UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Nonparametric Identification:

Identification of Dependent Multidimensional Unobserved Variables in a System of Linear Equations

Identification and Estimation for Regressions with Errors in All Variables

Identification of Nonparametrically Distributed Random Coefficients in Linear Panel Data Models

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Economics

by

Dan Ben-Moshe

2012

ABSTRACT OF THE DISSERTATION

Essays on Nonparametric Identification:
Identification of Dependent Multidimensional Unobserved Variables in a System of Linear Equations
Identification and Estimation for Regressions with Errors in All Variables
Identification of Nonparametrically Distributed Random Coefficients in Linear Panel Data Models

by

Dan Ben-Moshe
Doctor of Philosophy in Economics
University of California, Los Angeles, 2012
Professor Rosa Liliana Matzkin, Chair

In Chapter 1, I extend the techniques in Li and Vuong (1998), Schennach (2004a), and Bonhomme and Robin (2010) to identify nonparametric distributions of unobserved variables in a system of linear equations with more unobserved variables than outcome variables and with subsets of statistically dependent unobserved variables. I construct estimators of the distributions of unobserved variables and derive their uniform convergence rates. In Chapter 2, I develop a method for identification and estimation of coefficients in a linear regression model with measurement error in all the variables. The method is extended to identification in a system of linear equations in which only some of the coefficients on the unobserved variables are known. The estimator uses an assumption that is testable in the data and is in the class of Extremum estimators. The asymptotic distribution of the estimator is derived. In Chapter 3, I identify the nonparametric joint distribution of random coefficients in a linear panel data regression model. The distributions of the coefficients can depend on covariates, coefficients can be statistically dependent or equal in distribution, and there can be more coefficients than the fixed number of time periods. I construct estimators from the identification proofs. In finite sample simulations all the estimators have tight confidence bands around their theoretical counterparts.

The dissertation of Dan Ben-Moshe is approved.

Jinyong Hahn

Edward E. Leamer

Terence Tao

Raphael Thomadsen

Rosa Liliana Matzkin, Committee Chair

University of California, Los Angeles

2012

I dedicate this dissertation to my parents, Hanna and Yacov.

Contents

List of Figures

List of Tables

Preface

This thesis is concerned with identification in the system of linear equations

$$\begin{pmatrix} Y_{n1} \\ \vdots \\ Y_{nT} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{T1} & \vdots & a_{TM} \end{pmatrix} \begin{pmatrix} U_{n1} \\ \vdots \\ U_{nM} \end{pmatrix}$$

where $\vec{Y_n} = (Y_{n1}, \dots, Y_{nT})' \in \mathbb{R}^T$, $\vec{U_n}(U_{n1}, \dots, U_{nM})' \in \mathbb{R}^M$ and A is a $T \times M$ matrix with entries $\{a_{tm}\}$.[1]

Assume for now that the matrix A is known, the vector $\vec{Y_n}$ is known, and the vector $\vec{U_n}$ is unknown. If the dimension of $\vec{Y_n}$ is smaller than the dimension of $\vec{U_n}$ (i.e. $M > T$), then for any given value of $\vec{Y_n}$ there is in general no unique solution to $\vec{U_n}$. Usually, a system with fewer equations than unknown variables does not have a unique solution.

Now assume that $\vec{U_n}$, $n = 1, \dots, N$, are independent and identically distributed copies of an underlying nonparametrically distributed random vector $\vec{U}$. In this thesis I show that even when $M = P(P+1)/2$,

i. The joint distribution of $\vec{U}$ can be identified ("unique") and

ii. Some of the coefficients in the matrix A can be identified despite being unknown.

Kotlarski (1967) is the first person to identify nonparametric distributions in a system of linear equations with more unobserved variables than outcome variables. Consider

$$\begin{aligned} Y_1 &= U_1 + U_2 \\ Y_2 &= U_1 + U_3 \end{aligned} \tag{1}$$

He shows that if the distribution of $\vec{Y} = (Y_1, Y_2)$ is known and $\vec{U} = (U_1, U_2.U_3)$ is an unobserved independent random vector then $\vec{U}$ is identified.

[1] The subscript n represents the n^{th} observation or individual in the sample.

In Chapter 1, I prove that in a system of linear equations with 2 outcome variables, the maximum number of unobserved variables that are identified wihtout any additional information is 3. In Chapter 2, however, I show that the system in Equation (1) is still identified when it includes an unknown coefficient. Consider

$$Y_1 = U_1 + U_2$$
$$Y_2 = bU_1 + U_3$$

where $\vec{Y} = (Y_1, Y_2)$ is an observed random vector, $\vec{U} = (U_1, U_2.U_3)$ is an unobserved independent random vector, and b is an unknown coefficient. I show that b and the distribution of $\vec{U}$ are identified. In Chapter 3, I consider

$$Y_1 = U_1 + U_2 + U_3$$
$$Y_2 = aU_1 + U_2 + U_4 \qquad a^2 \neq 1$$

Assume that a is known and make the additional assumption that $U_3 \overset{d}{=} U_4$, then I show that all the distributions are still identified.

Chapter 1: Identification of Dependent Multidimensional Unobserved Variables in a System of Linear Equations

In Chapter 1, I study the system of linear equations

$$\vec{Y} = A\vec{U}$$

where $\vec{Y} \in \mathbb{R}^P$ is an observed random vector, $\vec{U} \in \mathbb{R}^M$ is an unobserved random vector, and A is a $P \times M$ matrix of known coefficients.

I identify the nonparametric distributions of the unobserved variables and explain the tradeoffs between the number of outcome variables, the number of unobserved variables, and the statistical dependence of the unobserved variables.

To illustrate the identification strategy I consider an earnings dynamics model from Bonhomme and Robin (2010) that is modeled as a system of linear equations with mutually independent unobserved variables. I relax various assumptions from Bonhomme and Robin (2010) and show identification. First, I allow a subset of unobserved variables to be arbitrarily dependent. Second, I assume that subsets of the unobserved variables are mean independent (but otherwise arbitrarily dependent). Third, I show that without adding additional equations or restrictions it is possible to include more unobserved variables and still identify all of the distributions.

Chapter 2: Identification and Estimation for Regressions with Errors in All Variables

In Chapter 2, I study the system of linear equations

$$\vec{Y} = \begin{pmatrix} A \\ B \end{pmatrix} \vec{U}$$

where $\vec{Y} \in \mathbb{R}^{T_A+T_B}$ is an observed random vector, $\vec{U} \in \mathbb{R}^M$ is an unobserved random vector, A is a $T_A \times M$ matrix of known coefficients, and B is a $T_B \times M$ matrix of unknown coefficients. In this chapter, I identify the coefficients in the matrix B.

I identify coefficients in three models:

i. Errors-in-Variables model:

$$Y = \beta_0 + \beta_1 X_1^* + \ldots + \beta_M X_M^* + \varepsilon$$

$$X_m = X_m^* + U_m \qquad\qquad m = 1, \ldots, M$$

where $(Y, X_1, \ldots, X_M)$ is an observed random vector and $(X_1^*, \ldots, X_M^*, U_1, \ldots, U_M, \varepsilon)$ is an unobserved mutually independent random vector. I identify $(\beta_0, \ldots, \beta_M)$ without any additional information.

ii. Moving-average process of order 1:

$$Y_1 = \varepsilon_1 - \theta\varepsilon_0$$
$$Y_2 = \varepsilon_2 - \theta\varepsilon_1$$

where (Y_1, Y_2) is an observed random vector and ε_0, ε_1, and ε_2 are unobserved mutually independent random variables. I identify θ without assuming that ε_0, ε_1, and ε_2 have equal variances.

iii. Simultaneous equations model:

$$Y_1 = \delta_1 Y_2 + \beta_1 X + \varepsilon_1$$
$$Y_2 = \delta_2 Y_1 + \varepsilon_2$$

where (Y_1, Y_2, X) is an observed random vector and ε_0 and ε_1 are conditionally independent unobserved random variables. I assume $E[X\varepsilon_2] = 0$ but do not place any restriction on the dependence of ε_1 on X. I identify the coefficients δ_1, δ_2, and β_1.

Chapter 3: Identification of Nonparametrically Distributed Random Coefficients in Linear Panel Data Models

In Chapter 3, I identify nonparametrically distributed random coefficients in the linear regression panel data model:

$$Y = X\beta + \varepsilon$$

where $Y \in \mathbb{R}^T$ is an observed random vector, X is a $P \times M$ matrix of covariates, $\varepsilon \in \mathbb{R}^T$ is a vector of errors, and β is a vector of random coefficients.

I identify the nonparametric joint distribution of the coefficients under various assumptions about the statistical dependence of coefficients on covariates, the conditional statistical relationship of coefficients (allowing them to be statistically dependent or equal in distribution), and the number of time periods per individual relative to the number of coefficients

I show how to identify random coefficients in a cross-sectional regression model with coefficients that are independent of covariates, in a panel data regression model with coefficients that are dependent on covariates, in a fixed effects regression model, and a first-order autoregressive panel data regression model.

Acknowledgments

It is with great pleasure that I thank the people who have been most influential in the preparation of this thesis. It is difficult to overstate how grateful I am to Rosa Matzkin, the Chair of my committee. At the beginning stages of my research she was an endless source of ideas and research topics. During our meetings, I spent hours telling her about what I was working on and she was always kind and patient even when my ideas were bad. When I hit roadblocks or wanted to give up, she offered encouragement and a contagious enthusiasm for economics and the research process. Without her I would never have completed this thesis.

I am in great debt to Jinyong Hahn who proved to be a perfect balance to Rosa. He provided the tough love by forcing me to think rigorously about my assumptions, motivations, and proofs and pushing me to produce my best. His fierce intelligence was evident by the speed in which he understood my research and solved problems that I spent many hours thinking about.

I owe much to Ed Leamer who challenged me to solve the problem in Chapter 2. My talks with him made me think about my research from different angles and why it is interesting to broader audiences.

I am also appreciative to the remaining members of my committee, Terence Tao and Raphael Thomadsen, who were open to meeting with me, offering support, and helping me with problems.

I thank Seth Kerstein for reading and offering extensive comments to many drafts of many parts of this dissertation and I thank Matthew Baird for also reading parts of this dissertation and for being an excellent officemate.

I was lucky to have Sanjay Unni as a boss for a few years. He taught me about perseverance and that there are both many obstacles towards a solution but also many paths that one can take towards it.

Curriculum Vitae

Education

- M.A. Economics, University of California Los Angeles 2009
- M.S. Financial Mathematics, Stanford University 2005
- B.S. Mathematics, B.A. Economics, Stanford University 2004

Work

- Teaching Assistant University of California Los Angeles, 2009-2011
- Research Assistant University of California Los Angeles, 2010
- Senior Associate LECG, 2005-2008
- Teaching Assistant Stanford University, 2005

Awards and Fellowships

- Dissertation Year Fellowship, University of California Los Angeles, 2011-2012
- NSF Fellow, 4th Lindau Nobel Laureate Meetings on Economic Sciences, 2011
- Welton Prize, Outstanding Student Paper in Econometrics, 2011
- Ettinger Prize, Best Pre-Job Market Paper, 2011
- Departmental Progress Award, 2008

Current Research

- Identification of Dependent Multidimensional Unobserved Variables in a System of Linear Equations
- Identification and Estimation for Regressions with Errors in All Variables
- Identification of Nonparametrically Distributed Random Coefficients in Linear Panel Data Models
- A Central Limit Theorem on Unobserved Random Variables
- Estimation of Nonparametric Distributions and Coefficients in an Earnings Dynamics Model (with Matthew Baird)

Chapter 1

Identification of Dependent Multidimensional Unobserved Variables in a System of Linear Equations

1.1 Introduction

Kotlarski (1967) studies identification of the unobserved variables in the system of linear equations

$$\begin{aligned} X_{n1} &= X_n^* + \varepsilon_{n1} \\ X_{n2} &= X_n^* + \varepsilon_{n2} \end{aligned} \tag{1.1}$$

where $(X_{n1}, X_{n2}) \in \mathbb{R}^2$ is a vector of observed outcomes and $(X_n^*, \varepsilon_{n1}, \varepsilon_{n2}) \in \mathbb{R}^3$ is a vector of unobserved variables.[1] This system has more unobserved variables than outcome equations so that for any observed (X_{n1}, X_{n2}) there is no unique solution of $(X_n^*, \varepsilon_{n1}, \varepsilon_{n2})$.[2] More

[1]The subscript n represents the n^{th} observation or individual in the sample.

[2]The solutions lie on the line $\varepsilon_{n1} - \varepsilon_{n2} = X_{n1} - X_{n2}$ in $\mathbb{R}^3$.

generally, for any system of linear equations with fewer equations than unknowns there are no unique solutions of the unknowns.

Kotlarski (1967), however, proved that if $(X_n^*, \varepsilon_{n1}, \varepsilon_{n2})$ are independent and identically distributed copies of an underlying independent random vector $(X^*, \varepsilon_1, \varepsilon_2)$, then its nonparametric distribution is identified ("uniquely determined") from the distributions of the observed outcome variables. In this paper I generalize Kotlarski (1967) in two ways:

i. Instead of a linear system with two outcome variables and three unobserved variables, I consider a general linear system with fewer outcome variables than unobserved variables;
ii. Instead of mutually independent unobserved variables, I allow the unobserved variables to be mean independent or arbitrarily dependent.

My aim is to identify the nonparametric distributions of the unobserved variables and to understand the tradeoffs between the number of outcome variables, the number of unobserved variables, and the statistical dependence of the unobserved variables.

To understand the tradeoffs I present three theorems. The first theorem extends the identification strategy in Li and Vuong (1998) from the system in Equation (1.1) with mutually independent unobserved variables to a system of equations with subsets of arbitrarily dependent unobserved variables.[3] The second theorem relaxes the mutual independence assumption from Bonhomme and Robin (2010) by providing necessary and sufficient conditions for identification in a system of linear equations in which subsets of the unobserved variables are arbitrarily dependent. The third theorem extends Schennach (2004a) from identification in the system in Equation (1.1) with mean independent unobserved variables to a system of equations.

My contributions are demonstrated in an earnings dynamics model from Bonhomme and Robin (2010) in which the unobserved variables are mutually independent permanent and transitory income shocks.[4] I solve this model relaxing various assumptions. First, I allow the

[3] Li and Vuong (1998) use the same identification strategy as Kotlarski (1967).

[4] With the exceptions of Horowitz and Markatou (1996) and Bonhomme and Robin (2010), the earnings dynamics literature assumes that the income shocks are jointly normal mutually independent unobserved variables. See Meghir and Pistaferri (2011) for a review of the earnings dynamics literature.

transitory shocks to be arbitrarily dependent. Second, I assume that the transitory shocks are mean independent (but otherwise arbitrarily dependent) and the permanent shocks are mean independent (but otherwise arbitrarily dependent). Third, I show that without adding additional equations or restrictions it is possible to include more unobserved variables and still identify all of the distributions.

The identification strategy takes advantage of the linearity of the system by a log characteristic function (CF) transformation. The result is an equation that expresses the log CF of a linear combination of the outcome variables in terms of additively separated log CFs of unobserved variables. Identification is achieved by taking partial derivatives and choosing a linear combination of outcome variables so that a single log CF of an unobserved variable is expressed in terms of observed quantities.

The estimators have closed form solutions coming from the identification results and are obtained by replacing population quantities with their sample analogs. I provide results on the uniform convergence rates of these estimators; similar to the estimators in Carroll and Hall (1988) and Fan (1991), these are relatively slow and depend on the smoothness of distributions of observed and unobserved variables.

In a Monte Carlo simulation, I compare several estimators of the distribution of X^* in Equations (1.1). The finite sample properties are encouraging with strong indications that the estimators should perform well in practice even with sample sizes of the outcome vector that are less than 100.

The literature on identification in models with more unobserved variables than outcome variables was initiated by Kotlarski (1967) and continued by Khatri and Rao (1968), Rao and Székely (2000), and others. Based on these papers, Li and Vuong (1998), Schennach (2004a,b), Bonhomme and Robin (2010), and others construct estimators.

The measurement error literature relaxes the additivity assumption by studying identification in nonlinear models.[5] Hu and Schennach (2007) are at the cutting edge of this

[5]See Schennach (2011) for a review of the measurement error literature in nonlinear models.

literature, using general operators to identify densities of unobserved variables. They use a completeness condition that requires strong restrictions on the dimension of the unobserved variables relative to the outcome variables.

This paper is organized as follows. Section 1.2 presents identification in an earnings dynamics model that explains the main identification ideas. Section 1.3 presents the general model, the assumptions, and the three main identification theorems. Section 1.4 presents an extension of the earnings dynamics model. Section 1.5 presents a few more illustrative examples that show how to use the identification theorems. Section 1.6 constructs the estimators and establishes their asymptotic properties. Section 1.7 presents Monte Carlo simulations. Section 1.8 concludes. The Appendix contains detailed solutions of the examples from Sections 1.2, 1.4, and 1.5 (Appendix A), the identification proofs from Section 1.3 (Appendix B), and proofs of the uniform convergence rates from Section 1.6 (Appendix C).

1.2 Example 1A: Earnings Dynamics Model

To explain the main identification ideas of this paper and compare them to the existing literature, consider the earnings dynamics model from Bonhomme and Robin (2010) on pages 494 and 495:

$$\begin{aligned} w_t &= f + y_t^P + y_t^T && t = 1,2,3,4 \\ y_t^P &= y_{t-1}^P + \varepsilon_t && t \geq 2 \\ y_t^T &= \eta_t \\ \eta_1 &= \eta_4 = 0 \end{aligned}$$

where w_t is the residual of a regression of individual log earnings on a set of strictly exogenous regressors, f is the unobserved fixed effect, y_t^P is the unobserved persistent component, y_t^T is the unobserved transitory shock, and ε_t and η_t are unobserved innovations that are mutually independent and independent over time. The system in matrix notation is

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} f \\ \eta_2 \\ \eta_3 \\ y_1^P \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} \tag{1.2}$$

The fixed effect f and the persistent component y_1^P, which are represented by the first and fourth columns in the matrix on the right of Equation (1.2), cannot be separately identified so Bonhomme and Robin (2010) difference out these effects. Let $\vec{Y} = (w_2 - w_1, w_3 - w_2, w_4 - w_3)'$ and $\vec{U} = (\eta_2, \eta_3, \varepsilon_2, \varepsilon_3, \varepsilon_4)'$. The system of equations on page 495 from Bonhomme and Robin (2010) is

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \end{pmatrix} \tag{1.3}$$

where Y_1, Y_2, and Y_3 are observed random variables and U_1, U_2, U_3, U_4, and U_5 are unobserved random variables with expected values equal to 0. Bonhomme and Robin (2010) assume that the unobserved random variables are mutually independent.

The first difference between the existing literature and my paper is that I relax the mutual independence assumption. Assume that U_1 and U_2 are arbitrarily dependent and (U_1, U_2), U_3, U_4, and U_5 are mutually independent.

I now solve for the joint distribution of the unobserved vector $\vec{U}$ in two ways. Solution

1, like Kotlarski (1967) and Li and Vuong (1998), uses first-order partial derivatives of log CFs. Solution 2, like Bonhomme and Robin (2010), uses second-order partial derivatives of log CFs.

Log CF transformation: The log CF of the observed vector $(Y_1, Y_2, Y_3) \in \mathbb{R}^3$ in terms of log CFs of unobserved variables is

$$\begin{aligned}
&\ln E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right] \\
&= \ln E\left[\exp\left(it_1(U_1 + U_3) + it_2(-U_1 + U_2 + U_4) + it_3(-U_2 + U_5)\right)\right] \\
&= \ln E\left[\exp\left(iU_1(t_1 - t_2) + iU_2(t_2 - t_3) + iU_3t_1 + iU_4t_2 + iU_5t_3\right)\right] \\
&= \ln E\left[\exp\left(iU_1(t_1 - t_2) + iU_2(t_2 - t_3)\right)\right] \\
&\quad + \ln E\left[\exp\left(iU_3t_1\right)\right] + \ln E\left[\exp\left(iU_4t_2\right)\right] + \ln E\left[\exp\left(iU_5t_3\right)\right] \qquad (1.4)
\end{aligned}$$

where the first equality follows by substituting $t_1Y_1 = t_1(U_1+U_3)$, $t_2Y_2 = t_2(-U_1+U_2+U_4)$, and $t_3Y_3 = t_3(-U_2 + U_5)$ and the last equality follows by the independence assumption.

The log CFs of the unobserved variables are additively separated because of the linearity in Equation (1.4) and the mutual independence of (U_1, U_2), U_3, U_4, and U_5. The random variables U_1 and U_2 are dependent so that their CFs cannot be separated and remain together in a multidimensional CF.

The notation I use in this paper is

$$\begin{aligned}
\varphi_{U_1,U_2}\left(t_1 - t_2, t_2 - t_3\right) &= \ln E\left[\exp\left(iU_1(t_1 - t_2) + iU_2(t_2 - t_3)\right)\right] \\
\varphi_{U_3}\left(t_1\right) &= \ln E\left[\exp\left(iU_3t_1\right)\right] \\
\varphi_{U_4}\left(t_2\right) &= \ln E\left[\exp\left(iU_4t_2\right)\right] \\
\varphi_{U_5}\left(t_3\right) &= \ln E\left[\exp\left(iU_5t_3\right)\right]
\end{aligned}$$

Using this notation

$$\ln E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right] = \varphi_{U_1,U_2}\left(t_1 - t_2, t_2 - t_3\right) + \varphi_{U_3}\left(t_1\right) + \varphi_{U_4}\left(t_2\right) + \varphi_{U_5}\left(t_3\right) \qquad (1.5)$$

For any $(s_1, s_2, s_3, s_4, s_5) \in \mathbb{R}^5$ there are in general no solutions $(t_1, t_2, t_3) \in \mathbb{R}^3$ that satisfy $\ln E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right] = \varphi_{U_1,U_2}(s_1, s_2) + \varphi_{U_3}(s_3) + \varphi_{U_4}(s_4) + \varphi_{U_5}(s_5)$.

1.2.1 Solution 1: First-Order Partial Derivatives

First-order partial derivative: The partial derivative of Equation (1.5) with respect to t_1 is

$$\begin{aligned} \frac{\partial \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1} &= \frac{iE\left[Y_1e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]} \\ &= \left.\frac{\partial \varphi_{U_1,U_2}(\omega_1, \omega_2)}{\partial \omega_1}\right|_{(t_1-t_2, t_2-t_3)} + \varphi'_{U_3}(t_1) \end{aligned} \tag{1.6}$$

Equation (1.6) has fewer functions than Equation (1.5). Only the log CFs containing U_1 and U_3 remain because of the substitution $t_1Y_1 = t_1(U_1 + U_3)$ into Equation (1.4). The log CFs of U_4 and U_5 vanish because of the additivity in Equation (1.5) and because their arguments do not contain t_1. The first-order partial derivative with respect to t_p reduces the equation to only contain log CFs of unobserved variables in the p^{th} equation. Hence, I refer to the partial derivative with respect to t_p as a "derivative with respect to the p^{th} equation." The effectiveness of the partial derivative depends on exclusion restrictions of unobserved variables from an equation.

Next, I show that for any $(s_1, s_2, s_3) \in \mathbb{R}^3$ there exists $(t_1, t_2, t_3) \in \mathbb{R}^3$ such that

$$\frac{iE\left[Y_1e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]} = \left.\frac{\partial \varphi_{U_1,U_2}(\omega_1, \omega_2)}{\partial \omega_1}\right|_{(s_1, s_2)} + \varphi'_{U_3}(s_3)$$

This means that $(t_1, t_2, t_3) \in \mathbb{R}^3$ solves

$$\begin{pmatrix} t_1 - t_2 \\ t_2 - t_3 \\ t_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}' \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \tag{1.7}$$

This matrix is the transpose of the first three columns of the matrix in Equation (1.3). These columns contain the coefficients of U_1, U_2, and U_3.

Choose (t_1, t_2, t_3)**:** For any $s_3 \in \mathbb{R}$ choose $(t_1, t_2, t_3) = (s_3, s_3, s_3)$ so that

$$\begin{pmatrix} t_1 - t_2 \\ t_2 - t_3 \\ t_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}' \begin{pmatrix} s_3 \\ s_3 \\ s_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ s_3 \end{pmatrix}$$

Substitute $(t_1, t_2, t_3) = (s_3, s_3, s_3)$ into Equation (1.6)

$$\begin{aligned} \frac{iE\left[Y_1 \exp\left(isY_1 + isY_2 + isY_3\right)\right]}{E\left[\exp\left(isY_1 + isY_2 + isY_3\right)\right]} &= \left.\frac{\partial \varphi_{U_1,U_2}\left(\omega_1, \omega_2\right)}{\partial \omega_1}\right|_{(0,0)} + \varphi'_{U_3}\left(s_3\right) \\ &= \varphi'_{U_3}\left(s_3\right) \end{aligned}$$

where the last equality follows from $\left.\frac{\partial \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1}\right|_{(0,0)} = iE\left[U_1\right]$ and the assumption that $E\left[U_1\right] = 0$.

The derivative of φ_{U_3} is now expressed in terms of observed quantities. The CF of U_3 is identified by integration:

$$E\left[\exp\left(iU_3 s_3\right)\right] = \exp\left(\int_0^{s_3} \frac{iE\left[Y_1 \exp\left(iu\left(Y_1 + Y_2 + Y_3\right)\right)\right]}{E\left[\exp\left(iu\left(Y_1 + Y_2 + Y_3\right)\right)\right]} du\right)$$

□[6]

The strategy in Solution 1 uses a first-order partial derivative of the log CF of $\vec{Y}$. The main assumption, Assumption 1i in Section 1.3, is that the image of the matrix transformation in Equation (1.7) contains either $(s_1, s_2, 0)'$ or $(0, 0, s_3)'$ where $(s_1, s_2) \in \mathbb{R}^2$ and $s_3 \in \mathbb{R}$.

[6]Appendix A identifies the rest of $\vec{U}$.

1.2.2 Solution 2: Second-Order Partial Derivatives

Second-order partial derivatives: The second-order partial derivative of Equation (1.5) with respect to t_1 and t_2 is

$$\frac{\partial^2 \ln E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]}{\partial t_1 \partial t_2} = -\left.\frac{\partial^2 \varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} + \left.\frac{\partial^2 \varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial \omega_1 \partial \omega_2}\right|_{(t_1-t_2,t_2-t_3)}$$

The log CFs of U_3, U_4, and U_5 vanish because of the additivity in Equation (1.5) and because their arguments do not contain t_1 and t_2. The second-order partial derivative with respect to t_{p_1} and t_{p_2} reduces the equation to only contain log CFs of unobserved variables in **both** the p_1^{th} and p_2^{th} equations.

All the second-order partial derivatives are

$$\begin{pmatrix} \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1^2} \\ \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2} \\ \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3} \\ \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2^2} \\ \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2 \partial t_3} \\ \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_3^2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1 \partial \omega_2}\right|_{(t_1-t_2,t_2-t_3)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} \\ \varphi''_{U_3}\left(t_1\right) \\ \varphi''_{U_4}\left(t_2\right) \\ \varphi''_{U_5}\left(t_3\right) \end{pmatrix} \tag{1.8}$$

It is instructive to set $(t_1, t_2, t_3) = (0, 0, 0)$ because the vector on the left hand side of Equation (1.8) will equal the vector of observed covariances, $Cov(Y_{p_1}, Y_{p_2})$, and the vector on the right hand side of Equation (1.8) will equal the vector of unobserved covariances, $Cov(U_{m_1}, U_{m_2})$. Hence, the entries in the matrix on the right hand side of Equation (1.8) are the same as the entries of the matrix that expresses $Cov(Y_{p_1}, Y_{p_2})$ in terms of $Cov(U_{m_1}, U_{m_2})$. Further, if

U_1 and U_2 are independent, then

$$\begin{aligned}
\left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} &= \varphi''_{U_1}(t_1-t_2) \\
\left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1 \partial \omega_2}\right|_{(t_1-t_2,t_2-t_3)} &= 0 \\
\left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_2^2}\right|_{(t_1-t_2,t_2-t_3)} &= \varphi''_{U_2}(t_2-t_3)
\end{aligned} \tag{1.9}$$

Equation (1.9) evaluated at $(t_1,t_2,t_3)=(0,0,0)$ becomes $Cov(U_1,U_2)=0$. The difference between Solution 2 and the solution from Bonhomme and Robin (2010) can be understood as the difference between allowing for $Cov(U_1,U_2)\neq 0$ and assuming $Cov(U_1,U_2)=0$.[7] The matrix in Equation (1.8) incorporates cross-partial derivatives and is the first step to dealing with the statistically dependent unobserved variables.

The matrix in Equation (1.8) is invertible so that all the second-order partial derivatives of log CFs of unobserved variables can be expressed in terms of observed quantities. For example,

$$\left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} = -\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2} - \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3} \tag{1.10}$$

Next, I show that for any $(s_1,s_2)\in\mathbb{R}^2$ there exists $(t_1,t_2,t_3)\in\mathbb{R}^3$ that

$$\left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(s_1,s_2)} = -\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2} - \frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3}$$

[7]The matrix in Equation (1.8) is the same as the one from Bonhomme and Robin (2010) when the unobserved variables are mutually independent.

This means that $(t_1, t_2, t_3) \in \mathbb{R}^3$ solves

$$\begin{pmatrix} t_1 - t_2 \\ t_2 - t_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}' \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \tag{1.11}$$

The matrix is the transpose of the first two columns of the matrix in Equation (1.3). These columns contain the coefficients of U_1 and U_2.

Choose (t_1, t_2, t_3): For any $(s_1, s_2) \in \mathbb{R}^2$ choose $(t_1, t_2, t_3) = (s_1, 0, -s_2)$ so that

$$\begin{pmatrix} t_1 - t_2 \\ t_2 - t_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}' \begin{pmatrix} s_1 \\ 0 \\ -s_2 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}$$

Substitute $(t_1, t_2, t_3) = (s_1, 0, -s_2)$ into Equation (1.10)

$$\left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1, \omega_2)}{\partial \omega_1^2}\right|_{(s_1,s_2)} = -\left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2}\right|_{(s_1,0,-s_2)} \left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3}\right|_{(s_1,0,-s_2)}$$

The CFs of the unobserved variables are identified by integration. For example,

$$\begin{aligned}\phi_{U_1,U_2}(s_1, s_2) = \exp\Bigg(& \int_0^{s_1}\int_0^{v} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1, \omega_2)}{\partial \omega_1^2}\right|_{(u,0)} \mathrm{d}u\mathrm{d}v \\ & + \int_0^{s_2}\int_0^{s_1} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1, \omega_2)}{\partial \omega_1 \omega_2}\right|_{(u,v)} \mathrm{d}u\mathrm{d}v + \int_0^{s_2}\int_0^{v} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1, \omega_2)}{\partial \omega_2^2}\right|_{(0,u)} \mathrm{d}u\mathrm{d}v \Bigg)\end{aligned}$$

□[8]

The strategy in Solution 2 uses the second-order partial derivatives of the log CF of $\vec{Y}$. The main assumptions, Assumptions 2i and 2ii in Section 1.3, are that the matrix in Equation (1.8) of all second-order partial derivatives is invertible and that the image of the matrix transformation in Equation (1.11) spans $\mathbb{R}^2$.

[8] Appendix A identifies the rest of $\vec{U}$.

1.3 Identification in the General Setup

An important aspect of this paper is that subsets of unobserved variables can be statistically dependent. To make this explicit, let $\vec{U} = (\vec{U}_1', \ldots, \vec{U}_M')'$ where $\vec{U}_m = (U_{m1}, \ldots, U_{mK_m})'$ is a vector of arbitrarily dependent real random variables. Assume that the vectors $\vec{U}_m \in \mathbb{R}^{K_m}$, $m = 1, \ldots, M$ are mutually independent. Let $\vec{Y} \in \mathbb{R}^P$ and consider the system of equations

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_P \end{pmatrix} = \begin{pmatrix} a^1_{11} & \ldots & a^1_{1K_1} \\ \vdots & \ddots & \vdots \\ a^1_{P1} & \ldots & a^1_{PK_1} \end{pmatrix} \begin{pmatrix} U_{11} \\ \vdots \\ U_{1K_1} \end{pmatrix} + \ldots + \begin{pmatrix} a^M_{11} & \ldots & a^M_{1K_M} \\ \vdots & \ddots & \vdots \\ a^M_{P1} & \ldots & a^M_{PK_M} \end{pmatrix} \begin{pmatrix} U_{M1} \\ \vdots \\ U_{MK_M} \end{pmatrix} \tag{1.12}$$

or $\vec{Y} = A_1\vec{U}_1 + \ldots + A_M\vec{U}_M = A\vec{U}$ where A_m is the $P \times K_m$ matrix with entries $\left\{a^m_{pk}\right\}_{p=1,k=1}^{P,K_m}$ and $A = (A_1, \ldots, A_M)$ is a partition of the $P \times \sum\limits_{m=1}^{M} K_m$ matrix A.

Define

$$A^{p^*} = \left(A_1^{p^*} \ \ldots \ A_M^{p^*}\right) = \left(A_1 \mathbf{I}\left(\bigcup_k a^1_{p^*k} \neq 0\right) \ \ldots \ A_M \mathbf{I}\left(\bigcup_k a^M_{p^*k} \neq 0\right)\right)$$

the matrix that includes the matrix A_m if and only if at least one of the columns of A_m has a nonzero coefficient in the p^{*th} row.[9,10] The image of $A^{p^*\prime}$ is a subspace with dimension equal to the number of unobserved variables that are dependent with unobserved variables in the p^{*th} equation. Assumption 1i, the main identifying assumption, is a condition on the span of the image of A^{p^*}.

If A is the matrix from the model in Equation (1.3) and $\vec{U} = (\vec{U}_1, U_3, U_4, U_5)$ where

[9] The function $\mathbf{I}(E)$ is the indicator function.
[10] Zero columns are removed from all matrices in this paper.

$\vec{U}_1 = (U_1, U_2)$ (the same dependence structure as in Example 1A), then

$$A^1 = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \qquad A^2 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \end{pmatrix} \qquad A^3 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

A^1 is the same matrix as in Equation (1.7) and contains the first three columns of A because $\vec{U}_1$ and U_3 have nonzero coefficients in the 1^{st} row.

Assumption 1. *There exists* $p_{k^*} \in \{1, \ldots, P\}$, *and* $\vec{t}_{m^*} = (t_{m^*1}, \ldots, t_{m^*P})'$ *for* $k^* = 1, \ldots, K_{m^*}$ *such that*

i. $$A^{p_{k^*}\prime}\vec{t}_{m^*} = \begin{pmatrix} A_1^{p_{k^*}\prime}\vec{t}_{m^*} \\ \vdots \\ A_M^{p_{k^*}\prime}\vec{t}_{m^*} \end{pmatrix} = \begin{pmatrix} \vec{0}_{\sum_{m<m^*} K_m} \\ \vec{s}_{m^*} \\ \vec{0}_{\sum_{m>m^*} K_m} \end{pmatrix}$$

ii. $a^{m^*}_{p_{k^*}k} = 0$ *for all* $k \neq k^*$

where $\vec{0}_J = (0, \ldots, 0)'$ *is a column vector with* J *zeros and* $\vec{s}_{m^*} = (s_{m^*1}, \ldots, s_{m^*K_{m^*}})'$.[11]

Assumption 1i implies that the image of $A^{p_{k^*}\prime}$ spans $(0, \ldots, 0, \vec{s}'_{m^*}, 0, \ldots, 0)'$ where $\vec{s}_{m^*} \in \mathbb{R}^{K_{m^*}}$. Assumption 1ii implies that coefficients of unobserved variables that are dependent with U_{m^*k} are zero in the p^{*th} equation.[12] Assumption 1ii is always satisfied when all the unobserved variables are mutually independent (i.e. $K_m = 1$ for all m).

[11] Assumption 1i can be replaced by several other assumptions. For example, if $\text{Rank}(A^{p_{k^*}}) \geq \sum_{m=1}^{M} \mathbf{I}(a_{p_{k^*}m} \neq 0)$ then the marginal distributions of $a_{p_{k^*}m}U_m$, $m = 1, \ldots, M$ are identified.

[12] In Solution 1: $A^1 = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}$ and Assumption 1 was satisfied by $(t_1, t_2, t_3) = (s_3, s_3, s_3)$ so that $\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}' \begin{pmatrix} s_3 \\ s_3 \\ s_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ s_3 \end{pmatrix}$.

Theorem 1. *If* $\int_0^{s_k} \left| \left(E\left[\exp i\left(U_{m^*1}s_1 + \ldots + U_{m^*k-1}s_{k-1} + U_{m^*k}u_k\right)\right]\right)^{-1} \right| \mathrm{d}u_k < \infty$ *for all fixed* $s_1, \ldots, s_{k-1}$ *and all* s_k *in the support of the CF of* $\vec{U}_{m^*}$*,* $E\left[|U_{m^*k}|\right] < \infty$*, and* $\vec{U}$ *has zero mean then the joint distribution of* $\vec{U}_{m^*}$ *is identified when Assumption 1 holds.*[13] *The CF of* $\vec{U}_{m^*}$ *is*

$$\phi_{m^*}(\vec{s}_{m^*}) =$$
$$\exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{a^{m^*}_{p_k k}} \int_0^{s_k} \frac{iE\left[Y_{p_{k^*}} \exp\left(i\vec{Y}' \left(A^{p_{k^*}}\right)^+ \left(\vec{0}_{\sum_{m<m^*} K_m}, s_1, \ldots, s_{k-1}, u_k, 0, \ldots, 0, \vec{0}_{\sum_{m>m^*} K_m}\right)'\right)\right]}{E\left[\exp\left(i\vec{Y}' \left(A^{p_{k^*}}\right)^+ \left(\vec{0}_{\sum_{m<m^*} K_m}, s_1, \ldots, s_{k-1}, u_k, 0, \ldots, 0, \vec{0}_{\sum_{m>m^*} K_m}\right)'\right)\right]} du_k\right) \tag{1.13}$$

where $\left(A^{p_{k^*}}\right)^+$ *is the Moore-Penrose pseudoinverse of* $A^{p_{k^*}}$*.*[14]

Remark 1. *Identification of* $\vec{U}$ *is achieved sequentially by:*

(1) Using Theorem 1 to identify unobserved variables,

(2) Moving the unobserved variables that are identified in step (1) (and that mutually independent of the other unobserved variables) to the left hand side of the equation and treating them as part of Y*.*

Remark 2. *Using Equation* (1.13)*, the mean and variance of* U_{mk} *are*

$$E\left[U_{mk}\right] = i^{-1}\phi'_{mk}(0) = \frac{E\left[Y_{p^*}\right]}{a^m_{p^*k}}$$
$$Var\left(U_{mk}\right) = \frac{1}{a^m_{p^*k}}\left(E\left[Y_{p^*}\vec{Y}'\vec{t}_m\right] - E\left[Y_{p^*}\right]E\left[\vec{Y}'\vec{t}_m\right]\right)$$

This implies that if $\widetilde{p} \neq p^*$ *then expectations and variances of estimators based on* $\widetilde{p}$ *and* p^* *may differ. Furthermore, if* $\vec{Y}'\vec{\widetilde{t}}_m \neq \vec{Y}'\vec{t}_m$ *then variances of estimators based on* $\vec{Y}'\vec{\widetilde{t}}_m$ *and* $\vec{Y}'\vec{t}_m$ *may differ. Hence, if the dependence structure of the unobserved variables is misspecified then an estimator of the distribution of* U_{mk} *based on Equation* (1.13) *will in*

[13] The condition that $\vec{U}$ has zero mean can be weakened to knowing or identifying $\sum_{(m,k)\neq(m^*,k^*)} a^m_{p_{k^*}k} E\left[U_{mk}\right]$.

[14] The proofs of the theorems in this section are in Appendix B.

general be inconsistent.

Remark 3. *The CF of U_{mk} is overidentified if A^{p_k} or $\vec{t}_m$ are not unique.*[15] *Overidentification suggests the possibility for testing and opens the possibility for a "best" estimator. Neither of these topics are studied in this paper.*

The theorem that follows relies on an assumption about a matrix that has the same coefficients as the matrix representation of the covariance of $\vec{Y}$ in terms of the covariance of $\vec{U}$:

$$\begin{aligned} Cov(Y_{p_1}, Y_{p_2}) &= Cov\left(\sum_{m_1=1}^{M}\sum_{k_1=1}^{K_{m_1}} a^{m_1}_{p_1k_1} U_{m_1k_1}, \sum_{m_2=1}^{M}\sum_{k_2=1}^{K_{m_2}} a^{m_2}_{p_2k_2} U_{m_2k_2}\right) \\ &= \sum_{m=1}^{M}\left(\sum_{k=1}^{K_m} a^m_{p_1k}a^m_{p_2k} Cov(U_{mk}, U_{mk}) + \sum_{k_1<k_2}\left(a^m_{p_1k_1}a^m_{p_2k_2} + a^m_{p_1k_2}a^m_{p_2k_1}\right) Cov(U_{mk_1}, U_{mk_2})\right) \end{aligned} \tag{1.14}$$

where the second equality follows because $Cov(U_{m_1k_1}, U_{m_2k_2}) = 0$ when $m_1 \neq m_2$ and $Cov(U_{mk_1}, U_{mk_2}) = Cov(U_{mk_2}, U_{mk_1})$. The coefficients are: $a^m_{p_1k}a^m_{p_2k}$ and $a^m_{p_1k_1}a^m_{p_2k_2} + a^m_{p_1k_2}a^m_{p_2k_1}$.

Let $A_m = (A^m_1, \ldots, A^m_{K_m})$ be a partition of the matrix A_m where A^m_k is the k^{th} column of A_m. Define the matrix multiplication

$$\begin{aligned} &A_m * A_m := \\ &\left(A^m_1 \otimes A^m_1,\ A^m_1 \otimes A^m_2 + A^m_2 \otimes A^m_1\ , \ldots,\ A^m_k \otimes A^m_k\ , \ldots,\ A^m_k \otimes A^m_{k+j} + A^m_{k+j} \otimes A^m_k\ , \ldots,\ A^m_{K_m} \otimes A^m_{K_m}\right) \end{aligned}$$

where $\otimes$ is the Kronecker product and $1 \leq j \leq K_m - k$. The matrix $A_m * A_m$ has dimension $P^2 \times K_m(K_m + 1)/2$.[16]

[15] Another reason for overidentification is that the system of equations $\vec{Y} = A\vec{U}$ is first transformed to $\vec{\tilde{Y}} = B\vec{Y} = BA\vec{U} = \tilde{A}\vec{U}$ and then unobserved variables are identified.

[16] The matrix $A_m * A_m$ has some repeated rows because the order of the scalar multiplication does not matter, that is $a^m_{p_1k_1}a^m_{p_2k_2} = a^m_{p_2k_2}a^m_{p_1k_1}$ so for calculation purposes I remove repeated rows and define the matrix $A_m \bar{*} A_m$ as the matrix $A_m * A_m$ without repeated rows so that a typical row looks like

$$\left[a^m_{p1}a^m_{p+j1}\ , \ldots,\ a^m_{pk_1}a^m_{p+jk_2} + a^m_{p+jk_1}a^m_{pk_2}\ , \ldots,\ a^m_{pK_m}a^m_{p+jK_m}\right]$$

where $0 \leq j \leq P - p$.The matrix $A_m \bar{*} A_m$ has dimension $(P + 1)P/2 \times K_m(K_m + 1)/2$.

Let $A = (A_1, \ldots, A_M)$ be a partition of the matrix A and define the matrix multiplication

$$A \odot A := (A_1 * A_1 \ , \ldots , \ A_M * A_M)$$

The matrix $A \odot A$ has dimension $P^2 \times \sum_{m=1}^{M} K_m(K_m + 1)/2$.[17]

When $K_m = 1$ then all the unobserved variables are mutually independent and

$$\begin{aligned} A_m &= A_1^m = (a_{11}^m \ \ldots \ a_{P1}^m)' \\ A_m * A_m &= (A_m \otimes A_m) \\ A \odot A &= (A_1 \otimes A_1 \ , \ldots , A_M \otimes A_M) \end{aligned}$$

$A \odot A$ is the same as the matrix Q from Bonhomme and Robin (2010) and the central part of their identification strategy. As mentioned earlier one of the contributions beyond Bonhomme and Robin (2010) is to show how to deal with dependent unobserved variables.

If A is the matrix from the model in Equation (1.3) and $\vec{U} = (\vec{U}_1, U_3, U_4, U_5)$ where $\vec{U}_1 = (U_1, U_2)$ (the same dependence structure as in Example 1A), then

$$A \mathbin{\bar{\odot}} A = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$A \mathbin{\bar{\odot}} A$ is the same matrix as in Equation (1.8). Inversion of this matrix was one of the steps towards identification in Solution 2.

---

[17] The matrix $A \odot A$ has some repeated rows so for calculation purposes define

$$A \mathbin{\bar{\odot}} A := (A_1 \mathbin{\bar{*}} A_1, \ \ldots, \ A_M \mathbin{\bar{*}} A_M)$$

This matrix $A \mathbin{\bar{\odot}} A$ has dimension $P(P+1)/2 \times \sum_{m=1}^{M} K_m(K_m + 1)/2$

Assumption 2.

i. $Rank(A \odot A) = \sum_{m=1}^{M} K_m(K_m+1)/2$

ii. $Rank(A_m) = K_m$ *for all* m

Theorem 2. *If* $\int_0^{s_{k_2}} \int_0^{s_{k_1}} \left(E\left[\exp\left(i \sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2} \right) \right] \right)^{-2} \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} < \infty$ *for all fixed* $s_1, \ldots, s_{k_1-1}$ *and all* s_{k_1}, s_{k_2} *in the support of the CF of* $\vec{U}_m$, $E\left[|U_{mk_1} U_{mk_2}|\right] < \infty$, *and* $\vec{U}$ *has zero mean then the joint distribution of* $\vec{U}_m$ *is identified* ***if and only if*** *Assumption 2 holds. The CF of* $\vec{U}_m$ *is*

$$\phi_m(\vec{s}_m) = \exp\left(\sum_{k=1}^{K_m} \int_0^{s_k} \int_0^{v_k} \left. \frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{mk}^2} \right|_{(0,\ldots,u_k,0,\ldots,0)} \mathrm{d}u_k \mathrm{d}v_k + \sum_{k_1<k_2} \int_0^{s_{k_2}} \int_0^{s_{k_1}} \left. \frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{mk_1} \partial \omega_{mk_2}} \right|_{(s_1,\ldots,s_{k_1-1},u_{k_1},0,\ldots,0,u_{k_2},0,\ldots,0)} \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} \right) \tag{1.15}$$

where

$$\left(\ldots \left. \frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{m1}^2} \right|_{\vec{s}'_m}, \ldots, \left. \frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{mK_m}^2} \right|_{\vec{s}'_m} \ldots \right)' = (A \odot A)^+ \left(\left. \frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1^2} \right|_{(A'_1)^+ \vec{s}_m}, \ldots, \left. \frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_P^2} \right|_{(A'_m)^+ \vec{s}_m} \right)'$$

and[18]

$$\left. \frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_{p_1} \partial t_{p_2}} \right|_{(A'_m)^+ \vec{s}_m} = \frac{E\left[Y_{p_1} e^{i\vec{Y}'(A'_m)^+ \vec{s}_m} \right] E\left[Y_{p_2} e^{i\vec{Y}'(A'_m)^+ \vec{s}_m} \right]}{\left(E\left[e^{i\vec{Y}'(A'_m)^+ \vec{s}_m} \right] \right)^2} - \frac{E\left[Y_{p_1} Y_{p_2} e^{i\vec{Y}'(A'_m)^+ \vec{s}_m} \right]}{E\left[e^{i\vec{Y}'(A'_m)^+ \vec{s}_m} \right]}$$

Assumptions 2i and 2ii are necessary and sufficient conditions for identification. They provide the connection between the number of outcome equations, P, the number of subsets that have arbitrarily dependent unobserved variables, M, and the number of unobserved variables in each subset, $K_1, \ldots, K_M$.

The number of linearly independent rows in $A \odot A$ is at most $P(P+1)/2$ and the number of linearly independent rows in A_m is at most P so by Assumptions 2i and 2ii respectively,

[18] $(A \odot A)^+$ is the Moore-Penrose pseudoinverse of $A \odot A$.

$\sum_{m=1}^{M} K_m(K_m+1) \leq P(P+1)$ and $K_m \leq P$ for all m. But $\sum_{m=1}^{M} K_m(K_m+1) \leq P(P+1)$ implies $K_m \leq P$ for all m. So for a given number of outcome variables, P, the number of subsets that have arbitrarily dependent unobserved variables, M, and the number of unobserved variables in each subset, $K_1, \ldots, K_M$ must satisfy

$$\sum_{m=1}^{M} \frac{K_m(K_m+1)}{2} \leq \frac{P(P+1)}{2} \tag{1.16}$$

When all the unobserved variables are mutually independent, for example, then $K_m = 1$ for all m and M must satisfy $M \leq \frac{P(P+1)}{2}$ for identification. The earnings dynamics model in Equation (1.2) has $P = 4$ so a maximum of $P(P+1)/2 = 10$ mutually independent unobserved variables can be identified. After differencing to the model in Equation (1.3) $P = 3$ so a maximum of $P(P+1)/2 = 6$ mutually independent unobserved variables can be identified. In Section 1.4 I extend the earnings dynamics model in Equation (1.2) from Bonhomme and Robin (2010) by identifying 10 mutually independent unobserved variables, the maximum number that is possible with $P = 4$.[19,20]

Remark 4. *The matrices A^{p^*} and $A \odot A$ are connected. Assume for this discussion that all the unobserved variables are mutually independent ($K_m = 1$ for all m) then*

$$A^{p^*} = \begin{pmatrix} A_1^{p^*} & \ldots & A_M^{p^*} \end{pmatrix} = \left(A_1 \mathbf{I}\left(a_{p^*1}^{1} \neq 0\right) \;\; \ldots \;\; A_M \mathbf{I}\left(a_{p^*1}^{M} \neq 0\right) \right)$$

and

$$A \odot A = (A_1 \otimes A_1 \;, \ldots, A_M \otimes A_M) = \begin{pmatrix} \vdots & \vdots & \vdots \\ a_{p^*1}A_1 & \ldots & a_{p^*M}A_M \\ \vdots & \vdots & \vdots \end{pmatrix}$$

[19] In Example 1A, $P = 3$ and $\vec{U} = ((U_{11}, U_{12}), U_2, U_3, U_4)$. So $K_1 = 2$, $K_2 = 1$, $K_3 = 1$, and $K_4 = 1$. $\sum_{m=1}^{4} K_m(K_m+1)/2 = P(P+1)/2 = 6$.

[20] There are some combinatorial questions that might be of interest. For example, for a given P, how many subsets $\{K_1, K_2, \ldots\}$ of positive integers with $K_1 \leq K_2 \leq \ldots$ satisfy $\sum_{m=1}^{M} K_m(K_m+1) = P(P+1)$?

$$= \begin{pmatrix} \vdots & \vdots & \vdots \\ a^1_{p^*1} A_1 \mathbf{I}\left(a^1_{p^*1} \neq 0\right) & \ldots & a^M_{p^*1} A_1 \mathbf{I}\left(a^M_{p^*1} \neq 0\right) \\ \vdots & \vdots & \vdots \end{pmatrix}$$

$$= \begin{pmatrix} \vdots & \vdots & \vdots \\ a^1_{p^*1} A^{p^*}_1 & \ldots & a^M_{p^*1} A^{p^*}_M \\ \vdots & \vdots & \vdots \end{pmatrix}$$

The part of $A \odot A$ *that is visible, call it* $(A \odot A)^{p^*} := (a^1_{p^*1} A^{p^*}_1 \ \ldots \ a^M_{p^*1} A^{p^*}_M)$, *is different from* $A^{p^*} = (A^{p^*}_1 \ \ldots \ A^{p^*}_M)$ *only by multiplication of each column by a nonzero constant. The properties of* $(A \odot A)^{p^*}$ *and* A^{p^*} *are identical. Hence, Assumption 1i, which is a condition on* A^{p^*}, *can be replaced by an equivalent condition on* $(A \odot A)^{p^*}$.

Identification in Theorem 1 uses the information contained in the partial derivatives separately and sequentially while Theorem 2 uses the information from all the partial derivatives together.[21]

Remark 5. *Theorems 1 and 3 provide sufficient conditions for identification while Theorem 2 provides* ***necessary*** *and sufficient conditions for identification. The drawback of Theorem 2 is that it uses second-order partial derivatives of the log CF instead of first-order partial derivatives of the log CF.*

Setting up a system of equations of third-order (or higher-order) partial derivatives of the log CF leads to more equations and can identify partial derivatives of more unobserved variables. The problem is that integrating out the derivatives requires knowing higher order moments of the unobserved variables.[22]

[21] The spatial model from Bonhomme and Robin (2010)

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1 & \rho & \rho & 1 & 0 & 0 \\ \rho & 1 & \rho & 0 & 1 & 0 \\ \rho & \rho & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{pmatrix}$$

is identified using Theorem 2 but not using Theorem 1.

[22] In Theorem 2, the assumption that $\vec{U}$ has zero mean is used to undo derivatives: the mean is the value

Remark 6. *When* $K_m = 1$ *for all* m *then Equation* (1.15) *simplifies to the solution from Bonhomme and Robin (2010):*

$$\phi_m(s_m) = \exp\left(\int_0^{s_m}\int_0^{v_m} \varphi''_m(u_m)\,\mathrm{d}u_m\mathrm{d}v_m\right)$$

The next Theorem identifies marginal distributions when arbitrary dependence is replaced by mean independence. Mean independence is a strong assumption that implies zero covariance but allows for the unobserved variables to be dependent in other ways. This theorem extends Theorem 1 in Schennach (2004a) and Theorem 1 in Cunha, Heckman and Schennach (2010) from the system in Equation (1.1) to a general system in Equation (1.12).

Let $A = (A_{11}, \ldots, A_{MK_M})$ be a partition of A where A_{mk} is the k^{th} column of the matrix A_m and define

$$A^{p^*m^*} = \left(A_{11}^{p^*m^*} \;\ldots\; A_{MK_M}^{p^*m^*}\right) = \left(A_{11}\mathbf{I}\left(\left\{a^1_{p^*1} \neq 0\right\} \cup \{m^* = 1\}\right) \;\ldots\; A_{MK_M}\mathbf{I}\left(\left\{a^M_{p^*K_M} \neq 0\right\} \cup \{m^* = M\}\right)\right)$$

the matrix that excludes the column A_{mk} if it has a zero coefficient in the p^{*th} row and U_{mk} and $U_{(mk)^*}$ are independent. The additional part in the Indicator function, $\{m^* = m\}$, will be used to condition on unobserved variables and lead to some terms vanishing because of mean independence.

Assumption 3. *There exists a* $p^* \in \{1, \ldots, P\}$ *and a vector* $\vec{t}_{(mk)^*} = \left(t_{(mk)^*1}, \ldots, t_{(mk)^*P}\right)'$, *where* $(mk)^* := \sum_{m<m^*} K_m + k^*$ *is an index, such that*

at the limit of an integral at 0. If higher-order moments are used for identification then the value of the integral at its limit is a variance (or higher-order moment).

In Theorems 1 and 3 the assumption that $\vec{U}$ has zero mean serves a different purpose: the value of the first-derivative of a log CF at 0 is its mean.

i. $A^{p^*m^*\prime}\vec{t}_{(mk)^*} = \begin{pmatrix} A_1^{p^*m^*\prime}\vec{t}_{(mk)^*} \\ \vdots \\ A_M^{p^*m^*\prime}\vec{t}_{(mk)^*} \end{pmatrix} = \vec{e}_{(mk)^*}$.[23]

ii. $E\left[U_{mk}|U_{-(mk)}\right] = 0$.[24]

Assumption 3i implies that the image of $A^{p^*\prime}$ spans $(0,\ldots,0,s,0,\ldots,0)'$ where $s \in \mathbb{R}$ is in the $(mk)^{*th}$ coordinate. Assumption 3ii is the mean independence assumption.[25] It replaces Assumption 1ii that required $a_{p^*k}^{m^*} = 0$ for all $k \neq k^*$ i.e. coefficients of dependent unobserved variables equal zero in the p^{*th} equation.

Theorem 3. *If* $\int\limits_0^{s_{(mk)^*}} \left|\left(E\left[\exp\left(iU_{(mk)^*}u\right)\right]\right)^{-1}\right| \mathrm{d}u < \infty$ *for all* $s_{(mk)^*}$ *in the support of the CF of* $U_{(mk)^*}$, $E\left[\left|U_{(mk)^*}\right|\right] < \infty$, *and* $\vec{U}$ *has zero mean then* $U_{(mk)^*}$ *is identified when Assumption 3 holds. The CF of* $U_{(mk)^*}$ *is*

$$\phi_{(mk)^*}(s_{(mk)^*}) = \;= \exp\left(\frac{1}{a_{p^*k^*}^{m^*}}\int_0^{s_{(mk)^*}} \frac{iE\left[Y_{p^*}\exp\left(iu\vec{Y}'\left(A^{p^*m^*\prime}\right)^+\vec{e}_{(mk)^*}\right)\right]}{E\left[\exp\left(iu\vec{Y}'\left(A^{p^*m^*\prime}\right)^+\vec{e}_{(mk)^*}\right)\right]}\mathrm{d}u\right) \quad (1.17)$$

Remark 7. *Several papers analyze the regularity conditions that impose restrictions on the CFs. The early measurement error literature (and literature on deconvolution) followed the Kotlarski (1967) assumption of nonvanishing CFs; Fan (1991) and Li and Vuong (1998) assume nonvanishing CFs on finite support while Schennach (2004a, 2004b) assumes nonvanishing CFs on infinite support. Bondesson (1974) was the first to prove identification when CFs satisfy a "short gap" condition, which meant that the CFs do not vanish on intervals of length L for all L > 0. More recently, Delaigle, Hall and Meister (2008), Carrasco and Flo-*

[23] The standard basis is denoted by $\vec{e}_{(mk)^*} = (0,\ldots,0,1,0,\ldots,0)'$ where 1 is in the $(mk)^{*th}$ coordinate.
[24] $U_{-(mk)} = \left(U_{(m1)},\ldots,U_{(mk-1)},U_{(mk+1)},\ldots,U_{(mK_m)}\right)$
[25] Assumption 3ii can be weakened by keeping track of the unobserved variables with zero coefficients in the p^{*th} row.

rens (2010) and Evdokimov and White (2011) restrict some of the CFs to have a countable number of isolated zeros on unbounded support and other CFs to have no regularity restrictions.[26] *In Theorems 1, 2, and 3, I impose an integrability condition that is motivated by the closed form expressions for the CFs of the unobserved variables.*[27] *The closed form solutions suggest that the weakest regularity condition would be based on the absolute continuity of a CF of an unobserved variable with respect to a CF of outcome variables.*

1.4 Example 1B: An Extension of the Earnings Dynamics Model

In this section I identify unobserved variables in an Earnings Dynamics model that extends the model from Bonhomme and Robin (2010) that was replicated in Equation (1.2). By conceding that the fixed effect and the persistent component are not separately identified, I show (without differencing) how to identify 10 unobserved variables instead of just 5.

