d_t=d}\|_\infty^2 f(d)^2}{|\phi_{x_t|d_t=d'}(s/h)|^2 \cdot |\phi_{x_t|d_t=d'}(r/h)|^2 \cdot f(d')^2} \times \\ \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\ \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}$$

$$\begin{aligned}
J_5 &= \frac{2 \|\phi_{x_t|d_t=d}\|_\infty^2 \left\| \phi'_{x_t^*|d_t=d'} \right\|_\infty f(d)^2}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(s_1)| \cdot |\phi_{x_t|d_t=d'}(r/h)| \cdot |\phi_{x_t|d_t=d'}(r_1)| \cdot f(d')^2 \cdot |\gamma^{d'}|} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
J_6 &= \frac{2 \|\phi_{x_t|d_t=d}\|_\infty f(d)}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(s_1)| \cdot |\phi_{x_t|d_t=d'}(r/h)| \cdot f(d') \cdot |\gamma^{d'}|} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
J_7 &= \frac{2 \|\phi_{x_t|d_t=d}\|_\infty^2 f(d)^2}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(s_1)| \cdot |\phi_{x_t|d_t=d'}(r/h)|^2 \cdot f(d')^2 \cdot |\gamma^{d'}|} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} (X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E}(X_{j,t+1} - \alpha^{d'}) e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
J_8 &= \frac{2 \|\phi_{x_t|d_t=d}\|_\infty \left\| \phi'_{x_t^*|d_t=d'} \right\|_\infty f(d)}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(s_1)| \cdot |\phi_{x_t|d_t=d'}(r/h)| \cdot f(d')} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
J_9 &= \frac{2 \|\phi_{x_t|d_t=d}\|_\infty^2 \|\phi'_{x_t^*|d_t=d'}\|_\infty f(d)^2}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(s_1)| \cdot |\phi_{x_t|d_t=d'}(r/h)|^2 \cdot f(d')^2} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{is_1 X_{jt}} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2} \\
J_{10} &= \frac{2 \|\phi_{x_t|d_t=d}\|_\infty f(d)}{|\phi_{x_t|d_t=d'}(s/h)| \cdot |\phi_{x_t|d_t=d'}(r/h)|^2 \cdot f(d')} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d\} - \mathbb{E} e^{is X_{jt}/h} \mathbb{1}\{D_{jt} = d\} \right]^2 \right)^{1/2} \times \\
&\quad \left(\mathbb{E} \left[\frac{1}{N(T-1)} \sum_{j=1}^N \sum_{t=1}^{T-1} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} - \mathbb{E} e^{ir X_{jt}/h} \mathbb{1}\{D_{jt} = d'\} \right]^2 \right)^{1/2}
\end{aligned}$$

Consequently, under Assumption 7 (n), the bandwidth parameter choice prescribed in Assumption 7 (o) equates the asymptotic orders of the squared bias and the variance of $\widehat{g_d(x^*)f(x^*)}$ as $n^{-1}h^{-4} |\phi_{x_t|d_t=d'}|^{-4} \sim h^{2\min\{k,l\}}$ holds if and only if $nh_x^{4+4q+2\min\{k,l\}} \sim 1$ holds. Substituting this asymptotic rate of the bandwidth parameter into the bias or the square-root of the variance, we obtain

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

By similar lines of argument, we have

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

Since the MSE of the CCP estimator is given by

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

it follows that

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

Now, notice that this convergence rate is faster than $n^{-1/4}$ as $\min\{k, l\}/(2(2+2q+\min\{k, l\})) > 1/4$ under Assumption 7 (p). Moreover, this rate is invariant across x^* over the assumed compact support, implying that the CCP estimator converges uniformly at the rate faster than $n^{-1/4}$. Similar calculations show that the same conclusion is true for the other components of the Markov kernel. By the continuity of $R(\rho_0, f; x^*)$ and $\xi(\rho_0, f; x^*)$ with respect to f under Assumption 7 (e) and (f), the asymptotic independence is satisfied.

