An Examination of Initial Condition Specification in the Structural Equations Modeling Framework

Diane Losardo

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology (Quantitative).

Chapel Hill
2012

Approved by

Sy-Miin Chow, Ph.D.

Patrick J. Curran, Ph.D.

Stephen du Toit, Ph.D.

Andrea Hussong, Ph.D.

Robert C. MacCallum, Ph.D.

ABSTRACT

DIANE LOSARDO: An Examination of Initial Condition Specification in the Structural Equations Modeling Framework
(Under the direction of Sy-Miin Chow)

A challenge when estimating time series models is deciding how to correctly specify an initial condition distribution, which describes the process prior to the first sampled observation. For example, the process may have started in the distant past, started exactly at the first observation point, or displayed a different structure before the first time point was collected. Time series models may be estimated within the structural equations modeling (SEM) framework (Browne & Nesselroade, 2005), and while some psychological research has focused on the issue of initial condition specification (e.g., Du Toit & Browne, 2007; Chow, Ho., Hamaker, & Dolan, 2010; Oud, Bercken, & Essers, 1990), a thorough examination of the consequences of using a misspecified initial condition distribution has not been conducted. As the number of time points increases, the consequences of a misspecified initial condition become less severe (i.e., parameter estimates and state estimates are not affected as noticeably, see Oud et al., 1990). If a process is not stationary (i.e., does not have the same mean and covariance structure over time), then conventional methods for initial condition specification may not be appropriate (De Jong, 1991; Harvey, 1991). Proper methods for such cases have been developed in the state-space literature (De Jong, 1991; Koopman, 1997). In this thesis I conducted a systematic examination comparing initial condition specifications for time series models estimated within the SEM framework. For stationary models, I considered three approaches, including (1) a model-implied initial condition, (2) a free-parameter condition, (3) and a null initial condition specification. For nonstationary models, I considered De Jong's augmented filtering ap-

proach (De Jong, 1991), which consists of augmenting the standard Kalman filter (KF) with computational products associated with nonstationary portions of the model, a modification by Koopman (1997), who developed an exact initial KF approach which removes the reliance of the filtering equations on the nonstationary portions, and a large κ approximation which is widely used in the time series literature but may lead to numerical inaccuracies. A Monte Carlo simulation was conducted to examine parameter and state estimate recovery given different types of initial condition specification for both intensive repeated measures data and panel data. Finally, each initial condition specification was estimated using an empirical data set.

Results suggest that, when the process is nonstationary and true initial condition is diffuse, the de Jong initial condition approach leads to proper point and standard error estimates with fewer numerical difficulties when compared to the other approaches. However, the Koopman and free-parameter approaches also performed well, but exhibited more severe computational problems in the estimation process, leading to convergence problems and biased point estimates of the variance parameters. Furthermore, results illustrate how using different initial condition specifications with real data may lead to different point and standard error estimates and thus different substantive conclusions. Implications of results and recommendations for practice are highlighted in the discussion section.

To Nana and Buddy.

ACKNOWLEDGEMENTS

I wish to thank everyone who has helped and supported me throughout this dissertation process. First, I thank my advisor, Sy-Miin Chow, for her perpetual guidance, support, and encouragment. I am grateful for the many hours she spent with me explaining difficult concepts and providing me with useful resources while allowing me to pursue research on such interesting topics.

I thank Patrick Curran, who provided me with the foundational knowledge for understanding key concepts in psychometrics. His mentorship was valuable not only in the process of my dissertation, but also in preparing me academically and mentally for the completion of such a task.

I also thank the rest of my dissertation committee members - Stephen du Toit, Andrea Hussong, and Bud MacCallum - for their helpful suggestions and creative ideas throughout this process. I thank all members of the L.L. Thurstone lab for providing me with such a great atmosphere for learning. I thank the former members of the Curran Army, Daniel, Jim, and Taehun, for providing me with so much knowledge and laughter.

I thank my family for support and encouragement throughout the years of this process. I am grateful for my friends who kept me sane when times were tough - thank you Kristi, Ben, Allison, Dana, Jolynn, Susan, Ellen, Andrea, and last but not least, John, whose support in the last few months leading up to my dissertation defense was invaluable.

Finally, I thank my cats, Leon, Booty, and Artemis, for being such wonderful companions and such excellent stress relievers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

While initial condition specification has not been extensively discussed within the psychological literature, there are both methodological and substantive reasons as to why it is an important topic to consider. Generally speaking, an initial condition distribution describes the structure of a process *before* the first observation was collected, or right on the first occasion. As an example, consider a situation where a psychologist has collected longitudinal data for a given time frame. A model may then be fit to the data in hopes of uncovering and understanding some underlying process. However, it is very possible that the process has started in the distant past, and furthermore that the structure of the process was different before the first collected data point. It is also possible that the overall process displays differing structures as a function of time, thus rendering it difficult to discern what the initial condition process was.

From a substantive viewpoint, the role of initial condition in certain psychological theories plays a major part. Several psychological theories concerning initial condition have been proposed within the context of chaotic processes that are highly nonlinear in nature, but are actually deterministic (i.e., a process that can be perfectly predicted given current and previous information) (see Robertson & Combs, 1995). Such a process has been hypothesized to exist in several areas of psychology, including developmental psychopathology, neuropsychology, cognitive psychology,

and developmental psychology (Ayers, 1997). One element of a chaotic process, as applied to explain a psychological process, is the idea that a process is very sensitive to its initial condition status. Consequently, small changes in initial condition may produce large effects over time. As an example, Gottschalk, Bauer, and Whybrow (1995) found evidence indicating that long-term mood variations for both bi-polar patients and normal patients are not periodic in nature, and may display some form of chaotic behavior.

Some theories in developmental psychology posit that early childhood events, which may be loosely thought of as initial conditions for a given individual, have a significant and long-lasting effect on an individual's trajectory into adulthood. Classic examples include attachment theory (Bowlby, 1982), social-cognitive theories (e.g., Vygotsky, 1978), and social-cognitive learning theory (Bandura, 1986). As cognitive growth tends to be rapid during early childhood, any environmental or biological influences may be more pronounced and affect the eventual progression to a more stable cognitive style (Sternberg, 1999). An individual's social development and social engagement have also been hypothesized to display rapid growth at early ages, and thus be sensitive to any internal or external events that eventually shape the person's more stable social behavior. Early psychological trauma, such as child maltreatment, may have a long term effect on a child's functioning with respect to both brain development (Cicchetti & Tucker, 1994) and psychological functioning (Maughan & Cicchetti, 2002). For example, Greenberg, Kusche, and Spelz (1991) discussed how children's early exposure to either witnessing or being victims of violence along with a display of negative affect may hinder a child's ability to effectively manage and process emotions. This may, in turn, lead to a larger risk of psychopathology in adulthood. In the substance use literature, early engagement in problem behaviors is hypothesized to predict early onset of alcohol use which in turn predicts later substance use diagnosis (Caspi, Moffitt, Newman, & Silva, 1996; Jessor, Donovan, &

Costa, 1991; Zucker, Chermack, & Curran, 2000).

Such theories indicate that the role of initial condition in psychological research is important to consider and may help shape a developmental pathway into adulthood. It is not always the case, however, that a researcher will know how to characterize the initial condition distribution. Consider the case where a researcher collects data on a sample of individuals with a certain anxiety disorder. One measured variable may be the degree to which a person displays a symptom of anxiety. If this variable is tracked over the course of a few weeks, a trajectory of anxiety levels for the sampled individuals will be revealed. However, it is unlikely that the given individuals began to display such a process at the exact time the first data point was collected. This process may have started long before the data were collected, and may even have displayed a different structure beforehand. In other words, the initial condition distribution may be characterized differently than the distribution of the process for the sampled time frame.

One way psychologists have attempted to explain initial condition status is by introducing covariates, such as gender, socioeconomic status, and age, to help explain interindividual differences in initial condition. One example is adding covariates into a growth curve model to explain interindividual differences in intercept. However, as stated, this prior knowledge is not always available or exhaustive. If there is no precise information available researchers may have to resort to other specifications of initial condition. In this thesis, I consider some of these possible specifications. I describe different procedures that have been used for initializing a process and evaluate a set of procedures for more accurately specifying initial condition, with a focus on types of data that are commonly collected in psychological research. I first describe stationary and nonstationary processes in more detail as such processes are integral in informing what type of initial condition to specify.

1.1 Stationary vs. Nonstationary Processes

To begin, I will introduce the more technical definitions that describe and categorize a process. Specifically, processes may be generally described as either stationary or nonstationary (Chatfield, 2004; Hamilton, 1994; Shumway & Stoffer, 2006; Wei, 1990), which are terms that arise from the time series literature. Broadly speaking, a stationary process implies that the statistical properties, including mean, variance, and higher order moments, of a process do not change over time regardless of when data were sampled. A nonstationary process, in contrast, may display a different process depending on a given time frame. This includes time frames that occurred before a given data point was sampled. Depending on the type of stationarity, different initial condition distributions will be needed to properly specify the model and to ensure that accurate latent variable scores and parameter estimates are obtained.

Time series processes and stationarity. To more extensively describe the stationarity of a process, I will illustrate how it is explained within the context of time series processes. In general, time series models allow for the examination of the relations between series of random variables indexed by time (Chatfield, 2004; Hamilton, 1994; Shumway & Stoffer, 2006; Wei, 1990). Specifically, a multiple-subject time series may be represented by the following equation:

$$x_{it} = g_{it} + v_{it} \tag{1.1}$$

where x_{it} represent outcome variables for a given person i ($i = 1, 2 \ldots N$) and time point t ($t = 1, 2 \ldots T$), g_{it} represents a time series process, such as an autoregressive (AR) or moving average (MA) process, and v_{it} represents noise or uncertainties in the change process, such as a white noise process (i.e., a process of uncorrelated variables with means of zero and some finite variance). A goal in time series analysis is to model the systematic process, g_{it}, and separate out the noise, v_{it}.

Browne and Nesselroade (2005) referred to g_{it} as the process variables and v_{it} as a series of random shocks (i.e.,"dynamic noise," often assumed to be white noise). The process variables are those that are changing over time, while the random shock variables are forces that cause uncertainties in the process variables. Two examples of time series processes are an AR(p) process and an MA(q) process

$$\text{AR(p): } x_{it} = \alpha_1 x_{i,t-1} + \alpha_2 x_{i,t-2} + ... + \alpha_p x_{i,t-p} + v_{it} \qquad v_{it} \sim N(0, \sigma_v^2) \tag{1.2}$$

$$\text{MA(q): } x_{it} = \beta_1 v_{i,t-1} + \beta_2 v_{i,t-2} + ... + \beta_q v_{i,t-q} + v_{it} \qquad v_{it} \sim N(0, \sigma_v^2) \tag{1.3}$$

where the x_{it} variables represent the process variables and the v_{it} variables represent the random shocks that are uncorrelated with each other over time and distributed normally with a mean zero and variance σ_v^2.

Multiple-subject time series models may be extended to include nx-vector (multivariate) autoregression with moving average (VARMA) processes (Shumway & Stoffer, 2006), where nx equals the number of x variables. In this case, the model is formulated as

$$\mathbf{x}_{it} = \sum_{r=1}^{p} \mathbf{A}_r \mathbf{x}_{i,t-1} + \sum_{j=1}^{q} \mathbf{B}_j \mathbf{v}_{i,t-j} + \mathbf{v}_{it} \tag{1.4}$$

where $\mathbf{x}_{it}$ represents a set of $nx \times 1$ vectors of process variables, $\mathbf{u}_{it}$ represents a set of $nx \times 1$ vectors of random shock variables which are again normally distributed with a mean vector of zero and covariance matrix Σ_v, $\mathbf{A}_1, \ldots, \mathbf{A}_p$ represent the $nx \times nx$ autoregression weight matrices, and $\mathbf{B}_1, \ldots, \mathbf{B}_q$ represent the $nx \times nx$ moving average weight matrices. There are now a total of $N \times T \times nx$ observations. To technically define a stationary process, consider the backshift operator B,

$$B\mathbf{x}_{it} = \mathbf{x}_{i,t-1} \tag{1.5}$$

such that the VARMA process may be re-written as

$$(\mathbf{I} - \mathbf{A}_1 B^1 - \ldots - \mathbf{A}_p B^p)\mathbf{x}_{it} = \mathbf{v}_{it}. \tag{1.6}$$

Then, the process is mathematically defined as stationary if the roots of the determinant of

$$(\mathbf{I} - \mathbf{A}_1 B^1 - \ldots - \mathbf{A}_p B^p) \tag{1.7}$$

lie outside of the unit circle (i.e., all roots have a modulus, or absolute value, greater than 1, Ltkepohl, 1993). This model may be particularly attractive to psychologists as they allow for a focus on understanding intra-individual (i.e., within person) changes while still retaining a multivariate orientation.

As the variables are dependent on each other with respect to time, in order to fully understand the concept of stationarity we must first understand the autocovariance function of a process. This function is a measure of the linear dependence between two observations from the same time series, measured as

$$\mathbf{\Gamma}(s,t) = E[(\mathbf{x}_{i,s} - \boldsymbol{\mu}_s)(\mathbf{x}_{i,t} - \boldsymbol{\mu}_t)] \tag{1.8}$$

where $s - t$ represents an arbitrary time lag. The smoother a time series is, the larger the autocovariance value will be. If the autocovariance function is zero, then there exists no linear dependency among $\mathbf{x}_{i,s}$ and $\mathbf{x}_{i,t}$, and the series will be choppier. Note that this function applies to all individuals within the time series.

Inherent within time series is the idea that there is some regularity to be modeled plus some noise, or "error." This regularity is tested by making use of the concept of stationarity. Stationarity implies the stability of a process over time, although stability of a process over time does not necessarily mean a series is stationary. This can

be formally expressed in mathematical terms by considering processes that display strong stationarity, weak stationarity, or nonstationarity.

Strong stationarity asserts that the probabilistic behavior of every set of observations remains the same when there is an arbitrary shift in time, expressed as $t + h$. Formally, this states that every set of observations $\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{ik}$ will have the same joint statistical distribution as a time shifted set $\mathbf{x}_{i,1+h}, \mathbf{x}_{i,2+h}, \ldots, \mathbf{x}_{i,k+h}$. Thus,

$$P(\mathbf{x}_{i1} \leq \mathbf{c}_1, \ldots, \mathbf{x}_{ik} \leq \mathbf{c}_k) = P(\mathbf{x}_{i,1+h} \leq \mathbf{c}_1, \ldots, \mathbf{x}_{i,k+h} \leq \mathbf{c}_k) \tag{1.9}$$

holds for all $k = 1, 2, \ldots$, all time points $t_1, t_2, \ldots, t_k$, all constants $c_1, c_2, \ldots, c_k$, and all time shifts $h = 0, \pm 1, \pm 2, \ldots$. This implies that the probability of observing a value at a given time period is the same as the probability of observing the value at a later or earlier time period. Also, the mean function does not change with time, so that $\mu_t = \mu_s = \mu$ for all s and t. Any process that involves a mean change over time is then not stationary. Additionally, the variance function, which is the autocovariance of a variable with itself, does not change with time. The autocovariance of a stationary process can be defined as:

$$\mathbf{\Gamma}(s, t) = \mathbf{\Gamma}(s + h, t + h) \tag{1.10}$$

Thus, the covariance depends on the lag, or time difference between s and t, only and not on the actual times. This strong stationarity incorporates all of the possible distributions of a time series. Thus, not only are the first two moments stationary, but all possible characteristics of the distribution, including higher-order moments. This assumption is very strict and difficult to meet, especially in data commonly collected in psychological research.

A more relaxed version of stationarity is termed weak stationarity. Instead of making assumptions about the entire probability distribution, this version focuses on the first two moments. Specifically, the mean and variance structures remain constant

across time points and the covariance structure depends only on the lag $s - t$, and given this lag remains constant over time. Also, if a process arises from a multivariate normal distribution and is weakly stationary, it is considered to formally exhibit a stationary process. To illustrate these points more concretely, I will now provide examples of both stationary and nonstationary processes.

Examples of stationary processes. One example of a model with a stationary process is an AR(1) model where $|\alpha_1| < 1$. A psychological process that might align well with this model could be an individual's overall mood sampled daily over the course of a few weeks. The model testifies that a person's current mood is affected by his or her mood on a previous day, and to a lesser extent by his or her mood two days ago, and to an even lesser extent by his or her mood three days ago, and so forth. This can be understood more precisely by examining the autocovariance function for an AR(1) process, expressed as

$$\frac{\sigma_v^2 \alpha^{s-t}}{1 - \alpha^2}. \tag{1.11}$$

The numerator of this term expresses the AR(1) weight α as being raised to the power of $s - t$, which represents a time lag. Thus, if $|\alpha_1| < 1$, larger lags in time will be characterized by an autocovariance closer to zero. In other words, the covariance between any two observations is attenuated exponentially (i.e., decays) as the two observations are further apart in time, either with (when $-1 < \alpha_1 \leq 0$) or without (when $0 \leq \alpha_1 < 1$) fluctuations. Thus, the linear dependence between a person's mood on, say, day 1 as compared to day 10 is lower than the linear dependence of a person's mood on day 1 as compared to day 3.

Another feature of a stationary AR(1) model where $|\alpha_1| < 1$ is that the random shock variables only affect the current time point process variables *directly*. Thus, a person's mood today is also affected by this random shock variable, which are

uncertainties that may be different from day to day. For example, a person's anxiety may be affected today by a pop-quiz in his or her calculus class and tomorrow by an unannounced visit from a family member. Still, the process is stationary in that it ensures that large effects of a process variable will become less pronounced with time instead of being intensified over time.

Another example of a stationary process is a MA(1) model, which is mathematically expressed as:

$$x_{it} = \beta_1 v_{i,t-1} + v_{it}. \tag{1.12}$$

In this model, the random shock variables for a given time point directly affect the process variable for that time point but also directly affect the process variable for the subsequent time point. A psychological process that might align well with this model is one where random day-to-day environmental or external influences, such as a rainy day or a pop quiz, are hypothesized to directly influence a person's current mood and also his or her mood the following day. However, the affect of the external influence will be gone the third day.

This process has an expected value of zero, a variance of $(1 + \beta^2)\sigma_v^2$ and an autocovariance function of:

$$\mathbf{\Gamma}(s,t) = \begin{cases} \beta\sigma_v^2, & \text{if } (s-t) = 1 \\ 0, & \text{if } (s-t) > 1 \end{cases} \tag{1.13}$$

Thus, the process is stationary as time does not appear in any of these equations. As discussed, however, not all processes are stationary. I will now discuss several models that describe nonstationary processes.

Examples of nonstationary processes. A nonstationary process is one that allows the behavior, including the mean and covariance structures, of a process to change given

an arbitrary shift in time. The process may become "explosive" in the sense that the effects of current process variables are magnified with time as opposed to decaying with time. An example of this process is the random walk model, which has the structure of an AR(1) process with $\alpha = 1$, expressed as

$$x_{it} = x_{i,t-1} + v_{it}. \tag{1.14}$$

In this model, the value of the x_i variable at time t is the value of the x_i variable at time $t-1$ plus a random movement determined by v_{it}. This process is non-stationary as the variance of the random walk is equal to $t\sigma_v^2$, which increases without bound as time increases. The autocovariance function is equal to $\min(s,t)\sigma_v^2$ and depends on the particular time values, s and t, and not just the time lag $s-t$. Finally, the mean function changes also changes with time. Another nonstationary process is a random walk with drift model, which is expressed as

$$x_{it} = \nu + x_{i,t-1} + v_{it} \tag{1.15}$$

where ν is a constant called the "drift", or stochastic trend. This process is nonstationary with respect to its mean structure, which is equal to νt. The mean of this process, then, increases as t increases. This may be more clearly illustrated by rewriting the random walk with drift process as a cumulative sum of the random shock variables, expressed as

$$x_{it} = \nu t + \sum_{j=1}^{T} v_{ij}. \tag{1.16}$$

Here we can see that the past influences of the process accumulate over time and the drift term increases as t increases. A psychological process that may follow this is a person's emotional stability over the course of a series of stressful events. With

time, this person might become more and more emotionally unstable, as the original effects of instability do not diminish with time, but rather accumulate with time. Also, as time increases, so does the variability of a person's emotional state, as the person becomes more and more unstable.

Another model that describes a nonstationary process is the latent curve model (LCM; Meredith & Tisak, 1990), which is a longitudinal model estimated within the structural equation modeling (SEM; see Bollen, 1989) framework. This model describes the process via a set of latent factors that are theorized to have given rise to set of repeated measures. For example, if a linear trend is hypothesized, the model implies that the process may be explained in part by an intercept latent factor (i.e., initial starting point) and linear slope latent factor. Furthermore, individual differences may be captured by allowing the two latent factors to have a group mean plus variability around this group mean. A larger degree of variability indicates that there are more individual differences in either the intercept or slope.

While psychologists sometimes implement such models, they may be unaware that the model is assuming a nonstationary process. To more clearly understand why such models are nonstationary consider the following modeling equations that comprise a linear LCM:

$$\begin{aligned} y_{it} &= x_{INTi} + TIME_t x_{SLPi} + u_{it} \\ x_{INTi} &= \mu_{x_{INT}} + v_{x_{INTi}} \\ x_{SLPi} &= \mu_{x_{SLP}} + v_{x_{SLPi}} \end{aligned} \tag{1.17}$$

where y_{it} represent the repeated measures variables for person i at time t, $TIME_t$ represents the specific value of time at time t, x_{INTi} is the random intercept term with a mean of $\mu_{x_{INT}}$ and individual-specific disturbance term $v_{x_{INTi}}$, x_{SLPi} is the random slope term with a mean of $\mu_{x_{SLP}}$ and individual-specific disturbance term

$v_{x_{SLPi}}$, and u_{it} are measurement error variables that are uncorrelated with each other, have means of zero, and a variance of σ_u^2 that remains the same over time. Also, the disturbance terms $v_{x_{INTi}}$ and $v_{x_{SLPi}}$ are multivariate normally distributed with means of zero and covariance matrix

$$\Sigma_v = \begin{bmatrix} \sigma^2_{v_{x_{INT}}} & \sigma_{v_{x_{INT}x_{SLP}}} \\ \sigma_{v_{x_{SLP}x_{INT}}} & \sigma^2_{v_{x_{SLP}}} \end{bmatrix}. \tag{1.18}$$

The expected value and variance of y_{it} are expressed as

$$E(y_{it}) = \mu_{x_{INT}} + TIME_t \mu_{x_{SLP}} \tag{1.19}$$

$$\text{VAR}(y_{it}) = \sigma^2_{v_{x_{INT}}} + TIME_t^2 \sigma^2_{v_{x_{SLP}}} + 2TIME_t \sigma^2_{v_{x_{INT}x_{SLP}}} + \sigma_u^2 \tag{1.20}$$

while the covariance of y_{it} and $y_{i,t+s}$ ($t \neq s$) is

$$\text{COV}(y_{it}, y_{i,t+s}) = \sigma^2_{v_{x_{INT}}} + TIME_t TIME_{(t+s)} \sigma^2_{v_{x_{SLP}}} + (TIME_t + TIME_{(t+s)}) \sigma_{v_{x_{INT}x_{SLP}}}. \tag{1.21}$$

From these equations we can see that the mean, variance, and covariance structures all depend on time, making the process nonstationary. Note that nonstationary processes are not necessarily more complicated than stationary processes. However, many time series modeling tools are designed to specifically handle stationary processes.

1.2 Initial Condition and Stationarity.

Thus far I have discussed the substantive importance of initial condition in psychological theories and defined stationarity with respect to time series processes. I will now give a more detailed explanation of an initial condition distribution and describe how it directly relates to stationarity. To begin, an initial condition distribution

may be formally defined as

$$\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}_0, \mathbf{P}_0) \tag{1.22}$$

where $\boldsymbol{\mu}_0$ is the mean vector and $\mathbf{P}_0$ the covariance matrix of the initial condition. To better understand the role of initial condition, consider a multiple-subject univariate AR(1) model:

$$x_{it} = \alpha_1 x_{i,t-1} + v_{it} \qquad\qquad v_{it} \sim N(0, \sigma_v^2). \tag{1.23}$$

Since the current process variables are affected by previous process variables, the model can be iterated backwards:

$$\begin{aligned} x_{it} &= \alpha_1 x_{i,t-1} + v_{it} \\ x_{i,t-1} &= \alpha_1 x_{i,t-2} + v_{i,t-1} \\ x_{it} &= \alpha_1(\alpha_1 x_{i,t-2} + v_{i,t-1}) + v_{it} \\ x_{it} &= \alpha_1(\alpha_1(\alpha_1 x_{i,t-3} + v_{i,t-2}) + v_{i,t-1}) + v_{it}. \end{aligned} \tag{1.24}$$

This process can recur indefinitely if one continues to go back in time. Going back in time b times leads to the following,

$$x_{it} = \alpha_1^b x_{i,t-b} + \sum_{j=0}^{b-1} \alpha_1^j v_{i,t-j}. \tag{1.25}$$

Now we can clearly seen that if $|\alpha| > 1$, the process will explode, i.e., the effect of the process variable will be magnified with time, and thus will not be stationary. Also, the model cannot go back in time indefinitely, thus there needs to be an initial starting point. In order to initiate this otherwise infinite recursion, a distribution for $\mathbf{x}_{i0}$ needs to be specified.

For a stationary process, the unconditional mean vector and covariance matrix

may be used as the initial condition distribution (Harvey, 1991). This is because a stationary process states that the mean and covariance structures do not change with time. For example, a stationary AR(1) process may have an initial condition distribution of

$$x_{i0} \sim N\left(0, \frac{\sigma_v^2}{1-\alpha_1^2}\right). \tag{1.26}$$

This holds because for the stationary AR(1) process the mean is equal to zero and the variance is equal to $\frac{\sigma_v^2}{1-\alpha_1^2}$. If a process is nonstationary, or even if it is hypothesized that the process prior to the first observation has a different structure, the unconditional mean and covariance matrix may not be the best choice for specifying the initial condition distribution.

There are a number of methods for specifying initial condition (the unconditional mean vector and covariance matrix specification is just one), some of which apply to stationary models and some that apply to nonstationary models. As psychologists are currently implementing models incorporating time series processes (Hamaker & Dolan, 2009), further research regarding the specification of initial condition becomes important. In time series analysis, nonstationary processes may be made stationary by either detrending or differencing (Shumway & Stoffer, 2006). Detrending removes a trend by subtracting an estimated trend component from the original process and working with the subsequent residuals. Differencing involves subtracting a previous observation from the current observation a specified number of times. If the trend is linear, then performing this calculation once, also called first differencing, will remove the trend. A second difference removes a quadratic trend, a third difference removes a cubic trend, and so forth. While these approaches work well with respect to making a nonstationary process stationary, in many psychological theories it is the trend that is of actual interest. For example, a trend in psychological data may reflect some sort of development which may be critical to a researcher's hypothesis.

Since the parameters that describe the development over time may be integral in understanding the psychological process, it is not always helpful to remove the trend. Therefore, when psychologists implement time series models it may be important to retain the nonstationarity of a process, rendering the proper specification of initial condition also important.

Browne and Nesselroade (2005) and Du Toit and Browne (2001, 2007) have discussed how multivariate time series models may be estimated as a longitudinal model in the SEM framework. This modeling framework is often used by psychologists as it can readily and flexibly model relations among both manifest and latent variables. Furthermore, latent variable measurement models may be easily estimated. However, when estimating any times series model in the SEM framework, several methodological issues must be addressed, including the proper specification of an initial condition distribution. While some methodologists have addressed this issue (Browne & Nesselroade, 2005; Du Toit & Browne, 2001, 2007), the focus has been on stationary processes.

A particularly flexible way to represent a time series model is to formulate them within the state-space modeling framework (Harvey, 1991). State-space models were originally derived for single-subject time series analysis but may also be used to estimate intensive repeated measures or panel data within the structural equation modeling framework. In this thesis I will formulate structural equation models containing time series processes within the state-space modeling framework. Given this set-up, I will describe how to incorporate different types of initial condition distributions. To facilitate a more technical discussion of initial condition specification given this set-up, I will first describe the state-space modeling framework in more detail followed by a description of the SEM framework.

1.3 State-Space Modeling Framework

State-space models have been more extensively used in econometrics and engineering fields of study. They represent a modeling framework which can incorporate a host of time series models, difference models, and, with appropriate constraints, differential equation models (Harvey, 1991). In the psychological literature, examples of what state-space modelers refer to as states include factor scores and true scores (Grice, 2001).

The general framework for linear state space models is as follows:

$$\text{Measurement equation: } \mathbf{y}_{it} = \boldsymbol{\tau} + \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it} \qquad \mathbf{u}_{it} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u) \tag{1.27}$$

$$\text{State or Transition equation: } \mathbf{x}_{it} = \boldsymbol{\pi} + \mathbf{T}\mathbf{x}_{i,t-1} + \mathbf{v}_{it} \qquad \mathbf{v}_{it} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_v) \tag{1.28}$$

where $\mathbf{x}_{it}$ is a $nx \times 1$ vector of latent state variables at time t, where nx now equals the number of latent variables (which are called state variables in the state-space literature), $\boldsymbol{\pi}$ is a $nx \times 1$ vector of intercepts for the transition equation, $\mathbf{T}$ is a $nx \times nx$ transition matrix which links the previous states to the current states, $\mathbf{v}_{it}$ is a $nx \times 1$ vector of process noise variables with means of zero and a $nx \times nx$ covariance matrix Σ_v, $\boldsymbol{\tau}$ is a $ny \times 1$ vector of measurement intercepts, where ny equals the number of observed variables, $\mathbf{Z}$ is a $ny \times nx$ matrix that links the measured $\mathbf{y}_{it}$ variables to the $\mathbf{x}_{it}$ states, and $\mathbf{u}_{it}$ is a $ny \times 1$ vector of measurement error variables with with means of zero and a $ny \times ny$ covariance matrix Σ_u. Further model assumptions we make include the process noise variables being uncorrelated with each other over time, the measurement error variables being uncorrelated with each other over time, and the process noise variables being uncorrelated with the measurement error variables.

This framework is very flexible and allows for a chosen number of state variables that affect each other via the transition matrix and are linked to the observed variables through the $\mathbf{Z}$ matrix. Covariates may be added to either equation. It is also possible

to incorporate higher order lags (i.e., by expanding the size of $\mathbf{x}_{it}$) and deal with missingness under assumptions of missing at random (MAR; Shafer & Graham, 2002; Little & Rubin, 1987). All of these features allow for a host of models that may be fit (Shumway & Stoffer, 2006).

Estimation in state-space models. To illustrate the role of initial condition in state space models, I will describe the estimation process in some detail. The estimation of state-space models may be accomplished using the Kalman Filter (KF; see Harvey, 1991), which is an iterative estimation procedure designed to estimate the conditional means and covariance matrix, $E(\mathbf{x}_{it}|\mathbf{Y}_{it})$, $COV(\mathbf{x}_{it}|\mathbf{Y}_{it})$, $\mathbf{Y}_{it} = \{\mathbf{y}_{ij}, j = 1, \ldots, t\}$. Also, by using by-products of the filter a raw data likelihood function known as the prediction error decomposition (PED; Schweppe, 1965) may be composed. This likelihood function can then be maximized via an optimization routine of choice to obtain a new set of parameter estimates.