Consider,

$$
\begin{aligned}
w_t &= f + y_t^P + m_t + y_t^T && t = 1, 2, 3, 4 \\
y_t^P &= y_{t-1}^P + \varepsilon_t && t \geq 2 \\
m_t &= \eta_t \\
y_t^T &= \zeta_t - \theta_1 \zeta_{t-1} - \theta_2 \zeta_{t-2} && \zeta_{-1} = \zeta_3 = \zeta_4 = 0 \\
\eta_4 &= 0
\end{aligned}
$$

The differences between this model and the one from Bonhomme and Robin (2010) are:

[26] Allowing for a countable number of zeros is important because some commonly used parametric distributions have CFs that cross the x-axis (for example the uniform and gamma distributions) but none of these disappear on a set of nonzero Lebesgue measure and then reappear.

[27] In Theorem 1, for example, in order for the CF of $\vec{U}_{m^*}$ in Equation (1.13) to be defined, I impose $\int_0^{s_k} \left| \left(E \left[\exp i \left(\ldots U_{m^* k} u_k \ldots \right) \right] \right)^{-1} \right| \mathrm{d}u_k < \infty$

- y_t^T is relabeled m_t and now represents measurement error,
- y_t^T follows a moving-average process of order 2,
- η_1 is no longer restricted to be equal to zero.

Let $Y = (w_1, w_2, w_3, w_4)'$ and $U = (f + y_1^P, \varepsilon_2, \varepsilon_3, \varepsilon_4, \eta_1, \eta_2, \eta_3, \zeta_0, \zeta_1, \zeta_2)'$ then in matrix notation

$$Y = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & -\theta_1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & -\theta_2 & -\theta_1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & -\theta_2 & -\theta_1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & -\theta_2 \end{pmatrix} U$$

Assume $E[U_m] = 0$ and assume all the unobserved variables are mutually independent.

Set $p^* = 1$. Then

$$A^1 = \begin{pmatrix} 1 & 1 & -\theta_1 & 1 \\ 1 & 0 & -\theta_2 & -\theta_1 \\ 1 & 0 & 0 & -\theta_2 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

where A^1 consists of the first, fifth, eighth, and ninth columns of A. When $t_1 = s_1(0, 0, 0, 1)$ then $A^{1\prime} t_1 = s\vec{e}_1$ where $s \in \mathbb{R}$ so Assumption 1i is satisfied. Using Equation (1.13), the CF of $\eta_{f+y_1^P}$ is

$$\phi_{f+y_1^P}(s_1) = \exp\left(\int_0^{s_1} \frac{iE\left[Y_1 \exp\left(iuY_4\right)\right]}{E\left[\exp\left(iuY_4\right)\right]} \mathrm{d}u\right)$$

Appendix A identifies the rest of $\vec{U}$.[28]

[28]The parameters θ_1 and θ_2 can be identified using Ben-Moshe (2012a).

1.5 A Few More Illustrative Examples

In this section I solve the earnings dynamics model one last time allowing for mean independence. I then provide two further examples to show that the methods in this paper can be used in a variety of settings and can allow for covariates.[29]

1.5.1 Example 1C: The Earnings Dynamics Model with Mean Independence

Consider the earnings dynamics model from Equation (1.3). Assume $\vec{U}_1 = (U_{11}, U_{12})$ and $\vec{U}_2 = (U_{21}, U_{22}, U_{23})$ are independent and assume $E[U_{mk}|U_{-(mk)} = 0]$ for all k and m.

Set $p^* = 2$ and $m^* = 1$. Then

$$A^{21} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

When $\vec{t}_{11} = s_{11}(1,0,0)$ then $A^{21\prime} = s_{11}\vec{e}_{11}$ where $s_{11} \in \mathbb{R}$ so Assumption 3i is satisfied. Using Equation (1.17) the CF of U_{11} is

$$\phi_{U_{11}}(s_{11}) = \exp\left(\int_0^{s_{11}} \frac{iE\left[Y_2 \exp\left(iuY_1\right)\right]}{E\left[\exp\left(iuY_1\right)\right]} \mathrm{d}u\right)$$

Appendix A identifies the rest of $\vec{U}$.

1.5.2 Example 2: Difference-in-Differences Model

Consider a difference-in-differences model with two periods and two groups. In the first period all individuals are in state 0 and in the second period individuals in group $g \in \{C, T\}$ (where C stands for control and T stands for treatment) go to state g. Hence, there are

[29] See also Ben-Moshe (2012b) for identification of random coefficients in linear regression models.

three states $t \in \{0, C, T\}$. Let Y_{gt} be the outcome for an individual in group g in state t. Assume that outcomes are represented by

$$Y_{C0} = m_C(X_C, \alpha_C) + h_0(W_0, \beta_0) + \varepsilon_{C0}$$
$$Y_{T0} = m_T(X_T, \alpha_T) + h_0(W_0, \beta_0) + \varepsilon_{T0}$$
$$Y_{CC} = m_C(X_C, \alpha_C) + h_C(W_C, \beta_C) + \varepsilon_{CC}$$
$$Y_{TT} = m_T(X_T, \alpha_T) + h_T(W_T, \beta_T) + \varepsilon_{TT}$$

where m_C and m_T are nonparametric production functions of individuals in groups C and T respectively, and h_0, h_C and h_T are nonparametric production functions of states 0, C and T respectively. The covariates X_g and W_t are observed variables, the random variables α_g and β_t are unobserved heterogeneity, and ε_{gt} is an unobserved idiosyncratic shock.

If an individual in the control group had instead been treated in the second period then the unobserved counterfactual outcome is assumed to be

$$Y^*_{CT} = m_C(X_C, \alpha_C) + h_T(W_T, \beta_T) + \varepsilon_{CT}$$

If an individual that is treated had instead been a part of the control group then the unobserved counterfactual outcome is assumed to be

$$Y^*_{TC} = m_T(X_T, \alpha_T) + h_C(W_C, \beta_C) + \varepsilon_{TC}$$

I focus on identifying the distribution of (Y^*_{CT}, Y^*_{TC}), the counterfactual outcomes, which are the objects of interest in the difference-in-differences literature.[30,31]

[30]For a review on the difference-in-differences literature see Angrist and Krueger (2000) and Blundell and MaCurdy (2000).

[31]Bonhomme and Sauder (2011) consider a similar model and apply it to compare the effects of different education systems. In their setup all students attend the same type of primary school but two different types of secondary schools. The outcome variables are test scores, one source of unobserved heterogeneity is child-specific ability that may be distributed differently for children in different groups, and another source of heterogeneity is a school-specific effect that may be distributed differently depending on the type of school.

Condition on $\vec{X} := (X_C, X_T, W_0, W_C, W_T) = (x_C, x_T, w_0, w_C, w_T) =: \vec{x}$ and let $\vec{Y} = (Y_{C0}, Y_{T0}, Y_{CC}, Y_{TT})'$ and $\vec{U} = (m_C, m_T, h_0, h_C, h_T, \varepsilon_{C0}, \varepsilon_{T0}, \varepsilon_{CC}, \varepsilon_{TT})'$. Then[32]

$$Y = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} U$$

Assume $E[U_m] = 0$ and assume U_1, U_2, U_3, U_4, U_5, (U_6, U_7) and (U_8, U_9) are mutually independent.[33]

As a preliminary step towards identifying counterfactuals, U_1, U_2, $U_4 + U_8$ and $U_5 + U_9$ are identified.[34,35] With one additional assumption that is defined later, the distribution of (Y^*_{CT}, Y^*_{TC}) is identified.

Set $p^* = 1$. Then

$$A^1 = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where A^1 consists of the first, third, sixth, and seventh columns of A. When $\vec{t_1} = s_1(0, 0, 1, 0)$ then $A^{1\prime}\vec{t_1} = s_1\vec{e}_1$ where $s_1 \in \mathbb{R}$ so Assumption 1i is satisfied (Assumption 1ii follows imme-

[32]To save on notation I denote $m_C(x_C, \alpha_C)$ by m_C, $m_T(x_T, \alpha_T)$ by m_T, $h_0(w_0, \beta_0)$ by h_0, $h_C(w_C, \beta_C)$ by h_C and $h_T(w_T, \beta_T)$ by h_T.

[33]The unobservables U_6 and U_7 are arbitrarily dependent. The unobservables U_8 and U_9 are arbitrarily dependent.

[34]It is impossible to separately identify U_4 and U_8 since they appear only once and in the same equation. Similarly, U_5 and U_9 are not separately identified.

[35]The distributions of U_3 and (U_6, U_7) are also identified but not needed for this example.

diately from mutual independence). Using Equation (1.13), the CF of $m_C(x_C, \alpha_C)$ is

$$\phi_{m_C}(s_1|\vec{X} = \vec{x}) = \exp\left(\int_0^{s_1} \frac{iE\left[Y_{C0}\exp\left(iuY_{CC}\right)|\vec{X} = \vec{x}\right]}{E\left[\exp\left(iuY_{CC}\right)|\vec{X} = \vec{x}\right]}\mathrm{d}u\right)$$

In Appendix A I identify the distributions of m_T and $(h_C + \varepsilon_{CC}, h_T + \varepsilon_{TT})$ and the counterfactual joint distribution of (Y^*_{CT}, Y^*_{TC}) with one of two possible assumptions

i. The joint distribution of $(\varepsilon_{CT}, \varepsilon_{TC})$ is the same as $(\varepsilon_{TT}, \varepsilon_{CC})$ or

ii. The joint distribution of $(\varepsilon_{CT}, \varepsilon_{TC})$ is the same as $(\varepsilon_{TT} - \varepsilon_{T0} + \varepsilon_{C0},\ \varepsilon_{CC} - \varepsilon_{C0} + \varepsilon_{T0})$

Remark 8. *Example 2 is related to models on wage decomposition in which an individual in group $g \in \{1, \ldots, G\}$ and job $t \in \{1, \ldots, T\}$ has wage*

$$W_{gt} = \Lambda^{gt}\left(m_g(X_g, \alpha_g) + h_t(W_t, \beta_t) + \varepsilon_{gt}\right)$$

where Λ^{gt} is a known invertible function, m_g and h_t are nonparametric production functions, X_g and W_t are observed covariates, α_g and β_t are unobserved heterogeneity, and ε_{gt} is an idiosyncratic shock. This can be used to estimate distributions of counterfactual wages for individuals in the same group but with different jobs like in an occupational choice model or for individuals in different groups but with the same job as in Juhn, Murphy, and Pierce (1991) who consider black-white wage differentials.

1.5.3 Example 3: Measurement Error Model With Three Measurements

Consider the measurement error model with three measurements

$$X_1 = X^* + \varepsilon_1$$

$$X_2 = X^* + \varepsilon_2$$

$$X_3 = X^* + \varepsilon_3$$

where (X_1, X_2, X_3) is observed and X^* and $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are unobserved.

Let $\vec{Y} = (X_1, X_2, X_3)'$ and $\vec{U} = (X^*, \varepsilon_1, \varepsilon_2, \varepsilon_3)'$ then

$$Y = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} U$$

Assume $E[U_m] = 0$ and assume (U_1, U_2), U_3 and U_4 are mutually independent (X^* and ε_1 are arbitrarily dependent). Then

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad A_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad A \bar{\odot} A = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$K_1 = 2$, $K_2 = 1$, and $K_3 = 1$ so $\sum_{m=1}^{3} K_m(K_m+1)/2 = (2\times 3)/2+1+1 = 5$. Rank$(A \bar{\odot} A) = 5$ so $(A \bar{\odot} A) = \sum_{m=1}^{3} K_m(K_m+1)/2 = 5$ and Assumption 2i is satisfied. Rank$(A_1) = K_1 = 2$, Rank$(A_2) = K_2 = 1$, and Rank$(A_3) = K_3 = 1$ so Assumption 2ii is satisfied. Using Equation (1.15) the CF of (X^*, ε_1) is

$$\begin{aligned}\phi_{X^*,\varepsilon_1}(s_0, s_1) = \exp\Bigg(& \int_0^{s_1}\int_0^{v} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(u, s_2)}{\partial \omega_1^2} \mathrm{d}u\mathrm{d}v + \int_0^{s_2}\int_0^{s_1} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(u, v)}{\partial \omega_1 \omega_2} \mathrm{d}u\mathrm{d}v \\ & + \int_0^{s_1}\int_0^{s_2} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(0, u)}{\partial \omega_1 \partial \omega_2} \mathrm{d}u\mathrm{d}v + \int_0^{s_2}\int_0^{v} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(0, u)}{\partial \omega_2^2} \mathrm{d}u\mathrm{d}v\Bigg)\end{aligned}$$

Appendix A identifies the rest of $\vec{U}$.

Remark 9. *Let* $X_p = X^* + \varepsilon_p$*,* $p = 1, \ldots, P$*, and* $P \geq 2$*.* X_p *is the* p^{th} *measurement of the unobserved variable* X^**. Assume all the unobserved variables are mutually independent. Then a solution for the CF of* X^* *that uses all the observations is*

$$\phi_{X^*}(s) = \exp\left(\int_0^s \frac{iE\left[X_1 \exp\left(iu\frac{1}{P-1}\sum_{p=2}^{P} X_p\right)\right]}{\phi_{\left(\frac{1}{P-1}\sum_{p=2}^{P} X_p\right)}(u)} \mathrm{d}u\right)$$

Remark 10. *The measurement error model with repeated measurements can be extended to a model with more than one unobserved covariate as follows*

$$X_p = \sum_{m=1}^{M} X_m^* \mathbf{I}\left(X_m^* \in \{\text{Measurement } p\text{'s information set}\}\right) + \varepsilon_p \qquad p = 1, \ldots, P$$

where X_p*,* $p = 1, \ldots, P$ *are* P *observed measurements,* X_m^**,* $m = 1, \ldots, M$ *are* M *unobserved covariates,* $\mathbf{I}\left(X_m^* \in \{\textit{Measurement p's information set}\}\right)$ *is an indicator that* X_m^* *is included in equation p, and* ε_p*,* $p = 1, \ldots, P$ *are measurement errors.*[36]

1.6 Estimation and Asymptotics

In this section estimators for densities are constructed using the closed form solutions from Theorems 1, 2, and 3. I show that the estimators are uniformly consistent.

[36] Li, Perrigne and Vuong (2000) use the results of the measurement literature and a solution mechanism in a first price auction to identify distributions when each bidder has valuation $U_0 + A_p$, $p = 1, \ldots, P$ where U_0 is the common value, and A_p is a private value. This can be extended to a model with more than one common value. Consider,

$$Y_p = \sum_{m=1}^{M} U_m \mathbf{I}\left(X_m^* \in \{\text{Bidder p's information set}\}\right) + A_p \qquad p = 1, \ldots, P$$

where Y_p is the observed bid of bidder p, U_m, $m = 1, \ldots, M$ are unobserved common values, $\mathbf{I}\left(X_m^* \in \{\text{Bidder p's information set}\}\right)$ is an indicator that bidder p's valuation includes the common value U_m and A_p, $p = 1, \ldots, P$ are unobserved private values.

Denote

$$\phi_{\prod_{p=1}^{P} Y_p^{\alpha_p}}(\vec{t}) = \frac{\partial^{|\alpha|}\phi_{\vec{Y}}(\vec{t})}{\prod_{p=1}^{P}\partial^{\alpha_p}t_p} = i^{|\alpha|}E\left[\prod_{p=1}^{P} Y_p^{\alpha_p}\exp\left(i\vec{Y}'\vec{t}\right)\right]$$

and estimate it by

$$\widehat{\phi}_{\prod_{p=1}^{P} Y_p^{\alpha_p}}(\vec{t}) = \frac{\widehat{\partial^{|\alpha|}\phi_{\vec{Y}}(\vec{t})}}{\prod_{p=1}^{P}\partial^{\alpha_p}t_p} = i^{|\alpha|}E_N\left[\prod_{p=1}^{P} Y_p^{\alpha_p}\exp\left(i\vec{Y}'\vec{t}\right)\right] = \frac{i^{|\alpha|}}{N}\sum_{n=1}^{N}\prod_{p=1}^{P} Y_{np}^{\alpha_p}\exp\left(i\vec{Y}_n'\vec{t}\right)$$

where $\alpha = (\alpha_1, \ldots, \alpha_P)$ is a multi-index of nonnegative integers with norm $|\alpha| = \sum_{p=1}^{P}\alpha_p$. When $|\alpha| = 0$ then the expression is the CF of $\vec{Y}$ denoted by

$$\phi_{\vec{Y}}(\vec{t}) = E\left[\exp\left(i\vec{Y}'\vec{t}\right)\right]$$

and estimated by

$$\widehat{\phi}_{\vec{Y}}(\vec{t}) = E_N\left[\exp\left(i\vec{Y}'\vec{t}\right)\right] = \frac{1}{N}\sum_{n=1}^{N}\exp\left(i\vec{Y}_n'\vec{t}\right)$$

Assume that U_{m^*} is a scalar. The CF of U_{m^*} in Theorems 1 and 3, up to a constant and for some $\vec{t}$, is

$$\phi_{m^*}(s) = \exp\left(i\int_0^s \frac{E\left[Y_{p^*}\exp(iu\vec{Y}'\vec{t})\right]}{E\left[\exp\left(iu\vec{Y}'\vec{t}\right)\right]}\mathrm{d}u\right) \tag{1.18}$$

and is estimated by

$$\widehat{\phi}_{m^*}(s) = \exp\left(i\int_0^s \frac{E_N\left[Y_{p^*}\exp\left(u\vec{Y}'\vec{t}\right)\right]}{E_N\left[\exp\left(iu\vec{Y}'\vec{t}\right)\right]}\mathrm{d}u\right)$$

The CF of U_{m^*} in Theorem 2, up to a constant and for some $\vec{t}$, is [37]

$$\phi_{m^*}(s) = \exp\left(\int_0^s \int_0^v \frac{E\left[Y_{p_1} e^{iu\vec{Y}'\vec{t}}\right] E\left[Y_{p_2} e^{iu\vec{Y}'\vec{t}}\right]}{\left(E\left[e^{iu\vec{Y}'\vec{t}}\right]\right)^2} - \frac{E\left[Y_{p_1} Y_{p_2} e^{iu\vec{Y}'\vec{t}}\right]}{E\left[e^{iu\vec{Y}'\vec{t}}\right]} \mathrm{d}u\mathrm{d}v\right) \quad (1.19)$$

and is estimated by

$$\widehat{\phi}_{m^*}(s) = \exp\left(\int_0^s \int_0^v \frac{E_N\left[Y_{p_1} e^{iu\vec{Y}'\vec{t}}\right] E_N\left[Y_{p_2} e^{iu\vec{Y}'\vec{t}}\right]}{\left(E_N\left[e^{iu\vec{Y}'\vec{t}}\right]\right)^2} - \frac{E_N\left[Y_{p_1} Y_{p_2} e^{iu\vec{Y}'\vec{t}}\right]}{E_N\left[e^{iu\vec{Y}'\vec{t}}\right]} \mathrm{d}u\mathrm{d}v\right)$$

The density of U_{m^*} is obtained by inverting the CF using the inverse Fourier transformation

$$f_{m^*}(u) = \frac{1}{2\pi} \int e^{-isu} \phi_{m^*}(s)\mathrm{d}s$$

This integral does not converge when the CF is replaced by its sample analog so the integral is weighted by the Fourier transform of a kernel. The density of U_{m^*} is estimated by

$$\widehat{f}_{m^*}(u) = \frac{1}{2\pi} \int e^{-isu} \widehat{\phi}_{m^*}(s) \phi_K\left(sh_N\right)\mathrm{d}s$$

where $\phi_K(s) = \int \exp(isu)H(u)\mathrm{d}u$ is the Fourier transform of a kernel K supported on $[-1, 1]$ and $h_N = \frac{1}{S_N}$ is the bandwidth of the kernel. The kernel leads to relatively slow convergence rates but solves any irregularity problems by smoothing the estimator. I use the commonly

[37] To be more exact the CF is

$$\phi_{m^*}(s) = \exp\left(\sum C_{p'_1 p'_2 m^*} \int_0^s \int_0^v \frac{E\left[Y_{p'_1} e^{iu\vec{Y}'\vec{t}}\right] E\left[Y_{p'_2} e^{iu\vec{Y}'\vec{t}}\right]}{\left(E\left[e^{iu\vec{Y}'\vec{t}}\right]\right)^2} - \frac{E\left[Y_{p'_1} Y_{p'_2} e^{iu\vec{Y}'\vec{t}}\right]}{E\left[e^{iu\vec{Y}'\vec{t}}\right]} \mathrm{d}u\mathrm{d}v\right)$$

but assume for clarity that $C_{p'_1 p'_2 m^*} = 1$ when $p'_1 = p_1$ and $p'_2 = p_2$ and $C_{p'_1 p'_2 m^*} = 0$ otherwise.

used second-order kernel[38]

$$K(u) = \frac{48\cos(u)}{\pi u^4}\left(1 - \frac{15}{u^2}\right) - \frac{144\sin(u)}{\pi u^5}\left(2 - \frac{5}{u^2}\right)$$

whose Fourier transform is

$$\phi_K(s) = (1 - s^2)^3 \mathbf{I}(s \in [-1, 1])$$

Lemma 1. *Let F denote the cumulative distribution function of Y and F_N the empirical cumulative distribution function corresponding to a sample $(Y_1, \ldots, Y_N)$ of N independent identically distributed random draws from F. Assume $E\left[\prod_{p=1}^{P} |Y_p|^{2\alpha_p}\right] < \infty$. Let*

$$\begin{aligned} T_N &= CN^{\delta/2} & 0 < \delta \\ \varepsilon_N &= C_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)} \left(\frac{\ln N}{N}\right)^{\frac{1}{2}} \end{aligned}$$

where $C > 0$ and $C_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)} > 0$ is a constant that may depend on the arguments in the subscript. Then

$$\sup_{\vec{t}\in[-T_N,T_N]^P} \left| E_N\left[\prod_{p=1}^{P} Y_p^{\alpha_p} \exp\left(i\vec{Y}'\vec{t}\right)\right] - E\left[\prod_{p=1}^{P} Y_p^{\alpha_p} \exp\left(i\vec{Y}'\vec{t}\right)\right]\right| < \varepsilon_N \qquad a.s.$$

when N tends to infinity.[39,40]

As $N \to \infty$, Lemma 1 uniformly bounds the estimation error on the compact interval $[-T_N, T_N]^P$ by $O\left(\frac{\ln N}{N}\right)^{\frac{1}{2}}$ provided that T_N does not grow faster than some power of N.[41]

[38] See Delaigle and Gijbels (2002).

[39] To simplify notation I suppress the subscript $\vec{t} \in [-T_N, T_N]^P$ in $\sup_{\vec{t}\in[-T_N,T_N]^P}$ unless there is some ambiguity or the sup is not over this region.

[40] The proofs of the lemma and theorems in this section are in Appendix C.

[41] $Z_N = O(a_N)$ is Big-O notation and means that there exists $C > 0$ such that $Z_N \leq Ca_N$.

The strategy in the proof is standard for finding uniform convergence rates in the empirical processes literature:[42]

1. Use the truncation trick to divide the random variable into $E_N\left[\prod_{p=1}^{P} Y_p^{\alpha_p}\exp\left(i\vec{Y}'\vec{t}\right)\right] \leq \kappa_N$ and the tail, $E_N\left[\prod_{p=1}^{P} Y_p^{\alpha_p}\exp\left(i\vec{Y}'\vec{t}\right)\right] > \kappa_N$, where κ_N is a truncation parameter to be chosen later,
2. Use Chebyshev's inequality to estimate the tail,
3. Use symmetrization, the L_1 covering number, and Bernstein's inequality to estimate the component that is smaller than the truncation parameter,
4. Combine the two components and use the Borel-Cantelli lemma to show that the sample analog approaches the population mean uniformly almost surely.

Theorem 4. *Choose ε_N and T_N according to Lemma 1. Assume $\int_{-S_N}^{S_N}\frac{1}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\mathrm{d}u < \infty$ and $E\left[|Y_p^2|\right] < \infty$. The CF from Theorems 1 and 3, in Equation* (1.18)*, is uniformly bounded by*

$$\begin{aligned}\sup_{s\in[-S_N,S_N]}\left|\widehat{\phi}_{m^*}(s)-\phi_{m^*}(s)\right| &= \sup_{s\in[-S_N,S_N]}\left|\exp\left(\int_0^s\frac{\widehat{\phi}_{Y_p}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)-\exp\left(\int_0^s\frac{\phi_{Y_p}(u\vec{t})}{|\phi_{\vec{Y}}(u\vec{t})|}\mathrm{d}u\right)\right| \\ &= O\left(\varepsilon_N E\left[|Y_p|\right]\int_{-S_N}^{S_N}\frac{1}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\mathrm{d}u\right)\end{aligned}$$

Assume $\int_{-S_N}^{S_N}\frac{1}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^3}\mathrm{d}u < \infty$, $E\left[|Y_{p1}^2|\right] < \infty$, $E\left[|Y_{p2}^2|\right] < \infty$, and $E\left[|Y_{p1}^2 Y_{p2}^2|\right] < \infty$. The CF from Theorems 2, in Equation (1.19)*, is uniformly bounded by*

$$\begin{aligned}&\sup_{s\in[-S_N,S_N]}\left|\widehat{\phi}_{m^*}(s)-\phi_{m^*}(s)\right| \\ &= \sup_{s\in[-S_N,S_N]}\left|\exp\left(\int_0^s\int_0^v\frac{\widehat{\phi}_{Y_{p_1}}(u\vec{t})\widehat{\phi}_{Y_{p_2}}(u\vec{t})}{\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right)\right. \\ &\qquad\left.-\exp\left(\int_0^s\int_0^v\frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right)\right|\end{aligned}$$

[42] The argument can be found in Pollard (1986) or Van den Geer (2006) and is used by Hu and Ridder (2012), Evdokimov (2011), Bonhomme and Robin (2010), and others.

$$= O\left(\varepsilon_N \left(E\left[|Y_{p_1}|\right] + E\left[|Y_{p_2}|\right] + E\left[|Y_{p_1}Y_{p_2}|\right]\right) \int_{-S_N}^{S_N} \int_0^v \frac{1}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^3} \mathrm{d}u\mathrm{d}v\right)$$

Theorem 5. *Choose ε_N and T_N according to Lemma 1 and assume the convergence rates from Theorem 4 apply. Then*

$$\sup_u \left|\widehat{f}_{m^*}(u) - f_{m^*}(u)\right|$$
$$= O\left(\sup_{s\in[-S_N,S_N]} \left|\widehat{\phi}_{m^*}(s) - \phi_{m^*}(s)\right| + \sup_{s\in[-1,1]} |m(s)|\, h_N^q \int_{-S_N}^{S_N} |s|^q |\phi_{m^*}(s)| \mathrm{d}s \right.$$
$$\left. + \int_{-\infty}^{-S_N} |\phi_{m^*}(s)| \mathrm{d}s + \int_{S_N}^{\infty} |\phi_{m^*}(s)| \mathrm{d}s\right)$$

The first term in the convergence rate of $\widehat{f}_{m^*}(u)$, in Theorem 5, comes from the estimation error of ϕ_{m^*}, from Theorem 4. The second, third and fourth terms in the convergence rate of $\widehat{f}_{m^*}(u)$ in Theorem 5 come from the Fourier transform inversion, and depend on the smoothing kernel ϕ_K and its bandwidth h_N, the limits of integration $-S_N$ and S_N, and the CF of the unobserved variable, ϕ_{m^*}.[43]

The uniform bounds on the convergence rates in Theorems 1 and 3 suggest that estimators based on first-order partial derivatives converge faster than estimators based on second-order partial derivatives. The bounds in these Theorems are worse than Li and Vuong (1998) who obtain $O\left(\frac{\ln\ln N}{N}\right)^{\frac{1}{2}}$ but assume bounded support.[44]

[43] The constant in the big-O notation does not depend on the dimension of the vector of unobserved variables, M, but depends on the dimension of the outcome vector, P.

[44] The literature has so far only found upper bounds on convergence rates of estimators based on partial derivatives of CF and so at this stage the bounds are only suggestive about which estimators have the fastest convergence rates. Schennach (2004) and Schennach, White, and Chalak (2010) find asymptotic distributions for these types of estimators, which may be a good way to find the best estimators.

1.7 Monte Carlo Simulations: Measurement Error Model with a Repeated Measurement

This section presents a Monte Carlo study of the finite sample properties of three estimators of the density of X^* in the measurement error model with a repeated measurement:

$$X_{n1} = X_n^* + \varepsilon_{n1}$$

$$X_{n2} = X_n^* + \varepsilon_{n2}$$

where X_{n1} and X_{n2} are observed measurements, X_n^* is an unobserved variable, and ε_{n1} and ε_{n2} are errors for $n = 1, \ldots, N$. Assume samples are independent and identically distributed.

Two of the estimators for the density of X^* are based on first-order partial derivatives and one of the estimators is based on second-order partial derivatives. All the estimators perform very well in the simulations with the median estimates almost indistinguishable from the underlying theoretical density of X^*. This is evidence that these estimators should perform well in practice.

The data is generated from one of the following specifications of the distributions of X^*, ε_1, and ε_2

Experiment	f_{X^*}	f_{ε_1}	f_{ε_2}
1	Norm(0,1)	Norm(0,1)	Norm(0,1)
2	Gamma(5,1)	Norm(0,1)	Norm(0,1)
3	$\frac{1}{2}N(-2,1) + \frac{1}{2}N(2,1)$	Norm(0,1)	Norm(0,1)
4	Unif(0,2)	0	0
5	Norm(0,1)	Norm(0,x^{*2})	Norm(0,1)

where x^{*2} (the variance of ε_1 in Experiment 5) is the square of the value that is attained by the random variable X^* in each trial. I compare three estimators of ϕ_{X^*}:

	Estimator
A	$\widehat{\phi}_{X^*}(s) = \exp\left(\int_0^s \frac{iE_N[X_1\exp(iuX_2)]}{E_N[\exp(iuX_2)]}\mathrm{d}u\right)$
B	$\widehat{\phi}_{X^*}(s) = \frac{\phi_{X_1}(s)}{\widehat{\phi}_{\varepsilon_1}(s)}$ where $\widehat{\phi}_{\varepsilon_1}(t) = \exp\left(\int_0^s \frac{iE_N[(X_1-X_2)\exp(iuX_1)]}{E_N[\exp(iuX_1)]}\mathrm{d}u\right)$
C	$\widehat{\phi}_{X^*}(s)$ $= \exp\left(\int_0^s\int_0^v\left(-\frac{iE_N\left[X_1X_2\exp\left(\frac{iu}{2}(X_1+X_2)\right)\right]}{E_N\left[\exp\left(\frac{iu}{2}(X_1+X_2)\right)\right]} + \frac{E_N\left[X_1\exp\left(\frac{iu}{2}(X_1+X_2)\right)\right]}{E_N\left[\exp\left(\frac{iu}{2}(X_1+X_2)\right)\right]}\frac{E_N\left[X_2\exp\left(\frac{iu}{2}(X_1+X_2)\right)\right]}{E_N\left[\exp\left(\frac{iu}{2}(X_1+X_2)\right)\right]}\right)\mathrm{d}u\mathrm{d}v\right)$

where the first two estimators are constructed using Equation (1.13), and the third using Equation (1.15). The first estimator is used by Li and Vuong (1998), the second estimator has not been used to my knowledge, and the third estimator is used by Bonhomme and Robin (2010). I present evidence that all three estimators have good finite sample properties.

I generate 100 simulations of sample size $N = 100$, $N = 1,000$ and $N = 10,000$. The grid on the x-axis is divided into $1,000$ equidistant grid points for integration in both the CF and density domains.

The results are summarized graphically in Figures 1.1 to 1.5. Figure 1.1 reports the outcomes of 100 simulations of sample size 100 where the data is generated according to Experiment 1. The first column represents the real part of the CF, the second column represents the imaginary part of the CF, and the third column represents the density. On each graph the solid red line represents population quantities, the solid blue line represents the median of the simulations and the dotted blue lines represent the 10-90% pointwise confidence bands. The first row depicts the results of Estimator A, the second row depicts the results of Estimator B, and the third row depicts the results of Estimator C. Figures 1.2 to 1.5 are the same as Figure 1.1 except for Experiments 1.2 to 1.5.

To provide an indication of relative finite sample efficiencies of the estimators, Tables 1.1, 1.2 and 1.3 report the mean integrated squared error (MISE) of each estimator for $N = 100$, $N = 1,000$ and $N = 10,000$ respectively where

$$\text{MISE} = E\left[\int\left(\widehat{f}_{X^*}(x) - f_{X^*}(x)\right)^2 dx\right]$$

The median estimates do very well, lying almost on top of the theoretical CFs and densities. As expected, only Estimator A is consistent in Experiment 5 (due to the dependence structure of unobserved variables). The MISE values suggest that Estimator C, which is based on second-order partial derivatives is the least robust.

1.8 Conclusion

I consider a system of linear equations in which each observed outcome variable is a linear combination of unobserved variables. I present techniques to identify nonparametric distributions of unobserved variables. The system has more unobserved variables than outcome variables and subsets of the unobserved variables can be statistically dependent (either arbitrarily dependent or mean independent). I establish a relationship between the number of outcome variables, the number of unobserved variables, and the dependence of the unobserved variables. The identification strategy involves taking partial derivatives of log CFs to reduce the number of log CFs of unobserved variables and using the arguments of a log CF of a linear combination of outcome variables to express log CFs of unobserved variables in terms of observed quantities. I analyze the identification strategy in an earnings dynamics model from Bonhomme and Robin (2010). The identification proofs are constructive so estimators replace population quantities with sample analogs. The estimators are part of a general class of estimators that use partial derivatives of log CFs. I show that these estimators are consistent. In finite sample simulations, estimators closely match their theoretical counterparts.

1.9 Appendix A

1.9.1 Example 1A: Earnings Dynamics Model (Solution 1)

As mentioned earlier the unobserved variables are identified sequentially. Following the proof for identification of U_3, the log CF of (Y_1, Y_2, Y_3) is

$$\ln E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right] = \varphi_{U_1,U_2}\left(t_1 - t_2, t_2 - t_3\right) + \varphi_{U_3}\left(t_1\right) + \varphi_{U_4}\left(t_2\right) + \varphi_{U_5}\left(t_3\right)$$

The CF of U_4: The partial derivative with respect to t_2 is

$$\frac{iE\left[Y_2\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]}{E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]} = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_1}\right|_{(t_1-t_2,t_2-t_3)} + \left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(t_1-t_2,t_2-t_3)} + \varphi'_{U_4}\left(t_2\right)$$

Set $(t_1, t_2, t_3) = (s_4, s_4, s_4)$. Then

$$\frac{iE\left[Y_2\exp\left(is_4Y_1 + is_4Y_2 + is_4Y_3\right)\right]}{E\left[\exp\left(is_4Y_1 + is_4Y_2 + is_4Y_3\right)\right]} = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_1}\right|_{(0,0)} + \left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(0,0)} + \varphi'_{U_4}\left(s_4\right) = \varphi'_{U_4}\left(s_4\right)$$

where the last equality follows from $\varphi'_{U_4}(0) = iE[U_4]$ and the assumption that $E[U_4] = 0$.

$$E\left[\exp\left(iU_4s_4\right)\right] = \exp\left(\int_0^{s_4}\frac{iE\left[Y_2\exp\left(iu\left(Y_1 + Y_2 + Y_3\right)\right)\right]}{E\left[\exp\left(iu\left(Y_1 + Y_2 + Y_3\right)\right)\right]}du\right)$$

The CF of U_5: The partial derivative with respect to t_3 is

$$\frac{iE\left[Y_3\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]}{E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]} = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(t_1-t_2,t_2-t_3)} + \varphi'_{U_5}\left(t_3\right)$$

Set $(t_1, t_2, t_3) = (s_3, s_3, s_3)$. Then

$$\frac{iE\left[Y_3\exp\left(is_3Y_1 + is_3Y_2 + is_3Y_3\right)\right]}{E\left[\exp\left(is_3Y_1 + is_3Y_2 + is_3Y_3\right)\right]} = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(0,0)} + \varphi'_{U_5}\left(s_3\right) = \varphi'_{U_5}\left(s_3\right)$$

where the last equality follows from $\varphi'_{U_5}(0) = iE[U_5]$ and the assumption that $E[U_5] = 0$.

$$E\left[\exp\left(iU_5s_3\right)\right] = \exp\left(\int_0^{s_3}\frac{iE\left[Y_3\exp\left(iu\left(Y_1 + Y_2 + Y_3\right)\right)\right]}{E\left[\exp\left(iu\left(Y_1 + Y_2 + Y_3\right)\right)\right]}du\right)$$

The CF of (U_1, U_2): The partial derivative with respect to t_1 is

$$\frac{iE\left[Y_1 \exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]}{E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]} = \left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_1}\right|_{(t_1-t_2,t_2-t_3)} + \varphi'_{U_3}\left(t_1\right)$$

Set $(t_1, t_2, t_3) = (0, -s_1, -s_1)$. Then

$$\frac{iE\left[Y_1 \exp\left(-is_1\left(Y_2+Y_3\right)\right)\right]}{E\left[\exp\left(-is_1\left(Y_2+Y_3\right)\right)\right]} = \left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_1}\right|_{(s_1,0)} + \varphi'_{U_3}(0) = \left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_1}\right|_{(s_1,0)}$$

where the last equality follows from $\varphi'_{U_3}(0) = iE[U_3]$ and the assumption that $E[U_3] = 0$.

The partial derivative with respect to t_3 is

$$\frac{iE\left[Y_3 \exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]}{E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right]} = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(t_1-t_2,t_2-t_3)} + \varphi'_{U_5}\left(t_3\right)$$

Set $(t_1, t_2, t_3) = (s_1 + s_2, s_2, 0)$. Then

$$\frac{iE\left[Y_3 \exp\left(iY_1\left(s_1+s_2\right)+is_2Y_2\right)\right]}{E\left[\exp\left(iY_1\left(s_1+s_2\right)+is_2Y_2\right)\right]} = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(s_1,s_2)} + \varphi'_{U_5}(0) = -\left.\frac{\partial\varphi_{U_1,U_2}\left(\omega_1,\omega_2\right)}{\partial\omega_2}\right|_{(s_1,s_2)}$$

where the last equality follows from $\varphi'_{U_5}(0) = iE[U_5]$ and the assumption that $E[U_5] = 0$. Integration leads to

$$\begin{aligned}&E\left[\exp\left(iU_1s_1 + iU_2s_2\right)\right]\\&= \exp\left(\int_0^{s_1} \frac{iE\left[Y_1 \exp\left(-iu_1\left(Y_2+Y_3\right)\right)\right]}{E\left[\exp\left(-iu_1\left(Y_2+Y_3\right)\right)\right]}du_1 - \int_0^{s_2} \frac{iE\left[Y_3 \exp\left(iY_1\left(s_1-u_2\right)+iu_2Y_2\right)\right]}{E\left[\exp\left(iY_1\left(s_1+u_2\right)+iu_2Y_2\right)\right]}du_2\right)\end{aligned}$$

where I used

$$\varphi_{U_1,U_2}(s_1,s_2) = \int_0^{s_1} \left.\frac{\partial\varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial\omega_1}\right|_{(u_1,0)} du_1 + \int_0^{s_2} \left.\frac{\partial\varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial\omega_2}\right|_{(s_1,u_2)} du_2$$

1.9.2 Example 1A: Earnings Dynamics Model (Solution 2)

The log CF of (Y_1, Y_2, Y_3) is

$$\ln E\left[\exp\left(it_1Y_1 + it_2Y_2 + it_3Y_3\right)\right] = \varphi_{U_1,U_2}\left(t_1-t_2, t_2-t_3\right) + \varphi_{U_3}\left(t_1\right) + \varphi_{U_4}\left(t_2\right) + \varphi_{U_5}\left(t_3\right)$$

All the second-order partial derivatives are

$$\begin{pmatrix} \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1^2} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2^2} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2 \partial t_3} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_3^2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1 \partial \omega_2}\right|_{(t_1-t_2,t_2-t_3)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} \\ \varphi''_{U_3}(t_1) \\ \varphi''_{U_4}(t_2) \\ \varphi''_{U_5}(t_3) \end{pmatrix}$$

The inverse is

$$\begin{pmatrix} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1 \partial \omega_2}\right|_{(t_1-t_2,t_2-t_3)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(t_1-t_2,t_2-t_3)} \\ \varphi''_{U_3}(t_1) \\ \varphi''_{U_4}(t_2) \\ \varphi''_{U_5}(t_3) \end{pmatrix} = \begin{pmatrix} 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1^2} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2^2} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2 \partial t_3} \\ \dfrac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_3^2} \end{pmatrix}$$

All the second-order partial derivatives of the log CF of unobserved variables are solved for in terms of observed quantities.

For any $(s_1, s_2) \in \mathbb{R}^2$ choose $(t_1, t_2, t_3) = (s_1, 0, -s_2)$. Then

$$\begin{aligned} \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1^2}\right|_{(s_1,s_2)} &= -\left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_2}\right|_{(s_1,0,-s_2)} - \left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3}\right|_{(s_1,0,-s_2)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_1 \partial \omega_2}\right|_{(s_1,s_2)} &= -\left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3}\right|_{(s_1,0,-s_2)} \\ \left.\frac{\partial^2 \varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial \omega_2^2}\right|_{(s_1,s_2)} &= -\left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1 \partial t_3}\right|_{(s_1,0,-s_2)} - \left.\frac{\partial^2 \ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2 \partial t_3}\right|_{(s_1,0,-s_2)} \end{aligned}$$

Integrating out

$$\phi_{U_1,U_2}(s_1,s_2)=\exp\left(\int_0^{s_1}\int_0^{v}\left.\frac{\partial^2\varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial\omega_1^2}\right|_{(u,0)}\mathrm{d}u\mathrm{d}v+\int_0^{s_2}\int_0^{s_1}\left.\frac{\partial^2\varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial\omega_1\omega_2}\right|_{(u,v)}\mathrm{d}u\mathrm{d}v\right.$$
$$\left.+\int_0^{s_2}\int_0^{v}\left.\frac{\partial^2\varphi_{U_1,U_2}(\omega_1,\omega_2)}{\partial\omega_2^2}\right|_{(0,u)}\mathrm{d}u\mathrm{d}v\right)$$

Similarly for U_3 let $(t_1,t_2,t_3)=(s_3,0,0)$, for U_4 let $(t_1,t_2,t_3)=(0,s_4,0)$, and for U_5 let $(t_1,t_2,t_3)=(0,0,s_5)$. Then

$$\varphi''_{U_3}(s_3)=$$
$$\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1^2}\right|_{(s_3,0,0)}+\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1\partial t_2}\right|_{(s_3,0,0)}+\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1\partial t_3}\right|_{(s_3,0,0)}$$
$$\varphi''_{U_4}(s_4)=$$
$$\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1\partial t_2}\right|_{(0,s_4,0)}+\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2^2}\right|_{(0,s_4,0)}+\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2\partial t_3}\right|_{(0,s_4,0)}$$
$$\varphi''_{U_5}(s_5)=$$
$$\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_1\partial t_3}\right|_{(0,0,s_5)}+\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_2\partial t_3}\right|_{(0,0,s_5)}+\left.\frac{\partial^2\ln E\left[e^{it_1Y_1+it_2Y_2+it_3Y_3}\right]}{\partial t_3^2}\right|_{(0,0,s_5)}$$

Integrating out

$$\phi_{U_3}(s_3)=\exp\left(\int_0^{s_2}\int_0^{v}\varphi''_{U_3}(v)\,\mathrm{d}u\mathrm{d}v\right)$$
$$\phi_{U_4}(s_4)=\exp\left(\int_0^{s_3}\int_0^{v}\varphi''_{U_4}(v)\,\mathrm{d}u\mathrm{d}v\right)$$
$$\phi_{U_5}(s_4)=\exp\left(\int_0^{s_3}\int_0^{v}\varphi''_{U_5}(v)\,\mathrm{d}u\mathrm{d}v\right)$$

1.9.3 Example 1B: Extension of the Earnings Dynamics Model

To identify U_1, U_5, U_8 and U_9 set $p^*=1$. Then

$$A^1=\begin{pmatrix}1&1&-\theta_1&1\\1&0&-\theta_2&-\theta_1\\1&0&0&-\theta_2\\1&0&0&0\end{pmatrix}$$

Set $t_1 = s_1(0,0,0,1)$, $t_5 = s_5\left(1, -\frac{\theta_1}{\theta_2}, \frac{\theta_1^2+\theta_2}{\theta_2^2}, -\frac{\theta_1^2-\theta_1\theta_2+\theta_2^2+\theta_2}{\theta_2^2}\right)$, $t_8 = s_8\left(0, -\frac{1}{\theta_2}, \frac{\theta_1}{\theta_2^2}, \frac{\theta_2-\theta_1}{\theta_2^2}\right)$, and $t_9 = s_9\left(0,0,-\frac{1}{\theta_2}, \frac{1}{\theta_2}\right)$ where $s_1, s_5, s_8, s_9 \in \mathbb{R}$. Then ${A^1}' t_1 = s_1 \vec{e}_1$, ${A^1}' t_5 = s_5 \vec{e}_2$, ${A^1}' t_8 = s_8 \vec{e}_3$ and ${A^1}' t_9 = s_9 \vec{e}_4$ so Assumption1i is satisfied. Using Equation (1.13), the CFs of $f + y_1^P\ \eta_1$, ζ_0 and ζ_1 are

$$\phi_{f+y_1^P}(s_1) = \exp\left(\int_0^{s_1} \frac{iE\left[Y_1 \exp\left(iuY_4\right)\right]}{E\left[\exp\left(iuY_4\right)\right]} \mathrm{d}u\right)$$

$$\phi_{\eta_1}(s_5) = \exp\left(\int_0^{s_5} \frac{iE\left[Y_1 \exp\left(\frac{iu}{\theta_2^2}\left(Y_1\theta_2^2 - Y_2\theta_1\theta_2 + Y_3\theta_1^2 + Y_3\theta_2 - Y_4\theta_1^2 + Y_4\theta_1\theta_2 - Y_4\theta_2^2 - Y_4\theta_2\right)\right)\right]}{E\left[\exp\left(\frac{iu}{\theta_2^2}\left(Y_1\theta_2^2 - Y_2\theta_1\theta_2 + Y_3\theta_1^2 + Y_3\theta_2 - Y_4\theta_1^2 + Y_4\theta_1\theta_2 - Y_4\theta_2^2 - Y_4\theta_2\right)\right)\right]} \mathrm{d}u\right)$$

$$\phi_{\zeta_0}(s_8) = \exp\left(-\frac{1}{\theta_1}\int_0^{s_8} \frac{iE\left[Y_1 \exp\left(-\frac{iu}{\theta_2^2}\left(Y_2\theta_2 - Y_3\theta_1 + Y_4\theta_1 - Y_4\theta_2\right)\right)\right]}{E\left[\exp\left(-\frac{iu}{\theta_2^2}\left(Y_2\theta_2 - Y_3\theta_1 + Y_4\theta_1 - Y_4\theta_2\right)\right)\right]} \mathrm{d}u\right)$$

$$\phi_{\zeta_1}(s_9) = \exp\left(\int_0^{s_9} \frac{iE\left[Y_1 \exp\left(-\frac{iu}{\theta_2}\left(Y_3 - Y_4\right)\right)\right]}{E\left[\exp\left(-\frac{iu}{\theta_2}\left(Y_3 - Y_4\right)\right)\right]} \mathrm{d}u\right)$$

The unobserved variables U_1, U_6, U_8 and U_9 are identified and satisfy independence assumptions that allow a rearrangement of the system so that the identified unobserved variables can be treated as part of Y. Let

$$\widetilde{Y} = Y - A_1' U_1 - A_5' U_5 - A_8' U_8 - A_9' U_9 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & -\theta_1 \\ 1 & 1 & 1 & 0 & 0 & -\theta_2 \end{pmatrix} \begin{pmatrix} U_2 \\ U_3 \\ U_4 \\ U_6 \\ U_7 \\ U_{10} \end{pmatrix} = \widetilde{A}\widetilde{U}$$

To identify U_2, U_6 and U_{10} set $p^* = 2$. Then

$$\widetilde{A}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & -\theta_1 \\ 1 & 0 & -\theta_2 \end{pmatrix}$$

Set $t_2 = s_2\left(0,0,-\frac{\theta_2}{\theta_1-\theta_2}, \frac{\theta_1}{\theta_1-\theta_2}\right)$, $t_6 = s_6\left(0,1,\frac{1+\theta_2}{\theta_1-\theta_2}, -\frac{1+\theta_1}{\theta_1-\theta_2}\right)$ and $t_{10} = s_{10}\left(0,0,-\frac{1}{\theta_1-\theta_2}, \frac{1}{\theta_1-\theta_2}\right)$ where $s_2, s_6, s_{10} \in \mathbb{R}$. Then ${\widetilde{A}^2}' t_2 = s_2 \vec{e}_1$, ${\widetilde{A}^2}' t_6 = s_6 \vec{e}_2$ and ${\widetilde{A}^2}' t_{10} = s_{10} \vec{e}_3$ so Assumption 1i is satisfied. Using

Equation (1.13), the CFs of ζ_2, ε_2 and η_2 are

$$\phi_{\varepsilon_2}(s_2) = \exp\left(\int_0^{s_2} \frac{iE\left[Y_2 \exp\left(-\frac{iu}{\theta_1-\theta_2}\left(Y_3\theta_2 - Y_4\theta_1\right)\right)\right]}{E\left[\exp\left(-\frac{iu}{\theta_1-\theta_2}\left(Y_3\theta_2 - Y_4\theta_1\right)\right)\right]}\mathrm{d}u\right) \Big/ \left(\phi_{f+y_1^P}\left(s_2\right)\phi_{\zeta_1}\left(\frac{-s_2\theta_2^2}{\theta_1-\theta_2}\right)\right)$$

$$\phi_{\eta_2}(s_6) =$$

$$\exp\left(\int_0^{s_6} \frac{iE\left[Y_2 \exp\left(\frac{iu}{\theta_1-\theta_2}(Y_2(\theta_1-\theta_2)+Y_3(1+\theta_2)+Y_4(1+\theta_1))\right)\right]}{E\left[\exp\left(\frac{iu}{\theta_1-\theta_2}(Y_2(\theta_1-\theta_2)+Y_3(1+\theta_2)+Y_4(1+\theta_1))\right)\right]}\mathrm{d}u\right) \Big/ \left(\phi_{f+y_1^P}\left(s_6\right)\phi_{\zeta_1}\left(s_6\frac{\theta_1(\theta_2-\theta_1)-\theta_2(1+\theta_2)}{\theta_1-\theta_2}\right)\right)$$

$$\phi_{\eta_2}(s_{10}) = \exp\left(\int_0^{s_{10}} \frac{iE\left[Y_2 \exp\left(\frac{iu}{\theta_1-\theta_2}\left(-Y_3+Y_4\right)\right)\right]}{E\left[\exp\left(\frac{iu}{\theta_1-\theta_2}\left(-Y_3+Y_4\right)\right)\right]}\mathrm{d}u\right) \Big/ \phi_{\zeta_1}\left(\frac{-s_{10}\theta_2}{\theta_1-\theta_2}\right)$$

The unobserved variables U_1, U_2, U_5, U_6, U_8, U_9 and U_{10} are identified and satisfy independence assumptions so that ε_3, ε_4 and η_3 are identified in a similar way to the unobserved variables above

$$\phi_{\varepsilon_3}(s_3) = \exp\left(\int_0^{s_3} \frac{iE\left[Y_3 \exp\left(iuY_4\right)\right]}{E\left[\exp\left(iuY_4\right)\right]}\mathrm{d}u\right) \Big/ \left(\phi_{\varepsilon_1}\left(s_3\right)\phi_{\varepsilon_2}\left(s_3\right)\phi_{\varepsilon_2}\left(-\theta_2 s_3\right)\right)$$

$$\phi_{\varepsilon_4}(s_4) = \exp\left(\int_0^{s_4} \frac{iE\left[Y_4 \exp\left(iuY_4\right)\right]}{E\left[\exp\left(iuY_4\right)\right]}\mathrm{d}u\right) \Big/ \left(\phi_{f+y_1^P}\left(s\right)\phi_{\varepsilon_2}\left(s_4\right)\phi_{\varepsilon_3}\left(s_4\right)\phi_{\varepsilon_2}\left(-\theta_2 s_4\right)\right)$$

$$\phi_{\eta_3}(s) = \exp\left(\int_0^{s} \frac{iE\left[Y_3 \exp\left(iu(Y_3-Y_4)\right)\right]}{E\left[\exp\left(iu(Y_3-Y_4)\right)\right]}\mathrm{d}u\right) \Big/ \left(\phi_{\zeta_1}\left(-\theta_2 s_3\right)\phi_{\zeta_2}\left(-\theta_1 s_3 + \theta_2 s_3\right)\right)$$

1.9.4 Example 1C: Earnings Dynamics Model with Mean Independence

Set $p^* = 2$ and $m^* = 1$. Then

$$A^{21} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

When $\vec{t}_{11} = s_{11}(1,0,0)$ then $A^{21'} = s_{11}\vec{e}_{11}$ where $s_{11} \in \mathbb{R}$ so Assumption 3i is satisfied. Using Equation (1.17) the CF of U_{11} is

$$\phi_{U_{11}}(s_{11}) = \exp\left(\int_0^{s_{11}} \frac{iE\left[Y_2 \exp\left(iuY_1\right)\right]}{E\left[\exp\left(iuY_1\right)\right]}\mathrm{d}u\right)$$

When $\vec{t}_{12} = s_{11}(0,0,-1)$ then $A^{21'} = s_{12}\vec{e}_{12}$ where $s_{12} \in \mathbb{R}$ so Assumption 3i is satisfied. Using Equation (1.17) the CF of U_{12} is

$$\phi_{U_{12}}(s_{12}) = \exp\left(\int_0^{s_{12}} \frac{iE\left[Y_2 \exp\left(-iuY_3\right)\right]}{E\left[\exp\left(iu-Y_3\right)\right]}\mathrm{d}u\right)$$

In Example 1B, it was possible to move identified unobserved variables to the left hand side of the equation because of the mutual independence assumption. In Example 1C, this is not possible because the joint distribution of $\vec{U}_1$ is not identified (only the marginal distributions U_{11} and U_{12} are identified). Identification comes from first manipulating the system from Equation (1.3) as follows

$$\begin{pmatrix} Y_1 \\ Y_1 + Y_2 + Y_3 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \end{pmatrix}$$

or $\vec{\widetilde{Y}} = \widetilde{A}\vec{U}$.

Set $p^* = 2$ and $m^* = 2$. Then

$$\widetilde{A}^{22} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

When $\vec{t}_{21} = s_{21}(1,0,0)$ then $\widetilde{A}^{22'} = s_{21}\vec{\vec{e}}_{21}$ where $s_{21} \in \mathbb{R}$ so Assumption 3i is satisfied. Using Equation (1.17) the CF of U_{21} is

$$\phi_{U_{21}}(s_{21}) = \exp\left(\int_0^{s_{21}} \frac{iE\left[(Y_1 + Y_2 + Y_3)\exp\left(iuY_1\right)\right]}{E\left[\exp\left(iuY_1\right)\right]} \mathrm{d}u\right)$$

When $\vec{t}_{22} = s_{22}(-1,1,-1)$ then $\widetilde{A}^{22'} = s_{22}\vec{\vec{e}}_{22}$ where $s_{22} \in \mathbb{R}$ so Assumption 3i is satisfied. Using Equation (1.17) the CF of U_{22} is

$$\phi_{U_{22}}(s_{22}) = \exp\left(\int_0^{s_{22}} \frac{iE\left[(Y_1 + Y_2 + Y_3)\exp\left(iu(-Y_1 + Y_2 - Y_3)\right)\right]}{E\left[\exp\left(iu(-Y_1 + Y_2 - Y_3)\right)\right]} \mathrm{d}u\right)$$

When $\vec{t}_{23} = s_{23}(0,0,1)$ then $\widetilde{A}^{22'} = s_{23}\vec{\vec{e}}_{23}$ where $s_{23} \in \mathbb{R}$ so Assumption 3i is satisfied. Using Equation (1.17) the CF of U_{23} is

$$\phi_{U_{23}}(s_{23}) = \exp\left(\int_0^{s_{23}} \frac{iE\left[(Y_1 + Y_2 + Y_3)\exp\left(iuY_3\right)\right]}{E\left[\exp\left(iuY_3\right)\right]} \mathrm{d}u\right)$$

1.9.5 Example 2: Difference-in-Differences Model

As a preliminary step I identify U_1, U_2, $U_4 + U_8$ and $U_5 + U_9$. With one additional assumption that is defined later, the distribution of (Y^*_{CT}, Y^*_{TC}) is identified.