With all these arguments, applying Andrews (1994) yields the desired asymptotic normality result for the estimator $\hat{\theta}$ of the structural parameters under the stated assumptions. \square

A.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.⁷ The transition rule for the unobserved state variable x_t^* is defined by

$$\begin{aligned} x_t^* &= 1.000 + 1.000 \cdot x_{t-1}^* + \eta_t^0 & \text{if } d = 0 \\ x_t^* &= 0.000 + 0.000 \cdot x_{t-1}^* + \eta_t^1 & \text{if } d = 1 \end{aligned}$$

⁷In 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 A.6 where large sample properties of the estimator are studied in a similarly simplified setup.

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

$$\begin{aligned} 1.000 - 0.015 \cdot x_t^* + \omega_{0t} & \quad \text{if } d = 0 \\ 0.000 \cdot x_t^* + \omega_{1t} & \quad \text{if } d = 1 \end{aligned}$$

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 4.2. 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

	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.

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 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.

A.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.

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

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

$$\begin{aligned} f(x_t^* | d_{t-1}, w_{t-1}, x_{t-1}^*) : \quad x_t^* &= \alpha^d + \beta^d w_{t-1} + \gamma^d x_{t-1}^* + \eta_t^d & \text{if } d_{t-1} = d \\ f(x_t | d_{t-1}, x_t^*) : \quad x_t &= \delta^d x_t^* + \varepsilon_t^d & \text{if } d_{t-1} = d \end{aligned}$$

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

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

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

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

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

Assumption 5'. The random variables w_t and x_t^* have bounded conditional 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 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 $(\delta^0, \dots, \delta^{\bar{d}})$ and distributions of $(\varepsilon_t^0, \dots, \varepsilon_t^{\bar{d}})$, unless we normalize δ^d for one of the choices d . We therefore make the following assumption in addition to the baseline assumptions.

Assumption 8. *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 8 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 A.10 in the appendix. 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.

A.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' :

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

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

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

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

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

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 $f(\varepsilon_t^d)$ and $f(\eta_t^d)$ for each d . Observe that

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

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

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

Therefore,

$$\begin{aligned}
& E \left[\exp \left(is \delta^{d'} x_{t-1}^* \right) | d_{t-1} = d, d_{t-2} = d' \right] \\
&= \exp \left[\int_0^s \left[\frac{\delta^{d'}}{\gamma^d \delta^d} \frac{\partial}{\partial s_2} \ln E [\exp (is_1 x_{t-1} + is_2 x_t) | d_{t-1} = d, d_{t-2} = d'] \right]_{s_2=0} ds_1 \right. \\
&\quad \left. - \int_0^s \frac{i\alpha^d \delta^{d'}}{\gamma^d} ds_1 - \int_0^s \frac{\beta^d \delta^{d'}}{\gamma^d} \frac{E[iw_{t-1} \exp(is_1 x_{t-1}) | d_{t-1} = d, d_{t-2} = d']}{E[\exp(is_1 x_{t-1}) | d_{t-1} = d, d_{t-2} = d']} ds_1 \right] \\
&= \exp \left[\int_0^s \frac{E \left[i \left(\frac{\delta^{d'}}{\delta^d} x_t - \alpha^d \delta^{d'} - \beta^d \delta^{d'} w_{t-1} \right) \exp (is_1 x_{t-1}) | d_{t-1} = d, d_{t-2} = d' \right]}{\gamma^d E [\exp (is_1 x_{t-1}) | d_{t-1} = d, d_{t-2} = d']} ds_1 \right].
\end{aligned}$$

From the proxy model and the independence assumption for ε_t ,

$$E [\exp (is x_{t-1}) | d_{t-1} = d, d_{t-2} = d'] = E \left[\exp \left(is \delta^{d'} x_{t-1}^* \right) | d_{t-1} = d, d_{t-2} = d' \right] E \left[\exp \left(is \varepsilon_{t-1}^{d'} \right) \right].$$

We then obtain the following result using any d .