The set of recursions for the KF is as follows, adopted from the modified recursions used in De Jong (2003). $\mathbf{x}_{itPRED}$ represents the expected value of the current value of $\mathbf{x}$ given $\mathbf{Y}_{i,t-1}$, where $\mathbf{Y}_{i,t-1} = \{\mathbf{y}_{ij}, j = 1 \ldots t-1\}$. $\mathbf{P}_{tPRED}$ represents the covariance matrix of $\mathbf{x}_{it}$ given $\mathbf{Y}_{t-1}$, $\mathbf{x}_{tFILT}$ represents the expected value of the current value of $\mathbf{x}$ given $\mathbf{Y}_{it}$, and $\mathbf{P}_{tFILT}$ represents the covariance matrix of $\mathbf{x}_{it}$ given $\mathbf{Y}_{it}$. The following equations initiate the recursions,

$$\mathbf{x}_{i0FILT} = \boldsymbol{\mu}_0 \tag{1.29}$$

$$\mathbf{P}_{0FILT} = \mathbf{P}_0 \tag{1.30}$$

$$\mathbf{x}_{i1PRED} = \mathbf{T}\boldsymbol{\mu}_0 \tag{1.31}$$

$$\mathbf{P}_{1PRED} = \mathbf{T}\mathbf{P}_0\mathbf{T}' + \boldsymbol{\Sigma}_v. \tag{1.32}$$

Then, for $t = 1, \ldots, T$,

$$\boldsymbol{\epsilon}_{it} = \mathbf{y}_{it} - \boldsymbol{\tau} - \mathbf{Z}\mathbf{x}_{itPRED} \tag{1.33}$$

$$\mathbf{F}_{it} = \mathbf{Z}\mathbf{P}_{tPRED}\mathbf{Z}' + \boldsymbol{\Sigma}_u \tag{1.34}$$

$$\mathbf{K}_t = \mathbf{P}_{tPRED}\mathbf{Z}'\mathbf{F}_{it}^{-1} \tag{1.35}$$

$$\mathbf{x}_{itFILT} = \mathbf{x}_{itPRED} + \mathbf{K}_t\boldsymbol{\epsilon}_{it} \tag{1.36}$$

$$\mathbf{P}_{tFILT} = \mathbf{P}_{tPRED} - \mathbf{P}_{tPRED}\mathbf{Z}'\mathbf{K}_t'. \tag{1.37}$$

$$\mathbf{x}_{i,(t+1)PRED} = \boldsymbol{\pi} + \mathbf{T}\mathbf{x}_{itFILT} \tag{1.38}$$

$$\mathbf{P}_{(t+1)PRED} = \mathbf{T}\mathbf{P}_{tFILT}\mathbf{T}' + \boldsymbol{\Sigma}_v. \tag{1.39}$$

From these equations it is clear to see how initial condition comes into play. The process must be initiated by specifying a distribution for the initial condition, which can be seen by examining the prediction equations for $\mathbf{x}_1$. The $\mathbf{K}_t$ function may be considered a variation of the Kalman gain function. The $\boldsymbol{\epsilon}_{it}$ represents a vector of innovations, namely discrepancies between $\mathbf{y}_{it}$ and $E(\mathbf{y}_{it}|\mathbf{Y}_{i,t-1})$, with $\mathbf{F}_{it}$ being the covariance function of the innovations. These can be thought of as the difference between the predicted measurement and the actual measurement (Chow et al., 2010). More generally, $\boldsymbol{\epsilon}_{it} = \mathbf{y}_{it} - E(\mathbf{y}_{it}|\mathbf{Y}_{i,t-1})$ and $\boldsymbol{\epsilon}_{it} \sim N(\mathbf{0}, \mathbf{F}_{it})$. Also notable is that $E(\mathbf{x}_{it}, \mathbf{y}_{is}) = \mathbf{0}$ for all $s < t$, which implies that the innovations are independent of all past observations. Thus, the innovations and the covariance function of the innovations can be used to formulate the likelihood function known as the PED (Schweppe, 1965). Minimizing negative two times the log of the likelihood is often completed, resulting in the function, where $\boldsymbol{\theta}$ is a vector that contains all of the parameters to be estimated (i.e., all parameters in the $\boldsymbol{\tau}$, $\boldsymbol{\pi}$, $\mathbf{T}$, $\mathbf{Z}$, $\boldsymbol{\Sigma}_v$, $\boldsymbol{\Sigma}_u$ vectors and

matrices)

$$-2\ln L(\boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{t=1}^{T} \mathrm{ny}\log(2\pi) + \log|\mathbf{F}_{it}| + \boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it}. \tag{1.40}$$

This likelihood function is computed from a conditional probability distribution function of the innovations, $\boldsymbol{\epsilon}_{it}$. The likelihoods are summed across time across individuals, which allows for the structure of psychological data with multiple time points and participants. Several versions of this likelihood with special recursions to handle the initial condition have been considered (Harvey, 1991; De Jong, 1988, 2003; Koopman & Durbin, 2003), which I will discuss in more detail later. For now, it is worth noting that after a sufficient number of time points the KF becomes independent of the initial state condition (see Jazwinski, 1970, pp. 239-243; Oud et al., 1990). Given the type of data commonly collected in psychological research, it becomes even more important to correctly specify initial condition.

A related set of recursions are called the Kalman Smoother (KF; Harvey, 1991; Shumway & Stoffer, 2006). While the KF provides state estimates based on observations up to time T, i.e., the filtered estimates $\mathbf{x}_{itFILT}$, the KS provides state estimates based on all of the observations, which I will call $\mathbf{x}_{itSMOOTH}$. This is accomplished by making a backwards pass through all observations from the last time point to the first, using $\mathbf{x}_{iTFILT}$ and $\mathbf{P}_{TFILT}$ as the initial condition values of the KS. The recursions are then, for $t = T, T-1, \ldots, 1$,

$$\mathbf{x}_{i,(t-1)SMOOTH} = \mathbf{x}_{i,(t-1)FILT} + \mathbf{J}_{i,t-1}(\mathbf{x}_{itSMOOTH} - \mathbf{x}_{itPRED}) \tag{1.41}$$

$$\mathbf{P}_{(t-1)SMOOTH} = \mathbf{P}_{(t-1)FILT} + \mathbf{J}_{i,t-1}(\mathbf{P}_{tSMOOTH} - \mathbf{P}_{tPRED})\mathbf{J}_{i,t-1}' \tag{1.42}$$

where

$$\mathbf{J}_{i,t-1} = \mathbf{P}_{(t-1)FILT}\mathbf{T}'(\mathbf{P}_{tPRED})^{-1}. \tag{1.43}$$

As I will be formulating structural equation models in the state-space framework, I will now describe the SEM framework in some detail, followed by a discussion of similarities and differences between the two modeling frameworks.

1.4 Structural Equation Modeling Framework

In general, SEM is concerned with understanding and evaluating the relations among both observed and latent variables (Bollen, 1989). Similar to state-space modeling, the model may be expressed using two equations - a measurement equation, linking the observed variables to the latent variables, and a structural equation, describing the relations among the latent variables. These equations are formally expressed as:

$$\text{Measurement equation: } \mathbf{y}_i = \boldsymbol{\tau}_S + \mathbf{Z}_S\mathbf{x}_i + \mathbf{u}_i \qquad \mathbf{u}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{u_S}) \tag{1.44}$$

$$\text{Structural equation: } \mathbf{x}_i = \boldsymbol{\pi}_S + \mathbf{T}_S\mathbf{x}_i + \mathbf{v}_i \qquad \mathbf{v}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{v_S}). \tag{1.45}$$

A subscript of S indicates that the parameter matrix contains elements associated with the SEM approach as opposed to the state-space modeling approach. Note here that the t subscript has been dropped from the equations. If longitudinal models are fit, variables associated with the T measurement equations must be incorporated into $\mathbf{y}_i$ and $\mathbf{x}_i$. Also, the structural equation contains the term $\mathbf{x}_i$ on both sides of the equation. Such a set up is well suited for examining structures of inter-individual (i.e., between people) differences (Chow et al., 2010).

Estimation in SEM. A raw data (i.e., full information) maximum likelihood (FIML; see Enders, 2001) estimation procedure may be used, which, under certain conditions,

is equivalent to using the PED likelihood function. Parameter estimation is accomplished by placing all parameters in a vector, $\boldsymbol{\theta}_S$, and choosing values such that the likelihood function is maximized (or the minimum of the negative likelihood function is minimized). Specifically, negative two times the log of the raw maximum likelihood function is expressed as (Bollen, 1989)

$$-2\ln LL_{FIML}(\boldsymbol{\theta}_S) = \sum_{i=1}^{N}[ny\log(2\pi) + \log|\boldsymbol{\Sigma}_S| + (\mathbf{y}_i - \boldsymbol{\mu}_S)'\boldsymbol{\Sigma}_S^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_S)] \tag{1.46}$$

where $\boldsymbol{\Sigma}_S$ is the model implied covariance matrix and $\boldsymbol{\mu}_S$ is the model implied mean vector containing all parameters. The goal is to find estimates for the parameters that minimizes the negative likelihood.

At this point, it may not be clear why a psychologist might wish to formulate models using the state-space modeling framework as opposed to the SEM framework. One important advantage of estimating models within the state-space modeling framework is that, given the KF and PED estimation procedure, data where T exceeds N may be directly accommodated. In the SEM framework there is no easy or direct estimation procedure when using data where T exceeds N, which I will discuss more extensively in Section 1.5. Given that psychologists have been using time series models to estimate their data (Hamaker & Dolan, 2009), it may be beneficial to formulate models in state-space form. To better understand the advantages offered from both the SEM and state-space modeling frameworks, I will now discuss the differences and similarities between the two approaches.

1.5 Similarities Between SEM and State-Space Modeling Frameworks

Given the unique advantages both techniques offer, researchers have discussed similarities and differences between the SEM and state-space modeling frameworks (Chow et al., 2010; MacCallum & Ashby, 1986; Oud et al., 1990). In fact, both frameworks may be thought of as special cases of each other (Chow et al., 2010). As

discussed, an advantage of state-space models is their ability to capture complicated intra-individual dynamics, as opposed to SEM where a main strength is the ability to capture group-based and in some instances, inter-individual dynamics using, e.g., panel data. It may be beneficial to use state-space models to answer specific hypotheses geared more toward understanding the change that occurs within individuals over many time points. Additionally, ARMA and VARMA time series models may be readily structured as state space models making it easier to account for missing data and to include complex multivariate systems, mixed effects, and certain types of nonstationarity (Shumway & Stoffer, 2006).

Several methodologists have commented on the correspondence between factor scores obtained from latent variables in SEM and state estimates obtained from the KF and related KS (see Shumway & Stoffer, 2006) in state-space models. If the transition matrix, $\mathbf{T}$, is null (i.e., there exist no structural relations about the state variables), then the state estimates obtained from the KF are equal to those obtained by the regression method for obtaining factor scores (Oud et al., 1990; Lawley & Maxwell, 1971). Furthermore, when $\mathbf{T}$ is null and $\mathbf{P}_0 \to \infty$, the state estimates obtained from the KF are equal to the Bartlett estimator of factor scores. This is also the case when a diffuse prior is used (to be described in detail later), namely, when the diagonals of the covariance matrix, $\mathbf{P}_0$, approach infinity. When $\mathbf{T}$ is not null, Dolan and Molenaar (1991) showed that the regression method produces more accurate factor score estimates than the KF; however, if the related KS is also applied to yield $E(\mathbf{y}_{it}|\mathbf{Y}_{iT})$, $\mathbf{Y}_{iT} = \{\mathbf{y}_{ij}, j = 1, \ldots, T\}$, then the state and factor score estimates are again equivalent (also see Chow et al., 2010).

Equivalences between SEM and PED estimation with respect to parameter estimates have been discussed by Chow et al. (2010) and MacCallum and Ashby (1986). When the transition matrix $\mathbf{T}$ is null Chow et al. (2010) show the exact analytical equivalence, for the cases where both $T = 1$ and $T > 1$, of the likelihood functions

for SEM and the PED function derived from using the KF. In this case, since they are equivalent initial condition specification is not an issue. However, when the transition matrix **T** is not null, then the parameter estimates from the SEM likelihood function will NOT be equivalent to those from the state-space likelihood function, unless the same initial condition distribution is incorporated as part of the model in the SEM framework. Thus, if a state-space model is estimated in the SEM framework, the transition matrix **T** is not null, and the same initial condition specification was not used, the parameter estimates will not be correct.

Issues in fitting longitudinal models with large T. A reason why psychologists may wish to formulate models in state-space form as opposed to SEM form is to handle data that have a large number of time points T and a relatively small sample size N. Estimating such data within the SEM framework poses several difficulties. For instance, when there are T measurement occasions along with ny manifest variables at each time point, the data input matrix will be of dimension $N \times Tny$. The associated data covariance matrix is singular (i.e., rank-deficient) if $T > N$, even when $ny = 1$, and thus cannot be handled using SEM. In addition, in SEM it is required that a $Tny \times Tny$ model-implied covariance matrix, $\mathbf{\Sigma}_S$, be inverted, which can be computationally unmanageable. Several methodologists have proposed solutions to circumvent the problems of a singular data matrix, including using a block Toeplitz approach (Molenaar, 1985; Nesselroade & Molenaar, 1999) and a FIML estimation procedure with specific constraints on the model-implied covariance structure (Hamaker, Dolan, & Molenaar, 2003), both of which pose their own challenges. Furthermore, all existing SEM approaches to handling data where T exceeds N deal only with stationary systems.

A block Toeplitz approach, proposed by Molenaar (1985), provides a way to reduce the input covariance matrix. Basically, a block Toeplitz matrix is formulated by lagging the observed data on themselves for as many lags as the model specifies.

Structurally, this matrix consists of the combination of lagged covariance matrices with a lag zero covariance matrix in the main block diagonal and the lagged covariance matrices in the off-diagonal block. If individuals are expected to display a similar structure, these matrices may be pooled (Nesselroade & Molenaar, 1999). While this approach has some benefits, there are also some drawbacks. The independence assumption that SEM relies on is violated due to the temporal dependence in the time series data and this results in pseudo-ML estimates (Molenaar & Nesselroade, 1998). There are also redundant elements in the block Toeplitz matrix which researchers need to be aware of so the degrees of freedom are properly adjusted. One option to resolve this issue is to just use the unique parts associated with the concurrent and lagged effects (see Browne & Zhang, 2005). Another limitation, however, is that neither the block Toeplitz formulation or the formulation by Browne and Zhang (2005) inherently address the concerns for initial condition specification and use least squares or generalized least squares approaches for parameter estimation (ML estimates cannot be obtained).

Hamaker et al. (2003) described a method that employs FIML estimation using the raw data. In this approach, special parameter constraints need to be introduced in order to make sure the model-implied covariance matrix is positive definite (non-singular). This procedure may be computationally difficult as a $Tny \times Tny$ covariance matrix still needs to be inverted at each iteration. They deal with stationary models and they describe procedures for computing initial condition values based on either 1) back-forecasting (Box & Jenking, 1976, pp. 213-217) to obtain the unconditional sum of squares, 2) the unconditional mean and covariance structures, or 3) setting the first p states equal to zero and calculating the initial condition based on the data up to p, then using the rest of the $p+1$ data to estimate parameter values. Thus, nonstationary systems are not dealt with because the constraints that have to be imposed to ensure the positive definiteness of the model-implied covariance matrix

can only be derived for stationary systems.

In summary, it may be beneficial to structure models within the state-space modeling framework as time series processes are easily incorporated, data where T exceeds N may be easily accommodated even for nonstationary processes, and intra-individual dynamics may be more easily modeled and captured. I now turn to a more technical discussion of initial condition specification in both the SEM and state-space modeling frameworks.

1.6 Technical Detail of Initial Condition Specification in the SEM Framework

In the past, psychologists have initialized their models in a number of ways when using the SEM framework. For example, as discussed, standard LCMs are typically initialized by freely estimating the parameters associated with the latent factors in the initial condition distribution. This amounts to freely estimating the means, variances, and covariances among the latent factors within the initial condition distribution.

Bollen and Curran (2004) discussed an auto-regression latent trajectory (ALT) model which incorporates features of both a latent curve model and autoregressive time series process of order 1. They also specify that the AR(1) weight, α, must be in between the values of -1 and 1, rendering this process to be non-explosive. In this case, the manifest variables must be initialized as they are the variables exhibiting the AR(1) process. One method they propose for initializing this process is specifying the first manifest variable, y_{i1}, to be exogenous. This first manifest variable is then not regressed on the latent states, although it may be regressed upon possible time-invariant covariates. It is, however, correlated with the latent states and all subsequent manifest variables. Therefore, the first manifest variable is distributed as

$$y_{i1} \sim N\left(\mu_{y_1}, \sigma^2_{u_1}\right) \tag{1.47}$$

with the option of adding covariates into the distribution. If no covariates are added,

then μ_{y_1} and $\sigma^2_{u_1}$ are simply the unconditional mean and variance of the observed variable y_{i1}. They describe another method for initializing the process that specifies y_{i1} to be endogenous and influenced by another variable, y_{i0}. They show that, if the α parameters are equal over time and $|\alpha| < 1$ (i.e., conditions for stationarity), then the factor loadings linking y_{i1} to the latent intercept and latent slope may be constrained to $(1-\alpha)^{-1}$ and $-\alpha(1-\alpha)^{-2}$, respectively. The initial state vector may then be expressed as,

$$y_{i0} = x_{INTi} + (1-\alpha)^{-1} - x_{SLPi}\alpha(1-\alpha)^2 + z_{i0} \tag{1.48}$$

where z_{i0} is an infinite weighted sum of previous u variables. Hamaker (2005) described how the z_{i0} term is actually just an AR(1) process. Thus, the initial condition distribution for y_{i1} is now much more complicated and expressed as

$$y_{i1} \sim N\left(\left[(1-\alpha)^{-1}\mu_{x_{INT}} - \alpha(1-\alpha)^{-2}\mu_{x_{SLP}}\right], \tag{1.49}$$
$$\left[((1-\alpha)^{-1})^2\sigma^2_{v_{x_{INT}}} + (\alpha(1-\alpha)^{-2})^2\sigma^2_{v_{x_{SLP}}} - 2(-\alpha(1-\alpha)^{-3})\sigma^2_{v_{x_{INT}x_{SLP}}} + \frac{\sigma^2_u}{1-\alpha^2}\right]\right).$$

While this works mathematically, it may be difficult for the model to converge as the initial condition is made up of complicated non-linear constraints that must be explicitly constructed into the initial condition distribution of the model. Also, these constraints are derived assuming the model is stationary, so nonstationary processes are not dealt with in this case.

Du Toit and Browne (2001, 2007) were among the first to explicitly consider initial condition specification in the SEM framework. Specifically, Du Toit and Browne (2001, 2007) discussed methods analyzing times series models in SEM by incorporating a covariance structure that allows for some specification of initial condition based on the stationarity of the model. The models they discussed are VARMA time series

models, as expressed in Equation 1.4, being estimated in the context of SEM with the option of using ny observed variables for all individuals instead of just a single time series (i.e., a single set of y_t variables). The VARMA equation formulation is reproduced here for convenience,

$$\mathbf{y}_{it} = \sum_{r=1}^{p} \mathbf{A}_r \mathbf{y}_{i,t-1} + \sum_{j=1}^{q} \mathbf{B}_j \mathbf{v}_{i,t-j} + \mathbf{v}_{it}$$

where $\mathbf{y}_{it}$ represent the $ny \times 1$ vector of observed variables over time.

When assuming the process is stationary, Du Toit and Browne (2007) used the unconditional mean and covariance matrix to specify the initial condition distribution (i.e., $\boldsymbol{\mu}_0 = \boldsymbol{\mu}$ $\mathbf{P}_0 = \mathbf{P}$). They also considered a condition where they allowed for a change of process prior to the first observation, but assumed a stationary process onward. This amounts to the solution that freely estimates the parameters of the initial condition distribution, i.e., parameters contained in $\boldsymbol{\mu}_0$ and $\mathbf{P}_0$ are freely estimated. The third method Du Toit and Browne (2007) consider is assuming the process truly started at $t = 1$. This amounts to using a null vector and matrix for the mean and covariance matrices of the initial condition distribution, i.e., $\boldsymbol{\mu}_0 = \mathbf{0}$ and $\mathbf{P}_0 = \mathbf{0}$.

Du Toit and Browne (2007) first explained how to estimate a VARMA model to a group of subjects using SEM by deriving the lagged covariance structure for the observed variables, $\mathbf{y}_{it}$. They explained that, in a VARMA(p,q) process, the first m variables are directly affected by the initial state variables where $m = \max(p, q)$. They created an initial state vector of dimension $nym \times 1$ that contains variables that affect the first m time points of the observed variables,

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}'_{11} & \mathbf{f}'_{21} & \mathbf{f}'_{31} & \dots & \mathbf{f}'_{m1} \end{bmatrix}' \tag{1.50}$$

An equation representing the effects of the initial state variables on the process vari-

ables at the first time point is formulated as

$$\mathbf{f}_{l1} = \sum_{j=1}^{m} \mathbf{A}_j^{[l]} \mathbf{y}_{l-1} + \mathbf{B}_j^{[l]} \mathbf{v}_{l-1} \qquad l = 1 \ldots m. \tag{1.51}$$

Given this, if we let $\mathbf{y}$ be a $nyT \times 1$ vector stacking all T of the $ny \times 1$ observed variable vectors and $\mathbf{v}$ be a $nyT \times 1$ vector stacking all T of the $ny \times 1$ random shock vectors, the overall model structure may now be defined as

$$\mathbf{y} = \mathbf{T}_{-A}^{-1}(\mathbf{I}_{T|m}\mathbf{f} + \mathbf{T}_B\mathbf{v}) \tag{1.52}$$

where $\mathbf{I}_{T|m}$ is a $nyT \times nym$ matrix formed from the first nym columns of the $nyT \times nyT$ Identity matrix, $\mathbf{T}_{-A}$ and $\mathbf{T}_B$ are both $nyT \times nyT$ lower triangular block Toeplitz matrices, with elements going back in time as far as m, where $\mathbf{T}_{-A}$ contains the negative autoregression parameters while $\mathbf{T}_B$ contains the moving average parameters (for a specific example, see Du Toit & Browne, 2007, Eq. 7 and Eq. 8),

$$\mathbf{T}_{-A} = \begin{bmatrix} \mathbf{I}_{ny} & 0 & 0 & 0 & \ldots & 0 \\ -\mathbf{A}_1 & \mathbf{I}_{ny} & 0 & 0 & \ddots & 0 \\ -\mathbf{A}_2 & -\mathbf{A}_1 & \mathbf{I}_{ny} & 0 & \ddots & 0 \\ -\mathbf{A}_3 & -\mathbf{A}_2 & -\mathbf{A}_1 & \mathbf{I}_{ny} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & -\mathbf{A}_m & \ldots & -\mathbf{A}_1 & \mathbf{I}_{ny} \end{bmatrix} \tag{1.53}$$

$$\mathbf{T}_B = \begin{bmatrix} \mathbf{I}_{ny} & 0 & 0 & 0 & \dots & 0 \\ \mathbf{B}_1 & \mathbf{I}_{ny} & 0 & 0 & \ddots & 0 \\ \mathbf{B}_2 & \mathbf{B}_1 & \mathbf{I}_{ny} & 0 & \ddots & 0 \\ \mathbf{B}_3 & \mathbf{B}_2 & \mathbf{B}_1 & \mathbf{I}_{ny} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \mathbf{B}_m & \dots & \mathbf{B}_1 & \mathbf{I}_{ny} \end{bmatrix}. \tag{1.54}$$

The covariance structure of $\mathbf{y}$ may then be represented as (Du Toit & Browne, 2007, Equation 31)

$$\mathbf{\Sigma}_{\mathbf{y}} = \text{Cov}(\mathbf{y}, \mathbf{y}') = \mathbf{T}_{-A}^{-1}(\mathbf{I}_{T|m}\mathbf{P}_{0f}\mathbf{I}_{T|m}' + \mathbf{T}_B(\mathbf{I}_T \otimes \mathbf{\Sigma}_v^2)\mathbf{T}_B')\mathbf{T}_{-A}^{-1'} \tag{1.55}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{P}_{0f} = \text{Cov}(\mathbf{f}, \mathbf{f}')$. Given this structure, the $\mathbf{P}_0$ matrix may be composed of extra parameters to be estimated. This allows for the possibility of a change of process before the first observation was collected, but assumes stationarity from the first observation to the final one.

Du Toit and Browne (2007) next introduced a state-space formulation to derive the specific structure of the initial condition distribution for a stationary VARMA model, and hence an initial condition distribution that contains parameters that are already in the model. Since stationarity is assumed throughout the whole process, the $\mathbf{f}$ variables may be extended to represent the whole process,

$$\mathbf{f}_{t+1} = \mathbf{A}\mathbf{f}_t + \mathbf{G}\mathbf{v}_t \tag{1.56}$$

$$\mathbf{y}_t = \mathbf{H}\mathbf{f}_t + \mathbf{v}_t \tag{1.57}$$

where Equation 1.56 is the transition equation and Equation 1.57 is the measurement

equation. $\mathbf{f}_t$ is now a $nym \times 1$ vector of the initial state variables given by

$$\mathbf{f}_t = \begin{bmatrix} \mathbf{f}_{1t} \\ \mathbf{f}_{2t} \\ \vdots \\ \mathbf{f}_{mt} \end{bmatrix} t = 1, 2, \ldots, T. \tag{1.58}$$

$\mathbf{A}$ is a $nym \times nym$ matrix given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{I}_{ny} & 0 & \ldots & 0 \\ \mathbf{A}_2 & 0 & \mathbf{I}_{ny} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ldots \\ \mathbf{A}_{m-1} & 0 & 0 & \ldots & \mathbf{I}_{ny} \\ \mathbf{A}_m & 0 & 0 & \ldots & 0 \end{bmatrix} \tag{1.59}$$

where $\mathbf{A}_1, \ldots \mathbf{A}_m$ are $ny \times ny$ autoregression weight matrices and $\mathbf{I}_{ny}$ is the identity matrix of dimension ny. $\mathbf{G}$ is a $nym \times ny$ matrix given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{A}_1 + \mathbf{B}_1 \\ \mathbf{A}_2 + \mathbf{B}_2 \\ \vdots \\ \mathbf{A}_m + \mathbf{B}_m \end{bmatrix} \tag{1.60}$$

where $\mathbf{B}_1, \ldots \mathbf{B}_m$ are $ny \times ny$ moving average weight matrices. For the measurement equation, $\mathbf{y}_t$ is a $nyT \times 1$ vector of observed variables, $\mathbf{H}$ is a $ny \times nym$ matrix given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_{ny} & \mathbf{0}_{ny} & \mathbf{0}_{ny} & \ldots & \mathbf{0}_{ny} \end{bmatrix}. \tag{1.61}$$

Finally, $\mathbf{v}_t$ is a $nym \times 1$ vector given by

$$\mathbf{v}_t = \begin{bmatrix} \mathbf{v}_{1t} \\ \mathbf{v}_{2t} \\ \vdots \\ \mathbf{v}_{mt} \end{bmatrix} \tag{1.62}$$

where $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_v)$. Now, the covariance matrix for the initial state vector may be expressed as a function of parameters already in the model,

$$\text{vec}(\mathbf{P}_{0f}) = (\mathbf{I} - \mathbf{A} \otimes \mathbf{A})^{-1} \text{vec}(\mathbf{G}\Sigma_v\mathbf{G}) \tag{1.63}$$

where the vec operator stacks all elements (column-wise) into a vector and the $\mathbf{A}$ and $\mathbf{G}$ matrices correspond to those from Equations 1.57, 1.59, and 1.60. This specification is for a process that has started in the distant past and satisfies the condition of weak stationarity. Since the process remains stable prior to the first observation, the initial condition distribution, specifically the covariance matrix $\mathbf{P}_{0f}$, is a function of parameters that are *already* in the model, and represents the unconditional covariance matrix. This can be seen in Equation 1.63, where the covariance function includes the AR parameters, MA parameters, and the process noise variance. Also, assuming stationarity ensures that the structure of Equation 1.55 holds.

Another specification for initial condition Du Toit and Browne (2007) discussed happens when a change in the process could have occurred before the first observation, x_1. In this case, the parameters in $\mathbf{P}_{0f}$ are considered additional parameters that are actually estimated, although the mean value of $\mathbf{x}_{i0}$ is not considered. In this case, there are $\frac{1}{2}nym(nym+1)$ additional parameters to be estimated. This specification may not produce the exact covariance structure as in Equation 1.55, with more distinct differences at earlier time points. If the vector AR part of the model is stationary

(i.e., as indicated by elements of the **A** matrices), however, this specification will tend toward the covariance structure as t increases. This type of initial condition distribution is also commonly implemented in growth curve and latent difference score models in the SEM literature. For example, Chow et al. (2010) formulated a dual change score model (McArdle & Hamagami, 2001) in state-space form and included initial condition by allowing the elements of the initial condition mean vector and covariance matrix to be parameters estimated in the model.

Another variation of stationarity Du Toit and Browne (2007) discuss is to assume the process started at the time of the first observation and that there are no prior influences. The choice of covariance matrix $\mathbf{P}_{0f}$ would then be a null matrix. However, they are assuming stationarity once the process has started. The structure of the overall covariance matrix thus may not be the structure presented in Equation 1.55, but rather it will approximate this structure more and more accurately with increasing time points.

One similar feature across their differing structures is that the process is considered stationary after the initial time point. Thus, while there may have been a change beforehand, the process must remain stationary throughout the time points after $t = 0$. Thus, the covariance structure differs as a function of lag but not over time.

Some quantitative methodologists have considered initial condition specification with respect to using the KF and PED estimation procedure. Oud et al. (1990) estimated a model by first using a ML estimation in the LISREL software program (Jöreskog & Sörbom, 2001) and subsequently using the parameter estimates to make a pass through the KF. When making a pass through the KF, Bartlett's estimates and associated covariance matrix of the cross-sectional factor scores (Lawley & Maxwell, 1971) for the first time point were used as the initial condition distribution. This leads

to the following equations for $\mathbf{x}_{i1PRED}$ and $\mathbf{P}_{1PRED}$:

$$\mathbf{x}_{i1PRED} = (\mathbf{Z}'(\mathbf{\Sigma}_u)^{-1}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{\Sigma}_u^{-1})\mathbf{y}_{i1} \tag{1.64}$$

$$\mathbf{P}_{1PRED} = (\mathbf{Z}'(\mathbf{\Sigma}_u)^{-1}\mathbf{Z})^{-1}. \tag{1.65}$$

Oud and Jansen (1996) performed a similar procedure, but this time implemented the expectation-maximization (EM; Dempster, Laird, & Rubin, 1977) estimation procedure iterating between performing the M-step in LISREL and the E-step using the KF until some convergence criterion was met. In this case, however, they used the cross-sectional regression estimator for factor scores (Lawley & Maxwell, 1971) at the first time point as opposed to Bartlett's estimates, leading to the following equations for $\mathbf{x}_{i1PRED}$ and $\mathbf{P}_{1PRED}$:

$$\mathbf{x}_{i1PRED} = \mathbf{\Sigma}_v\mathbf{Z}'(\mathbf{Z}\mathbf{\Sigma}_v\mathbf{Z}' + \mathbf{\Sigma}_u)^{-1}\mathbf{y}_{i1} \tag{1.66}$$

$$\mathbf{P}_{1PRED} = \mathbf{\Sigma}_v(\mathbf{I} + \mathbf{Z}'\mathbf{\Sigma}_u^{-1}\mathbf{Z}\mathbf{\Sigma}_v)^{-1}. \tag{1.67}$$

Once again, however, these approaches are for stationary models that assume the process has started in the distant past.

In summary, quantitative psychologists have considered several different initial condition specifications. Assuming a stationary process, the unconditional mean and covariance matrix may be used for the initial condition distribution, as considered by Du Toit and Browne (2007). Bollen and Curran (2004) considered a similar procedure when they derived what the factor loadings would be given a non-explosive process. Another specification is freely estimating parameters that are included in the initial condition distribution. Du Toit and Browne (2007) considered this in the context of VARMA models, Bollen and Curran (2004) considered this in the context of ALT models, Chow et al. (2010) considered this in the context of a dual change

score model, and this is the general procedure used when estimating a conventional latent growth curve model. Finally, treating the initial condition as a null vector was considered by Du Toit and Browne (2007).

1.7 Technical Detail of Initial Condition Specification in the State-Space Modeling Framework

While some authors in psychological research have discussed initial condition specification in depth (e.g., Du Toit & Browne, 2007), most of the methodological work on this subject has been conducted in the state-space literature. Harvey (1991) explains that if the transition equation is stationary, then the distribution of initial condition is given by the mean and covariance matrix of the unconditional distribution of the state vector. This amounts to a mean of $E(\mathbf{x}_{i0}) = (\mathbf{I} - \mathbf{T})^{-1}\boldsymbol{\tau}$, where $\boldsymbol{\tau}$ is a vector of intercepts (see Equation 1.27), and covariance matrix which is the unique solution to the equation: $\mathbf{P}_t = \mathbf{T}\mathbf{P}_t\mathbf{T}' + \boldsymbol{\Sigma}'_v$. A solution to this equation given more complicated models has a similar form as found in Du Toit and Browne (2007), here shown in Equation 1.63,

$$\text{vec}(\mathbf{P}_0) = (\mathbf{I} - \mathbf{T} \otimes \mathbf{T})^{-1}\text{vec}(\boldsymbol{\Sigma}_v). \tag{1.68}$$

This holds for all state-space models, of which VARMA processes can be structured as a special case. However, this is also only for stationary models.

If the transition equation is not stationary, the unconditional distribution of the state vector is not defined (Harvey, 1991). If prior information is available for the distribution of $\mathbf{x}_{i0}$, then the KF via the PED will yield the exact likelihood function, as there will be a known $\boldsymbol{\mu}_0$ and bounded $\mathbf{P}_0$. However, there must be realistic prior information for this to be the case; otherwise a diffuse, or non-informative prior, must be used.

The idea of a diffuse initial condition has been discussed in depth in the state-

space and econometric literature (De Jong, 1991; De Jong & Chu-Chun Lin, 1994; De Jong, 2003; Koopman & Durbin, 2000, 2003). De Jong (1991) explains that a state is diffuse if the covariance matrix associated with it has arbitrarily large values in its diagonal entries. More formally, this is defined as $\mathbf{P}_0 = \kappa\mathbf{I}$ where κ is a positive scalar, with a diffuse prior having $\kappa \to \infty$ (Harvey, 1991). Another way of interpreting this is to say $\mathbf{P}_0^{-1}$ approaches zero. Such a diffuse initial state is appropriate to use when there is either uncertainty in the initial condition or the model is nonstationary (De Jong, 1991).

The KF in its original set up cannot handle incorporating a diffuse initial state (De Jong & Chu-Chun Lin, 1994). It would need to be initiated with $\mathbf{P}_0 = \kappa\mathbf{I}$, and plugging this into the prediction equation for the covariance matrix (Equation 1.39) would amount to $\mathbf{P}_{1PRED} = \mathbf{T}\kappa\mathbf{I}\mathbf{T}' + \mathbf{\Sigma}_v$, which cannot be used unless κ takes on a finite value. Also, since the initial condition distribution will enter into the likelihood function, κ will be present in this function as well. Thus, several methodologists have developed methods for incorporating a diffuse initial state. Schweppe (1973), Harvey and Phillips (1979), and Chow et al. (2010) have dealt with a diffuse initial state by initiating the KF with a very large covariance matrix. De Jong (1991) discussed how this can lead to numerical problems and furthermore, the question of the existence of diffuse constructs is not addressed. Koopman (1997) commented that this is not an exact solution and may lead to inaccurate results due to rounding errors.