To identify U_1 set $p^* = 1$ and $\vec{t_1} = (0,0,1,0)$. Then

$$A^1 = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and $A^{1\prime}\vec{t_1} = s_1\vec{e_1}$ so Assumption 2 is satisfied for identification of U_1. Using Equation (1.13), the CF of m_C is

$$\phi_{m_C}(s_1|\vec{X} = \vec{x}) = \exp\left(\int_0^{s_1} \frac{iE\left[Y_{C0}\exp\left(iuY_{CC}\right)|\vec{X} = \vec{x}\right]}{E\left[\exp\left(iuY_{CC}\right)|\vec{X} = \vec{x}\right]}\mathrm{d}u\right)$$

Similarly, U_2 is identified by setting $p^* = 2$ and $\vec{t_2} = (0,0,0,1)$. Then

$$A^1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$A^{2\prime}\vec{t_2} = s_2\vec{e_2}$ so Assumption 2 is satisfied for identification of U_2. Using Equation (1.13) the CF of m_T is

$$\phi_{m_T}(s_2|\vec{X} = \vec{x}) = \exp\left(\int_0^{s_2} \frac{iE\left[Y_{T0}\exp\left(iuY_{TT}\right)|\vec{X} = \vec{x}\right]}{E\left[\exp\left(iuY_{TT}\right)|\vec{X} = \vec{x}\right]}\mathrm{d}u\right)$$

Next, identify $(U_4 + U_8, U_5 + U_9)$ by

$$\begin{aligned}\phi_{Y_{CC},Y_{TT}}(s_4, s_5|\vec{X} = \vec{x}) &= \phi_{m_C+h_C+\varepsilon_{CC},\ m_T+h_T+\varepsilon_{TT}}(s_4, s_5|\vec{X} = \vec{x}) \\ &= \phi_{m_C}(s_4|\vec{X} = \vec{x}) \cdot \phi_{m_T}(s_5|\vec{X} = \vec{x}) \cdot \phi_{h_C+\varepsilon_{CC},\ h_T+\varepsilon_{TT}}(s_4, s_5|\vec{X} = \vec{x})\end{aligned}$$

where the second equality follows from the independence assumptions. I already identified m_C and m_T so by rearranging the above equation $h_C + \varepsilon_{CC},\ h_T + \varepsilon_{TT}$ is identified by

$$\phi_{h_C+\varepsilon_{CC},\ h_T+\varepsilon_{TT}}(s_4, s_5|\vec{X} = \vec{x}) = \frac{\phi_{Y_{CC},Y_{TT}}(s_4, s_5|\vec{X} = \vec{x})}{\phi_{m_C}(s_4|\vec{X} = \vec{x}) \cdot \phi_{m_T}(s_5|\vec{X} = \vec{x})}$$

Finally, the distribution of (Y^*_{CT}, Y^*_{TC}) is identified with one of two possible assumptions

i. Assume $(\varepsilon_{CT}, \varepsilon_{TC})$ has the same distribution as $(\varepsilon_{TT}, \varepsilon_{CC})$, then

$$\begin{aligned}
&\phi_{Y^*_{CT},Y^*_{TC}}\left(s_4, s_5|\vec{X}=\vec{x}\right)\\
&=\phi_{m_C+h_T+\varepsilon_{CT},\ m_T+h_C+\varepsilon_{TC}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{m_C}(s_4|\vec{X}=\vec{x})\cdot\phi_{m_T}(s_5|\vec{X}=\vec{x})\cdot\phi_{h_T}(s_4|\vec{X}=\vec{x})\cdot\phi_{h_C}(s_5|\vec{X}=\vec{x})\cdot\phi_{\varepsilon_{CT},\ \varepsilon_{TC}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{m_C}(s_4|\vec{X}=\vec{x})\cdot\phi_{m_T}(s_5|\vec{X}=\vec{x})\cdot\phi_{h_T}(s_4|\vec{X}=\vec{x})\cdot\phi_{h_C}(s_5|\vec{X}=\vec{x})\cdot\phi_{\varepsilon_{TT},\ \varepsilon_{CC}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{m_C}(s_4|\vec{X}=\vec{x})\cdot\phi_{m_T}(s_5|\vec{X}=\vec{x})\cdot\phi_{h_T+\varepsilon_{TT},\ h_C+\varepsilon_{CC}}(s_4, s_5|\vec{X}=\vec{x})
\end{aligned}$$

where the second and fourth equalities follow by independence, and the third equality follows from the assumption that $(\varepsilon_{CT}, \varepsilon_{TC})$ and $(\varepsilon_{TT}, \varepsilon_{CC})$ are equally distributed. We already identified m_C, m_T and $(h_C+\varepsilon_{CC}, h_T+\varepsilon_{TT})$ so (Y^*_{CT}, Y^*_{TC}) is also identified.

ii. Assume $(\varepsilon_{CT}, \varepsilon_{TC})$ has the same distribution as $(\varepsilon_{TT}-\varepsilon_{T0}+\varepsilon_{C0},\ \varepsilon_{CC}-\varepsilon_{C0}+\varepsilon_{T0})$, then

$$\begin{aligned}
&\phi_{Y^*_{CT},Y^*_{TC}}\left(s_4, s_5|\vec{X}=\vec{x}\right)\\
&=\phi_{m_C+h_T+\varepsilon_{CT},\ m_T+h_C+\varepsilon_{TC}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{m_C+h_T}(s_4|\vec{X}=\vec{x})\cdot\phi_{m_T+h_C}(s_5|\vec{X}=\vec{x})\cdot\phi_{\varepsilon_{CT},\ \varepsilon_{TC}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{m_C+h_T}(s_4|\vec{X}=\vec{x})\cdot\phi_{m_T+h_C}(s_5|\vec{X}=\vec{x})\cdot\phi_{\varepsilon_{TT}-\varepsilon_{T0}+\varepsilon_{C0},\ \varepsilon_{CC}-\varepsilon_{C0}+\varepsilon_{T0}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{m_C+h_T+\varepsilon_{TT}-\varepsilon_{T0}+\varepsilon_{C0},\ m_T+h_C+\varepsilon_{CC}-\varepsilon_{C0}+\varepsilon_{T0}}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{(m_T+h_T+\varepsilon_{TT})-(m_T+h_0+\varepsilon_{T0})+(m_C+h_0+\varepsilon_{C0}),\ (m_C+h_C+\varepsilon_{CC})-(m_C+h_0+\varepsilon_{C0})+(m_T+h_0+\varepsilon_{T0})}(s_4, s_5|\vec{X}=\vec{x})\\
&=\phi_{Y_{TT}-Y_{T0}+Y_{C0},\ Y_{CC}-Y_{C0}+Y_{T0}}(s_4, s_5|\vec{X}=\vec{x})
\end{aligned}$$

where the second and fourth equalities follow by independence, and the third equality follows from the assumption that $(\varepsilon_{TT}-\varepsilon_{T0}+\varepsilon_{C0},\ \varepsilon_{CC}-\varepsilon_{C0}+\varepsilon_{T0})$. The distribution of $(Y_{TT}-Y_{T0}+Y_{C0},\ Y_{CC}-Y_{C0}+Y_{T0})$ is observed so (Y^*_{CT}, Y^*_{TC}) is identified.

1.9.6 Example 3: Measurement Error Model with Three Measurements

$$A \bar{\odot} A = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The second-order partial derivatives are

$$\begin{aligned}
\frac{\partial^2 \varphi_{X^*,\varepsilon_1}(s_0,s_1)}{\partial \omega_1^2} &= \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_2 \partial t_3}\right|_{(s_0-s_1,s_1,0,0)} \\
\frac{\partial^2 \varphi_{X^*,\varepsilon_1}(s_0,s_1)}{\partial \omega_1 \omega_2} &= \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1 \partial t_2}\right|_{(s_0-s_1,s_1,0,0)} - \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_2 \partial t_3}\right|_{(s_0-s_1,s_1,0,0)} \\
\frac{\partial^2 \varphi_{X^*,\varepsilon_1}(s_0,s_1)}{\partial \omega_1 \omega_2} &= \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1^2}\right|_{(s_0-s_1,s_1,0,0)} - 2\left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1 \partial t_2}\right|_{(s_0-s_1,s_1,0,0)} + \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_2 \partial t_3}\right|_{(s_0-s_1,s_1,0,0)} \\
\varphi''_{\varepsilon_2}(s_2) &= \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1^2}\right|_{(0,0,s_2,0)} - \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_2 \partial t_3}\right|_{(0,0,s_2,0)} \\
\varphi''_{\varepsilon_3}(s_3) &= \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1^3}\right|_{(0,0,0,s_3)} - \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_2 \partial t_3}\right|_{(0,0,0,s_3)}
\end{aligned}$$

Using these relationships and Equation (1.15), the CFs are

$$\begin{aligned}
\phi_{X^*,\varepsilon_1}(s_0,s_1) = \exp\Bigg(&\int_0^{s_1}\int_0^{v} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(u,s_2)}{\partial \omega_1^2} \mathrm{d}u\mathrm{d}v + \int_0^{s_2}\int_0^{s_1} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(u,v)}{\partial \omega_1 \omega_2} \mathrm{d}u\mathrm{d}v \\
&+ \int_0^{s_1}\int_0^{s_2} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(0,u)}{\partial \omega_1 \partial \omega_2} \mathrm{d}u\mathrm{d}v + \int_0^{s_2}\int_0^{v} \frac{\partial^2 \varphi_{X^*,\varepsilon_1}(0,u)}{\partial \omega_2^2} \mathrm{d}u\mathrm{d}v\Bigg) \\
\phi_{\varepsilon_2}(s_2) = \exp\Bigg(&\int_0^{s_2}\int_0^{v} \varphi''_{\varepsilon_2}(v)\,\mathrm{d}u\mathrm{d}v\Bigg) \\
\phi_{\varepsilon_3}(s_3) = \exp\Bigg(&\int_0^{s_3}\int_0^{v} \varphi''_{\varepsilon_3}(v)\,\mathrm{d}u\mathrm{d}v\Bigg)
\end{aligned}$$

1.10 Appendix B

1.10.1 Proof of Theorem 1

Let $\phi_{Y_1,\ldots,Y_P}$ denote the CF of $\vec{Y}$ and $\phi_{\vec{U}_m}$ denote the CF of $\vec{U}_m$ for $1 \le m \le M$. Then,

$$\begin{aligned}
\phi_{Y_1,\ldots,Y_P}(t_1,\ldots,t_P) &= E\left[\exp\left(iY_1t_1+\ldots+iY_Pt_P\right)\right] \\
&= E\left[\exp\left(i(a_{11}^1U_{11}+\ldots+a_{1K_M}^MU_{MK_M})t_1+\ldots+i(a_{P1}^1U_{11}+\ldots+a_{PK_M}^MU_{MK_M})t_P\right)\right] \\
&= E\left[\exp\left(i(a_{11}^1t_1+\ldots+a_{P1}^1t_P)U_{11}+\ldots+i(a_{1K_M}^Mt_1+\ldots+a_{PK_M}^Mt_P)U_{MK_M}\right)\right] \\
&= \prod_{m=1}^{M} E\left[\exp\left(iU_{m1}\sum_{p=1}^{P}a_{p1}^mt_p+\ldots+iU_{mK_m}\sum_{p=1}^{P}a_{pK_m}^mt_p\right)\right]
\end{aligned}$$

where the second equality follows by substituting $Y_p = a_{p1}^1U_{11}+\ldots+a_{pK_M}^MU_{MK_M}$ and the fourth equality follows from the independence assumptions.

Let $\varphi_{\vec{Y}}(\vec{t}) = \varphi_{Y_1,\ldots,Y_P}(t_1,\ldots,t_P) = \ln\phi_{\vec{Y}}(\vec{t})$ and

$$\varphi_m\left(\vec{\omega}_m\right) = \varphi_{U_{m1},\ldots,U_{mK_m}}\left(\omega_{m1},\ldots,\omega_{mK_m}\right) = \ln E\left[\exp\left(iU_{m1}\omega_{m1}+\ldots+iU_{mK_m}\omega_{mK_m}\right)\right]$$

then

$$\varphi_{\vec{Y}}(\vec{t}) = \sum_{m=1}^{M}\varphi_m\left(\sum_{p=1}^{P}a_{p1}^mt_p,\ldots,\sum_{p=1}^{P}a_{pK_m}^mt_p\right) = \sum_{m=1}^{M}\varphi_m\left(A_1^{m\prime}\vec{t},\ldots,A_{K_m}^{m\prime}\vec{t}\right) = \sum_{m=1}^{M}\varphi_m\left(\left(A_m'\vec{t}\right)'\right)$$

where $A = (A_1,\ldots,A_M)$ partitions A. The partial derivative with respect to t_p is

$$\frac{\partial\varphi_{\vec{Y}}(\vec{t})}{\partial t_p} = \sum_{m=1}^{M}\sum_{k=1}^{K_m}a_{pk}^m\left.\frac{\partial\varphi_m\left(\vec{\omega}_m\right)}{\partial\omega_{mk}}\right|_{\left(A_m'\vec{t}\right)'}$$

In matrix notation the first-order partial derivatives are

$$\begin{pmatrix}\dfrac{\partial\varphi_{\vec{Y}}(\vec{t})}{\partial t_1}\\ \vdots \\ \dfrac{\partial\varphi_{\vec{Y}}(\vec{t})}{\partial t_P}\end{pmatrix} = \sum_{m=1}^{M}\begin{pmatrix}a_{11}^m & \ldots & a_{1K_m}^m\\ \vdots & \ddots & \vdots\\ a_{P1}^m & \ldots & a_{PK_m}^m\end{pmatrix}\begin{pmatrix}\left.\dfrac{\partial\varphi_m\left(\vec{\omega}_m\right)}{\partial\omega_{m1}}\right|_{\left(A_m'\vec{t}\right)'}\\ \vdots \\ \left.\dfrac{\partial\varphi_m\left(\vec{\omega}_m\right)}{\partial\omega_{mK_m}}\right|_{\left(A_m'\vec{t}\right)'}\end{pmatrix}$$

The new system of equations is identical to Equation (1.12) except the unobserved random variable U_{mk} is replaced by the first-order partial derivative $\left.\dfrac{\partial\varphi_m\left(\vec{\omega}_m\right)}{\partial\omega_{mk}}\right|_{\left(A_m'\vec{t}\right)'}$.

The first-order partial derivative with respect to $t_{p_{k^*}}$ is

$$\begin{aligned}\frac{\partial \varphi_{\vec{Y}}(\vec{t})}{\partial t_{p_{k^*}}} &= \sum_{m=1}^{M}\sum_{k=1}^{K_m} a^m_{p_{k^*}k} \left.\frac{\partial \varphi_m(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{(A'_m\vec{t})'} \\ &= \sum_{m=1}^{M}\sum_{k=1}^{K_m} a^m_{p_{k^*}k} \left.\frac{\partial \varphi_m(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{\left(\mathbf{I}\left(\bigcup_k a^m_{p_{k^*}k}\neq 0\right)(A'_m\vec{t})'\right)} \\ &= \sum_{m=1}^{M}\sum_{k=1}^{K_m} a^m_{p_{k^*}k} \left.\frac{\partial \varphi_m(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{\left(A^{p_{k^*}\prime}_m\vec{t}\right)'}\end{aligned}$$

where $A^{p_{k^*}} = (A^{p_{k^*}}_1, \dots, A^{p_{k^*}}_M)$ partitions $A^{p_{k^*}}$.

By Assumption 1i, there exists $\vec{t}_{m^*}$ such that $A^{p_{k^*}\prime}_m\vec{t}_{m^*} = \vec{0}_{K_m}$ for all $m \neq m^*$ and $A^{p_{k^*}\prime}_{m^*}\vec{t}_{m^*} = \vec{s}_{m^*} \in \mathbb{R}^{K^*_m}$. One solution is $\vec{t}_{m^*} = (A^{p_{k^*}})^+\left(\vec{0}'_{\sum_{m<m^*}K_m}\ ,\ \vec{s}'_{m^*}\ ,\ \vec{0}'_{\sum_{m>m^*}K_m}\right)'$. To save on notation I denote this solution as $\vec{t}_{m^*} = (A^{p_{k^*}})^+(\vec{0}', \vec{s}'_{m^*}, \vec{0}')'$. Then

$$\begin{aligned}\left.\frac{\partial \varphi_{\vec{Y}}(\vec{t})}{\partial t_{p_{k^*}}}\right|_{(A^{p_{k^*}})^+(\vec{0}',\vec{s}'_{m^*},\vec{0}')'} &= \sum_{k=1}^{K_{m^*}} a^{m^*}_{p_{k^*}k} \left.\frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k}}\right|_{\vec{s}_{m^*}} + \sum_{m\neq m^*}\sum_{k=1}^{K_m} \left.\frac{\partial \varphi_m(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{\vec{0}'_{K_m}} \\ &= a^{m^*}_{p_{k^*}k^*} \left.\frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k^*}}\right|_{\vec{s}_{m^*}} + \sum_{m\neq m^*}\sum_{k=1}^{K_m} a^m_{p_{k^*}k}E[U_{mk}] \\ &= a^{m^*}_{p_{k^*}k^*} \left.\frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k^*}}\right|_{\vec{s}_{m^*}} \end{aligned} \tag{1.20}$$

where the second equality follows from Assumption 1ii that $a^{m^*}_{p_{k^*}k} = 0$ for all $k \neq k^*$ and the last equality because $E[U_{mk}] = 0$.

The CF of U_{m^*} is expressed in terms of its first-order partial derivatives

$$\begin{aligned}\phi_{m^*}(\vec{s}_{m^*}) &= \exp\left(\sum_{k=1}^{K_{m^*}} \int_0^{s_k} \left.\frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k}}\right|_{(s_1,\dots,s_{k-1},u_k,0,\dots,0)} du_k\right) \\ &= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{a^{m^*}_{p_k k}} \int_0^{s_k} \left.\frac{\partial \varphi_{\vec{Y}}(\vec{t})}{\partial t_{p_k}}\right|_{(A^{p_{k^*}})^+(\vec{0}',s_1,\dots,s_{k-1},u_k,0,\dots,0,\vec{0}')'} du_k\right) \\ &= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{a^{m^*}_{p_k k}} \int_0^{s_k} \left.\frac{\partial \ln E\left[\exp\left(i\vec{Y}'\vec{t}\right)\right]}{\partial t_{p_{k^*}}}\right|_{(A^{p_{k^*}})^+(\vec{0}',s_1,\dots,s_{k-1},u_k,0,\dots,0,\vec{0}')'} du_k\right) \\ &= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{a^{m^*}_{p_k k}} \int_0^{s_k} \frac{iE\left[Y_{p_{k^*}}\exp\left(i\vec{Y}'(A^{p_{k^*}})^+(\vec{0}',s_1,\dots,s_{k-1},u_k,0,\dots,0,\vec{0}')'\right)\right]}{E\left[\exp\left(i\vec{Y}'(A^{p_{k^*}})^+(\vec{0}',s_1,\dots,s_{k-1},u_k,0,\dots,0,\vec{0}')'\right)\right]} du_k\right)\end{aligned}$$

where the first equality uses the Fundamental Theorem of Calculus and the second equality follows by substituting Equation (1.20).

The CF of $\vec{U}_{m^*}$ is defined by bounding:

$$\begin{aligned}
&\left|\int_0^{s_k} \frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k}}\bigg|_{(s_1,\ldots,s_{k-1},u_k,0,\ldots,0)} \mathrm{d}u_k\right| \\
&= \left|\int_0^{s_k} \frac{iE\left[U_{m^*k}\exp i\left(U_{m^*1}s_1+\ldots+U_{m^*k-1}s_{k-1}+U_{m^*k}u_k\right)\right]}{E\left[\exp i\left(U_{m^*1}s_1+\ldots+U_{m^*k-1}s_{k-1}+U_{m^*k}u_k\right)\right]}\mathrm{d}u_k\right| \\
&\leq E\left[|U_{m^*k}|\right]\int_0^{s_k} \frac{1}{\left|E\left[\exp i\left(U_{m^*1}s_1+\ldots+U_{m^*k-1}s_{k-1}+U_{m^*k}u_k\right)\right]\right|}\mathrm{d}u_k \\
&< \infty
\end{aligned}$$

where the first inequality follows from the triangle inequality and $|\exp(\cdot)| \leq 1$ and the second inequality follows from the assumptions $\int\limits_0^{s_k}\left|\left(E\left[\exp i\left(U_{m^*1}s_1+\ldots+U_{m^*k-1}s_{k-1}+U_{m^*k}u_k\right)\right]\right)^{-1}\right|\mathrm{d}u_k < \infty$ and $E\left[|U_{m^*k}|\right] < \infty$ for $k = 1, \ldots, K_{m^*}$.

This shows that the CF of $\vec{U}_{m^*}$ is identified. The joint density of $\vec{U}_{m^*}$ is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_{m^*}(\vec{u}_{m^*}) = \frac{1}{2\pi}\int e^{-i\vec{s}_{m^*}'\vec{u}_{m^*}}\phi_{m^*}(\vec{s}_{m^*})\mathrm{d}\vec{s}_{m^*}$$

1.10.2 Proof of Theorem 2

The CF of $\vec{Y}$ is

$$\begin{aligned}
\phi_{Y_1,\ldots,Y_P}(t_1,\ldots,t_P) &= E\left[\exp\left(iY_1t_1+\ldots+iY_Pt_P\right)\right] \\
&= E\left[\exp\left(i(a_{11}^1U_{11}+\ldots+a_{1K_M}^MU_{MK_M})t_1+\ldots+i(a_{P1}^1U_{11}+\ldots+a_{PK_M}^MU_{MK_M})t_P\right)\right] \\
&= E\left[\exp\left(i(a_{11}^1t_1+\ldots+a_{P1}^1t_P)U_{11}+\ldots+i(a_{1K_M}^Mt_1+\ldots+a_{PK_M}^Mt_P)U_{MK_M}\right)\right] \\
&= \prod_{m=1}^{M} E\left[\exp\left(iU_{m1}\sum_{p=1}^{P}a_{p1}^mt_p+\ldots+iU_{mK_m}\sum_{p=1}^{P}a_{pK_m}^mt_p\right)\right]
\end{aligned}$$

where the second equality follows by substituting $Y_p = a_{p1}^1U_{11}+\ldots+a_{pK_M}^MU_{MK_M}$ and the fourth equality follows from the independence assumptions.

Let $\varphi_{\vec{Y}}(\vec{t}) = \varphi_{Y_1,\ldots,Y_P}(t_1,\ldots,t_P) = \ln\phi_{\vec{Y}}(\vec{t})$ and

$$\varphi_m\left(\vec{\omega}_m\right) = \varphi_{U_{m1},\ldots,U_{mK_m}}\left(\omega_{m1},\ldots,\omega_{mK_m}\right) = \ln E\left[\exp\left(iU_{m1}\omega_{m1}+\ldots+iU_{mK_m}\omega_{mK_m}\right)\right]$$

then

$$\varphi_{\vec{Y}}(\vec{t}) = \sum_{m=1}^{M} \varphi_m \left(\sum_{p=1}^{P} a_{p1}^m t_p, \ldots, \sum_{p=1}^{P} a_{pK_m}^m t_p \right) = \sum_{m=1}^{M} \varphi_m \left(A_1^{m\prime} \vec{t}, \ldots, A_{K_m}^{m\prime} \vec{t} \right) = \sum_{m=1}^{M} \varphi_m \left(\left(A_m' \vec{t} \right)' \right)$$

where $A = (A_1, \ldots, A_M)$ partitions A.

Necessity: Assume Assumption 2i does not hold. Let $\tilde{\vec{U}}_1, \ldots, \tilde{\vec{U}}_M$ and $\vec{U}_1, \ldots, \vec{U}_M$ be observationally equivalent. Then

$$\varphi_{\vec{Y}}(\vec{t}) = \sum_{m=1}^{M} \varphi_m \left(\left(A_m' \vec{t} \right)' \right) = \sum_{m=1}^{M} \widetilde{\varphi}_m \left(\left(A_m' \vec{t} \right)' \right)$$

where φ_m is the log CF of $\vec{U}_m$ and $\widetilde{\varphi}_m$ is the log CF of $\tilde{\vec{U}}_m$ for $m = 1, \ldots, M$. Then

$$\sum_{m=1}^{M} \varphi_m \left(\left(A_m' \vec{t} \right)' \right) - \sum_{m=1}^{M} \widetilde{\varphi}_m \left(\left(A_m' \vec{t} \right)' \right) = 0$$

The partial derivative with respect to t_p is

$$\sum_{m=1}^{M} \sum_{k=1}^{K_m} a_{pk}^m \left(\left. \frac{\partial \varphi_m \left(\vec{\omega}_m \right)}{\partial \omega_{mk}} \right|_{\left(A_m' \vec{t} \right)'} - \left. \frac{\partial \widetilde{\varphi}_m \left(\vec{\omega}_m \right)}{\partial \omega_{mk}} \right|_{\left(A_m' \vec{t} \right)'} \right) = 0$$

In matrix notation the first-order partial derivatives are

$$\sum_{m=1}^{M} \begin{pmatrix} a_{11}^m & \ldots & a_{1K_m}^m \\ \vdots & \ddots & \vdots \\ a_{P1}^m & \ldots & a_{PK_m}^m \end{pmatrix} \begin{pmatrix} \frac{\partial \varphi_m}{\partial \omega_{m1}} - \frac{\partial \widetilde{\varphi}_m}{\partial \omega_{m1}} \\ \vdots \\ \frac{\partial \varphi_m}{\partial \omega_{mK_m}} - \frac{\partial \widetilde{\varphi}_m}{\partial \omega_{mK_m}} \end{pmatrix} = \left(A_1 \ \cdots \ A_M \right) \begin{pmatrix} \frac{\partial \varphi_1}{\partial \omega_{11}} - \frac{\partial \widetilde{\varphi}_1}{\partial \omega_{11}} \\ \vdots \\ \frac{\partial \varphi_m}{\partial \omega_{mk}} - \frac{\partial \widetilde{\varphi}_m}{\partial \omega_{mk}} \\ \vdots \\ \frac{\partial \varphi_M}{\partial \omega_{MK_M}} - \frac{\partial \widetilde{\varphi}_M}{\partial \omega_{MK_M}} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

where for clarity of notation the arguments of the CFs are omitted. The second-order partial derivative with respect to t_{p_1} and t_{p_2} is

$$\begin{aligned} \frac{\partial \varphi_{\vec{Y}}^2(t)}{\partial t_{p_1} t_{p_2}} &= \sum_{m=1}^{M} \sum_{k_1=1}^{K_m} a_{p_1 k_1}^m \sum_{k_2=1}^{K_m} a_{p_2 k_2}^m \left(\frac{\partial^2 \varphi_m}{\partial \omega_{mk_1} \partial \omega_{mk_2}} - \frac{\partial^2 \widetilde{\varphi}_m}{\partial \omega_{mk_1} \partial \omega_{mk_2}} \right) \\ &= \sum_{m=1}^{M} \sum_{k=1}^{K_m} a_{p_1 k}^m a_{p_2 k}^m \left(\frac{\partial^2 \varphi_m}{\partial \omega_{mk}^2} - \frac{\partial^2 \widetilde{\varphi}_m}{\partial \omega_{mk}^2} \right) + \sum_{m=1}^{M} \sum_{k_1 \neq k_2}^{K_m} a_{p_1 k_1}^m a_{p_2 k_2}^m \left(\frac{\partial^2 \varphi_m}{\partial \omega_{mk_1} \partial \omega_{mk_2}} - \frac{\partial^2 \widetilde{\varphi}_m}{\partial \omega_{mk_1} \partial \omega_{mk_2}} \right) \end{aligned}$$

$$= \sum_{m=1}^{M}\sum_{k=1}^{K_m} a^m_{p_1k}a^m_{p_2k}\left(\frac{\partial^2\varphi_m}{\partial\omega^2_{mk}} - \frac{\partial^2\widetilde{\varphi}_m}{\partial\omega^2_{mk}}\right)$$
$$+ \sum_{m=1}^{M}\sum_{k_1<k_2}^{K_m}\left(a^m_{p_1k_1}a^m_{p_2k_2} + a^m_{p_1k_2}a^m_{p_2k_1}\right)\left(\frac{\partial^2\varphi_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}} - \frac{\partial^2\widetilde{\varphi}_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}}\right)$$

where the third equality follows because $\frac{\partial^2\varphi_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}} = \frac{\partial^2\varphi_m}{\partial\omega_{mk_2}\partial\omega_{mk_1}}$. In matrix notation the second-order partial derivatives are

$$(A \odot A)\begin{pmatrix} \frac{\partial^2\varphi_1}{\partial\omega^2_{11}} - \frac{\partial^2\widetilde{\varphi}_1}{\partial\omega^2_{11}} \\ \vdots \\ \frac{\partial^2\varphi_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}} - \frac{\partial^2\widetilde{\varphi}_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}} \\ \vdots \\ \frac{\partial^2\varphi_M}{\partial\omega^2_{MK_M}} - \frac{\partial^2\widetilde{\varphi}_M}{\partial\omega^2_{MK_M}} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{1.21}$$

where $k_1 \leq k_2$.

The matrix $(A \odot A)$ is of dimension $P^2 \times \sum_{m=1}^{M} K_m(K_m+1)/2$. If Assumption 2i does not hold then $\text{Rank}(A \odot A) < \sum_{m=1}^{M} K_m(K_m+1)/2$ and there are nonzero solutions to Equation (1.21). Say one such solution is $\frac{\partial^2\varphi_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}} - \frac{\partial^2\widetilde{\varphi}_m}{\partial\omega_{mk_1}\partial\omega_{mk_2}} = c_{mk_1k_2}$ then $\varphi_m\left(\sum_{p=1}^{P} a^m_{p1}t_p, \ldots, \sum_{p=1}^{P} a^m_{pK_m}t_p\right)$ and $\widetilde{\varphi}_m\left(\sum_{p=1}^{P} a^m_{p1}t_p, \ldots, \sum_{p=1}^{P} a^m_{pK_m}t_p\right) = \varphi_m\left(\sum_{p=1}^{P} a^m_{p1}t_p, \ldots, \sum_{p=1}^{P} a^m_{pK_m}t_p\right) - \sum_{k_1,k_2} c_{mk_1k_2}t_{k_1}t_{k_2}$ are observationally equivalent. This implies that $\phi_m(t)$ is observationally equivalent to $\widetilde{\phi}_m(\vec{t}) = \phi_m(\vec{t})\exp(\widetilde{c}_m + \sum_k \widetilde{c}_{mk}t_k + \sum_{k_1,k_2}\widetilde{c}_{mk_1k_2}t_{k_1}t_{k_2})$ (a shift by some polynomial of degree two) and hence that $\vec{\widetilde{U}}_1, \ldots \vec{\widetilde{U}}_M$ and $\vec{U}_1, \ldots \vec{U}_M$ are observationally equivalent.

The matrix A_m is of dimension $P \times K_m$. If Assumption 2ii does not hold then $\text{Rank}(A_m) < K_m$ for some m. Without loss of generality let Assumption 2ii not hold when $m = m^*$, then there exists a nonzero $\delta \in \mathbb{R}^{K_{m^*}}$ that satisfies

$$A_{m^*}\delta = \begin{pmatrix} \sum_{k=1}^{K_{m^*}} a^{m^*}_{1k}\delta_k \\ \vdots \\ \sum_{k=1}^{K_{m^*}} a^{m^*}_{Pk}\delta_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Let $\left(\widetilde{U}_{m^*1}, \ldots \widetilde{U}_{m^*K_{m^*}}\right) := (U_{m^*1} + \delta_1 U_{m^*1}, \ldots, U_{m^*K_{m^*}} + \delta_{K_{m^*}}U_{m^*1})$. The CF of $\vec{Y}$ is

$$\phi_{\vec{Y}}(\vec{t}) = E\left[\exp\left(i(a^1_{11}t_1 + \ldots + a^1_{P1}t_P)U_{11} + \ldots + i(a^M_{1K_M}t_1 + \ldots + a^M_{PK_M}t_P)U_{MK_M}\right)\right]$$

$$
\begin{aligned}
&= \prod_{m=1}^{M} E\left[\exp\left(i\sum_{k=1}^{K_m} U_{mk}\sum_{p=1}^{P} a_{pk}^{m} t_p\right)\right] \\
&= \prod_{m=1}^{M} \phi_m\left(\sum_{p=1}^{P} a_{pk}^{m} t_p\right) \\
&= E\left[\exp\left(i\sum_{k=1}^{K_{m^*}} U_{m^*k}\sum_{p=1}^{P} a_{pk}^{m^*} t_p\right)\right]\prod_{m\neq m^*} E\left[\exp\left(i\sum_{k=1}^{K_m} U_{mk}\sum_{p=1}^{P} a_{pk}^{m} t_p\right)\right] \\
&= E\left[\exp\left(i\sum_{k=1}^{K_{m^*}} U_{m^*k}\sum_{p=1}^{P} a_{pk}^{m^*} t_p + U_{m^*1}\sum_{p=1}^{P} t_p\sum_{k=1}^{K_{m^*}} a_{pk}^{m^*}\delta_k\right)\right]\prod_{m\neq m^*} E\left[\exp\left(i\sum_{k=1}^{K_m} U_{mk}\sum_{p=1}^{P} a_{pk}^{m} t_p\right)\right] \\
&= E\left[\exp\left(i\sum_{k=1}^{K_{m^*}} \left(U_{m^*k} + U_{m^*1}\delta_k\right)\sum_{p=1}^{P} a_{pk}^{m^*} t_p\right)\right]\prod_{m\neq m^*} E\left[\exp\left(i\sum_{k=1}^{K_m} U_{mk}\sum_{p=1}^{P} a_{pk}^{m} t_p\right)\right] \\
&= E\left[\exp\left(i\sum_{k=1}^{K_{m^*}} \widetilde{U}_{m^*k}\sum_{p=1}^{P} a_{pk}^{m^*} t_p\right)\right]\prod_{m\neq m^*} E\left[\exp\left(i\sum_{k=1}^{K_m} U_{mk}\sum_{p=1}^{P} a_{pk}^{m} t_p\right)\right] \\
&= \prod_{m=1}^{M} \widetilde{\phi}_m\left(\sum_{p=1}^{P} a_{pk}^{m} t_p\right)
\end{aligned}
$$

where the fifth equality follows because $\sum_{k=1}^{K_{m^*}} a_{pk}^{m^*}\delta_k = 0$ for all p and the second to last equality holds from the definition of $\left(\widetilde{U}_{m^*1},\ldots,\widetilde{U}_{m^*K_{m^*}}\right)$. Hence the CFs of $(\vec{U}_1,\ldots,\vec{\widetilde{U}}_{m^*},\ldots,\vec{U}_M)$ and $(\vec{U}_1,\ldots,\vec{U}_{m^*},\ldots,\vec{U}_M)$ are observationally equivalent, which implies that $(\vec{U}_1,\ldots,\vec{\widetilde{U}}_{m^*},\ldots,\vec{U}_M)$ and $(\vec{U}_1,\ldots,\vec{U}_{m^*},\ldots,\vec{U}_M)$ are observationally equivalent.

Sufficiency: Assume Assumption 2 holds. The second-order partial derivatives of $\varphi_{\vec{Y}}(\vec{t})$ are

$$
\begin{pmatrix}
\frac{\partial^2\varphi_{\vec{Y}}(\vec{t})}{\partial t_1^2} \\
\vdots \\
\frac{\partial^2\varphi_{\vec{Y}}(\vec{t})}{\partial t_{p_1}\partial t_{p_2}} \\
\vdots \\
\frac{\partial^2\varphi_{\vec{Y}}(\vec{t})}{\partial t_P^2}
\end{pmatrix}
= (A\odot A)
\begin{pmatrix}
\left.\frac{\partial\varphi_1^2(\vec{\omega}_1)}{\partial\omega_{11}^2}\right|_{(A_1'\vec{t})'} \\
\vdots \\
\left.\frac{\partial\varphi_m^2(\vec{\omega}_m)}{\partial\omega_{mk_1}\partial\omega_{mk_2}}\right|_{(A_m'\vec{t})'} \\
\vdots \\
\left.\frac{\partial\varphi_M^2(\vec{\omega}_M)}{\partial\omega_{MK_M}^2}\right|_{(A_M'\vec{t})'}
\end{pmatrix}
$$

$k_1 \leq k_2$.

By Assumption 2i

$$
\left(\left.\frac{\partial\varphi_1^2(\vec{\omega}_1)}{\partial\omega_{11}^2}\right|_{(A_1'\vec{t})'},\ldots,\left.\frac{\partial\varphi_M^2(\vec{\omega}_M)}{\partial\omega_{MK_M}^2}\right|_{(A_M'\vec{t})'}\right)' = (A\odot A)^{+}\left(\frac{\partial^2\varphi_{\vec{Y}}(\vec{t})}{\partial t_1^2},\ldots,\frac{\partial^2\varphi_{\vec{Y}}(\vec{t})}{\partial t_P^2}\right)'
$$

By Assumption 2ii, for all $\vec{s}_m \in \mathbb{R}^{K_m}$ there exists a $\vec{t}_m \in \mathbb{R}^P$ that solves $A_m'\vec{t}_m = \vec{s}_m$. One solution is

$\vec{t}_m = (A'_m)^+ \vec{s}_m$. Then

$$\left(\cdots \left.\frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{m1}^2}\right|_{\vec{s}'_m}, \ldots, \left.\frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{mK_m}^2}\right|_{\vec{s}'_m} \cdots \right)' = (A \odot A)^+ \left(\left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_1^2}\right|_{(A'_m)^+ \vec{s}_m}, \ldots, \left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_P^2}\right|_{(A'_m)^+ \vec{s}_m} \right)'$$

where

$$\left.\frac{\partial^2 \varphi_{\vec{Y}}(\vec{t})}{\partial t_{p_1} \partial t_{p_2}}\right|_{(A'_m)^+ \vec{s}_m} = \frac{E\left[Y_{p_1} e^{i\vec{Y}'(A'_m)^+ \vec{s}_m}\right] E\left[Y_{p_2} e^{i\vec{Y}'(A'_m)^+ \vec{s}_m}\right]}{\left(E\left[e^{i\vec{Y}'(A'_m)^+ \vec{s}_m}\right]\right)^2} - \frac{E\left[Y_{p_1} Y_{p_2} e^{i\vec{Y}'(A'_m)^+ \vec{s}_m}\right]}{E\left[e^{i\vec{Y}'(A'_m)^+ \vec{s}_m}\right]}$$

The CF of U_m is expressed in terms of second-order partial derivatives

$$\begin{aligned} \phi_m(\vec{s}_m) = \exp \Bigg(& \sum_{k=1}^{K_m} \int_0^{s_k} \int_0^{v_k} \left.\frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{mk}^2}\right|_{(0,\ldots,u_k,0,\ldots,0)} \mathrm{d}u_k \mathrm{d}v_k \\ & + \sum_{k_1 < k_2} \int_0^{s_{k_2}} \int_0^{s_{k_1}} \left.\frac{\partial \varphi_m^2(\vec{\omega}_m)}{\partial \omega_{mk_1} \partial \omega_{mk_2}}\right|_{(s_1,\ldots,s_{k_1-1},u_{k_1},0,\ldots,0,u_{k_2},0,\ldots,0)} \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} \Bigg) \end{aligned}$$

The CF of $\vec{U}_m$ is defined by bounding:

$$\begin{aligned} & \left| \int_0^{s_{k_2}} \int_0^{s_{k_1}} \left.\frac{\partial^2 \varphi_m(\vec{\omega}_m)}{\partial \omega_{mk_1} \partial \omega_{mk_2}}\right|_{(\ldots,u_{k_1},\ldots,u_{k_2},\ldots)} \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} \right| \\ = & \left| \int_0^{s_{k_2}} \int_0^{s_{k_1}} \left(\frac{E\left[U_{mk_1} e^{i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}}\right] E\left[U_{mk_2} e^{i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}}\right]}{\left(E\left[e^{i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}}\right]\right)^2} \right.\right. \\ & \left.\left. - \frac{E\left[U_{mk_1} U_{mk_2} e^{i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}}\right]}{E\left[e^{i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}}\right]} \right) \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} \right| \\ \le & E\left[|U_{mk_1} U_{mk_2}|\right] \int_0^{s_{k_2}} \int_0^{s_{k_1}} \frac{1}{\left(E\left[\exp\left(i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}\right)\right]\right)^2} \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} \\ < & \infty \end{aligned}$$

where the first inequality follows from the triangle inequality and $|\exp(\cdot)| \le 1$ and the second inequality follows from the assumptions $\int_0^{s_{k_2}} \int_0^{s_{k_1}} \left(E\left[\exp\left(i\sum_{k=1}^{k_1-1} U_{mk} s_k + iU_{mk_1} u_{k_1} + iU_{mk_2} u_{k_2}\right)\right]\right)^{-2} \mathrm{d}u_{k_1} \mathrm{d}u_{k_2} < \infty$ and $E\left[|U_{mk_1} U_{mk_2}|\right] < \infty$ for $k_1, k_2 = 1, \ldots, K_m$.

This shows that the CF of $\vec{U}_m$ is identified. The joint density of $\vec{U}_m$ is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_m(\vec{u}_m) = \frac{1}{2\pi} \int e^{-i\vec{s}'_m \vec{u}_m} \phi_m(\vec{s}_m) \mathrm{d}\vec{s}_m$$

1.10.3 Proof of Theorem 3

The CF of $\vec{Y}$ is

$$\begin{aligned}
\phi_{Y_1,\ldots,Y_P}(t_1,\ldots,t_P) &= E\left[\exp\left(iY_1t_1+\ldots+iY_Pt_P\right)\right] \\
&= E\left[\exp\left(i(a^1_{11}U_{11}+\ldots+a^M_{1K_M}U_{MK_M})t_1+\ldots+i(a^1_{P1}U_{11}+\ldots+a^M_{PK_M}U_{MK_M})t_P\right)\right] \\
&= E\left[\exp\left(i(a^1_{11}t_1+\ldots+a^1_{P1}t_P)U_{11}+\ldots+i(a^M_{1K_M}t_1+\ldots+a^M_{PK_M}t_P)U_{MK_M}\right)\right] \\
&= \prod_{m=1}^{M} E\left[\exp\left(iU_{m1}\sum_{p=1}^{P}a^m_{p1}t_p+\ldots+iU_{mK_m}\sum_{p=1}^{P}a^m_{pK_m}t_p\right)\right]
\end{aligned}$$

where the second equality follows by substituting $Y_p = a^1_{p1}U_{11}+\ldots+a^M_{pK_M}U_{MK_M}$ and the fourth equality follows from the independence assumptions.

Let $\varphi_{\vec{Y}}(\vec{t}) = \varphi_{Y_1,\ldots,Y_P}(t_1,\ldots,t_P) = \ln\phi_{\vec{Y}}(\vec{t})$ then

$$\varphi_{\vec{Y}}(\vec{t}) = \sum_{m=1}^{M}\ln E\left[\exp\left(iU_{m1}\sum_{p=1}^{P}a^m_{p1}t_p+\ldots+iU_{mK_m}\sum_{p=1}^{P}a^m_{pK_m}t_p\right)\right]$$

The first-order partial derivative with respect to t_{p^*} is

$$\frac{\partial\varphi_{\vec{Y}}(\vec{t})}{\partial t_{p^*}} = i\sum_{m=1}^{M}\sum_{k=1}^{K_m}a^m_{p^*k}\frac{E\left[U_{mk}\exp\left(i\sum_{k=1}^{K_m}U_{mk}\sum_{p=1}^{P}a^m_{pk}t_p\right)\right]}{E\left[\exp\left(i\sum_{k=1}^{K_m}U_{mk}\sum_{p=1}^{P}a^m_{pk}t_p\right)\right]}$$

By Assumption 3i, there exists $\vec{t}_{(mk)^*}$ such that $A^{p^*m^*\prime}_{m}\vec{t}_{(mk)^*} = \vec{0}_{K_m}$ for all $m\neq m^*$ and $A^{p^*m^*\prime}_{m^*}\vec{t}_{(mk)^*} = \vec{e}_{k^*}$. This means that

$$\begin{aligned}
&A_{mk}\mathbf{I}\left(\left\{a^m_{p^*k}\neq 0\right\}\cup\{m^*=m\}\right)\vec{t}_{(mk)^*} \\
&= \mathbf{I}\left(\left\{a^m_{p^*k}\neq 0\right\}\cup\{m^*=m\}\right)\sum_{p=1}^{P}a^m_{pk}t_{(mk)^*p} \\
&= \begin{cases} 1 & \text{if } m=m^* \text{ and } k=k^* \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

One solution is $\vec{t}_{(mk)^*} = \left(A^{p^*m^*\prime}\right)^+\vec{e}_{(mk)^*}$. Let $s_{(mk)^*}\in\mathbb{R}$, then

$$\begin{aligned}
&\left.\frac{\partial\varphi_{\vec{Y}}\left(\vec{t}\right)}{\partial t_{p^*}}\right|_{\left(A^{p^*m^*\prime}\right)^+\vec{e}_{(mk)^*}s_{(mk)^*}} \\
&= i\sum_{k=1}^{K_{m^*}}a^{m^*}_{p^*k}\frac{E\left[U_{m^*k}\exp\left(iU_{(mk)^*}s_{(mk)^*}\right)\right]}{E\left[\exp\left(iU_{(mk)^*}s_{(mk)^*}\right)\right]}
\end{aligned} \tag{1.22}$$

$$+ i \sum_{m \neq m^*} \sum_{k=1}^{K_m} a_{p^*k}^{m} \frac{E\left[U_{mk} \exp\left(i \sum_{k=1}^{K_m} U_{mk} \mathbf{I}\left(\left\{a_{p^*k}^{m} = 0\right\}\right) \sum_{p=1}^{P} a_{pk}^{m} t_{(mk)^*p}\right)\right]}{E\left[\exp\left(i \sum_{k=1}^{K_m} U_{mk} \mathbf{I}\left(\left\{a_{p^*k}^{m} = 0\right\}\right) \sum_{p=1}^{P} a_{pk}^{m} t_{(mk)^*p}\right)\right]}$$

$$= i \sum_{k=1}^{K_{m^*}} a_{p^*k}^{m^*} \frac{E\left[E\left[U_{m^*k} | U_{(mk)^*}\right] \exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]}{E\left[\exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]}$$

$$+ i \sum_{m \neq m^*} \sum_{k=1}^{K_m} a_{p^*k}^{m} \frac{E\left[E\left[U_{mk} | U_{m\bar{k}}\right] \exp\left(i \sum_{k=1}^{K_m} U_{mk} \mathbf{I}\left(\left\{a_{p^*k}^{m} = 0\right\}\right) \sum_{p=1}^{P} a_{pk}^{m} t_{(mk)^*p}\right)\right]}{E\left[\exp\left(i \sum_{k=1}^{K_m} U_{mk} \mathbf{I}\left(\left\{a_{p^*k}^{m} = 0\right\}\right) \sum_{p=1}^{P} a_{pk}^{m} t_{(mk)^*p}\right)\right]}$$

$$= \frac{i a_{p^*k^*}^{m^*} E\left[U_{(mk)^*} \exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]}{E\left[\exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]} \tag{1.23}$$

where the first equality follows from 3i, the second equality follows by assuming, with out loss of generality, that $a_{p^*\bar{k}}^{m} = 0$, and the third equality by Assumption 3ii (mean independence). Let $\varphi_{(mk)^*}(s_{(mk)^*})$ be the log CF of $U_{(mk)^*}$, then

$$\begin{aligned}
a_{p^*k^*}^{m^*} \varphi'_{(mk)^*}(s_{(mk)^*}) &= a_{p^*k^*}^{m^*} \frac{\partial \ln E\left[\exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]}{\partial s_{(mk)^*}} \\
&= \frac{i a_{p^*k^*}^{m^*} E\left[U_{(mk)^*} \exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]}{E\left[\exp\left(iU_{(mk)^*} s_{(mk)^*}\right)\right]} \\
&= \left.\frac{\partial \varphi_{\vec{Y}}\left(\vec{t}\right)}{\partial t_{p^*}}\right|_{\left(A^{p^*m^*\prime}\right)^{+} \vec{e}_{(mk)^*} s_{(mk)^*}} \\
&= \frac{iE\left[Y_{p^*} \exp\left(i s_{(mk)^*} \vec{Y}' \left(A^{p^*m^*\prime}\right)^{+} \vec{e}_{(mk)^*}\right)\right]}{E\left[\exp\left(i s_{(mk)^*} \vec{Y}' \left(A^{p^*m^*\prime}\right)^{+} \vec{e}_{(mk)^*}\right)\right]}
\end{aligned} \tag{1.24}$$

where the last equality follows by substituting in Equation (1.23). By the Second Fundamental Theorem of Calculus:

$$\begin{aligned}
\phi_{(mk)^*}(s_{(mk)^*}) = \exp\left(\varphi_{(mk)^*}(s_{(mk)^*})\right) &= \exp\left(\varphi_{(mk)^*}(0) + \int_0^{s_{(mk)^*}} \varphi'_{(mk)^*}(u) \mathrm{d}u\right) \\
&= \exp\left(\frac{1}{a_{p^*k^*}^{m^*}} \int_0^{s_{(mk)^*}} \left.\frac{\partial \varphi_{\vec{Y}}\left(\vec{t}\right)}{\partial t_{p^*}}\right|_{\left(A^{p^*m^*\prime}\right)^{+} \vec{e}_{(mk)^*} u} \mathrm{d}u\right) \\
&= \exp\left(\frac{1}{a_{p^*k^*}^{m^*}} \int_0^{s_{(mk)^*}} \frac{iE\left[Y_{p^*} \exp\left(iu \vec{Y}' \left(A^{p^*m^*\prime}\right)^{+} \vec{e}_{(mk)^*}\right)\right]}{E\left[\exp\left(iu \vec{Y}' \left(A^{p^*m^*\prime}\right)^{+} \vec{e}_{(mk)^*}\right)\right]} \mathrm{d}u\right)
\end{aligned}$$

where the second equality follows by substituting in Equation (1.24) and $\varphi_{(mk)^*}(0) \ln E[\exp(0)] = 0$.

The CF of $U_{(mk)^*}$ is defined by bounding:

$$\begin{aligned}\left|\int_0^{s_{(mk)^*}} \varphi'_{(mk)^*}(u)\mathrm{d}u\right| &= \left|\int_0^{s_{(mk)^*}} \frac{iE\left[U_{(mk)^*}\exp\left(iU_{(mk)^*}u\right)\right]}{E\left[\exp\left(iU_{(mk)^*}u\right)\right]}\mathrm{d}u\right| \\ &\leq E\left[\left|U_{(mk)^*}\right|\right]\int_0^{s_{(mk)^*}} \frac{1}{\left|E\left[\exp\left(iU_{(mk)^*}u\right)\right]\right|}\mathrm{d}u \\ &< \infty\end{aligned}$$

where the first inequality follows from the triangle inequality and $|\exp(\cdot)| \leq 1$ and the second inequality follows from the assumptions $\int\limits_0^{s_{(mk)^*}} \left|\left(E\left[\exp\left(iU_{(mk)^*}u\right)\right]\right)^{-1}\right| \mathrm{d}u < \infty$ and $E\left[\left|U_{(mk)^*}\right|\right] < \infty$.

This shows that the CF of $U_{(mk)^*}$ is identified. The marginal density of $U_{(mk)^*}$ is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_{(mk)^*}(u_{(mk)^*}) = \frac{1}{2\pi}\int e^{-is'_{(mk)^*}u_{(mk)^*}}\phi_{(mk)^*}(s_{(mk)^*})\mathrm{d}s_{(mk)^*}$$

1.11 Appendix C

1.11.1 Proof of Lemma 1

Let $g_{\vec{t}}(\vec{Y}) = \prod_{p=1}^{P} Y_p^{\alpha_p}\exp\left(i\vec{Y}'\vec{t}\right)$

$$\begin{aligned}&\Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon\right) \\ &= \Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon \,\middle|\, E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] \geq \kappa\right)\cdot\Pr\left(E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] \geq \kappa\right) \\ &\qquad + \Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon \,\middle|\, E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] < \kappa\right)\cdot\Pr\left(E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] < \kappa\right) \\ &\leq \Pr\left(E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] \geq \kappa\right) + \Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon \,\middle|\, E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] < \kappa\right) \\ &= A_1 + A_2\end{aligned}$$

(i) Consider A_1

$$\Pr\left(E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right] \geq \kappa\right) \leq \frac{Var\left(E_N\left[\prod_{p=1}^{P}|Y_p|^{\alpha_p}\right]\right)}{\kappa^2} \leq \frac{E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]}{N\kappa^2}$$

where the first inequality follows by Chebyshev's inequality.

(ii) To bound A_2 I will use an argument that is similar to Pollard (1984) and Van De Geer (2006) but instead of using Hoeffding's inequality I use Bernstein's inequality as in Evdokimov (2010).

Define the L_1-covering number, $N_1(\varepsilon, \mathcal{Q}, \mathcal{G})$, as the smallest L for which there exist functions $g_1 \ldots, g_L$ such that $\min_l E_{\mathcal{Q}}||g - g_l|| \leq \varepsilon$ for all $g \in \mathcal{G}$ (e.g. Pollard (1984)).[45] I show that $N_1(\varepsilon, \mathcal{P}_N, \mathcal{G}) \lesssim \left(\frac{TE_N\left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right]}{\varepsilon}\right)^P$ where $\mathcal{P}_N$ is the empirical probability measure and $\mathcal{G}$ is the class of functions defined as $\mathcal{G} = \{g_{\vec{t}}(\vec{Y}) : \vec{t} \in [-T,T]^P\}$ where as before $g_{\vec{t}}(\vec{Y}) = \prod_{p=1}^P Y_p^{\alpha_p} \exp\left(i\vec{Y}'\vec{t}\right) \exp\left(i\vec{Y}\vec{t}\right)$, $p = 1, \ldots, P$.[46] Discretize $[-T,T]^P$ into $L = \left(\frac{4TPE_N\left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right]}{\varepsilon}\right)^P$ points, $\vec{t}_1, \ldots, \vec{t}_L$, by cutting $[-T,T]$ in each dimension into equidistant segments of length $\frac{\varepsilon}{2PE_N \prod_{p=1}^P |Y_p|^{2\alpha_p}}$. Let $g_l(\vec{Y}) = \prod_{p=1}^P Y_p^{\alpha_p} \exp\left(i\vec{Y}'\vec{t}\right) \exp\left(i\vec{Y}\vec{t}_l\right)$ for $\vec{t}_1, \ldots, \vec{t}_L$ chosen above. For any $\vec{t} \in [-T,T]^P$ there exists an l such that

$$
\begin{aligned}
& E_N \left| \prod_{p=1}^P Y_p^{\alpha_p} \exp\left(i\vec{Y}\vec{t}\right) - \prod_{p=1}^P Y_p^{\alpha_p} \exp\left(i\vec{Y}\vec{t}_l\right) \right| \\
&= E_N \left| \prod_{p=1}^P Y_p^{\alpha_p} \cos\left(\vec{Y}\vec{t}\right) + i \prod_{p=1}^P Y_p^{\alpha_p} \sin\left(\vec{Y}\vec{t}\right) - \prod_{p=1}^P Y_p^{\alpha_p} \cos\left(\vec{Y}\vec{t}_l\right) - i \prod_{p=1}^P Y_p^{\alpha_p} \sin\left(\vec{Y}\vec{t}_l\right) \right| \\
&\leq E_N \left| \prod_{p=1}^P Y_p^{\alpha_p} \cos\left(\vec{Y}\vec{t}\right) - \prod_{p=1}^P Y_p^{\alpha_p} \cos\left(\vec{Y}\vec{t}_l\right) \right| + E_N \left| i \prod_{p=1}^P Y_p^{\alpha_p} \sin\left(\vec{Y}\vec{t}\right) - i \prod_{p=1}^P Y_p^{\alpha_p} \sin\left(\vec{Y}\vec{t}_l\right) \right| \\
&\leq 2P \max_l \left\{\left|\vec{t} - \vec{t}_l\right|\right\} \cdot E_N \left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right] \\
&\leq \varepsilon
\end{aligned}
$$

It follows that the L_1-covering number satisfies $N_1(\varepsilon, P_N, \mathcal{G}) \lesssim \left(\frac{TE_N\left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right]}{\varepsilon}\right)^P$.

A_2 is now bounded using a symmetrization argument (e.g. Pollard (1984)), Bernstein's inequality, and the L_1-covering number:

$$
\begin{aligned}
& \Pr\left(\sup \left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon \,\middle|\, E_N\left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right] < \kappa\right) \\
&\leq 8N_1\left(\varepsilon/8, P_N, \mathcal{G}\right) \exp\left(-\frac{N\varepsilon^2}{64} \Big/ \left(2E\left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right] + \frac{2}{3}\varepsilon\kappa\right)\right) \\
&\lesssim \left(\frac{T\kappa}{\varepsilon}\right)^P \exp\left(-\frac{N\varepsilon^2}{64} \Big/ \left(2E\left[\prod_{p=1}^P |Y_p|^{2\alpha_p}\right] + \frac{2}{3}\varepsilon\kappa\right)\right)
\end{aligned}
$$

[45] $\mathcal{Q}$ is a probability measure and $\mathcal{G}$ is a class of functions in $\mathcal{L}^1(\mathcal{Q})$

[46] $Z_N \lesssim a_N$ means that there exists $C > 0$ such that $Z_N \leq Ca_N$.

For N large enough the bounds for A_1 and A_2 imply

$$\Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon\right) \leq A_1 + A_2$$
$$\lesssim \frac{E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]}{N\kappa^2} + \left(\frac{T\kappa}{\varepsilon}\right)^P \exp\left(-\frac{N\varepsilon^2}{64}\Big/\left(2E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right] + \frac{2}{3}\varepsilon\kappa\right)\right)$$

(iii) The last step is to apply the Borel-Cantelli Lemma. Index ε, T and κ by N and let

$$T_N = CN^{\delta/2} \qquad 0 < \delta$$
$$\varepsilon_N = C_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)}\left(\frac{\ln N}{N}\right)^{\frac{1}{2}}$$
$$\kappa_N = \left(N^{\delta_\kappa}\ln N\right)^{\frac{1}{2}} \qquad 0 < \delta_\kappa < 1$$

where $C_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)}$ is a constant that may depend on the arguments in the subscript. To simplify the notation a little denote $E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]$ by σ^2 and $C_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)}$ by C_ε. For N large enough

$$\begin{aligned}
&\Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right] - E\left[g_{\vec{t}}\right]\right| > \varepsilon_N\right)\\
&\lesssim \frac{\sigma^2}{N\kappa_N^2} + \left(\frac{T_N\kappa_N}{\varepsilon_N}\right)^P \exp\left(-\frac{N\varepsilon_N^2}{64}\Big/\left(2\sigma^2 + \frac{2}{3}\varepsilon_N\kappa_N\right)\right)\\
&= \frac{\sigma^2}{N\kappa_N^2} + \exp\left(P\ln\left(\frac{T_N\kappa_N}{\varepsilon_N}\right) - \frac{N\varepsilon_N^2}{64}\Big/\left(2\sigma^2 + \frac{2}{3}\varepsilon_N\kappa_N\right)\right)\\
&\leq \frac{\sigma^2}{N^{1+\delta_\kappa}\ln N} + \exp\left(P\ln\left(\frac{C\left(N^\delta N^{\delta_\kappa}\ln N\right)^{\frac{1}{2}}}{C_\varepsilon\left(\frac{\ln N}{N}\right)^{\frac{1}{2}}}\right) - \frac{N\left(C_\varepsilon^2\frac{\ln N}{N}\right)}{64}\Big/\left(2\sigma^2 + \frac{2}{3}C_\varepsilon\frac{\ln N}{N^{(1-\delta_\kappa)/2}}\right)\right)\\
&= \frac{\sigma^2}{N^{1+\delta_M}\ln N} + \exp\left(P\ln\left(\frac{C}{C_\varepsilon}N^{(\delta+\delta_\kappa+1)/2}\right) - \frac{C_\varepsilon^2\ln N}{64}\Big/\left(2\sigma^2 + \frac{2}{3}C_\varepsilon\frac{\ln N}{N^{(1-\delta_\kappa)/2}}\right)\right)\\
&\leq \frac{\sigma^2}{N^{1+\delta_M}\ln N} + \exp\left(P\ln\left(\frac{C}{C_\varepsilon}N^{(\delta+\delta_\kappa+1)/2}\right) - \frac{C_\varepsilon^2\ln N}{128(\sigma^2+1)}\right)\\
&= \frac{\sigma^2}{N^{1+\delta_M}\ln N} + \exp\left(P\ln\left(\frac{C}{C_\varepsilon}\right) + \left[\frac{P(\delta+1)}{2} - \frac{C_\varepsilon^2}{128(\sigma^2+1)}\right]\ln N\right)\\
&\lesssim_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)} \frac{\sigma^2}{N\ln N} + \exp\left(\left[\frac{P(\delta+\delta_\kappa+1)}{2} - \frac{C_\varepsilon^2}{128(\sigma^2+1)}\right]\ln N\right)\\
&= \frac{\sigma^2}{N^{1+\delta_\kappa}\ln N} + \frac{1}{N^{1+\beta}}
\end{aligned}$$

where $\lesssim_{\left(P,\delta,E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]\right)}$ means that the constant depends on P, δ, and $E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]$ and C_ε^2 is chosen so that $\beta = -\frac{P(\delta+\delta_\kappa+1)}{2} + \frac{C_\varepsilon^2}{128(\sigma^2+1)} - 1 > 0$ so that C_ε depends on P, δ, δ_κ, and $E\left[\prod_{p=1}^{P}|Y_p|^{2\alpha_p}\right]$.