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

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

$$E [\exp (is \varepsilon_t^d)] = \frac{E [\exp (is x_t) | d_t = d', d_{t-1} = d]}{\exp \left[\int_0^s \frac{E \left[i \left(\frac{\delta^d}{\delta^{d'}} x_{t+1} - \alpha^{d'} \delta^d - \beta^{d'} \delta^d w_t \right) \exp (is_1 x_t) | d_t = d', d_{t-1} = d \right]}{\gamma^{d'} E [\exp (is_1 x_t) | d_t = d', d_{t-1} = d]} ds_1 \right]}. \quad (\text{A.9})$$

using any d' .

In order to identify $f(\eta_t^d)$ for each d , consider

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

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

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

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

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

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

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

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

$$\begin{aligned} \phi_{\eta_t^d}(s) &= \frac{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^d} x_t \right) \mid d_{t-1} = d, d_{t-2} = d' \right]}{\mathbb{E} \left[\exp \left(i s (\alpha^d + \beta^d w_{t-1} + \gamma^d \frac{1}{\delta^{d'}} x_{t-1}) \right) \mid d_{t-1} = d, d_{t-2} = d' \right]} \\ &\quad \times \frac{\exp \left[\int_0^{s/\delta^d} \frac{\mathbb{E} \left[i \left(\frac{\delta^d}{\delta^{d'}} x_{t+1} - \alpha^{d'} \delta^d - \beta^{d'} \delta^d w_t \right) \exp(i s_1 x_t) \mid d_t = d', d_{t-1} = d \right]}{\gamma^{d'} \mathbb{E}[\exp(i s_1 x_t) \mid d_t = d', d_{t-1} = d]} ds_1 \right]}{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^d} x_t \right) \mid d_t = d', d_{t-1} = d \right]} \\ &\quad \times \frac{\mathbb{E} \left[\exp \left(i s \gamma^d \frac{1}{\delta^{d'}} x_{t-1} \right) \mid d_{t-1} = d, d_{t-2} = d' \right]}{\exp \left[\int_0^{s \gamma^d / \delta^{d'}} \frac{\mathbb{E} \left[i \left(\frac{\delta^{d'}}{\delta^d} x_t - \alpha^d \delta^{d'} - \beta^d \delta^{d'} w_{t-1} \right) \exp(i s_1 x_{t-1}) \mid d_{t-1} = d, d_{t-2} = d' \right]}{\gamma^d \mathbb{E}[\exp(i s_1 x_t) \mid d_{t-1} = d, d_{t-2} = d']} ds_1 \right]}. \end{aligned}$$

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

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

In summary, we obtain the closed-form expression

$$\begin{aligned} f(x_t^* \mid d_{t-1}, x_{t-1}, x_{t-1}^*) &= \sum_d \mathbb{1}\{d_{t-1} = d\} \left(\mathcal{F} \phi_{\eta_t^d} \right) (x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*) \\ &= \sum_d \frac{\mathbb{1}\{d_{t-1} = d\}}{2\pi} \int \exp(-i s (x_t^* - \alpha^d - \beta^d w_{t-1} - \gamma^d x_{t-1}^*)) \times \\ &\quad \frac{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^d} x_t \right) \mid d_{t-1} = d, d_{t-2} = d' \right]}{\mathbb{E} \left[\exp \left(i s (\alpha^d + \beta^d w_{t-1} + \gamma^d \frac{1}{\delta^{d'}} x_{t-1}) \right) \mid d_{t-1} = d, d_{t-2} = d' \right]} \times \\ &\quad \frac{\exp \left[\int_0^{s/\delta^d} \frac{\mathbb{E} \left[i \left(\frac{\delta^d}{\delta^{d'}} x_{t+1} - \alpha^{d'} \delta^d - \beta^{d'} \delta^d w_t \right) \exp(i s_1 x_t) \mid d_t = d', d_{t-1} = d \right]}{\gamma^{d'} \mathbb{E}[\exp(i s_1 x_t) \mid d_t = d', d_{t-1} = d]} ds_1 \right]}{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^d} x_t \right) \mid d_t = d', d_{t-1} = d \right]} \times \\ &\quad \frac{\mathbb{E} \left[\exp \left(i s \gamma^d \frac{1}{\delta^{d'}} x_{t-1} \right) \mid d_{t-1} = d, d_{t-2} = d' \right]}{\exp \left[\int_0^{s \gamma^d / \delta^{d'}} \frac{\mathbb{E} \left[i \left(\frac{\delta^{d'}}{\delta^d} x_t - \alpha^d \delta^{d'} - \beta^d \delta^{d'} w_{t-1} \right) \exp(i s_1 x_{t-1}) \mid d_{t-1} = d, d_{t-2} = d' \right]}{\gamma^d \mathbb{E}[\exp(i s_1 x_t) \mid d_{t-1} = d, d_{t-2} = d']} ds_1 \right]} \end{aligned}$$