Another way to deal with a diffuse initial condition is to use the Information filter, which is a variant of the KF. This procedure adapts recursions that are based on $\mathbf{P}^{-1}$ instead of $\mathbf{P}$, making it useful for specifying a diffuse prior of $\mathbf{P}^{-1} = 0$ (Kitagawa, 1981). However, Ansley and Kohn (1985) reported that this procedure breaks down in some cases and may be numerically inefficient. Also, it cannot be used in all modeling situations. Ansley and Kohn (1985) also developed a modified KF which allows for the implementation of a diffuse prior by allowing the covariance matrix of the

initial state vector, $\mathbf{P}_0$, to approach infinity. This method employs a transformation eliminating the dependence on the initial state vector. While this method works, De Jong (1988, 1991) developed an augmented KF, called the diffuse KF (DKF), that augments the original KF by adding an extra set of recursions. This procedure is easier to implement and more efficient than the one proposed by Ansley and Kohn (1985). Also, it does not require any transformation of the data and the diffuse elements of the model are estimated simultaneously via the extended recursions. This approach can also handle missing data and does not require a marked degree of added complexity with respect to computation. In fact, the DKF implicitly computes initial condition estimates from a set of initial observations and uses these to initiate the regular KF. Koopman (1997) developed a method allowing for a diffuse initial condition that is a modification of the Ansley and Kohn (1985) approach. This method expressed the KF in terms of the constant κ and then allows $\kappa \to \infty$. He explains that this approach is more computationally efficient than De Jong's (1991) DKF approach. As these two approaches are currently considered to be both efficient and easy to implement, I will discuss both in greater detail.

De Jong's DKA. To circumvent the problem of κ appearing in the initial condition distribution and likelihood function, De Jong (1991) proposed a solution that essentially takes the first nd observations of the data, where nd equals the number of diffuse states, and uses them to construct an estimate of the mean and covariance matrix for the diffuse elements. To better illustrate the derivation of the DKA, consider the following state vector at time $t = 0$:

$$\mathbf{x}_{i0} = \boldsymbol{\mu} + \mathbf{A}_\infty \boldsymbol{\delta} + \mathbf{A}_* \boldsymbol{\eta}, \qquad \boldsymbol{\delta} \sim N(0, \kappa \mathbf{I}_{nd}), \text{with } \kappa \to \infty, \qquad \boldsymbol{\eta} \sim N(0, \mathbf{P}_*) \tag{1.69}$$

where $\boldsymbol{\mu}$ is a $nx \times 1$ vector of fixed and known elements corresponding to the unconditional mean for a stationary process or zero for a nonstationary process, $\mathbf{A}_\infty$

and $\mathbf{A}_*$ are selection or indicator matrices that correspond to the diffuse part of the initial state vector and the stationary part of the initial state vector, respectively, where $\mathbf{A}_\infty$ is of dimension $nx \times nd$ and $\mathbf{A}_*$ is of dimension $nx \times ns$ where ns equals the number of stationary states. The random vector $\boldsymbol{\delta}$ represents the diffuse initial condition random vector and is a column vector of dimension $nd \times 1$. The random vector $\boldsymbol{\eta}$ is a column vector of dimension $ns \times 1$ and represents the non-diffuse or stationary part of the initial state vector and has a covariance matrix that may be equal to the unconditional covariance matrix of the non-diffuse or stationary series, which we will call $\mathbf{P}_*$. De Jong uses the augmented equations from $t = 1, \ldots, t = nd$ to calculate the expected value of the diffuse state vector given the data, $\hat{\boldsymbol{\delta}} = E(\boldsymbol{\delta}|\text{data from } t = 1, \ldots, t = nd)$, the covariance matrix of the diffuse state vector given the data, $\text{Cov}(\boldsymbol{\delta}|\text{data from } t = 1, \ldots, t = nd)$, and related quantities which are then used to refine the filtered estimates, the associated covariance matrix, and the log-likelihood function.

The overall distribution of the initial state vector is then

$$\mathbf{x}_0 \sim N(\boldsymbol{\mu}, \kappa \mathbf{A}_\infty \mathbf{A}_\infty' + \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*'). \tag{1.70}$$

The DKF will be the same as the ordinary KF if $\boldsymbol{\delta}$ is a null vector and thus $\mathbf{A}_\infty$ is a null matrix, making the whole model non-diffuse (De Jong & Chu-Chun Lin, 1994). With the existence of diffuse initial states, the original KF would be initialized using $\mathbf{P}_0 = \kappa \mathbf{A}_\infty \mathbf{A}_\infty' + \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*'$. In the diffuse case, $\kappa \to \infty$, and thus the initialization breaks down since κ appears in the equation for the initial covariance matrix. In order to implement the DKA, there must be a solution to the fact that when $\kappa \to \infty$, the original recursions break down and cannot be initialized. The solution De Jong (2003) described consists of adding several vector augmentations of the original KF and an extra set of recursions which are used to calculate all quantities associated

with δ until $t = nd$ (or, given missing data in the initial stretch of data, until a later time). This augmented filter may then be collapsed to the original KF after a certain period of time, usually equal to the number of diffuse states, nd (however, this is not always the case, due to data-irregularities such as missing data). The by-products of these augmented equations will then result in a set of estimates for the mean and covariance matrix of the diffuse initial condition as well as the terms for the log-likelihood function.

The augmented equations are as follows. First, the observed $\mathbf{y}_{it}$ vector must be augmented with nd columns of zeros. With the extra consideration that there can be multiple numbers of $\mathbf{y}$ variables, where ny again equals the number of y or observed variables, this amounts to:

$$\mathbf{y}_{it}^{*} = [\mathbf{0}_{nd} \quad \mathbf{y}_{it}] \tag{1.71}$$

where $\mathbf{0}_{nd}$ is an $ny \times nd$ zero matrix. The vector of innovations and the vector of predicted values are also augmented as follows:

$$\boldsymbol{\epsilon}_{it}^{*} = [\mathbf{E}_{it} \quad \boldsymbol{\epsilon}_{it}] = \mathbf{y}_{it}^{*} - \mathbf{Z}\mathbf{x}_{itPRED}^{*} \tag{1.72}$$

$$\mathbf{x}_{itPRED}^{*} = [\mathbf{A}_{\infty it} \quad \mathbf{x}_{itPRED}] \tag{1.73}$$

$$\mathbf{x}_{i,t+1,PRED}^{*} = [\mathbf{A}_{\infty i,t+1} \quad \mathbf{x}_{i,t+1,PRED}] = \mathbf{T}[\mathbf{A}_{\infty it} \quad \mathbf{x}_{itPRED}] + \mathbf{K}_{t}[\mathbf{E}_{it} \quad \boldsymbol{\epsilon}_{it}] \tag{1.74}$$

where $\mathbf{x}_{itPRED}^{*}$ is of dimension $nx \times (nd + 1)$ and $\boldsymbol{\epsilon}_{it}^{*}$ is $ny \times (nd + 1)$, with $\mathbf{E}_{it}$ and $\mathbf{A}_{\infty it}$ having dimension $nx \times nd$, and $\boldsymbol{\epsilon}_{it}$ and $\mathbf{x}_{itPRED}$ having dimension $nx \times 1$. $\boldsymbol{\epsilon}_{it}$ are the usual innovations while $\mathbf{E}_{it}$ now contains information needed to estimate the diffuse constructs, δ, and this information is also needed in the log-likelihood equation. The process is initialized with $\mathbf{x}_{i1PRED}^{*} = \mathbf{T}[\mathbf{A}_{\infty} \quad \boldsymbol{\mu}]$, $\mathbf{Q}_1 = 0$, and $\mathbf{P}_{1PRED} = \mathbf{T}\mathbf{A}_{*}\mathbf{P}_{*}\mathbf{A}_{*}\mathbf{T}' + \boldsymbol{\Sigma}_{v}$ where $\mathbf{P}_{*}$ again represents the covariance matrix associated with the

stationary part of the model. In addition, the following recursion is added:

$$\mathbf{Q}_{t+1} = \mathbf{Q}_t + \boldsymbol{\epsilon}_{it}^{*'}\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it}^{*} \tag{1.75}$$

such that $\mathbf{Q}_{t+1}$ is a $(nd+1) \times (nd+1)$ matrix with the following elements:

$$\mathbf{Q}_{t+1} \quad = \quad \mathbf{Q}_t + \begin{bmatrix} \mathbf{E}_{it}'\mathbf{F}_{it}^{-1}\mathbf{E}_{it} & \mathbf{E}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it} \\ \boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\mathbf{E}_{it} & \boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it} \end{bmatrix} \tag{1.76}$$

where $\mathbf{E}_{it}'\mathbf{F}_{it}^{-1}\mathbf{E}_{it}$ is an $nd \times nd$ matrix, $\mathbf{E}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it}$ is an $nd \times 1$ matrix, $\boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\mathbf{E}_{it}$ is a $1 \times nd$ matrix, and $\boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it}$ is a scalar. Thus, if there are no stationary elements and all states in the model are diffuse, this term will be zero. If there are stationary states, then the $\mathbf{P}_*$ matrix may be the unconditional covariance matrix, corresponding to $\mathbf{A}_*\mathbf{P}_*\mathbf{A}_*'$. If no proper priors are used this is a null matrix. Note that none of these initial conditions are dependent on κ, which allows the DKF to be initialized without explicitly incorporating κ into the initial condition distribution.

After an initial stretch of the data, De Jong and Chu-Chun Lin (1994) and De Jong (2003) described how the DKF can be collapsed to the regular KF. This is because the recursions are no longer dependent on κ as the elements of the equations up to this point will have provided estimates of the initial state and associated covariance matrix. Specifically, after an initial stretch of the data, $t = 1, 2, \ldots, d \leq n$, $\mathbf{x}_{itPRED}^{*}$ and $\boldsymbol{\epsilon}_t^{*}$ may be collapsed out as they can be reduced to vectors and $\mathbf{Q}_t$ can be reduced to a scalar value, after which the ordinary KF can be implemented. At $t = d$, the $\mathbf{Q}_t$ matrix is:

$$\mathbf{Q}_d \quad = \quad \begin{bmatrix} \mathbf{S} & \mathbf{s} \\ \mathbf{s}' & q_d \end{bmatrix} \tag{1.77}$$

where $\mathbf{S} = \sum_{k=1}^{d}\sum_{i=1}^{N} \mathbf{E}'_{k-1,i}\mathbf{F}^{-1}_{k-1,i}\mathbf{E}_{k-1,i}$, $\mathbf{s} = \sum_{k=1}^{d}\sum_{i=1}^{N} \mathbf{E}'_{k-1,i}\mathbf{F}^{-1}_{k-1,i}\boldsymbol{\epsilon}_{k-1,i}$, and $q_d = \sum_{k=1}^{d}\sum_{i=1}^{N} \boldsymbol{\epsilon}'_{k-1,i}\mathbf{F}^{-1}_{k-1,i}\boldsymbol{\epsilon}_{k-1,i}$, where q_d contains the usual terms that are input into the log likelihood function in the regular KF. Once $\mathbf{S}^{-1}$ is nonsingular a collapse can occur such that the DKF can become the KF, which usually occurs when $d = nd$. However, when there is missingness in the initial stretch of the data, the collapse may occur at a different time.

At the point of the collapse, maximum likelihood estimates of the mean and covariance matrix of the diffuse state vector may be calculated as follows,

$$\hat{\delta} = E(\delta|Y_d) = \mathbf{S}^{-1}\mathbf{s} \tag{1.78}$$

$$\text{VAR}(\hat{\delta}) = \frac{(q_d - \mathbf{s}'\mathbf{S}^{-1}\mathbf{s})}{t-d}\mathbf{S}^{-1}. \tag{1.79}$$

The filter equations may then be calculated as

$$\mathbf{x}_{i,d+1,FILT} = \mathbf{a}_d + \mathbf{A}_d\mathbf{S}^{-1}\mathbf{s} \tag{1.80}$$

$$\mathbf{P}_{d+1,FILT} = \mathbf{P}_d - \mathbf{A}_d\mathbf{S}^{-1}\mathbf{A}'_d. \tag{1.81}$$

Then, from $t = d+1, \ldots, T$, the regular KF is employed which now incorporates estimates of elements from the mean and covariance matrix of the diffuse states. Thus, the effects of initial condition are estimated from an initial stretch of the data, added to the regular KF, and then the regular KF is employed for the rest of the data.

Another challenge when using a diffuse initial condition is using a likelihood function that does not depend on κ. The term $\mathbf{P}_{tPRED}$ needed to compute $\mathbf{F}_{it}$ in Equation 1.40 contains an infinity term. De Jong (1991) derives a formation of a

likelihood as $\kappa \to \infty$ that can be expressed as,

$$-2 \ln L(\boldsymbol{\theta}) \propto \sum_{i=1}^{N} \sum_{t=1}^{T} \log |\mathbf{F}_{it}| + q_T + \log |\mathbf{S}_T^{-1}| + \mathbf{s}_T' \mathbf{S}_T \mathbf{s}_T. \tag{1.82}$$

Shown another way, if the DKF can be collapsed at $t = nd$, such that $\mathbf{S}^{-1}$ is non-singular at $t = nd$, then the likelihood is as follows:

$$\begin{aligned} -2 \ln L(\boldsymbol{\theta}) \propto \sum_{i=1}^{N} \sum_{t=1}^{nd} \log |\mathbf{F}_{it}| + q_{nd} + \log |\mathbf{S}_{nd}^{-1}| - \mathbf{s}_{nd}' \mathbf{S}_{nd} \mathbf{s}_{nd} \\ + \sum_{i=1}^{N} \sum_{t=nd+1}^{T} \log |\mathbf{F}_{it}| + \sum_{i=1}^{N} \sum_{t=nd+1}^{T} \boldsymbol{\epsilon}_{it}' \mathbf{F}_{it}^{-1} \boldsymbol{\epsilon}_{it}. \end{aligned} \tag{1.83}$$

Koopman's approach. The approach initially described by Koopman (1997) does not require the augmented matrix operations that the De Jong (1991) approach requires. Instead, Koopman developed what he terms an exact initial KF that breaks up the KF into two parts: recursions for the diffuse elements and recursions for the non-diffuse elements. Koopman (1997) termed this the exact initial KF because the KF equations are expressed explicitly in terms of κ and then a solution is obtained by allowing κ to approach infinity (see p. 1631-1632 for the proof).

To describe this approach, first denote $\mathbf{M}_t = \mathbf{P}_{tPRED} \mathbf{Z}'$ so $\mathbf{K}_t = \mathbf{M}_t \mathbf{F}_{it}^{-1}$. Koopman then structures the covariance matrix of the initial condition as follows,

$$\mathbf{P}_0 = \mathbf{P}_* + \kappa \mathbf{P}_\infty \tag{1.84}$$

where $\mathbf{P}_*$ represents the non-diffuse part of the initial condition covariance matrix, i.e., $\mathbf{A}_* \mathbf{P}_* \mathbf{A}_*'$ using the notation from Equation 1.69, and $\mathbf{P}_\infty$ represents the diffuse part of the initial condition covariance matrix, i.e., $\mathbf{A}_\infty \mathbf{A}_\infty'$. This procedure is then initialized with $\mathbf{P}_{\infty,0} = \mathbf{A}_\infty \mathbf{A}_\infty'$ and $\mathbf{P}_{*,0} = \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*'$.

The exact initial KF then modifies Equation 1.34 and Equation 1.35 to include

recursions for the diffuse elements as follows,

$$\mathbf{F}_{it} = \mathbf{F}_{*,it} + \kappa \mathbf{F}_{\infty,it} \tag{1.85}$$

$$\mathbf{M}_{t} = \mathbf{M}_{*,t} + \kappa \mathbf{M}_{\infty,t} \tag{1.86}$$

where

$$\mathbf{F}_{*,it} = \mathbf{Z}\mathbf{M}_{*,t} + \mathbf{\Sigma}_{u} \qquad \mathbf{F}_{\infty,it} = \mathbf{Z}\mathbf{M}_{\infty,t} \tag{1.87}$$

$$\mathbf{M}_{*,t} = \mathbf{P}_{*,t}\mathbf{Z}' \qquad \mathbf{M}_{\infty,t} = \mathbf{P}_{\infty,t}\mathbf{Z}'. \tag{1.88}$$

Koopman (1997, p. 1631-1632) provided explicit solutions to the exact initial KF recursions. Koopman and Durbin (2003) suggested only using solutions resulting from the two special cases of $\mathbf{F}_{\infty,it}$, that is, when $\mathbf{F}_{\infty,it}$ is non-singular and when $\mathbf{F}_{\infty,it}$ is equal to zero. When $\mathbf{F}_{\infty,it}$ is a nonzero singular matrix, Koopman and Durbin (2003) suggested using the approach described by Koopman and Durbin (2000), which involves converting the multivariate series to a univariate series under the assumptions that the measurement errors and process noise are uncorrelated and $\mathbf{\Sigma}_{u}$ is a diagonal matrix (Koopman, Shephard, & Doornik, 2008, p. 100).

For most well-defined state-space models, $\mathbf{F}_{\infty,it}$ is non-singular. $\mathbf{F}_{\infty,it}$ will be equal to zero when there is all missing data in $\mathbf{y}_{it}$ for a particular occasion t. For all cases where $\mathbf{F}_{\infty,it}$ is singular, there is usually something peculiar with the model, such as having redundant latent variables or manifest variables. Overall, it makes sense to use the univariate approach in case of a singular $\mathbf{F}_{\infty,it}$ matrix. Also, Koopman and Durbin (2000) described how using their univariate method improves computational efficiency, especially for more complicated models (see Table 1, p. 287, Koopman & Durbin, 2000).

Koopman (1997) derived the form of Equation 1.85 as $\kappa \to \infty$. In order to com-

plete this, Koopman (1997) first partially diagonalized the $\mathbf{F}_\infty$ and $\mathbf{F}_*$ matrices (see Equation 7 in Koopman, 1997). He next expanded the inverse of $\mathbf{F}$ in Equation 1.85 to arrive at the equation,

$$\mathbf{F}^{-1} = \mathbf{F}_*^- + \frac{1}{\kappa}\mathbf{F}_\infty^- - \frac{1}{\kappa^2}\mathbf{F}_\infty^-\mathbf{F}_*\mathbf{F}_\infty^- + \mathcal{O}(\frac{1}{\kappa^3}) \tag{1.89}$$

where $\mathbf{F}_*^-$ and $\mathbf{F}_\infty^-$ are derived from elements of the partial diagonalization of $\mathbf{F}_\infty$ and $\mathbf{F}_*$. Taking the limit of Equation 1.89 as κ approaches infinity will allow the terms involving κ terms to drop out of the log-likelihood as they are now all in a denominator, and the whole term will become infinitely small. The diffuse log-likelihood is then,

$$-2\ln L(\boldsymbol{\theta}) \propto \sum_{i=1}^{N} nd\log|k| + \sum_{t=1}^{T} ny\log 2\pi + \log|\mathbf{F}_{it}| + \boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it} \tag{1.90}$$

and Koopman (1997) further showed that the -2 log-likelihood becomes,

$$-2\ln L(\boldsymbol{\theta}) \propto \sum_{i=1}^{N}\sum_{t=1}^{T} \log|\mathbf{F}_{*,it}^- + \mathbf{F}_{\infty,it}^-| + \boldsymbol{\epsilon}_{it}'\mathbf{F}_{*,it}^-\boldsymbol{\epsilon}_{it} \tag{1.91}$$

as $\sum_{i=1}^{N}\sum_{t=1}^{T} \boldsymbol{\epsilon}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{\epsilon}_{it} = \sum_{i=1}^{N}\sum_{t=1}^{T} \boldsymbol{\epsilon}_{it}'\mathbf{F}_{*,it}^-\boldsymbol{\epsilon}_{it}$ as $\kappa \to \infty$. For the cases where $\mathbf{F}_{\infty,it}$ equals zero or is non-singular, the exact prediction equations are

$$\mathbf{x}_{i,(t+1)PRED} = \mathbf{T}\mathbf{x}_{itFILT} \tag{1.92}$$

$$\mathbf{P}_{(*,t+1)PRED} = \mathbf{T}\mathbf{P}_{*,tPRED}\mathbf{T}' + \boldsymbol{\Sigma}_v \tag{1.93}$$

$$\mathbf{P}_{(\infty,t+1)PRED} = \mathbf{T}\mathbf{P}_{\infty,tPRED}\mathbf{T}'. \tag{1.94}$$

For the case where $\mathbf{F}_{\infty,it}$ is non-singular, the filtering equations are

$$\mathbf{x}_{itFILT} = \mathbf{x}_{itPRED} + \mathbf{K}_{\infty,t}\boldsymbol{\epsilon}_{it} \tag{1.95}$$

$$\mathbf{P}_{(*,t)FILT} = \mathbf{P}_{*,tPRED} - \mathbf{K}_{\infty,t}\mathbf{M}'_{*,t} - \mathbf{K}_{*,t}\mathbf{M}'_{\infty,t} \tag{1.96}$$

$$\mathbf{P}_{(\infty,t)FILT} = \mathbf{P}_{\infty,tPRED} - \mathbf{K}_{\infty,t}\mathbf{M}'_{\infty,t} \tag{1.97}$$

where

$$\mathbf{K}_{*,t} = (\mathbf{M}_{*,t} - \mathbf{K}_{\infty,t}\mathbf{F}_{*,it})\mathbf{F}^{-1}_{\infty,it} \tag{1.98}$$

$$\mathbf{K}_{\infty,t} = \mathbf{M}_{\infty,t}\mathbf{F}^{-1}_{\infty,it}. \tag{1.99}$$

For the case where $\mathbf{F}_{\infty,it}$ is zero, the filtering equations reduce to

$$\mathbf{x}_{itFILT} = \mathbf{x}_{itPRED} + \mathbf{K}_{*,t}\boldsymbol{\epsilon}_{it} \tag{1.100}$$

$$\mathbf{P}_{(*,t)FILT} = \mathbf{P}_{*,tPRED} - \mathbf{K}_{*,t}\mathbf{M}'_{*,t} \tag{1.101}$$

$$\mathbf{P}_{(\infty,t)FILT} = \mathbf{P}_{\infty,tPRED} \tag{1.102}$$

where

$$\mathbf{K}_{*,t} = \mathbf{M}_{*,t}\mathbf{F}^{-1}_{\infty,it}. \tag{1.103}$$

Koopman (1997) next explained that the extra equations do not need to be used after a time period d when $\mathbf{P}_{\infty,d+1}$ becomes zero. This is because $\mathbf{P}_t$ is no longer dependent on κ after time $t = d$. In this sense, the extra equations may be thought of as providing the exact initialization of the regular KF at time $t = d + 1$.

The log-likelihood functions used in this approach depend on whether $\mathbf{F}_{\infty,it}$ is non-singular, zero, or singular. Here I will present the functions for the non-singular and zero case, as Koopman and Durbin (2003) state that these two cases "apply to

nearly all time series occurring in practice" (p. 88). When $\mathbf{F}_{\infty,it}$ is non-singular, $\mathbf{F}^{-}_{*,t} = 0$ and $\mathbf{F}^{-}_{\infty} = \mathbf{F}^{-}_{\infty}$. When $\mathbf{F}_{\infty,it}$ is zero, $\mathbf{F}^{-}_{*,t} = \mathbf{F}^{-1}_{*,t}$ and $\mathbf{F}^{-}_{\infty} = 0$. Given this, the log-likelihoods are, aside from a constant,

$$\text{Non-Singular Case: } -2\ln L(\boldsymbol{\theta}) \propto \sum_{i=1}^{N}\sum_{t=1}^{d} \log|\mathbf{F}^{-1}_{\infty,it}| + \sum_{i=1}^{N}\sum_{t=d+1}^{T} \log|\mathbf{F}_{it}| + \boldsymbol{\epsilon}'_{it}\mathbf{F}^{-1}_{it}\boldsymbol{\epsilon}_{it} \tag{1.104}$$

$$\text{Zero Case: } -2\ln L(\boldsymbol{\theta}) \propto \sum_{i=1}^{N}\sum_{t=1}^{d} \log|\mathbf{F}^{-1}_{*,it}| + \boldsymbol{\epsilon}'_{it}\mathbf{F}^{-1}_{*,it}\boldsymbol{\epsilon}_{it} + \sum_{i=1}^{N}\sum_{t=d+1}^{T} \log|\mathbf{F}_{it}| + \boldsymbol{\epsilon}'_{it}\mathbf{F}^{-1}_{it}\boldsymbol{\epsilon}_{it}. \tag{1.105}$$

Comparison of diffuse approaches. To better understand the differences between the methods for using a diffuse initial state, consider Table 1.1, which compares three methods for using a diffuse initial state: 1. the original KF, 2. DFK by De Jong (1991), and 3. exact initial KF with a non-singular $\mathbf{F}_{\infty,it}$ by Koopman (1997). To aid presentation, I used the same initial state equation for deriving the KF initial conditions for the prediction equations. κ appears only in the original KF, as the DFK and exact initial KF were derived such that the recursions would be invariant of κ. For the DKF there are extra dimensions for the matrices in the measurement equation, prediction equation for the expected value of the states, and innovation equation. There is also the extra recursion of $\mathbf{Q}_t$, which is not used in the other two approaches. For the Koopman (1997) approach, there are extra equations for the prediction equation for the covariance matrix of the states and the equation for the covariance matrix of the innovations. A similarity between the DKF and the exact initial KF is that they both eventually "collapse" to the original KF. For the DKF, this occurs when an element of the $\mathbf{Q}_t$ matrix, $\mathbf{S}^{-1}$, is nonsingular, which is usually at the time point $t = nd$. For the exact initial KF, this occurs when $\mathbf{P}_{\infty,t}$ is a zero matrix, which usually occurs at time point $t = nd$.

Koopman et al. (2008) explained that the de Jong and Koopman approaches yield numerically equivalent results. Thus, the log-likelihood values would ultimately be the same. However, de Jong's approach may be easier to implement as it is simply the ordinary KF with some augmented vectors and matrices. Also, Koopman et al. (2008) explained that de Jong's approach might be more useful when dealing with fixed regression effects, which I have not considered. However, Koopman (1997) stated that de Jong's approach may be more computationally inefficient due to the extra recursions yielding the matrix $\mathbf{Q}$ that needs to be inverted at each step until a collapse. Furthermore, Koopman (1997) explained that the equations for the KS, which I did not show, are more complicated for de Jong's approach.

In summary, de Jong's approach computes initial values estimates of the diffuse states from an initial stretch of the data via the use of the augmented equations, and then proceeds with the ordinary KF. Koopman's approach explicitly writes out the KF equations and provides an alternative method for taking the inverse of the innovation covariance matrix, $\mathbf{F}$, that involves a partial diagonalization procedure. Despite both approaches yielding identical results, it is not clear which approach will be computationally most efficient and easiest to implement given the models I will consider.

In general, depending on the type of initial condition specification used, state and parameter estimates may be affected. The degree to which this will affect estimation results depends on the number of time points such that with fewer time point there will be more noticeable discrepancies. However, the degree to which parameter estimates are affected due to using a misspecified initial condition distribution is not currently known. In this thesis, investigated the effects of using different initial condition specifications given cases where the number of time points is relatively small and the number of time points is relatively large.

Table 1.1: Kalman Filter Comparisons

Initial State Equation:	$\mathbf{x}_0 = \boldsymbol{\mu} + \mathbf{A}_\infty \boldsymbol{\delta} + \mathbf{A}_* \boldsymbol{\eta} \qquad \boldsymbol{\eta} \sim (0, \mathbf{P}_*)$		$\boldsymbol{\delta} \sim (0, \kappa \mathbf{I})$, with $\kappa \to \infty$
Equation	Regular KF	Diffuse KF De Jong	Exact Initial KF Koopman, $\mathbf{F}_{\infty,t}$ non-singular
y_{it}	$\boldsymbol{\tau} + \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it}$	$[\mathbf{0}_{nd} \quad \boldsymbol{\tau} + \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it}]$	$\boldsymbol{\tau} + \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it}$
$\mathbf{x}_{i1PRED}$	$\mathbf{T}\boldsymbol{\mu}$	$\mathbf{T}[\mathbf{A}_\infty \quad \boldsymbol{\mu}]$	$\mathbf{T}\boldsymbol{\mu}$
$\mathbf{P}_{1PRED}$	$\mathbf{T}(\kappa \mathbf{A}_\infty \mathbf{A}_\infty' + \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*')\mathbf{T}' + \boldsymbol{\Sigma}_v$	$\mathbf{T}\mathbf{A}_* \mathbf{P}_* \mathbf{A}_*' \mathbf{T}' + \boldsymbol{\Sigma}_v$	$\mathbf{P}_{\infty,1PRED} = \mathbf{T}\mathbf{A}_\infty \mathbf{A}_\infty' \mathbf{T}'$ $\mathbf{P}_{*,1PRED} = \mathbf{T}\mathbf{A}_* \mathbf{P}_* \mathbf{A}_*' \mathbf{T}' + \boldsymbol{\Sigma}_v$
$\boldsymbol{\epsilon}_{it}$	$\mathbf{y}_{it} - \boldsymbol{\tau} - \mathbf{Z}\mathbf{x}_{itPRED}$	$[\mathbf{0}_{nd} \quad \boldsymbol{\tau} + \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it}] - \mathbf{Z}[\mathbf{A}_{\infty t} \quad \mathbf{x}_{itPRED}]$	$\mathbf{y}_{it} - \boldsymbol{\tau} - \mathbf{Z}\mathbf{x}_{itPRED}$
$\mathbf{F}_t$	$\mathbf{Z}\mathbf{P}_{tPRED}\mathbf{Z}' + \boldsymbol{\Sigma}_u$	$\mathbf{Z}\mathbf{P}_{tPRED}\mathbf{Z}' + \boldsymbol{\Sigma}_u$	$\mathbf{F}_{*,t} = \mathbf{Z}\mathbf{P}_{*,t}\mathbf{Z}' + \boldsymbol{\Sigma}_u$ $\mathbf{F}_{\infty,t} = \mathbf{Z}\mathbf{P}_{\infty,t}\mathbf{Z}'$
$\mathbf{K}_t$	$\mathbf{P}_{tPRED}\mathbf{Z}'\mathbf{F}_{it}^{-1}$	$\mathbf{P}_{tPRED}\mathbf{Z}'\mathbf{F}_{it}^{-1}$	$\mathbf{K}_{*,t} = (\mathbf{M}_{*,t} - \mathbf{K}_{\infty,t}\mathbf{F}_{*,t})\mathbf{F}_{\infty,it}^{-1}$ $\mathbf{K}_{\infty,t} = \mathbf{M}_{\infty,t}\mathbf{F}_{\infty,it}^{-1}$
$\mathbf{Q}_{t+1}$	- -	$\mathbf{Q}_t + [\mathbf{0}_{nd} \quad \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it}] - \mathbf{Z}[\mathbf{A}_{\infty t} \quad \mathbf{x}_{itPRED}]'$ $\mathbf{F}_{it}^{-1}[\mathbf{0}_{nd} \quad \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it}] - \mathbf{Z}[\mathbf{A}_{\infty t} \quad \mathbf{x}_{itPRED}]$	-
$\mathbf{x}_{i,(t+1)PRED}$	$\boldsymbol{\pi} + \mathbf{T}\mathbf{x}_{itFILT}$	$\mathbf{T}[\mathbf{A}_t \quad \boldsymbol{\pi} + \mathbf{x}_{itFILT}]$	$\boldsymbol{\pi} + \mathbf{T}\mathbf{x}_{itFILT}$
$\mathbf{P}_{(t+1)PRED}$	$\mathbf{T}\mathbf{P}_{tFILT}\mathbf{T}' + \boldsymbol{\Sigma}_v$	$\mathbf{T}\mathbf{P}_{tFILT}\mathbf{T}' + \boldsymbol{\Sigma}_v$	$\mathbf{T}\mathbf{P}_{*,tPRED}\mathbf{T}' + \boldsymbol{\Sigma}_v$ $\mathbf{T}\mathbf{P}_{\infty,tPRED}\mathbf{T}'$

Note KF=Kalman Filter, $\boldsymbol{\mu}$ and $\mathbf{P}$ are the unconditional mean vector and covariance matrix for a stationary time-series

1.8 Examples of Initial Condition Specification in the Literature

As discussed in Section 1.2, researchers implementing models for analyzing longitudinal data have used a variety of methods for initializing the process. To review, there are four main general approaches to initializing a process: 1. using the unconditional mean vector and covariance matrix, 2. freely estimating parameters in the initial distribution, 3. setting the initial condition parameters to zero, 4. using a diffuse initial condition approach. In this section I will revisit some of the examples of initial condition implementation described in earlier sections and explicitly show how initial condition was specified, with an emphasis on showing what elements are populated in the matrices of Equation 1.69. I will also describe examples from the state-space literature.