For the above choices of ε_N, T_N, and κ_N

$$\sum_{N=1}^{\infty} \Pr\left(\sup\left|E_N\left[g_{\vec{t}}\right]-E\left[g_{\vec{t}}\right]\right|>\varepsilon_N\right) \lesssim \sum_{N=1}^{\infty}\left(\frac{\sigma^2}{N^{1+\delta_\kappa}\ln N}+\frac{1}{N^{1+\beta}}\right)<\infty$$

The Borel-Cantelli lemma then implies that

$$\sup\left|E_N\left[g_{\vec{t}}\right]-E\left[g_{\vec{t}}\right]\right| \leq \varepsilon_N \qquad \text{a.s}$$

for N large enough.

1.11.2 Proof of Theorem 4

I use Lemmas 1 and 2 and a Taylor expansion. For N large enough

$$\begin{aligned}
&\sup_{s\in[-S_N,S_N]}\left|\widehat{\phi}_{m^*}(s)-\phi_{m^*}(s)\right| \\
&=\sup_{s\in[-S_N,S_N]}\left|\exp\left(\int_0^s \frac{\widehat{\phi}_{Y_p}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)-\exp\left(\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)\right| \\
&=\sup_{s\in[-S_N,S_N]}\left|\exp\left(\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)\left[\left(\int_0^s \frac{\widehat{\phi}_{Y_p}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u-\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)\right.\right. \\
&\qquad\qquad\left.\left.+o\left(\int_0^s \frac{\widehat{\phi}_{Y_p}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u-\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)\right]\right| \\
&=\sup_{s\in[-S_N,S_N]}\left|\exp\left(\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)\left[\int_0^s \frac{1}{\phi_{\vec{Y}}(u\vec{t})}\left(\widehat{\phi}_{Y_p}(u\vec{t})-\phi_{Y_p}(u\vec{t})\right)\mathrm{d}u\right.\right. \\
&\qquad\qquad-\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})-\phi_{\vec{Y}}(u\vec{t})\right)\mathrm{d}u \\
&\qquad\left.\left.+o\left(\left|\int_0^s \frac{1}{\phi_{\vec{Y}}(u\vec{t})}\left(\widehat{\phi}_{Y_p}(u\vec{t})-\phi_{Y_p}(u\vec{t})\right)\mathrm{d}u\right|+\left|\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})-\phi_{\vec{Y}}(u\vec{t})\right)\mathrm{d}u\right|\right)\mathrm{d}u\right]\right| \\
&\lesssim \sup_{s\in[-S_N,S_N]}\int_0^s \frac{1}{|\phi_{\vec{Y}}(u\vec{t})|}\left|\widehat{\phi}_{Y_p}(u\vec{t})-\phi_{Y_p}(u\vec{t})\right|\mathrm{d}u+\sup_{s\in[-S_N,S_N]}\int_0^s \frac{|\phi_{Y_p}(u\vec{t})|}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left|\widehat{\phi}_{\vec{Y}}(u\vec{t})-\phi_{\vec{Y}}(u\vec{t})\right|\mathrm{d}u \\
&\leq \sup_{s\in[-S_N,S_N]}\left|\widehat{\phi}_{Y_p}(u\vec{t})-\phi_{Y_p}(u\vec{t})\right|\int_{-S_N}^{S_N}\left|\frac{1}{\phi_{\vec{Y}}(u\vec{t})}\right|\mathrm{d}u+\sup_{s\in[-S_N,S_N]}\left|\widehat{\phi}_{\vec{Y}}(u\vec{t})-\phi_{\vec{Y}}(u\vec{t})\right|\int_{-S_N}^{S_N}\frac{|\phi_{Y_p}(u\vec{t})|}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\mathrm{d}u \\
&\lesssim \varepsilon_N E\left[|Y_p|\right]\int_{-S_N}^{S_N}\frac{1}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\mathrm{d}u
\end{aligned}$$

where the second equality uses the Taylor expansion $e^x = e^{x_0} + e^{x_0}(x-x_0) + e^{x_0}o|x-x_0|$, the third equality uses the Taylor expansion $\frac{x}{y} = \frac{x_0}{y_0} + \frac{1}{y_0}(x-x_0) - \frac{x_0}{y_0^2}(y-y_0) + o\left(|\frac{1}{y_0}(x-x_0)| + |\frac{x_0}{y_0^2}(y-y_0)|\right)$, the first $\lesssim$ by the triangle inequality, $\left|\exp\left(\int_0^s \frac{\phi_{Y_p}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\right)\right| \leq 1$ because it is a CF, and the implications of the little-o notion, and the last inequality from Lemma 1.[47]

As before, use Lemmas 1 and 2 and a Taylor expansion. For N large enough

$$
\begin{aligned}
&\sup_{s\in[-S_N,S_N]}\left|\widehat{\phi}_{m^*}(s)-\phi_{m^*}(s)\right| \\
&=\sup_{s\in[-S_N,S_N]}\Bigg|\exp\left(\int_0^s\int_0^v \frac{\widehat{\phi}_{Y_{p_1}}(u\vec{t})\widehat{\phi}_{Y_{p_2}}(u\vec{t})}{\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right) \\
&\qquad -\exp\left(\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right)\Bigg| \\
&=\sup_{s\in[-S_N,S_N]}\Bigg|\exp\left(\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right)\times \\
&\qquad\Bigg[\int_0^s\int_0^v \frac{\widehat{\phi}_{Y_{p_1}}(u\vec{t})\widehat{\phi}_{Y_{p_2}}(u\vec{t})}{\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v-\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v \\
&+o\left(\left|\int_0^s\int_0^v \frac{\widehat{\phi}_{Y_{p_1}}(u\vec{t})\widehat{\phi}_{Y_{p_2}}(u\vec{t})}{\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\widehat{\phi}_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right|+\left|\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right|\right)\Bigg]\Bigg| \\
&=\sup_{s\in[-S_N,S_N]}\Bigg|\exp\left(\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})}\mathrm{d}u\mathrm{d}v\right) \\
&\times\Bigg[\int_0^s\int_0^v \frac{\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left(\widehat{\phi}_{Y_{p_1}}(u\vec{t})-\phi_{Y_{p_1}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v+\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left(\widehat{\phi}_{Y_{p_2}}(u\vec{t})-\phi_{Y_{p_2}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v \\
&\qquad -\int_0^s\int_0^v \frac{1}{\phi_{\vec{Y}}(u\vec{t})}\left(\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})-\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v \\
&\qquad +\int_0^s\int_0^v \left(\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{2\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^3}\right)\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})-\phi_{\vec{Y}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v \\
&+o\Bigg(\left|\int_0^s\int_0^v \frac{\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left(\widehat{\phi}_{Y_{p_1}}(u\vec{t})-\phi_{Y_{p_1}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v\right|+\left|\int_0^s\int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}\left(\widehat{\phi}_{Y_{p_2}}(u\vec{t})-\phi_{Y_{p_2}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v\right| \\
&\qquad +\left|\int_0^s\int_0^v \frac{1}{\phi_{\vec{Y}}(u\vec{t})}\left(\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})-\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v\right| \\
&\qquad +\left|\int_0^s\int_0^v \left(\frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2}-\frac{2\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^3}\right)\left(\widehat{\phi}_{\vec{Y}}(u\vec{t})-\phi_{\vec{Y}}(u\vec{t})\right)\mathrm{d}u\mathrm{d}v\right|\Bigg)\Bigg]\Bigg| \\
&\lesssim\sup_{s\in[-S_N,S_N]}\Bigg[\int_0^s\int_0^v \frac{\left|\phi_{Y_{p_2}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^2}\left|\widehat{\phi}_{Y_{p_1}}(u\vec{t})-\phi_{Y_{p_1}}(u\vec{t})\right|\mathrm{d}u\mathrm{d}v+\int_0^s\int_0^v \frac{\left|\phi_{Y_{p_1}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^2}\left|\widehat{\phi}_{Y_{p_2}}(u\vec{t})-\phi_{Y_{p_2}}(u\vec{t})\right|\mathrm{d}u\mathrm{d}v \\
&+\int_0^s\int_0^v \frac{1}{\left|\phi_{\vec{Y}}(u\vec{t})\right|}\left|\widehat{\phi}_{Y_{p_1}Y_{p_2}}(u\vec{t})-\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})\right|\mathrm{d}u\mathrm{d}v
\end{aligned}
$$

[47] $d_N=o(e_N)$ is Little-o notation and means that for every $\delta>0$ there exists N large enough so that $d_n\leq\delta e_n$ for all $n>N$.

$$
+\int_0^s \int_0^v \left(\frac{\left|\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^2} + \frac{\left|\phi_{Y_{p_1}}(u\vec{t})\right| \left|\phi_{Y_{p_2}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^3} \right) \left|\widehat{\phi}_{\vec{Y}}(u\vec{t}) - \phi_{\vec{Y}}(u\vec{t})\right| \mathrm{d}u\mathrm{d}v \Bigg]
$$

$$
\le \varepsilon_N \left(\int_{-S_N}^{S_N} \int_0^v \frac{\left|\phi_{Y_{p_2}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^2} \mathrm{d}u\mathrm{d}v + \int_{-S_N}^{S_N} \int_0^v \frac{\left|\phi_{Y_{p_1}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^2} \mathrm{d}u\mathrm{d}v + \int_{-S_N}^{S_N} \int_0^v \frac{1}{\left|\phi_{\vec{Y}}(u\vec{t})\right|} \mathrm{d}u\mathrm{d}v \right.
$$

$$
\left. + \int_{-S_N}^{S_N} \int_0^v \left(\frac{\left|\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^2} + \frac{\left|\phi_{Y_{p_1}}(u\vec{t})\right| \left|\phi_{Y_{p_2}}(u\vec{t})\right|}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^3} \right) \mathrm{d}u\mathrm{d}v \right)
$$

$$
\lesssim \varepsilon_N \left(E\left[|Y_{p_1}|\right] + E\left[|Y_{p_2}|\right] + E\left[|Y_{p_1}Y_{p_2}|\right] \right) \int_{-S_N}^{S_N} \int_0^v \frac{1}{\left|\phi_{\vec{Y}}(u\vec{t})\right|^3} \mathrm{d}u\mathrm{d}v
$$

where the second equality uses the Taylor expansion $e^x = e^{x_0} + e^{x_0}(x - x_0) + e^{x_0} o(x - x_0)$, the third equality uses the Taylor expansion $\frac{x}{y} = \frac{x_0}{y_0} + \frac{1}{y_0}(x - x_0) - \frac{x_0}{y_0^2}(y - y_0) + o\left(\left|\frac{1}{y_0}(x - x_0)\right| + \left|\frac{x_0}{y_0^2}(y - y_0)\right|\right)$, the first $\lesssim$ by the triangle inequality, $\left|\exp\left(\int_0^s \int_0^v \frac{\phi_{Y_{p_1}}(u\vec{t})\phi_{Y_{p_2}}(u\vec{t})}{\left(\phi_{\vec{Y}}(u\vec{t})\right)^2} - \frac{\phi_{Y_{p_1}Y_{p_2}}(u\vec{t})}{\phi_{\vec{Y}}(u\vec{t})} \mathrm{d}u\mathrm{d}v\right)\right| \le 1$ because it is a CF, and the implications of the little-o notion, and the inequality from Lemma 1.

1.11.3 Proof of Theorem 5

For all u in the support of U_{m^*} and for N large enough

$$
\begin{aligned}
&\left|\widehat{f}_{m^*}(u) - f_{m^*}(u)\right| \\
&= \left| \frac{1}{2\pi} \int e^{-isu} \widehat{\phi}_{m^*}(s) \phi_K(sh_N)\,\mathrm{d}s - \frac{1}{2\pi} \int e^{-isu} \phi_{m^*}(s) \mathrm{d}s \right| \\
&= \left| \frac{1}{2\pi} \int e^{-isu} \left(\widehat{\phi}_{m^*}(s)\phi_K(sh_N) - \phi_{m^*}(s)\phi_K(sh_N) + \phi_{m^*}(s)\phi_K(sh_N) - \phi_{m^*}(s) \right) \mathrm{d}s \right| \\
&= \left| \frac{1}{2\pi} \int e^{-isu} \phi_K(sh_N) \left(\widehat{\phi}_{m^*}(s) - \phi_{m^*}(s) \right) \mathrm{d}s + \frac{1}{2\pi} \int e^{-isu} \phi_{m^*}(s) \left(\phi_K(sh_N) - 1 \right) \mathrm{d}s \right| \\
&\le \frac{1}{2\pi} \int \left|\phi_K(sh_N)\right| \left|\widehat{\phi}_{m^*}(s) - \phi_{m^*}(s)\right| + \frac{1}{2\pi} \int \left|\phi_{m^*}(s)\right| \left|\phi_K(sh_N) - 1\right| \mathrm{d}s \\
&\le \frac{1}{2\pi} \int_{-S_N}^{S_N} \left|\widehat{\phi}_{m^*}(s) - \phi_{m^*}(s)\right| \mathrm{d}s + \frac{1}{2\pi} \int_{-S_N}^{S_N} \left|\phi_{m^*}(s)\right| \left|m(sh_N)\,(sh_N)^q\right| \mathrm{d}s \\
&\qquad + \frac{1}{2\pi} \int_{S_N}^{\infty} \left|\phi_{m^*}(s)\right| \mathrm{d}s + \frac{1}{2\pi} \int_{-\infty}^{-S_N} \left|\phi_{m^*}(s)\right| \mathrm{d}s \\
&\lesssim S_N \sup_{s \in [-S_N, S_N]} \left|\widehat{\phi}_{m^*}(s) - \phi_{m^*}(s)\right| + \sup_{s \in [-1,1]} \left|m(s)\right| h_N^q \int_{-S_N}^{S_N} \left|\phi_{m^*}(s)\right| |s|^q \mathrm{d}s \\
&\qquad + \int_{-\infty}^{-S_N} \left|\phi_{m^*}(s)\right| \mathrm{d}s + \int_{S_N}^{\infty} \left|\phi_{m^*}(s)\right| \mathrm{d}s
\end{aligned}
$$

where the second inequality follows because $|\phi_K(s)| < 1$, $\phi_K(s) = 1 + m(s)s^q$ for $s \in [-1, 1]$ and $\phi_K(s) = 0$ otherwise and $m(s)$ is continuous for $s \in [-1, 1]$.

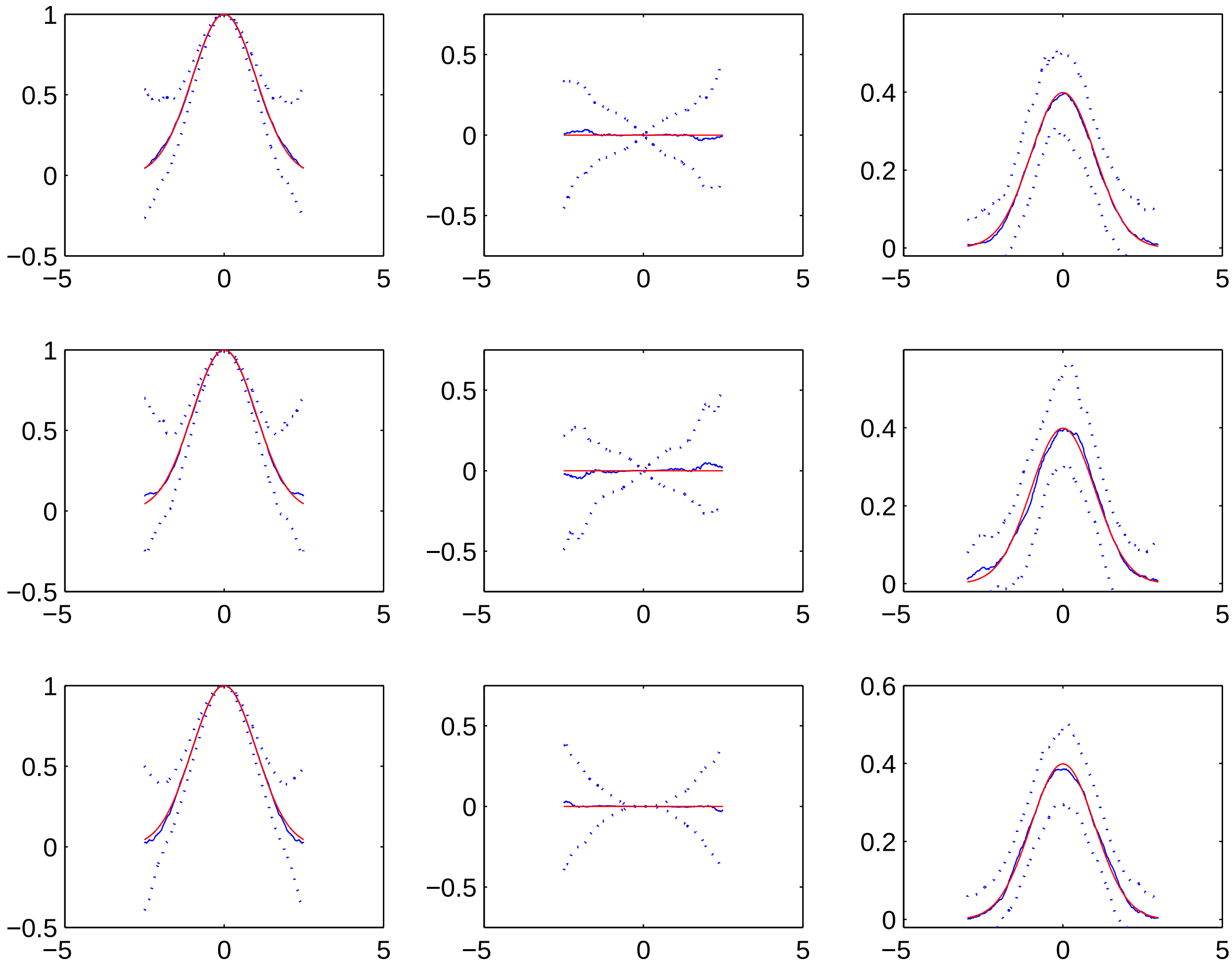

Figure 1.1: Experiment 1: $X^* \sim$ Normal$(0, 1)$, $\varepsilon_1 \sim$ Normal$(0, 1)$, $\varepsilon_2 \sim$ Normal$(0, 1)$ with $N = 100$

The left column is the real part of the characteristic function, the middle column is the imaginary part of the characteristic function and the right column is the density. The first through third rows are estimators A though C, respectively.

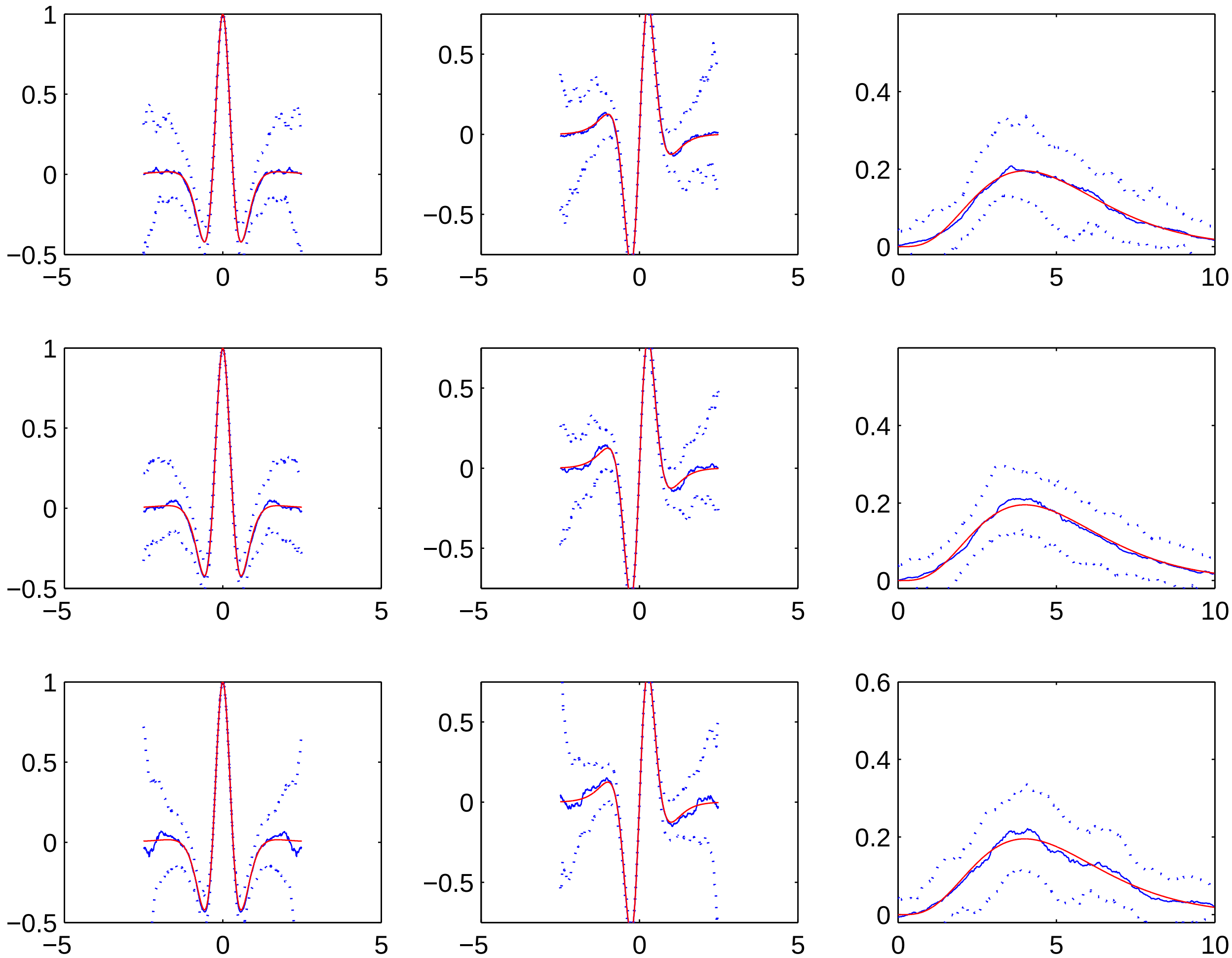

Figure 1.2: Experiment 2: $X^* \sim \text{Gamma}(5,1)$, $\varepsilon_1 \sim \text{Normal}(0,1)$, $\varepsilon_2 \sim \text{Normal}(0,1)$ with $N = 100$

The left column is the real part of the characteristic function, the middle column is the imaginary part of the characteristic function and the right column is the density. The first through third rows are estimators A though C, respectively.

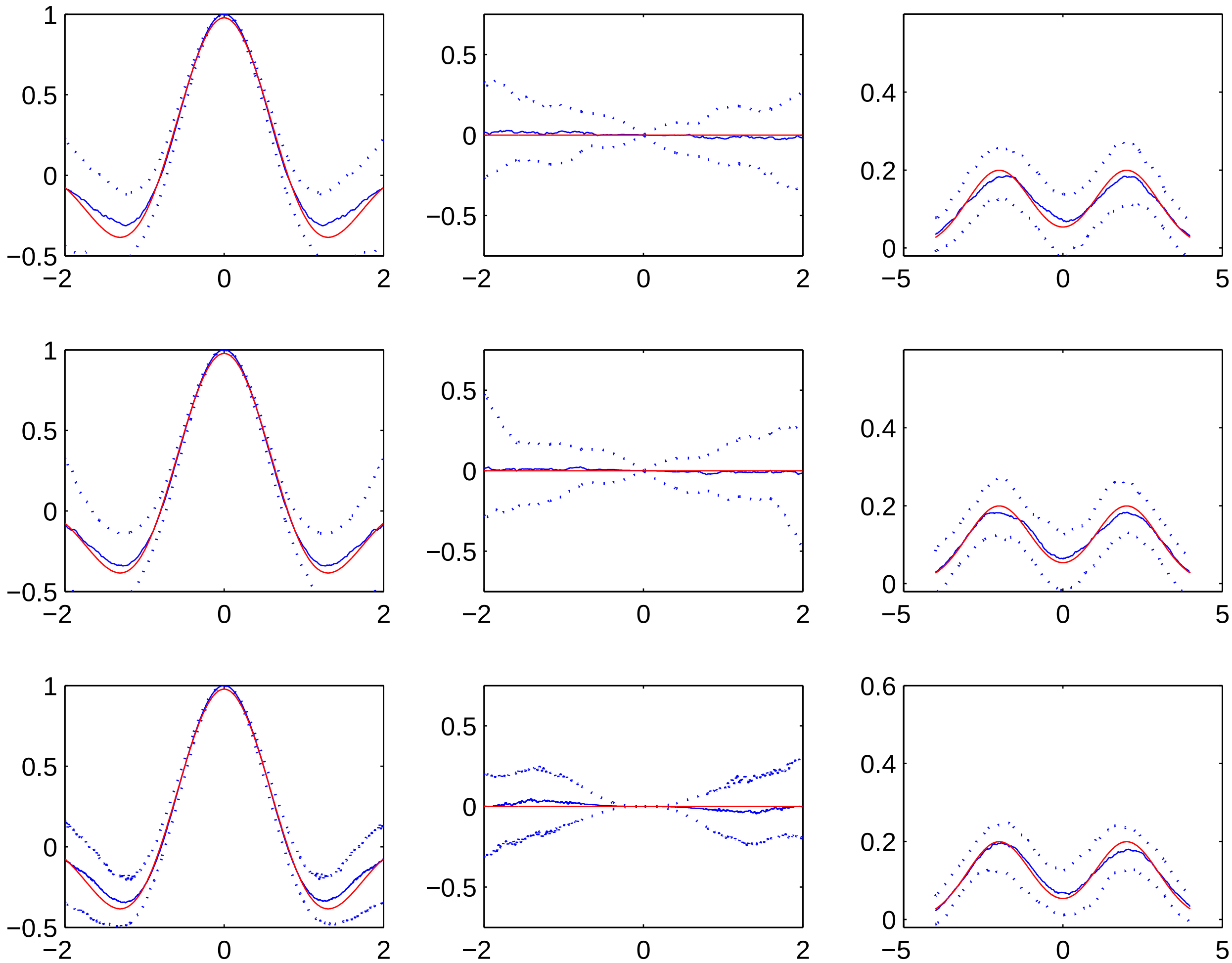

Figure 1.3: Experiment 3: $X^* \sim \frac{1}{2}N(-2,1) + \frac{1}{2}N(2,1)$ (Bimodal), $\varepsilon_1 \sim \text{Normal}(0,1)$, $\varepsilon_2 \sim \text{Normal}(0,1)$ with $N = 100$
The left column is the real part of the characteristic function, the middle column is the imaginary part of the characteristic function and the right column is the density. The first through third rows are estimators A though C, respectively.

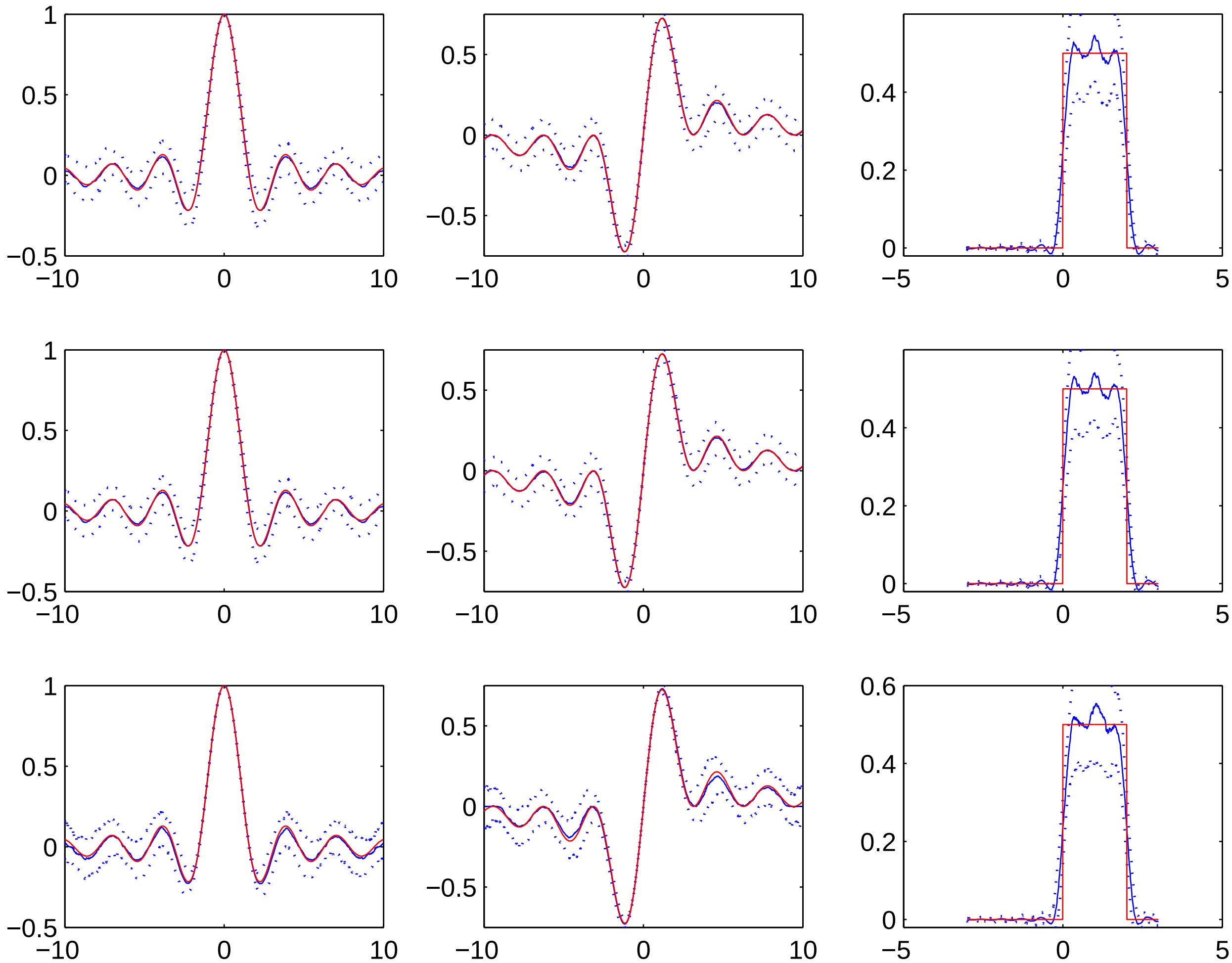

Figure 1.4: Experiment 4: $X^* \sim$Unif$(0, 1)$, $\varepsilon_1 \equiv 0$, $\varepsilon_2 \equiv 0$ with $N = 100$
The left column is the real part of the characteristic function, the middle column is the imaginary part of the characteristic function and the right column is the density. The first through third rows are estimators A though C, respectively.

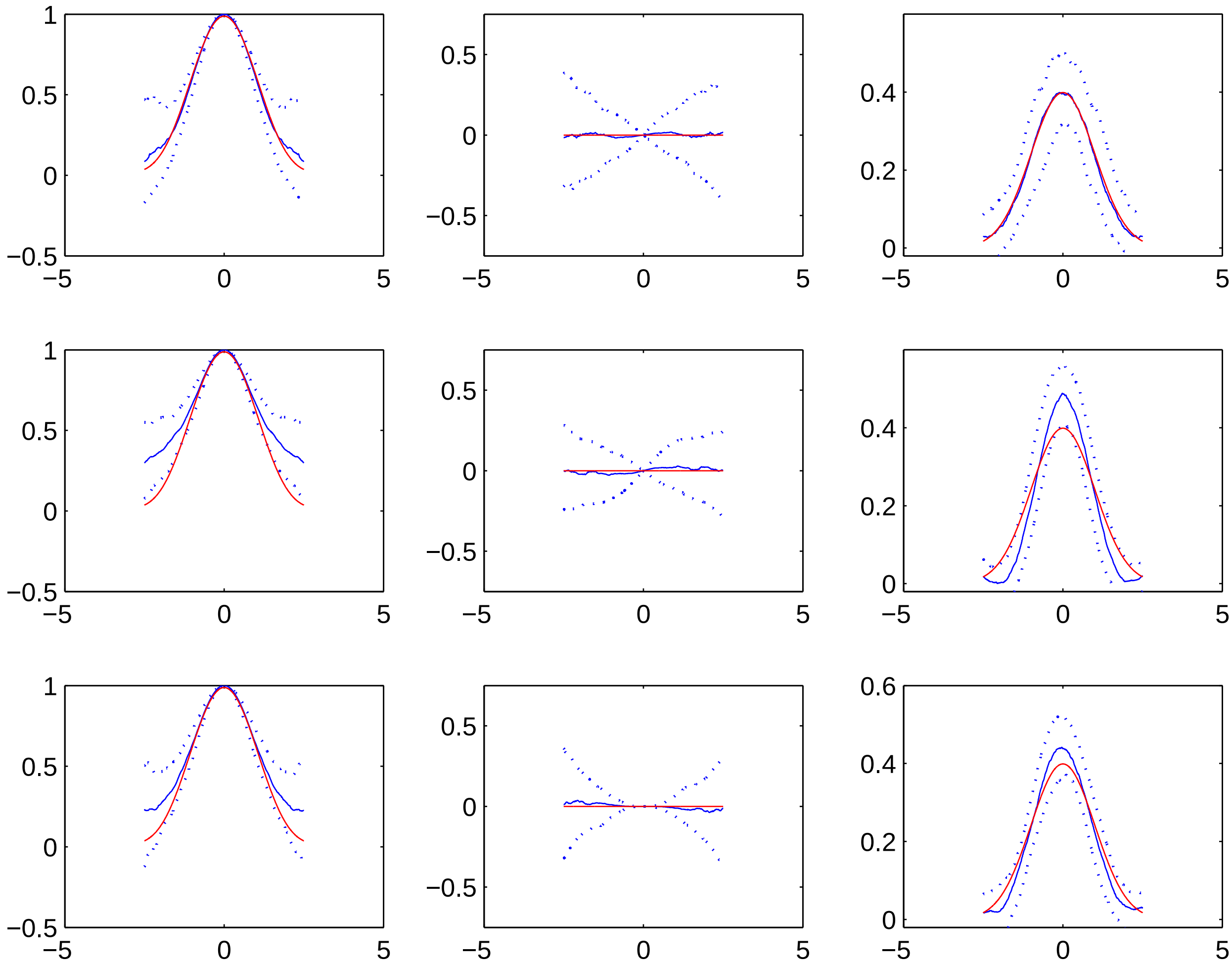

Figure 1.5: Experiment 5: $X^* \sim \text{Normal}(0,1)$ (X^* and ε_1 dependent), $\varepsilon_1 \sim \text{Normal}(0, x^{*2})$, $\varepsilon_2 \sim \text{Normal}(0,1)$ with $N = 100$
The left column is the real part of the characteristic function, the middle column is the imaginary part of the characteristic function and the right column is the density. The first through third rows are estimators A though C, respectively.

Table 1.1: Comparing Estimators in Measurement Error Model With a Repeated Measurement with N=100

Experiment		Estimator A	Estimator B	Estimator C
Norm(0,1)	MISE	0.0429	0.0672	0.0391
Gamma(5,1)	MISE	0.2104	0.0393	>1,000
Bimodal	MISE	0.0326	0.0324	>1,000
Norm(0,1) (Depend)	MISE	0.0404	0.0348	>1,000
Unif(0,1)	MISE	0.0292	0.0300	0.0195

Table 1.2: Comparing Estimators: Measurement Error Model With a Repeated Measurement with N=1,000

Experiment		Estimator A	Estimator B	Estimator C
Norm(0,1)	MISE	0.0066	0.0071	0.0025
Gamma(5,1)	MISE	0.0365	0.0048	>1,000
Bimodal	MISE	0.0124	0.0024	>1,000
Norm(0,1) (Depend)	MISE	6.3110	0.0201	>1,000
Unif(0,1)	MISE	0.0039	0.0155	0.0058

Table 1.3: Comparing Estimators: Measurement Error Model With a Repeated Measurement with N=10,000

Experiment		Estimator A	Estimator B	Estimator C
Norm(0,1)	MISE	0.0008	0.0007	0.0004
Gamma(5,1)	MISE	0.0127	0.0007	>1,000
Bimodal	MISE	13.9634	0.0003	>1,000
Norm(0,1) (Depend)	MISE	>1,000	0.0187	>1,000
Unif(0,1)	MISE	0.0005	0.0148	0.0044

Bibliography

[1] ANGRIST, J. and KRUEGER, A. (2000), "Empirical Strategies in Labor Economics," *Handbook of Labor Economics*, 1277-1366.

[2] ARELLANO, M. and HONORÉ, B. (2001), "Panel Data Models: Some Recent Developments," *Handbook of Econometrics*, Vol. 5, ed. by J. J. Heckman, and E. Leamer. North-Holland.

[3] ARELLANO, M. and HAHN, J. (2005), "Understanding Bias in Nonlinear Panel Models: Some Recent Developments," Invited Lecture, Econometric Society World Congress, London.

[4] ALTONJIl, J. G. and MATZKIN, R. L. (2005), "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, **73**, 1053-1102.

[5] BESTER, A. C. and HANSEN, C. (2011), "Identification of Marginal Effects in a Nonparametric Correlated Random Effects Model,", *Journal of Business and Economic Statistics*, forthcoming.

[6] BONDESSON, L. (1974), "Characterizations of Probability Laws Through Constant Regression," *Z. Wahrsch. v. Geb*, **29**, 93-115.

[7] BONHOMME, S., and SAUDER, U. (2010), "Recovering Distributions in Difference-in-Differences: A Comparison of Selective and Comprehensive Schooling," *Review of Economics and Statistics*, forthcoming.

[8] BONHOMME, S. and ROBIN, J.M. (2010), "Generalized Non-Parametric Deconvolu-

tion with an Application to Earnings Dynamics," *Review of Economic Studies*, **77** (2), 491-533.

[9] BOUND, J., BROWN C., DUNCAN G., and RODGERS W., (1994), "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, **12**, 345-368.

[10] BLUNDELL, R. and MACURDY, T. (2000), "Labour Supply: A Review of Alternative Approaches," *Handbook of Labor Economics*, Vol 3a, North Holland.

[11] BROWNING, M. and CARRO, J. (2007), "Heterogeneity and Microeconometrics Modelling," in Blundell, R., W.K. Newey, T. Persson (eds.), *Advances in Theory and Econometrics*, Vol. 3 ; Cambridge: Cambridge University Press.

[12] CARNEIRO, P., HANSEN, K., and HECKMAN, J.J. (2003), "Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice," *International Economic Review* **44 (2)**, 361-422

[13] CARRASCO, M., FLORENS J.-P. (2010), "Spectral Method for Deconvolving a Density," *Econometric Theory*, forthcoming

[14] CARROLL, R.J., RUPPERT, D., STEFANSKI, L.A., CRAINICEANU, C. (2006), *Measurement Error in Nonlinear Models: A Modern Perspective, Second Edition* (Chapman and Hall).

[15] CARROLL, R. J. and STEFANSKI, L. A. (1990). Carroll, "Approximate quasi-likelihood estimation in models with surrogate predictors," *Journal of the American Statistical Association*, **85**, 652-663.

[16] CHEN, X., HU, Y. and LEWBEL, A. (2009), "Nonparametric Identification and Estimation of Nonclassical Errors-in-Variables Models Without Additional Information", *Statistica Sinica*, **19**,949-968.

[17] CHEN, X., HONG, H. and NEKIPELOV, D. (2011), "Nonlinear Models of Measurement Errors," *Journal of Economic Literature*, forthcoming

[18] CHEN, X., HONG, H. and TAMER, E. (2005), "Measurement Error Models with Auxiliary Data." *Review of Economic Studies*, **72**, No. 2.

[19] CHERNOZHUKOV, V., FERNANDEZ-VAL, I., HAHN, J. and NEWEY, W. (2010), "Average and Quantile Effects in Nonseparable Panel Models," working paper.

[20] CHESHER, A. (2007), "Instrumental Values," *Journal of Econometrics*, **139**, 15-34.

[21] CHESHER, A. (2009), "Excess Heterogeneity, Endogeneity and Index Restrictions," *Journal of Econometrics*, **152**, 37-45.

[22] CSÖRGO, S. (1981), "Limit Behaviour of the Empirical Characteristic Function," *Annals of Probability*, **9** (1), 130-144.

[23] CUNHA, F., HECKMAN, J., SCHENNACH, S. M., (2010), "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*. **78**, 883-931.

[24] DELAIGLE, A. and GIJBELS, I. (2002), "Estimation of Integrated Squared Density Derivatives from a Contaminated Sample," *Journal of the Royal Statistical Society, Series B*, **64**, 869-886.

[25] DELAIGLE, A., HALL, P. and MEISTER, A. (2008), "On Deconvolution with Repeated Measurements," *Annals of Statistics*, **36** (2), 665-685.

[26] EVDOKIMOV, K. (2011), "Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity," working paper.

[27] EVDOKIMOV, K. WHITE, H. (2011). "An Extension of a Lemma of Kotlarski," *Econometric Theory*, forthcoming.

[28] FAN, J. Q. (1991), "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *Annals of Statistics*, **19**, 1257-1272.

[29] HAUSMAN, J.A., NEWEY W.K., and POWELL J.L., 1995, "Nonlinear errors in variables: estimation of some Engel curves," *Journal of Econometrics*, **65**, 205-233.

[30] HECKMAN, J. MATZKIN, R. L. and NESHEIM, L. (2010), "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, **78**, 1569-1591.

[31] HODERLEIN, S., and WHITE, H. (2009), "Nonparametric Identification in Nonsepa-

rable Panel Data Models with Generalized Fixed Effects", working paper.

[32] HOROWITZ, J. L. and MARKATOU, M. (1996), "Semiparametric Estimation of Regression Models for Panel Data," Review of Economic Studies, 63, 145-168.

[33] HSIAO, C. (1986). *Analysis of Panel Data* (Cambridge: Cambridge University Press).

[34] HSIAO, C. and WANG, Q.K., (2000), "Estimation of structural nonlinear errors-in-variables models by simulated least-squares method," *International Economic Review*, Vol. 41, No. 2, 523-542.

[35] HU, Y. (2008), "Identification and Estimation of Nonlinear Models with Misclassification Error using Instrumental Variables: A General Solution," *Journal of Econometrics*, vol. 144 , issue 1, pages 27-61.

[36] HU, Y. and RIDDER, G. (2010), "On Deconvolution as a First Stage Nonparametric Estimator," *Econometric Reviews*, **29**, 1-32.

[37] HU, Y. and RIDDER, G. (2011), "Estimation of Nonlinear Models with Mismeasured Regressors Using Marginal Information," *Journal of Applied Econometrics*, forthcoming

[38] HU, Y. and SCHENNACH, S. M. (2007), "Instrumental variable treatment of nonclassical measurement error models," *Econometrica*, **76**, 195-216.

[39] JUHN, C., MURPHY, K., and PIERCE, B. (1991), "Accounting for the Slowdown in Black-WhiteWage Convergence," *Workers and Their Wages*, pages 107-143.

[40] KLEPPER, S. and LEAMER, E. (1984), "Consistent sets of estimates for regressions with errors in all variables," *Econometrica*, **52**, 163-183.

[41] KHATRI, C. G. and RAO, C. R. (1968), "Solutions to Some Functional Equations and their Applications to Characterization of Probability Distributions," *Sankhyä*, **30**, 167-180.

[42] KHATRI, C. G. and RAO, C. R. (1972), "Functional Equations and Characterization of Probability Laws Through Luinear Functions of Random Variables," *Journal of Multivariate Analysis*, **2**, 162-173.

[43] KOTLARSKI, I. (1967), "On Characterizing the Gamma and Normal Distribution,"

Pacific Journal of Mathematics, **20**, 69-76.

[44] LEVINSOHN, J. and PETRIN, A.(2003), "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 317-342.

[45] LI, T. (2002),"Robust and Consistent Estimation of Nonlinear Errors-in-Variables Models," *Journal of Econometrics*, **110**, 126.

[46] LI, T., PERRIGNE, I. and VUONG, Q. (2000), "Conditionally Independent Private Information in OSC Wildcat Auctions," *Journal of Econometrics*, **98**, 129-161.

[47] LI, T. and VUONG, Q. (1998), "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, **65**, 139-165.

[48] MATZKIN, R. L. (2003), "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, **71** (5), 1339-1375.

[49] MATZKIN, R. L. (2007), "Nonparametric identification," *Handbook of Econometrics*, **6**, 5307-5368.

[50] MEGHIR, C. and PISTAFERRI, L. (2004),"Income Variance Dynamics and Heterogeneity", *Econometrica*, **72**, 132.

[51] MEGHIR, C. and PISTAFERRI, L. (2011),"Earnings, Consumption and Life-Cycle Choices", *Handbook of Labor Economics*, **9**, 773854.

[52] MEISTER, A. (2007), "Deconvolving Compactly Supported Densities." *Mathematical Methods of Statistics*, **16**, 63-76.

[53] NEWEY, W.K. (1994), "Kernel Estimation of Partial Means and a General Variance Estimator," *Econometric Theory*, **10**, 233-253.

[54] OLLEY, S. and PAKES, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, **64**, 1263-1295.

[55] POLLARD, D. (1984), *Convergence of Stochastic Processes* (New York: Springer).

[56] RAO, C. R. (1971), "Characterization of Probability Laws by Linear Functions," *Sankhyä*, **33**, 265-270.

[57] SCHENNACH, S. M. (2004a), "Estimation of Nonlinear Models with Measurement

Error," *Econometrica*, **72** (1), 33-75.

[58] SCHENNACH, S. M. (2004b), "Nonparametric Estimation in the Presence of Measurement Error," *Econometric Theory* **20**, 1046-109.

[59] SCHENNACH, S. M. and HU, Y. (2007), "Nonparametric Identification of the Classical Errors-in-Variables Model Without Side Information," working paper.

[60] SZÉKELY, G. J. and RAO, C. R. (2000), "Identifiability of Distributions of Independent Random Variables by Linear Combinations and Moments," *Sankhyä*, **62**, 193-202.

Chapter 2

Identification and Estimation for Regressions with Errors in All Variables

2.1 Introduction

In this paper I study identification of the coefficients, $\beta_1, \ldots, \beta_M$, in the linear regression model with measurement error in all the variables

$$\begin{aligned} Y &= \beta_0 + \beta_1 X_1^* + \ldots + \beta_M X_M^* + \varepsilon \\ X_m &= X_m^* + U_m \qquad\qquad m = 1, \ldots, M \end{aligned} \tag{2.1}$$

where Y is an observed outcome, X_m is an observed measurement of the unobserved explanatory variable X_m^*, and ε and U_m are measurement errors.

Estimation techniques that ignore the measurement errors in the explanatory variables, such as Ordinary Least Squares, lead to biased estimates of the coefficients. Solutions in the literature have focused on using additional information such as repeated measurements (Li and Vuong (1998), Schennach (2004a)), instrumental variables (Hausman, Ichimura, Newey, and Powell (1991), Carroll and Stefanski (1996)), signal-to-noise ratio (Fuller (1986)), known measurement error distributions (Hu and Ridder (2012)) validation data (Chen, Hong, and Tamer (2005)), or bounding the coefficients (Klepper and Leamer (1984)).

I develop a new method that identifies the coefficients under an assumption about a characteristic function (CF) that is testable in the data.[1] This method uses a CF transformation of the data, which contains more information than the moments of the observed variables. The main idea is to view the partial derivatives of a log CF as a moment adjusted by a direction. Thus, instead of the moment $E[YX_1]$ I use $E[YX_1e^{is_0Y+is_1X_1}]$, where $(s_0Y + s_1X_1)$ is the direction of the moment. The coefficients are identified by minimizing a distance between two of these partial derivatives evaluated at two different choices of (s_0, s_1).

I show how to use this method to identify the coefficients in the Errors-in-Variables model from Equation (2.1) without additional information, the parameter in a moving-average process in a panel data with only two time periods and without restricting shocks to have equal variance, and the coefficients in a simultaneous equations model from Hausman and Taylor (1983) without restricting one of the error terms to be mean independent. I then extend the methods to identification of coefficients in a system of linear equations in which only some of the coefficients on the unobserved variables are known.

The estimator is in the class of Extremum estimators. I show that the estimator is consistent and derive its asymptotic distribution. In finite sample simulations of the Errors-in-Variables model in Equation (2.1), the estimates have small variances and are close to the values of the underlying coefficients.

This paper is organized as follows. Section 2.2 proves identification in the Errors-in-Variables model. Section 2.3 proves identification in a moving-average process of order 1. Section 2.4 proves identification in a simultaneous equations model. Section 2.5 presents identification in the general setup. Section 2.6 presents the asymptotic results. Section 2.7 presents Monte Carlo simulations. Section 2.8 concludes. Appendix A contains the

[1] Consistent with Klepper and Leamer (1984) and Schennach and Hu (2007), the assumption fails when the unobserved variables are jointly normal.

identification proofs and Appendix B contains proofs of the asymptotic results.

2.2 Errors-in-Variables Model

In this section I identify the coefficients in the Errors-in-Variables model

$$\begin{aligned} Y &= \beta_0 + \beta_1 X_1^* + \ldots + \beta_M X_M^* + \varepsilon \\ X_m &= X_m^* + U_m \qquad\qquad m = 1, \ldots, M \end{aligned}$$

where $(Y, X_1, \ldots, X_M)$ is an observed random vector, $(X_1^*, \ldots, X_M^*, U_1, \ldots, U_M, \varepsilon)$ is an unobserved mutually independent random vector, and $(\beta_0, \ldots, \beta_M)$ are unknown nonzero coefficients.

Assumption 4. *There exists $\mathcal{U} \subseteq \mathbb{R}$ with nonzero Lebesgue measure such that for all $u \in \mathcal{U}$ and all $b \neq \beta_m$*

$$\varphi_m''(bu) \neq \varphi_m''(\beta_m u)$$

where

$$\begin{aligned} \varphi_m''(u) &= \frac{\partial^2 \ln E\left[\exp\left(iuX_m^*\right)\right]}{\partial u^2} \\ &= \left(\frac{E\left[X_m^* \exp\left(iuX_m^*\right)\right]}{E\left[\exp\left(iuX_m^*\right)\right]}\right)^2 - \frac{E\left[\left(X_m^*\right)^2 \exp\left(iuX_m^*\right)\right]}{E\left[\exp\left(iuX_m^*\right)\right]} \end{aligned}$$

is the second derivative of the log CF of X_m^*.

Theorem 6. *If $\varphi_m''(\beta_m u) < \infty$ for all $u \in \mathcal{U}$ and $\beta_m \neq 0$, then β_m is identified when Assumption 1 holds and is the unique solution to*

$$\beta_m = \underset{b \in \mathbb{R}}{argmin} \int_{\mathcal{U}} \left(\left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m}\right|_{(0,\ldots,0,bu,0,\ldots,0)} - \left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m}\right|_{(u,0,\ldots,0)} \right)^2 w(u) du$$

where $w(u)$ is a weight function that satisfies $\int_{\mathcal{U}} w(u)du = 1$ and

$$\begin{aligned}\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m} &= \frac{\partial^2 \ln E\left[\exp\left(is_0 Y + is_1 X_1 + \ldots + is_M X_M\right)\right]}{\partial s_0 \partial s_m} \\ &= \frac{E\left[Ye^{is_0Y+is_1X_1+\ldots+is_MX_M}\right]E\left[X_m e^{is_0Y+is_1X_1+\ldots+is_MX_M}\right]}{\left(E\left[e^{is_0Y+is_1X_1+\ldots+is_MX_M}\right]\right)^2} - \frac{E\left[YX_m e^{is_0Y+is_1X_1+\ldots+is_MX_M}\right]}{E\left[e^{is_0Y+is_1X_1+\ldots+is_MX_M}\right]}\end{aligned}$$

is the second-order partial derivative of the log CF of $(Y, X_1, \ldots, X_M)$ with respect to s_0 and s_m.

The main insight in this paper is that for all $u \in \mathbb{R}$

$$\begin{aligned}\frac{\partial^2 \ln E\left[\exp\left(iu\beta_m X_m^*\right)\right]}{\partial u^2} &= \left.\frac{\partial^2 \ln E\left[\exp\left(is_0 Y + is_m X_m\right)\right]}{\partial s_0 \partial s_m}\right|_{(s_0,s_m)=(0,\beta_m u)} \\ &= \left.\frac{\partial^2 \ln E\left[\exp\left(is_0 Y + is_m X_m\right)\right]}{\partial s_0 \partial s_m}\right|_{(s_0,s_m)=(u,0)} \qquad (2.2)\end{aligned}$$

This has two important implications: First, $\partial^2 \ln E\left[\exp\left(iu\beta_m X_m^*\right)\right]/\partial u^2$ is expressed in terms of observables. Second, $\partial^2 \ln E\left[\exp\left(is_0 Y + is_m X_m\right)\right]/\partial s_0 \partial s_m$ is the same when evaluated in the two directions: (1) $(s_0, s_m) = (0, \beta_m u)$ and (2) $(s_0, s_m) = (u, 0)$.

Remark 11. *If $\varphi_m''(u) = a$ for all $u \in \mathbb{R}$ then Assumption 4 fails (and Equation (2.2) equals a constant) because $\varphi_m''(bu) = \varphi_m''(\beta_m u)$ for all $b \in \mathbb{R}$.*

$$\varphi_m''(u) = a \;\Rightarrow\; E\left[\exp\left(iuX_m\right)\right] = \exp(au^2 + bu + c)$$

Let $a = -\sigma^2/2$, $b = i\mu$ and $c = 0$, then $E\left[\exp\left(iuX_m\right)\right] = \exp(i\mu u - \sigma^2 u^2/2)$ is the CF of a Normal distribution with mean μ and variance σ^2. Let $a = 0$, $b = i\mu$, and $c = 0$, then $E\left[\exp\left(iuX_m\right)\right] = \exp(i\mu u)$ is the CF of a Degenerate distribution with mass at μ.

This is consistent with Klepper and Leamer (1984) and Schennach and Hu (2007) who show that coefficients are not identified when unobservables are jointly normal.

While Assumption 4 fails when X_m^* is normal or has a point mass, it is satisfied, for *example, when X_m^* is Gamma(5,1), Uniform(0,1) or Laplace(0,1) (see Figure 2.1).*[2]

Assumption 8 in the Estimation and Asymptotics section is an alternative to Assumption 4 that can be checked in the data.

Remark 12. *The unobserved covariates X_m^* can be identified using Bonhomme and Robin (2010) or Ben-Moshe (2012a).*

Remark 13. *Let $M = 1$ and relabel the variables so that the model is*

$$W_1 = \beta W^* + U_1$$

$$W_2 = W^* + U_2$$

which is a measurement error model with repeated measurements without the assumption that β is known.

2.3 Moving-Average Process of Order 1

In this section I identify the parameter θ in the moving-average model

$$Y_1 = \varepsilon_1 - \theta\varepsilon_0$$

$$Y_2 = \varepsilon_2 - \theta\varepsilon_1$$

where (Y_1, Y_2) is an observed random vector, ε_0, ε_1, and ε_2 are unobserved mutually independent random variables, and θ is an unknown nonzero coefficient.[3]

[2] When X_m^* is Laplace(0,1) then Assumption 4 is modified to $\varphi_m'''(bu) \neq \varphi_m'''(\beta_m u)$ and Theorem 6 minimizes a third-order partial derivative (see Section 2.5 for details).

[3] A common way to identify θ is by the system of second-order moments

$$\begin{aligned} E[Y_1^2] &= E[\varepsilon_1^2] + \theta^2 E[\varepsilon_0^2] \\ E[Y_1 Y_2] &= -\theta E[\varepsilon_1^2] \\ E[Y_2^2] &= E[\varepsilon_2^2] + \theta^2 E[\varepsilon_1^2] \end{aligned}$$

which does not work without an additional assumption about the variances of the unobserved variables and / or $T > 2$.

Assumption 5. *There exists $\mathcal{U} \subseteq \mathbb{R}$ with nonzero Lebesgue measure such that for all $u \in \mathcal{U}$ and all $b \neq \theta$*

$$\varphi''_{\varepsilon_1}(bu) \neq \varphi''_{\varepsilon_1}(\theta u)$$

where

$$\begin{aligned}\varphi''_{\varepsilon_1}(u) &= \frac{\partial^2 \ln E\left[\exp\left(iu\varepsilon_1\right)\right]}{\partial u^2} \\ &= \left(\frac{E\left[\varepsilon_1 \exp\left(iu\varepsilon_1\right)\right]}{E\left[\exp\left(iu\varepsilon_1\right)\right]}\right)^2 - \frac{E\left[\varepsilon_1{}^2 \exp\left(iu\varepsilon_1\right)\right]}{E\left[\exp\left(iu\varepsilon_1\right)\right]}\end{aligned}$$

is the second derivative of the log CF of ε_1.

Theorem 7. *If $\varphi''_{\varepsilon_1}(\theta u) < \infty$ for all $u \in \mathcal{U}$ and $\theta \neq 0$, then θ is identified when Assumption 2 holds and is the unique solution to*

$$\theta = \underset{b \in \mathbb{R}}{argmin} \int_{\mathcal{U}} \left(\left.\frac{\partial^2 \varphi_{Y_1,Y_2}\left(s_1, s_2\right)}{\partial s_1 \partial s_2}\right|_{(bu,0)} - \left.\frac{\partial^2 \varphi_{Y_1,Y_2}\left(s_1, s_2\right)}{\partial s_1 \partial s_2}\right|_{(0,u)} \right)^2 w(u)du$$

where $w(u)$ is a weight function that satisfies $\int_{\mathcal{U}} w(u)du = 1$ and

$$\begin{aligned}\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2} &= \frac{\partial^2 \ln E\left[\exp\left(is_1Y_1 + is_2Y_2\right)\right]}{\partial s_1 \partial s_2} \\ &= \frac{E\left[Y_1 e^{is_1Y_1 + is_2Y_2}\right] E\left[Y_2 e^{is_1Y_1 + is_2Y_2}\right]}{\left(E\left[e^{is_1Y_1 + is_2Y_2}\right]\right)^2} - \frac{E\left[Y_1 Y_2 e^{is_1Y_1 + is_2Y_2}\right]}{E\left[e^{is_1Y_1 + is_2Y_2}\right]}\end{aligned}$$

is the second-order partial derivative of the log CF of (Y_1, Y_2) with respect to s_1 and s_2.

Remark 14. *The distributions of ε_1 and ε_2 can be estimated using Bonhomme and Robin (2010) or Ben-Moshe (2012a).*

Remark 15. *The techniques can also be applied to a times-series with the additional assumption $\varepsilon_t \stackrel{d}{=} \varepsilon_{t-2}$.*

Remark 16. *The same techniques can be used to identify* γ_m *and* θ_m *in a moving-average process of order (p,q)*

$$Y_t = c + \varepsilon_t + \sum_{m=1}^{p} \gamma_m Y_{t-m} + \sum_{m=1}^{q} \theta_m \varepsilon_{t-m}$$

See Ben-Moshe (2012b) for identification in an Autoregressive Process of order 1.