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 (A.9), we

can write the density of ε_t^d by

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

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

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

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

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

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

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

using any d' . This completes Step 2.

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

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

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

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

by the independence assumption for ε_t^d , and so

$$\begin{aligned} \mathbb{E} \left[\exp \left(i s_1 \delta^d x_{t-1}^* + i s_2 w_t \right) \mid d_{t-1}, w_{t-1}, d_{t-2} = d \right] &= \frac{\mathbb{E} \left[\exp \left(i s_1 x_{t-1} + i s_2 w_t \right) \mid d_{t-1}, w_{t-1}, d_{t-2} = d \right]}{\mathbb{E} \left[\exp \left(i s_1 \varepsilon_{t-1}^d \right) \right]} \\ &= \mathbb{E} \left[\exp \left(i s_1 x_{t-1} + i s_2 w_t \right) \mid d_{t-1}, w_{t-1}, d_{t-2} = d \right] \\ &\quad \times \frac{\exp \left[\int_0^{s_1} \frac{\mathbb{E} \left[i \left(\frac{\delta^d}{\delta^{d'}} x_t - \alpha^{d'} \delta^d - \beta^{d'} \delta^d w_{t-1} \right) \exp \left(i s'_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]}{\gamma^{d'} \mathbb{E} \left[\exp \left(i s'_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]} ds'_1 \right]}{\mathbb{E} \left[\exp \left(i s_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]} \end{aligned}$$

follows with any choice of d' . Rescaling s_1 yields

$$\begin{aligned} &\mathbb{E} \left[\exp \left(i s_1 x_{t-1}^* + i s_2 w_t \right) \mid d_{t-1}, w_{t-1}, d_{t-2} = d \right] \\ &= \mathbb{E} \left[\exp \left(i s_1 \frac{1}{\delta^d} x_{t-1} + i s_2 w_t \right) \mid d_{t-1}, w_{t-1}, d_{t-2} = d \right] \times \\ &\quad \frac{\exp \left[\int_0^{s_1/\delta^d} \frac{\mathbb{E} \left[i \left(\frac{\delta^d}{\delta^{d'}} x_t - \alpha^{d'} \delta^d - \beta^{d'} \delta^d w_{t-1} \right) \exp \left(i s'_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]}{\gamma^{d'} \mathbb{E} \left[\exp \left(i s'_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]} ds'_1 \right]}{\mathbb{E} \left[\exp \left(i s_1 \frac{1}{\delta^d} x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]}. \end{aligned}$$

We can then express the conditional density as

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

where the characteristic function is defined by

$$\begin{aligned} \phi_{w_t, x_{t-1}^* \mid d_{t-1}, w_{t-1}, d_{t-2} = d}(s_1, s_2) &= \mathbb{E} \left[\exp \left(i s_1 \frac{1}{\delta^d} x_{t-1} + i s_2 w_t \right) \mid d_{t-1}, w_{t-1}, d_{t-2} = d \right] \times \\ &\quad \frac{\exp \left[\int_0^{s_1/\delta^d} \frac{\mathbb{E} \left[i \left(\frac{\delta^d}{\delta^{d'}} x_t - \alpha^{d'} \delta^d - \beta^{d'} \delta^d w_{t-1} \right) \exp \left(i s'_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]}{\gamma^{d'} \mathbb{E} \left[\exp \left(i s'_1 x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]} ds'_1 \right]}{\mathbb{E} \left[\exp \left(i s_1 \frac{1}{\delta^d} x_{t-1} \right) \mid d_{t-1} = d', d_{t-2} = d \right]}. \end{aligned}$$