Examples from the psychological literature. As discussed, traditional growth curve models structured in the SEM framework (Meredith & Tisak, 1990; Bollen & Curran, 2006) have initialized processes by estimating the parameters in the initial condition distribution. This can be more clearly illustrated if we formulate the model in state-space form. Drawing from the SEM formulation using the set of equations in Equation 1.17, the LCM may be expressed in state-space form as

$$y_{it} = \begin{bmatrix} 1 & TIME_{it} \end{bmatrix} \begin{bmatrix} x_{INTit} \\ x_{SLPit} \end{bmatrix} + \begin{bmatrix} u_{it} \end{bmatrix} \tag{1.106}$$

$$\begin{bmatrix} x_{INTit} \\ x_{SLPit} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{INTi,t-1} \\ x_{SLPi,t-1} \end{bmatrix}. \tag{1.107}$$

When expressed this way, an initial condition distribution for the two latent states,

x_{INTit} and x_{SLPit}, is specified as

$$\mathbf{x}_0 = \begin{bmatrix} x_{INTi0} \\ x_{SLPi0} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{x_{INT}} \\ \mu_{x_{SLP}} \end{bmatrix}, \begin{bmatrix} \sigma^2_{x_{INT}} & \sigma^2_{x_{INT}x_{SLP}} \\ \sigma^2_{x_{SLP}x_{INT}} & \sigma^2_{x_{SLP}} \end{bmatrix} \right) \tag{1.108}$$

such that this model now corresponds exactly to the latent growth curve model expressed in the SEM framework. From Equation 1.69, we can see that the following matrices would be specified:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_{x_{INT}} \\ \mu_{x_{SLP}} \end{bmatrix}, \quad \mathbf{A}_\infty = \mathbf{0}, \quad \mathbf{A}_* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{P}_0 = \begin{bmatrix} \sigma^2_{x_{INT}} & \sigma_{x_{INT}x_{SLP}} \\ \sigma_{x_{SLP}x_{INT}} & \sigma^2_{x_{SLP}} \end{bmatrix}. \tag{1.109}$$

The dual change score model used by Chow et al. (2010) also used the approach of freely estimating the parameters in the initial condition distribution. The model expressed in state-space form has a transition equation expressed as

$$\begin{bmatrix} \mathbf{x}_{INTit} \\ \mathbf{x}_{SLPit} \end{bmatrix} = \begin{bmatrix} 1+\beta & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{INTi,t-1} \\ \mathbf{x}_{SLPi,t-1} \end{bmatrix} \tag{1.110}$$

with the following measurement model:

$$\mathbf{y}_{it} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{INTit} \\ \mathbf{x}_{SLPit} \end{bmatrix} + u_{it}. \tag{1.111}$$

To initialize this process, Chow et al. (2010) constructed an initial condition distribution for the two state variables, $\mathbf{x}_{INTit}$ and $\mathbf{x}_{SLPit}$ with a structure of

$$\mathbf{x}_0 \sim N\left(\begin{bmatrix} \mu_{x_{INT}} \\ \mu_{x_{SLP}} \end{bmatrix}, \begin{bmatrix} \sigma^2_{x_{INT}} & \sigma_{x_{INT}x_{SLP}} \\ \sigma_{x_{SLP}x_{INT}} & \sigma^2_{x_{SLP}} \end{bmatrix} \right). \tag{1.112}$$

The elements in this initial condition matrix are then freely estimated, which corresponds to the solution of freely estimating parameters in the initial condition covariance matrix. The elements in the matrices of Equation 1.69 are as follows:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_{x_{INT}} \\ \mu_{x_{SLP}} \end{bmatrix}, \quad \mathbf{A}_{\infty} = \mathbf{0}, \quad \mathbf{A}_{*} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{P}_0 = \begin{bmatrix} \sigma^2_{x_{INT}} & \sigma_{x_{INT}x_{SLP}} \\ \sigma_{x_{SLP}x_{INT}} & \sigma^2_{x_{SLP}} \end{bmatrix}. \quad (1.113)$$

Chow et al. (2010) also considered a direct autoregressive factor score (DAFS; see Nesselroade, McArdle, Aggen, & Meyers, 2001) model which contains two sets of AR(1) processes at the latent level. Each latent state displays an AR(1) process and cross regressions among the states are estimated. The model in state-space form is expressed as:

$$\begin{bmatrix} x_{1it} \\ x_{2it} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} x_{1i,t-1} \\ x_{2i,t-1} \end{bmatrix} + \begin{bmatrix} v_{1it} \\ v_{2it} \end{bmatrix} \quad (1.114)$$

with the following measurement model:

$$\begin{bmatrix} y_{1it} \\ y_{2it} \\ y_{3it} \\ y_{4it} \\ y_{5it} \\ y_{6it} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ Z_{21} & 0 \\ Z_{31} & 0 \\ 0 & 1 \\ 0 & Z_{42} \\ 0 & Z_{52} \end{bmatrix} \begin{bmatrix} x_{1it} \\ x_{2it} \end{bmatrix} + \begin{bmatrix} u_{1it} \\ u_{2it} \\ u_{3it} \\ u_{4it} \\ u_{5it} \\ u_{6it} \end{bmatrix}. \quad (1.115)$$

The authors then specified that a large constant be used for κ, which specifies the

elements of the matrices in Equation 1.69 to be:

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{A}_{\infty} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{A}_{*} = \mathbf{0}, \mathbf{P}_0 = \begin{bmatrix} 1000 & 0 \\ 0 & 1000 \end{bmatrix}. \tag{1.116}$$

This is one of the more commonly used approaches in the state-space literature, especially for nonstationary processes. For large T, the parameter estimates appear to be satisfactory. However, this may not be the case for data with a relatively small T, which is a feature of panel data studies in psychological research. In this thesis I will explore this issue and determine whether a large κ approach produces satisfactory results when using small T, or whether a different diffuse approach, such as de Jong's DKF and Koopman's exact initial KF approach, would produce more acceptable results.

Du Toit and Browne (2007) implemented and compared three procedures for initializing a VARMA process estimated in the SEM framework. When assuming the process is stationary, they used Equation 1.63 to derive the unconditional mean and covariance matrix to be used as initial condition (i.e., $\boldsymbol{\mu}_0 = \boldsymbol{\mu}$ and $\mathbf{P}_0 = \mathbf{P}$). They also considered a condition where they allowed for a change of process prior to the first observation, but assumed a stationary process onward. This is equal to the solution that freely estimates the parameters of the initial condition distribution, i.e., parameters contained in $\boldsymbol{\mu}_0$ and $\mathbf{P}_0$ are freely estimated. The third solution (Du Toit & Browne, 2007) consider is assuming the process truly started at $t = 1$. This amounts to using a null vector and matrix for the mean and covariance matrices of the initial condition distribution, i.e., $\boldsymbol{\mu}_0 = 0$ and $\mathbf{P}_0 = 0$. To compare the three initialization methods, they fit three VARMA(1,1) models, each one using a differen initial condition specification, to an empirical data set and examined the results. The specification where the parameters are freely estimated in the initial condition distribution provided the best fit, while the null condition provided the worst fit. This may be be-

cause the process is not stationary, and the freely estimated specification corresponds most closely with a nonstationary process. They also conducted a χ^2 difference test as the model-implied specification and null condition are nested under the freely estimated condition. The results indicated that the freely estimated condition provided better fit above and beyond what would be expected by change alone, thus indicating that there was most likely a change in process before the first observation. However, the use of a diffuse prior was not considered, and may have provided even better fit if the process is truly nonstationary.

Examples from the state-space literature. Initializing procedures from the state-space modeling literature more frequently include some variation of a diffuse initial condition. For example, De Jong (2003, see p. 144) considered a single-subject model that contains one process that is stationary and one process that is nonstationary, expressed as,

$$y_t = \begin{bmatrix} 1 & 1 \end{bmatrix} \mathbf{x}_t + v_t \tag{1.117}$$

$$\mathbf{x}_{t+1} = \begin{bmatrix} 1 & 0 \\ 0 & e^{\phi h_t} \end{bmatrix} \mathbf{x}_t + \mathbf{u}_t. \tag{1.118}$$

De Jong (2003) describe this model as being a random walk plus an AR(1) process where the error terms are correlated. The random walk is then nonstationary while the AR(1) process is stationary. They initialize the nonstationary process using the DKF method and the stationary process using the unconditional mean and covariance matrix method. Given Equation 1.69, they initialize the procedure with

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{A}_\infty = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{A}_* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{P}_{\infty,0} = 0, \mathbf{P}_{*,0} = \frac{1}{-2\phi^3} \tag{1.119}$$

where zero and $\frac{1}{-2\phi^3}$ represent the unconditional mean and variance for the station-

ary process of the model.

Koopman (1997) illustrated several models where he initializes the initial condition using his exact KF approach. One such model is a single-subject time series model with a local linear trend component (see p. 1634). The transition equation is expressed as

$$\begin{bmatrix} \mathbf{x}_{INTit} \\ \mathbf{x}_{SLPit} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{INTi,t-1} \\ \mathbf{x}_{SLPi,t-1} \end{bmatrix} + \begin{bmatrix} v_{x_{INTit}} \\ v_{x_{SLPit}} \end{bmatrix} \tag{1.120}$$

and measurement model is expressed as

$$\mathbf{y}_{it} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{INTit} \\ \mathbf{x}_{SLPit} \end{bmatrix} + u_{it}. \tag{1.121}$$

Given that both states, $\mathbf{x}_{INTit}$ and $\mathbf{x}_{SLPit}$, in this model are nonstationary, Koopman (1997) initiates the process by assuming a diffuse prior for both states. Thus, when using the exact KF, the process is initialized with the following:

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{A}_{\infty} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{A}_{*} = \mathbf{0}, \mathbf{P}_{\infty,0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{P}_{*,0} = 0. \tag{1.122}$$

1.9 Current Investigation

Past research (i.e., Oud et al., 1990; Du Toit & Browne, 2001, 2007; De Jong, 1991; Koopman, 1997; Harvey, 1991) and the analytical expressions indicate that it is important, especially with few time points and nonstationarity, to correctly specify initial condition. Furthermore, no one has explicitly shown the consequences of using different specifications for initial condition when there are both stationary and nonstationary elements in the model. In psychological research there exist both

nonstationary processes and a design containing a relatively small number of time points, which makes it even more important to consider these issues. Past research highlights the difficulties in estimating state space models and time series models as SEMs. The different approaches all carry some advantages and disadvantages, although no systematic and thorough examination of the effects of initial condition specification have been conducted in the context of using these models in SEM. Thus, it is not known whether using a diffuse prior, either with the original KF, DKF, or exact initial KF, would be beneficial. Also, it is unclear how this might be initialized, as doing so using the KF and PED may require adding the augmented equations or extra recursions (De Jong, 1991; Koopman, 1997). The overall goal of this thesis was to conduct an examination of the effects of using different initial condition specifications for both stationary and nonstationary models that may be formulated as state-space models and estimated within the SEM framework. Unique contributions of this thesis include (1) comparing different ways of initializing state-space models and equivalent structural equation models for both intensive repeated measures data and panel data using the raw data likelihood approach and (2) comparing methods proposed in the state-space literature within the context of latent variable models within the SEM framework.

CHAPTER 2

Methods

For this thesis I evaluated and compared different methods for initializing SEM models incorporating time series processes formulated as state-space models. I conducted a Monte Carlo simulation study varying the type of initial condition specification administered for both a stationary and non-stationary process. The two models used were process factor analysis (PFA; Browne & Nesselroade, 2005) models with vector AR processes and no MA process, also called a direct autoregressive factor score (DAFS; see Nesselroade et al., 2001) model, which is actually a special case of the PFA model with an AR(p) process and no MA process. For each initial condition specification and model, I examined two data structures: one that resembles an intensive repeated measures study with a small sample size, $N = 20$, and large number of time points, $T = 50$, and one that resembles a panel data study with a large sample size, $N = 200$, and small number of time points, $T = 5$. Table 2.1 contains information regarding the different manipulated conditions in the simulation study, each of which will be described in greater detail.

I also illustrated and compared the implementation of all approaches using an empirical example. Specifically, I analyzed an empirical data set from Van Vuuren, De Beer, and Du Toit (1982) that was also used by Du Toit and Browne (2007) using some of the initial condition specifications considered in this manuscript. Unlike Du Toit and Browne (2007), however, I fitted a PFA(1,0) model to the empirical data.

Table 2.1: Summary of Simulation Design

	Model	True IC Specification	Fitted IC Specification	NT
Stationary	Model 1: PFA(1,0) with 2 stationary vector AR(1) processes	Model-Implied	1. Model-Implied	$N = 20, T = 50$
			2. Free-Parameter	$N = 200, T = 5$
		Free-Parameter	3. Null Initial Condition	
			4. De Jong (1991) DKF	
		Null Initial Condition	5. Koopman (1997) exact initial KF	
			7. Large κ approximation	
Non-Stationary	Model 2: PFA(1,0) with 2 nonstationary vector AR(1) processes	Free-Parameter	2. Free-Parameter	$N = 20, T = 50$
			3. Null initial condition	$N = 200, T = 5$
		Null initial condition	4. De Jong (1991) DKF	
			5. Koopman (1997) exact initial KF	
		Diffuse: Mild Stationarity	6. De Jong (1991) DKF	
			7. Large κ approximation	
		Diffuse: Moderate Stationarity		$N = 200, T = 5$

Note PFA=Process factor analysis, AR=autoregression

2.1 Models

I examined two dynamic factor analysis (DFA; Molenaar, 1985; Nesselroade et al., 2001) models; one that is stationary and one that is non-stationary. The models considered were PFA models, which combine a factor analytic model and a VARMA(p,q) time series model to describe the dynamics of the latent variable series. These models will not have a moving average process but will contain a vector autoregressive (VAR) process.

Model 1: PFA with stationary VAR(1) process. The first model examined was a PFA(1,0) model that contains a vector autoregressive process of order 1. In state-space modeling form, this model may be expressed as

$$\mathbf{y}_{it} = \mathbf{Z}\mathbf{x}_{it} + \mathbf{u}_{it} \tag{2.1}$$

$$\mathbf{x}_{it} = \mathbf{A}\mathbf{x}_{i,t-1} + \mathbf{v}_{it} \tag{2.2}$$

where $\mathbf{y}_{it}$ contains the sets of observed variables, $\mathbf{Z}$ is a $ny \times nx$ matrix of factor loadings where ny is the number of observed variables and nx is the number of sets of latent series, $\mathbf{x}_{it}$ contains the sets of latent time series, or factors, $\mathbf{A}$ is the $nx \times nx$ transition matrix containing the autoregression and cross-regression weights, $\mathbf{u}_{it}$ is a $ny \times 1$ vector of measurement error variables, and $\mathbf{v}_{it}$ is a $nx \times 1$ vector of process noise variables. Both models will have three indicator variables per latent variable.

The population values for the stationary model are chosen to be

$$\begin{bmatrix} \mathbf{y}_{1,it} \\ \mathbf{y}_{2,it} \\ \mathbf{y}_{3,it} \\ \mathbf{y}_{4,it} \\ \mathbf{y}_{5,it} \\ \mathbf{y}_{6,it} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1.2 & 0 \\ .8 & 0 \\ 0 & 1 \\ 0 & .9 \\ 0 & 1.1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1,it} \\ \mathbf{x}_{2,it} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{1,it} \\ \mathbf{u}_{2,it} \\ \mathbf{u}_{3,it} \\ \mathbf{u}_{4,it} \\ \mathbf{u}_{5,it} \\ \mathbf{u}_{6,it} \end{bmatrix}, \quad \mathbf{u}_t \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} .8 & 0 & 0 & 0 & 0 & 0 \\ 0 & .6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & .8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & .4 \end{bmatrix} \right) \tag{2.3}$$

$$\begin{bmatrix} \mathbf{x}_{1,it} \\ \mathbf{x}_{2,it} \end{bmatrix} = \begin{bmatrix} .5 & -.1 \\ -.3 & .6 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1i,t-1} \\ \mathbf{x}_{2i,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{1,it} \\ \mathbf{v}_{2,it} \end{bmatrix} \quad \mathbf{v}_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & .4 \\ .4 & 1 \end{bmatrix} \right). \tag{2.4}$$

This model has two AR processes of order 1 for two distinct latent series, which allows for the previous latent states to affect the current latent states. Also, there are cross loadings from one set of latent factors to the other such that a latent factor at time t predicts the latent factor from the other series at time $t+1$. In order to decipher whether the model is stationary, I calculated the roots of the determinant of Equation 1.7, which were 1.37 and 2.70. They both have modulus greater than unity making this a stationary process. Using Equation 1.69, the following matrices would be specified for the model-implied, free-parameter, and null initial condition specifications,

$$\mathbf{A}_\infty = \mathbf{0}, \quad \mathbf{A}_* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.5}$$

Table 2.3 illustrates the elements that would populate the initial condition mean vector, $\boldsymbol{\mu}_0$, and initial condition covariance matrix, $\mathbf{P}_0$, given the different initial condition specifications.

Model 2: PFA with nonstationary VAR(1) process. The second model was also a PFA(1,0) process, but this time the vector autoregressive process is nonstationary. The measurement model contains the same structure and population values as Equation 2.3, however the transition equation now contains the following population values,

$$\begin{bmatrix} \mathbf{x}_{1,it} \\ \mathbf{x}_{2,it} \end{bmatrix} = \begin{bmatrix} 1.2 & -.4 \\ .6 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1i,t-1} \\ \mathbf{x}_{2i,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{1,it} \\ \mathbf{v}_{2,it} \end{bmatrix} \qquad \mathbf{v}_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & .4 \\ .4 & 1 \end{bmatrix} \right). \tag{2.6}$$

Calculating the roots of the determinant of Equation 1.7 led to values of .833 and .833, and the absolute value of both values is not greater than 1, verifying that this model is nonstationary. Using Equation 1.69, the following matrices would be specified for the model-implied, free-parameter, and null initial condition specifications,

$$\mathbf{A}_\infty = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{A}_* = \mathbf{0}. \tag{2.7}$$

The top half of Table 2.4 illustrates the elements that would populate the initial condition mean vector, $\boldsymbol{\mu}_0$, and initial condition covariance matrix, $\mathbf{P}_0$, given the different initial condition specifications.

Because most replications failed to converge in the condition with intensive longitudinal data (i.e., $N = 20$ and $T = 50$), a simulation condition was added where the population generating values yielded simulated trajectories that were not as explosive and exhibited mild nonstationarity; i.e., the values towards the end of the time frame were not exceedingly low or high when compared to the initial values. Still, the process is nonstationary as the roots of the determinant of Equation 1.7 both have

modulus less than unity. The values used in this mild nonstationarity case were

$$\begin{bmatrix} \mathbf{x}_{1,it} \\ \mathbf{x}_{2,it} \end{bmatrix} = \begin{bmatrix} 1.2 & -.4 \\ .6 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1i,t-1} \\ \mathbf{x}_{2i,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{1,it} \\ \mathbf{v}_{2,it} \end{bmatrix} \qquad \mathbf{v}_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} .05 & .02 \\ .02 & .05 \end{bmatrix} \right). \tag{2.8}$$

As this set of parameter values produces a true trajectory that is less explosive, I will refer to it as the condition with mild nonstationarity and the nonstationary model described in Equation 2.6 as the condition with moderate nonstationarity.

In summary, there are two overall models being considered; a stationary PFA(1,0) model a nonstationary PFA(1,0) model. The nonstationary model has two sets of parameter values; one with mild nonstationarity and one with moderate nonstationarity. Note that all models have the same form, the difference in stationarity and degree of nonstationarity is a result of the varying true parameter values.

To further highlight the differences between stationary and nonstationary states with respect to the chosen models, consider Figure 2.1. In the top panel, row (A), the two stationary states are displayed from Model 1. Notice that the values of the y-axis do not have a large range as the process does not explode over time. The second panel, row (B), displays the two states from Model 2 for the moderate nonstationary case. Here, the process is clearly explosive and increases without bound as time increases. The third panel, row (C), displays the two states from Model 2 for the mildly nonstationary case. While still increasing without bound over time, this process is not as explosive as the values on the y-axis range from -6 to 6 for state 1 and -10 to 5 for state 2, as opposed to the moderately nonstationary condition where the values on the y-axis range from -200 to 200 for both states.

Figure 2.2 and Figure 2.3 display the autocorrelation and partial autocorrelation functions (ACF and PACFs) for a random person for each of the states of the stationary model, moderately nonstationary model, and mildly nonstationary model in panels (A), (B), and (C), respectively. The ACF is a standardized version of the au-

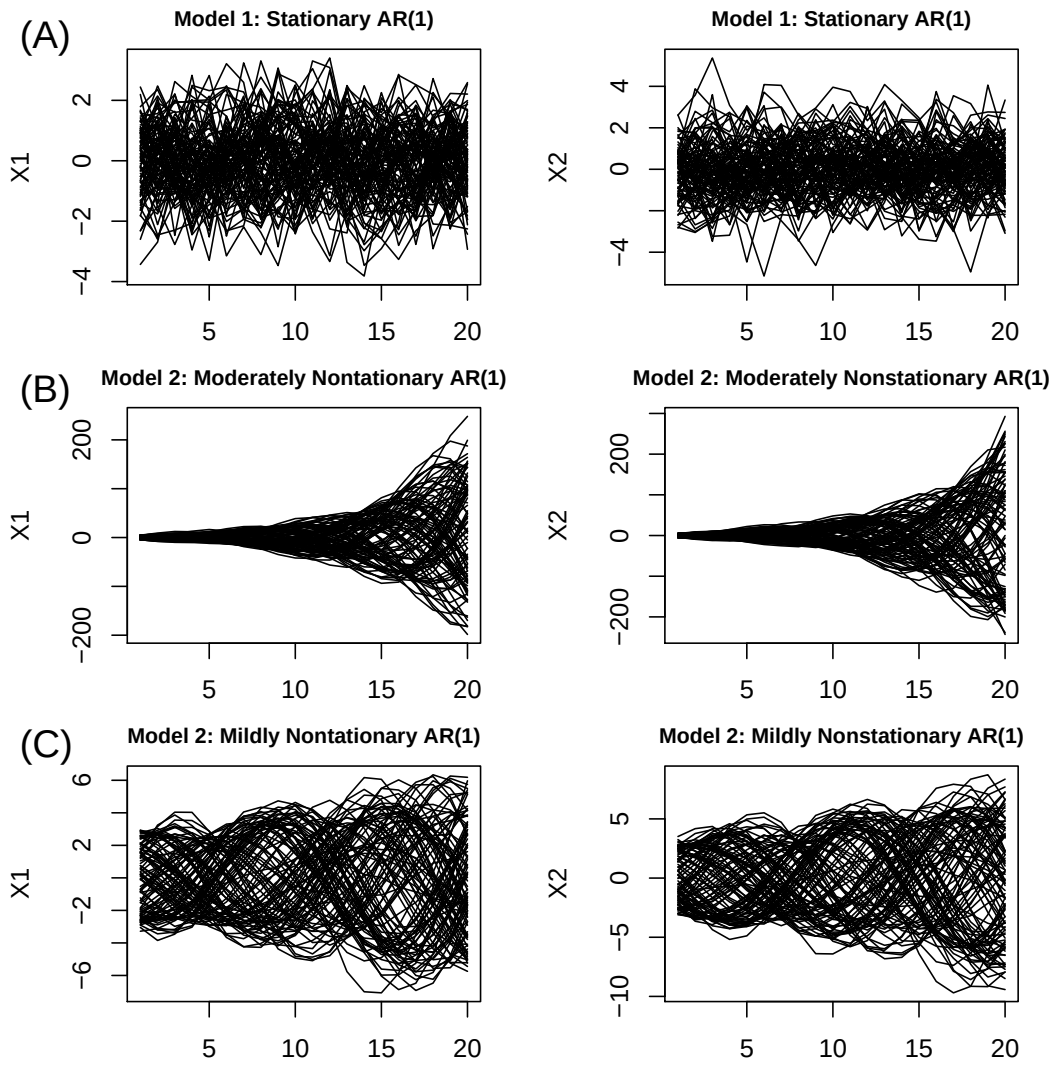

Figure 2.1: Time series plots of states for Model 1 (A), Model 2 with moderate nonstationarity (B), and Model 2 with mild nonstationarity. Time is displayed on the x-axis.

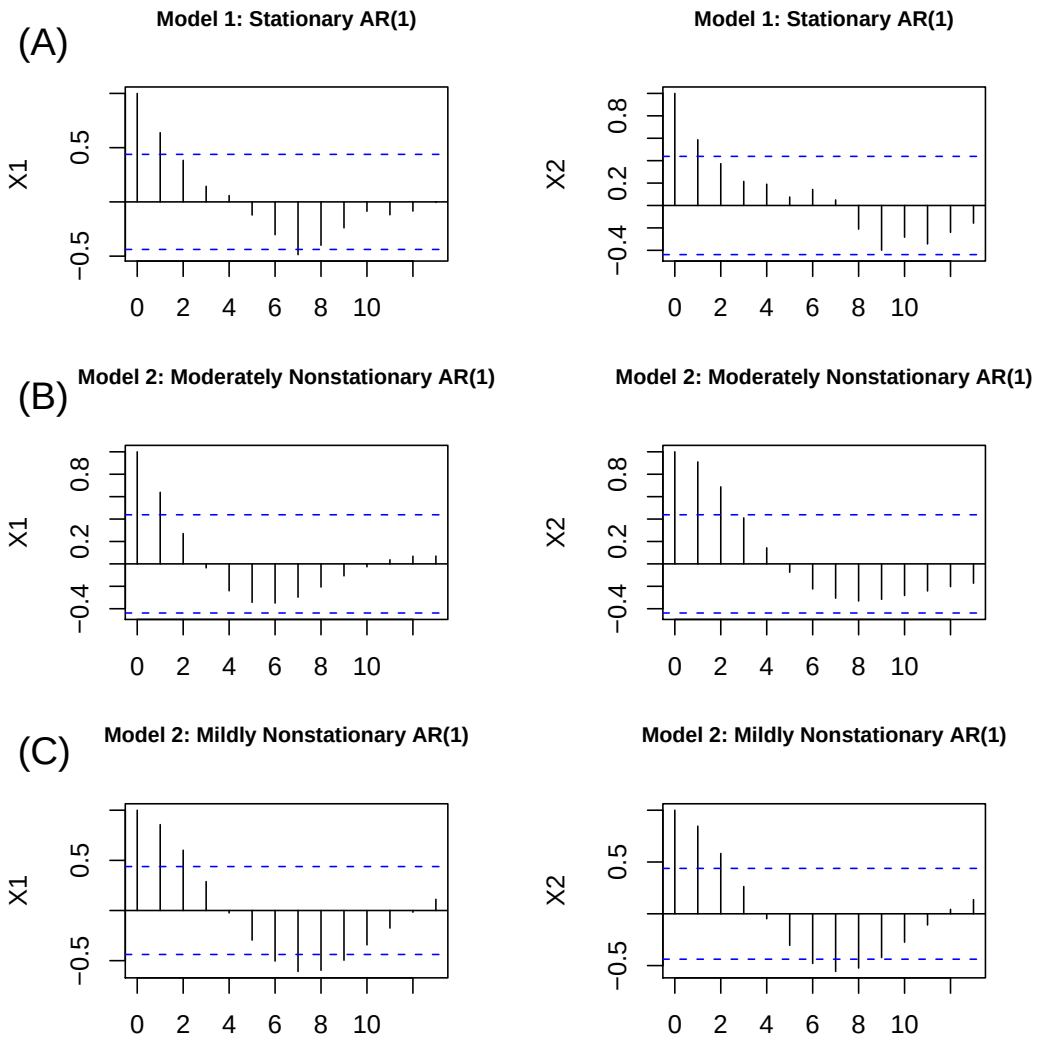

Figure 2.2: Autocorrelation plots of states for one individual for Model 1 (A), Model 2 with moderate nonstationarity (B), and Model 2 with mild nonstationarity. Lag is displayed on the x-axis.

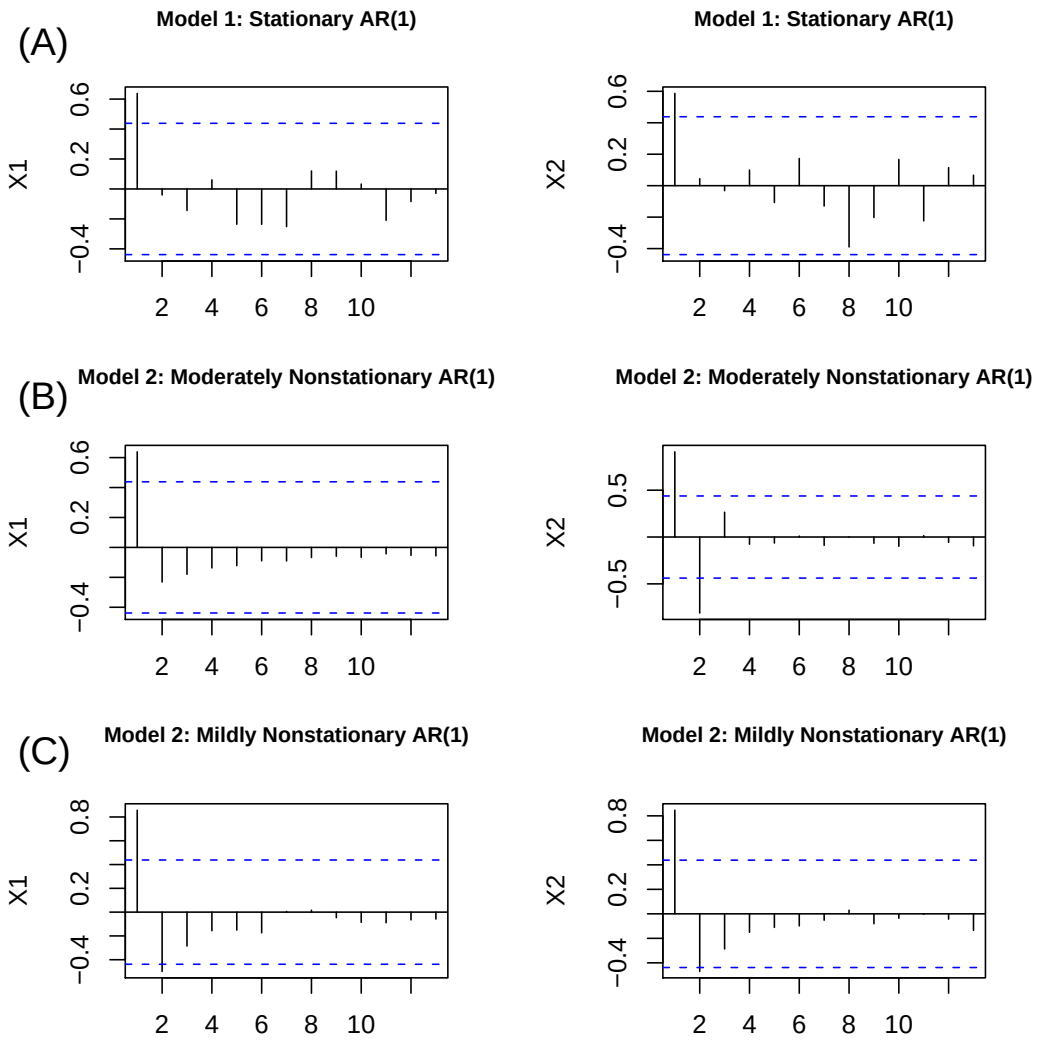

Figure 2.3: Partial autocorrelation plots of states for one individual for Model 1 (A), Model 2 with moderate nonstationarity (B), and Model 2 with mild nonstationarity. Lag is displayed on the x-axis.

tocovariance function given in Equation 1.8 from Section 1.1. It is a measure of the linear dependence between two points on the same time series observed at different times, and the values must be between -1 and 1. The PACF is a measure of the linear dependence between two points on the same time series with the linear dependence of previous lags removed. The PACF is particularly useful for examining a stationary AR time series process as the PACF value should be zero starting at lag $p + 1$. Examination of Figure 2.3 reveals that, given this person's observations, the stationary states do indeed follow such a structure as the values are statistically zero starting with a lag of 2.

Table 2.2: Summary of Initial Condition Specifications

Specification	Initial Distribution
1. Model-Implied: Assuming stationarity throughout	Unconditional model implied moments: $\mu_0 = \mu$ $\mathbf{P}_0 = \mathbf{P}$
2. Free-Parameter: Allow change in process before 1st observation	parameters of μ_0 and $\mathbf{P}_0$ freely estimated
3. Null Initial Condition: Assuming process started at 1st observation	$\mu_0 = \mathbf{0}$ and $\mathbf{P}_0 = \mathbf{0}$
4. De Jong (1991) DKF	$\mu_0 = [\mathbf{A}_\infty \quad \mu]$ $\mathbf{P}_0 = \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*'$
5. Koopman (1997) exact initial KF	$\mu_0 = \mu$ $\mathbf{P}_{\infty,0} = \mathbf{A}_\infty \mathbf{A}_\infty'$, $\mathbf{P}_{*,0} = \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*'$
6. Large κ approximation	$\mu_0 = \mathbf{0}$ $\mathbf{P}_0 = \kappa \mathbf{I}$ where κ is a large finite value

Note FIML=full information maximum likelihood, KF=Kalman filter, DKF=Diffuse Kalman filter

Note Initial condition distributions for Specification 5 and 6 are obtained from using the initial state equation from Equation 1.69

2.2 Initial Condition Specification

The different methods I considered for specifying a fitted initial condition depend on the stationarity of the model. For the stationary PFA, I examined three methods for specifying initial condition. See Table 2.2 for a list of the specifications and corresponding initial condition state vectors. I called the first specification a "model-implied" specification as it uses the unconditional model-implied moments as initial condition, i.e., $\mathbf{P}_0 = \mathbf{P}$ and $\boldsymbol{\mu}_0 = \boldsymbol{\mu}$. The specification assumes the data are stationary and is the method imposed by Du Toit and Browne (2001) using Equation 1.63 and Harvey (1991) using Equation 1.68. I called the second specification the "free-parameter condition" condition as it freely estimates parameters inside the initial condition matrices, i.e., parameters in $\mathbf{P}_0$ and $\boldsymbol{\mu}_0$ are freely estimated. This specification allows for a change in process to have occurred before the first observation, but assumes stationarity onward. This approach was employed by both Chow et al. (2010) and Du Toit and Browne (2001). De Jong (1991) also considered this case when he assumes the initial state vector is fixed but unknown. I called the third specification the "null initial condition" as it specifies the initial condition distribution to be zero. This specification assumes the process started at the time of the first observation and there are no prior influences, i.e., the model implied moments of the initial condition are set to zero ($\mathbf{P}_0$ and $\boldsymbol{\mu}_0 = 0$). This approach was employed by Du Toit and Browne (2001).