2.4 Simultaneous Equations Model

Consider the simultaneous equations model in Hausman and Taylor (1983)

$$Y_1 = \delta_1 Y_2 + \beta_1 X + \varepsilon_1$$
$$Y_2 = \delta_2 Y_1 + \varepsilon_2$$

where (Y_1, Y_2, X) is an observed random vector and ε_0 and ε_1 are unobserved random variables. Hausman and Taylor (1983) identify the coefficients δ_1, δ_2, and β_1 under the assumptions $E[X\varepsilon_1] = 0$, $E[X\varepsilon_2] = 0$, and $E[\varepsilon_1\varepsilon_2] = 0$. I allow ε_1 and X to be arbitrarily dependent, I assume $E[X\varepsilon_2] = 0$, and I assume ε_1 and ε_2 are mutually independent conditional on the scalar X.

Assumption 6. *There exists* $\mathcal{U} \subseteq \mathbb{R}$ *with nonzero Lebesgue measure such that for all* $u \in \mathcal{U}$ *and all* $b \neq \delta_1$

$$\varphi''_{\varepsilon_2}\left(\frac{bu}{1-\delta_1\delta_2}\right) \neq \varphi''_{\varepsilon_2}\left(\frac{\delta_1 u}{1-\delta_1\delta_2}\right)$$

where

$$\begin{aligned}\varphi''_{\varepsilon_2}(u) &= \frac{\partial^2 \ln E[\exp(iu\varepsilon_2)]}{\partial u^2} \\ &= \left(\frac{E[\varepsilon_2 \exp(iu\varepsilon_2)]}{E[\exp(iu\varepsilon_2)]}\right)^2 - \frac{E\left[(\varepsilon_2)^2 \exp(iu\varepsilon_2)\right]}{E[\exp(iu\varepsilon_2)]}\end{aligned}$$

is the second derivative of the log CF of ε_2.

Theorem 8. *If $E[XY_1] \neq 0$, then δ_2 is identified. Furthermore, if $\varphi_2''(\theta bu) < \infty$ for all $u \in \mathcal{U}$, $\delta_1\delta_2 \neq 1$, and $\delta_1 \neq 0$, then δ_1 is identified when Assumption 2 holds and is the unique solution to*

$$\theta = \underset{b\in\mathbb{R}}{argmin} \int_{\mathcal{U}} \Bigg[\left(\delta_2 \cdot \left. \frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2} \right|_{(u,0)} - \left. \frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1 \partial s_2} \right|_{(u,0)} \right) - \left(\delta_2 \cdot \left. \frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2} \right|_{(0,bu)} - \left. \frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1 \partial s_2} \right|_{(0,bu)} \right) \Bigg]^2 w(u)du$$

where $w(u)$ is a weight function that satisfies $\int_{\mathcal{U}} w(u)du = 1$ and

$$\begin{aligned}
\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1,s_2)}{\partial s_1 \partial s_2} &= \frac{\partial^2 \ln E\left[\exp\left(is_1Y_1 + is_2Y_2\right)\right]}{\partial s_1 \partial s_2} \\
&= \frac{E\left[Y_1 e^{is_1Y_1+is_2Y_2}\right] E\left[Y_2 e^{is_1Y_1+is_2Y_2}\right]}{\left(E\left[e^{is_1Y_1+is_2Y_2}\right]\right)^2} - \frac{E\left[Y_1Y_2 e^{is_1Y_1+is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]} \\
\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1,s_2)}{\partial s_1^2} &= \frac{\partial^2 \ln E\left[\exp\left(is_1Y_1 + is_2Y_2\right)\right]}{\partial s_1^2} \\
&= \left(\frac{E\left[Y_1 e^{is_1Y_1+is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]} \right)^2 - \frac{E\left[Y_1^2 e^{is_1Y_1+is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]}
\end{aligned}$$

is the second-order partial derivative of the log CF of (Y_1, Y_2) with respect to s_1 and s_2. Furthermore, if $E[\varepsilon_1] = 0$ and $E[X] \neq 0$, then β_1 is identified.

Remark 17. *Identification of δ_1 and δ_2 is still possible when β_1 is a random coefficient.*

2.5 Identification in the General Setup

Let $U_m \in \mathbb{R}$, $m = 1, \ldots, M$ be unobserved mutually independent random variables, let A be a $T_A \times M$ matrix of nonzero known coefficients, let B be a $T_B \times M$ matrix of unknown nonzero coefficients, and consider the observed vector $Y \in \mathbb{R}^{T_A+T_B}$ such that

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_{T_A} \\ Y_{T_A+1} \\ \vdots \\ Y_{T_A+T_B} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{T_A1} & \dots & a_{T_AM} \\ b_{11} & \dots & b_{1M} \\ \vdots & \ddots & \vdots \\ b_{T_B1} & \dots & b_{T_BM} \end{pmatrix} \begin{pmatrix} U_1 \\ \vdots \\ U_M \end{pmatrix}$$

which can be represented as $Y = \begin{pmatrix} A \\ B \end{pmatrix} U$.[4]

Define the matrix A^D by

$$A^D = \begin{pmatrix} \prod_{t=1}^{T_A} a_{t1}^{\alpha_t^1} & \dots & \prod_{t=1}^{T_A} a_{tM}^{\alpha_t^1} \\ \vdots & \ddots & \vdots \\ \prod_{t=1}^{T_A} a_{t1}^{\alpha_t^R} & \dots & \prod_{t=1}^{T_A} a_{tM}^{\alpha_t^R} \end{pmatrix}$$

where D is a nonnegative integer and $(\alpha_1^r, \dots, \alpha_{T_A}^r)$ is a vector of nonnegative integers such that $D = \alpha_1^r + \dots + \alpha_{T_A}^r$ for $r = 1, \dots, R$ and $(\alpha_1^r, \dots, \alpha_{T_A}^r) \neq (\alpha_1^{r'}, \dots, \alpha_{T_A}^{r'})$ for $r \neq r'$. The matrix A^D contains all products of entries in the same column of A with the restriction that the sum of the exponents is exactly equal to D. The matrix A^D has dimension $R = \begin{pmatrix} D + T_A - 1 \\ D \end{pmatrix} \times M$.

Assumption 7. *There exists a positive integer D and a subset $\mathcal{U} \subset \mathbb{R}$ of nonzero Lebesgue measure such that*

i. $Rank(A^D) = M$

[4]The assumptions that entries are nonzero and that known coefficients can be separated from unknown coefficients, into matrices A and B respectively, are done for clarity. The proof is similar if for every b_{tm} that is unknown, and is to be identified, there is at least one coefficient in the m^{th} column that is known and nonzero. The proof fails when an a coefficient is unknown **and** equal to 0.

ii. For all $u \in \mathcal{U}$ and all $b \neq b_{t^*m}$,

$$\varphi_m^{D+T_B}(bu) \neq \varphi_m^{D+T_B}(b_{t^*m}u)$$

where $\varphi_m^j(u) = \partial^j E\left[\exp(iU_m u)\right]/\partial u^j$ is the j^{th} derivative of φ_m.[5]

Theorem 9. *If $\int_{\mathbb{R}} \left(\varphi_m^{D+T_B}(b_{t^*m}u)\right)^2 w(u)du < \infty$ and $b_{t^*m} \neq 0$, then the unknown coefficient b_{t^*m} is identified when Assumption 7 holds. The unknown coefficient satisfies*

$$b_{t^*m} = \underset{b\in\mathbb{R}}{argmin} \int_{\mathcal{U}} \left(\sum_{r=1}^{R} a_{mr}^{D+} \left[\left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(bu\vec{s}_A^m,\vec{0})} - \left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(\vec{0},u\vec{e}_m)} \right] \right)^2 w(u)du$$

where $\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s}) / \prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}$ is a partial derivative of $\varphi_{\vec{Y}}(\vec{s}) = \ln E[\exp(i\vec{Y}'\vec{s})]$, $\{a_{mr}^{D+}\}_{m,r}$ are the entries in $(A^D)^+$, the Moore-Penrose pseudoinverse of A^D, $\vec{e}_m = (0,\ldots,1,0,\ldots,0)$ with 1 in the m^{th} coordinate, and $w(u)$ is a weight function that satisfies $\int_{\mathcal{U}} w(u)du = 1$.[6]

The proof sets up a system of equations of all $D + T_B$-order partial derivatives of $\ln E[\exp(i\vec{Y}'\vec{s})]$. In parametric settings this is analogous to setting up a system of equations of all $D + T_B$-order moments (i.e. all moments of the form $E\left[\prod_{t=1}^{T_A} Y_t^{\alpha_t^r} \prod_{t=1}^{T_B} Y_{T_A+t}\right]$ where $\sum_{t=1}^{T_A} \alpha_t^r = D$). By Assumption 7i this system can be inverted to solve for $\varphi_m^{D+T_B}$.

[5]Identification is also possible under the weaker condition: For some nonzero $\bar{u} \in \mathbb{R}$ and all $b \neq b_{t^*m}$

$$\varphi_m^{D+T_B}(b\bar{u}) \neq \varphi_m^{D+T_B}(b_{t^*m}\bar{u})$$

but for estimation this is harder to use.

[6]Instead of the L_2 norm in Theorem 3 other measures of distance can be used.

This implies that

$$\varphi_m^{D+T_B}(\cdot) = \text{linear combination of observed partial derivatives of } \ln E[\exp(i\vec{Y}'\vec{s})]$$

Two different choices of directions: (1) $(\vec{s}_A, \vec{s}_B) = (bu\vec{s}_A^m, \vec{0})$ and (2) $(\vec{s}_A, \vec{s}_B) = (\vec{0}, u\vec{e}_m)$ correspond to different choices of linear combinations of $Y_1, \ldots, Y_{T_A+T_B}$. By Assumption 7ii

$$\begin{aligned}\prod_{t=1}^{T_B} b_{tm}\varphi^{D+T_B}(b_{t^*m}u) &= \left.\frac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^r}\prod_{t=1}^{T_B}\partial s_{T_A+t}}\right|_{(\vec{s}_A,\vec{s}_B)=(bu\vec{s}_A^m,\vec{0})} \\ &= \left.\frac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^r}\prod_{t=1}^{T_B}\partial s_{T_A+t}}\right|_{(\vec{s}_A,\vec{s}_B)=(\vec{0},u\vec{e}_m)}\end{aligned}$$

for all $u \in \mathbb{R}$ if and only if $b = b_{t^*m}$.

Remark 18. *If (1) $\varphi_m^{D+T_B}(u) = a$ for $a \in \mathbb{R}$ then Assumption 7ii fails for all $b \in \mathbb{R}$ and if (2) $\varphi_m^{D+T_B}(u) = \varphi_m^{D+T_B}(au)$ for $a \in \mathbb{R}$ then Assumption 7ii fails for all $b = a^K b_{t^*m}$ where K is an integer.*

Remark 19. *Assumption 7ii can be modified as follows: Let $D = 1, 2, \ldots$ and assume that for $u \in \mathcal{U}^D \subset \mathbb{R}$*

$$\varphi_m^{D+T_B}(bu) = \varphi_m^{D+T_B}(b_{t^*m}u)$$

if only if $b \in \mathcal{B}^D$. Then $b_{t^*m} \in \cap_D \mathcal{B}^D$. Assumption 8 in the next section can be used to check which of these conditions holds in the data and once this is established different D's can be used simultaneously to make estimators more robust, test the validity of an estimator, or tighten a partially identified set.

Remark 20. *Theorem 9 can be modified to allow for subsets of unobservables to be statistically dependent. This somewhat complicates the proof because dependent unobservables cannot be separated into different CFs that are added together. Ben-Moshe (2012a) solves* *this problem by keeping dependent unobservables in a single multidimensional CF and including another rank condition on the matrices of coefficients of dependent unobservables. A similar approach is possible here.*

2.6 Estimation and Asymptotics

In this section I show that an estimator of β_m in the Errors-in-Variables model considered in Section 2.2 is consistent and asymptotically normal. Deriving the asymptotic properties of estimators of coefficients in the general setup in Section 2.5 is similar but more tedious. I also show that Assumption 4 can be checked using the data.

Let $\{Y_n, X_{n1}, \ldots, X_{nM}\}_{n=1}^N$ denote independent identically distributed observations of the random vector $(Y, X_1, \ldots, X_M) \in \mathbb{R}^{M+1}$ and let $\beta_m \in \mathcal{B} \subset \mathbb{R}$ denote the parameter of interest. Let

$$\widehat{Q}_N(b) = \int_{\mathcal{U}} \left[\left(\frac{E_N\left[Y e^{ibuX_m}\right] E_N\left[X_m e^{ibuX_m}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2} - \frac{E_N\left[Y X_m e^{ibuX_m}\right]}{E_N\left[e^{ibuX_m}\right]} \right) - \left(\frac{E_N\left[Y e^{iuY}\right] E_N\left[X_m e^{iuY}\right]}{\left(E_N\left[e^{iuY}\right]\right)^2} - \frac{E_N\left[Y X_m e^{iuY}\right]}{E_N\left[e^{iuY}\right]} \right) \right]^2 w(u) du$$

where $w(u)$ is a positive bounded weight function that satisfies $\int_{\mathcal{U}} w(u)du = 1$, $\mathcal{U}$ is compact, and

$$E_N\left[Y^\alpha X_m^\gamma e^{is_0 Y + is_m X_m}\right] = \frac{1}{N}\sum_{n=1}^N Y_n^\alpha X_{nm}^\gamma e^{is_0 Y_n + is_m X_{nm}} \qquad \alpha, \gamma \in \{0, 1, 2, \ldots\}$$

is the sample analog of the population quantity $E\left[Y^\alpha X_m^\gamma e^{is_0 Y + is_m X_m}\right]$.

The Extremum estimator I consider is defined as

$$\widehat{\beta}_m = \operatorname*{argmin}_{b \in \mathcal{B}} \widehat{Q}_N(b)$$

Its consistency and asymptotic normality are proved by checking the conditions listed by Newey and McFadden (1994):

Condition 1. ***(Consistency)***

(i) $Q_0(b)$ is uniquely minimized at $b = \beta_m$ where

$$Q_0(b) = \int_{\mathcal{U}} \left[\left(\frac{E\left[Ye^{ibuX_m}\right] E\left[X_m e^{ibuX_m}\right]}{\left(E\left[e^{ibuX_m}\right]\right)^2} - \frac{E\left[YX_m e^{ibuX_m}\right]}{E\left[e^{ibuX_m}\right]} \right) - \left(\frac{E\left[Ye^{iuY}\right] E\left[X_m e^{iuY}\right]}{\left(E\left[e^{iuY}\right]\right)^2} - \frac{E\left[YX_m e^{iuY}\right]}{E\left[e^{iuY}\right]} \right) \right]^2 w(u)du$$

(ii) $\beta_m \in \mathcal{B}$ where $\mathcal{B} \subset \mathbb{R}$ is a compact set

(iii) $Q_0(b)$ is continuous

(iv) $Q_N(b)$ converges uniformly in probability to $Q_0(b)$

Theorem 10. ***(Consistency)*** *Assume $E[Y^2] < \infty$, $E[X_m^2] < \infty$, $E[(YX_m)^2] < \infty$, $\int_{\mathcal{U}} \left|E\left[e^{iuY}\right]\right|^{-5} w(u)du < \infty$, $\int_{\mathcal{U}} \left|E\left[e^{ibuX_m}\right]\right|^{-5} w(u)du < \infty$ for all $b \in \mathcal{B}$, Assumption 4 holds, and $(\mathcal{U}, \mathcal{B}) \subset \mathbb{R}^2$ is compact, then $\widehat{\beta}_m \xrightarrow{p} \beta_m$.*

Assumption 4 is assumed to hold. Then by Theorem 6 $Q_0(b) = 0$ if and only if $b = \beta_m$. Hence, condition 1(i) is satisfied. Condition 1(ii) is assumed to hold. Condition 1(iii) is satisfied because of the bounds on the moments so that $Q_0(b) < \infty$ and continuity is checked. Condition 1(iv) is shown to hold in the Appendix by linearization through a Taylor series expansion.

Condition 2. ***(Asymptotic Normality)*** *Suppose $\widehat{\beta}_m \xrightarrow{p} \beta_m$ and*

(i) β_m is an interior point of $\mathcal{B}$

(ii) $\widehat{Q}_N(b)$ is twice continuously differentiable in a neighborhood of β_m

(iii) $\sqrt{N}Q'_N(\beta_m) \xrightarrow{d} N(0, \Omega(\beta_m))$

(iv) $H_n(b) := Q''_N(b)$ *converges uniformly in probability to* $H_0(b)$ *and* $H_0(\beta_m)$ *is nonsingular*

Theorem 11. ***(Asymptotic Normality)*** *Assume* $E[Y^2] < \infty$, $E[X_m^6] < \infty$, $E[(YX_m^3)^2] < \infty$, $\int_{\mathcal{U}} u|E[e^{iuY}]|^{-4}|E[e^{i\beta_m uX_m}]|^{-3}w(u)du < \infty$, $\int_{\mathcal{U}} u|E[e^{i\beta_m uX_m}]|^{-7}w(u)du < \infty$, $\int_{\mathcal{U}} u^2|E[e^{iuY}]|^{-2}|E[e^{ibuX_m}]|^{-4}w(u)du < \infty$, $\int_{\mathcal{U}} u^2|E[e^{ibuX_m}]|^{-6}w(u)du < \infty$ *for all* $b \in \mathcal{B}$, *Assumption 4 holds, and* $(\mathcal{U}, \mathcal{B}) \subset \mathbb{R}^2$ *is compact, then* $\sqrt{N}(\widehat{\beta}_m - \beta_m) \xrightarrow{d} N(0, (H_0(\beta_m))^{-2}\Omega(\beta_m))$ *where*

$$\Omega(\beta_m) = O\left(\int_{\mathcal{U}}\int_{\mathcal{U}} uv\left[Cov\left(Ye^{i\beta_m uX_m}, Ye^{i\beta_m vX_m}\right) + \ldots + Cov\left(e^{iuY}, e^{ivY}\right)\right]w(u)w(v)dudv\right)$$

and[7]

$$H_0(\beta_m) := -2\int_{\mathcal{U}} u^2\left(\frac{2E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} + \frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right.$$
$$\left. -\frac{2E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right)^2 w(u)du$$

$\widehat{\beta}_m \xrightarrow{p} \beta_m$ because the conditions for Theorem 10 hold. Condition 2(i) is assumed to hold. Condition 2(ii) is satisfied because of the bounds on the moments so that $Q''_0(b) < \infty$ and continuity is checked. Condition 2(iii) is shown to hold in the Appendix by linearization through a Taylor series expansion. The linear terms satisfy the central limit theorem while higher order terms are negligible. Condition 2(iv) is proved in a similar way to condition 1(iv).

This estimation procedure only works as long as Assumption 4 holds. Assumption 4 places a condition on an unobserved variable so consider instead the following alternative assumption whose validity can be checked in the data.

Assumption 8. *There exist compact sets* $\widetilde{\mathcal{U}} \subset \mathbb{R}$ *and* $\mathcal{B} \subset \mathbb{R}$ *such that for all* $u \in \widetilde{\mathcal{U}}$ *and*

[7] $Z_N = O(a_N)$ is Big-O notation and means that there exists $C > 0$ such that $Z_N \leq Ca_N$.

$b \in \mathcal{B}$

$$\left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m}\right|_{(u,0,\ldots,0)} \neq \left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m}\right|_{(bu,0,\ldots,0)}$$

Assumption 8 checks that the function $\partial^2 \varphi_{Y,\vec{X}}(\vec{s})/\partial s_0 \partial s_m$ is not constant or log periodic. Assumption 8 implies Assumption 4 as follows:

$$\begin{aligned}
& \left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m}\right|_{(u,0,\ldots,0)} \neq \left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m}\right|_{(bu,0,\ldots,0)} && \forall u \in \widetilde{\mathcal{U}}, b \in \mathcal{B} \\
\Rightarrow \quad & \beta_m \varphi_m''(\beta_m u) \neq \beta_m \varphi_m''(\beta_m b u) && \forall u \in \widetilde{\mathcal{U}}, b \in \mathcal{B} \\
\Rightarrow \quad & \beta_m \varphi_m''(\beta_m u) \neq \beta_m \varphi_m''(b u) && \forall u \in \mathcal{U}, b \in \mathcal{B}
\end{aligned}$$

where the first "$\Rightarrow$" follows from Equation (2.2) and the last "$\Rightarrow$" by letting $\beta_m \widetilde{\mathcal{U}} = \mathcal{U}$.

2.7 Monte Carlo Simulations: Errors-in-Variables

This section presents a Monte Carlo study on the finite sample properties of estimators of β_1 in the Errors-in-Variables model

$$\begin{aligned}
Y &= \beta_0 + \beta_1 X_1^* + \beta_2 X_2^* + \beta_3 X_3^* + \varepsilon \\
X_m &= X_m^* + U_m && m = 1,2,3
\end{aligned}$$

where (Y, X_1, X_2, X_3) is observed, $(X_1^*, X_2^*, X_3^*, U_1, U_2, U_3, \varepsilon)$ is an unobserved mutually independent random vector, and $(\beta_0, \beta_1, \beta_2, \beta_3)$ are unknown coefficients. The random variables ε and U_m, $m = 1,2,3$ are i.i.d $N(1,1)$.

The data is generated using the following four configurations

Experiment	$\left(f_{X_1^*}, f_{X_2^*}, f_{X_3^*}\right)$	$(\beta_0, \beta_1, \beta_2, \beta_3)$
i	χ^2_2, Unif(0,1), Unif(0,1)	$(3, 2, 1, -1)$
ii	exp(1), Unif(0,1), Norm(1,1)	$(3, 2, -1, -1)$
iii	Gamma(5,1), exp(1), Poiss(1)	$(3, -2, 1, 1)$
iv	Gamma(5,1), Norm(1,1), Norm(1,1)	$(3, -2, -1, 1)$

I estimate $\hat{\beta}_1$ as the solution to

$$\widehat{\beta}_1 = \underset{b\in[-4,4]}{\text{argmin}} \int_{[-0.3,0.3]} \left[\left(\frac{E\left[Ye^{ibuX_1}\right] E\left[X_m e^{ibuX_1}\right]}{\left(E\left[e^{ibuX_1}\right]\right)^2} - \frac{E\left[YX_m e^{ibuX_1}\right]}{E\left[e^{ibuX_1}\right]} \right) - \left(\frac{E\left[Ye^{iuY}\right] E\left[X_1 e^{iuY}\right]}{\left(E\left[e^{iuY}\right]\right)^2} - \frac{E\left[YX_1 e^{iuY}\right]}{E\left[e^{iuY}\right]} \right) \right]^2 w(u)du$$

I generate 100 simulations of sample size $N = 100$, $N = 1,000$ and $N = 10,000$. The x-axis is divided into 100 equidistant grid points. The results are summarized in Tables 2.1, 2.2, and 2.3. The estimates of $\widehat{\beta}_1$ are close to β_1 in all the experiments.

Figure 2.2 shows that Assumption 8 is satisfied by plotting $\partial^2\varphi_{Y,\vec{X}}(\vec{s})/\partial s_0 \partial s_m$, and $\beta_1\varphi''_{X_1^*}(\beta_1 u)$ for a Gamma(5,1) distribution using the configuration in Experiment iv with $N = 100$.

2.8 Conclusion

I minimize the distance between partial derivatives of log CFs in two different directions to identify the coefficients of the matrix B in the system of linear equations

$$\vec{Y} = \begin{pmatrix} A \\ B \end{pmatrix} \vec{U}$$

where $\vec{Y} \in \mathbb{R}^{T_A+T_B}$ is an observed random vector, $\vec{U} \in \mathbb{R}^M$ is an unobserved random vector, A is a $T_A \times M$ matrix of known coefficients, and B is a $T_B \times M$ matrix of unknown coefficients.

I show how to use the identification strategy in three models:

i. Errors-in-Variables model:

$$\begin{aligned} Y &= \beta_0 + \beta_1 X_1^* + \ldots + \beta_M X_M^* + \varepsilon \\ X_m &= X_m^* + U_m \qquad\qquad m = 1, \ldots, M \end{aligned}$$

where $(Y, X_1, \ldots, X_M)$ is an observed random vector and $(X_1^*, \ldots, X_M^*, U_1, \ldots, U_M, \varepsilon)$ is an unobserved mutually independent random vector. I identify $(\beta_0, \ldots, \beta_M)$ without any additional information.

ii. Moving-average process of order 1:

$$\begin{aligned} Y_1 &= \varepsilon_1 - \theta\varepsilon_0 \\ Y_2 &= \varepsilon_2 - \theta\varepsilon_1 \end{aligned}$$

where (Y_1, Y_2) is an observed random vector and ε_0, ε_1, and ε_2 are unobserved mutually independent random variables. I identify θ without assuming that ε_0, ε_1, and ε_2 have equal variances.

iii. Simultaneous equations model:

$$\begin{aligned} Y_1 &= \delta_1 Y_2 + \beta_1 X + \varepsilon_1 \\ Y_2 &= \delta_2 Y_1 + \varepsilon_2 \end{aligned}$$

where (Y_1, Y_2, X) is an observed random vector and ε_0 and ε_1 are conditionally independent unobserved random variables. I assume $E[X\varepsilon_2] = 0$ but do not place any restriction on the dependence of ε_1 on X. I identify the coefficients δ_1, δ_2, and β_1.

2.9 Appendix A

2.9.1 Proof of Theorem 6

Let $\phi_{Y,X_1...,X_M}$ denote the CF of $(Y, X_1, \ldots, X_M)$, ϕ_{X^*m} the CF of X_m^* for $1 \leq m \leq M$, ϕ_{U_m} the CF of U_m for $1 \leq m \leq M$, and ϕ_ε the CF of ε. Then,

$$
\begin{aligned}
&\phi_{Y,X_1...,X_M}(s_0, s_1, \ldots, s_M) \\
&= E\left[\exp\left(iYs_0 + iX_1 s_1 + \ldots + iX_M s_M\right)\right] \\
&= E\left[\exp\left(i(\beta_0 + \beta_1 X_1^* + \ldots + \beta_M X_M^* + \varepsilon)s_0 + i(X_1^* + U_1)s_1 + \ldots + i(X_M^* + U_M)s_M\right)\right] \\
&= E\left[\exp i\left(\beta_0 s_0 + (\beta_1 s_0 + s_1)X_1^* + \ldots + (\beta_M s_0 + s_M)X_M^* + s_1 U_1 + \ldots + s_M U_M + s_0\varepsilon\right)\right] \\
&= \exp\left(i\beta_0 s_0\right) E\left[\exp\left(is_0\varepsilon\right)\right] \prod_{m=1}^{M} E\left[\exp\left(i\left(\beta_m s_0 + s_m\right)X_m^*\right)\right] \prod_{m=1}^{M} E\left[\exp\left(is_m U_m\right)\right] \\
&= \exp\left(i\beta_0 s_0\right) \phi_\varepsilon\left(s_0\right) \prod_{m=1}^{M} \phi_{X_m^*}\left(\beta_m s_0 + s_m\right) \prod_{m=1}^{M} \phi_{U_m}\left(s_m\right)
\end{aligned}
$$

where the second equality follows by substituting $Y = \beta_0 + \beta_1 X_1^* + \ldots + \beta_M X_M^* + \varepsilon$ and $X_m = X_m^* + U_m$ for $m = 1, \ldots, M$ and the fourth equality follows from the mutual independence of the unobserved variables.

Let $\varphi_{Y,\vec{X}}(\vec{s}) = \varphi_{Y,X_1...,X_M}(s_0, s_1, \ldots, s_M) = \ln \phi_{Y,\vec{X}}(\vec{s})$, $\varphi_m(u) = \ln \phi_{X_m^*}(u)$, $\varphi_{U_m}(u) = \ln \phi_{U_m}(u)$, and $\varphi_\varepsilon(u) = \ln \phi_\varepsilon(u)$ where $\vec{s} \in \mathbb{R}^{M+1}$ and $u \in \mathbb{R}$, then

$$
\varphi_{Y,\vec{X}}(\vec{s}) = i\beta_0 s_0 + \varphi_\varepsilon\left(s_0\right) + \sum_{m=1}^{M} \varphi_m\left(\beta_m s_0 + s_m\right) + \sum_{m=1}^{M} \varphi_{U_m}\left(s_m\right)
$$

The second-order partial derivative with respect to s_0 and s_{m^*} is

$$
\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_{m^*}} = \beta_{m^*} \varphi''_{m^*}\left(\beta_{m^*} s_0 + s_{m^*}\right) \tag{2.3}
$$

where

$$
\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_m} = \frac{E\left[Ye^{is_0Y + is_1X_1 + \ldots + is_MX_M}\right] E\left[X_m e^{is_0Y + is_1X_1 + \ldots + is_MX_M}\right]}{\left(E\left[e^{is_0Y + is_1X_1 + \ldots + is_MX_M}\right]\right)^2} - \frac{E\left[YX_m e^{is_0Y + is_1X_1 + \ldots + is_MX_M}\right]}{E\left[e^{is_0Y + is_1X_1 + \ldots + is_MX_M}\right]}
$$

$$
\varphi''_m(u) = \left(\frac{E\left[X_m^* \exp\left(iuX_m^*\right)\right]}{E\left[\exp\left(iuX_m^*\right)\right]}\right)^2 - \frac{E\left[\left(X_m^*\right)^2 \exp\left(iuX_m^*\right)\right]}{E\left[\exp\left(iuX_m^*\right)\right]}
$$

Evaluate Equation (2.3) at $(0,\ldots,0,bu,0,\ldots,0)$ and $(u,0\ldots,0)$

$$\left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_{m^*}}\right|_{(0,\ldots,0,bu,0,\ldots,0)} = \beta_{m^*}\varphi''_{m^*}(bu) \tag{2.4}$$

$$\left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_{m^*}}\right|_{(u,0,\ldots,0)} = \beta_{m^*}\varphi''_{m^*}(\beta_{m^*}u) \tag{2.5}$$

where by assumption $\varphi''_m(\beta_{m^*}) < \infty$ for all $u \in \mathcal{U}$. Define

$$\begin{aligned} R_0(b,u) :&= \left(\left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_{m^*}}\right|_{(0,\ldots,0,bu,0,\ldots,0)} - \left.\frac{\partial^2 \varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0 \partial s_{m^*}}\right|_{(u,0,\ldots,0)} \right)^2 \\ &= \beta_{m^*}^2 \left(\varphi''_{m^*}(bu) - \varphi''_{m^*}(\beta_{m^*}u)\right)^2 \end{aligned}$$

where the second equality follows by substituting in Equations (2.4) and (2.5).

Let $b = \beta_{m^*}$, then $R_0(\beta_{m^*},u) = 0$ for all $u \in \mathcal{U}$ and by Assumption 4 $R_0(b,u) > 0$ for all $b \neq \beta_{m^*}$ and all $u \in \mathcal{U}$. The coefficient β_{m^*} is identified as the unique solution to

$$\beta_{m^*} = \operatorname*{argmin}_{b\in\mathbb{R}} \int_{\mathcal{U}} R_0(b,u)w(u)du$$

Assume $E[\varepsilon] = E[U_1] = \ldots E[U_M] = 0$, then after identifying $\{\beta_m\}_{m=1}^M$

$$\beta_0 = E[Y] - \sum_{m=1}^M \beta_m E[X_m^*] = E[Y] - \sum_{m=1}^M \beta_m E[X_m]$$

2.9.2 Proof of Theorem 7

The log CF of (Y_1, Y_2) is

$$\begin{aligned} \ln E\left[\exp\left(iY_1 s_1 + iY_2 s_2\right)\right] &= \ln E\left[\exp\left(i(\varepsilon_1 - \theta\varepsilon_0)s_1 + i(\varepsilon_2 - \theta\varepsilon_1)s_2\right)\right] \\ &= \ln E\left[\exp\left(-i\theta s_1\varepsilon_0 + i\left(s_1 - \theta s_2\right)\varepsilon_1 + is_2\varepsilon_2\right)\right] \\ &= \ln E\left[\exp\left(-i\theta s_1\varepsilon_0\right)\right] + \ln E\left[\exp\left(i\left(s_1 - \theta s_2\right)\varepsilon_1\right)\right] + \ln E\left[\exp\left(is_2\varepsilon_2\right)\right] \end{aligned}$$

where the first equality follows by substituting $Y_1 = \varepsilon_1 - \theta\varepsilon_0$ and $Y_2 = \varepsilon_2 - \theta\varepsilon_1$ and the last equality follows from the mutual independence of the unobserved variables.

Let φ_{Y_1,Y_2} denote the log CF of (Y_1, Y_2) and φ_m the log CF of ε_m. Then

$$\varphi_{Y_1,Y_2}(s_1, s_2) = \varphi_0(-\theta s_1) + \varphi_1(s_1 - \theta s_2) + \varphi_2(s_2)$$

The second-order partial derivative with respect to s_1 and s_2 is

$$\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2} = -\theta \varphi_1''(s_1 - \theta s_2) \tag{2.6}$$

where

$$\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2} = \frac{E\left[Y_1 e^{is_1Y_1+is_2Y_2}\right] E\left[Y_2 e^{is_1Y_1+is_2Y_2}\right]}{\left(E\left[e^{is_1Y_1+is_2Y_2}\right]\right)^2} - \frac{E\left[Y_1Y_2 e^{is_1Y_1+is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]}$$
$$\varphi_{\varepsilon_1}''(u) = \left(\frac{E\left[\varepsilon_1 \exp(iu\varepsilon_1)\right]}{E\left[\exp(iu\varepsilon_1)\right]}\right)^2 - \frac{E\left[\varepsilon_1{}^2 \exp(iu\varepsilon_1)\right]}{E\left[\exp(iu\varepsilon_1)\right]}$$

Evaluate Equation (2.6) at $(0, bu)$ and $(u, 0)$

$$\left.\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2}\right|_{(0,bu)} = -\theta \varphi_1''(bu) \tag{2.7}$$

$$\left.\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2}\right|_{(u,0)} = -\theta \varphi_1''(\theta u) \tag{2.8}$$

where by assumption $\varphi_1''(\theta u) < \infty$ for all $u \in \mathcal{U}$. Define

$$\begin{aligned} R_0(b, u) :&= \left(\left.\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2}\right|_{(0,bu)} - \left.\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1, s_2)}{\partial s_1 \partial s_2}\right|_{(u,0)}\right)^2 \\ &= \theta^2 \left(\varphi_1''(bu) - \varphi_1''(\theta u)\right)^2 \end{aligned}$$

where the second equality follows by substituting in Equations (2.7) and (2.8).

Let $b = \theta$, then $R_0(\theta, u) = 0$ for all $u \in \mathcal{U}$ and by Assumption 5 $R_0(b, u) > 0$ for all $b \neq \theta$ and all $u \in \mathcal{U}$. The parameter θ is identified as the unique solution to

$$\theta = \underset{b \in \mathbb{R}}{\operatorname{argmin}} \int_{\mathcal{U}} R_0(b, u) w(u) du$$

2.9.3 Proof of Theorem 8

The parameter $\delta_2 = E[XY_2]/E[XY_1]$ is identified using the condition $0 = E[X\varepsilon_2] = E[XY_2 - X\delta_2 Y_1]$.

The structural system is now rewritten in its reduced form

$$Y_1 = \frac{X}{1-\delta_1\delta_2}\cdot\beta_1 + \frac{1}{1-\delta_1\delta_2}\cdot\varepsilon_1 + \frac{\delta_1}{1-\delta_1\delta_2}\cdot\varepsilon_2$$
$$Y_2 = \frac{\delta_2 X}{1-\delta_1\delta_2}\cdot\beta_1 + \frac{\delta_2}{1-\delta_1\delta_2}\cdot\varepsilon_1 + \frac{1}{1-\delta_1\delta_2}\cdot\varepsilon_2$$

Let $\theta = 1/(1-\delta_1\delta_2)$. The log CF of (Y_1, Y_2) conditional on $X = x$ is

$$\begin{aligned}
&\ln E\left[\exp\left(iY_1 s_1 + iY_2 s_2\right)|X = x\right] \\
&= \ln E\left[\exp\left(i(X\theta\beta_1 + \theta\varepsilon_1 + \delta_1\theta\varepsilon_2)s_1 + i(X\delta_2\theta\beta_1 + \delta_2\theta\varepsilon_1 + \theta\varepsilon_2)s_2\right)|X = x\right] \\
&= ix\theta(s_1 + \delta_2 s_2)\beta_1 + \ln E\left[\exp\left(i\theta(s_1 + \delta_2 s_2)\varepsilon_1 + i\theta(s_1\delta_1 + s_2)\varepsilon_2\right)|X = x\right] \\
&= ix\theta(s_1 + \delta_2 s_2)\beta_1 + \ln E\left[\exp\left(i\theta(s_1 + \delta_2 s_2)\varepsilon_1\right)|X = x\right] + \ln E\left[\exp\left(i\theta(\delta_1 s_1 + s_2)\varepsilon_2\right)|X = x\right]
\end{aligned}$$

where the first equality follows by substituting $Y_1 = X\theta\beta_1 + \theta\varepsilon_1 + \delta_1\theta\varepsilon_2$ and $Y_2 = X\delta_2\theta\beta_1 + \delta_2\theta\varepsilon_1 + \theta\varepsilon_2$ and the last equality follows from the mutual independence of the unobserved variables.

Let $\varphi_{Y_1,Y_2|X}$ denote the log CF of $(Y_1, Y_2|X = x)$ and $\varphi_{m|X}$ the log CF of $\varepsilon_m|X = x$. Then

$$\varphi_{Y_1,Y_2|X}(s_1, s_2) = ix\theta(s_1 + \delta_2 s_2)\beta_1 + \varphi_{1|X}\left(\theta s_1 + \theta\delta_2 s_2\right) + \varphi_{2|X}\left(\theta\delta_1 s_1 + \theta s_2\right)$$

where the equality follows from the independence assumptions. The second order partial derivatives are

$$\begin{pmatrix} \dfrac{\partial\varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2} \\ \dfrac{\partial\varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1\partial s_2} \\ \dfrac{\partial\varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_2^2} \end{pmatrix} = \begin{pmatrix} \theta^2 & \theta^2\delta_1^2 \\ \theta^2\delta_2 & \theta^2\delta_1 \\ \theta^2\delta_2^2 & \theta^2 \end{pmatrix} \begin{pmatrix} \varphi''_{1|X}\left(\theta s_1 + \theta\delta_2 s_2\right) \\ \varphi''_{2|X}\left(\theta\delta_1 s_1 + \theta s_2\right) \end{pmatrix}$$

Hence,[8]

$$\delta_2\cdot\frac{\partial\varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2} - \frac{\partial\varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1\partial s_2} = \theta^2\delta_1\left(\delta_1\delta_2 - 1\right)\varphi''_{2|X}\left(\theta\delta_1 s_1 + \theta s_2\right) \tag{2.9}$$

where

$$\frac{\partial^2\varphi_{Y_1,Y_2}(s_1,s_2)}{\partial s_1^2} = \left(\frac{E\left[Y_1 e^{is_1Y_1 + is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]}\right)^2 - \frac{E\left[Y_1^2 e^{is_1Y_1+is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]}$$

[8]Other identification strategies are possible.

$$\frac{\partial^2 \varphi_{Y_1,Y_2}(s_1,s_2)}{\partial s_1 \partial s_2} = \frac{E\left[Y_1 e^{is_1Y_1+is_2Y_2}\right]E\left[Y_2 e^{is_1Y_1+is_2Y_2}\right]}{\left(E\left[e^{is_1Y_1+is_2Y_2}\right]\right)^2} - \frac{E\left[Y_1Y_2 e^{is_1Y_1+is_2Y_2}\right]}{E\left[e^{is_1Y_1+is_2Y_2}\right]}$$
$$\varphi_2''(u) = \left(\frac{E\left[\varepsilon_2 \exp\left(iu\varepsilon_2\right)\right]}{E\left[\exp\left(iu\varepsilon_2\right)\right]}\right)^2 - \frac{E\left[\varepsilon_2{}^2 \exp\left(iu\varepsilon_2\right)\right]}{E\left[\exp\left(iu\varepsilon_2\right)\right]}$$

Evaluate Equation (2.9) at $(u,0)$ and $(0,du)$

$$\left(\delta_2 \cdot \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2}\right|_{(u,0)} - \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{(u,0)}\right) = \theta^2\delta_1\left(\delta_1\delta_2 - 1\right)\varphi''_{2|X}\left(\theta\delta_1 u\right) \tag{2.10}$$

$$\delta_2 \cdot \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2}\right|_{(0,bu)} - \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{(0,bu)} = \theta^2\delta_1\left(\delta_1\delta_2 - 1\right)\varphi''_{2|X}\left(\theta b u\right) \tag{2.11}$$

where by assumption $\varphi_2''(\theta bu) < \infty$ for all $u \in \mathcal{U}$.

Define

$$\begin{aligned} R(b,u) = \Bigg[& \left(\delta_2 \cdot \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2}\right|_{(u,0)} - \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{(u,0)}\right) \\ & - \left(\delta_2 \cdot \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1^2}\right|_{(0,bu)} - \left.\frac{\partial \varphi^2_{Y_1,Y_2|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{(0,bu)}\right)\Bigg]^2 \\ = {} & \theta^4\delta_1^2\left(\delta_1\delta_2 - 1\right)^2\left(\varphi''_{\varepsilon_2|X}\left(\theta\delta_1 u\right) - \varphi''_{\varepsilon_2|X}\left(\theta b u\right)\right)^2 \end{aligned}$$

where the second equality follows by substituting in Equations (2.10) and (2.11).

Let $b = \delta_1$, then $R_0(\delta_1, u) = 0$ for all $u \in \mathcal{U}$ and by Assumption 6 $R_0(b,u) > 0$ for all $b \neq \delta_1$ and all $u \in \mathcal{U}$. The parameter δ_1 is identified as the unique solution to

$$\delta_1 = \operatorname*{argmin}_{b\in\mathbb{R}} \int_{\mathcal{U}} R_0(b,u)w(u)du$$

Assume $E[\varepsilon_1] = 0$ and $E[X] \neq 0$, then $\beta_1 = 1 - \delta_1\delta_2/E[X]$.

2.9.4 Proof of Theorem 9

The CF of $(Y_1, \ldots, Y_T)$ is

$$\begin{aligned} \phi_{Y_1,\ldots,Y_T}(s_1,\ldots,s_T) &= E\left[\exp\left(iY_1s_1 + \ldots + iY_{T_A+T_B}s_{T_A+T_B}\right)\right] \\ &= E\left[\exp\left(i(a_{11}U_1 + \ldots + a_{1M}U_M)s_1 + \ldots + i(b_{T_B1}U_1 + \ldots + b_{T_BM}U_M)s_{T_A+T_B}\right)\right)\right] \end{aligned}$$

$$= E\left[\exp\left(i(a_{11}s_1 + \ldots + b_{T_B1}s_{T_A+T_B})U_1 + \ldots + i(a_{1M}s_1 + \ldots + b_{T_BM}s_{T_A+T_B})U_M\right)\right]$$

$$= \prod_{m=1}^{M} E\left[\exp\left(i(a_{1m}s_1 + \ldots + b_{T_Bm}s_{T_A+T_B})U_m\right)\right]$$

where the second equality follows by substituting $Y_t = a_{t1}U_1 + \ldots + a_{tM}U_M$ and the fourth equality follows from the mutual independence of the unobserved variables.

Let $\varphi_{\vec{Y}}(\vec{s}) = \varphi_{Y_1,\ldots,Y_T}(s_1,\ldots,s_T) = \ln\phi_{\vec{Y}}(\vec{s})$ and $\varphi_m(u) = \ln\phi_{U_m}(u) = \ln E\left[\exp\left(iuU_m\right)\right]$, $m = 1,\ldots,M$ where $\vec{s} \in \mathbb{R}^T$ and $u \in \mathbb{R}$. Then

$$\varphi_{\vec{Y}}(\vec{s}) = \sum_{m=1}^{M} \varphi_m\left(\sum_{t=1}^{T_A} a_{tm}s_t + \sum_{t=1}^{T_B} b_{tm}s_{T_A+t}\right) = \sum_{m=1}^{M} \varphi_m\left(A_m'\vec{s}_A + B_m'\vec{s}_B\right)$$

where A_m is the m^{th} column of A, B_m is the m^{th} column of B, $\vec{s}_A = (s_1,\ldots,s_{T_A})'$ and $\vec{s}_B = (s_{T_A+1},\ldots,s_{T_A+T_B})'$.

Let $(\alpha_1^r,\ldots,\alpha_{T_A}^r)$ be a multi-index T_A-tuple of nonnegative integers. The norm of the multi-index is defined by $|\alpha^r| = \alpha_1^r + \ldots + \alpha_{T_A}^r$. For all multi-indexes with $|\alpha^r| = D$ the partial derivative of $\varphi_{\vec{Y}}(\vec{s})$ with respect to $s_1^{\alpha_1^r}$, …, $s_{T_A}^{\alpha_{T_A}^r}$, s_{T_A+1}, …, $s_{T_A+T_B}$ is

$$\frac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^r}\prod_{t=1}^{T_B}\partial s_{T_A+t}} = \sum_{m=1}^{M}\prod_{t=1}^{T_A} a_{tm}^{\alpha_t^r}\left[\prod_{t=1}^{T_B} b_{tm}\varphi_m^{D+T_B}\left(A_m'\vec{s}_A + B_m'\vec{s}_B\right)\right]$$

where $\varphi_m^j(\cdot)$ is the j^{th} derivative of $\varphi_m(\cdot)$. This is represented in matrix notation by

$$\begin{pmatrix} \dfrac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^1}\prod_{t=1}^{T_B}\partial s_{T_A+t}} \\ \vdots \\ \dfrac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^R}\prod_{t=1}^{T_B}\partial s_{T_A+t}} \end{pmatrix} = \begin{pmatrix} \prod_{t=1}^{T_A} a_{t1}^{\alpha_t^1} & \ldots & \prod_{t=1}^{T_A} a_{tM}^{\alpha_t^1} \\ \vdots & \ddots & \vdots \\ \prod_{t=1}^{T_A} a_{t1}^{\alpha_t^R} & \ldots & \prod_{t=1}^{T_A} a_{tM}^{\alpha_t^R} \end{pmatrix} \begin{pmatrix} \prod_{t=1}^{T_B} b_{t1}\varphi_1^{D+T_B}\left(A_1'\vec{s}_A + B_1'\vec{s}_B\right) \\ \vdots \\ \prod_{t=1}^{T_B} b_{tM}\varphi_M^{D+T_B}\left(A_M'\vec{s}_A + B_M'\vec{s}_B\right) \end{pmatrix}$$

where $R = \begin{pmatrix} D + T_A - 1 \\ D \end{pmatrix}$.

By Assumption 7i

$$\begin{pmatrix} \prod_{t=1}^{T_B} b_{t1}\varphi_1^{D+T_B}\left(A_1'\vec{s}_A + B_1'\vec{s}_B\right) \\ \vdots \\ \prod_{t=1}^{T_B} b_{tM}\varphi_M^{D+T_B}\left(A_M'\vec{s}_A + B_M'\vec{s}_B\right) \end{pmatrix} = \begin{pmatrix} \prod_{t=1}^{T_A} a_{t1}^{\alpha_t^1} & \ldots & \prod_{t=1}^{T_A} a_{tM}^{\alpha_t^1} \\ \vdots & \ddots & \vdots \\ \prod_{t=1}^{T_A} a_{t1}^{\alpha_t^R} & \ldots & \prod_{t=1}^{T_A} a_{tM}^{\alpha_t^R} \end{pmatrix}^{+} \begin{pmatrix} \dfrac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^1}\prod_{t=1}^{T_B}\partial s_{T_A+t}} \\ \vdots \\ \dfrac{\partial\varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A}\partial s_t^{\alpha_t^R}\prod_{t=1}^{T_B}\partial s_{T_A+t}} \end{pmatrix}$$

where $\left(A^D\right)^+$ is the Moore-Penrose pseudoinverse of A^D with entries $\{a_{mr}^{D+}\}_{m,r}$.

Let $\vec{s}_A^m$ satisfy $A_m' \vec{s}_A^m = 1$. For $u \in \mathbb{R}, b \in \mathbb{R}$

$$\sum_{r=1}^{R} a_{mr}^{D+} \left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(bu\vec{s}_A^m,\vec{0})} = \prod_{t=1}^{T_B} b_{tm} \varphi_m^{D+T_B}(bu) \tag{2.12}$$

$$\sum_{r=1}^{R} a_{mr}^{D+} \left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(\vec{0},u\vec{e}_m)} = \prod_{t=1}^{T_B} b_{tm} \varphi_m^{D+T_B}(b_{t^*m} u) \tag{2.13}$$

where $\vec{e}_m = (0,\dots,0,1,0,\dots,0)$ with 1 in the m^{th} coordinate.

Define

$$Q_0(b) := \int_{\mathcal{U}} \left(\sum_{r=1}^{R} a_{mr}^{D+} \left[\left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(bu\vec{s}_A^m,\vec{0})} - \left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(\vec{0},u\vec{e}_m)} \right] \right)^2 w(u)du$$

where $w(u)$ is a weight function that satisfies $\int_{\mathcal{U}} w(u)du = 1$.

I show that $Q_0(b_{t^*m}) = 0$ and $Q_0(b) > 0$ for all $b \neq b_{t^*m}$:

$$\begin{aligned}
Q_0(b_{t^*m}) &= \int_{\mathcal{U}} \left(\sum_{r=1}^{R} a_{mr}^{D+} \left[\left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(b_{t^*m}u\vec{s}_A^m,\vec{0})} - \left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(\vec{0},u\vec{e}_m)} \right] \right)^2 w(u)du \\
&= \int_{\mathcal{U}} \left(\prod_{t=1}^{T_B} b_{tm} \varphi_m^{D+T_B}(b_{t^*m}u) - \prod_{t=1}^{T_B} b_{tm} \varphi_m^{D+T_B}(b_{t^*m}u) \right)^2 w(u)du \\
&= 0
\end{aligned}$$

where the second equality follows by substituting in Equations (2.12) and (2.13) and the last equality follows by the assumption that $\int_{\mathcal{U}} \left(\varphi_m^{D+T_B}(b_{t^*m}u)\right)^2 w(u)du < \infty$.

$$\begin{aligned}
Q_0(b) &:= \int_{\mathcal{U}} \left(\sum_{r=1}^{R} a_{mr}^{D+} \left[\left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(bu\vec{s}_A^m,\vec{0})} - \left. \frac{\partial \varphi_{\vec{Y}}^{D+T_B}(\vec{s})}{\prod_{t=1}^{T_A} \partial s_t^{\alpha_t^r} \prod_{t=1}^{T_B} \partial s_{T_A+t}} \right|_{(\vec{s}_A,\vec{s}_B)=(\vec{0},u\vec{e}_m)} \right] \right)^2 w(u)du \\
&= \int_{\mathcal{U}} \left(\prod_{t=1}^{T_B} b_{tm} \varphi_m^{D+T_B}(bu) - \prod_{t=1}^{T_B} b_{tm} \varphi_m^{D+T_B}(b_{t^*m}u) \right)^2 w(u)du
\end{aligned}$$

$$= \left(\prod_{t=1}^{T_B} b_{tm}\right)^2 \int_{\mathbb{R}} \left(\varphi_m^{D+T_B}(bu) - \varphi_m^{D+T_B}(b_{t^*m}u)\right)^2 w(u)du$$

$$> 0$$

where the second equality follows by substituting in Equations (2.12) and (2.13) and the last inequality follows by Assumption 7ii. Hence, b_{t^*m} uniquely minimizes Q_0 and is identified.

2.10 Appendix B

2.10.1 Proof of Condition 1(iv): $Q_N(b)$ Converges Uniformly in Probability to $Q_0(b)$

Lemma 2. *Let F denote the cumulative distribution function of $(Y, X_1, \ldots, X_M)$ and F_N the empirical cumulative distribution function corresponding to a sample $\{Y_n, X_{n1}, \ldots, X_{nM}\}_{n=1}^{N}$ of N independent identically distributed random draws from F. Assume $E\left[Y^{2\alpha}X_m^{2\gamma}\right] < \infty$. Let*

$$\varepsilon_N = C_{\left(M, E\left[Y^{2\alpha}X_m^{2\gamma}\right]\right)} \left(\frac{\ln N}{N}\right)^{\frac{1}{2}}$$

where $C > 0$ and $C_{\left(M, E\left[Y^{2\alpha}X_m^{2\gamma}\right]\right)} > 0$ is a constant that may depend on the arguments in the subscript. Then

$$\sup_{(s_0, s_m) \in [-S_0, S_0] \times [-S_m, S_m]} \left| E_N\left[Y^\alpha X_m^\gamma e^{is_0 Y + is_m X_m}\right] - E\left[Y^\alpha X_m^\gamma e^{is_0 Y + is_m X_m}\right]\right| < \varepsilon_N \qquad a.s.$$

when N tends to infinity.

Proof: See Lemma 1 in Ben-Moshe (2012a).

Let

$$R_0(b, u)$$

$$= \left[\left(\frac{E\left[Ye^{ibuX_m}\right]E\left[X_m e^{ibuX_m}\right]}{\left(E\left[e^{ibuX_m}\right]\right)^2} - \frac{E\left[YX_m e^{ibuX_m}\right]}{E\left[e^{ibuX_m}\right]}\right) - \left(\frac{E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]}{\left(E\left[e^{iuY}\right]\right)^2} - \frac{E\left[YX_m e^{iuY}\right]}{E\left[e^{iuY}\right]}\right)\right]^2$$

and

$$\widehat{R}_N(b,u) = \left[\left(\frac{E_N\left[Ye^{ibuX_m}\right]E_N\left[X_me^{ibuX_m}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2} - \frac{E_N\left[YX_me^{ibuX_m}\right]}{E_N\left[e^{ibuX_m}\right]}\right)\right.$$
$$\left. - \left(\frac{E_N\left[Ye^{iuY}\right]E_N\left[X_me^{iuY}\right]}{\left(E_N\left[e^{iuY}\right]\right)^2} - \frac{E_N\left[YX_me^{iuY}\right]}{E_N\left[e^{iuY}\right]}\right)\right]^2$$

Then

$$Q_0(b) = \int_{\mathcal{U}} R_0(b,u)w(u)du$$
$$\widehat{Q}_N(b) = \int_{\mathcal{U}} R_N(b,u)w(u)du$$

Expand the brackets in $\widehat{R}_N(b,u)$ and use a Taylor expansion

$$\begin{aligned}
&\widehat{R}_N(b,u) \\
&= \frac{\left(E_N\left[Ye^{ibuX_m}\right]\right)^2\left(E_N\left[X_me^{ibuX_m}\right]\right)^2}{\left(E_N\left[e^{ibuX_m}\right]\right)^4} - \frac{2E_N\left[Ye^{ibuX_m}\right]E_N\left[X_me^{ibuX_m}\right]E_N\left[YX_me^{ibuX_m}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^3} \\
&- \frac{2E_N\left[Ye^{ibuX_m}\right]E_N\left[X_me^{ibuX_m}\right]E_N\left[Ye^{iuY}\right]E_N\left[X_me^{iuY}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2\left(E_N\left[e^{iuY}\right]\right)^2} \\
&+ \frac{2E_N\left[Ye^{ibuX_m}\right]E_N\left[X_me^{ibuX_m}\right]E_N\left[YX_me^{iuY}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2E_N\left[e^{iuY}\right]} + \frac{\left(E_N\left[YX_me^{iuY}\right]\right)^2}{\left(E_N\left[e^{iuY}\right]\right)^2} \\
&- \frac{2E_N\left[Ye^{iuY}\right]E_N\left[X_me^{iuY}\right]E_N\left[YX_me^{iuY}\right]}{\left(E_N\left[e^{iuY}\right]\right)^3} + \frac{2E_N\left[YX_me^{ibuX_m}\right]E_N\left[Ye^{iuY}\right]E_N\left[X_me^{iuY}\right]}{E_N\left[e^{ibuX_m}\right]\left(E_N\left[e^{iuY}\right]\right)^2} \\
&- \frac{2E_N\left[YX_me^{ibuX_m}\right]E_N\left[YX_me^{iuY}\right]}{E_N\left[e^{ibuX_m}\right]E_N\left[e^{iuY}\right]} + \frac{\left(E_N\left[Ye^{iuY}\right]\right)^2\left(E_N\left[X_me^{iuY}\right]\right)^2}{\left(E_N\left[e^{iuY}\right]\right)^4} + \frac{\left(E_N\left[YX_me^{ibuX_m}\right]\right)^2}{\left(E_N\left[e^{ibuX_m}\right]\right)^2} \\
&= R_0(b,u) + g_0^1(b,u)\left(E_N\left[Ye^{ibuX_m}\right] - E\left[Ye^{ibuX_m}\right]\right) + g_0^2(b,u)\left(E_N\left[X_me^{ibuX_m}\right] - E\left[X_me^{ibuX_m}\right]\right) \\
&+ g_0^3(b,u)\left(E_N\left[YX_me^{ibuX_m}\right] - E\left[YX_me^{ibuX_m}\right]\right) + g_0^4(b,u)\left(E_N\left[e^{ibuX_m}\right] - E\left[e^{ibuX_m}\right]\right) \\
&+ g_0^5(b,u)\left(E_N\left[Ye^{iuY}\right] - E\left[Ye^{iuY}\right]\right) + g_0^6(b,u)\left(E_N\left[X_me^{iuY}\right] - E\left[X_me^{iuY}\right]\right) \\
&+ g_0^7(b,u)\left(E_N\left[YX_me^{iuY}\right] - E\left[YX_me^{iuY}\right]\right) + g_0^8(b,u)\left(E_N\left[e^{iuY}\right] - E\left[e^{iuY}\right]\right) \\
&+ o\left[\left|g_0^1(b,u)\left(E_N\left[Ye^{ibuX_m}\right] - E\left[Ye^{ibuX_m}\right]\right)\right| + \ldots + \left|g_0^8(b,u)\left(E_N\left[e^{iuY}\right] - E\left[e^{iuY}\right]\right)\right|\right]
\end{aligned}$$

where the second equality follows by a Taylor expansion and

$$g_0^1(b,u) =$$
$$\frac{2E\left[X_me^{ibuX_m}\right]}{\left(E\left[e^{ibuX_m}\right]\right)^4\left(E\left[e^{iuY}\right]\right)^2}\left(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right] - E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2\right.$$

$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^2(b,u)=$$

$$\frac{2E\left[Ye^{ibuX_m}\right]}{\left(E\left[e^{ibuX_m}\right]\right)^4\left(E\left[e^{iuY}\right]\right)^2}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^3(b,u)=$$

$$\frac{-2}{\left(E\left[e^{ibuX_m}\right]\right)^3\left(E\left[e^{iuY}\right]\right)^2}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^4(b,u)=$$

$$\frac{2}{\left(E\left[e^{ibuX_m}\right]\right)^5\left(E\left[e^{iuY}\right]\right)^2}\left(E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]-2E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\right)\times$$
$$\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^5(b,u)=$$

$$\frac{-2E\left[X_me^{iuY}\right]}{\left(E\left[e^{ibuX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^4}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^6(b,u)=$$

$$\frac{-2E\left[Ye^{iuY}\right]}{\left(E\left[e^{ibuX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^4}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^7(b,u)=$$

$$\frac{2}{\left(E\left[e^{ibuX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^3}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$g_0^8(b,u)=$$

$$-\frac{2}{\left(E\left[e^{ibuX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^5}\left(E\left[YX_me^{iuY}\right]E\left[e^{iuY}\right]-2E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\right)\times$$
$$\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2E\left[e^{iuY}\right]-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{ibuX_m}\right]\right)^2$$
$$-E\left[YX_me^{ibuX_m}\right]E\left[e^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2+E\left[Ye^{ibuX_m}\right]E\left[X_me^{ibuX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

Substitute $\widehat{R}_N(b,u)$ into $\sup_b \left|\widehat{Q}_N(b) - Q_0(b)\right|$

$$
\begin{aligned}
&\sup_b \left|\widehat{Q}_N(b) - Q_0(b)\right| \\
&= \sup_b \left| \int_{\mathcal{U}} \left(\widehat{R}_N(b,u) - R_0(b,u) \right) w(u) du \right| \\
&= \sup_b \left| \int_{\mathcal{U}} g_0^1(b,u) \left(E_N \left[Y e^{ibuX_m} \right] - E\left[Y e^{ibuX_m} \right] \right) + g_0^2(b,u) \left(E_N \left[X_m e^{ibuX_m} \right] - E\left[X_m e^{ibuX_m} \right] \right) \right. \\
&+ g_0^3(b,u) \left(E_N \left[Y X_m e^{ibuX_m} \right] - E\left[Y X_m e^{ibuX_m} \right] \right) + g_0^4(b,u) \left(E_N \left[e^{ibuX_m} \right] - E\left[e^{ibuX_m} \right] \right) \\
&+ g_0^5(b,u) \left(E_N \left[Y e^{iuY} \right] - E\left[Y e^{iuY} \right] \right) + g_0^6(b,u) \left(E_N \left[X_m e^{iuY} \right] - E\left[X_m e^{iuY} \right] \right) \\
&+ g_0^7(b,u) \left(E_N \left[Y X_m e^{iuY} \right] - E\left[Y X_m e^{iuY} \right] \right) + g_0^8(b,u) \left(E_N \left[e^{iuY} \right] - E\left[e^{iuY} \right] \right) \\
&\left. + o\left[\left| g_0^1(b,u) \left(E_N \left[Y e^{ibuX_m} \right] - E\left[Y e^{ibuX_m} \right] \right) \right| + \ldots + \left| g_0^8(b,u) \left(E_N \left[e^{iuY} \right] - E\left[e^{iuY} \right] \right) \right| \right] w(u) du \right| \\
&\lesssim \varepsilon_N \int_{\mathcal{U}} \left(|g_0^1(b,u)| + |g_0^2(b,u)| + |g_0^3(b,u)| + |g_0^4(b,u)| + |g_0^5(b,u)| + |g_0^6(b,u)| + |g_0^7(b,u)| + |g_0^8(b,u)| \right) w(u) du \\
&\lesssim \left(\frac{\ln N}{N} \right)^{\frac{1}{2}} \left(E[|Y|] + E[|X_m|] + E[|Y X_m|] \right) \int_{\mathcal{U}} \left(\frac{1}{\left| E\left[e^{iuY} \right] \right|^5} + \frac{1}{\left| E\left[e^{ibuX_m} \right] \right|^5} \right) w(u) du
\end{aligned}
$$

where the "$\lesssim$"s follow by Lemma 2.[9] By the assumptions $E\left[Y^2\right] < \infty$, $E\left[X_m^2\right] < \infty$, $E\left[(YX_m)^2\right] < \infty$, $\int_{\mathcal{U}} \left| E\left[e^{iuY} \right] \right|^{-5} w(u) du < \infty$, and $\int_{\mathcal{U}} \left| E\left[e^{ibuX_m} \right] \right|^{-5} w(u) du < \infty$ for all $b \in \mathcal{B}$ so $Q_N(b)$ converges uniformly to $Q_0(b)$.