Using this conditional density, we nonparametrically identify the transition rule

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

In summary, we obtain the closed-form expression

$$\begin{aligned}
f(w_t | d_{t-1}, w_{t-1}, x_{t-1}^*) &= \sum_d \mathbb{1}\{d_{t-1} = d\} \times \\
&\frac{\sum_{d'} \left(\mathcal{F}_2 \phi_{x_{t-1}^*, w_t | d_{t-1}=d, w_{t-1}, d_{t-2}=d'} \right) (w_t, x_{t-1}^*) \cdot \Pr(d_{t-2} = d' | d_{t-1} = d, w_{t-1})}{\int \sum_{d'} \left(\mathcal{F}_2 \phi_{x_{t-1}^*, w_t | d_{t-1}=d, w_{t-1}, d_{t-2}=d'} \right) (w_t, x_{t-1}^*) \cdot \Pr(d_{t-2} = d' | d_{t-1} = d, w_{t-1}) dw_t} \\
&= \sum_d \mathbb{1}\{d_{t-1} = d\} \left\{ \sum_{d'} \Pr(d_{t-2} = d' | d_{t-1} = d, w_{t-1}) \int \int \exp(-is_1 w_t - is_2 x_{t-1}^*) \times \right. \\
&\frac{\mathbb{E} \left[\exp \left(is_1 \frac{1}{\delta^{d'}} x_{t-1} + is_2 w_t \right) | d_{t-1} = d, w_{t-1}, d_{t-2} = d' \right]}{\mathbb{E} \left[\exp \left(is_1 \frac{1}{\delta^{d'}} x_{t-1} \right) | d_{t-1} = d'', d_{t-2} = d' \right]} \times \\
&\exp \left[\int_0^{s_1/\delta^{d'}} \frac{\mathbb{E} \left[i \left(\frac{\delta^{d'}}{\delta^{d''}} x_t - \alpha^{d''} \delta^{d'} - \beta^{d''} \delta^{d'} w_{t-1} \right) \exp(is'_1 x_{t-1}) | d_{t-1} = d'', d_{t-2} = d' \right]}{\gamma^{d''} \mathbb{E} [\exp(is'_1 x_{t-1}) | d_{t-1} = d'', d_{t-2} = d']} ds'_1 \right] ds_1 ds_2 \Bigg\} / \\
&\left\{ \sum_{d'} \int \Pr(d_{t-2} = d' | d_{t-1} = d, w_{t-1}) \int \int \exp(-is_1 w_t - is_2 x_{t-1}^*) \times \right. \\
&\frac{\mathbb{E} \left[\exp \left(is_1 \frac{1}{\delta^{d'}} x_{t-1} + is_2 w_t \right) | d_{t-1} = d, w_{t-1}, d_{t-2} = d' \right]}{\mathbb{E} \left[\exp \left(is_1 \frac{1}{\delta^{d'}} x_{t-1} \right) | d_{t-1} = d'', d_{t-2} = d' \right]} \times \\
&\exp \left[\int_0^{s_1/\delta^{d'}} \frac{\mathbb{E} \left[i \left(\frac{\delta^{d'}}{\delta^{d''}} x_t - \alpha^{d''} \delta^{d'} - \beta^{d''} \delta^{d'} w_{t-1} \right) \exp(is'_1 x_{t-1}) | d_{t-1} = d'', d_{t-2} = d' \right]}{\gamma^{d''} \mathbb{E} [\exp(is'_1 x_{t-1}) | d_{t-1} = d'', d_{t-2} = d']} ds'_1 \right] ds_1 ds_2 dw_t \Bigg\}
\end{aligned}$$