For the nonstationary PFA model, I examined the free-parameter and null initial condition specifications plus three extra methods that allow for a diffuse initial condition. The first diffuse method is using the DKF described and developed by De Jong and colleagues (1991, 1994, 2003). The second diffuse method is using the exact KF approach described by Koopman (1997). The third method is using a diffuse initial condition specification without using the appropriate likelihood functions, i.e., a "large κ approximation". This condition will serve as a comparison to the De Jong

(1991) and Koopman (1997) approaches that use likelihood functions derived for the purpose of accurately including a diffuse initial state.

Table 2.3 displays the values that will populate the initial condition distribution for Model 1 (the stationary PFA(1,0) model), while Table 2.4 displays the values that will populate the initial condition distribution for Model 2 given the diffuse approaches.

Data simulation procedure. Simulating stationary states was accomplished by imposing three true initial condition specifications: 1. populating the initial condition matrices, $\mathbf{P}_0$ and $\boldsymbol{\mu}_0$, with the unconditional mean and variance matrices to start the process, i.e., using the model-implied initial condition specification as the true initial condition specification, 2. populating the initial condition matrices, $\mathbf{P}_0$ and $\boldsymbol{\mu}_0$, with true population vales that are different from the unconditional mean and covariance matrices, i.e., using the free-parameter initial condition specification as the true initial condition specification, and 3. populating the initial condition matrices, $\mathbf{P}_0$ and $\boldsymbol{\mu}_0$, with values of zero, i.e., using the null initial condition specification as the true initial condition specification. When simulating data according to the free-parameter initial condition, the following true values were used to populate the initial condition matrices

$$\boldsymbol{\mu}_0 = \begin{bmatrix} 1 \\ .5 \end{bmatrix}, \mathbf{P}_0 = \begin{bmatrix} 1.2 & .3 \\ .3 & .7 \end{bmatrix}. \tag{2.9}$$

Simulating the nonstationary states was accomplished by imposing three true initial condition specifications. The first two are identical to the last two stationary specifications (i.e., initializing with free-parameter and null initial condition matrices, respectively), while the last one allows for the entire process to be fully diffuse. This last simulation process involved simulating a total of $t = 50$ time points *before* the actually observations are collected. Next, state values were standardized within

person across the 50 time points and each person's respective standardized value at the 50th time point to simulate a new process starting at $t = 1$. Specifically, each person's $x1_{i1}$ and $x2_{i1}$ values were assigned to be the that person's respective standardized state value at the 50th time point. This is because, since the process is so explosive over time (see Figure 2.1 panel (B) and (C)), it is likely that the estimation procedures I intend on using will pose computational difficulties when simply using the raw data. Also, this procedure allows the process to have started in the distant past, without imposing an exact distribution for initial condition.

Nature of model misspecification. Table 2.5 displays the different methods used for misspecifying modeling conditions. The misspecification of fitted models arises from simulating the data using a distinct initial condition specification. For the stationary model, data were simulated according to three true initial condition specifications: 1. model-implied moments specification, 2. null initial condition specification, and 3. free parameter specification. All initial condition specifications were fit to each of the simulated data sets. For the nonstationary models, data were simulated according to three true true initial condition specifications: 1. null initial condition, 2. free parameter specification, and 3. diffuse initial condition (procedure described above). With the exception of the model-implied moments approach, all initial condition specifications were fit to each of the simulated data sets.

Table 2.3: Initial Condition Specifications for Model 1: Stationary PFA(1,0) Model

Model-Implied	Free Parameter	Null Initial Condition
$\boldsymbol{\mu}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_0 = \begin{bmatrix} \mu_{x_{1,0}} \\ \mu_{x_{2,0}} \end{bmatrix}$	$\boldsymbol{\mu}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$\mathbf{P}_0 =$	$\mathbf{P}_0 =$	$\mathbf{P}_0 =$
$\begin{bmatrix} (\alpha_{21}\sigma^2_{x2}\alpha^3_{12} - \alpha_1\sigma^2_{x2}\alpha^2_{12}\alpha_2 - \sigma^2_{x2}\alpha^2_{12} + \alpha_{21}\sigma^2_{x1}\alpha_{12}\alpha^2_2 + \alpha_{21}\sigma^2_{x1}\alpha_{12} - \alpha_1\sigma^2_{x1}\alpha^3_2 + \sigma^2_{x1}\alpha^2_2 + \alpha_1\sigma^2_{x1}\alpha_2 - \sigma^2_{x1})/((\alpha_1\alpha_2 - \alpha_2 - \alpha_1 - \alpha_{12}\alpha_{21} + 1)(\alpha_1\alpha_2 - \alpha_{12}\alpha_{21} - 1)(\alpha_1 + \alpha_2 + \alpha_1\alpha_2 - \alpha_{12}\alpha_{21} + 1)) & SYM \\ (\sigma^2_{x2}\alpha^2_1\alpha_{12}\alpha_2 - \sigma^2_{x2}\alpha_1\alpha^2_{12}\alpha_{21} + \sigma^2_{x1}\alpha_1\alpha_{21}\alpha^2_2 - \sigma^2_{x1}\alpha_1\alpha_{21} - \sigma^2_{x1}\alpha_{12}\alpha^2_{21}\alpha_2 - \sigma^2_{x2}\alpha_{12}\alpha_2)/((\alpha_1 + \alpha_2 + \alpha_1\alpha_2 - \alpha_{12}\alpha_{21} + 1)(\alpha_1\alpha_2 - \alpha_2 - \alpha_1 - \alpha_{12}\alpha_{21} + 1)(\alpha_1\alpha_2 - \alpha_{12}\alpha_{21} - 1)) & (\alpha_2\sigma^2_{x2}\alpha^3_1 - \alpha_{12}\sigma^2_{x2}\alpha^2_1\alpha_{21} - \sigma^2_{x2}\alpha^2_{11} + \alpha_2\sigma^2_{x1}\alpha_1\alpha^2_{21} - \alpha_2\sigma^2_{x2}\alpha_1 - \alpha_{12}\sigma^2_{x1}\alpha^3_{21} + \sigma_{x1}\alpha^2_{21} - \alpha_{12}\sigma^2_{x2}\alpha_{21} + \sigma^2_{x2})/((\alpha_1 + \alpha_2 - \alpha_1\alpha_2 + \alpha_{12}\alpha_{21} - 1)(\alpha_1\alpha_2 - \alpha_{12}\alpha_{21} - 1)(\alpha_1 + \alpha_2 + \alpha_1\alpha_2 - \alpha_{12}\alpha_{21} + 1)) \end{bmatrix}$	$\begin{bmatrix} \sigma^2_{x_{1,0}} & \sigma_{x_{2,0},x_{1,0}} \\ \sigma_{x_{1,0},x_{2,0}} & \sigma^2_{x_{2,0}} \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Note PFA=Process factor analysis

Table 2.4: Initial Condition Specifications for Model 2: Nonstationary PFA(1,0)

	DKF	Exact Initial Condition	Large κ Approximation
Model 2: Nonstationary PFA(1,0)	$\boldsymbol{\mu}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\mathbf{P}_{*,0} = 0$ $\mathbf{P}_{\infty,0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\boldsymbol{\mu}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\mathbf{P}_{*,0} = 0$ $\mathbf{P}_{\infty,0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\boldsymbol{\mu}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\mathbf{P}_{*,0} = 0$ $\mathbf{P}_{\infty,0} = \begin{bmatrix} 1000 & 0 \\ 0 & 1000 \end{bmatrix}$

Note PFA=Process factor analysis

Table 2.5: True vs. Fitted Initial Condition Specifications

True Initial Condition Specification	Initial Condition in Fitted Model
1. Null Initial Condition Specification $\mathbf{x}_{i0} \sim N(\mathbf{0}, \mathbf{0})$	Correctly Specified Model:$\mathbf{x}_{i0} \sim N(\mathbf{0}, \mathbf{0})$ Misspecified Models: (a) $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ Model-Implied Moments (b) $\mathbf{x}_{i0} \sim N(\hat{\boldsymbol{\mu}}*, \hat{\boldsymbol{\Sigma}}*)$ Free-Parameter (c) $\mathbf{x}_{i0} \sim N([\mathbf{A}_\infty \boldsymbol{\mu}], \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*')$ De Jong DKF (d) $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*' \text{and} \mathbf{A}_\infty \mathbf{A}_\infty')$ Koopman exact initial KF (e) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \kappa \mathbf{I})$ Large κ approximation
2. Free Parameter Specification $\mathbf{x}_{i0} \sim N\left(\begin{bmatrix}1\\.5\end{bmatrix}, \begin{bmatrix}1.2 & .3\\.3 & .7\end{bmatrix}\right)$	Correctly Specified Model:$\mathbf{x}_{i0} \sim N(\hat{\boldsymbol{\mu}}*, \hat{\boldsymbol{\Sigma}}*)$ Misspecified Models: (a) $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ Model-Implied Moments (b) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \mathbf{0})$ Null Initial Condition (c) $\mathbf{x}_{i0} \sim N([\mathbf{A}_\infty \boldsymbol{\mu}], \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*')$ De Jong DKF (d) $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*' \text{and} \mathbf{A}_\infty \mathbf{A}_\infty')$ Koopman exact initial KF (e) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \kappa \mathbf{I})$ Large κ approximation
3. Model-Implied Moments Specification $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu} = \mathbf{0}$ and $\text{vec}(\boldsymbol{\Sigma}) = (\mathbf{I} - \mathbf{T} \otimes \mathbf{T})^{-1} \text{vec}(\boldsymbol{\Sigma}_v))$	Correctly Specified Model: $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ Misspecified Models: (a) $\mathbf{x}_{i0} \sim N(\hat{\boldsymbol{\mu}}*, \hat{\boldsymbol{\Sigma}}*)$ Free-Parameter (b) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \mathbf{0})$ Null Initial Condition (c) $\mathbf{x}_{i0} \sim N([\mathbf{A}_\infty \boldsymbol{\mu}], \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*')$ De Jong DKF (d) $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*' \text{and} \mathbf{A}_\infty \mathbf{A}_\infty')$ Koopman exact initial KF (e) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \kappa \mathbf{I})$ Large κ approximation
4. Diffuse Initial Condition Distant past procedure: Simulate t=50 data points, standardize state values within person across time points. Standardized values at t=50 used as t state values a first time point.	Correctly Specified Models: $\mathbf{x}_{i0} \sim N([\mathbf{A}_\infty \boldsymbol{\mu}], \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*')$ De Jong DKF $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \mathbf{A}_* \mathbf{P}_* \mathbf{A}_*' \text{and} \mathbf{A}_\infty \mathbf{A}_\infty')$ Koopman exact initial KF Misspecified Models: (a) $\mathbf{x}_{i0} \sim N(\hat{\boldsymbol{\mu}}*, \hat{\boldsymbol{\Sigma}}*)$ Free-Parameter (b) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \mathbf{0})$ Null Initial Condition (c) $\mathbf{x}_{i0} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ Model-Implied Moments (d) $\mathbf{x}_{i0} \sim N(\mathbf{0}, \kappa \mathbf{I})$ Large κ approximation

Note KF=Kalman filter, DKF=Diffuse Kalman filter
Note

2.3 Software Packages

The specifications using the De Jong (1991) and Koopman and Durbin (2000) approaches (specifications 4 and 5, respectively) require a modification of the likelihood function. Access to this likelihood function is not currently available via in some common canned software packages such as Mplus (Muthén & Muthén, 1998–2007). Ox version 5.1 (Doornik, 2007) in combination with SSfpack version 3.0 (Koopman et al., 2008) allows for the specification of both the De Jong (1991) and Koopman and Durbin (2000) approaches. Ox is an object-oriented programming language and the software package SSFpack allows models to be formulated as state-space models and uses an estimation procedure incorporating the KF and PED. Using this package also allows for the estimation of data where the number of time points exceeds the number of observations.

I also programmed the likelihood functions for both the De Jong (1991) and Koopman and Durbin (2000) approaches using the SAS proc IML software package. However, the estimation procedure was unable to find a minimum of the negative log-likelihood in certain conditions, especially when the number of time points was large in comparison to the sample size. Specifically, the optimization procedure was unable to successfully converge and would fluctuate at a given log-likelihood value indefinitely.

2.4 Summary of Simulation Design

The simulation design contained four conditions that were varied: 1. model type, 2. true initial condition specification, 3. fitted initial condition specification, and 3. type of data. Two models were considered: 1. PFA stationary model and 2. PFA nonstationary model. For the nonstationary model, two degrees of nonstationarity were considered: 1. moderate nonstationarity and 2. mild nonstationarity. For the stationary model, three true initial condition specifications were used: 1. Model-implied

moments initial condition, 2. Free parameter initial condition, and 3. Null initial condition. For the nonstationary models, three true initial condition specifications were used: 1. Free parameter initial condition, 2. null initial condition, 3. diffuse initial condition. The fitted initial condition specification will have six conditions for the stationary model (i.e., all discussed conditions) and five conditions for the non-stationary models (i.e., all but the model implied condition). Finally, two types of data will be considered: 1. intensive repeated measures data ($T = 50, N = 20$) and 2. panel data ($T = 5, N = 200$). The moderately nonstationary model was only estimated with panel data. Thus, there will be a total of 66 cells in the simulation design. For each cell I completed a total of 500 Monte Carlo replications. Table 2.1 summarizes the simulation design.

2.5 Outcome Measures

To assess the consequences of the different initial condition specifications, I determined how well the true latent variable scores were recovered, how often the models converged, how well the parameter estimates were recovered, and computational time. I also compared fit indices across approaches and indicated how many replications produced strange or outlying results.

To compare the latent variable scores obtained from the model fitting procedure to the true scores obtained from the simulation specifications I will calculate the root mean square error (RMSE) between these two values,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{itLVS} - x_{itTRUE})^2}{NT}} \tag{2.10}$$

where i represents a given individual, t represents a given time point, x_{itLVS} represent the latent variable score estimates, and x_{itTRUE} represent the true values. Smaller values indicate a smaller discrepancy between the latent variable score estimates

and true scores. To assess the point estimates of the parameters I will calculate the relative bias, RMSE, confidence interval coverage, and power for detecting non-zero parameters. To assess the standard errors of the parameter estimates I calculated the estimated standard errors, i.e., the mean of the estimated standard errors ($\hat{SE}$) and compared them to the empirical or "true" standard errors, i.e., the standard deviation of the parameter estimates (SE_θ). To more directly evaluate the consistency of the standard error estimation, I compared the estimated standard errors against the empirical standard errors by computing the difference $SEDIFF = SE_\theta - \hat{SE}$. Lower values indicated better consistency of standard error estimation. Also, the percentage of iterations where models converged was assessed.

Relative bias was calculated as:

$$\text{Relative Bias } = \frac{\frac{\sum_{r=1}^{R} \hat{\theta}_r}{R} - \theta}{\theta} \tag{2.11}$$

where r represents a distinct Monte Carlo run, R is the total number of replications, and θ is the true parameter value. This index provides information concerning the difference between the average estimate of the parameter and the population parameter, with smaller values indicating greater parameter accuracy. RMSE was calculated as:

$$\text{RMSE } = \sqrt{\frac{\sum_{r=1}^{R} (\hat{\theta}_r - \theta)^2}{R}} \tag{2.12}$$

where smaller values indicate greater parameter precision. Confidence interval coverage was calculated as the percentage of replications in which the population parameter falls between the lower confidence limit and the upper confidence limit with an alpha of 0.05. Power was calculated as the percentage of replications in which the null hypothesis of $\theta = 0$ is rejected. I also assessed the confidence interval coverage rates which were calculated as the percentage of replications in which the population

parameter falls between the lower confidence limit and the upper confidence limit with an alpha of 0.05

I compared values of information fit indices including the Bayesian Information Criterion (BIC; Raftery, 1995), which is defined as

$$BIC = -2LL + (nparm)\ln(N) \tag{2.13}$$

, and the Akaike information criterion (AIC Akaike, 1974), defined as

$$AIC = -2LL + 2(nparm) \tag{2.14}$$

where $nparm$ is the number of parameters estimated in the model. The form of these equations are identical to those used in the software package Mplus (Muthén & Muthén, 1998–2004).

2.6 Hypotheses

Based on past research and current knowledge, I made the following hypotheses for the stationary modeling condition:

- When the model-implied moments true specification is implemented, the fitted model-implied moments condition will perform the best with respect to simulation outcomes. The free-parameter condition will perform almost as well, but may not be chosen by the AIC or BIC as the degrees of freedom is greater.

- When the free-parameter true specification is implemented, the fitted free-parameter condition will perform the best with respect to simulation outcomes.

- When the null true specification is implemented, the fitted null initial condition will perform the best with respect to simulation outcomes. The free-parameter condition will perform almost as well, but may not be chosen by the AIC or BIC

as the degrees of freedom is greater.

- For all of the true initial condition specifications, the fitted diffuse methods (de Jong DKF, Koopman exact initial KF, and large κ approaches) will perform the worst although similar to each other, with the de Jong DKF and Koopman exact initial KF slightly outperforming the large κ condition.

I also made the following hypotheses for the nonstationary modeling condition:

- When the free-parameter true specification is implemented with mild nonstationarity, the fitted free-parameter condition will perform the best with respect to simulation outcomes, followed by the de Jong and Koopman approaches, large κ approach, and null approach.

- When the null true specification is implemented with mild nonstationarity, the fitted null initial condition will perform the best with respect to simulation outcomes. The free-parameter condition will perform almost as well, but may not be chosen by the AIC or BIC as the degrees of freedom is greater. The de Jong and Koopman approaches will perform similarly, but not as well as the null or free-parameter approach, while the large κ approach will perform the worst.

- When the true initial condition is diffuse (mildly or moderately stationary), Koopman's approach will be comparable to de Jong's approach and both perform the best, followed by the free-parameter condition, large κ approximation, and null initial condition with respect to all simulation outcomes.

- When the free-parameter true specification is implemented with moderate nonstationarity, the fitted free-parameter condition will perform the best with respect to simulation outcomes, followed by the de Jong and Koopman approaches, large κ approach, and null approach. However, the degree to which

the free-parameter approach outperforms both the de Jong and Koopman approaches will be less than in the mild nonstationarity case.

- When the null true specification is implemented with moderate nonstationarity, the fitted null initial condition will perform the best with respect to simulation outcomes. The free-parameter condition will perform almost as well, but may not be chosen by the AIC or BIC as the degrees of freedom is greater. The de Jong and Koopman approaches will perform similarly, but not as well as the null or free-parameter approach, while the large κ approach will perform the worst. However, the degree to which the free-parameter approach outperforms both the de Jong and Koopman approaches will be less than in the mild nonstationarity case.

-

In addition, I made the following general hypotheses:

- The de Jong approach and Koopman approach will produce almost identical results.

- Both the de Jong and Koopman approaches compared to the free-parameter, null, and large κ conditions will have a greater impact on results for the panel data condition versus the intensive repeated measures condition. That is, the free-parameter, null, and large κ approaches will produce poorer results for the panel data condition when compared to their performance in the intensive repeated measures condition, and the improvement in results when using either the de Jong or Koopman approaches will be more noticeable and substantial for the panel data condition.

CHAPTER 3

Results

Given the large amount of results, this section is broken down by data generating model. Tables are presented followed by figures that capture some significant findings. For each data generating model (i.e., stationary PFA, mildly nonstationary PFA, and moderately stationary PFA) I will describe results associated with model convergence, model fit and comparison, and simulation outcomes. To formally quantify which results should be considered significant findings, I applied a series of meta-models as suggested by Skrondal (2000). Specifically, I ran a series of Analysis of Variance (ANOVA) models for each true initial condition specification with simulation outcome measures serving as dependent variables. The factors were data type (two levels: the $N = 20/T = 50$ intensive repeated measures condition and the $N = 200/T = 5$ panel design condition) and fitted initial condition specification (five levels for stationary model, six levels for nonstationary model). The models were expressed as

$$\text{ANOVA Model:SimOutcome}_i = \text{DataType}_i + \text{FittedIC}_i + \text{DataType}_i * \text{FittedIC}_i + \epsilon_i \tag{3.1}$$

where SimOutcome is any given simulation outcome for a given true initial condition specification, DataType is a factor for data type, FittedIC is a factor for fitted initial condition specification. As there was a large number of parameters per estimated

model, I aggregated results into the following categories: 1. measurement model parameters, which consisted of averaging factor loading parameters and measurement error variance parameters, i.e., $Z21$, $Z31$, $Z52$, $Z62$, $U11$, $U22$, $U33$, $U44$, $U55$, and $U66$, 2. time series parameters, which consisted of averaging across parameters in the transition matrix, i.e., $T11$, $T21$, $T12$, and $T22$, and 3. process noise parameters, which consisted of average across parameters in the process noise covariance matrix, i.e., $V11$, $V21$, and $V22$.

Given the large number of replications, most effects were significant. Due to this, ANOVA main effect and interaction results that have an R^2 value of at least .1 were considered to have relatively strong effects. Thus, effects that had an $R^2 \geq .1$ were emphasized. While this cut-off value is arbitrary it coincides with Cohen's (1992) popular effect size guidelines.

3.1 Model Convergence Issues

Given the complicated nature of the models and initial condition specifications, several computational problems emerged. The percentage of times a model converged to a proper solution is important when considering which initial specification to choose as a specification that is too computationally burdensome, despite being accurate when it does converge, may not be optimal for noisy empirical data sets. Time to convergence is also an issue to be considered, but given the relatively short time spans to converge across all simulation conditions this may not be a deciding factor when choosing an initial condition specification.

Tables 3.1, 3.2, and 3.3 display model performance results with respect to convergence, model selection, latent variable score recovery, and average model estimation time. The first column of results is of particular interest as it reports how many replications were retained after removing cases that did not converge to a proper solution (i.e., those that reported both weak convergence and no convergence) and cases that contained extreme outliers. Outliers were detected via visual inspection and omitted

if they drastically deviated from the both the true population value and the majority of estimates. Figure 3.1 contains plots illustrating typical outliers that were removed. This plot displays parameter estimates of $Z52$, i.e., a factor loading for the second set of manifest variables, from fitting the free-parameter initial condition specification to moderately nonstationary data with a true diffuse initial condition specification. The true parameter value is .9, and clearly there are two outlying cases that have a magnitude greater than absolute 30. Thus, these two cases were removed.

The next three columns of Tables 3.1, 3.2, and 3.3 report the number of cases that produced strong convergence, weak convergence, and no convergence, respectively. Weak and strong convergence are defined as iterations with error tolerance values that are less than $eps_1 = 1e^{-4}$ and $eps_2 = 5e^{-3}$, respectively (Doornik, 2007). Only cases with strong convergence were considered. The fifth column of results reports the number of outliers followed by a column that reports the number of replications did converged but failed to produce standard error values. The next two columns report the number of replications where the AIC and BIC, respectively, would choose a given fitted initial condition specification. The next column reports the RMSE of the latent variable scores and the final column reports the average time it took to estimate the models, whether the models converged or not.

When inverting the Hessian matrix (i.e., the matrix of second derivatives) to obtain standard error estimates, there were some replications where a generalized symmetric inverse using singular value decomposition was implemented as the regular inversion procedure failed. Upon close inspection of the point estimates and standard error estimates produced, there were certain times when the variances of either process noise variable were very close to zero, which indicates that the optimizer hit a boundary condition, despite these cases strongly converging. Figure 3.2 displays a typical plot illustrating some replications that hit a boundary condition. This plot illustrates the process noise variance for the second state when the free-parameter

Table 3.1: Rates of Convergence to a Proper Solution, AIC and BIC Model Selection, Latent Variable Recovery, and Average Total Time Obtained from Stationary PFA Model.

N/T	True IC	Fitted IC	Final # Reps	Strong Conv	Weak Conv	No Conv	Outliers	no SEs	AIC	BIC	$\overline{RMSE}$lvs[a]	$\overline{Time}$[b]
N=200 T=5	Model Implied	Model Implied	36	386	0	45	353	3	0	0	0.460	14.318
		Free Parameter	454	480	9	11	25	0	0	0	0.430	11.599
		Null	424	484	1	15	60	0	0	0	0.690	7.273
		de Jong DKF	499	500	0	0	0	0	278	278	0.469	16.633
		Koopman exact initial KF	459	494	1	2	34	0	222	222	0.475	7.803
		Large κ	463	496	1	3	32	0	0	0	0.475	7.426
	Free Parameter	Model Implied	23	370	1	63	353	5	0	0	0.452	19.809
		Free Parameter	451	476	9	15	25	0	0	0	0.426	14.774
		Null	445	479	0	21	34	0	0	0	0.703	9.221
		de Jong DKF	500	500	0	0	0	0	278	278	0.464	21.702
		Koopman exact initial KF	447	488	0	8	41	0	222	222	0.469	10.334
		Large κ	450	492	0	8	42	0	0	0	0.469	9.941
	Null	Model Implied	51	365	0	80	322	6	0	0	0.42	15.856
		Free Parameter	412	427	35	38	12	0	1	1	0.61	19.041
		Null	454	488	0	12	34	0	0	0	0.381	8.13
		de Jong DKF	499	499	0	1	0	1	269	269	0.494	19.651
		Koopman exact initial KF	457	481	1	16	24	0	229	229	0.51	8.868
		Large κ	461	484	0	16	23	0	0	0	0.51	8.474
N=20 T=50	Model Implied	Model Implied	405	478	0	13	73	0	32	32	0.426	4.204
		Free Parameter	473	492	3	5	25	0	16	16	0.425	6.433
		Null	476	497	2	1	21	0	3	3	0.458	3.652
		de Jong DKF	499	500	0	0	1	0	229	229	0.427	6.947
		Koopman exact initial KF	474	489	6	5	20	0	211	211	0.426	3.681
		Large κ	474	490	4	6	20	0	0	0	0.426	3.667
	Free Parameter	Model Implied	404	482	0	11	79	0	45	45	0.424	4.358
		Free Parameter	477	494	0	6	19	0	71	71	0.424	6.497
		Null	478	494	3	3	17	0	1	1	0.459	3.749
		de Jong DKF	497	500	0	0	3	0	179	179	0.425	7.347
		Koopman exact initial KF	477	498	1	1	21	0	203	203	0.429	3.849
		Large κ	477	496	1	3	19	0	1	1	0.429	3.773
	Null	Model Implied	418	471	0	11	53	0	47	47	0.426	6.582
		Free Parameter	464	490	5	5	19	0	90	90	0.423	13.686
		Null	477	494	3	3	16	0	80	80	0.421	5.602
		de Jong DKF	497	499	0	1	1	0	130	130	0.428	10.728
		Koopman exact initial KF	476	496	2	2	20	0	150	150	0.443	5.665
		Large κ	477	497	3	0	19	0	1	1	0.445	5.569

[a] $\overline{RMSE}$lvs represents the average RMSE value (see Equation 2.10) across replications for true latent variable scores versus estimated latent variable scores. Smaller values indicate a greater correspondence between the true and estimated scores.

[b] $\overline{Time}$ represents the average time across replications taken to estimate the model, whether convergence was reached or not.

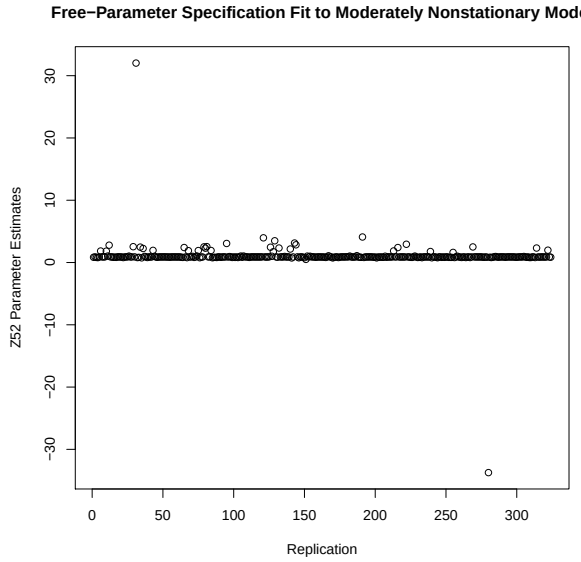

Figure 3.1: Sample of typical outlier cases removed. The free-parameter model is fitted to the moderately nonstationary model using a diffuse true initial condition specification. The x-axis displays parameter estimates of $Z52$, and the values displaying an absolute value greater than 30 were removed.

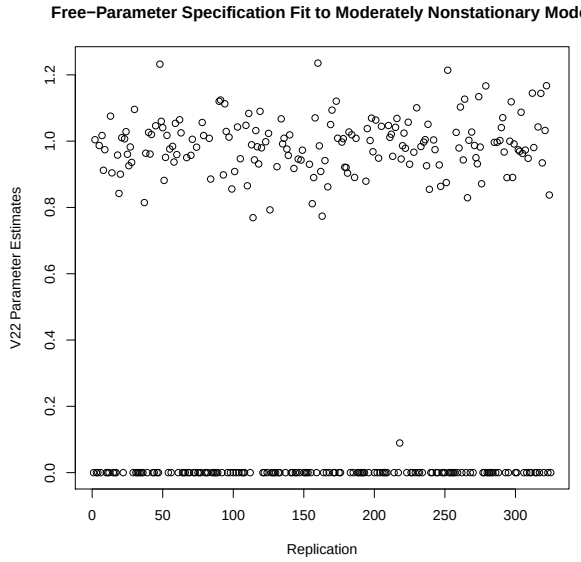

Figure 3.2: Sample of typical cases hitting a boundary condition. The free-parameter model is fitted to the moderately nonstationary model using a diffuse true initial condition specification. The x-axis displays parameter estimates of $V22$, and the values that are very close to zero were removed for the second set of results.

approach is fitted to moderately nonstationary simulated data using a true diffuse initial condition specification. The replications that hit a boundary condition are apparent as they are all very closet to zero, while cases that did not cluster closer closer to the true value of 1. Given the moderate frequency of such cases, results are also presented without these boundary cases.

Tables 3.4, 3.5, and 3.6 display model performance results for the stationary model, mildly nonstationary model, and moderately nonstationary model, respectively, when the boundary replications are removed. Most boundary condition occurred when the free-parameter approach was fitted to any true initial condition specification, particularly in the nonstationary model. The fitted Koopman exact initial KF and large κ approaches also displayed a relatively large number of removed boundary cases. Interestingly, when the boundary cases were removed, the three diffuse approaches, i.e., de Jong DKF, Koopman exact initial KF, and large κ approaches, were more similar to each other.

The remainder of the results section is structured as follows. First, results where the boundary cases were not removed are discussed. The stationary model is discussed first, followed by the mildly nonstationary model, ending with the moderately nonstationary model. Next, results are displayed with the removal of boundary cases. Again, the stationary model is discussed first, followed by the mildly nonstationary model, ending with the moderately nonstationary model. General simulation conclusions are next presented, followed by the results from the empirical example.

3.2 Stationary PFA

General model performance for stationary PFA. As illustrated in Table 3.1, for the stationary PFA, the number of replications retained was lowest in the panel data condition ($T = 5$ and $N = 200$) when a diffuse initial condition was specified. Surprisingly, the number of retained cases was very low when fitting the model implied initial condition specification with panel data to any true specification. This may be due to the

rather complicated constraints that are placed in the initial condition covariance matrix, as shown in Table 2.3. Also, there were not any stationarity constraints imposed on the time series parameters, which may have led the model-implied method to produce highly outlying estimated AR values. In fact, most cases were omitted due to extreme outliers in the parameter results rather than issues of non-convergence. Other than this, for the stationary model, the number of retained cases was relatively large across true initial condition specification conditions. When using the null initial condition specification as the true one, however, there were somewhat lower rates of retained cases across fitted initial conditions, particularly in the panel data condition. Fitting the de Jong DKF initialization approach produced the largest retained cases across the different true initial condition specifications. Fitting the Koopman exact initial KF approach, however, did not produce as many retained replications as the de Jong DKF approach. It is important to note that, under ideal modeling circumstances, the results produced by these two models would be identical (Koopman et al., 2008). This was corroborated as the results generated from both conditions when there was both strong convergence and no extreme outlying cases were found to be identical. Thus, any differences in results are a function of model convergence and outlier issues.

Somewhat surprisingly, even when fitting the de Jong DKF and Koopman exact initial KF approaches did not produce optimal results with respect to parameter estimates, the AIC and BIC tended to choose these two approaches. In general, fitting the free-parameter approach produced the smallest RMSE of latent variable scores, indicating that this approach produced estimated latent variable scores that best approximated the true latent variable scores. However, the superior performance was marginal and all fitted initial condition specifications across all conditions produced adequate RMSE values. On average, the estimation of the models did not take a very long time, with the panel data condition taking slightly longer than the intensive

repeated measures condition and fitting the de Jong initial condition specification taking slightly longer than the other fitted initial condition specification.