2.10.2 Proof of Condition 2(iii): $\sqrt{N} Q_N'(\beta_m) \xrightarrow{d} N(0, \Omega(\beta_m))$

The derivative $Q_N'(\beta_m)$ is

$$
\begin{aligned}
&\widehat{Q}_N'(\beta_m) \\
&= 2i \int_{\mathcal{U}} u \left(\frac{E_N \left[Y e^{i\beta_m u X_m} \right] E_N \left[X_m e^{i\beta_m u X_m} \right]}{\left(E_N \left[e^{i\beta_m u X_m} \right] \right)^2} - \frac{E_N \left[Y X_m e^{i\beta_m u X_m} \right]}{E_N \left[e^{i\beta_m u X_m} \right]} \right. \\
&\qquad \left. + \frac{E_N \left[Y e^{iuY} \right] E_N \left[X_m e^{iuY} \right]}{\left(E_N \left[e^{iuY} \right] \right)^2} - \frac{E_N \left[Y X_m e^{iuY} \right]}{E_N \left[e^{iuY} \right]} \right) \\
&\quad \times \left(\frac{2 E_N \left[Y X_m e^{i\beta_m u X_m} \right] E_N \left[X_m e^{i\beta_m u X_m} \right]}{\left(E_N \left[e^{i\beta_m u X_m} \right] \right)^2} + \frac{E_N \left[Y e^{i\beta_m u X_m} \right] E_N \left[X_m^2 e^{i\beta_m u X_m} \right]}{\left(E_N \left[e^{i\beta_m u X_m} \right] \right)^2} \right. \\
&\qquad \left. - \frac{2 E_N \left[Y e^{i\beta_m u X_m} \right] \left(E_N \left[X_m e^{i\beta_m u X_m} \right] \right)^2}{\left(E_N \left[e^{i\beta_m u X_m} \right] \right)^3} - \frac{E_N \left[Y X_m^2 e^{i\beta_m u X_m} \right]}{\left(E_N \left[e^{i\beta_m u X_m} \right] \right)^2} \right) w(u) du
\end{aligned}
$$

[9] $Z_N \lesssim a_N$ means that there exists $C > 0$ such that $Z_N \leq C a_N$.

Let

$$\widehat{P}_N(\beta_m,u) = \left(\frac{E_N\left[Ye^{i\beta_m uX_m}\right]E_N\left[X_m e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{E_N\left[YX_m e^{i\beta_m uX_m}\right]}{E_N\left[e^{i\beta_m uX_m}\right]} + \frac{E_N\left[Ye^{iuY}\right]E_N\left[X_m e^{iuY}\right]}{\left(E_N\left[e^{iuY}\right]\right)^2} - \frac{E_N\left[YX_m e^{iuY}\right]}{E_N\left[e^{iuY}\right]}\right)$$
$$\times\left(\frac{2E_N\left[YX_m e^{i\beta_m uX_m}\right]E_N\left[X_m e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2} + \frac{E_N\left[Ye^{i\beta_m uX_m}\right]E_N\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{2E_N\left[Ye^{i\beta_m uX_m}\right]\left(E_N\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E_N\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2}\right)$$

and

$$\begin{aligned}
&P_0(\beta_m,u)\\
&= \left(\frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{E\left[YX_m e^{i\beta_m uX_m}\right]}{E\left[e^{i\beta_m uX_m}\right]} + \frac{E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]}{\left(E\left[e^{iuY}\right]\right)^2} - \frac{E\left[YX_m e^{iuY}\right]}{E\left[e^{iuY}\right]}\right)\\
&\quad\times\left(\frac{2E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} + \frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{2E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right)\\
&= \left(\left.\frac{\partial^2\varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0\partial s_m}\right|_{(0,\ldots,0,\beta_m u,0,\ldots,0)} - \left.\frac{\partial^2\varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0\partial s_m}\right|_{(u,0,\ldots,0)}\right)\cdot\left(\frac{2E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right.\\
&\quad\left.+\frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{2E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right)\\
&= \left(\beta_m\varphi_m''\left(\beta_m u\right) - \beta_m\varphi_m''\left(\beta_m u\right)\right)\cdot\left(\frac{2E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right.\\
&\quad\left.+\frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{2E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right)\\
&= 0
\end{aligned}$$

where the third equality follows from Equations (2.4) and (2.5).

Expand the brackets of $\widehat{P}_N(\beta_m,u)$ and se a Taylor expansion

$$\begin{aligned}
&\widehat{P}_N(\beta_m,u)\\
&= \frac{E\left[YX_m e^{i\beta_m uX_m}\right]E\left[YX_m^2 e^{i\beta_m uX_m}\right]}{E\left[e^{i\beta_m uX_m}\right]^2} - \frac{2E\left[Ye^{i\beta_m uX_m}\right]^2 E\left[X_m e^{i\beta_m uX_m}\right]^3}{E\left[e^{i\beta_m uX_m}\right]^5}
\end{aligned}$$

$$
\begin{aligned}
&- \frac{2E\left[X_m e^{i\beta_m u X_m}\right] E\left[Y X_m e^{i\beta_m u X_m}\right]^2}{E\left[e^{i\beta_m u X_m}\right]^3} - \frac{E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m e^{i\beta_m u X_m}\right] E\left[Y X_m^2 e^{i\beta_m u X_m}\right]}{d^3} \\
&- \frac{E\left[Y e^{i\beta_m u X_m}\right] E\left[Y X_m e^{i\beta_m u X_m}\right] E\left[X_m^2 e^{i\beta_m u X_m}\right]}{E\left[e^{i\beta_m u X_m}\right]^3} + \frac{E\left[Y X_m^2 e^{i\beta_m u X_m}\right] E\left[Y e^{iuY}\right] E\left[X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right] E\left[e^{iuY}\right]^2} \\
&+ \frac{E\left[Y e^{i\beta_m u X_m}\right]^2 E\left[X_m e^{i\beta_m u X_m}\right] E\left[X_m^2 e^{i\beta_m u X_m}\right]}{E\left[e^{i\beta_m u X_m}\right]^4} - \frac{E\left[Y X_m^2 e^{i\beta_m u X_m}\right] E\left[Y X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right] E\left[e^{iuY}\right]} \\
&- \frac{2E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m e^{i\beta_m u X_m}\right]^2 E\left[Y X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right]^3 E\left[e^{iuY}\right]} + \frac{2E\left[X_m e^{i\beta_m u X_m}\right] E\left[Y X_m e^{i\beta_m u X_m}\right] E\left[Y X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right]^2 E\left[e^{iuY}\right]} \\
&+ \frac{E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m^2 e^{i\beta_m u X_m}\right] E\left[Y X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right]^2 E\left[e^{iuY}\right]} + \frac{4E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m e^{i\beta_m u X_m}\right]^2 E\left[Y X_m e^{i\beta_m u X_m}\right]}{E\left[e^{i\beta_m u X_m}\right]^4} \\
&+ \frac{2E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m e^{i\beta_m u X_m}\right]^2 E\left[Y e^{iuY}\right] E\left[X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right]^3 E\left[e^{iuY}\right]^2} \\
&- \frac{2E\left[X_m e^{i\beta_m u X_m}\right] E\left[Y X_m e^{i\beta_m u X_m}\right] E\left[Y e^{iuY}\right] E\left[X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right]^2 E\left[e^{iuY}\right]^2} \\
&- \frac{E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m^2 e^{i\beta_m u X_m}\right] E\left[Y e^{iuY}\right] E\left[X_m e^{iuY}\right]}{E\left[e^{i\beta_m u X_m}\right]^2 E\left[e^{iuY}\right]^2} \\
&= P_0(\beta_m, u) \\
&+ h_0^1(\beta_m, u)\left(E_N\left[Y e^{i\beta_m u X_m}\right] - E\left[Y e^{i\beta_m u X_m}\right]\right) + h_0^2(\beta_m, u)\left(E_N\left[X_m e^{i\beta_m u X_m}\right] - E\left[X_m e^{i\beta_m u X_m}\right]\right) \\
&+ h_0^3(\beta_m, u)\left(E_N\left[Y X_m e^{i\beta_m u X_m}\right] - E\left[Y X_m e^{i\beta_m u X_m}\right]\right) + h_0^4(\beta_m, u)\left(E_N\left[e^{i\beta_m u X_m}\right] - E\left[e^{i\beta_m u X_m}\right]\right) \\
&+ h_0^5(\beta_m, u)\left(E_N\left[Y X_m^2 e^{i\beta_m u X_m}\right] - E\left[Y X_m^2 e^{i\beta_m u X_m}\right]\right) \\
&+ h_0^6(\beta_m, u)\left(E_N\left[X_m^2 e^{i\beta_m u X_m}\right] - E\left[X_m^2 e^{i\beta_m u X_m}\right]\right) \\
&+ h_0^7(\beta_m, u)\left(E_N\left[Y e^{iuY}\right] - E\left[Y e^{iuY}\right]\right) + h_0^8(\beta_m, u)\left(E_N\left[X_m e^{iuY}\right] - E\left[X_m e^{iuY}\right]\right) \\
&+ h_0^9(\beta_m, u)\left(E_N\left[Y X_m e^{iuY}\right] - E\left[Y X_m e^{iuY}\right]\right) + h_0^{10}(\beta_m, u)\left(E_N\left[e^{iuY}\right] - E\left[e^{iuY}\right]\right) \\
&+ \left(E[|Y|] + E[|X_m^2|] + E[|Y X_m^2|]\right)\left(\frac{1}{\left|E\left[e^{iuY}\right]\right|^4 \left|E\left[e^{i\beta_m u X_m}\right]\right|^3} + \frac{1}{\left|E\left[e^{i\beta_m u X_m}\right]\right|^7}\right) \times \\
&O\Big[\left(E_N\left[Y e^{i\beta_m u X_m}\right] - E\left[Y e^{i\beta_m u X_m}\right]\right)^2 + \dots \\
&+ \left|E_N\left[e^{i\beta_m u X_m}\right] - E\left[e^{i\beta_m u X_m}\right]\right| \left|E_N\left[X_m e^{iuY}\right] - E\left[X_m e^{iuY}\right]\right| + \dots \\
&+ \left|E_N\left[Y e^{i\beta_m u X_m}\right] - E\left[Y e^{i\beta_m u X_m}\right]\right| \left|E_N\left[Y X_m^2 e^{i\beta_m u X_m}\right] - E\left[Y X_m^2 e^{i\beta_m u X_m}\right]\right| + \dots \\
&+ \left(E_N\left[e^{iuY}\right] - E\left[e^{iuY}\right]\right)^2\Big]
\end{aligned}
$$

where the second equality follows by a Taylor expansion and

$$
\begin{aligned}
h_0^1(\beta_m, u) = &-\frac{1}{\left(E\left[e^{i\beta_m u X_m}\right]\right)^5 \left(E\left[e^{iuY}\right]\right)^2}\Big(4E\left[Y e^{i\beta_m u X_m}\right] E\left[X_m e^{i\beta_m u X_m}\right]^3 \left(E\left[e^{iuY}\right]\right)^2 \\
&+ 2E\left[Y X_m e^{iuY}\right] E\left[X_m e^{i\beta_m u X_m}\right]^2 \left(E\left[e^{i\beta_m u X_m}\right]\right)^2 E\left[e^{iuY}\right]
\end{aligned}
$$

$$- 2E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2\left(E\left[e^{i\beta_m uX_m}\right]\right)^2$$

$$- 4E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2 E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$+ E\left[YX_m^2 e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$- 2E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$- E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[YX_m e^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3 E\left[e^{iuY}\right]$$

$$+ E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3$$

$$+E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$h_0^2(\beta_m,u) = -\frac{1}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^5\left(E\left[e^{iuY}\right]\right)^2}\Big(6E\left[Ye^{i\beta_m uX_m}\right]^2 E\left[X_m e^{i\beta_m uX_m}\right]^2\left(E\left[e^{iuY}\right]\right)^2$$

$$-E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[Ye^{i\beta_m uX_m}\right]^2 E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$-8E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]E\left[YX_m e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$+4E\left[YX_m e^{iuY}\right]E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2 E\left[e^{iuY}\right]$$

$$-4E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2$$

$$+ E\left[YX_m^2 e^{i\beta_m uX_m}\right]E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$+ 2E\left[YX_m e^{i\beta_m uX_m}\right]^2\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$- 2E\left[YX_m e^{iuY}\right]E\left[YX_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3 E\left[e^{iuY}\right]$$

$$+2E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]E\left[YX_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3\Big)$$

$$h_0^3(\beta_m,u) = \frac{1}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^4\left(E\left[e^{iuY}\right]\right)^2}\Big(4E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2\left(E\left[e^{iuY}\right]\right)^2$$

$$+ 2E\left[YX_m e^{iuY}\right]E\left[X_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2 E\left[e^{iuY}\right]$$

$$- 2E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]E\left[X_m e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2$$

$$- 4E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$+ E\left[YX_m^2 e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$-E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$h_0^4(\beta_m,u) = \frac{1}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^6\left(E\left[e^{iuY}\right]\right)^2}\Big(10E\left[Ye^{i\beta_m uX_m}\right]^2 E\left[X_m e^{i\beta_m uX_m}\right]^3\left(E\left[e^{iuY}\right]\right)^2$$

$$-4E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[Ye^{i\beta_m uX_m}\right]^2 E\left[X_m e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$-16E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2 E\left[YX_m e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$+6E\left[YX_m e^{iuY}\right]E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2\left(E\left[e^{i\beta_m uX_m}\right]\right)^2 E\left[e^{iuY}\right]$$

$$-6E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2\left(E\left[e^{i\beta_m uX_m}\right]\right)^2$$

$$+3E\left[YX_m^2e^{i\beta_m uX_m}\right]E\left[Ye^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$+3E\left[X_m^2e^{i\beta_m uX_m}\right]E\left[Ye^{i\beta_m uX_m}\right]E\left[YX_me^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$-2E\left[X_m^2e^{i\beta_m uX_m}\right]E\left[YX_me^{iuY}\right]E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3E\left[e^{iuY}\right]$$

$$+2E\left[X_m^2e^{i\beta_m uX_m}\right]E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3$$

$$+6E\left[X_me^{i\beta_m uX_m}\right]E\left[YX_me^{i\beta_m uX_m}\right]^2\left(E\left[e^{i\beta_m uX_m}\right]\right)^2\left(E\left[e^{iuY}\right]\right)^2$$

$$-4E\left[YX_me^{iuY}\right]E\left[X_me^{i\beta_m uX_m}\right]E\left[YX_me^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3E\left[e^{iuY}\right]$$

$$+4E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]E\left[X_me^{i\beta_m uX_m}\right]E\left[YX_me^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3$$

$$-2E\left[YX_m^2e^{i\beta_m uX_m}\right]E\left[YX_me^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^3\left(E\left[e^{iuY}\right]\right)^2$$

$$+E\left[YX_m^2e^{i\beta_m uX_m}\right]E\left[YX_me^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^4E\left[e^{iuY}\right]$$

$$-E\left[YX_m^2e^{i\beta_m uX_m}\right]E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^4\Big)$$

$$h_0^5(\beta_m,u)=\frac{E\left[Ye^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^4\left(E\left[e^{iuY}\right]\right)^2}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2E\left[e^{iuY}\right]$$

$$-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2-E\left[YX_me^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$+E\left[Ye^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$h_0^6(\beta_m,u)=-\frac{1}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3\left(E\left[e^{iuY}\right]\right)^2}\Big(E\left[YX_me^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2E\left[e^{iuY}\right]$$

$$-E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2-E\left[YX_me^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2$$

$$+E\left[Ye^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]\left(E\left[e^{iuY}\right]\right)^2\Big)$$

$$h_0^7(\beta_m,u)=\frac{E\left[X_me^{iuY}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3\left(E\left[e^{iuY}\right]\right)^2}\Big(2E\left[Ye^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]^2$$

$$+E\left[YX_m^2e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2-2E\left[YX_me^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]$$

$$-E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right])$$

$$h_0^8(\beta_m,u)=\frac{E\left[Ye^{iuY}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3\left(E\left[e^{iuY}\right]\right)^2}\Big(2E\left[Ye^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]^2$$

$$+E\left[YX_m^2e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2-2E\left[YX_me^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]$$

$$-E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right])$$

$$h_0^9(\beta_m,u)=\frac{-1}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3E\left[e^{iuY}\right]}\Big(2E\left[Ye^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]^2$$

$$+E\left[YX_m^2e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2-2E\left[YX_me^{i\beta_m uX_m}\right]E\left[X_me^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]$$

$$-E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right])$$

$$h_0^{10}(\beta_m,u)=\frac{1}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3\left(E\left[e^{iuY}\right]\right)^3}\left(E\left[YX_me^{iuY}\right]E\left[e^{iuY}\right]-2E\left[Ye^{iuY}\right]E\left[X_me^{iuY}\right]\right)\times$$

$$\Big(2E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]^2 - 2E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]$$
$$+E\left[YX_m^2 e^{i\beta_m uX_m}\right]\left(E\left[e^{i\beta_m uX_m}\right]\right)^2 - E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]E\left[e^{i\beta_m uX_m}\right]\Big)$$

Substitute $\widehat{P}_N(\beta_m, u)$ and $P_0(\beta_m, u) = 0$ into $\sqrt{N}\widehat{Q}'_N(\beta_m)$

$$\sqrt{N}\widehat{Q}'_N(\beta_m) = 2i\int_{\mathcal{U}} u\widehat{P}_N(\beta_m, u)w(u)du$$
$$= \sqrt{N}2i\int_{\mathcal{U}} u\left\{h_0^1(\beta_m, u)\left(E_N\left[Ye^{i\beta_m uX_m}\right] - E\left[Ye^{i\beta_m uX_m}\right]\right)\right.$$
$$+ h_0^2(\beta_m, u)\left(E_N\left[X_m e^{i\beta_m uX_m}\right] - E\left[X_m e^{i\beta_m uX_m}\right]\right) + h_0^3(\beta_m, u)\left(E_N\left[YX_m e^{i\beta_m uX_m}\right] - E\left[YX_m e^{i\beta_m uX_m}\right]\right)$$
$$+ h_0^4(\beta_m, u)\left(E_N\left[e^{i\beta_m uX_m}\right] - E\left[e^{i\beta_m uX_m}\right]\right) + h_0^5(\beta_m, u)\left(E_N\left[YX_m^2 e^{i\beta_m uX_m}\right] - E\left[YX_m^2 e^{i\beta_m uX_m}\right]\right)$$
$$+ h_0^6(\beta_m, u)\left(E_N\left[X_m^2 e^{i\beta_m uX_m}\right] - E\left[X_m^2 e^{i\beta_m uX_m}\right]\right) + h_0^7(\beta_m, u)\left(E_N\left[Ye^{iuY}\right] - E\left[Ye^{iuY}\right]\right)$$
$$+ h_0^8(\beta_m, u)\left(E_N\left[X_m e^{iuY}\right] - E\left[X_m e^{iuY}\right]\right) + h_0^9(\beta_m, u)\left(E_N\left[YX_m e^{iuY}\right] - E\left[YX_m e^{iuY}\right]\right)$$
$$\left.+h_0^{10}(\beta_m, u)\left(E_N\left[e^{iuY}\right] - E\left[e^{iuY}\right]\right)\right\} w(u)du$$
$$+ \sqrt{N}2i\int_{\mathcal{U}} u\left\{\left(E[|Y|] + E[|X_m^2|] + E[|YX_m^2|]\right)\left(\frac{1}{\left|E\left[e^{iuY}\right]\right|^4\left|E\left[e^{i\beta_m uX_m}\right]\right|^3} + \frac{1}{\left|E\left[e^{i\beta_m uX_m}\right]\right|^7}\right)\times\right.$$
$$O\Big[\left(E_N\left[Ye^{i\beta_m uX_m}\right] - E\left[Ye^{i\beta_m uX_m}\right]\right)^2 + \dots$$
$$+ \left|E_N\left[e^{i\beta_m uX_m}\right] - E\left[e^{i\beta_m uX_m}\right]\right|\left|E_N\left[X_m e^{iuY}\right] - E\left[X_m e^{iuY}\right]\right| + \dots$$
$$+ \left|E_N\left[Ye^{i\beta_m uX_m}\right] - E\left[Ye^{i\beta_m uX_m}\right]\right|\left|E_N\left[YX_m^2 e^{i\beta_m uX_m}\right] - E\left[YX_m^2 e^{i\beta_m uX_m}\right]\right| + \dots$$
$$\left.+ \left(E_N\left[e^{iuY}\right] - E\left[e^{iuY}\right]\right)^2\Big]\right\} w(u)du$$
$$= 2i\sqrt{N}\frac{1}{N}\sum_{n=1}^{N}\int_{\mathcal{U}} u\left\{h_0^1(\beta_m, u)\left(Y_n e^{i\beta_m uX_{nm}} - E\left[Ye^{i\beta_m uX_m}\right]\right)\right.$$
$$+ h_0^2(\beta_m, u)\left(X_{nm}e^{i\beta_m uX_{nm}} - E\left[X_{nm}e^{i\beta_m uX_{nm}}\right]\right)$$
$$+ h_0^3(\beta_m, u)\left(Y_n X_{nm}e^{i\beta_m uX_{nm}} - E\left[YX_{nm}e^{i\beta_m uX_m}\right]\right) + h_0^4(\beta_m, u)\left(e^{i\beta_m uX_{nm}} - E\left[e^{i\beta_m uX_{nm}}\right]\right)$$
$$+ h_0^5(\beta_m, u)\left(Y_n X_{nm}^2 e^{i\beta_m uX_{nm}} - E\left[YX_{nm}^2 e^{i\beta_m uX_m}\right]\right) + h_0^6(\beta_m, u)\left(X_{nm}^2 e^{i\beta_m uX_{nm}} - E\left[X_{nm}^2 e^{i\beta_m uX_m}\right]\right)$$
$$+ h_0^7(\beta_m, u)\left(Y_n e^{iuY_n} - E\left[Ye^{iuY}\right]\right) + h_0^8(\beta_m, u)\left(X_{nm}e^{iuY_n} - E\left[X_m e^{iuY}\right]\right)$$
$$\left.+h_0^9(\beta_m, u)\left(Y_n X_{nm}e^{iuY_n} - E\left[YX_m e^{iuY}\right]\right) + h_0^{10}(\beta_m, u)\left(e^{iuY_n} - E\left[e^{iuY}\right]\right)\right\} w(u)du + o(1)$$
$$= 2i\sqrt{N}\frac{1}{N}\sum_{n=1}^{N} G(Y_n, X_n; \beta_m) + o(1)$$

where I denoted[10]

$$G(Y_n, X_n; \beta_m) =$$

$$\int_{\mathcal{U}} u \left\{ h_0^1(\beta_m, u) \left(Y_n e^{i\beta_m u X_{nm}} - E\left[Y e^{i\beta_m u X_m} \right] \right) + h_0^2(\beta_m, u) \left(X_{nm} e^{i\beta_m u X_{nm}} - E\left[X_{nm} e^{i\beta_m u X_{nm}} \right] \right) \right.$$

$$+ h_0^3(\beta_m, u) \left(Y_n X_{nm} e^{i\beta_m u X_{nm}} - E\left[Y X_{nm} e^{i\beta_m u X_m} \right] \right) + h_0^4(\beta_m, u) \left(e^{i\beta_m u X_{nm}} - E\left[e^{i\beta_m u X_{nm}} \right] \right)$$

$$+ h_0^5(\beta_m, u) \left(Y_n X_{nm}^2 e^{i\beta_m u X_{nm}} - E\left[Y X_{nm}^2 e^{i\beta_m u X_m} \right] \right) + h_0^6(\beta_m, u) \left(X_{nm}^2 e^{i\beta_m u X_{nm}} - E\left[X_{nm}^2 e^{i\beta_m u X_m} \right] \right)$$

$$+ h_0^7(\beta_m, u) \left(Y_n e^{iuY_n} - E\left[Y e^{iuY} \right] \right) + h_0^8(\beta_m, u) \left(X_{nm} e^{iuY_n} - E\left[X_m e^{iuY} \right] \right)$$

$$\left. + h_0^9(\beta_m, u) \left(Y_n X_{nm} e^{iuY_n} - E\left[Y X_m e^{iuY} \right] \right) + h_0^{10}(\beta_m, u) \left(e^{iuY_n} - E\left[e^{iuY} \right] \right) \right\} w(u) du$$

the second equality follows because $P_0(\beta_m, u) = 0$ and the Taylor expansion of $\widehat{P}_N(\beta_m, u)$, the third equality follows by using the linearity of $E_N := \frac{1}{N} \sum_{n=1}^{N}$ and

$$\sqrt{N} \int_{\mathcal{U}} u \left\{ \left(E[|Y|] + E[|X_m^2|] + E[|Y X_m^2|] \right) \left(\frac{1}{\left| E\left[e^{iuY} \right] \right|^4 \left| E\left[e^{i\beta_m u X_m} \right] \right|^3} + \frac{1}{\left| E\left[e^{i\beta_m u X_m} \right] \right|^7} \right) \times \right.$$

$$O\left[\left(E_N\left[Y e^{i\beta_m u X_m} \right] - E\left[Y e^{i\beta_m u X_m} \right] \right)^2 + \dots \right.$$

$$+ \left| E_N\left[e^{i\beta_m u X_m} \right] - E\left[e^{i\beta_m u X_m} \right] \right| \left| E_N\left[X_m e^{iuY} \right] - E\left[X_m e^{iuY} \right] \right| + \dots$$

$$+ \left| E_N\left[Y e^{i\beta_m u X_m} \right] - E\left[Y e^{i\beta_m u X_m} \right] \right| \left| E_N\left[Y X_m^2 e^{i\beta_m u X_m} \right] - E\left[Y X_m^2 e^{i\beta_m u X_m} \right] \right| + \dots$$

$$\left. + \left(E_N\left[e^{iuY} \right] - E\left[e^{iuY} \right] \right)^2 \right]$$

$$\le \sqrt{N} \varepsilon_N^2 \left(E[|Y|] + E[|X_m^2|] + E[|Y X_m^2|] \right) \left. \left(\frac{1}{\left| E\left[e^{iuY} \right] \right|^4 \left| E\left[e^{i\beta_m u X_m} \right] \right|^3} + \frac{1}{\left| E\left[e^{i\beta_m u X_m} \right] \right|^7} \right) \right\} w(u) du$$

$$\le \frac{\ln N}{\sqrt{N}} \left(E[|Y|] + E[|X_m^2|] + E[|Y X_m^2|] \right) \int_{\mathcal{U}} u \left(\frac{1}{\left| E\left[e^{iuY} \right] \right|^4 \left| E\left[e^{i\beta_m u X_m} \right] \right|^3} + \frac{1}{\left| E\left[e^{i\beta_m u X_m} \right] \right|^7} \right) w(u) du$$

$$= o(1)$$

where the second inequality follows by Lemma 2 and the last equality follows because $\frac{\ln N}{\sqrt{N}} \overset{n \to \infty}{\to} 0$ and the assumptions $E\left[Y^2 \right] < \infty$, $E\left[X_m^4 \right] < \infty$, $E\left[(Y X_m^2)^2 \right] < \infty$, $\int_{\mathcal{U}} u \left| E\left[e^{iuY} \right] \right|^{-4} \left| E\left[e^{i\beta_m u X_m} \right] \right|^{-3} w(u) du < \infty$, $\int_{\mathcal{U}} u \left| E\left[e^{i\beta_m u X_m} \right] \right|^{-7} w(u) du < \infty$, and $\int_{\mathcal{U}} u^2 \left| E\left[e^{i\beta_m u X_m} \right] \right|^{-6} w(u) du < \infty$.

Therefore $\sqrt{N} \widehat{Q}'_N(\beta_m)$ is the sample average of independent identically distributed random variables multiplied by a constant so by the Classical Central Limit

$$\sqrt{N} \widehat{Q}'_N(\beta_m) \overset{d}{\to} N(0, 4\Omega(\beta_m))$$

[10] $d_N = o(e_N)$ is Little-o notation and means that for every $\delta > 0$ there exists N large enough so that $d_n \le \delta e_n$ for all $n > N$.

where by linearity and the Dominated Convergence theorem $E\left[G(Y_n, X_n; \beta_m)\right] = 0$ and

$$\begin{aligned}\Omega(\beta_m) &= E\left[G(Y_n, X_n; \beta_m)^2\right] \\ &= \int_{\mathcal{U}}\int_{\mathcal{U}} uv\left\{h_0^1(\beta_m, u)h_0^1(\beta_m, v)Cov\left(Ye^{i\beta_m uX_m}, Ye^{i\beta_m vX_m}\right)\right. \\ &\qquad + h_0^1(\beta_m, u)h_0^2(\beta_m, v)Cov\left(Ye^{i\beta_m uX_m}, X_m e^{i\beta_m vX_m}\right) + \dots \\ &\qquad + h_0^7(\beta_m, u)h_0^4(\beta_m, v)Cov\left(Ye^{iuY}, e^{i\beta_m vX_m}\right) + \dots \\ &\qquad \left. + h_0^{10}(\beta_m, u)h_0^{10}(\beta_m, v)Cov\left(e^{iuY}, e^{ivY}\right)\right\} w(u)w(v)dudv\end{aligned}$$

2.10.3 Proof of Condition 2(iv): $Q''_N(b)$ Converges Uniformly in Probability to $H_0(b)$ and $H_0(\beta_m)$ is Nonsingular

To prove that $\widehat{Q}''_N(\beta_m)$ converges uniformly to $H_0(b)$ use a Taylor expansion and Lemma 2 along with the assumptions $E\left[Y^2\right] < \infty$, $E\left[X_m^6\right] < \infty$, $E\left[(YX_m^3)^2\right] < \infty$, $\int_{\mathcal{U}} u^2 \left|E\left[e^{iuY}\right]\right|^{-2}\left|E\left[e^{ibuX_m}\right]\right|^{-4} w(u)du < \infty$, $\int_{\mathcal{U}} u^2\left|E\left[e^{ibuX_m}\right]\right|^{-6} w(u)du < \infty$ for all $b \in \mathcal{B}$ (The proof is similar to the proof of 1(iv). A detailed proof is available upon request).

Finally,

$$\begin{aligned}H_0(\beta_m) &:= \lim_{N\to\infty} \widehat{Q}''_N(\beta_m) \\ &= -2\lim_{N\to\infty}\int_{\mathcal{U}} u^2\left(\frac{2E_N\left[YX_m e^{i\beta_m uX_m}\right]E_N\left[X_m e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2} + \frac{E_N\left[Ye^{i\beta_m uX_m}\right]E_N\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2}\right. \\ &\qquad \left. - \frac{2E_N\left[Ye^{i\beta_m uX_m}\right]\left(E_N\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E_N\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2}\right)^2 w(u)du \\ &+ 2i\int_{\mathcal{U}} u^2\left(\frac{E_N\left[Ye^{i\beta_m uX_m}\right]E_N\left[X_m e^{i\beta_m uX_m}\right]}{\left(E_N\left[e^{i\beta_m uX_m}\right]\right)^2} - \frac{E_N\left[YX_m e^{i\beta_m uX_m}\right]}{E_N\left[e^{i\beta_m uX_m}\right]}\right. \\ &\qquad \left. + \frac{E_N\left[Ye^{iuY}\right]E_N\left[X_m e^{iuY}\right]}{\left(E_N\left[e^{iuY}\right]\right)^2} - \frac{E_N\left[YX_m e^{iuY}\right]}{E_N\left[e^{iuY}\right]}\right) \times \\ &\frac{\partial}{\partial b}\left(\frac{2E_N\left[YX_m e^{ibuX_m}\right]E_N\left[X_m e^{ibuX_m}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2} + \frac{E_N\left[Ye^{ibuX_m}\right]E_N\left[X_m^2 e^{ibuX_m}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2}\right. \\ &\qquad \left.\left. - \frac{2E_N\left[Ye^{ibuX_m}\right]\left(E_N\left[X_m e^{ibuX_m}\right]\right)^2}{\left(E_N\left[e^{ibuX_m}\right]\right)^3} - \frac{E_N\left[YX_m^2 e^{ibuX_m}\right]}{\left(E_N\left[e^{ibuX_m}\right]\right)^2}\right)\right|_{b=\beta_m} w(u)du \\ &= -2\int_{\mathcal{U}} u^2\left(\frac{2E\left[YX_m e^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2} + \frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right. \\ &\qquad \left. - \frac{2E\left[Ye^{i\beta_m uX_m}\right]\left(E\left[X_m e^{i\beta_m uX_m}\right]\right)^2}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^3} - \frac{E\left[YX_m^2 e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}\right)^2 w(u)du\end{aligned}$$

where the last equality follows because of uniform convergence and

$$\begin{aligned}
&\frac{E\left[Ye^{i\beta_m uX_m}\right]E\left[X_m e^{i\beta_m uX_m}\right]}{\left(E\left[e^{i\beta_m uX_m}\right]\right)^2}-\frac{E\left[YX_m e^{i\beta_m uX_m}\right]}{E\left[e^{i\beta_m uX_m}\right]}+\frac{E\left[Ye^{iuY}\right]E\left[X_m e^{iuY}\right]}{\left(E\left[e^{iuY}\right]\right)^2}-\frac{E\left[YX_m e^{iuY}\right]}{E\left[e^{iuY}\right]}\\
&=\left.\frac{\partial^2\varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0\partial s_m}\right|_{(0,\ldots,0,\beta_m u,0,\ldots,0)}-\left.\frac{\partial^2\varphi_{Y,\vec{X}}(\vec{s})}{\partial s_0\partial s_m}\right|_{(u,0,\ldots,0)}\\
&=\beta_m\varphi_m''\left(\beta_m u\right)-\beta_m\varphi_m''\\
&=0
\end{aligned}$$

where the second equality follows from Equations (2.4) and (2.5).

The assumption $\int_{\mathcal{U}} u^2\big(E\left[e^{i\beta_m uX_m}\right]\big)^{-6}w(u)du<\infty$ implies that $0<H_0(\beta_m)<\infty$ so $H_0(\beta_m)$ is non-singular and

$$\sqrt{N}\left(\widehat{\beta}_m-\beta_m\right)\xrightarrow{d}N\left(0,\left(H_0(\beta_m)\right)^{-2}\Omega(\beta_m)\right)$$

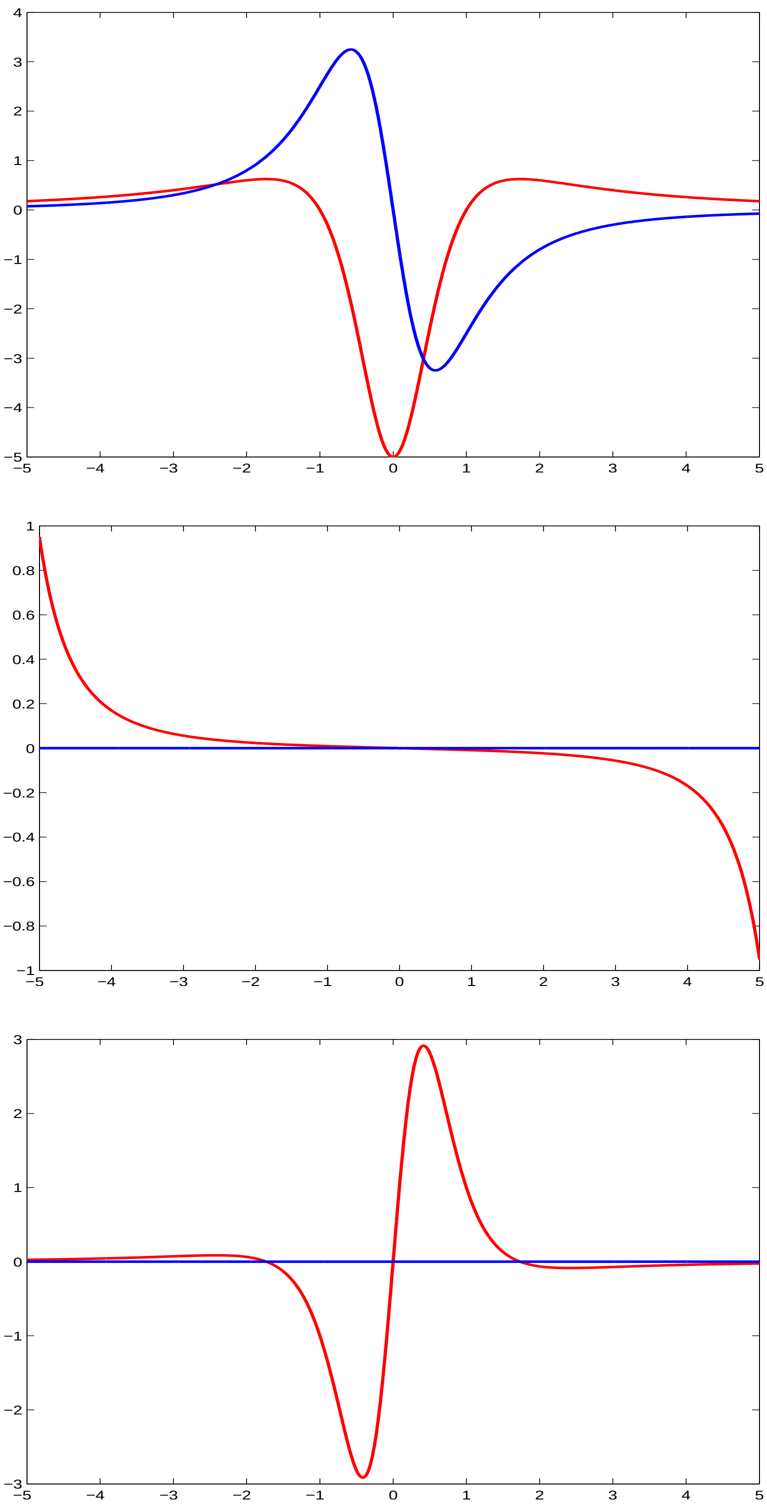

Figure 2.1: Top graph: $\varphi''_{m}(u) = -5/(iu-1)^2$ when $X^*_{m^*} \sim$ Gamma$(5,1)$
Middle graph: $\varphi''_{m^*}(u) = (2i + it^2e^{it} - 2ie^{it} - 2te^{it})/t^3$ when $X^*_m \sim$ Uniform$(0,1)$
Bottom graph: $\varphi'''_{m^*}(u) = -4u(u^2-3)/(u^2+1)^3$ when $X^*_m \sim$ Laplace$(0,1)$
The real parts are the red lines and the imaginary parts are the blue lines.

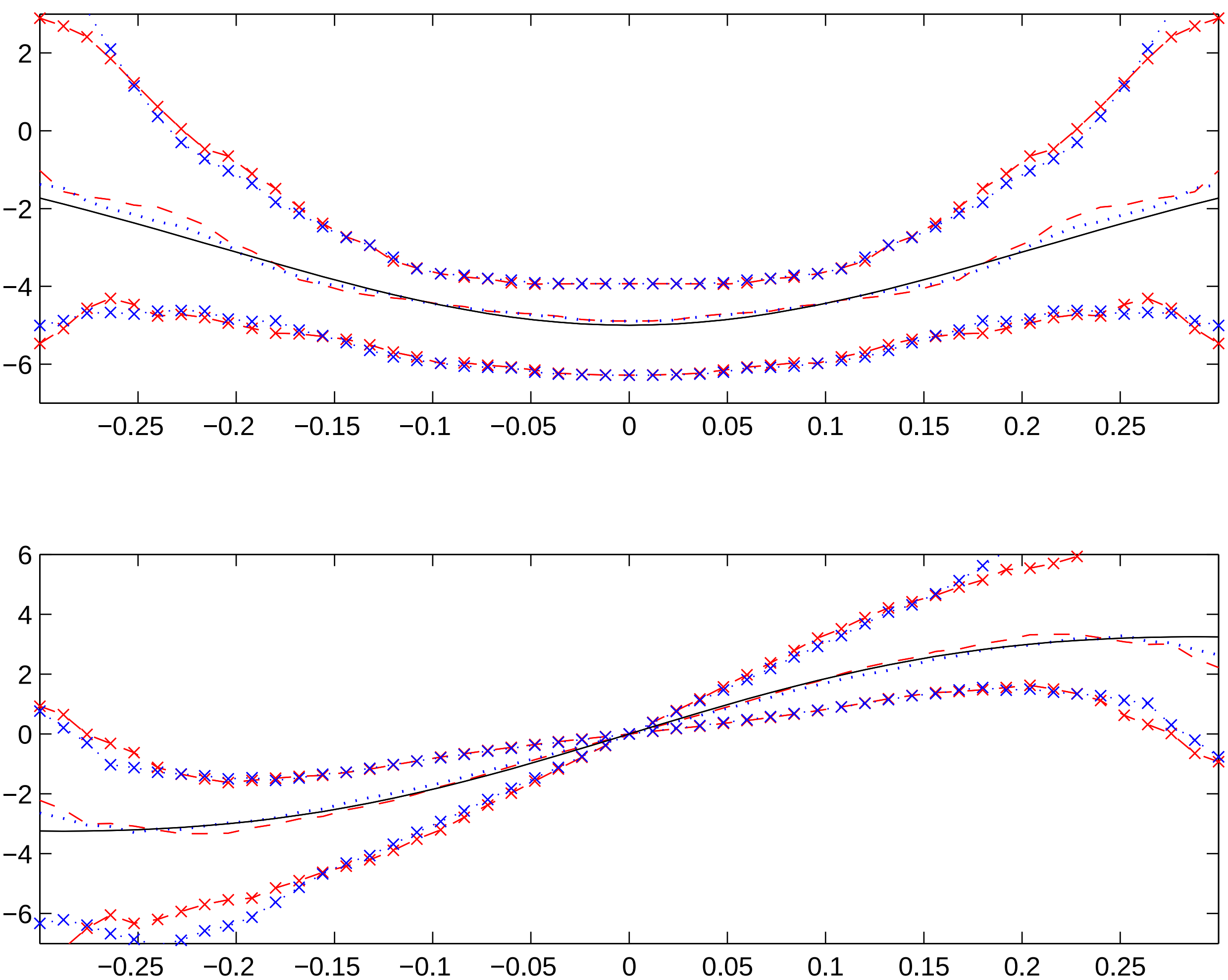

Figure 2.2: Errors-in-Variables. Experiment iv: $\left(f_{X_1^*}, f_{X_2^*}, f_{X_3^*}\right)$=(Gamma(5,1), Norm(1,1), Norm(1,1)) and $(\beta_0, \beta_1, \beta_2, \beta_3) = (3, -2, -1, 1)$ with $N = 100$

The top and bottom graphs depict the real and imaginary parts respectively of $\beta_1 \varphi''_{X_1^*}(\beta_1 u)$ (black solid line), the median of $\partial^2 \varphi_{Y,\vec{X}}(\vec{s})/\partial s_0 \partial s_m\Big|_{(0,\beta u)}$ (blue dotted line), its 10-90% confidence bands (blue dotted line with x's), the median of $\partial^2 \varphi_{Y,\vec{X}}(\vec{s})/\partial s_0 \partial s_m\Big|_{(u,0)}$ (red dashed line), and its 10-90% confidence bands (red dashed line with x's).

Table 2.1: Estimates for β_1 in the Errors-in Variables Model with N=100

Experiment	$(f_{X_1^*}, f_{X_2^*}, f_{X_3^*})$	$(\beta_0, \beta_1, \beta_2, \beta_3)$	Mean($\hat{\beta}_1$)	Stdev($\hat{\beta}_1$)
i	χ_2^2, Unif(0,1), Unif(0,1)	(3,2,1,-1)	2.0008	0.1645
ii	exp(1), Unif(0,1), Norm(1,1)	(3,2,-1,-1)	2.0066	0.1787
iii	Gamma(5,1), exp(1), Poiss(1)	(3,-2,1,1)	-1.9708	0.2084
iv	Gamma(5,1), Norm(1,1), Norm(1,1)	(3,-2,-1,1)	-1.9636	0.1225

Table 2.2: Estimates for β_1 in the Errors-in Variables Model with N=1,000

Experiment	$(f_{X_1^*}, f_{X_2^*}, f_{X_3^*})$	$(\beta_0, \beta_1, \beta_2, \beta_3)$	Mean($\hat{\beta}_1$)	Stdev($\hat{\beta}_1$)
i	χ_2^2, Unif(0,1), Unif(0,1)	(3,2,1,-1)	1.9961	0.0385
ii	exp(1), Unif(0,1), Norm(1,1)	(3,2,-1,-1)	1.9977	0.0515
iii	Gamma(5,1), exp(1), Poiss(1)	(3,-2,1,1)	-1.9963	0.0484
iv	Gamma(5,1), Norm(1,1), Norm(1,1)	(3,-2,-1,1)	-1.9968	0.0352

Table 2.3: Estimates for β_1 in the Errors-in Variables Model with N=10,000

Experiment	$(f_{X_1^*}, f_{X_2^*}, f_{X_3^*})$	$(\beta_0, \beta_1, \beta_2, \beta_3)$	Mean($\hat{\beta}_1$)	Stdev($\hat{\beta}_1$)
i	χ_2^2, Unif(0,1), Unif(0,1)	(3,2,1,-1)	1.9996	0.0085
ii	exp(1), Unif(0,1), Norm(1,1)	(3,2,-1,-1)	1.9983	0.0143
iii	Gamma(5,1), exp(1), Poiss(1)	(3,-2,1,1)	-1.9994	0.0139
iv	Gamma(5,1), Norm(1,1), Norm(1,1)	(3,-2,-1,1)	-2.0002	0.0128

Bibliography

[1] ANGRIST, J. and KRUEGER, A. (2000), "Empirical Strategies in Labor Economics," *Handbook of Labor Economics*, 1277-1366.

[2] ARELLANO, M. and HONORÉ, B. (2001), "Panel Data Models: Some Recent Developments," *Handbook of Econometrics*, Vol. 5, ed. by J. J. Heckman, and E. Leamer. North-Holland.

[3] ARELLANO, M. and HAHN, J. (2005), "Understanding Bias in Nonlinear Panel Models: Some Recent Developments," Invited Lecture, Econometric Society World Congress, London.

[4] ALTONJIl, J. G. and MATZKIN, R. L. (2005), "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, **73**, 1053-1102.

[5] BESTER, A. C. and HANSEN, C. (2011), "Identification of Marginal Effects in a Nonparametric Correlated Random Effects Model,", *Journal of Business and Economic Statistics*, forthcoming.

[6] BONDESSON, L. (1974), "Characterizations of Probability Laws Through Constant Regression," *Z. Wahrsch. v. Geb*, **29**, 93-115.

[7] BONHOMME, S., and SAUDER, U. (2010), "Recovering Distributions in Difference-in-Differences: A Comparison of Selective and Comprehensive Schooling," *Review of Economics and Statistics*, forthcoming.

[8] BONHOMME, S. and ROBIN, J.M. (2010), "Generalized Non-Parametric Deconvolu-

tion with an Application to Earnings Dynamics," *Review of Economic Studies*, **77** (2), 491-533.

[9] BOUND, J., BROWN C., DUNCAN G., and RODGERS W., (1994), "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, **12**, 345-368.

[10] BLUNDELL, R. and MACURDY, T. (2000), "Labour Supply: A Review of Alternative Approaches," *Handbook of Labor Economics*, Vol 3a, North Holland.

[11] BROWNING, M. and CARRO, J. (2007), "Heterogeneity and Microeconometrics Modelling," in Blundell, R., W.K. Newey, T. Persson (eds.), *Advances in Theory and Econometrics*, Vol. 3 ; Cambridge: Cambridge University Press.

[12] CARNEIRO, P., HANSEN, K., and HECKMAN, J.J. (2003), "Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice," *International Economic Review* **44 (2)**, 361-422

[13] CARRASCO, M., FLORENS J.-P. (2010), "Spectral Method for Deconvolving a Density," *Econometric Theory*, forthcoming

[14] CARROLL, R.J., RUPPERT, D., STEFANSKI, L.A., CRAINICEANU, C. (2006), *Measurement Error in Nonlinear Models: A Modern Perspective, Second Edition* (Chapman and Hall).

[15] CARROLL, R. J. and STEFANSKI, L. A. (1990). Carroll, "Approximate quasi-likelihood estimation in models with surrogate predictors," *Journal of the American Statistical Association*, **85**, 652-663.

[16] CHEN, X., HU, Y. and LEWBEL, A. (2009), "Nonparametric Identification and Estimation of Nonclassical Errors-in-Variables Models Without Additional Information", *Statistica Sinica*, **19**,949-968.

[17] CHEN, X., HONG, H. and NEKIPELOV, D. (2011), "Nonlinear Models of Measurement Errors," *Journal of Economic Literature*, forthcoming

[18] CHEN, X., HONG, H. and TAMER, E. (2005), "Measurement Error Models with Auxiliary Data." *Review of Economic Studies*, **72**, No. 2.

[19] CHERNOZHUKOV, V., FERNANDEZ-VAL, I., HAHN, J. and NEWEY, W. (2010), "Average and Quantile Effects in Nonseparable Panel Models," working paper.

[20] CHESHER, A. (2007), "Instrumental Values," *Journal of Econometrics*, **139**, 15-34.

[21] CHESHER, A. (2009), "Excess Heterogeneity, Endogeneity and Index Restrictions," *Journal of Econometrics*, **152**, 37-45.

[22] CSÖRGO, S. (1981), "Limit Behaviour of the Empirical Characteristic Function," *Annals of Probability*, **9** (1), 130-144.

[23] CUNHA, F., HECKMAN, J., SCHENNACH, S. M., (2010), "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*. **78**, 883-931.

[24] DELAIGLE, A. and GIJBELS, I. (2002), "Estimation of Integrated Squared Density Derivatives from a Contaminated Sample," *Journal of the Royal Statistical Society, Series B*, **64**, 869-886.

[25] DELAIGLE, A., HALL, P. and MEISTER, A. (2008), "On Deconvolution with Repeated Measurements," *Annals of Statistics*, **36** (2), 665-685.

[26] EVDOKIMOV, K. (2011), "Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity," working paper.

[27] EVDOKIMOV, K. WHITE, H. (2011). "An Extension of a Lemma of Kotlarski," *Econometric Theory*, forthcoming.

[28] FAN, J. Q. (1991), "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *Annals of Statistics*, **19**, 1257-1272.

[29] HAUSMAN, J.A., NEWEY W.K., and POWELL J.L., 1995, "Nonlinear errors in variables: estimation of some Engel curves," *Journal of Econometrics*, **65**, 205-233.

[30] HECKMAN, J. MATZKIN, R. L. and NESHEIM, L. (2010), "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, **78**, 1569-1591.

[31] HODERLEIN, S., and WHITE, H. (2009), "Nonparametric Identification in Nonsepa-

rable Panel Data Models with Generalized Fixed Effects", working paper.

[32] HOROWITZ, J. L. and MARKATOU, M. (1996), "Semiparametric Estimation of Regression Models for Panel Data," Review of Economic Studies, 63, 145-168.

[33] HSIAO, C. (1986). *Analysis of Panel Data* (Cambridge: Cambridge University Press).

[34] HSIAO, C. and WANG, Q.K., (2000), "Estimation of structural nonlinear errors-in-variables models by simulated least-squares method," *International Economic Review*, Vol. 41, No. 2, 523-542.

[35] HU, Y. (2008), "Identification and Estimation of Nonlinear Models with Misclassification Error using Instrumental Variables: A General Solution," *Journal of Econometrics*, vol. 144 , issue 1, pages 27-61.

[36] HU, Y. and RIDDER, G. (2010), "On Deconvolution as a First Stage Nonparametric Estimator," *Econometric Reviews*, **29**, 1-32.

[37] HU, Y. and RIDDER, G. (2012), "Estimation of Nonlinear Models with Mismeasured Regressors Using Marginal Information," *Journal of Applied Econometrics*, forthcoming

[38] HU, Y. and SCHENNACH, S. M. (2007), "Instrumental variable treatment of nonclassical measurement error models," *Econometrica*, **76**, 195-216.

[39] JUHN, C., MURPHY, K., and PIERCE, B. (1991), "Accounting for the Slowdown in Black-WhiteWage Convergence," *Workers and Their Wages*, pages 107-143.

[40] KLEPPER, S. and LEAMER, E. (1984), "Consistent sets of estimates for regressions with errors in all variables," *Econometrica*, **52**, 163-183.

[41] KHATRI, C. G. and RAO, C. R. (1968), "Solutions to Some Functional Equations and their Applications to Characterization of Probability Distributions," *Sankhyä*, **30**, 167-180.

[42] KHATRI, C. G. and RAO, C. R. (1972), "Functional Equations and Characterization of Probability Laws Through Luinear Functions of Random Variables," *Journal of Multivariate Analysis*, **2**, 162-173.

[43] KOTLARSKI, I. (1967), "On Characterizing the Gamma and Normal Distribution,"

Pacific Journal of Mathematics, **20**, 69-76.

[44] LEVINSOHN, J. and PETRIN, A.(2003), "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 317-342.

[45] LI, T. (2002),"Robust and Consistent Estimation of Nonlinear Errors-in-Variables Models," *Journal of Econometrics*, **110**, 126.

[46] LI, T., PERRIGNE, I. and VUONG, Q. (2000), "Conditionally Independent Private Information in OSC Wildcat Auctions," *Journal of Econometrics*, **98**, 129-161.

[47] LI, T. and VUONG, Q. (1998), "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, **65**, 139-165.

[48] MATZKIN, R. L. (2003), "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, **71** (5), 1339-1375.

[49] MATZKIN, R. L. (2007), "Nonparametric identification," *Handbook of Econometrics*, **6**, 5307-5368.

[50] MEGHIR, C. and PISTAFERRI, L. (2004),"Income Variance Dynamics and Heterogeneity", *Econometrica*, **72**, 132.

[51] MEGHIR, C. and PISTAFERRI, L. (2011),"Earnings, Consumption and Life-Cycle Choices", *Handbook of Labor Economics*, **9**, 773854.

[52] MEISTER, A. (2007), "Deconvolving Compactly Supported Densities." *Mathematical Methods of Statistics*, **16**, 63-76.

[53] NEWEY, W.K. (1994), "Kernel Estimation of Partial Means and a General Variance Estimator," *Econometric Theory*, **10**, 233-253.

[54] OLLEY, S. and PAKES, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, **64**, 1263-1295.

[55] POLLARD, D. (1984), *Convergence of Stochastic Processes* (New York: Springer).

[56] RAO, C. R. (1971), "Characterization of Probability Laws by Linear Functions," *Sankhyä*, **33**, 265-270.

[57] SCHENNACH, S. M. (2004a), "Estimation of Nonlinear Models with Measurement

Error," *Econometrica*, **72** (1), 33-75.

[58] SCHENNACH, S. M. (2004b), "Nonparametric Estimation in the Presence of Measurement Error," *Econometric Theory* **20**, 1046-109.

[59] SCHENNACH, S. M. and HU, Y. (2007), "Nonparametric Identification of the Classical Errors-in-Variables Model Without Side Information," working paper.

[60] SZÉKELY, G. J. and RAO, C. R. (2000), "Identifiability of Distributions of Independent Random Variables by Linear Combinations and Moments," *Sankhyä*, **62**, 193-202.