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

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

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

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

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

and rescaling s yields

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

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

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

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

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

where the characteristic functions are defined by

$$\begin{aligned} \phi_{(d)x_t^* | w_t(d')}(s) &= \frac{\mathbb{E} \left[\mathbb{1}\{d_t = d\} \exp \left(i s \frac{1}{\delta^{d'}} x_t \right) | w_t, d_{t-1} = d' \right]}{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^{d'}} \varepsilon_t^{d'} \right) \right]} \\ &= \mathbb{E} \left[\mathbb{1}\{d_t = d\} \exp \left(i s \frac{1}{\delta^{d'}} x_t \right) | w_t, d_{t-1} = d' \right] \\ &\quad \times \frac{\exp \left[\int_0^{s/\delta^{d'}} \frac{\mathbb{E} \left[i \left(\frac{\delta^{d'}}{\delta^{d''}} x_{t+1} - \alpha^{d''} \delta^{d'} - \beta^{d''} \delta^{d'} w_t \right) \exp(i s_1 x_t) | d_t = d'', d_{t-1} = d' \right]}{\gamma^{d''} \mathbb{E}[\exp(i s_1 x_t) | d_t = d'', d_{t-1} = d']} ds_1 \right]}{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^{d'}} x_t \right) | d_t = d', d_{t-1} = d'' \right]} \end{aligned}$$

and

$$\begin{aligned} \phi_{x_t^* | w_t(d')}(s) &= \frac{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^{d'}} x_t \right) | w_t \right]}{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^{d'}} \varepsilon_t^{d'} \right) \right]} \\ &= \frac{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^{d'}} x_t \right) | w_t \right] \cdot \exp \left[\int_0^{s/\delta^{d'}} \frac{\mathbb{E} \left[i \left(\frac{\delta^{d'}}{\delta^{d''}} x_{t+1} - \alpha^{d''} \delta^{d'} - \beta^{d''} \delta^{d'} w_t \right) \exp(i s_1 x_t) | d_t = d'', d_{t-1} = d' \right]}{\gamma^{d''} \mathbb{E}[\exp(i s_1 x_t) | d_t = d'', d_{t-1} = d']} ds_1 \right]}{\mathbb{E} \left[\exp \left(i s \frac{1}{\delta^{d'}} x_t \right) | d_t = d', d_{t-1} = d'' \right]} \end{aligned}$$

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

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

This completes Step 4. □

References

- Abbring, Jaap H. and Jeffrey R. Campbell (2004) “A Firm’s First Year,” Working Paper.
- Abbring, Jaap H., and James J. Heckman (2007) “Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects and Dynamic Discrete Choice, and General Equilibrium Policy Evaluation,” in J.J. Heckman and E.E. Leamer (ed.) *Handbook of Econometrics*, Vol. 6, Ch. 72.
- 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.

- Andrews, Donald W.K. (1994) “Asymptotics for Semiparametric Econometric Models via Stochastic Equicontinuity,” *Econometrica*, Vol. 62, No.1, pp. 43–72.
- 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.
- Chen, Xiaohong, Oliver Linton and Ingrid van Keilegom (2003) “Estimation of Semiparametric Models when the Criterion Function Is Not Smooth,” *Econometrica*, Vol. 71, pp. 1591–1608.
- Cragg, John G. and Stephen G.Donald (1997) “Inferring the Rank of a Matrix,” *Journal of Econometrics*, Vol. 76, pp. 223–250.
- Cunha, Flavio and James J. Heckman (2008) “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation,” *Journal of Human Resources*, Vol. 43, pp. 738–782.
- 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.
- 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.
- Heckman, James J. and Salvador Navarro (2007) “Dynamic Discrete Choice and Dynamic Treatment Effects,” *Journal of Econometrics*, Vol. 136 (2), pp. 341–396.
- 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 (2013) “Closed-Form Estimation of Nonparametric Models with Non-Classical Measurement Errors,” Working Paper, Johns Hopkins University.
- 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 (1987) “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, Vol. 55, No. 5, pp. 999–1033.

- 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.
- Schennach, Susanne (2007) “Instrumental Variable Estimation of Nonlinear Errors-in-Variables Models,” *Econometrica*, Vol. 75, No. 1, pp. 201–239.
- Todd, Petra E. and Kenneth I. Wolpin (2012) “Estimating a Coordination Game in the Classroom,” Working Paper, University of Pennsylvania.
- Wilhelm, Daniel (2013) “Identification and Estimation of Nonparametric Panel Data Regressions with Measurement Error,” Working Paper
- 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.