Absolute biases, absolute coverage rates minus .95, and the absolute discrepancy between empirical and estimated standard errors (i.e., $SEDIFF$) are illustrated in Figure 3.3, Figure 3.4, and Figure 3.5 when the model-implied moments approach, free parameter approach, and null initial condition approaches were used as the true data generating specifications, respectively. These results are aggregated across measurement parameters, process noise parameters, and time series parameters, as described above, which necessitated taking the absolute value of outcome measures. Results for individual parameters may be found in Appendix A, Tables 5.1 – 5.36.

Model implied moments approach as true initial condition specification. In general, the pattern of simulation outcome results across fitted initial condition specifications were similar, as illustrated in Figure 3.3. Significant differences in fitted initial condition specifications were found, for the measurement parameters, in absolute bias ($R^2 = .14$, see Figure 3.3A) and $SEDIFF$ ($R^2 = .142$, see Figure 3.3C), for the process noise parameters, absolute bias ($R^2 = .12$, see Figure 3.3D) and $SEDIFF$ ($R^2 = .14$, see Figure 3.3F), and for the times series parameters, $SEDIFF$ ($R^2 = .13$, see Figure 3.3I) and confidence interval width ($R^2 = .12$). In terms of absolute bias, the fitted free-parameter approach outperforms the others for all parameter groupings. However, the null initial condition outperformed the others with respect to $SEDIFF$, meaning that the empirical standard errors were close to the estimated standard errors. Upon further inspection of the estimates, a few cases in the all fitted conditions were found that, while not extreme outliers, still deviated from the cluster of the majority of values, which would explain why the free-parameter condition displayed better absolute bias but worse $SEDIFF$ than the null condition.

The main effect of data type produced values of R^2 high in magnitude for several outcome measures including, for the measurement parameters, $SEDIFF$ ($R^2 = .12$,

see Figure 3.3C), for the process noise parameters, absolute bias ($R^2 = .10$, see Figure 3.3D), relative bias ($R^2 = .14$), coverage rates ($R^2 = .15$, see Figure 3.3E), and $SEDIFF$ ($R^2 = .13$, see Figure 3.3F), and for the autoregressive parameters, confidence interval width ($R^2 = .11$). These effects all indicate that, for this particular modeling condition, the panel data condition did not perform as well as the intensive repeated measures condition across all fitted initial conditions. This result makes sense because, as the number of time points increases, the degree to which the initial condition specification matters diminishes. When there are more time points, there are more data for the models to estimate the trajectory.

Several interactions between fitted initial condition and data type were also apparent. For instance, as illustrated in Figure 3.3G, there is an interaction effect ($R^2 = .12$) between fitted initial condition and data type for the time series parameters such that fitting the model implied method results in poor absolute bias only for the panel data condition. This same type of interaction is found for coverage rates ($R^2 = .27$), see Figure 3.3H), and $SEDIFF$ ($R^2 = .61$, see Figure 3.3I). The implications of this effect are simply that all fitted initial conditions performed well with respect to recovering the point and standard error estimates of the time series parameters, with the exception of the model-implied fitted initial condition. However, as stated, this may be due to the complicated constraints involved in the model-implied condition and may be alleviated if stationarity constraints are imposed on the AR and cross-regression parameters.

Free parameter approach as true initial condition specification. When using the free parameter approach as the true initial condition specification, the findings were similar to true model-implied condition. However, for the process noise parameters, R^2 values for the main effect of initial condition specification were larger in magnitude for absolute bias ($R^2 = .17$, see Figure 3.4D), absolute relative bias ($R^2 = .12$), and $SEDIFF$ ($R^2 = .19$, see Figure 3.4F). The fitted free parameter specification performs

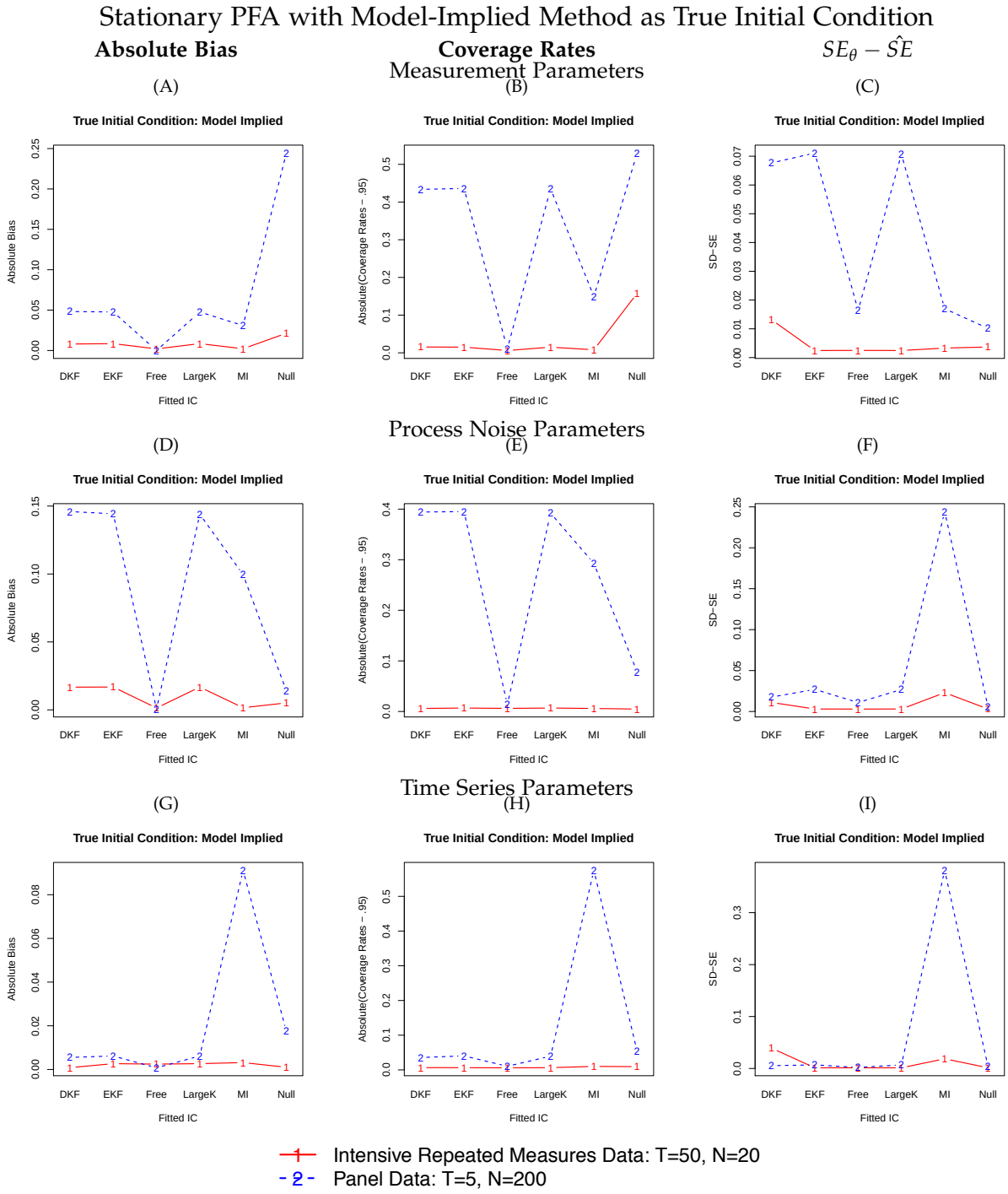

Figure 3.3: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the stationary PFA model simulated using the model-implied moments method as the true initial condition specification. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

the best, followed by the fitted null initial condition, while, in terms of biases, the fitted de Jong DKF, Koopman exact initial KF, and large κ conditions perform worst, although similarly to each other. In terms of standard error estimation, however, the model-implied method performed the worst. The larger effect size of fitted initial condition for the true free-parameter condition versus the true model-implied condition suggests that the performance of fitted initial conditions was more discrepant with respect to the process noise variables in this case.

There were several large R^2 values for the main effect of data type, including, for the process noise parameters, absolute bias ($R^2 = .11$, see Figure 3.4D) and absolute relative bias ($R^2 = .15$), and for the time series parameters, coverage rates ($R^2 = .15$, see Figure 3.4H), and confidence interval width ($R^2 = .19$). Again, the intensive repeated measures data consistently outperformed the panel data condition. An interaction effect between data type and fitted initial condition was found, for the time series parameters, for coverage rates ($R^2 = .17$, see Figure 3.4H). This interaction shows that the free parameter fitted initial condition performs best for both data types, while the model-implied initial condition performs very poorly for the panel data condition, much like the interaction effects found when the true initial condition specification was the model-implied moments specification.

Null approach as true initial condition specification. Results when using the null initial condition as the true initial condition specification revealed a similar pattern to the two previously discussed true initial specification conditions, as illustrated in Figure 3.5. One big difference is that the fitted null initial condition performed best in most cases, which makes sense given that the true initial condition specification was null. However, keep in mind that the proper convergence rates for this condition were smaller on average than the other true initial specification conditions.

The main effect of fitted initial condition specification was strong, for the process noise parameters, absolute bias ($R^2 = .2$, see Figure 3.5D), absolute relative bias

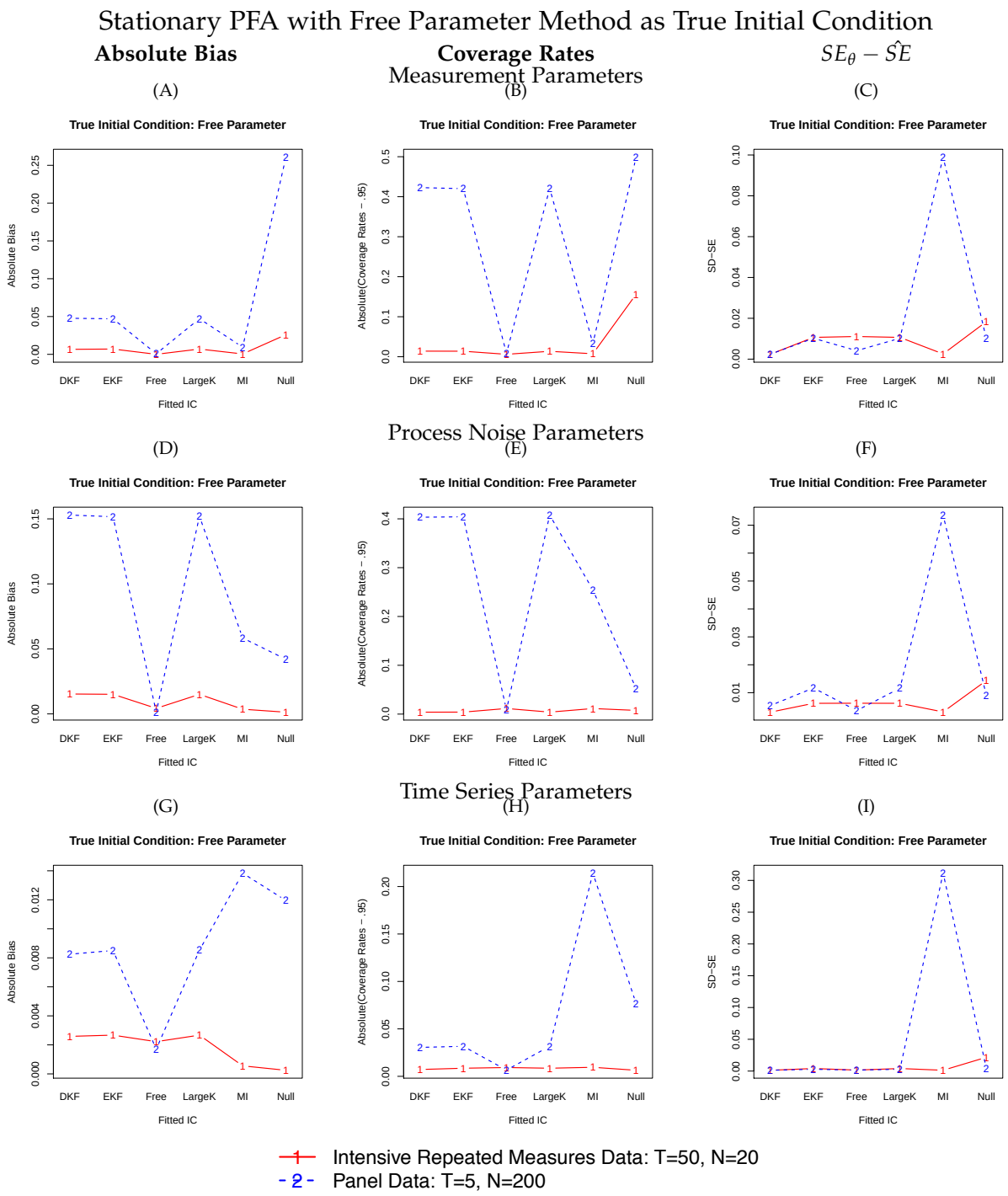

Figure 3.4: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the stationary PFA model simulated using the free parameter moments method as the true initial condition specification.

($R^2 = .15$), $SEDIFF$ ($R^2 = .28$, see Figure 3.5F), and for the time series parameters, $SEDIFF$ ($R^2 = .17$, see Figure 3.5I). While the fitted null initial condition performed well overall, the free parameter condition still also performed well. An unexpected finding was that the discrepancy between empirical and estimated standard errors was large for the null initial condition for the measurement and time series parameters in the intensive repeated measures condition. This pattern was also true for the de Jong DKF. Upon further inspection of estimates, for both fitted approaches, there were a few replications that produced values that were somewhat outside of the range of the majority of values. More specifically, these values were not identified as outlying values, but still somewhat deviated from the main cluster of points. When removed, the standard error estimates were a lot closer to the empirical standard errors. Also, despite the underestimation of the true standard errors of the parameters, the magnitudes of the bias were still small.

There were again several large R^2 values for the main effect of data type, including, for the process noise parameters, absolute bias ($R^2 = .17$, see Figure 3.5D) and absolute relative bias ($R^2 = .26$), MSE ($R^2 = .11$), and confidence interval width ($R^2 = .16$), and for the time series parameters, and confidence interval width ($R^2 = .60$). Several interaction effects were found for $SEDIFF$, and these simply illustrate the discrepant findings with respect to the de Jong and null approaches between the panel data case and intensive repeated measures data case. Other than that, the same pattern of the intensive repeated measure data condition outperforming the other true initial condition specification conditions was found. For the process noise parameters, an interaction effect of absolute bias was found ($R^2 = .15$) which indicates that the degree to which the biases deviate between data types only for the diffuse initial condition approaches is quite large. The other approaches, however, maintain adequate bias across data type conditions.

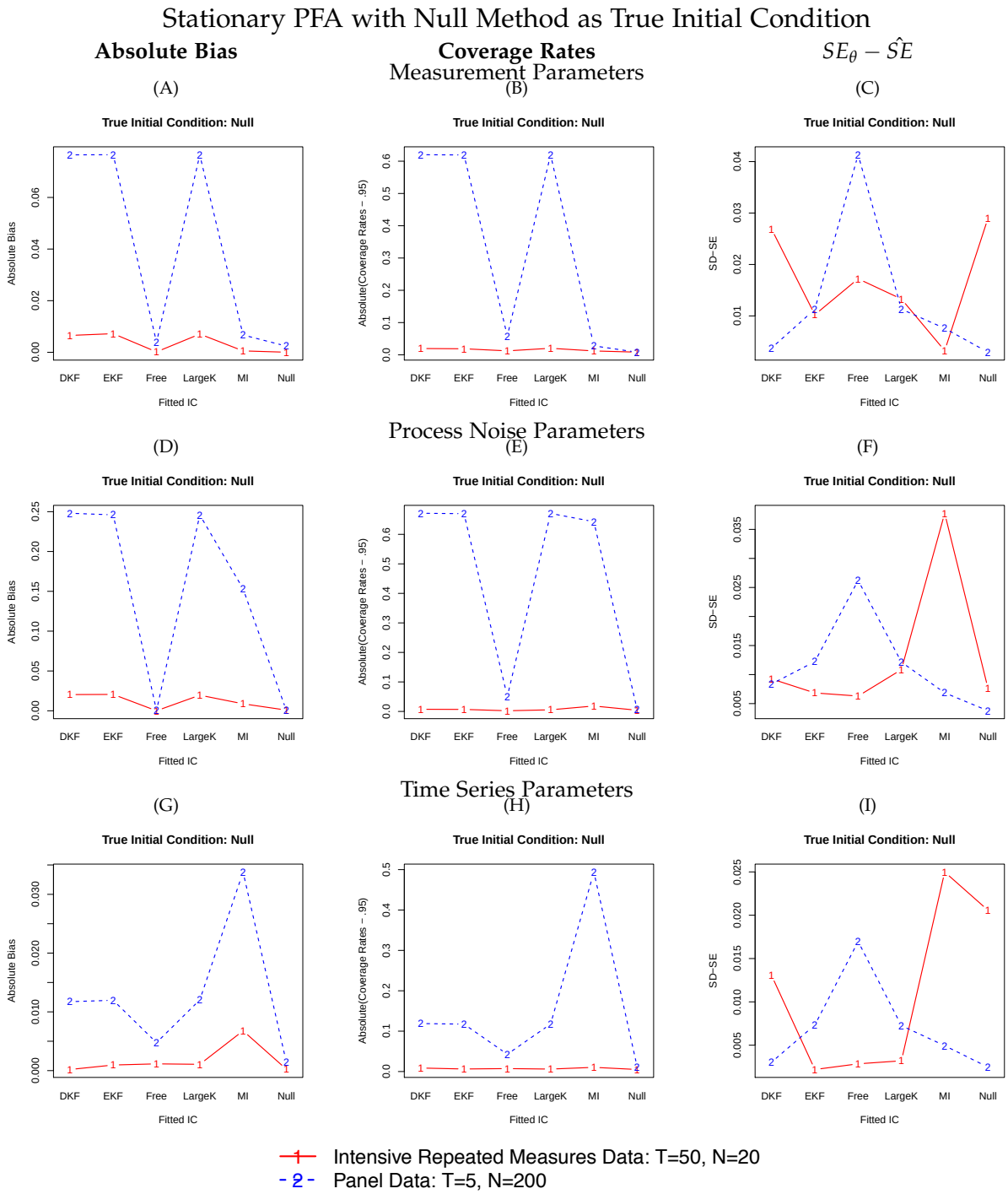

Figure 3.5: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the stationary PFA model simulated using the null method as the true initial condition specification. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

3.3 Summary of Stationary Results

Some general trends were observed across the different true initial condition specifications for the stationary model. First, when the true process is stationary, results suggest that the diffuse initial condition specifications should not be used to fit the model. Second, the fitted free-parameter approach appeared to have good versatility across the differing true initial condition specifications. Third, the process noise parameters were most affected when differing fitted initial conditions were implemented. Fourth, in the panel data condition, the AIC and BIC consistently chose either the de Jong or Koopman approaches. In the intensive repeated measures condition, while the de Jong and Koopman approaches were still generally chosen by the AIC and BIC, the other conditions were also chosen some of the time. Disregarding the de Jong and Koopman approaches, the AIC and BIC correctly picked out the fitted initial conditions that corresponded to the true initial condition specification a large percentage of the time, with the exception of the true null condition where the free-parameter condition was chosen most of the time. However, the null initial condition was the second most chosen condition. Given that researchers generally do not use a diffuse initial condition specification when they believe the hypothesized model has parameters that are in the stationary range, AIC and BIC can still legitimately be used to help select the optimal initial condition specification for a set of data. Fifth, when the true initial condition specifications was free-parameter the most optimal results were produced from the fitted free-parameter condition and when the true initial condition was null the most optimal results were produced from the fitted null condition. However, this was not the case with the model-implied condition, which may work better if appropriate stationary constraints placed on the parameters. Finally, the estimated standard errors tended to underestimate the true standard errors.

3.4 Nonstationary PFA

General model performance for nonstationary PFA. The general modeling performance results for the mildly nonstationary condition are reported in Table 3.2 while the results for the moderately nonstationary condition are reported in Table 3.3. Overall, fitting the de Jong DKF produced the largest number of retained cases, with the exception when the null initial condition was used as the true initial condition specification in the mildly nonstationary case for panel data. This is an interesting result as fitting the Koopman exact initial KF approach resulted in a much higher replication retention rate, which is counter to previous results. Also interesting is that many of the de Jong results produced no standard errors while still reporting strong convergence. Upon closer inspection, the parameter estimates for the measurement error variances of $U11$ and $U44$ yielded in such cases were very close to zero. One reason why the de Jong approach failed in this situation may be due inverting issues concerning the Hessian matrix. Specifically, at the first time point, there is no process noise variance added to the DKF equations, as the initial covariance matrix is a null matrix. At the second time point, there is only a small amount of process noise added (as the process noise variances in this case are small at .05), and since previous true states were zero, no other information from the first time point is being carried through. That, coupled with the fact that the panel data case does not have a lot of subsequent data, gave rise to a Hessian matrix with values close to zero, and thus inversion problems.

As in the stationary modeling condition, the AIC and BIC frequently chose either the de Jong DKF or Koopman exact initial KF initialization approaches over the other fitted approaches. The de Jong and free parameter fitted initial conditions tended to take the longest to be estimated and also tended to produce latent variable score estimates that best approximated the true latent variable scores. One notable finding is that, in the panel data case, when fitting the de Jong DKF to data generated from

Table 3.2: Rates of Convergence to a Proper Solution, AIC and BIC Model Selection, Latent Variable Recovery, and Average Total Time Obtained from Nonstationary PFA Model Mild Nonstationarity.

N/T	True IC	Fitted IC	Final # Reps	Strong Conv	Weak Conv	No Conv	Outliers	no SEs	AIC	BIC	$\overline{RMSE}$lvs[a]	$\overline{Time}$[b]
N=200 T=5	Free Parameter	Free Parameter	419	441	17	42	29	0	8	8	0.453	17.001
		Null	417	459	3	38	42	0	0	0	0.821	9.848
		de Jong DKF	462	484	0	16	22	15	264	264	0.298	23.11
		Koopman exact initial KF	429	488	0	6	59	1	219	219	0.479	11.692
		Large κ	432	493	1	6	61	0	1	1	0.476	11.272
	Null	Free Parameter	288	330	45	125	40	0	23	23	1.046	46.468
		Null	303	320	18	162	18	1	45	45	1.878	21.245
		de Jong DKF	79	256	142	102	3	390	47	47	0.752	34.437
		Koopman exact initial KF	307	350	20	114	43	15	230	230	1.238	23.354
		Large κ	315	352	12	136	37	1	5	5	1.263	22.184
	Diffuse	Free Parameter	360	386	30	84	20	0	3	3	0.334	17.618
		Null	391	443	0	57	52	0	0	0	0.966	8.353
		de Jong DKF	435	465	0	35	33	23	264	264	0.286	21.989
		Koopman exact initial KF	386	447	5	44	61	2	195	195	0.442	11.374
		Large κ	395	450	2	48	56	0	1	1	0.458	10.967
N=20 T=50	Free Parameter	Free Parameter	250	296	76	128	69	3	32	32	0.342	17.315
		Null	249	258	64	178	44	0	0	0	0.87	9.103
		de Jong DKF	387	424	17	59	1	104	239	239	0.231	20.133
		Koopman exact initial KF	285	340	74	76	79	5	116	116	0.842	12.891
		Large κ	279	342	87	71	92	3	1	1	0.466	13.05
	Null	Free Parameter	314	357	42	101	35	1	67	67	0.247	14.532
		Null	361	374	23	103	26	1	101	101	0.237	6.908
		de Jong DKF	425	463	3	34	2	70	140	140	0.238	13.021
		Koopman exact initial KF	379	424	20	48	59	6	119	119	0.479	7.989
		Large κ	378	429	15	56	62	2	0	0	0.626	7.773
	Diffuse	Free Parameter	190	236	70	194	87	0	13	13	0.561	15.049
		Null	192	211	53	236	42	0	0	0	0.391	7.735
		de Jong DKF	335	365	26	109	1	151	220	220	0.231	17.53
		Koopman exact initial KF	224	295	88	107	100	9	100	100	0.988	12.206
		Large κ	226	311	96	93	115	4	2	2	1.084	11.638

[a] $\overline{RMSE}$lvs represents the average RMSE value (see Equation 2.10) across replications for true latent variable scores versus estimated latent variable scores. Smaller values indicate a greater correspondence between the true and estimated scores.

[b] $\overline{Time}$ represents the average time across replications taken to estimate the model, whether convergence was reached or not.

Table 3.3: Rates of Convergence to a Proper Solution, AIC and BIC Model Selection, Latent Variable Recovery, and Average Total Time Obtained from Nonstationary PFA Model with Moderate Nonstationarity.

N/T	True IC	Fitted IC	Final # Reps	Strong Conv	Weak Conv	No Conv	Outliers	no SEs	AIC	BIC	$\overline{RMSE}$lvs[a]	$\overline{Time}$[b]
N=200 T=5	Free Parameter	Free Parameter	383	436	13	51	50	0	5	5	0.417	14.662
		Null	385	467	0	33	82	0	0	0	0.697	9.204
		de Jong DKF	467	491	0	9	24	4	298	298	0.419	18.732
		Koopman exact initial KF	370	471	1	26	102	0	187	187	0.467	11.651
		Large κ	372	470	1	29	98	0	0	0	0.469	10.463
	Null	Free Parameter	416	432	29	39	16	0	3	3	0.384	20.08
		Null	440	476	0	24	36	0	2	2	0.365	8.749
		de Jong DKF	487	497	0	3	11	1	261	261	0.428	19.915
		Koopman exact initial KF	451	488	0	9	37	0	230	230	0.45	9.417
		Large κ	452	491	0	9	39	0	0	0	0.451	9.099
	Diffuse large	Free Parameter	213	334	17	149	121	0	10	10	0.642	21.042
		Null	344	426	2	72	426	0	5	5	1.745	9.568
		de Jong DKF	416	436	0	64	22	31	337	337	0.415	23.292
		Koopman exact initial KF	197	353	9	123	160	2	83	83	0.791	15.295
		Large κ	194	367	11	122	173	0	0	0	0.98	15.354

[a] $\overline{RMSE}$lvs represents the average RMSE value (see Equation 2.10) across replications for true latent variable scores versus estimated latent variable scores. Smaller values indicate a greater correspondence between the true and estimated scores.

[b] $\overline{Time}$ represents the average time across replications taken to estimate the model, whether convergence was reached or not.

the moderately nonstationary process, there was a much larger percentage of retained replications than when fitting the other initial condition specifications. This was still true when the process was mildly nonstationary, but to a lesser extent.

For the moderately nonstationary condition, biases, coverage rates, discrepancies between empirical and estimated standard errors are illustrated in Figure 3.9, Figure 3.10, and Figure 3.11 when the free parameter approach, null initial condition approach, and diffuse approaches were used as the true data generating specifications, respectively. The corresponding results for mildly nonstationary condition are illustrated in Figure 3.6, Figure 3.7, and Figure 3.8. Given that no cases converged for the intensive repeated measures data condition in the moderately nonstationary case, only the panel data condition is displayed in Figures 3.9, 3.10, and 3.11. These results are aggregated across measurement parameters, process noise parameters, and autoregressive parameters. Results for individual parameters may be found in Appendix A, Tables 6.1 – 6.45.

3.5 Mildly Nonstationary PFA

Free parameter approach as true initial condition specification: Mild nonstationarity. Main effects with large effect sizes for fitted initial condition included, for the measurement parameters, $SEDIFF$ ($R^2 = .13$), see Figure 3.6C), for both the process noise parameters, absolute bias ($R^2 = .27$), see Figure 3.6D), absolute relative bias ($R^2 = .14$), and $SEDIFF$ ($R^2 = .17$), see Figure 3.6F), and for the time series parameters,$SEDIFF$ ($R^2 = .18$, see Figure 3.6I). While the free-parameter approach yielded the best performance in terms of biases, the de Jong DKF condition yielded the the most consistent estimation of standard errors. The differences between data types were not very salient, with one strong effect for confidence interval width ($R^2 = .13$), where the intensive repeated measures data displayed more narrow confidence intervals than the panel data condition.

Null approach as true initial condition specification: Mild nonstationarity. The strong

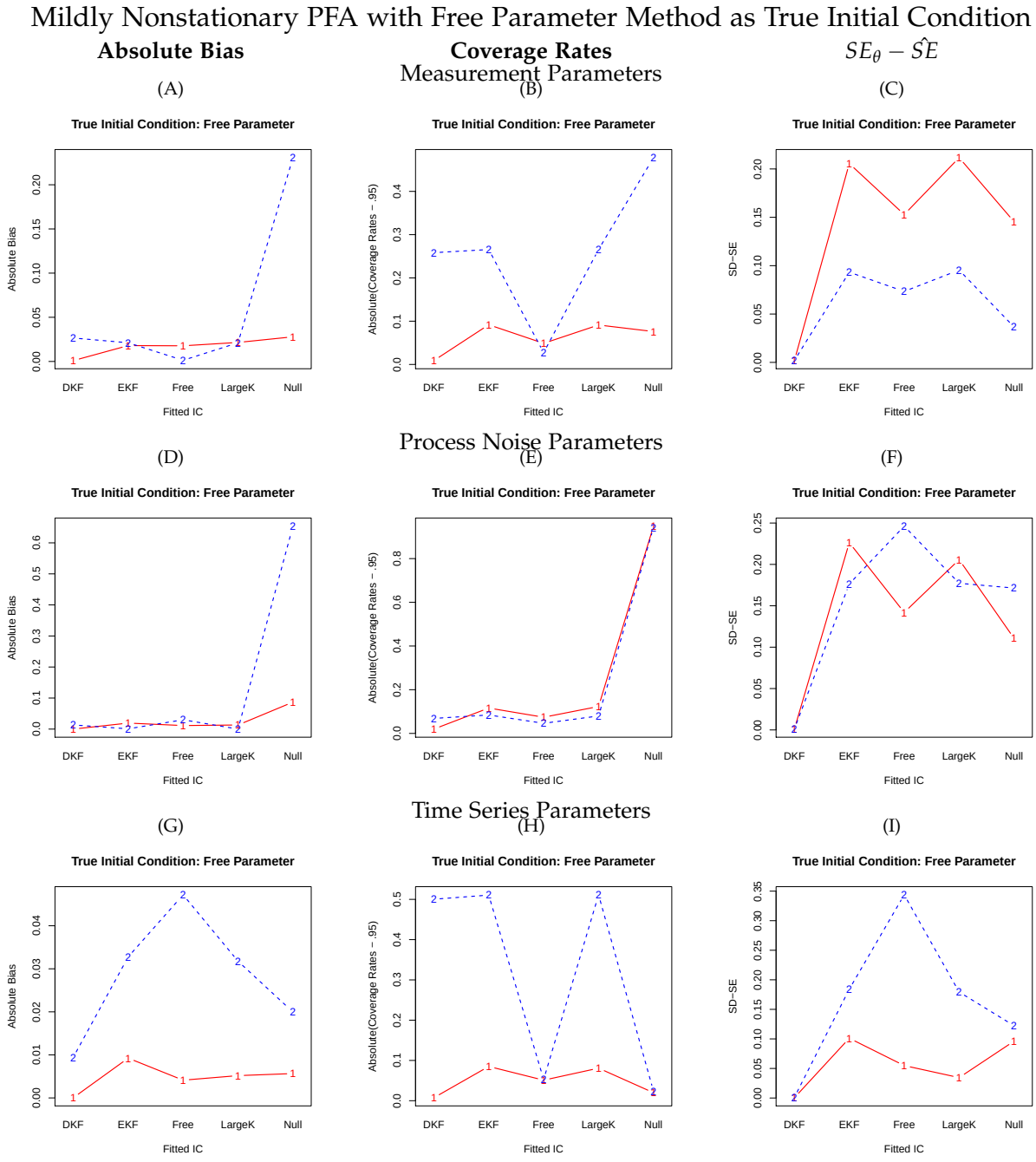

Figure 3.6: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the nonstationary PFA model simulated using the free parameter moments method as the true initial condition specification with mildly nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

main effects for fitted initial condition included, for the process noise parameters, absolute bias ($R^2 = .12$, see Figure 3.7D), and for the time series parameters, $SEDIFF$ ($R^2 = .10$, see Figure 3.7I). For the intensive repeated measures condition, all fitted initial condition approaches performed very well with respect to the outcome measures, except for consistency of standard error estimation. Examining the estimates revealed that there were a few cases that were not identified as outliers but still deviated from the main cluster of points, especially for the measurement error variances. On average, the standard error estimation was not very consistent across parameters, and the de Jong approach displayed the best performance in general. In the panel data case, both the null and free-parameter initial condition performed best with respect to biases and MSE while the de Jong approach performed best with respect to $SEDIFF$.

For the main effect of data type, several strong effects were found for the measurement parameters (absolute bias, $R^2 = .12$, see Figure 3.7A, absolute relative bias, $R^2 = .10$, and confidence interval width, $R^2 = .16$), the process noise parameters (absolute bias, $R^2 = .12$, see Figure 3.7D, absolute relative bias, $R^2 = .13$, and the time series parameters (absolute bias, $R^2 = .12$, see Figure 3.7G, absolute relative bias, $R^2 = .29$, power, $R^2 = .28$, and confidence interval width, $R^2 = .16$). Power has yet to vary much between fitted initial conditions or data type. In this case, the intensive repeated measures condition displayed high power for all fitted initial conditions, while the panel data case displayed high power only for the free-parameter and null initial condition specifications.