Chapter 3

Identification of Nonparametrically Distributed Random Coefficients in Linear Panel Data Models

3.1 Introduction

In this paper I consider the panel data linear regression model

$$Y_{nt} = X_{nt}'\beta_n + \varepsilon_{nt} \qquad t = 1, \ldots, T \quad n = 1, \ldots, N \tag{3.1}$$

where Y_{nt} is an outcome variable, X_{nt} is a vector of covariates, ε_{nt} is an error, and β_n is a vector of coefficients. My main objective is to show that identification is possible even when the coefficients are not fixed across individuals ($\beta_n = b$ for all n) and instead are nonparametrically distributed random variables. To illustrate this, I identify nonparametrically distributed random coefficients in a cross-sectional regression model, a panel data regression model, a fixed effects regression model from Maddala (1971), Chamberlain (1982), Arellano and Bover (1995), and Wooldridge (2005), and a first-order autoregressive panel data regression model from Alvarez and Arellano (2002), Bond and Windmeijer (2002)., and Arellano and Bonhomme (2011).

I identify the nonparametric joint distribution of the coefficients under various assumptions about the statistical dependence of coefficients on covariates, the conditional statistical relationship of coefficients (allowing them to be statistically dependent or equal in distribution), and the number of time periods per individual relative to the number of coefficients.

Linear regression models with fixed coefficients include unobserved heterogeneity only through the scalar error term. On the other hand, linear regression models with random coefficients can have multiple sources of unobserved heterogeneity through the random coefficients. In contrast to linear regression models with fixed coefficients, and more in line with reality, these random coefficients allow observationally equivalent individuals to respond differently to identical changes in covariates. For example, Card (2001) analyzes returns to schooling using a linear regression model with random coefficients. One of the aims of his research is to show that the marginal returns to schooling, as reflected by the random coefficient on education, are heterogeneous across the population. The focus in Foster and Hahn (2000) is not the distribution of unobserved heterogeneity but rather the expected value of consumer surplus, $E\left[S\left(\beta,\cdot\right)\right]=\int_b S(b,\cdot)f_\beta(b)db$. In order to estimate this expected value they first estimate the density of the coefficients, f_β.

Beran, Feuerverger, and Hall (1996) and Hoderlein, Klemela, and Mammen (2010) study linear models with nonparametrically distributed random coefficients that are independent of covariates. They use a Radon transform to estimate the distributions of coefficients. I take another approach to identification (and estimation) of the nonparametric distributions that uses the derivative of a log characteristic function (CF) of outcome variables with respect to a covariate. This is analogous to identification of a fixed coefficient by taking the derivative of an expected outcome variable with respect to a covariate. Identification is possible even when the data comes from a cross-section of the population and there are a countably infinite number of coefficients.

Arellano and Bonhomme (2011) "regard individual specific parameters as random draws from an unrestricted conditional distribution given regressors."[1] I deal with the dependence of the coefficients on the covariates by either introducing an instrumental variable or using the variation across time for each individual within a panel dataset. My contributions relative to Arellano and Bonhomme (2011), who use the panel data approach, are: (i) to allow coefficients to be statistically related either because they are conditionally arbitrary dependent or because they come from the same underlying distributions (for example, error terms in different periods can be modeled as homogeneous), and (ii) to allow the number of coefficients to be larger than the number of time periods.

The identification strategy uses a CF transformation to take advantage of the linear structure of the model. The main identification steps are to: 1) take partial derivatives of a log CF of a linear combination of outcome variables and 2) choose the arguments of this log CF. Specifically, the linearity in Equation (3.1) is exploited by a log CF transformation that retains the additivity:

$$\log \mathrm{CF}_{\sum Y_T}(\cdot) = \log \mathrm{CF}_{\beta_1}(\cdot) + \log \mathrm{CF}_{\beta_2}(\cdot) + \dots,$$

The separability of the $\log \mathrm{CF}_{\beta_m}(\cdot)$'s is exploited by partial derivatives with respect to covariates or arguments. This reduces the number of $\log \mathrm{CF}_{\beta_m}(\cdot)$'s on the right side of the equation. Then choices of arguments remove all but one of the log CFs of coefficients on the right hand side. This log CF is now expressed in terms of an observed partial derivative of a $\log \mathrm{CF}_{\sum Y_T}(\cdot)$.

Estimators are constructed from the identification proofs by replacing population quantities with sample analogs. The estimators are related to deconvolution estimators, which have slow convergence rates because of an ill-posed inverse problem and requirement of uniform convergence rates, and the Nadaraya-Watson kernel estimator, which is a locally weighted

[1] Arellano and Bonhomme (2011) view this method as a fixed effects approach because there are no restrictions on the distributions of the coefficients conditioned on covariates. Graham and Powell (2011) view this method as a correlated random coefficients approach because the 'random' coefficients can vary across individuals and the covariates can be 'correlated' with coefficients.

estimator that suffers from the curse of dimensionality.[2] Evdokimov (2011) shows that these estimators are consistent but optimal rates of convergence and asymptotic distributions as of yet have not been derived. The finite sample properties of the estimators are tested in Monte Carlo simulations and have tight confidence bands around their theoretical counterparts.

The literature on linear models is extensive. Linear panel data models with random coefficients are primarily concerned with expectations and variances (see Hsiao and Pesaran (2008) for a good review). Linear panel data models with fixed effects are analyzed by Maddala (1971), Mundlak (1978), and Chamberlain (1982). Linear panel data models with correlated random coefficients are analyzed by Graham and Powell (2011), who identify the expected value of the coefficients but not their distributions. Hoderlein, Nesheim, and Simoni (2012) analyze identification of nonparametrically distributed parameters conditioned on covariates in nonlinear models. They use a completeness condition that requires strong restrictions on the dimensionality of parameters relative to outcome variables.

The identification framework of this paper is based on the literature on linear models with multidimensional unobservables. The first paper in this literature is Kotlarski (1967). Subsequent papers include Khatri and Rao (1968), Székely and Rao (2003), Bonhomme and Robin (2011), and Ben-Moshe (2012a). In these papers the covariates are fixed across individuals and they do not show how to deal with unobserved variables that are homogeneous.

This paper is organized as follows. Section 3.2 presents the model, its assumptions, and the identification results. Section 3.3 presents examples that illustrate how to use the identification techniques from Section 2. Section 3.4 constructs the estimators. Section 3.5 presents Monte Carlo simulations. Section 3.6 concludes. Appendix A contains all the proofs from Section 2 and Appendix B contains detailed solutions to the examples in Section 3.

[2]When coefficients and covariates are dependent and covariates are continuous I believe the curse of dimensionality is unavoidable without additional restrictions. The reason is that the procedure is local so that estimating the density of $\beta|X = \bar{x}$ requires a lot of data near $\bar{x}$.

3.2 Identification

Consider the linear panel data model,

$$Y = X\beta$$

where $Y \in \mathbb{R}^T$ is an observed vector of outcomes, $\beta \in \mathbb{R}^M$ is an unobserved random vector of coefficients, and X is a $T \times M$ matrix of observed covariates. The goal in this paper is to identify the nonparametric joint distribution of β.[3, 4]

A general setup used in the handbook chapter of econometrics on panel data models by Arellano and Honoré (2001) is to let $\beta = (\gamma', \theta_1', \ldots, \theta_T', \alpha, \varepsilon_1, \ldots, \varepsilon_T)'$ and $X_t = (W_t', 0, \ldots, 0, Z_t', 0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)'$. The model is then rewritten as

$$Y_t = W_t'\gamma + Z_t'\theta_t + \alpha + \varepsilon_t \qquad\qquad t = 1, \ldots, T$$

where the unobservables γ and α are realized in T equations (per individual) while the unobservables θ_t and ε_t are realized in just a single equation (per individual).

3.2.1 Identification Using the Change of Variables Theorem

In this subsection I establish identification of the joint distribution of β using the well-known change of variables theorem. This method allows the components of β to be arbitrarily dependent but requires β to be independent of X and $dim(\beta) \leq T$.

Recall the change of variables theorem: Let $\beta \in \mathbb{R}^M$ be an unobserved arbitrarily dependent random vector, let $g : \mathbb{R}^M \rightarrow \mathbb{R}^T$ be a known, bijective, and differentiable function, and

[3] Each individual makes a random draw from the random matrix $\{Y, X, \beta\}$. The matrix $\{Y_n, X_n\}_{n=1}^N$ is observed while the vector $\{\beta_n\}_{n=1}^N$ is unobserved. For identification purposes, the joint distribution of $\{Y, X\}$ and the linear relationship $Y = X\beta$ is assumed known.

[4] Some of the covariates can be intercepts so that the model is rewritten as $Y = X\beta + \varepsilon$.

consider the observed vector $Y \in \mathbb{R}^T$ such that

$$Y = g(\beta)$$

then the change of variables formula for the density of β is

$$f_\beta(b) = f_Y(y) \left|\mathbf{det}\left(\frac{\mathbf{d}y}{\mathbf{d}b}\right)\right|$$

where $y = g(b)$ and $\left|\mathbf{det}\left(\frac{\mathbf{d}y}{\mathbf{d}b}\right)\right|$ is the absolute value of the determinant of the Jacobian.

The following is a straightforward modification of the change of variables theorem

Proposition 1. *Let $\beta \in \mathbb{R}^M$ be an unobserved arbitrarily dependent random vector, let $g_j : \mathbb{R}^M \to \mathbb{R}^T$ be known, bijective, and differentiable functions, and consider the observed vectors $Y_j \in \mathbb{R}^T$ such that*

$$Y_j = g_j(\beta) \qquad\qquad j = 1, \ldots$$

then the density of β can be expressed as

$$f_\beta(b) = \lim_{J\to\infty} \frac{1}{J} \sum_{j=1}^{J} f_{Y_j}(y) \left|\mathbf{det}\left(\frac{\mathbf{d}y_j}{\mathbf{d}b}\right)\right|$$

where $y_j = g_j(b)$ and $\left|\mathbf{det}\left(\frac{\mathbf{d}y_j}{\mathbf{d}b}\right)\right|$ is the absolute value of the determinant of the Jacobian.

Consider the linear panel data model

$$Y_j = X_j\beta$$

where $Y_j \in \mathbb{R}^T$ is a vector of observed outcomes, $\beta \in \mathbb{R}^M$ is a vector of arbitrarily dependent unobserved random coefficients, and X_j is a $T \times M$ matrix of observed covariates.

Corollary 1. *Assume* $X = (X_1, \ldots)$ *and* β *are independent.*[5] *If* X_j, $j = 1, \ldots$ *are square invertible matrices, then*

$$f_\beta(b) = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} f_{Y_j}\left(X_j b\right) \left|\mathbf{det}\left(X_j\right)\right|$$

Corollary 1 follows immediately from Proposition 1.

3.2.2 Identification Using Characteristic Functions

In this subsection I establish identification of the distribution of β conditioned on X using CF transformations. These methods allow β to be dependent on X and $T < dim(\beta)$.

I first explicitly describe the dependence of the unobserved coefficients β. Let $\beta = (\beta_1', \ldots, \beta_M')'$ and assume that conditional on X the unobserved vectors $\beta_m \in \mathbb{R}^{K_m}$, $m = 1, \ldots, M$ are mutually independent but $\beta_m = (\beta_{m1}, \ldots, \beta_{mK_m})$ are arbitrarily dependent. Let $X = (X_1, \ldots, X_M)$ with X_m a $T \times K_m$ matrix of observed covariates, and consider the observed vector $Y \in \mathbb{R}^T$ such that

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_T \end{pmatrix} = \begin{pmatrix} X_{11}^1 & \ldots & X_{1K_1}^1 \\ \vdots & \ddots & \vdots \\ X_{T1}^1 & \ldots & X_{TK_1}^1 \end{pmatrix} \begin{pmatrix} \beta_{11} \\ \vdots \\ \beta_{1K_1} \end{pmatrix} + \ldots + \begin{pmatrix} X_{11}^M & \ldots & X_{1K_M}^M \\ \vdots & \ddots & \vdots \\ X_{T1}^M & \ldots & X_{TK_M}^M \end{pmatrix} \begin{pmatrix} \beta_{M1} \\ \vdots \\ \beta_{MK_M} \end{pmatrix} \tag{3.2}$$

which can be represented as $Y = X_1\beta_1 + \ldots + X_M\beta_M$.

The following theorem uses the partial derivative of the log CF of a linear combination of outcome variables with respect to $x_{tk^*}^{m^*}$, $t = 1, \ldots, T$, which exploits the independence of coefficients and covariates. This method allows the dimension of the coefficients to be

[5]If X and β are dependent then identification is possible by first conditioning on conditioning $X_j = x$ and then applying the change of variables theorem.

countably infinite and subsets of β to be arbitrarily dependent but requires that the covariates and β be independent.

Condition on $X := (X_1, \ldots, X_M) = (x_1, \ldots, x_M) := x$

Assumption 9.

i. X and β are independent

ii. Span$(x'_{m^*}) = K_{m^*}$

Theorem 12. *If $E[|\beta_{m^*k}|] < \infty$ and $\int_0^{u_k} |(E[\exp i(\beta_{m^*1}u_1 + \ldots + \beta_{m^*k-1}u_{k-1} + \beta_{m^*k}v_k)])^{-1}|\mathrm{d}v_k < \infty$ for all fixed $u_1, \ldots, u_{k-1}$ and all u_k in the support of the CF of β_{m^*}, then β_{m^*} is identified when Assumption 9 holds. The CF of β_{m^*} is*

$$\phi_{\beta_{m^*}}(\vec{u}_{m^*}) = \sum_{t=1}^{T} E\left[\exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{s_{x_{m^*}kt}} \int_0^{u_k} \frac{E\left[\exp\left(\vec{Y}'\left(x'_{m^*}\right)^+ (u_1, \ldots, u_{k-1}, v_k, 0, \ldots, 0)'\right) \frac{\partial \ln f_{Y|X}(x)}{\partial x_{tk}^{m^*}}\right]}{E\left[\exp\left(\vec{Y}'\left(x'_{m^*}\right)^+ (u_1, \ldots, u_{k-1}, v_k, 0, \ldots, 0)'\right) |X = x\right]} dv_k\right)\right] w(t)$$

where $w(t)$ is a weight function that satisfies $\sum_{t=1}^{T} w(t) = 1$ and $w(t) \geq 0$.

The theorem uses the partial derivative of the log CF of $\vec{Y}$ with respect to X_{tk}^m, $\partial \ln E[\exp(i\vec{Y}\vec{s})]/\partial X_{tk}^m$ and the independence of X_{tk}^m and β. This is analogous to $\beta_{mk} = \partial E[Y_t]/\partial X_{tk}^m$ in the fixed coefficient framework. This approach no longer works if the unobserved heterogeneity (ε in the fixed coefficients framework and β in the random coefficients framework) depends on X_{tk}^m. When β is dependent on X_{tk}^m then the partial derivative of the log CF of Y with respect to $X_{t^*k^*}^{m^*}$ includes two terms: (1) the effects of the change on Y and (2) the effects on the density of β

$$\begin{aligned}\frac{\partial \varphi_{\beta_{m^*}|X}(s_{t^*}x_{t^*m^*})}{\partial x_{t^*m^*}} &= \frac{\partial \ln E\left[\exp\left(i\beta_{m^*}x_{t^*m^*}s_{t^*}\right)\right]}{\partial x_{t^*m^*}} \\ &= \frac{is_{t^*}E\left[\beta_{m^*}\exp\left(i\beta_{m^*}x_{t^*m^*}s_{t^*}\right)|X = x\right]}{E\left[\exp\left(i\beta_{m^*}x_{t^*m^*}s_{t^*}\right)|X = x\right]}\end{aligned}$$

$$+\frac{E\left[\exp\left(i\beta_{m^*}x_{t^*m^*}s_{t^*}\right)\frac{\partial\ln f_{\beta_{m^*}|X}(b)}{\partial x_{t^*m^*}}\middle| X=x\right]}{E\left[\exp\left(i\beta_{m^*}x_{t^*m^*}s_{t^*}\right)|X=x\right]}$$

When X and β are independent then the second term equals 0 and Theorem 12 follows. When X and β are dependent then the second term is not 0 and different techniques need to be used. Corollary 2 identifies β by an instrumental variable approach and Theorems 13, 14, and 15 identify β by using partial derivatives with respect to s_t, which will not include the second term.

Corollary 2. *Assume β_{mk} is dependent on $X_k^m = (X_{1k}^m, \ldots, X_{Tk}^m)'$ but there exists an instrumental variable $Z = (Z_1, \ldots, Z_T)'$ such that $X_k^m = Z\gamma$ where $\gamma \in \mathbb{R}$. If γ and β are independent and $(Z, X_1^1, \ldots, X_{k-1}^m, X_{k+1}^m, \ldots, X_{K_M}^M)$ is independent of γ and β, then the joint distribution of β_m is identified. If β_m is dependent on more covariates then β can still identified if there are more instrumental variables.*

Before stating Theorem 13 the following definition is needed:[6, 7, 8]

$$x^{t^*} = \begin{pmatrix} x_1^{t^*} & \ldots & x_M^{t^*} \end{pmatrix} = \left(x_1 \mathbf{I}\left(\bigcup_k x_{t^*k}^1 \neq 0\right) \;\ldots\; x_M \mathbf{I}\left(\bigcup_k x_{t^*k}^M \neq 0\right)\right)$$

Assumption 10. *There exists a $t_{k^*} \in \{1, \ldots, T\}$, and a vector $\vec{s}_{m^*} = (s_{m^*1}, \ldots, s_{m^*T})'$ for $k^* = 1, \ldots, K_{m^*}$ such that*

i. $x^{t_{k^*}\prime}\vec{s}_{m^*} = \begin{pmatrix} x_1^{t_{k^*}\prime}\vec{s}_{m^*} \\ \vdots \\ x_M^{t_{k^*}\prime}\vec{s}_{m^*} \end{pmatrix} = \begin{pmatrix} \vec{0}_{\sum_{m<m^*} K_m} \\ \vec{s}_{m^*} \\ \vec{0}_{\sum_{m>m^*} K_m} \end{pmatrix}$

[6] The function $\mathbf{I}(E)$ is the indicator function.
[7] Zero columns are removed from all matrices in this paper.
[8] Theorems 13 and 14 are very similar to theorems in Ben-Moshe (2012a), who has some further details and discussion on these theorems.

ii. $a^{m^*}_{t_{k^*}k} = 0$ *for all* $k \neq k^*$

where $\vec{0}_J = (0, \ldots, 0)'$ *is a column vector with* J *zeros and* $\vec{s}_{m^*} = (s_{m^*1}, \ldots, s_{m^*K_{m^*}})'$*.*

Theorem 13. *If* $E[|\beta_{m^*k}|] < \infty$ *and* $\int_0^{u_k} |(E[\exp i(\beta_{m^*1}u_1 + \ldots + \beta_{m^*k-1}u_{k-1} + \beta_{m^*k}v_k)])^{-1}| \mathrm{d}v_k < \infty$ *for all fixed* $u_1, \ldots, u_{k-1}$ *and all* u_k *in the support of the CF of* β_{m^*}*, then* β_{m^*} *is identified when Assumption 10 holds. The CF of* β_{m^*} *is*

$$\phi_{m^*|X}(\vec{u}_{m^*}) = \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{x^{m^*}_{t_k k}} \int_0^{u_k} \frac{iE\left[Y_{t_{k^*}} \exp\left(iY'\left(x^{t_{k^*}}\right)^+ (\vec{0}', u_1, \ldots, u_{k-1}, v_k, 0, \ldots, 0, \vec{0}')'\right)\right]}{E\left[\exp\left(iY'\left(x^{t_{k^*}}\right)^+ (\vec{0}', u_1, \ldots, u_{k-1}, v_k, 0, \ldots, 0, \vec{0}')'\right)\right]} dv_k \right.$$
$$\left. - \sum_{k=1}^{K_{m^*}} \frac{u_k}{x^{m^*}_{t_k k}} \sum_{m \neq m^*} \sum_{k'=1}^{K_m} x^m_{t_k k'} E\left[\beta_{mk'}|X = x\right] \right)$$

Remark 21. *The distributions of the coefficients in Corollary 1, Theorem 12, and Theorem 13 can be a point mass. This is the fixed coefficient linear regression model.*

Theorem 14 establishes identification of the joint distribution of β by solving a system of equations of second-order partial derivatives of the log CF of a linear combination of outcome variables. This method allows X and β to be arbitrarily dependent and $K_m \geq 1$, $m = 1, \ldots, M$ so that conditional on X subsets of β can be arbitrarily dependent. The model is described as in Equation (3.2), $Y = \beta_1 X_1 + \ldots + \beta_M X_M$.

Condition on $X := (X_1, \ldots, X_M) = (x_1, \ldots, x_M) := x$. Let $x_m = (x^m_1, \ldots, x^m_{K_m})$ be a partition of the matrix x_m where x^m_k is the k^{th} column of x_m. Define the matrix multiplication

$$x_m * x_m :=$$
$$\left(x^m_1 \otimes x^m_1,\ x^m_1 \otimes x^m_2 + x^m_2 \otimes x^m_1\ , \ldots,\ x^m_k \otimes x^m_k\ , \ldots,\ x^m_k \otimes x^m_{k+j} + x^m_{k+j} \otimes x^m_k\ , \ldots,\ x^m_{K_m} \otimes x^m_{K_m}\right)$$

where $\otimes$ is the Kronecker product and $1 \leq j \leq K_m - k$. The matrix $x_m * x_m$ has dimension $T^2 \times K_m(K_m + 1)/2$. Now, let $x = (x_1, \ldots, x_M)$ be a partition of the matrix x and define the matrix multiplication

$$x \odot x := (x_1 * x_1 \ , \ldots , \ x_M * x_M)$$

where $x \odot x$ is has dimension $T^2 \times K_m(K_m+1)/2$.

Assumption 11.

i. $Rank(x \odot x) = \sum_{m=1}^{M} K_m(K_m+1)/2$

ii. $Rank(x_m) = K_m$ *for all* m

Theorem 14. *If* $\int_0^{u_{k_2}} \int_0^{u_{k_1}} \left(E\left[\exp\left(i \sum_{k=1}^{k_1-1} \beta_{mk} u_k + i\beta_{mk_1} v_{k_1} + i\beta_{mk_2} v_{k_2} \right) \right] \right)^{-2} \mathrm{d}v_{k_1} \mathrm{d}v_{k_2} < \infty$ *for all fixed* $s_1, \ldots, s_{k_1-1}$ *and all* s_{k_1}, s_{k_2} *in the support of the CF of* $\vec{\beta}_m$ *and* $E\left[|\beta_{mk_1}\beta_{mk_2}| \right] < \infty$ *for* $k_1, k_2 = 1, \ldots, K_m$*, then the joint distribution of* β *conditional on* X *is identified when Assumption 11 holds. The CF of* β_{m^*} *is*

$$\phi_{m|X}(\vec{u}_m) = \exp\left(\sum_{k=1}^{K_m} \int_0^{u_k} \int_0^{w_k} \left. \frac{\partial \varphi^2_{m|X}(\vec{\omega}_m)}{\partial \omega^2_{mk}} \right|_{(0,\ldots,v_k,0,\ldots,0)} \mathrm{d}v_k \mathrm{d}w_k \right.$$
$$+ \sum_{k_1<k_2} \int_0^{u_{k_2}} \int_0^{u_{k_1}} \left. \frac{\partial \varphi^2_{m|X}(\vec{\omega}_m)}{\partial \omega_{mk_1} \partial \omega_{mk_2}} \right|_{(u_1,\ldots,u_{k_1-1},v_{k_1},0,\ldots,0,v_{k_2},0,\ldots,0)} \mathrm{d}v_{k_1} \mathrm{d}v_{k_2}$$
$$\left. + \sum_{k=1}^{K_m} u_k E\left[\beta_{mk} | X = x \right] \right)$$

Let $K_m = 1$, $m = 1, \ldots, M$ so that each matrix $X_m = (X^m_{11}, \ldots, X^m_{t1}, \ldots, X^m_{T1})'$ has only one column. The system is represented as

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_T \end{pmatrix} = \begin{pmatrix} X_{11} & \ldots & X_{1M} \\ \vdots & \ddots & \vdots \\ X_{T1} & \ldots & X_{TM} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_M \end{pmatrix}$$

where β_1, ... , β_{M-1}, and β_M are mutually independent.

In Theorem 4, I allow coefficients conditioned on X to be equal in distribution.[9] This allows homogeneity in the unobservables so that unobserved variables are drawn from the same distributions but do not need to be identical. To be specific define the equivalence classes

$$[\beta_{\widetilde{m}}] = \left\{\beta_m : \beta_m \overset{d}{=} \beta_{\widetilde{m}}\right\}$$

where $''\beta_m \overset{d}{=} \beta_{\widetilde{m}}\,''$ means β_m is equal in distribution to $\beta_{\widetilde{m}}$

$$f_{\beta_m|X}(b) = f_{\beta_{\widetilde{m}}|X}(b) \qquad\qquad \forall\ b \in \mathbb{R}$$

Let $\left\{[\beta_1], \ldots, \left[\beta_{\widetilde{M}}\right]\right\}$ be the equivalence classes, which are disjoint and partition $(\beta_1, \ldots, \beta_M)$.

Now, condition on $X := (X_1, \ldots, X_M) = (x_1, \ldots, x_M) := x$ and let $x = (x_1, \ldots, x_M)$ be a partition of x where x_m is the m^{th} column of x and define[10, 11]

$$\begin{aligned}
\widetilde{x}^{\widetilde{m}} &:= \left(\widetilde{x}_1^{\widetilde{m}} \ \ldots \ \widetilde{x}_M^{\widetilde{m}}\right) = \left(x_1\mathbf{I}\left(\beta_1 \in [\beta_{\widetilde{m}}]\right) \ \ldots \ \ x_M\mathbf{I}\left(\beta_M \in [\beta_{\widetilde{m}}]\right)\right) \\
\widetilde{x} &:= \left(\widetilde{x}^1 \ \ldots \ \widetilde{x}^{\widetilde{M}}\right) \\
\widetilde{x} \star \widetilde{x} &:= \left(\sum_{m=1}^{M} \widetilde{x}_m^1 \otimes \widetilde{x}_m^1 \ \ldots \ \sum_{m=1}^{M} \widetilde{x}_m^{\widetilde{M}} \otimes \widetilde{x}_m^{\widetilde{M}}\right)
\end{aligned}$$

Assumption 12.

i. $K_m = 1$ *so* β *is mutually independent*

[9] A similar kind of relationship structure on β can be used to modify Theorem 12.

[10] Columns of $\widetilde{x}^{\widetilde{m}}$ equal to the zero vector are removed.

[11] The matrix $\widetilde{x}\star\widetilde{x}$ has some repeated rows because the order of the scalar multiplication does not matter, that is $x_{t_1m}x_{t_2m} = x_{t_2m}x_{t_1m}$, so for calculation purposes I remove repeated rows and define the matrix $\widetilde{x}\bar{\star}\widetilde{x}$ as the matrix $\widetilde{x} \star \widetilde{x}$ without repeated rows so that a typical row looks like

$$\left[\sum_{m=1}^{M} x_{t1}x_{t+j1} \ , \ldots, \ \sum_{m=1}^{M} x_{tm}x_{t+jm}\right]$$

where $0 \leq j \leq T - t$. The matrix $x_m\bar{\star}x_m$ has dimension $(T+1)T/2 \times \widetilde{M}$.

ii. *The equivalence classes* $[\beta_{\widetilde{m}}]_{\widetilde{m}=1}^{\widetilde{M}}$ *are known*

iii. *There exists a vector* $\vec{\widetilde{s}} \in \mathbb{R}^T$ *such that*

$$\widetilde{x}'\vec{\widetilde{s}} = \vec{\widetilde{u}}$$

where $\vec{\widetilde{u}} = \left(\widetilde{u}_1 l_1, \ldots, \widetilde{u}_{\widetilde{M}} l_{\widetilde{M}}\right)' \in \mathbb{R}^M$, $u_{\widetilde{m}} \in \mathbb{R}$ *with* $u_{\widetilde{m}}$ *and* $u_{\widetilde{m}'}$ *not necessarily distinct and* $l_{\widetilde{m}} = (1, \ldots, 1)'$ *is a column vector of 1's of dimension* $|[\beta_{\widetilde{m}}]| \times 1$ *with* $|[\beta_{\widetilde{m}}]| = \sum_{m=1}^{M} \mathbf{I}(\beta_m \in [\beta_{\widetilde{m}}])$ *is the size of the equivalence class.*

iv. *Rank*$(\widetilde{x} \star \widetilde{x}) = \widetilde{M}$

Theorem 15. *If* $E[\beta_{\widetilde{m}}^2] < \infty$ *and* $\int_0^{u_{\widetilde{m}}} \int_0^w (E[\exp(iv\beta_{\widetilde{m}})])^{-2} \mathrm{d}v\mathrm{d}w < \infty$ *for all* $u_{\widetilde{m}}$ *in the support of* $\beta_{\widetilde{m}}$, *then the joint distribution of* β *conditional on* X *is identified when Assumption 12 holds. The CF of* $\beta_{\widetilde{m}}$ *is*

$$\phi_{\widetilde{m}|X}(u_{\widetilde{m}}) = \exp\left(\int_0^{u_{\widetilde{m}}} \int_0^w \varphi''_{\widetilde{m}|X}(v) \,\mathrm{d}v\mathrm{d}w + u_{\widetilde{m}} E[\beta_{\widetilde{m}}|X = x]\right)$$

Remark 22. *Assumption 12ii can be generalized by* $\beta_m \stackrel{d}{=} \sum a_{m\widehat{m}} \widetilde{\beta}_{\widetilde{m}}$ *but more caution is needed because equivalence classes might not be disjoint.*

Theorems 13, 14, and 15 assume that the conditional expectations of some unobservables are known. This is a strong assumption. There are at least two ways to deal with this: (1) Assume some unobservables have known expectations and use the formula $E[\beta|X] = E[(X'X)^{-1}X'Y|X]$ to identify the other expectations. As a rule of thumb the number of expectations that can be identified is less than or equal to the number of outcome variables (so that Xhas a pseudoinverse). Graham and Powell (2011), however, identify conditional expectations in a similar model under weaker condition can also be used (2) Concede that the expectations are not identified; and instead assume $E[\beta|X] = 0$ and identify the parameter

$b = E[Y]$ in the model $Y = b + X\beta$ (or $E[\varepsilon|X] = E[Y|X]$ in the model $Y = X\beta + \varepsilon$), which is similar to not being able to identify both the intercept and mean of the error in a fixed coefficient linear regression model.

3.3 Illustrative Examples

The following illustrative examples demonstrate how to use the Theorems in Section 3.2.2.[12]

3.3.1 Example 1: Cross Sectional and Panel Data Model

Consider the linear panel data model with random coefficients[13]

$$Y_t = \alpha + X_t'\beta + \varepsilon_t \qquad t = 1, \ldots, T$$

i. Let $T = 1$ and $\beta \in \mathbb{R}^M$ so that the data comes from a cross section of the population

$$Y_1 = \alpha + X_1'\beta + \varepsilon_1$$

Assume X and β are independent, and $(\alpha, \beta_1, \ldots, \beta_M, \varepsilon_1)$ is independent. Using Theorem 12,

$$\phi_{\beta_m}(u) = \exp\left(E\left[x_{1m}\int_0^u \frac{E\left[\exp\left(iY_1 v/x_{1m}\right)\dfrac{\partial \ln f_{Y_1|X}}{\partial x_{1m}}\right]}{vE\left[\exp\left(iY_1 v/x_{1m}\right)\right]}\mathrm{d}v\right]\right) \qquad m = 1, \ldots, M$$

[12] More detailed explanations of the examples are in Appendix B.

[13] Some of the papers that consider this setup are: Maddala (1971), Chamberlain (1982), Arellano and Bover (1995), and Wooldridge (2005).

The unobservables α and ε_1 are not separately identified but

$$\phi_{\alpha+\varepsilon_1}(u) = E\left[\frac{\phi_{Y_1|X}(u)}{\prod_{m=1}^{M}\phi_{\beta_m}(x_{1m}u)}\right]$$

ii. As in Example 1i, let $T = 1$ and $\beta \in \mathbb{R}^M$. Assume $X_1 = i_M$ (the $M \times 1$ vector of 1s)

$$Y_1 = \alpha + \beta_1 + \ldots \beta_M + \varepsilon_1$$

Assume $(\alpha, \beta_1, \ldots, \beta_M, \varepsilon_1)$ is independent but $\alpha \overset{d}{=} \beta_1 \overset{d}{=} \ldots \overset{d}{=} \beta_M \overset{d}{=} \varepsilon_1$ and assume without loss of generality that $E[\alpha] = E[\beta_1] = \ldots = E[\beta_M] = E[\varepsilon_1] = 0$ (otherwise normalize by subtracting $E[Y_1]$).

The CF of β_m is

$$\phi_{\beta_m}(s_1) = [\phi_{Y_1}(s_1)]^{\frac{1}{M+2}}$$

Remark 23. *When $Y_1 = \alpha + \varepsilon_1$ then this is the deconvolution problem with the assumption that $\alpha \overset{d}{=} \varepsilon_1$. When $M \to \infty$ then this is the start of the proof of the central limit theorem, which uses a Taylor expansion of the CF and further assumptions about existence of higher order moments.*

iii. Let $T = 2$ and $\beta \in \mathbb{R}$

$$Y_1 = \alpha + X_1\beta_1 + \varepsilon_1$$

$$Y_2 = \alpha + X_2\beta_1 + \varepsilon_2$$

Assume $\varepsilon_1 \overset{d}{=} \varepsilon_2 | X$ and $(\alpha, \beta_1, \varepsilon_1, \varepsilon_2)$ are mutually independent conditional on X, assume $E[\varepsilon_1|X] = E[\varepsilon_2|X] = 0$, and assume X and β are arbitrarily dependent.[14] When $x_1 \neq x_2$

[14] As mentioned earlier $E[\varepsilon_1|X] = E[\varepsilon_2|X] = 0$ is a strong assumption (notice $E[\alpha|X] \neq 0$ and $E[\beta_1|X] \neq 0$). This can be replaced with other perhaps weaker assumptions. Graham and Powell (2011) analyze the

then the expectation of (α, β_1) conditional on X is

$$E\left[\begin{array}{c|c} \alpha & \\ & X = x \\ \beta_1 & \end{array}\right] = \begin{pmatrix} -\frac{x_2}{x_1-x_2} & \frac{x_1}{x_1-x_2} \\ \frac{1}{x_1-x_2} & -\frac{1}{x_1-x_2} \end{pmatrix} E\left[\begin{array}{c|c} Y_1 & \\ & X = x \\ Y_2 & \end{array}\right]$$

Let $(\beta_1, \beta_2, \beta_3, \beta_4) = (\alpha, \beta_1, \varepsilon_1, \varepsilon_2)$, then

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & X_1 & 1 & 0 \\ 1 & X_2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

I now check Assumptions 12iii and 12iv. The details and explicit formulas for the CFs are left to Appendix B.

$$\widetilde{x}^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \widetilde{x}^2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \widetilde{x}^3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\widetilde{x}\bar{\star}\widetilde{x} = \begin{pmatrix} 1 & x_1^2 & 1 \\ 1 & x_1x_2 & 0 \\ 1 & x_2^2 & 1 \end{pmatrix}$$

Set $\vec{\vec{s}} = (1, 1)$ then $\widetilde{x}^{1\prime}\vec{\vec{s}} = 1$, $\widetilde{x}^{2\prime}\vec{\vec{s}} = x_1 + x_2$, $\widetilde{x}_1^{3\prime}\vec{s} = \widetilde{x}_2^{3\prime}\vec{\vec{s}} = (1, 1)'$ so Assumption 12iii is satisfied. Assume $|x_1| \neq |x_2|$ then Rank$(\widetilde{x}\bar{\star}\widetilde{x}) = 3$ and Assumption 12iv holds. Theorem 15 identifies the joint distribution of $(\alpha, \beta_1, \varepsilon_1, \varepsilon_2)$.

Remark 24. *By relabeling the variables Example 1iii can be viewed as an extension of*

same system of equations with $\varepsilon_2 = 0$.

a measurement error model with a repeated measurement[15]

$$X_1 = X^* + W^* + \varepsilon_1$$
$$X_2 = X^* + aW^* + \varepsilon_2 \qquad a^2 \neq 1$$

where X_1 and X_2 are two observed measurements. X^* and W^* are unobserved true variables, ε_1 and ε_2 are independent and identically distributed measurement errors, and a is a known constant.

3.3.2 Example 2: First-Order Autoregressive Process

The approach in this paper can be used under the more general formulation

$$Y = A(X, \delta)\beta$$

where $A(\cdot)$ is a $T \times M$ matrix of continuously differentiable functions that are known up to a vector of unknown common parameters δ.

Consider, for example, the first-order autoregressive panel data model

$$Y_t = \delta Y_{t-1} + X_t'\beta + \varepsilon_t \qquad |\delta| < 1$$

This model is considered, for example, by Maddala (1971), Alvarez and Arellano (2002), Bond and Windmeijer (2002), and Arellano and Bonhomme (2011). These papers assume δ and β are fixed parameters, $T \geq 3$, and $E[\varepsilon_1\varepsilon_2] = E[X\varepsilon_1] = E[X\varepsilon_2] = 0$. I assume that δ is a fixed parameter, β is a random variable, and $T = 2$. I require ε_1 and ε_2 to be independent conditional on X and $\varepsilon_1 \overset{d}{=} \varepsilon_2|X$.[16]

[15] The measurement error model with repeated measurements is analyzed for example by Kotlarski (1967) and Li and Vuong (1998).

[16] ε_1 and ε_1 do not need to be equal in distribution for identification of δ.

To be specific assume X_t is a scalar and $T = 2$, then

$$Y_1 = X_1\beta_1 + \delta Y_0 + \varepsilon_1$$
$$Y_2 = X_2\beta_1 + \delta X_1\beta_1 + \delta^2 Y_0 + \delta\varepsilon_1 + \varepsilon_2$$

where δ is an unknown fixed parameter and β_1 is a nonparametrically distributed random coefficient. Assume $\varepsilon_1 \stackrel{d}{=} \varepsilon_2|X$ and $(\beta_1, Y_0, \varepsilon_1, \varepsilon_2)$ are random variables that are mutually independent conditional on X, assume $E[\varepsilon_1|X] = E[\varepsilon_2|X] = 0$, and assume X and β_1 are arbitrarily dependent.[17] The fixed parameter is identified in Appendix B using a technique from Ben-Moshe (2012b).

When $x_2 \neq 0$, then

$$E[\beta_1|X] = \frac{E[Y_2|X] - \delta E[Y_1|X]}{x_2}$$
$$E[Y_0|X] = \frac{(x_2 + x_1\delta)E[Y_1|X] - x_1 E[Y_2|X]}{x_2\delta}$$

Let $(\beta_1, \beta_2, \beta_3, \beta_4) = (\beta_1, Y_0, \varepsilon_1, \varepsilon_2)$ then

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_2 + \delta X_1 & \delta^2 & \delta & 1 \\ X_1 & \delta & 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

I now check Assumptions 12iii and 12iv

$$\widetilde{x}^1 = \begin{pmatrix} x_1 \\ x_2 + \delta x_1 \end{pmatrix} \qquad \widetilde{x}^2 = \begin{pmatrix} \delta \\ \delta^2 \end{pmatrix} \qquad \widetilde{x}^3 = \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix}$$

[17]Despite $E[\varepsilon_1|X] = E[\varepsilon_2|X] = 0$, the covariates can be dependent with ε_1 and ε_2 in other ways.

$$\widetilde{x}\bar{\star}\widetilde{x} = \left(\begin{array}{ccc} x_1^2 & \delta^2 & 1 \\ (x_2 + \delta x_1)\, x_1 & \delta^3 & \delta \\ (x_2 + \delta x_1)^2 & \delta^4 & \delta^2 + 1 \end{array} \right)$$

Set $\vec{s} = (u(1-\delta), u)$ then $\widetilde{x}^{1\prime}\vec{s} = u(x_1 + x_2)$, $\widetilde{x}^{2\prime}\vec{s} = u\delta$, $\widetilde{x}_1^{3\prime}s = \widetilde{x}_2^{3\prime}s = u$ so Assumption 12iii is satisfied. Assume $x_1 \neq 0, x_2 \neq 0, \delta \neq 0$ then Rank$(\widetilde{x}\bar{\star}\widetilde{x}) = 3$ and Assumption 12iv holds. Theorem 15 identifies the joint distribution of $(\beta_1, Y_0, \varepsilon_1, \varepsilon_2)$.

Remark 25. *Similar techniques can be used to identify fixed parameters and unobserved distributions when Y_t follows an Autoregressive Process of order P (see for example Maddala (1971))*

$$Y_t = \sum_{p=1}^{P} \theta_p Y_{t-p} + X_t'\beta + \varepsilon_t$$

3.4 Estimation

Given i.i.d observations $\{Y_n, X_n\}_{n=1}^N$, estimators use the identification results by replacing population quantities with sample analogs.

When X and β are independent and $T = Dim(\beta)$ then an estimator is based on Corollary 1. Estimate $F_\beta(b)$ by the empirical distribution function

$$\widehat{F}_\beta(b) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{I}\left(X_n^{-1} Y_n \leq b\right)$$

This method is attractive when it can be used because estimators use densities of observed variables rather than CFs, which suffer from slow convergence rates and unknown asymptotic distributions.

When X and β are independent and $T < Dim(\beta)$ then an estimator is based on Theorem

12. Estimate $\phi_{\beta_{(mk)^*}}(u)$ by replacing population quantities with sample analogs

$$\widehat{\phi}_{\beta_{(mk)^*}}(u) = \exp\left(\frac{1}{N} \sum_{n=1}^{N} \frac{1}{s_{t^* x_n}^{(mk)^*}} \int_0^u \frac{\widehat{E}\left[\exp\left(ivY' s_x^{(mk)^*} \right) \frac{\partial \ln f_{Y|X}}{\partial x_{t^* k^*}^{m^*}} \right]}{v \widehat{E}\left[\exp\left(ivY' s_x^{(mk)^*} \right) \right]} \mathrm{d}v \right)$$

where for a function $g(x, y)$

$$\widehat{E}\left[g(y, x) \right] = \frac{\sum_n K_X(x - x_n) g(y_n, x_n)}{\sum_n K_X(x - x_n)}$$

is the Nadaraya-Watson kernel estimator and K_X is a Kernel that weights the observations x_n based on how close they are to x.

The density is identified using the inverse Fourier transformation and is estimated by a nonparametric kernel deconvolution estimator

$$\widehat{f}_{\beta_{(mk)^*}|X}(b) = \frac{1}{2\pi} \int \phi_K(u h_N) e^{-iub} \widehat{\phi}_{\beta_{(mk)^*}|X}(u) du$$

where ϕ_K is the Fourier transform of a kernel K supported on $[-1, 1]$ and h_N is the bandwidth of the kernel. In the Simulations section I use a second-order kernel[18]

$$K(b) = \frac{48 \cos(b)}{\pi b^4} \left(1 - \frac{15}{b^2} \right) - \frac{144 \sin(b)}{\pi b^5} \left(2 - \frac{5}{b^2} \right)$$

whose Fourier transform is

$$\phi_K(u) = (1 - u^2)^3 \mathbf{I}(u \in [-1, 1])$$

Estimators based on Theorems 13 to 15 replace population quantities with sample analogs and are constructed in a similar way to the estimator above for Theorem 12.

I do not prove consistency, which can be obtained from the existing literature. In par-

[18]See Delaigle and Gijbels (2002).

ticular, Evdokimov (2011) derives uniform convergence rates for a conditional distribution using partial derivatives of CFs.[19] Estimators that use deconvolutions are well-known to have slow convergence rates (see Carroll and Hall (1988) and Fan (1991)). The kernel-based estimator is a local estimator that weights data around x and will suffer from the curse of dimensionality.

3.5 Simulations

In this section, I study the finite sample behavior of the estimators obtained from Corollary 1, Theorem 12, and Theorem 15. The estimators of the densities have tight confidence bands around their underlying counterparts.

3.5.1 Estimator Using Corollary 1

Consider the linear panel data model with random coefficients,

$$Y_1 = \beta_1 X_{11} + \beta_2 X_{12}$$
$$Y_2 = \beta_1 X_{21} + \beta_2 X_{22}$$

Assume $(X_{11}, X_{12}, X_{21}, X_{22})$ and β are independent and

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \right) \quad \begin{pmatrix} X_{11} \\ X_{12} \\ X_{21} \\ X_{22} \end{pmatrix} \sim N\left(\begin{bmatrix} 5 \\ 5 \\ 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 & 0.3 & 0.3 & 0.3 \\ 0.3 & 1 & 0.3 & 0.3 \\ 0.3 & 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 0.3 & 1 \end{bmatrix} \right)$$

Based on Corollary 1, I estimate the marginal densities of β_1 and β_2 by generating 100 simulations each of sample size 100 and I estimate the joint density of (β_1, β_2) by generating

[19] Consistency in models without covariates can be found in, for example, Bonhomme and Robin (2010) and Ben-Moshe (2012a).

100 simulations each of sample size 500. The results are summarized graphically in Figures 3.1 and 3.2. Figure 3.1 represents the results for the marginal densities and figure 3.2 represents the results for the joint density.

3.5.2 Estimator Using Theorem 12

Consider the cross-sectional linear regression model with random coefficients

$$Y_1 = \beta_1 X_1 + \beta_2 X_2 + \varepsilon_1$$

Assume (X_1, X_2) and (β_1, β_2) are independent and

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \quad \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}\right) \quad \varepsilon_1 \sim N(0, 1)$$

Based on Theorem 12, I estimate the marginal densities of β_1 and β_2 by generating 100 simulations each of sample size 500. The results are summarized graphically in Figure 3.3.

3.5.3 Estimator Using Theorem 15

Consider the linear panel data model with random coefficients as in Example 1iii,

$$Y_1 = \alpha + \beta X_1 + \varepsilon_1$$

$$Y_2 = \alpha + \beta X_2 + \varepsilon_2$$

Assume

$$\begin{pmatrix} \alpha \\ \beta \\ X_1 \\ X_2 \end{pmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 5 \\ 10 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0.3 & 0.3 \\ 0 & 1 & 0.3 & 0.3 \\ 0.3 & 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 0.3 & 1 \end{bmatrix} \right) \qquad \begin{matrix} \varepsilon_1 \sim N(0,1) \\ \varepsilon_2 \sim N(0,1) \end{matrix}$$

so that (X_1, X_2) and (α, β) are dependent, the distributions of α and β are mutually independent conditional on X_1, and ε_1 and ε_2 are equally distributed and independent of X_1, X_2, α, and β.

Based on Theorem 15, I estimate the marginal density of β by generating 100 simulations each of sample size 500. The result is summarized graphically in Figure 4.

3.6 Conclusion

I study a linear model with nonparametrically distributed random coefficients. I identify the nonparametric distributions of these coefficients. The distributions of the coefficients can depend on covariates, coefficients can be conditionally statistically dependent or have homogeneous distributions, and the number of coefficients can be larger than the number of time periods per individual. I present examples to illustrate how the identification results can be used in practice and test their finite sample properties using Monte Carlo simulations, which suggest a practical estimation procedure.

3.7 Appendix A

3.7.1 Proof of Proposition 1

$$\lim_{J\to\infty}\frac{1}{J}\sum_{j=1}^{J}f_{Y_j}\left(y\right)\left|\mathbf{det}\left(\frac{\mathbf{d}y_j}{\mathbf{d}b}\right)\right| = \lim_{J\to\infty}\frac{1}{J}\sum_{j=1}^{J}f_{\beta}(b) = f_{\beta}(b)$$

where the first equality follows by the change of variables theorem.

3.7.2 Proof of Theorem 12

Let $\phi_{Y|X}$ denote the CF of Y conditioned on $X := (X_1, \ldots, X_M) = (x_1, \ldots, x_M) := x$ and let $\vec{s} = (s_1, \ldots, s_T)$. Then

$$\begin{aligned}
\phi_{Y|X}(\vec{s}) &= E\left[\exp\left(iY_1 s_1 + \ldots + iY_T s_T\right)|X = x\right] \\
&= E\left[\exp\left(i(x_{11}^1\beta_{11} + \ldots + x_{1K_M}^M\beta_{1K_M})s_1 + \ldots + i(x_{T1}^1\beta_{11} + \ldots + x_{TK_M}^M\beta_{MK_M})s_T\right)|X = x\right] \\
&= E\left[\exp\left(i(x_{11}^1 s_1 + \ldots + x_{T1}^1 s_T)\beta_{11} + \ldots + i(x_{1K_M}^M s_1 + \ldots + x_{TK_M}^M s_T)\beta_{MK_M}\right)|X = x\right] \\
&= \prod_{m=1}^{M} E\left[\exp\left(i\beta_{m1}\sum_{t=1}^{T}x_{t1}^m s_t + \ldots + i\beta_{mK_m}\sum_{t=1}^{T}x_{tK_m}^m s_t\right)\right]
\end{aligned}$$

where the second equality follows by substituting $Y_t = x_{t1}^1\beta_{11} + \ldots + x_{tK_M}^M\beta_{MK_M}$ and the last equality follows from the independence assumptions on β and the independence of X and β.

Let $\varphi_{Y|X}(\vec{s}) = \ln\phi_{Y|X}(\vec{s})$ and

$$\varphi_m\left(\vec{\omega}_m\right) = \varphi_{\beta_{m1},\ldots,\beta_{mK_m}}\left(\omega_{m1}, \ldots, \omega_{mK_m}\right) = \ln E\left[\exp\left(i\beta_{m1}\omega_{m1} + \ldots + i\beta_{mK_m}\omega_{mK_m}\right)\right]$$

then

$$\varphi_{Y|X}(\vec{s}) = \sum_{m=1}^{M}\varphi_m\left(\sum_{t=1}^{T}x_{t1}^m s_t, \ldots, \sum_{t=1}^{T}x_{tK_m}^m s_t\right) = \sum_{m=1}^{M}\varphi_m\left(x_1^{m\prime}\vec{s}, \ldots, x_{K_m}^{m\prime}\vec{s}\right) = \sum_{m=1}^{M}\varphi_m\left((x_m'\vec{s})'\right)$$

where $x = (x_1, \ldots, x_M)$ partitions x and $x_k^m = (x_{1k}^m, \ldots, x_{Tk}^m)'$ is the k^{th} column of x_m.

The partial derivative with respect to $x_{t^*k}^{m^*}$ is

$$\frac{\partial\varphi_{Y|X}(\vec{s})}{\partial x_{t^*k}^{m^*}} = s_{t^*} \times \left.\frac{\partial\varphi_{m^*}\left(\vec{\omega}_{m^*}\right)}{\partial\omega_k}\right|_{(x_{m^*}'s)'}$$

By Assumption 9i, $\text{span}(x'_{m^*}) = K_{m^*}$ so for any $\vec{u}_{m^*} \in \mathbb{R}^{K_{m^*}}$ there exists $\vec{s}_{x_{m^*}k} \in \mathbb{R}^T$ that solves $x'_{m^*}\vec{s}_{x_{m^*}k} = \vec{u}_{m^*}$. One solution is $\vec{s}_{x_{m^*}k} = (x'_{m^*})^+ \vec{u}_{m^*}$. Then

$$\left.\frac{\partial \varphi_{Y|X}(\vec{s})}{\partial x^{m^*}_{t^*k}}\right|_{(x'_{m^*})^+\vec{u}_{m^*}} = s_{x_{m^*}kt^*} \times \left.\frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_k}\right|_{\vec{u}_{m^*}} \tag{3.3}$$

The CF of β_{m^*} is expressed in terms of its first-order partial derivatives

$$\begin{aligned}
\phi_{\beta_{m^*}}(\vec{u}_{m^*}) &= \exp\left(\sum_{k=1}^{K_{m^*}} \int_0^{s_k} \left.\frac{\partial \varphi_{m^*}(\vec{\omega}_{m^*})}{\partial \omega_k}\right|_{(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)} dv_k\right) \\
&= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{s_{x_{m^*}kt^*}} \int_0^{u_k} \left.\frac{\partial \varphi_{Y|X}(\vec{s})}{\partial x^{m^*}_{t^*k}}\right|_{(x'_{m^*})^+(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)'} dv_k\right) \\
&= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{s_{x_{m^*}kt^*}} \int_0^{u_k} \left.\frac{\partial \ln E\left[\exp\left(\vec{Y}'\vec{s}\right)|X=x\right]}{\partial x^{m^*}_{t^*k}}\right|_{(x'_{m^*})^+(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)'} dv_k\right) \\
&= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{s_{x_{m^*}kt^*}} \int_0^{u_k} \frac{E\left[\exp\left(\vec{Y}'(x'_{m^*})^+(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)'\right)\frac{\partial \ln f_{Y|X}(x)}{\partial x^{m^*}_{t^*k}}\right]}{E\left[\exp\left(\vec{Y}'(x'_{m^*})^+(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)'\right)|X=x\right]} dv_k\right)
\end{aligned}$$

where the first equality uses the Fundamental Theorem of Calculus and the second equality follows by substituting Equation (3.3).

For estimation purposes, expectation is taken over X and weighted for each t

$$\begin{aligned}
&\phi_{\beta_{m^*}}(\vec{u}_{m^*}) \\
&= \sum_{t=1}^{T} E\left[\exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{s_{x_{m^*}kt}} \int_0^{u_k} \frac{E\left[\exp\left(\vec{Y}'(x'_{m^*})^+(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)'\right)\frac{\partial \ln f_{Y|X}(x)}{\partial x^{m^*}_{tk}}\right]}{E\left[\exp\left(\vec{Y}'(x'_{m^*})^+(u_1,\ldots,u_{k-1},v_k,0,\ldots,0)'\right)|X=x\right]} dv_k\right)\right] w(t)
\end{aligned}$$

where $w(t)$ is a weight function that satisfies $\sum_{t=1}^{T} w(t) = 1$ and $w(t) \geq 0$.

The CF of $\vec{U}_{m^*}$ is bounded using the regularity conditions: $E[|\beta_{m^*k}|] < \infty$ and $\int_0^{u_k} |(E[\exp i(\beta_{m^*1}u_1 + \ldots + \beta_{m^*k-1}u_{k-1} + \beta_{m^*k}v_k)])^{-1}|\mathrm{d}v_k < \infty$ for $k = 1, \ldots, K_{m^*}$.

This shows that the CF of β_{m^*} is identified. The density of β_{m^*} is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_{m^*}(\vec{b}_{m^*}) = \frac{1}{2\pi}\int e^{-i\vec{u}'_{m^*}\vec{b}_{m^*}}\phi_{m^*}(\vec{u}_{m^*})\mathrm{d}\vec{u}_{m^*}$$

This identifies the joint distribution of β_m for all m and in turn the joint distribution of β by the mutual independence assumption.

3.7.3 Proof of Corollary 2

The proof uses Theorem 12 twice and the Change of Variables Theorem: The distribution of $\gamma \in \mathbb{R}$ is identified from $X_k^m = Z\gamma$ using Theorem 12. Substitute $X_k^m = Z\gamma$ into

$$\begin{aligned} Y &= X\beta \\ &= (X_1^1, \ldots, X_{k-1}^m, Z, X_{k+1}^m, \ldots, X_{K_M}^M)(\beta_{11}, \ldots, \beta_{mk-1}, \gamma\beta_{mk}, \beta_{mk+1}, \ldots, \beta_{MK_M})' \end{aligned}$$

The joint distribution of $(\beta_{m1}, \ldots, \beta_{mk-1}, \gamma\beta_{mk}, \beta_{mk+1}, \ldots, \beta_{mK_m})$ is identified using Theorem 12. The joint distribution of β_m is identified from the joint distribution of $(\beta_{m1}, \ldots, \beta_{mk-1}, \gamma\beta_{mk}, \beta_{mk+1}, \ldots, \beta_{mK_m})$ and γ using their independence and the Change of Variables Theorem.

3.7.4 Proof of Theorem 13

Let $\phi_{Y|X}$ denote the CF of Y conditioned on $X := (X_1, \ldots, X_M) = (x_1, \ldots, x_M) := x$ and let $\vec{s} = (s_1, \ldots, s_T)$. Then

$$\begin{aligned} \phi_{Y|X}(\vec{s}) &= E\left[\exp\left(iY_1 s_1 + \ldots + iY_T s_T\right) |X = x\right] \\ &= E\left[\exp\left(i(x_{11}^1\beta_{11} + \ldots + x_{1K_M}^M\beta_{1K_M})s_1 + \ldots + i(x_{T1}^1\beta_{11} + \ldots + x_{TK_M}^M\beta_{MK_M})s_T\right) |X = x\right] \\ &= E\left[\exp\left(i(x_{11}^1 s_1 + \ldots + x_{T1}^1 s_T)\beta_{11} + \ldots + i(x_{1K_M}^M s_1 + \ldots + x_{TK_M}^M s_T)\beta_{MK_M}\right) |X = x\right] \\ &= \prod_{m=1}^{M} E\left[\exp\left(i\beta_{m1}\sum_{t=1}^{T} x_{t1}^m s_t + \ldots + i\beta_{mK_m}\sum_{t=1}^{T} x_{tK_m}^m s_t\right) |X = x\right] \end{aligned}$$

where the second equality follows by substituting $Y_t = x_{t1}^1\beta_{11} + \ldots + x_{tK_M}^M\beta_{MK_M}$ and the last equality follows from the independence assumptions.

Let $\varphi_{Y|X}(\vec{s}) = \ln \phi_{Y|X}(\vec{s})$ and

$$\varphi_{m|X}\left(\vec{\omega}_m\right) = \varphi_{\beta_{m1},\ldots,\beta_{mK_m}}\left(\omega_{m1}, \ldots, \omega_{mK_m}|X\right) = \ln E\left[\exp\left(i\beta_{m1}\omega_{m1} + \ldots + i\beta_{mK_m}\omega_{mK_m}\right)|X = x\right]$$

then

$$\varphi_{Y|X}(\vec{s}) = \sum_{m=1}^{M} \varphi_{m|X}\left(\sum_{t=1}^{T} x_{t1}^m s_t, \ldots, \sum_{t=1}^{T} x_{tK_m}^m s_t\right) = \sum_{m=1}^{M} \varphi_{m|X}\left(x_1^{m\prime}\vec{s}, \ldots, x_{K_m}^{m\prime}\vec{s}\right) = \sum_{m=1}^{M} \varphi_{m|X}\left((x_m'\vec{s})'\right)$$

where $x = (x_1, \dots, x_M)$ partitions x and $x_k^m = (x_{1k}^m, \dots, x_{Tk}^m)'$ is the k^{th} column of x_m.

The first-order partial derivative with respect to $s_{t_{k^*}}$ is

$$\begin{aligned}\frac{\partial \varphi_{Y|X}(\vec{s})}{\partial s_{t_{k^*}}} &= \sum_{m=1}^{M}\sum_{k=1}^{K_m} x_{t_{k^*}k}^m \left.\frac{\partial \varphi_{m|X}(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{(x_m'\vec{s})'} \\ &= \sum_{m=1}^{M}\sum_{k=1}^{K_m} x_{t_{k^*}k}^m \left.\frac{\partial \varphi_{m|X}(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{\left(\mathbf{I}\left(\bigcup_k x_{t_{k^*}k}^m \neq 0\right)(x_m'\vec{s})'\right)} \\ &= \sum_{m=1}^{M}\sum_{k=1}^{K_m} x_{t_{k^*}k}^m \left.\frac{\partial \varphi_{m|X}(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{\left(x_m^{t_{k^*}\prime}\vec{s}\right)'}\end{aligned}$$

where $x^{t_{k^*}} = \left(x_1^{t_{k^*}}, \dots, x_M^{t_{k^*}}\right)$ partitions $x^{t_{k^*}}$.