Some notable data type by fitted initial condition interaction effects were found for absolute bias for the process noise parameters ($R^2 = .13$, see 3.7F), and for the time series parameters, $SEDIFF$ ($R^2 = .27$, see 3.7I). For the absolute bias of the process noise parameters, the fitted de Jong, exact initial KF, and large κ initial condition approaches all performed well for the intensive repeated measures data but

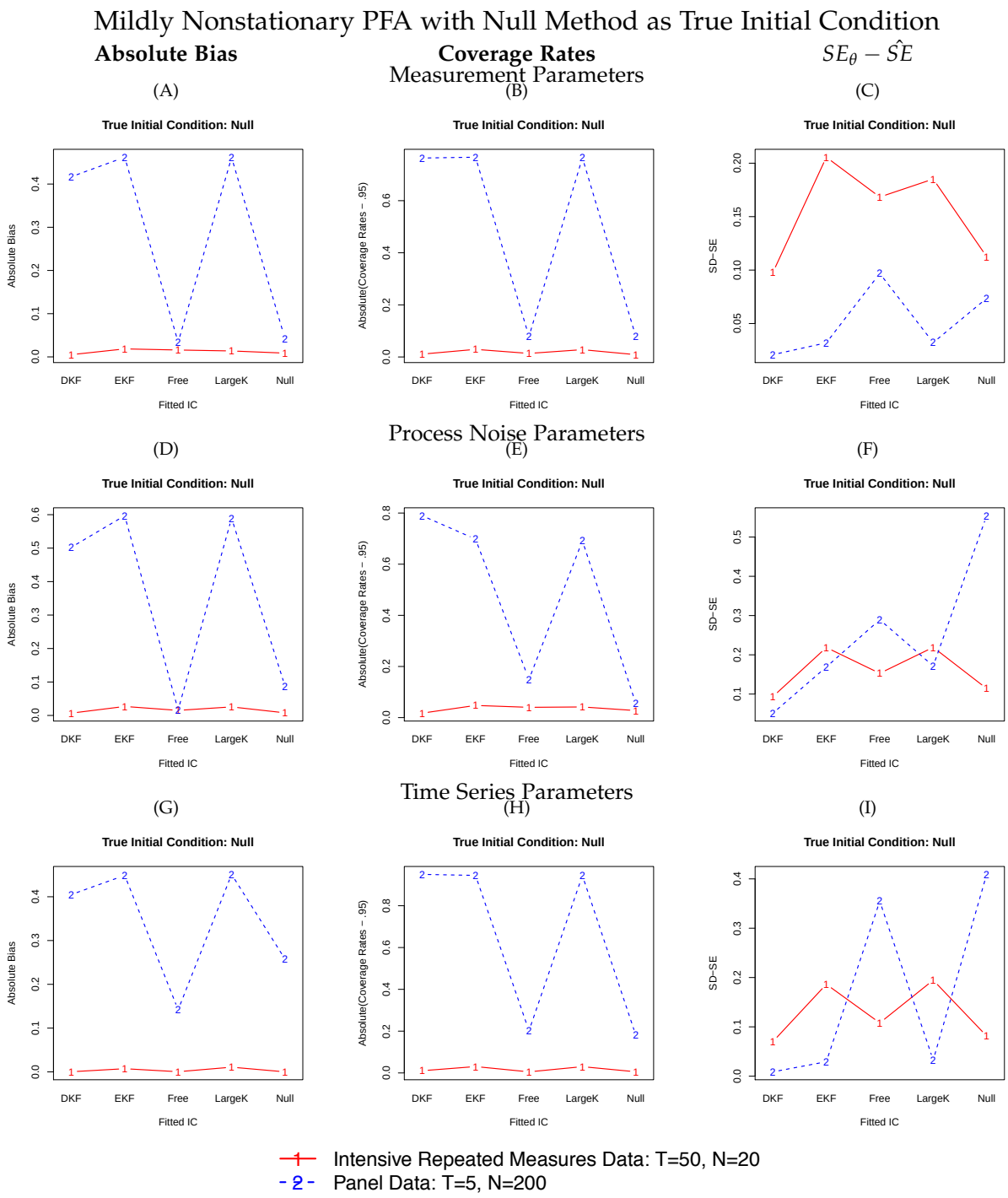

Figure 3.7: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the non-stationary PFA model simulated using the null method as the true initial condition specification with mildly nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

very poorly for the panel data, while the free parameter and null initial condition approaches performed well across both data conditions. For the $SEDIFF$ of the time series parameters, the opposite was true as all initial conditions performed relatively well in the intensive repeated measures case but the null and free-parameter conditions did not perform as well in the panel data case. Again, there were a few replications that produced values that were not detected as outliers but still deviated from the main cluster of values.

Diffuse approach as true initial condition specification: Mild nonstationarity. For this condition, strong main effects for the measurement parameters included absolute bias ($R^2 = .18$, see Figure 3.8A), absolute relative bias, $R^2 = .10$, and $SEDIFF$ ($R^2 = .21$, see Figure 3.8C)). For the process noise parameters, several substantial main effects for fitted initial condition included for absolute bias ($R^2 = .46$, see Figure 3.8D), absolute relative bias ($R^2 = .23$), MSE ($R^2 = .14$), and $SEDIFF$ ($R^2 = .35$, see Figure 3.8F). With respect to biases, the null initial condition did not perform well for the measurement and process noise parameters, but did for the time series parameters. For standard error consistency, however, the null initial condition performed well for all parameters, as did the de Jong approach. For both data types, the free-parameter fitted initial condition displayed good coverage rates for all parameters. The de Jong approach displayed the most optimal results regarding the process noise parameters across all conditions.

One strong data type by fitted initial condition interaction effect was for the absolute bias of the process noise variables ($R^2 = .16$, see Figure 3.8D). Examination of the figure reveals that the null initial condition yielded very poor results only for the panel data condition and only satisfactory bias in the intensive repeated measures condition. All other initial condition specifications and data types yielded good biases.

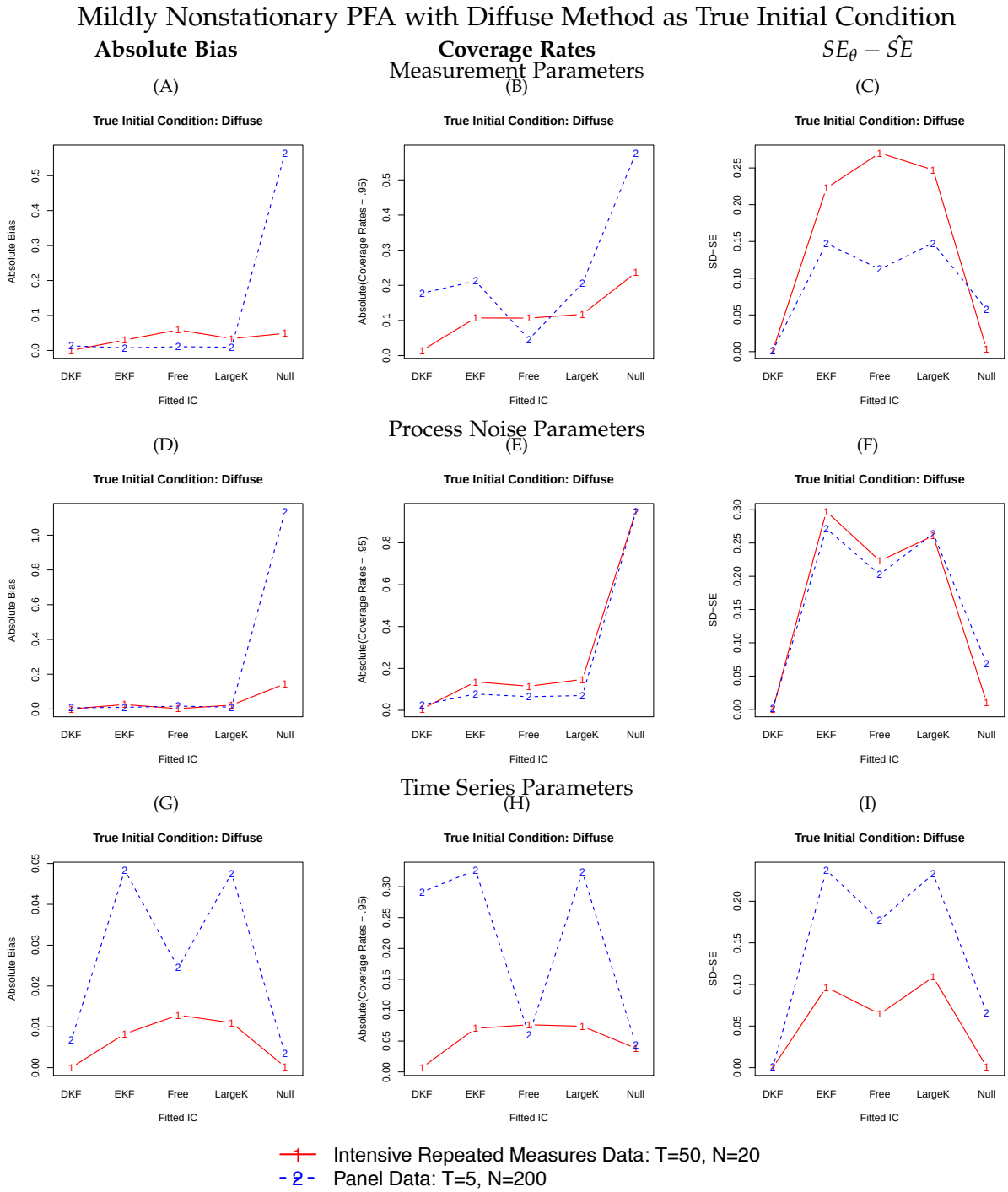

Figure 3.8: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the nonstationary PFA model simulated using the diffuse method as the true initial condition specification with mildly nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

3.6 Moderately Nonstationary PFA

Because the panel data condition is the only data type for the moderately nonstationary condition, the ANOVA model is modified accordingly,

$$\text{ANOVA Model: SimOutcome}_i = \text{FittedIC}_i + \epsilon_i \tag{3.2}$$

so that it is a one-way ANOVA comparing only the mean differences between fitted initial conditions.

When examining the panel data case, the differences between fitted initial condition specifications for mildly nonstationary case were less substantial compared to the moderately nonstationary condition. Although the pattern of results is similar to the moderate stationarity condition, This suggests that as the process becomes less nonstationary, the differences in approaches become less noticeable.

Free parameter approach as true initial condition specification: Moderate nonstationarity. When the true initial condition was specified as having free parameters, several simulation outcomes displayed a large main effect of fitted initial condition. Specifically, the measurement parameters displayed a large R^2 value for the main effect of absolute bias ($R^2 = .17$, see Figure 3.9A), absolute bias ($R^2 = .13$), and $SEDIFF$ ($R^2 = .18$, see Figure 3.9C); the process noise parameters displayed a large R^2 value for the main effect for absolute bias ($R^2 = .64$, see Figure 3.9D), relative bias ($R^2 = .49$), MSE ($R^2 = .15$), and $SEDIFF$ ($R^2 = .75$, see Figure 3.9F); and the time series parameters displayed a large R^2 value for the main effect for $SEDIFF$ ($R^2 = .22$, see Figure 3.9I). With respect to biases, coverage rates, the fitted free parameter approach performed the best, especially for the time series parameters. The de Jong DKF displayed the most consistent standard error estimation and both the free-parameter and de Jong approach displayed the smallest MSE values.

Null approach as true initial condition specification: Moderate nonstationarity. When

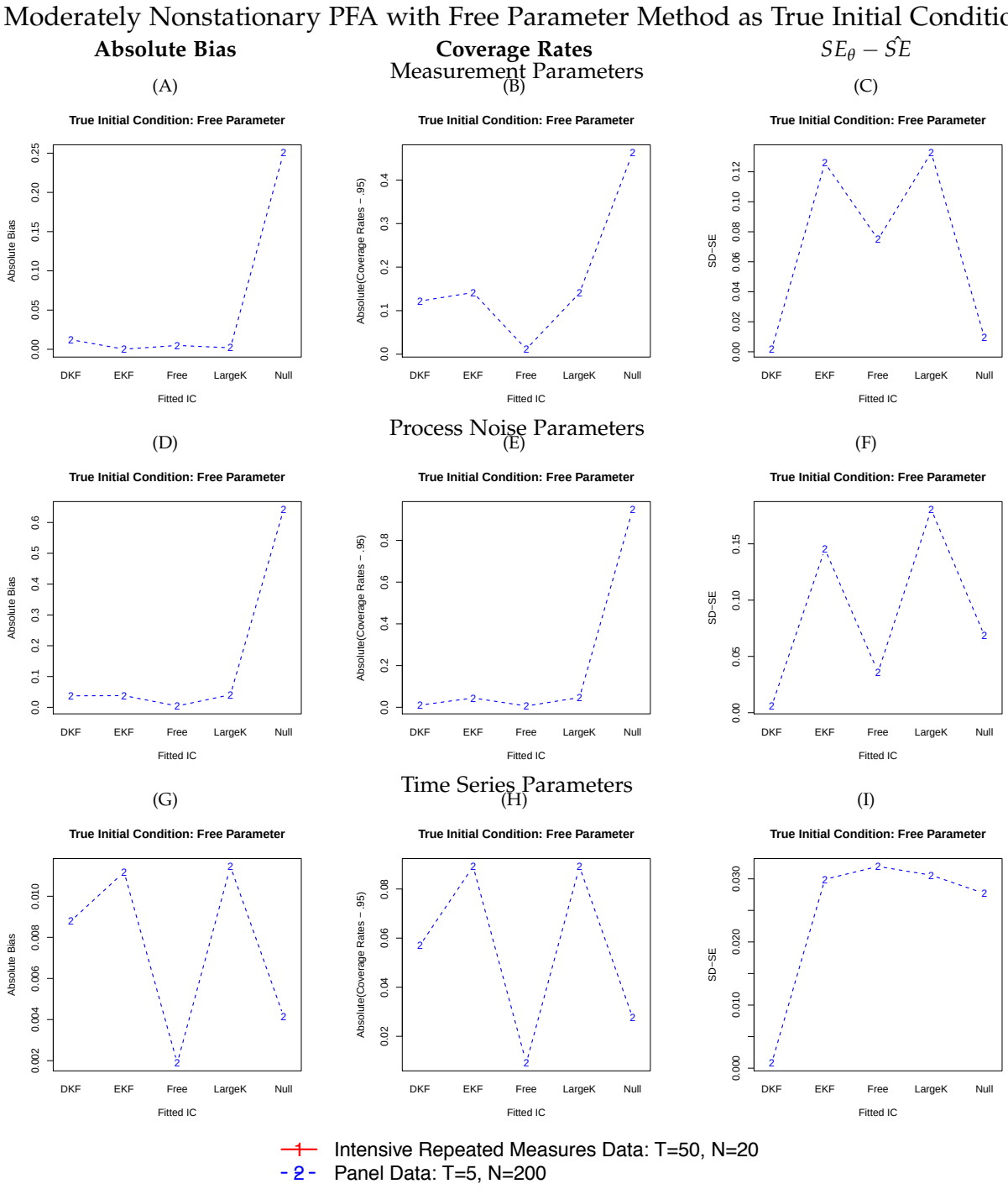

Figure 3.9: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the non-stationary PFA model simulated using the free parameter moments method as the true initial condition specification with moderately nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

the true initial condition specification were null matrices, the measurement parameters displayed a strong main effect of fitted initial condition for $SEDIFF$ ($R^2 = .27$, see Figure 3.10C). Fitting the large κ and Koopman exact initial KF specifications resulted in the least consistent standard error estimation, following by fitting the free-parameter approach, de Jong DKF, while fitting both the null approach yielded very consistent standard error estimation. The process noise parameters displayed very strong R^2 values for the main effect of fitted initial condition for several simulation outcome measures, including absolute bias ($R^2 = .11$, see Figure 3.10D), coverage rates($R^2 = .22$, see Figure 3.10E), $SEDIFF$ ($R^2 = .61$, see Figure3.10F), and confidence interval width ($R^2 = .10$).

When examining the plots, it is clear that the null condition performed the best with respect to absolute bias across the simulation outcomes. With respect to standard error consistency, the de Jong DKF approach also performed well. As far as biases and coverage rates, the free-parameter condition also performed well. Both the Koopman exact initial KF and the large κ fitted initial condition performed the worst. Thus far, the de Jong DKF fitted approach has regularly produced very consistent standard error estimation results, especially for the process noise parameters.

Diffuse approach as true initial condition specification: Moderate nonstationarity. The results from this condition are of particular interest as an important goal of this research was to determine whether fitting either the de Jong DKF or the Koopman exact initial KF to panel data simulated from a process that is nonstationary would result in more accurate, efficient, and consistent parameter estimation than other known stationary approaches. Because the fitted free parameter initial condition approach has proven to work well thus far, it is important to carefully compare this approach to the de Jong DKF and Koopman exact initial KF approaches.

The strong main effects of fitted initial condition for the measurement parameters were for absolute bias ($R^2 = .41$, see Figure 3.11A), absolute relative bias ($R^2 = .28$),

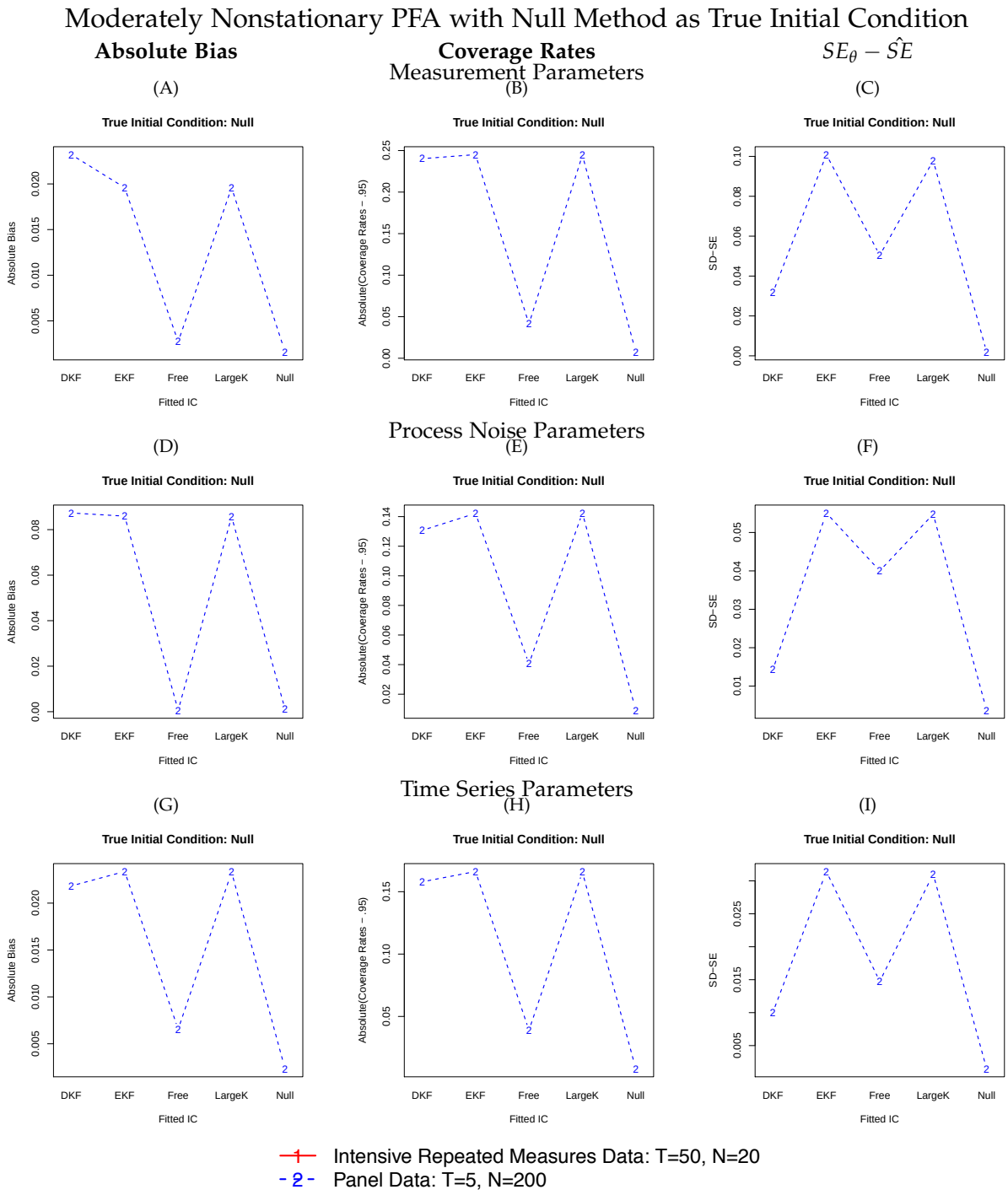

Figure 3.10: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the non-stationary PFA model simulated using the null method as the true initial condition specification with moderately nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

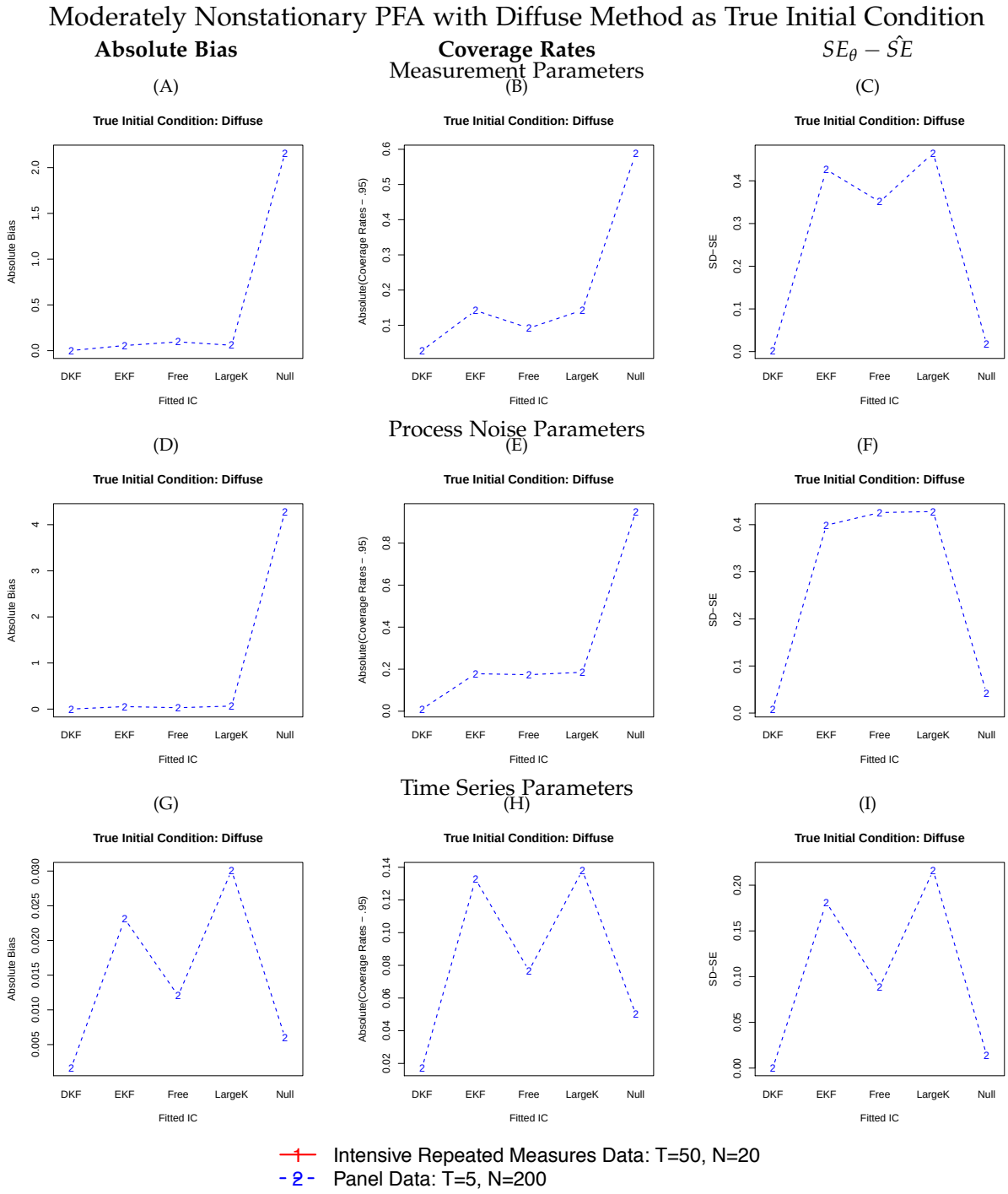

Figure 3.11: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the nonstationary PFA model simulated using the diffuse method as the true initial condition specification with moderately nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

coverage rates ($R^2 = .48$, see Figure 3.11B), MSE ($R^2 = .30$), $SEFIFF$ ($R^2 = .27$, see Figure 3.11C), and confidence interval width ($R^2 = .32$). The fitted de Jong DKF initial condition approach performed best with respect to these outcomes. With respect to confidence interval coverage, the fitted Koopman exact initial KF approach and free parameter approach did not perform as well, and the large κ approach displayed the poorest results. The null initial condition displayed poor bias and poor coverage rates, but good standard error consistency. This indicates that the null initial condition is providing consistent standard error estimates, but they are clustered around the wrong value.

For the process noise parameters, most simulation outcome measures displayed a large effect size, including absolute bias ($R^2 = .80$, see Figure 3.11D), absolute relative bias ($R^2 = .90$), coverage rates ($R^2 = .97$, see Figure 3.11E), MSE ($R^2 = .64$), $SEDIFF$ ($R^2 = .85$, see Figure 3.11F), and confidence interval width ($R^2 = .91$). With respect to relative bias, absolute relative bias, coverage rates, and confidence interval width, the de Jong DKF yielded the best results, followed closely by the free parameter condition, Koopman exact initial KF approach, and large κ approach, with the null initial condition performing by far the worst. For standard error estimation consistency, however, the de Jong DKF outperformed all approaches, followed by the null initial condition. Strong main effects for the time series parameters included coverage rates ($R^2 = .41$, see Figure 3.11H) and $SEDIFF$ ($R^2 = .29$, see Figure 3.11I). The de Jong DKF approach yielded the best coverage rates followed by the null initial condition approach.

One interesting result is that the fitted null initial condition performed well with respect to time series parameters. In fact, for these two sets of parameters, it performed the second best followed by the de Jong approach. However, for the measurement and process noise parameters, the fitted null initial condition did not perform well at all with respect to both biases and coverage rates, despite still displaying good

standard error consistency.

3.7 Summary of Nonstationary Results

As in the stationary condition, some general trends were observed across the different true initial condition specifications. First, the AIC and BIC tended to choose the de Jong and Koopman approaches, especially in the panel data condition. Second, the intensive repeated measures condition displayed better results, on average, than the panel data condition, in terms of both point and standard error estimates. Third, given the moderately nonstationary model, the fitted initial condition specifications performed worse, on average, regardless of true initial condition. Also, in this more explosive condition the fitted de Jong DKF outperformed the other models to a larger degree both in terms of 1. proper convergence rates and 2. parameter accuracy, efficiency, and consistency. This suggests that, as the process becomes more nonstationary (i.e., more explosive), the degree to which the de Jong DKF approach outperforms the other approaches increases. Fourth, the parameters that were affected most substantially by both misspecification in initial condition and a larger degree of nonstationarity were the process noise parameters. Fifth, when the null initial condition was the true initial condition, two interesting results emerged: 1. the de Jong approach did not do well in terms of both convergence rates and simulation outcomes, and 2. the free-parameter approach underestimated the standard errors for all parameters.

Also, the free-parameter approach performed well, with the exception of posing convergence issues and sometimes producing very biased process noise variance point estimates. When compared to the de Jong approach, the large κ approach did not perform as well, particularly with respect to standard error estimates. This is an important finding given that, in the time series literature, this approach is widely used when there is a diffuse initial condition.

3.8 General Simulation Conclusions

Overall, several trends were revealed across both data type and fitted initial condition specifications. In the stationary conditions, the free parameter approach worked very well most of the time, followed by the null initial condition approach. The model-implied moments condition did not perform as expected, perhaps due to the complicated constraints placed on the initial condition covariance matrix and not applying stationary constraints in the estimation procedure. When using intensive repeated measures data all fitted initial condition specifications worked relatively well for all true initial condition data generating cases.

For the nonstationary conditions, the initial condition specifications performed better when the simulated process was mildly nonstationary. The de Jong DKF approach worked exceptionally well with the intensive repeated measures data, as biases were very nearly zero, MSE values were very small, coverage rates very close to .95%, and power was very high for all parameters, indicating good parameter accuracy, efficiency, and consistency. However, in the panel data case, when using the null true initial condition specification, the de Jong DKF approach did not perform well, nor did the Koopman exact initial KF and large κ approaches. The free parameter approach, however, worked well in this case. However, when data were simulated according to a diffuse initial condition specification, the fitted de Jong DKF performed just as well if not better than both the free parameter condition and the null initial condition (with a few exceptions in the mildly nonstationary case). Also, on average, fitting the de Jong DKF approach resulted in very consistent standard error estimates. As expected, the large κ approach did not work well under most circumstances.

Somewhat surprisingly, the Koopman exact initial KF did not work very well. As the de Jong DKF and Koopman exact initial KF should in theory produce the same results, this must be due to convergence and outlier issues. The next section discusses results obtained when removing the boundary cases and sheds some light

on this discrepancy.

3.9 Results Removing Boundary Cases

The next set of results displays simulation outcomes when boundary cases, i.e., cases where variance parameter estimates were very close to zero, were removed (see Section 3.1 for a more in depth discussion of this issue). Results of key point and standard error estimates are presented in Figure 3.12 to Figure 3.20. One noticeable change is that the estimates are now more similar to each other, especially those derived from the de Jong DKF, Koopman initial KF, and large κ conditions. Also, the free-parameter approach works very well under most circumstances now, even when the data are moderately nonstationary and the diffuse initialization procedure was used to generated the data (see Figure 3.20). This suggests that, when the replications properly converge and a generalized inverse is not used, then the free-parameter approach performs as well as the de Jong approach. However, there were a large number of replications in the free-parameter approach where a boundary condition was observed for the process noise variables leaving only 173 retained replications. The de Jong DKF, however, still produced the most number of retained cases with 415.

Instead of detailing all results, I will highlight the results that differed substantially from the results retaining the boundary cases, starting with the stationary model. For this model, one interesting result is that, the standard error consistency tended to improve across conditions. When the true initial condition was null, the standard error consistency for the de Jong and null approaches improved when the boundary cases were removed, as can be seen by comparing Figure 3.13C, F, and I to Figure 3.13C, F, and I. This trend was also revealed when the model-implied approach served as true initial condition for the standard error consistent of the measurement variables for the de Jong, Koopman, free parameter, and large κ approaches.

Removing the boundary cases had a greater impact on results for the nonstation-

Table 3.4: Rates of Convergence to a Proper Solution, AIC and BIC Model Selection, Latent Variable Recovery, and Average Total Time Obtained from Stationary PFA Model Removing Boundary Cases.

N/T	True IC	Fitted IC	Final # Reps	Strong Conv	Weak Conv	No Conv	Outliers	no SEs	AIC	BIC	$\overline{RMSE}$lvs[a]	$\overline{Time}$[b]
		Model Implied	36	386	0	45	353	3	0	0	0.46	14.318
		Free Parameter	446	480	9	11	34	0	0	0	0.429	11.599
	Model	Null	424	484	1	15	60	0	0	0	0.69	7.273
	Implied	de Jong DKF	496	500	0	0	4	0	279	279	0.464	16.633
		Koopman exact initial KF	455	494	1	2	39	0	221	221	0.465	7.803
		Large κ	459	496	1	3	37	0	0	0	0.465	7.426
		Model Implied	22	370	1	63	354	5	0	0	0.431	19.809
	Free	Free Parameter	446	476	9	15	30	0	0	0	0.426	14.774
N=200	Parameter	Null	445	479	0	21	34	0	0	0	0.703	9.221
T=5		de Jong DKF	500	500	0	0	0	0	280	280	0.464	21.702
		Koopman exact initial KF	445	488	0	8	43	0	220	220	0.464	10.334
		Large κ	448	492	0	8	44	0	0	0	0.464	9.941
		Model Implied	51	365	0	80	322	6	0	0	0.42	15.856
		Free Parameter	378	427	35	38	49	0	0	0	0.391	19.041
	Null	Null	454	488	0	12	34	0	0	0	0.381	8.13
		de Jong DKF	499	499	0	1	0	1	272	272	0.494	19.651
		Koopman exact initial KF	455	481	1	16	26	0	227	227	0.493	8.868
		Large κ	459	484	0	16	25	0	0	0	0.493	8.474
		Model Implied	402	478	0	13	76	0	32	32	0.425	4.204
		Free Parameter	472	492	3	5	26	0	16	16	0.425	6.433
	Model	Null	476	497	2	1	21	0	3	3	0.458	3.652
	Implied	de Jong DKF	498	500	0	0	2	0	232	232	0.426	6.947
		Koopman exact initial KF	474	489	6	5	20	0	217	217	0.426	3.681
		Large κ	474	490	4	6	20	0	0	0	0.426	3.667
		Model Implied	404	482	0	11	79	0	46	46	0.424	4.358
	Free	Free Parameter	476	494	0	6	20	0	53	53	0.423	6.497
N=20	Parameter	Null	477	494	3	3	18	0	1	1	0.458	3.749
T=50		de Jong DKF	497	500	0	0	3	0	186	186	0.425	7.347
		Koopman exact initial KF	476	498	1	1	22	0	214	214	0.425	3.849
		Large κ	476	496	1	3	20	0	0	0	0.425	3.773
		Model Implied	416	471	0	11	55	0	49	49	0.425	6.582
		Free Parameter	448	490	5	5	42	0	87	87	0.421	13.686
	Null	Null	476	494	3	3	18	0	81	81	0.42	5.602
		de Jong DKF	495	499	0	1	4	0	132	132	0.426	10.728
		Koopman exact initial KF	475	496	2	2	22	0	149	149	0.426	5.665
		Large κ	475	497	3	0	22	0	0	0	0.426	5.569

[a] $\overline{RMSE}$lvs represents the average RMSE value (see Equation 2.10) across replications for true latent variable scores versus estimated latent variable scores. Smaller values indicate a greater correspondence between the true and estimated scores.