By Assumption 10i, there exists $\vec{s}_{m^*}$ such that $x_m^{t_{k^*}\prime}\vec{s}_{m^*} = \vec{0}_{K_m}$ for all $m \neq m^*$ and $x_{m^*}^{t_{k^*}\prime}\vec{s}_{m^*} = \vec{u}_{m^*} \in \mathbb{R}^{K_m^*}$. One solution is $\vec{s}_{m^*} = (x^{t_{k^*}})^+ \left(\vec{0}'_{\sum_{m<m^*} K_m}\ ,\ \vec{u}'_{m^*}\ ,\ \vec{0}'_{\sum_{m>m^*} K_m}\right)'$. Denote this solution as $\vec{s}_{m^*} = (x^{t_{k^*}})^+ (\vec{0}', \vec{u}'_{m^*}, \vec{0}')'$. Then

$$\begin{aligned}\left.\frac{\partial \varphi_{Y|X}(\vec{s})}{\partial s_{t_{k^*}}}\right|_{(x^{t_{k^*}})^+(\vec{0}',\vec{u}'_{m^*},\vec{0}')'} &= \sum_{k=1}^{K_{m^*}} x_{t_{k^*}k}^{m^*} \left.\frac{\partial \varphi_{m^*|X}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k}}\right|_{\vec{u}_{m^*}} + \sum_{m\neq m^*}\sum_{k=1}^{K_m} \left.\frac{\partial \varphi_{m|X}(\vec{\omega}_m)}{\partial \omega_{mk}}\right|_{\vec{0}'_{K_m}} \\ &= x_{t_{k^*}k^*}^{m^*} \left.\frac{\partial \varphi_{m^*|X}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k^*}}\right|_{\vec{u}_{m^*}} + \sum_{m\neq m^*}\sum_{k=1}^{K_m} x_{t_{k^*}k}^m E\left[\beta_{mk}|X=x\right] \qquad (3.4)\end{aligned}$$

where the second equality follows from Assumption 10ii that $x_{t_{k^*}k}^{m^*} = 0$ for all $k \neq k^*$, and the assumption $\sum_{m\neq m^*}\sum_{k=1}^{K_m} x_{t_{k^*}k}^m E\left[\beta_{mk}|X=x\right]$ is previously identified or assumed known. The CF of $\beta_{m^*|X}$ is expressed in terms of its first-order partial derivatives

$$\begin{aligned}&\phi_{m^*|X}(\vec{u}_{m^*}) \\ &= \exp\left(\sum_{k=1}^{K_{m^*}} \int_0^{u_k} \left.\frac{\partial \varphi_{m^*|X}(\vec{\omega}_{m^*})}{\partial \omega_{m^*k}}\right|_{(u_1,\dots,u_{k-1},v_k,0,\dots,0)} dv_k\right) \\ &= \exp\left(\sum_{k=1}^{K_{m^*}} \left(\frac{1}{x_{t_kk}^{m^*}} \int_0^{u_k} \left.\frac{\partial \varphi_{Y|X}(\vec{s})}{\partial s_{t_k}}\right|_{(x^{t_{k^*}})^+(\vec{0}',u_1,\dots,u_{k-1},v_k,0,\dots,0,\vec{0}')'} dv_k \right.\right. \\ &\qquad\qquad \left.\left. -u_k \sum_{m\neq m^*}\sum_{k'=1}^{K_m} x_{t_kk'}^m E\left[\beta_{mk'}|X=x\right]\right)\right) \\ &= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{x_{t_kk}^{m^*}} \int_0^{u_k} \left.\frac{\partial \ln E\left[\exp\left(iY'\vec{s}\right)\right]}{\partial t_{p_{k^*}}}\right|_{(x^{t_{k^*}})^+(\vec{0}',u_1,\dots,u_{k-1},v_k,0,\dots,0,\vec{0}')'} dv_k \right. \\ &\qquad\qquad \left. - \sum_{k=1}^{K_{m^*}} \frac{u_k}{x_{t_kk}^{m^*}} \sum_{m\neq m^*}\sum_{k'=1}^{K_m} x_{t_kk'}^m E\left[\beta_{mk'}|X=x\right]\right)\end{aligned}$$

$$= \exp\left(\sum_{k=1}^{K_{m^*}} \frac{1}{x_{t_k k}^{m^*}} \int_0^{u_k} \frac{iE\left[Y_{t_{k^*}} \exp\left(iY' \left(x^{t_{k^*}}\right)^+ (\vec{0}', u_1, \ldots, u_{k-1}, v_k, 0, \ldots, 0, \vec{0}')'\right)\right]}{E\left[\exp\left(iY' \left(x^{t_{k^*}}\right)^+ (\vec{0}', u_1, \ldots, u_{k-1}, v_k, 0, \ldots, 0, \vec{0}')'\right)\right]} dv_k \right.$$
$$\left. - \sum_{k=1}^{K_{m^*}} \frac{u_k}{x_{t_k k}^{m^*}} \sum_{m \neq m^*} \sum_{k'=1}^{K_m} x_{t_k k'}^{m} E\left[\beta_{mk'} | X = x\right]\right)$$

where the first equality uses the Fundamental Theorem of Calculus and the second equality follows by substituting Equation (3.4).

The CF of β_{m^*} is bounded using the regularity conditions: $E\left[|\beta_{m^*k}|\right] < \infty$ and $\int_0^{u_k} |(E[\exp i(\beta_{m^*1}u_1 + \ldots + \beta_{m^*k-1}u_{k-1} + \beta_{m^*k}v_k)])^{-1}|\mathrm{d}v_k < \infty$ for $k = 1, \ldots, K_{m^*}$.

This shows that the CF of $\beta_{m^*}|X$ is identified. The density of $\beta_{m^*}|X$ is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_{m^*|X}(\vec{b}_{m^*}) = \frac{1}{2\pi} \int e^{-i\vec{u}_{m^*}\vec{b}_{m^*}} \phi_{m^*|X}(\vec{u}_{m^*}) \mathrm{d}\vec{u}_{m^*}$$

3.7.5 Proof of Theorem 14

Let $\phi_{Y|X}$ denote the CF of Y conditioned on $X := (X_1, \ldots, X_M) = (x_1, \ldots, x_M) := x$ and let $\vec{s} = (s_1, \ldots, s_T)$. Then

$$\begin{aligned}
\phi_{Y|X}(\vec{s}) &= E\left[\exp\left(iY_1 s_1 + \ldots + iY_T s_T\right)|X = x\right] \\
&= E\left[\exp\left(i(x_{11}^1\beta_{11} + \ldots + x_{1K_M}^M \beta_{1K_M})s_1 + \ldots + i(x_{T1}^1\beta_{11} + \ldots + x_{TK_M}^M \beta_{MK_M})s_T\right)|X = x\right] \\
&= E\left[\exp\left(i(x_{11}^1 s_1 + \ldots + x_{T1}^1 s_T)\beta_{11} + \ldots + i(x_{1K_M}^M s_1 + \ldots + x_{TK_M}^M s_T)\beta_{MK_M}\right)|X = x\right] \\
&= \prod_{m=1}^{M} E\left[\exp\left(i\beta_{m1}\sum_{t=1}^{T} x_{t1}^m s_t + \ldots + i\beta_{mK_m}\sum_{t=1}^{T} x_{tK_m}^m s_t\right)|X = x\right]
\end{aligned}$$

where the second equality follows by substituting $Y_t = x_{t1}^1\beta_{11} + \ldots + x_{tK_M}^M \beta_{MK_M}$ and the last equality follows from the independence assumptions.

Let $\varphi_{Y|X}(\vec{s}) = \ln \phi_{Y|X}(\vec{s})$ and

$$\varphi_{m|X}\left(\vec{\omega}_m\right) = \varphi_{\beta_{m1},\ldots,\beta_{mK_m}}\left(\omega_{m1}, \ldots, \omega_{mK_m}|X\right) = \ln E\left[\exp\left(i\beta_{m1}\omega_{m1} + \ldots + i\beta_{mK_m}\omega_{mK_m}\right)|X = x\right]$$

then

$$\varphi_{Y|X}(\vec{s}) = \sum_{m=1}^{M} \varphi_{m|X}\left(\sum_{t=1}^{T} x_{t1}^m s_t, \ldots, \sum_{t=1}^{T} x_{tK_m}^m s_t\right) = \sum_{m=1}^{M} \varphi_{m|X}\left(x_1^{m\prime}\vec{s}, \ldots, x_{K_m}^{m\prime}\vec{s}\right) = \sum_{m=1}^{M} \varphi_{m|X}\left((x_m'\vec{s})'\right)$$

where $x = (x_1, \ldots, x_M)$ partitions x and $x_k^m = (x_{1k}^m, \ldots, x_{Tk}^m)'$ is the k^{th} column of x_m.

The second-order partial derivatives of $\varphi_{Y|X}(\vec{s})$ are

$$\begin{pmatrix} \frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_1^2} \\ \vdots \\ \frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_{t_1} \partial s_{t_2}} \\ \vdots \\ \frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_T^2} \end{pmatrix} = (x \odot x) \begin{pmatrix} \left. \frac{\partial \varphi_{1|X}^2 (\vec{\omega}_1)}{\partial \omega_{11}^2} \right|_{(x_1' \vec{s})'} \\ \vdots \\ \left. \frac{\partial \varphi_{m|X}^2 (\vec{\omega}_m)}{\partial \omega_{mk_1} \partial \omega_{mk_2}} \right|_{(x_m' \vec{s})'} \\ \vdots \\ \left. \frac{\partial \varphi_{M|X}^2 (\vec{\omega}_M)}{\partial \omega_{MK_M}^2} \right|_{(x_M' \vec{t})'} \end{pmatrix}$$

$k_1 \leq k_2$.

By Assumption 11i

$$\left(\left. \frac{\partial \varphi_{1|X}^2 (\vec{\omega}_1)}{\partial \omega_{11}^2} \right|_{(x_1' \vec{s})'} \cdots \left. \frac{\partial \varphi_{M|X}^2 (\vec{\omega}_M)}{\partial \omega_{MK_M}^2} \right|_{(x_M' \vec{s})'} \right)' = (x \odot x)^+ \left(\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_1^2}, \ldots, \frac{\partial^2 \varphi_{Y}(\vec{s})}{\partial s_T^2} \right)'$$

By Assumption 11ii, for all $\vec{u}_m \in \mathbb{R}^{K_m}$ there exists a $\vec{s}_m \in \mathbb{R}^P$ that solves $x_m' \vec{s}_m = \vec{u}_m$. One solution is $\vec{s}_m = (x_m')^+ \vec{u}_m$. Then

$$\left(\cdots \left. \frac{\partial \varphi_{m|X}^2 (\vec{\omega}_m)}{\partial \omega_{m1}^2} \right|_{\vec{u}_m'} \cdots \left. \frac{\partial \varphi_{m|X}^2 (\vec{\omega}_m)}{\partial \omega_{mK_m}^2} \right|_{\vec{u}_m'} \cdots \right)' = (x \odot x)^+ \left(\left. \frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_1^2} \right|_{(x_m')^+ \vec{u}_m} \cdots \left. \frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_T^2} \right|_{(x_m')^+ \vec{u}_m} \right)'$$

where

$$\left. \frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_{t_1} \partial s_{t_2}} \right|_{(x_m')^+ \vec{u}_m} = \frac{E\left[Y_{t_1} e^{iY'(x_m')^+ \vec{u}_m} | X = x \right] E\left[Y_{t_2} e^{iY'(x_m')^+ \vec{u}_m} | X = x \right]}{\left(E\left[e^{iY'(x_m')^+ \vec{u}_m} | X = x \right] \right)^2} - \frac{E\left[Y_{t_1} Y_{t_2} e^{iY'(x_m')^+ \vec{u}_m} | X = x \right]}{E\left[e^{iY'(x_m')^+ \vec{u}_m} | X = x \right]}$$

The CF of U_m is expressed in terms of second-order partial derivatives

$$\phi_{m|X}(\vec{u}_m) = \exp\left(\sum_{k=1}^{K_m}\int_0^{u_k}\int_0^{w_k} \left.\frac{\partial\varphi^2_{m|X}(\vec{\omega}_m)}{\partial\omega^2_{mk}}\right|_{(0,\ldots,v_k,0,\ldots,0)} \mathrm{d}v_k\mathrm{d}w_k \right.$$
$$+\sum_{k_1<k_2}\int_0^{u_{k_2}}\int_0^{u_{k_1}} \left.\frac{\partial\varphi^2_{m|X}(\vec{\omega}_m)}{\partial\omega_{mk_1}\partial\omega_{mk_2}}\right|_{(u_1,\ldots,u_{k_1-1},v_{k_1},0,\ldots,0,v_{k_2},0,\ldots,0)} \mathrm{d}v_{k_1}\mathrm{d}v_{k_2}$$
$$\left. +\sum_{k=1}^{K_m} u_k E\left[\beta_{mk}|X=x\right]\right)$$

The CF is defined using the regularity conditions: $E[|\beta_{mk_1}\beta_{mk_2}|]<\infty$ and $\int_0^{u_{k_2}}\int_0^{u_{k_1}}(E[\exp(i\sum_{k=1}^{k_1-1}\beta_{mk}u_k + i\beta_{mk_1}v_{k_1}+i\beta_{mk_2}v_{k_2})])^{-2}\mathrm{d}v_{k_1}\mathrm{d}v_{k_2}<\infty$ for $k_1, k_2 = 1,\ldots,K_m$.

This shows that the CF of $\beta_m|X$ is identified. The density of $\beta_m|X$ is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_{m|X}(\vec{b}_m) = \frac{1}{2\pi}\int e^{-i\vec{u}_m\vec{b}_m}\phi_{m|X}(\vec{u}_m)\mathrm{d}\vec{u}_m$$

3.7.6 Proof of Theorem 15

The CF of Y conditioned on $X = x$ is

$$\begin{aligned}
\phi_{Y|X}(s_1,\ldots,s_T) &= E\left[\exp\left(iY_1s_1+\ldots+iY_Ts_T\right)|X=x\right]\\
&= E\left[\exp\left(i(x_{11}\beta_1+\ldots+x_{1M}\beta_M)s_1+\ldots+i(x_{T1}\beta_1+\ldots+x_{TM}\beta_M)s_T\right)|X=x\right]\\
&= E\left[\exp\left(i(x_{11}s_1+\ldots+x_{T1}s_T)\beta_1+\ldots+i(x_{1M}s_1+\ldots+x_{TM}s_T)\beta_M\right)|X=x\right]\\
&= \prod_{m=1}^{M} E\left[\exp\left(i\beta_m\sum_{t=1}^{T}x_{tm}s_t\right)|X=x\right]
\end{aligned}$$

where the second equality follows by substituting $Y_t = x_{t1}\beta_1+\ldots+x_{tM}\beta_M$ and the last equality follows from mutual independence.

Let $\varphi_{Y|X}(\vec{s}) = \ln\phi_{Y|X}(\vec{s})$ and $\varphi_{m|X}(u_m) = \ln E\left[\exp\left(i\beta_m u_m\right)|X=x\right]$, $m = 1,\ldots,M$ then

$$\varphi_{Y|X}(\vec{s}) = \sum_{m=1}^{M}\varphi_{m|X}\left(\sum_{t=1}^{T}x_{tm}s_t\right) = \sum_{m=1}^{M}\varphi_{m|X}\left(x'_m\vec{s}\right)$$

where $x_m = (x_{1m},\ldots,x_{Tm})'$ is the m^{th} column of x.

The second-order partial derivative with respect to s_{t_1} and s_{t_2} is

$$\begin{aligned}\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_{t_1} \partial s_{t_2}} &= \sum_{m=1}^{M} x_{t_1 m} x_{t_2 m} \varphi''_{m|X}\left(x'_m \vec{s}\right) \\ &= \sum_{m=1}^{M} \sum_{\widetilde{m}=1}^{\widetilde{M}} \mathbf{I}\left(\beta_m \in [\beta_{\widetilde{m}}]\right) x_{t_1 m} x_{t_2 m} \varphi''_{\widetilde{m}|X}\left(\mathbf{I}\left(\beta_m \in [\beta_{\widetilde{m}}]\right) x'_m \vec{s}\right) \\ &= \sum_{\widetilde{m}=1}^{\widetilde{M}} \left(\sum_{m=1}^{M} \widetilde{x}^{\widetilde{m}}_{t_1 m} \widetilde{x}^{\widetilde{m}}_{t_2 m} \right) \varphi''_{\widetilde{m}|X}\left(\widetilde{x}^{\widetilde{m}\prime}_{m} \vec{s}\right)\end{aligned}$$

where the second equality follows by Assumption 12ii.[20]

By Assumption 12iii there exists $\vec{\tilde{s}} \in \mathbb{R}^T$ such that $\widetilde{x}^{\widetilde{m}\prime}_m \vec{\tilde{s}} = \widetilde{u}_{\widetilde{m}}$ where $\widetilde{u}_{\widetilde{m}} \in \mathbb{R}$. $\widetilde{u}_{\widetilde{m}'}$ and $\widetilde{u}_{\widetilde{m}}$ do not need to be distinct. One solution is $\vec{\tilde{s}} = \left(\widetilde{x}'\right)^+ \vec{\tilde{u}}$. Then

$$\left.\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_{t_1} \partial s_{t_2}}\right|_{(\widetilde{x}')^+ \vec{\tilde{u}}} = \sum_{\widetilde{m}=1}^{\widetilde{M}} \left(\sum_{m=1}^{M} \widetilde{x}^{\widetilde{m}}_{t_1 m} \widetilde{x}^{\widetilde{m}}_{t_2 m} \right) \varphi''_{\widetilde{m}|X}\left(\widetilde{u}_{\widetilde{m}}\right)$$

In matrix notation the second-order partial derivatives can be represented as

$$\begin{pmatrix} \left.\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_1^2}\right|_{(\widetilde{x}')^+ \vec{\tilde{u}}} \\ \vdots \\ \left.\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_T^2}\right|_{(\widetilde{x}')^+ \vec{\tilde{u}}} \end{pmatrix} = \left(\widetilde{x} \star \widetilde{x}\right) \begin{pmatrix} \varphi''_{1|X}\left(\widetilde{u}_1\right) \\ \vdots \\ \varphi''_{\widetilde{M}|X}\left(\widetilde{u}_{\widetilde{M}}\right) \end{pmatrix}$$

By Assumption 12iv

$$\left(\varphi''_{1|X}\left(\widetilde{u}_1\right), \ldots, \varphi''_{\widetilde{M}|X}\left(\widetilde{u}_{\widetilde{M}}\right)\right)' = \left(\widetilde{x} \star \widetilde{x}\right)^+ \left(\left.\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_1^2}\right|_{(\widetilde{x}')^+ \vec{\tilde{u}}}, \ldots, \left.\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_T^2}\right|_{(\widetilde{x}')^+ \vec{\tilde{u}}} \right)'$$

where

$$\left.\frac{\partial^2 \varphi_{Y|X}(\vec{s})}{\partial s_{t_1} \partial s_{t_2}}\right|_{(\widetilde{x}')^+ \vec{\tilde{u}}} = \frac{E\left[Y_{t_1} e^{iY'(\widetilde{x}')^+ \vec{\tilde{u}}} | X = x\right] E\left[Y_{t_2} e^{iY'(\widetilde{x}')^+ \vec{\tilde{u}}} | X = x\right]}{\left(E\left[e^{iY'(\widetilde{x}')^+ \vec{\tilde{u}}} | X = x\right]\right)^2} - \frac{E\left[Y_{t_1} Y_{t_2} e^{iY'(\widetilde{x}')^+ \vec{\tilde{u}}} | X = x\right]}{E\left[e^{iY'(\widetilde{x}')^+ \vec{\tilde{u}}} | X = x\right]}$$

[20] For all $\beta_m \in [\beta_{\widetilde{m}}]$

$$\varphi_{\beta_m|X}(\omega_m) = \ln\left(\int \exp\left(ib\omega_m\right) f_{\beta_m|X}(b) db\right) = \ln\left(\int \exp\left(ib\omega_m\right) f_{\beta_{\widetilde{m}}|X}(b) db\right) = \varphi_{\beta_{\widetilde{m}}|X}\left(\omega_m\right)$$

where the second equality follows because $f_{\beta_m|X}(b) = f_{\beta_{\widetilde{m}}|X}(b)$ for all $b \in \mathbb{R}$.

Applying the Second Fundamental Theorem of calculus twice

$$\phi_{\widetilde{m}|X}(u_{\widetilde{m}}) = \exp\left(\int_0^{u_{\widetilde{m}}}\int_0^{w}\varphi''_{\widetilde{m}|X}(v)\,\mathrm{d}v\mathrm{d}w + u_{\widetilde{m}}E\left[\beta_{\widetilde{m}}|X=x\right]\right)$$

The CF is defined using the regularity conditions: $E\left[\beta_{\widetilde{m}}^2\right] < \infty$ and $\int_0^{u_{\widetilde{m}}}\int_0^{w}\left(E\left[\exp\left(iv\beta_{\widetilde{m}}\right)\right]\right)^{-2}\mathrm{d}v\mathrm{d}w < \infty$.

This shows that the CF of $\beta_{\widetilde{m}}|X$ is identified. The density of $\beta_{\widetilde{m}}|X$ is identified using the bijection between densities and CFs by the inverse Fourier transform

$$f_{\widetilde{m}|X}(b_{\widetilde{m}}) = \frac{1}{2\pi}\int e^{-iu_{\widetilde{m}}b_{\widetilde{m}}}\phi_{\widetilde{m}|X}(u_{\widetilde{m}})\mathrm{d}u_{\widetilde{m}}$$

3.8 Appendix B

3.8.1 Example 1i: Cross-Sectional Linear Regression Model

The log CF of Y conditional on X is

$$\begin{aligned}\varphi_{Y_1|X}(s_1) &= \varphi_{\alpha+\varepsilon_1}(s_1) + \sum_{m=1}^{M}\varphi_m(x_{1m}s_1)\\ \frac{\partial\varphi_{Y_1|X}(s_1)}{\partial x_{1m^*}} &= s_1\varphi'_{m^*}(x_{1m^*}s_1)\end{aligned}$$

where the first equality follows by the linearity, mutual independence, and independence of X and β. The result now follows by the Second Fundamental Theorem of Calculus.

3.8.2 Example 1ii: Cross-Sectional Linear Regression Model with only Intercepts

The log CF of Y is

$$\begin{aligned}\varphi_{Y_1}(s_1) &= \varphi_{\alpha}(s_1) + \varphi_{\beta_1}(s_1) + \ldots + \varphi_{\beta_M}(s_1) + \varphi_{\varepsilon_1}(s_1)\\ &= (M+2)\varphi_{\beta_m}(s_1)\end{aligned}$$

where the first equality follows from the mutual independence assumption and the second equality follows from the equality in distribution assumption. Then

$$\phi_{\beta_m}(s_1) = [\phi_{Y_1}(s_1)]^{1/M+2}$$

3.8.3 Example 1iii: Panel Data Linear Regression Model

The log CF of Y conditional on X is

$$\varphi_{Y|X}(s_1, s_2) = \varphi_{\alpha|X}(s_1 + s_2) + \varphi_{\beta_1|X}(x_1 s_1 + x_2 s_2) + \varphi_{\varepsilon_1|X}(s_1) + \varphi_{\varepsilon_2|X}(s_2)$$

$$\begin{pmatrix} \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2} \\ \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1 \partial s_2} \\ \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2} \end{pmatrix} = \begin{pmatrix} 1 & x_1^2 & 1 & 0 \\ 1 & x_1 x_2 & 0 & 0 \\ 1 & x_2^2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \varphi''_{\alpha|X}(s_1 + s_2) \\ \varphi''_{\beta_1|X}(s_1 x_1 + s_2 x_2) \\ \varphi''_{\varepsilon_1|X}(s_1) \\ \varphi''_{\varepsilon_2|X}(s_2) \end{pmatrix}$$

Set $s_1 = s_2 = u$ then by the equality in distribution assumption $\varphi_{\varepsilon_1|X}(u) = \varphi_{\varepsilon_2|X}(u)$. Hence,

$$\begin{pmatrix} \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{(u,u)} \\ \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{(u,u)} \\ \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2}\right|_{(u,u)} \end{pmatrix} = \begin{pmatrix} 1 & x_1^2 & 1 \\ 1 & x_1 x_2 & 0 \\ 1 & x_2^2 & 1 \end{pmatrix} \begin{pmatrix} \varphi''_{\alpha|X}(2u) \\ \varphi''_{\beta_1|X}((x_1 + x_2)u) \\ \varphi''_{\varepsilon_1|X}(u) \end{pmatrix}$$

Under the assumption that $x_1^2 \neq x_2^2$,

$$\begin{pmatrix} \varphi''_{\alpha|X}(2u) \\ \varphi''_{\beta_1|X}((x_1 + x_2)u) \\ \varphi''_{\varepsilon_1|X}(u) \end{pmatrix} = \begin{pmatrix} -\frac{x_1 x_2}{x_1^2 - x_2^2} & 1 & \frac{x_1 x_2}{x_1^2 - x_2^2} \\ \frac{1}{x_1^2 - x_2^2} & 0 & -\frac{1}{x_1^2 - x_2^2} \\ \frac{x_2}{x_1 + x_2} & -1 & \frac{x_1}{x_1 + x_2} \end{pmatrix} \begin{pmatrix} \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{(u,u)} \\ \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{(u,u)} \\ \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2}\right|_{(u,u)} \end{pmatrix}$$

so

$$\varphi''_{\alpha|X}(u) = -\frac{x_1 x_2}{x_1^2 - x_2^2} \cdot \left.\frac{\partial\varphi^2_{Y|X}(s_1, s_2)}{\partial s_1^2}\right|_{(u/2,u/2)} + \left.\frac{\partial\varphi^2_{Y|X}(s_1, s_2)}{\partial s_1 \partial s_2}\right|_{(u/2,u/2)} + \left.\frac{\partial\varphi^2_{Y|X}(s_1, s_2)}{\partial s_2^2}\right|_{(u/2,u/2)}$$

$$\varphi''_{\beta_1|X}(u) = \frac{1}{x_1^2 - x_2^2} \cdot \left.\frac{\partial\varphi^2_{Y|X}(s_1, s_2)}{\partial s_1^2}\right|_{(u/(x_1+x_2),u/(x_1+x_2))} - \frac{1}{x_1^2 - x_2^2} \cdot \left.\frac{\partial\varphi^2_{Y|X}(s_1, s_2)}{\partial s_2^2}\right|_{(u/(x_1+x_2),u/(x_1+x_2))}$$

$$\varphi''_{\varepsilon_1|X}(u) = \frac{x_2}{x_1+x_2}\cdot\left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{(u,u)} - \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1\partial s_2}\right|_{(u,u)} + \frac{x_1}{x_1+x_2}\cdot\left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2}\right|_{(u,u)}$$

and now use the Second Fundamental Theorem of Calculus to obtain the CFs

$$\phi_{\widetilde{m}^*|X}(u) = \exp\left(\int_0^u\int_0^w \varphi''_{\widetilde{m}^*|X}(v)\,\mathrm{d}v\mathrm{d}w + iuE[\beta_{\widetilde{m}^*}|X]\right) \qquad \widetilde{m}^* = \alpha,\beta_1,\varepsilon_1,\varepsilon_2$$

3.8.4 Example 2: First-Order Autoregressive Process

The log CF of Y conditional on X is

$$\varphi_{Y|X}(s_1,s_2) = \varphi_{\beta_1|X}(x_1s_1+(x_2+\delta x_1)s_2) + \varphi_{Y_0|X}(\delta s_1+\delta^2 s_2) + \varphi_{\varepsilon_1|X}(s_1+\delta s_2) + \varphi_{\varepsilon_2|X}(s_2)$$

where the equality follows from the independence assumptions. The second order partial derivatives are

$$\begin{pmatrix} \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2} \\ \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1\partial s_2} \\ \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2} \end{pmatrix} = \begin{pmatrix} x_1^2 & \delta^2 & 1 & 0 \\ (x_2+\delta x_1)x_1 & \delta^3 & \delta & 0 \\ (x_2+\delta x_1)^2 & \delta^4 & \delta^2 & 1 \end{pmatrix} \begin{pmatrix} \varphi''_{\beta_1|X}(x_1s_1+(x_2+\delta x_1)s_2) \\ \varphi''_{Y_0|X}(\delta s_1+\delta^2 s_2) \\ \varphi''_{\varepsilon_1|X}(s_1+\delta s_2) \\ \varphi''_{\varepsilon_2|X}(s_2) \end{pmatrix} \tag{3.5}$$

To identify the parameter δ I employ a technique from Ben-Moshe (2012b). For all $d\in\mathbb{R}$

$$\begin{aligned} d\cdot\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2} - \frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1\partial s_2} &= (dx_1-\delta x_1-x_2)x_1\varphi''_{\beta_1|X}(x_1s_1+(x_2+\delta x_1)s_2) \\ &\quad + (d\delta^2-\delta^3)\varphi''_{Y_0|X}(\delta s_1+\delta^2 s_2) + (d-\delta)\varphi''_{\varepsilon_1|X}(s_1+\delta s_2) \end{aligned} \tag{3.6}$$

Define

$$\begin{aligned} R(d,u) =: &\left(d\cdot\left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{((x_2+dx_1)u,0)} - \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1\partial s_2}\right|_{((x_2+dx_1)u,0)}\right) \\ &\qquad - \left(d\cdot\left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{(0,x_1u)} - \left.\frac{\partial\varphi^2_{Y|X}(s_1,s_2)}{\partial s_1\partial s_2}\right|_{(0,x_1u)}\right) \\ &= (dx_1-\delta x_1-x_2)x_1\left(\varphi''_{\beta_1|X}(x_1(x_2+dx_1)u) - \varphi''_{\beta_1|X}((x_2+\delta x_1)x_1u)\right) \\ &+ (d\delta^2-\delta^3)\left(\varphi''_{Y_0|X}(\delta(x_2+dx_1)u) - \varphi''_{Y_0|X}(\delta^2 x_1u)\right) + (d-\delta)\left(\varphi''_{\varepsilon_1|X}((x_2+dx_1)u) - \varphi''_{\varepsilon_1|X}(\delta x_1u)\right) \end{aligned}$$

where the second equality follows by substituting in Equation (3.6) evaluated in two directions: $(s_1,s_2) =$

$((x_2 + dx_1)u, 0)$ and $(s_1, s_2) = (0, x_1 u)$.

Notice that $R(\delta, u) = 0$. Assume there exists $\mathcal{U} \subset \mathbb{R}$ with nonzero Lebesgue measure such that for all $u \in \mathcal{U}$ and all $d \neq \delta$

$$R(d, u) \neq 0$$

The coefficient $\delta \neq 0$ is identified as the unique solution to

$$\delta = \operatorname*{argmin}_{d \in \mathbb{R}} \int_{\mathcal{U}} \left(R(d, u)\right)^2 w(u) du$$

where $w(u)$ is a weight function that satisfies $\int_{\mathcal{U}} w(u) du = 1$.

In Equation (3.5) set $s_1 = u$, $s_2 = u(1-\delta)$ then by the equality in distribution assumption $\varphi_{\varepsilon_1|X}(u) = \varphi_{\varepsilon_2|X}(u)$. Hence,

$$\begin{pmatrix} \left.\frac{\partial \varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{u(1-\delta),u} \\ \left.\frac{\partial \varphi^2_{Y|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{u(1-\delta),u} \\ \left.\frac{\partial \varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2}\right|_{u(1-\delta),u} \end{pmatrix} = \begin{pmatrix} x_1^2 & \delta^2 & 1 \\ (x_2 + \delta x_1) x_1 & \delta^3 & \delta \\ (x_2 + \delta x_1)^2 & \delta^4 & \delta^2 + 1 \end{pmatrix} \begin{pmatrix} \varphi''_{\beta_1|X}((x_1 + x_2)u) \\ \varphi''_{Y_0|X}(\delta u) \\ \varphi''_{\varepsilon_1|X}(u) \end{pmatrix}$$

Assume $x_1 \neq 0$, $x_2 \neq 0$, and $\delta \neq 0$. Then

$$\begin{pmatrix} \varphi''_{\beta_1|X}((x_1 + x_2)u) \\ \varphi''_{Y_0|X}(\delta u) \\ \varphi''_{\varepsilon_1|X}(u) \end{pmatrix} = \begin{pmatrix} -\frac{\delta}{x_1 x_2} & \frac{1}{x_1 x_2} & 0 \\ \frac{-\delta^2 x_1 x_2 + \delta x_1^2 - \delta x_2^2 + x_1 x_2}{\delta^2 x_1 x_2} & \frac{-x_1^2 + 2\delta x_1 x_2 + x_2^2}{\delta^2 x_1 x_2} & -\frac{1}{\delta^2} \\ \frac{x_1 \delta^2 + x_2 \delta}{x_1} & -\frac{x_2 + 2\delta x_1}{x_1} & 1 \end{pmatrix} \begin{pmatrix} \left.\frac{\partial \varphi^2_{Y|X}(s_1,s_2)}{\partial s_1^2}\right|_{u(1-\delta),u} \\ \left.\frac{\partial \varphi^2_{Y|X}(s_1,s_2)}{\partial s_1 \partial s_2}\right|_{u(1-\delta),u} \\ \left.\frac{\partial \varphi^2_{Y|X}(s_1,s_2)}{\partial s_2^2}\right|_{u(1-\delta),u} \end{pmatrix}$$

Then

$$\begin{aligned} \varphi''_{\beta_1|X}(u) &= -\frac{\delta}{x_1 x_2} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_1^2}\right|_{u(1-\delta)/(x_1+x_2), u/(x_1+x_2)} + \frac{1}{x_1 x_2} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_1 \partial s_2}\right|_{u(1-\delta)/(x_1+x_2), u/(x_1+x_2)} \\ \varphi''_{Y_0|X}(u) &= \frac{-\delta^2 x_1 x_2 + \delta x_1^2 - \delta x_2^2 + x_1 x_2}{\delta^2 x_1 x_2} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_1^2}\right|_{\frac{u(1-\delta)}{\delta}, \frac{u}{\delta}} - \frac{x_1^2 - 2\delta x_1 x_2 - x_2^2}{\delta^2 x_1 x_2} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_1 \partial_2}\right|_{\frac{u(1-\delta)}{\delta}, \frac{u}{\delta}} \\ &\quad - \frac{1}{\delta^2} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_2^2}\right|_{\frac{u(1-\delta)}{\delta}, \frac{u}{\delta}} \end{aligned}$$

$$\varphi''_{\varepsilon_1|X}(u) = \frac{x_1\delta^2 + x_2\delta}{x_1} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_1^2}\right|_{u(1-\delta),u} - \frac{x_2 + 2\delta x_1}{x_1} \cdot \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_1 \partial s_2}\right|_{u(1-\delta),u} + \left.\frac{\partial \varphi^2_{Y|X}(\vec{s})}{\partial s_2^2}\right|_{u(1-\delta),u}$$

and now use the Second Fundamental Theorem of Calculus to obtain the CFs

$$\phi_{\widetilde{m}^*|X}(u) = \exp\left(\int_0^u \int_0^w \varphi''_{\widetilde{m}^*|X}(v)\,\mathrm{d}v\mathrm{d}w + iuE[\beta_{\widetilde{m}^*}|X]\right) \qquad \widetilde{m}^* = \beta_1, Y_0, \varepsilon_1$$

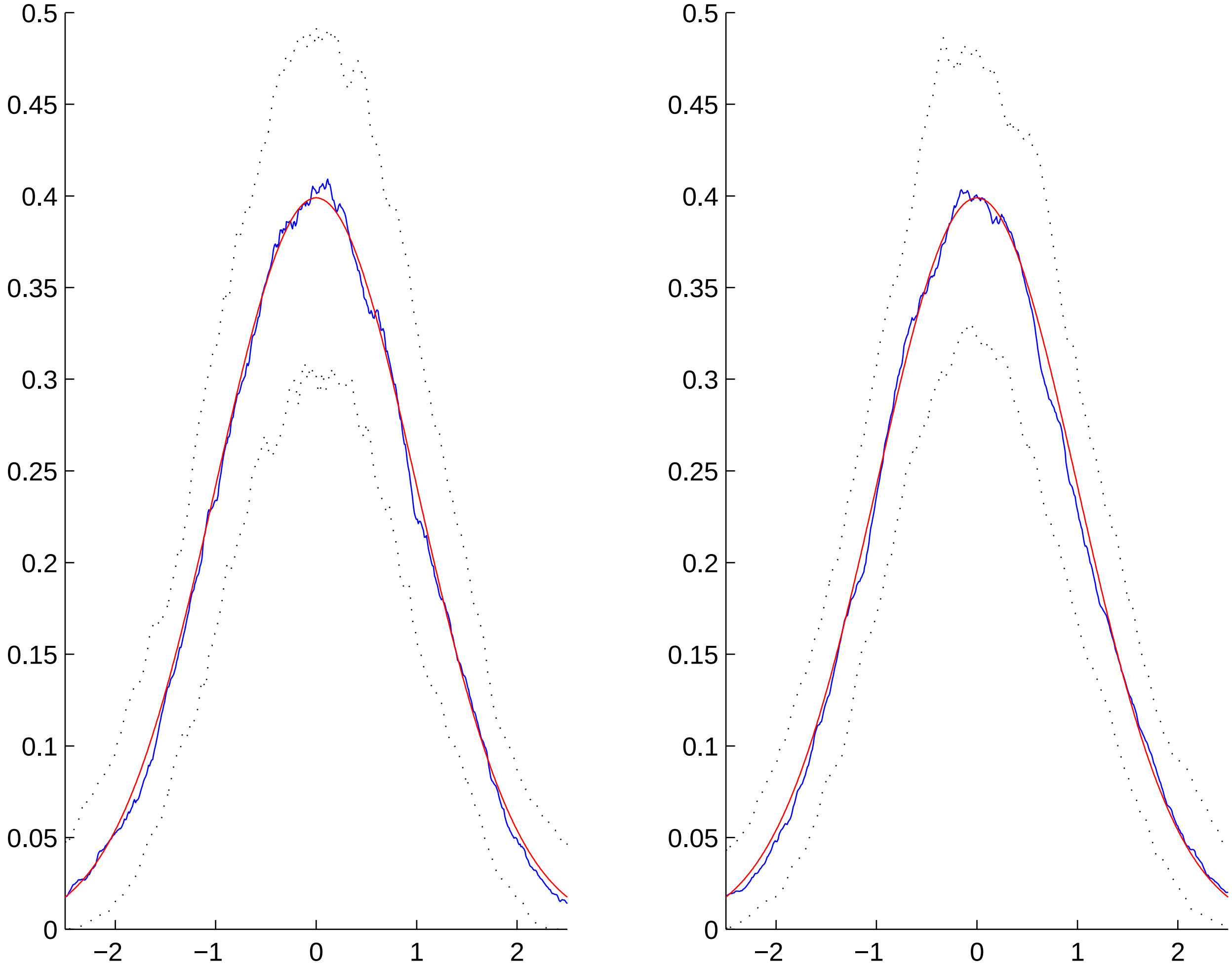

Figure 3.1: The marginal densities of β_1 and β_2 using Corollary 1
The left graph depicts the marginal distribution of β_1 and the right graph depicts the marginal distribution of β_2. The solid red lines are the underlying theoretical distributions, the solid blue lines are the medians of the estimates and the dotted black lines are the 10-90% confidence bands of the estimates. The mean squared error of the marginal density of β_1 is 0.0175. The mean squared error of the marginal density of β_2 is 0.0252.

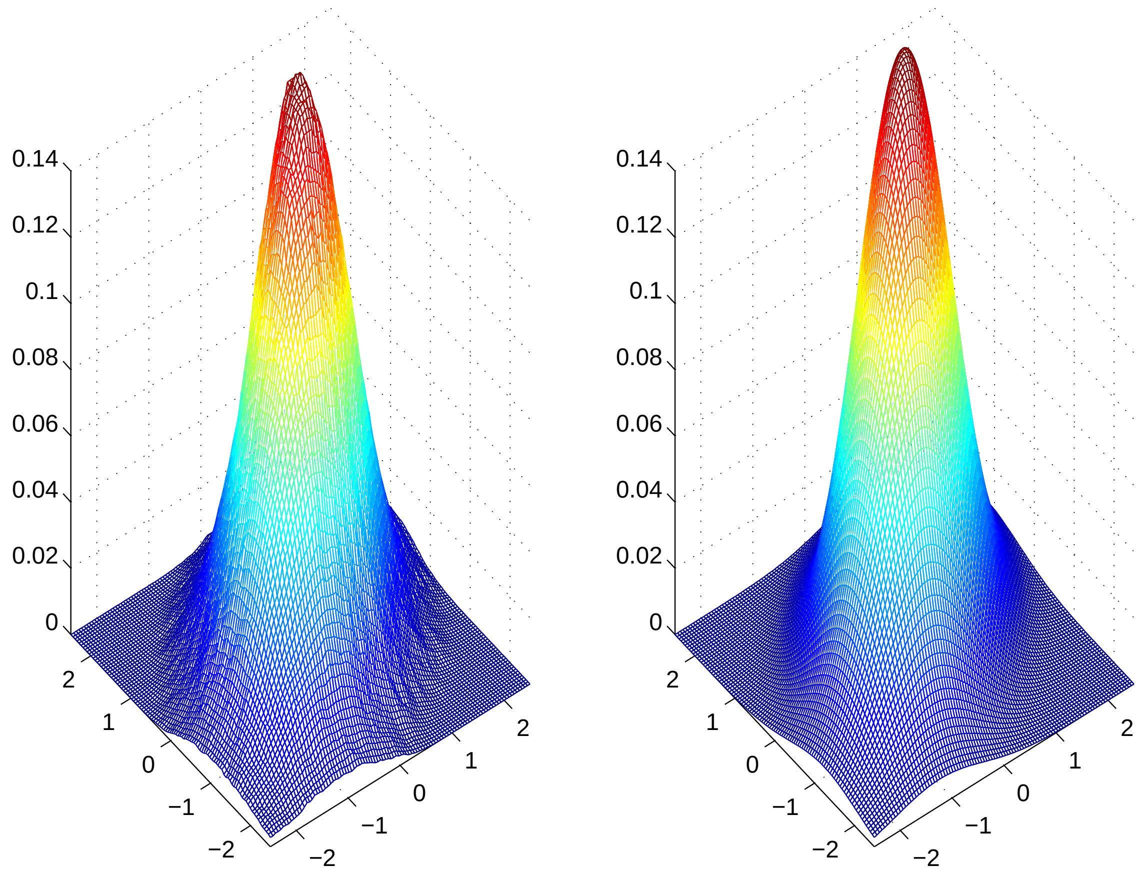

Figure 3.2: The joint density of (β_1, β_2) using Corollary 1
The left graph depicts the median of the estimates for the joint distribution of (β_1, β_2) and the right graph depicts the underlying theoretical joint distribution of (β_1, β_2). The mean squared error of the joint density of (β_1, β_2) is 0.0629.

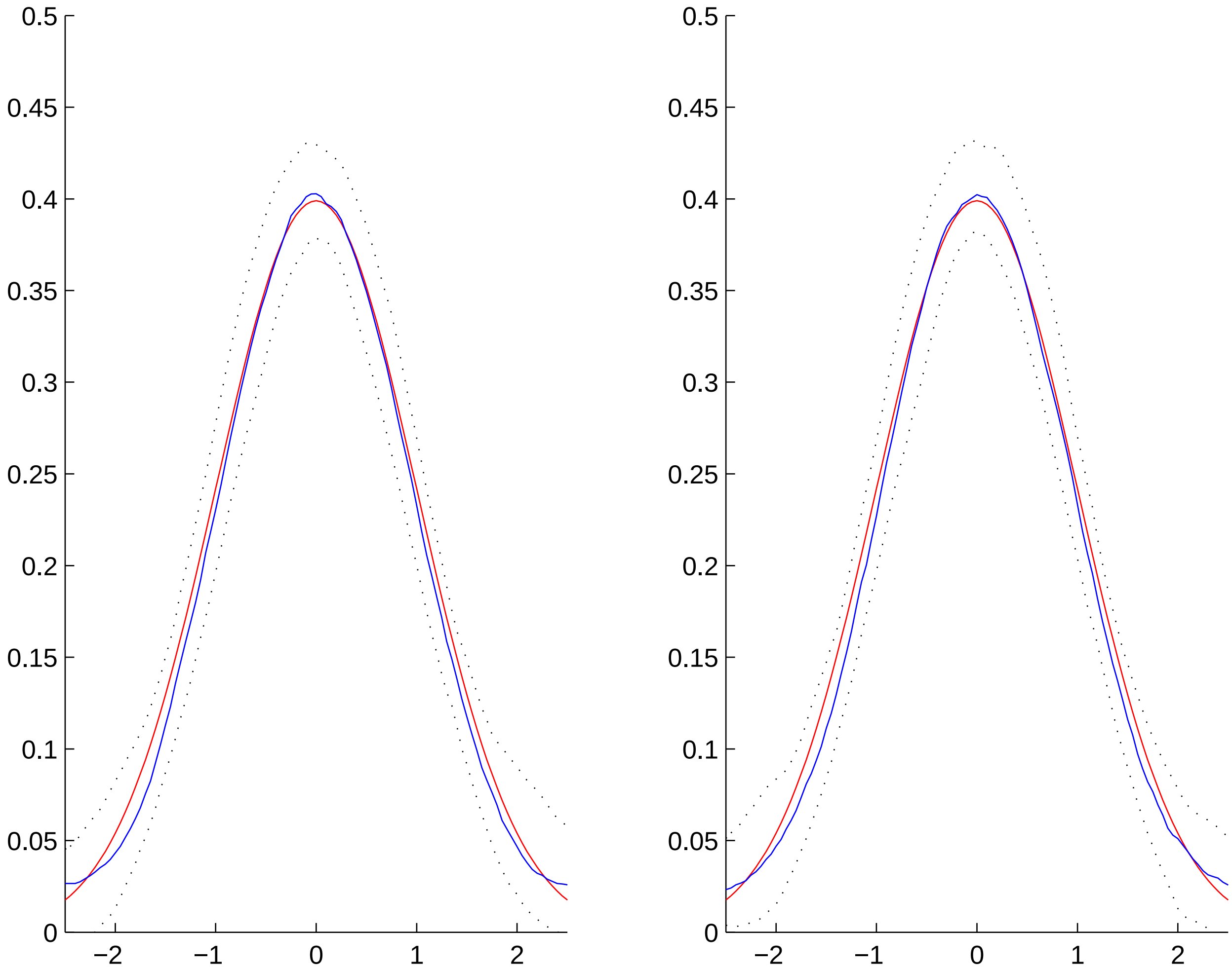

Figure 3.3: The marginal densities of β_1 and β_2 using Theorem 12
The left graph depicts the marginal distribution of β_1 and the right graph depicts the marginal distribution of β_2. The solid red lines are the underlying theoretical distributions, the solid blue lines are the medians of the estimates and the dotted black lines are the 10-90% confidence bands of the estimates. The mean squared error of the marginal density of β_1 is 0.0087. The mean squared error of the marginal density of β_2 is 0.0092.

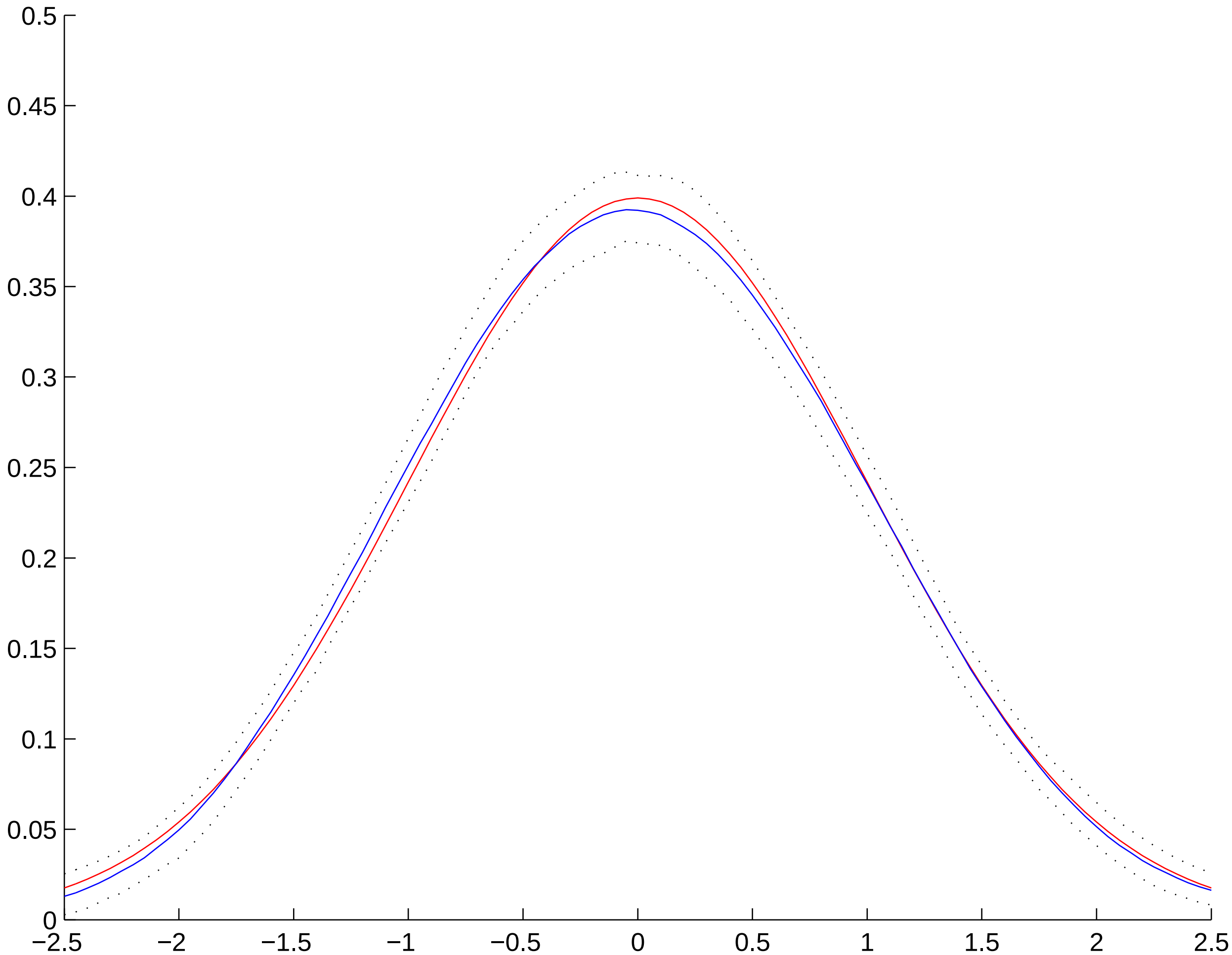

Figure 3.4: The marginal density of β using Theorem 15
The graph depicts the marginal distribution of β. The solid red line is the underlying theoretical distribution, the solid blue line is the median of the estimates and the dotted black lines are the 10-90% confidence bands of the estimates. The mean squared error of the marginal density of β is 0.0025.

Bibliography

[1] ACKERBERG, D., BENKARD, L., BERRY, S., PAKES, A. (2007), "Econometric Tools for Analyzing Market Outcomes," *Handbook of Econometrics*, Vol 6, ed by J. J. Heckman

[2] AHN, S.C. and SCHMIDT, P. (1997), "Efficient Estimation of Dynamic Panel Data Models Under Alternative Sets of Assumptions," *Journal of Econometrics*, **76**, 309-321.

[3] ALVAREZ, J. and ARELLANO, M. (2003), "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators," *Econometrica*, Vol. 71 No. 4, 1121-1159.

[4] ARELLANO, M. and BONHOMME, S. (2011), "Identifying Distributional Characteristics in Random Coefficients Panel Data Models," *Review of Economic Studies*, forthcoming.

[5] ARELLANO, M. and BOVER, O. (1995), "Another Look at the Instrumental-Variable Estimation of Error-Components Models," *Journal of Econometrics*, **68**, 29-51.

[6] ARELLANO, M. and HONORÉ, B. (2001), "Panel Data Models: Some Recent Developments," *Handbook of Econometrics*, Vol. 5, ed. by J. J. Heckman, and E. Leamer. North-Holland.

[7] ARELLANO, M. and HAHN, J. (2005), "Understanding Bias in Nonlinear Panel Models: Some Recent Developments," Invited Lecture, Econometric Society World Congress, London.

[8] BELZIL, C. and HANSEN J. (2007), "A Structural Analysis of the Correlated Random

Coefficient Wage Regression Model with an Application to the OLS-IV Puzzle" *Journal of Econometrics*, 140 (2), 333-948

[9] BERAN, R., FEUERVERGER A., and HALL, P. (1996), "On Nonparametric Estimation of Intercept and Slope Distributions in Random Coefficient Regression," *Annals of Statistics*, **24**, 2569-2592.

[10] BERAN, R. and HALL, P. (1992), "Estimating Coefficient Distributions in Random Coefficient Regressions," *Annals of Statistics* **20**, 19701984.

[11] BERAN, R. and MILLAR, P. (1994), "Minimum distance estimation in random coefficient regression models," *Annals of Statistics* **22**, 19761992.

[12] BESTER, C.A. and HANSEN C. (2009), "Identification of Marginal Effects in a Nonparametric Correlated Random Effects Model," *Journal of Business and Economic Statistics*, 27(2): 235-250.

[13] BESTER, C.A. and HANSEN C. (2007), "Flexible Correlated Randome Effects Estimation in Panel Models with Unobserved Heterogeneity," working paper

[14] BOND, S. (2002), "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice," CEMMAP working paper

[15] BOND, S. and WINDWEIJER, F. (2002), "Projection Estimators for Autoregressive Panel Data Models," *Econometrics Journal*, **5**, 457-479.

[16] BONHOMME, S. and ROBIN, J.-M. (2010), "Generalized Non-Parametric Deconvolution with an Application to Earnings Dynamics," *Review of Economic Studies*, **77** (2), 491-533.

[17] BROWNING, M. and CARRO, J. (2007), "Heterogeneity and Microeconometrics Modelling," in Blundell, R., W.K. Newey, T. Persson (eds.), *Advances in Theory and Econometrics*, Vol. 3 ; Cambridge: Cambridge University Press.

[18] CARD, D. (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, Vol. 69 No. 5, 1127-1160.

[19] CARROLL, R. J., and HALL, P. (1988), "Optimal rates of Convergence for Deconvo-

luting a Density," *Journal of the American Statistical Association*, **83**, 1184-1186.

[20] CHAMBERLAIN, G (1982), "Multivariate regression models for panel data," *Journal of Econometrics* **18**, 5-46.

[21] CHAMBERLAIN, G (1984), "Panel data," *Handbook of Econometrics*, Vol 2 ed. by Z Griliches and M D Intriligator. Amsterdam.

[22] DELAIGLE, A. and GIJBELS, I. (2002), "Estimation of Integrated Squared Density Derivatives from a Contaminated Sample," *Journal of the Royal Statistical Society, Series B*, **64**, 869-886.

[23] EVDOKIMOV, K. (2011), "Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity," working paper.

[24] FAN, J. Q. (1991), "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *Annals of Statistics*, **19**, 1257-1272.

[25] FOSTER, A. and HAHN, J. (2000) "A Consistent Semiparametric Estimation of the Consumer Surplus Distribution" *Economics Letters*, **69**, 245-251.

[26] GAUTIER, E. and KITAMURA, Y. (2009), "Nonparametric Estimation in Random Coefficients Binary Choice Models," *Econometrica*, forthcoming

[27] GRAHAM, B.S. and POWELL, J.L. (2011), "Identification and Estimation of Irregular Correlated Random Coefficient Models," Working Paper

[28] HAUSMAN, J.A. and TAYLOR, W.E. (1983), "Identification in Linear Simultaneous Equations Models with Covariance Restrictions: An Instrumental Variables Interpretation," *Econometrica*, **51**, 1527-1549.

[29] HECKMAN, J. MATZKIN, R. L. and NESHEIM, L. (2010), "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, **78**, 1569-1591.

[30] HODERLEIN, S., KLEMELA, J., and MAMMEN, E. (2010), "Reconsidering the Random Coefficient Model," *Econometric Theory*, **26(3)**, 804-837.

[31] HSIAO, C. and PESARAN, M.H. (2008), "Random Coefficients Models," *The Econometrics of Panel Data*, 3rd edition, edited by L. Matayas and P. Sevestre, Kluwer,

187-216.

[32] KHATRI, C. G. and RAO, C. R. (1968), "Solutions to Some Functional Equations and their Applications to Characterization of Probability Distributions," *Sankhyä*, **30**, 167-180.

[33] KHATRI, C. G. and RAO, C. R. (1972), "Functional Equations and Characterization of Probability Laws Through Luinear Functions of Random Variables," *Journal of Multivariate Analysis*, **2**, 162-173.

[34] KOTLARSKI, I. (1967), "On Characterizing the Gamma and Normal Distribution," *Pacific Journal of Mathematics*, **20**, 69-76.

[35] LI, T. and VUONG, Q. (1998), "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, **65**, 139-165.

[36] MADDALA, G. S. (1971), "The Use of Variance Components Models in Pooling Cross Section and Time Series Data," *Econometrica*, **39**, 341-358.

[37] MCFADDEN, D. (1974), "The Measurement of Urban Travel Demand," [PDF file, 1.8M] *Journal of Public Economics*, Vol. 3, No. 4, 303-328

[38] MURTAZASHVILI, I. (2007), "An Alternative Measure of Intergenerational Income Mobility Based on a Random Coefficient Model, *Journal of Applied Econometrics*,

[39] MUNDLAK, Y. (1961), "Empirical Production Function Free of Management Bias," *Journal of Farm Economics*, **43**, 44-56.

[40] MUNDLAK, Y. (1978), "On the Pooling of Time Series and Cross Section Data," *Econometrica*, **33**, 265-270.

[41] RAO, C. R. (1971), "Characterization of Probability Laws by Linear Functions," *Sankhyä*, **33**, 265-270.

[42] SWAMY, P.A.V.B. (1970), "Efficient Inference in a Random Coefficient Regression Model," **38**, 2, 311-323.

[43] SZÉKELY, G. J. and RAO, C. R. (2000), "Identifiability of Distributions of Independent Random Variables by Linear Combinations and Moments," *Sankhyä*, **62**, 193-202.

[44] WOOLDRIDGE, J.M. (2005) ‘Fixed Effects and Related Estimators in Correlated Random Coefficient and Treatment Effect Panel Data Models,” *Review of Economics and Statistics*, **87** (2), 385-390.