[b] $\overline{Time}$ represents the average time across replications taken to estimate the model, whether convergence was reached or not.

Table 3.5: Rates of Convergence to a Proper Solution, AIC and BIC Model Selection, Latent Variable Recovery, and Average Total Time Obtained from Nonstationary PFA Model Mild Nonstationarity Removing Boundary Cases.

N/T	True IC	Fitted IC	Final # Reps	Strong Conv	Weak Conv	No Conv	Outliers	no SEs	AIC	BIC	$\overline{RMSE}$lvs[a]	$\overline{Time}$[b]
N=200 T=5	Free Parameter	Free Parameter	387	441	17	42	54	0	9	9	0.268	17.001
		Null	407	459	3	38	52	0	0	0	0.654	9.848
		de Jong DKF	462	484	0	16	22	15	269	269	0.298	23.11
		Koopman exact initial KF	415	488	0	6	73	1	213	213	0.298	11.692
		Large κ	417	493	1	6	76	0	1	1	0.298	11.272
	Null	Free Parameter	229	330	45	125	103	0	9	9	0.216	46.468
		Null	299	320	18	162	22	1	52	52	1.761	21.245
		de Jong DKF	79	256	142	102	3	390	48	48	0.752	34.437
		Koopman exact initial KF	301	350	20	114	49	15	235	235	0.88	23.354
		Large κ	306	352	12	136	46	1	6	6	0.88	22.184
	Diffuse	Free Parameter	330	386	30	84	56	0	4	4	0.271	17.618
		Null	388	443	0	57	55	0	1	1	0.948	8.353
		de Jong DKF	435	465	0	35	33	23	273	273	0.286	21.989
		Koopman exact initial KF	355	447	5	44	92	2	185	185	0.286	11.374
		Large κ	366	450	2	48	85	0	0	0	0.286	10.967
N=20 T=50	Free Parameter	Free Parameter	228	296	76	128	102	3	32	32	0.231	17.315
		Null	247	258	64	178	46	0	0	0	0.311	9.103
		de Jong DKF	387	424	17	59	1	104	239	239	0.231	20.133
		Koopman exact initial KF	245	340	74	76	119	5	116	116	0.231	12.891
		Large κ	239	342	87	71	132	3	1	1	0.231	13.05
	Null	Free Parameter	246	357	42	101	126	1	60	60	0.225	14.532
		Null	358	374	23	103	29	1	102	102	0.224	6.908
		de Jong DKF	423	463	3	34	4	70	147	147	0.23	13.021
		Koopman exact initial KF	364	424	20	48	74	6	118	118	0.23	7.989
		Large κ	363	429	15	56	77	2	0	0	0.23	7.773
	Diffuse	Free Parameter	158	236	70	194	133	0	13	13	0.231	15.049
		Null	192	211	53	236	42	0	0	0	0.391	7.735
		de Jong DKF	335	365	26	109	1	151	220	220	0.231	17.53
		Koopman exact initial KF	183	295	88	107	141	9	100	100	0.231	12.206
		Large κ	180	311	96	93	161	4	2	2	0.231	11.638

[a] $\overline{RMSE}$lvs represents the average RMSE value (see Equation 2.10) across replications for true latent variable scores versus estimated latent variable scores. Smaller values indicate a greater correspondence between the true and estimated scores.

[b] $\overline{Time}$ represents the average time across replications taken to estimate the model, whether convergence was reached or not.

Table 3.6: Rates of Convergence to a Proper Solution, AIC and BIC Model Selection, Latent Variable Recovery, and Average Total Time Obtained from Nonstationary PFA Model with Moderate Nonstationarity Removing Boundary Cases.

N/T	True IC	Fitted IC	Final # Reps	Strong Conv	Weak Conv	No Conv	Outliers	no SEs	AIC	BIC	$\overline{RMSE}$lvs[a]	$\overline{Time}$[b]
N=200 T=5	Free Parameter	Free Parameter	363	436	13	51	72	0	5	5	0.405	14.662
		Null	384	467	0	33	83	0	0	0	0.687	9.204
		de Jong DKF	467	491	0	9	24	4	302	302	0.419	18.732
		Koopman exact initial KF	354	471	1	26	117	0	183	183	0.418	11.651
		Large κ	355	470	1	29	114	0	0	0	0.418	10.463
	Null	Free Parameter	365	432	29	39	67	0	3	3	0.366	20.08
		Null	440	476	0	24	36	0	2	2	0.365	8.749
		de Jong DKF	486	497	0	3	12	1	269	269	0.426	19.915
		Koopman exact initial KF	442	488	0	9	46	0	222	222	0.426	9.417
		Large κ	443	491	0	9	48	0	0	0	0.426	9.099
	Diffuse large	Free Parameter	173	334	17	149	161	0	9	9	0.413	21.042
		Null	344	426	2	72	82	0	8	8	1.745	9.568
		de Jong DKF	415	436	0	64	23	31	341	341	0.415	23.292
		Koopman exact initial KF	158	353	9	123	85	2	77	77	0.414	15.295
		Large κ	155	367	11	122	82	0	0	0	0.414	15.354

[a] $\overline{RMSE}$lvs represents the average RMSE value (see Equation 2.10) across replications for true latent variable scores versus estimated latent variable scores. Smaller values indicate a greater correspondence between the true and estimated scores.

[b] $\overline{Time}$ represents the average time across replications taken to estimate the model, whether convergence was reached or not.

ary models. For the mildly and moderately nonstationary conditions, the standard error consistency improved for the free-parameter fitted initial condition across all true initial condition specifications. For the Koopman and large κ conditions, an improvement was recorded for all conditions except with a true null initial condition in the moderately nonstationary model. However, the fitted null initial condition did not display as large an improvement in standard error consistency as the other conditions. Thus, the relatively good standard error consistency the null initial condition approach displayed in the results that included the boundary cases may have been due to the other approaches displaying poor standard error consistency due to boundary conditions. Still, the de Jong approach did not lose as many cases due to have a boundary condition, and displayed only a slight improvement in standard error consistency.

The free-parameter approach displayed an improvement in absolute bias for several conditions including, for the mildly nonstationary model, the time series parameters (for all true initial condition specifications) and, in the moderately nonstationary model, for the time series parameters only when the true process was diffuse. For the moderately nonstationary condition, coverage rates for all parameters improved for the free-parameter, Koopman, and large κ approaches. As illustrated in Figure 3.20, the optimal performance of the de Jong approach has disappeared as the free-parameter, Koopman, and large κ approaches all perform consistently well. Given that the correspondence between the de Jong and Koopman approach is now much greater, it is likely that the cases with boundary conditions were causing the discrepancy between these two diffuse approaches.

An inspection of the estimated values in the free-parameter initial condition matrices may help to explain why the free-parameter approach both 1. displayed a large number of boundary cases and 2. improved so much with the removal of such boundary cases. When boundary cases were removed and the true initial condition

specification was diffuse, for the mildly nonstationary model the free-parameter estimated the following values in the mean and covariance matrix of the initial condition, averaged across replications,

$$\text{T=5: } \hat{\boldsymbol{\mu}}_0 = \begin{bmatrix} .01 \\ 0 \end{bmatrix}, \hat{\mathbf{P}}_0 = \begin{bmatrix} 3.694 & 1.908 \\ 1.908 & 3.684 \end{bmatrix}, \text{T=50: } \hat{\boldsymbol{\mu}}_0 = \begin{bmatrix} .043 \\ .043 \end{bmatrix}, \hat{\mathbf{P}}_0 = \begin{bmatrix} 3.572 & 1.811 \\ 1.811 & 3.455 \end{bmatrix} \quad (3.3)$$

When the boundary cases were retained, the estimated values were,

$$\text{T=5: } \hat{\boldsymbol{\mu}}_0 = \begin{bmatrix} .003 \\ .075 \end{bmatrix}, \hat{\mathbf{P}}_0 = \begin{bmatrix} 3.696 & 2.518 \\ 2.518 & 3.591 \end{bmatrix}, \text{T=50: } \hat{\boldsymbol{\mu}}_0 = \begin{bmatrix} .013 \\ .025 \end{bmatrix}, \hat{\mathbf{P}}_0 = \begin{bmatrix} 3.60 & 1.745 \\ 1.745 & 3.17 \end{bmatrix} \quad (3.4)$$

In the moderately nonstationary condition, the estimated values were,

$$\text{Boundary Cases Removed: } \hat{\boldsymbol{\mu}}_0 = \begin{bmatrix} .036 \\ .026 \end{bmatrix}, \hat{\mathbf{P}}_0 = \begin{bmatrix} 13.316 & 4.02 \\ 4.02 & 12.872 \end{bmatrix} \quad (3.5)$$

$$\text{Boundary Cases Retained: } \hat{\boldsymbol{\mu}}_0 = \begin{bmatrix} .024 \\ -.014 \end{bmatrix}, \hat{\mathbf{P}}_0 = \begin{bmatrix} 13.295 & 3.498 \\ 3.498 & 14.072 \end{bmatrix}. \quad (3.6)$$

Interestingly, as the process becomes more explosive, as in the moderately diffuse case, the estimated variances and covariances become larger. This is because the diffuse part is being estimated freely, and the more diffuse a process is the more variability will be displayed. Thus, if the variability is being modeled in the initial condition matrices, perhaps the estimated values for the measurement errors is becoming very small. This is a problem when using this approach that the de Jong approach avoids by more directly estimating the mean and covariance of the diffuse process.

Overall, it is possible that the discrepancy between fitted initial condition and true initial condition would cause more replications to fail in the regular inversion

of the Hessian matrix, especially if the process is highly nonstationary. This makes the use of de Jong DKF seem more viable in practice because it was the approach that produced the most retained replications over all, except when the true initial condition was null. As empirical data sets tend to be noisier than simulated data sets, it's possible that the de Jong approach would be the only one to converge properly and not hit a boundary condition.

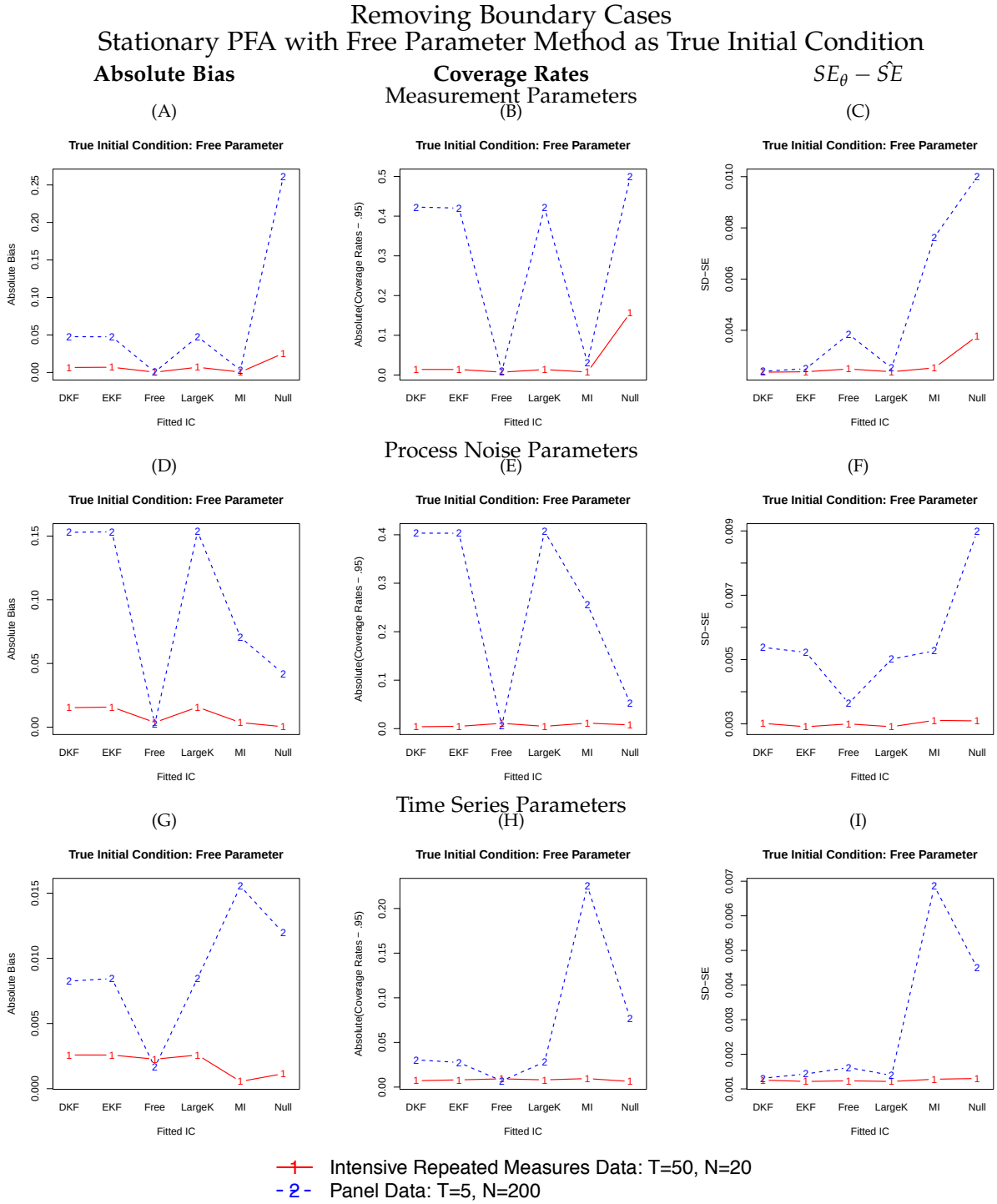

Figure 3.12: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the stationary PFA model simulated using the free parameter moments method as the true initial condition specification. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

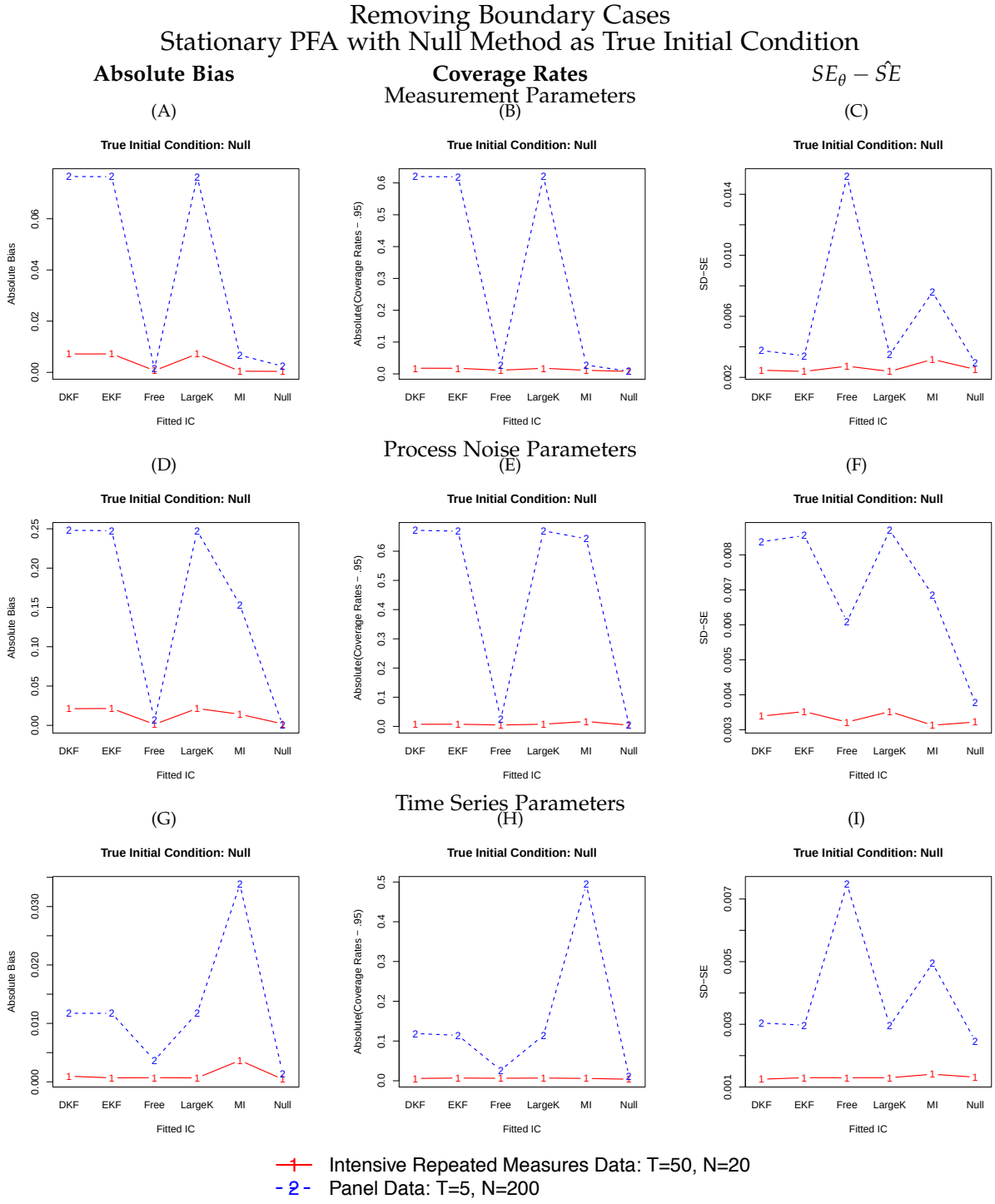

Figure 3.13: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the stationary PFA model simulated using the null method as the true initial condition specification. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

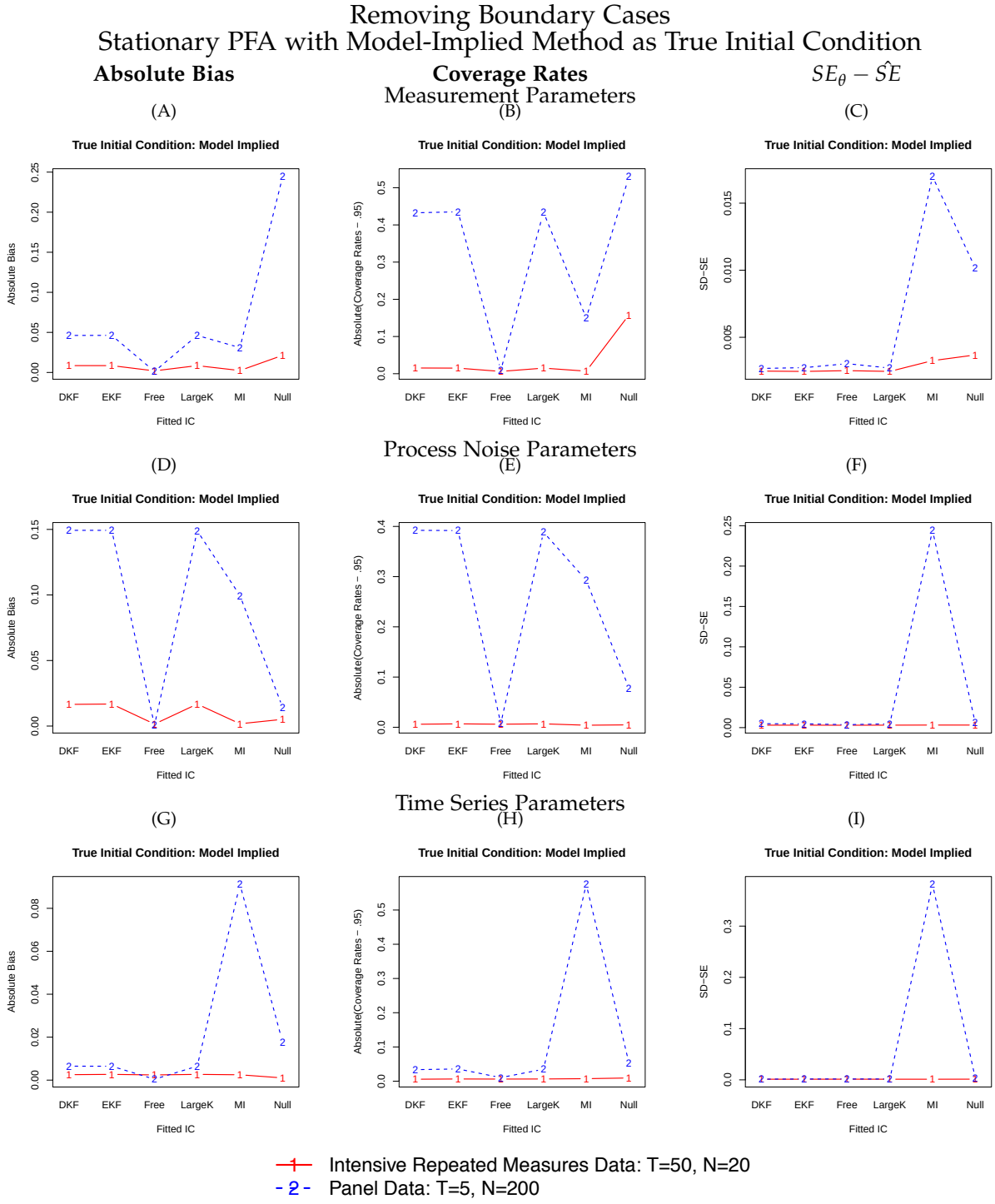

Figure 3.14: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the stationary PFA model simulated using the model-implied moments method as the true initial condition specification. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

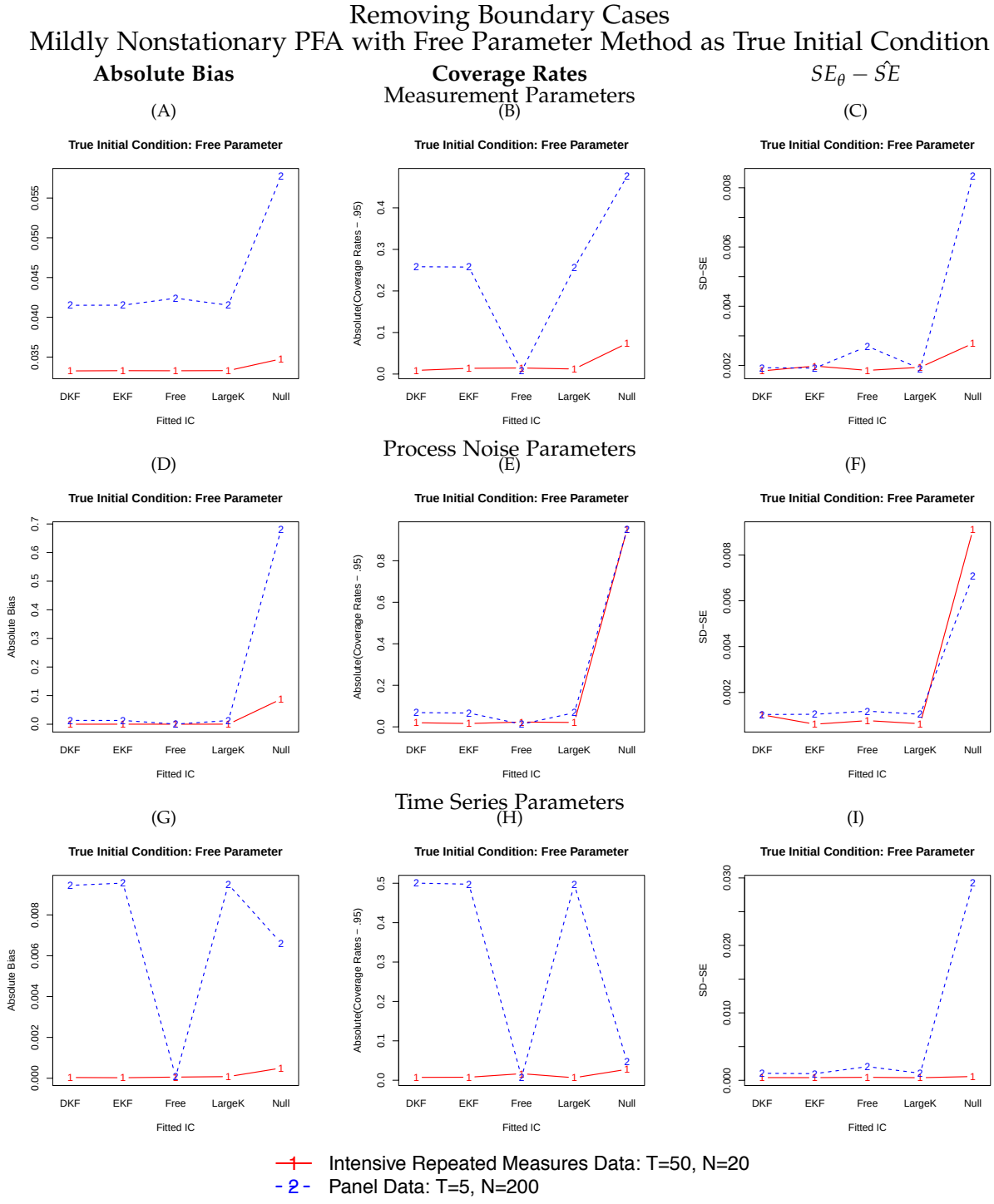

Figure 3.15: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the nonstationary PFA model simulated using the free parameter moments method as the true initial condition specification with mildly nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

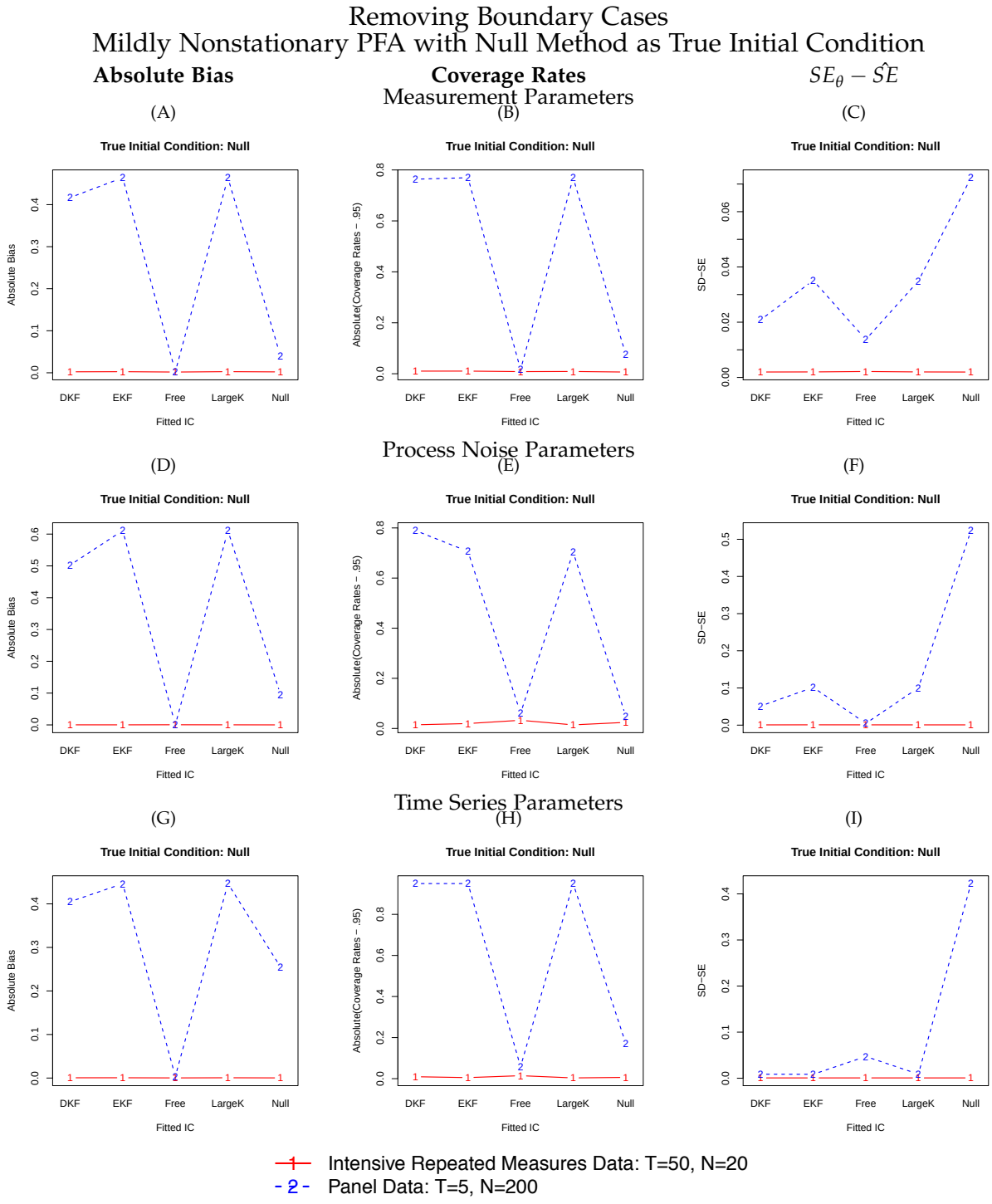

Figure 3.16: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the non-stationary PFA model simulated using the null method as the true initial condition specification with mildly nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

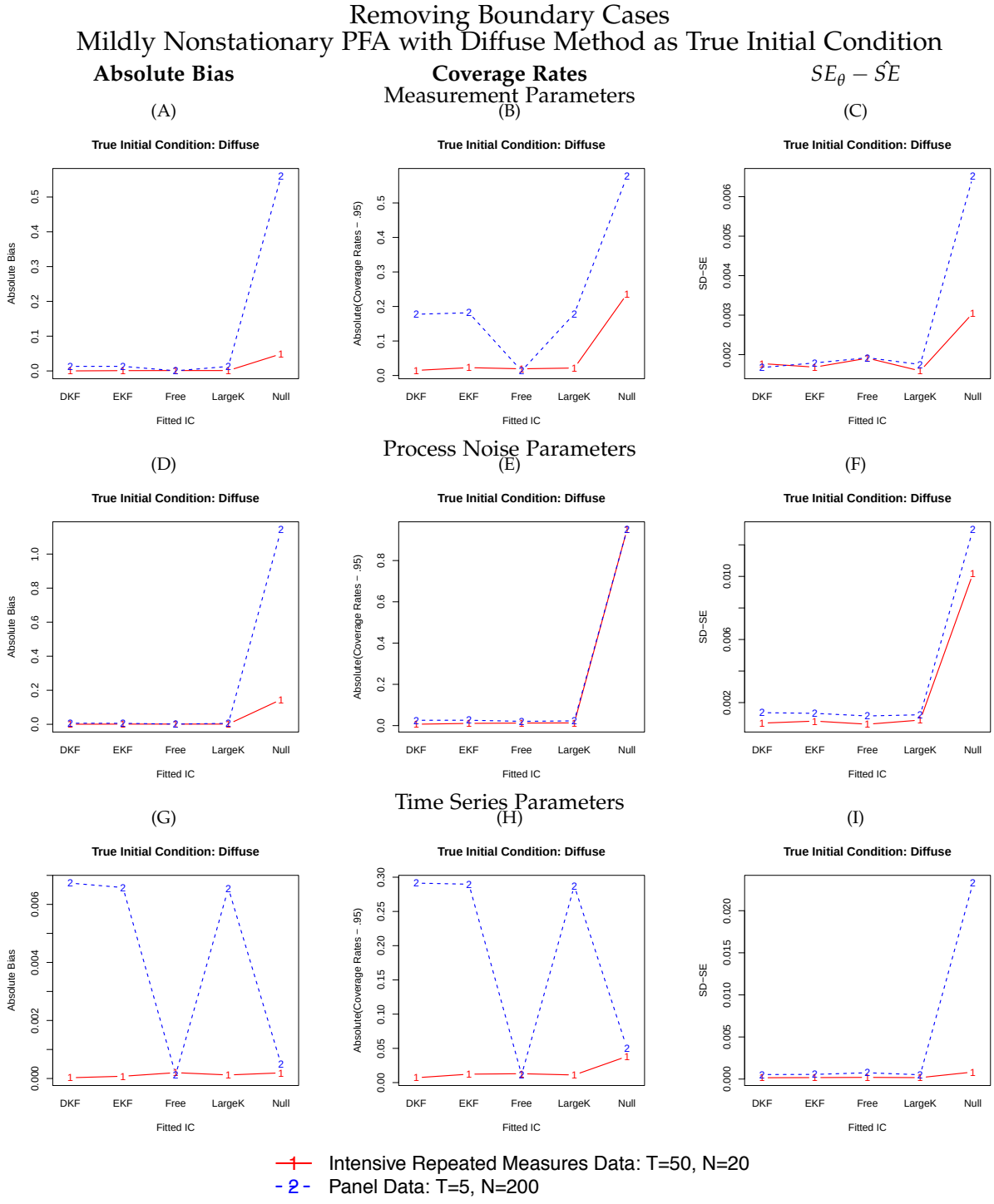

Figure 3.17: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the nonstationary PFA model simulated using the diffuse method as the true initial condition specification with mildly nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

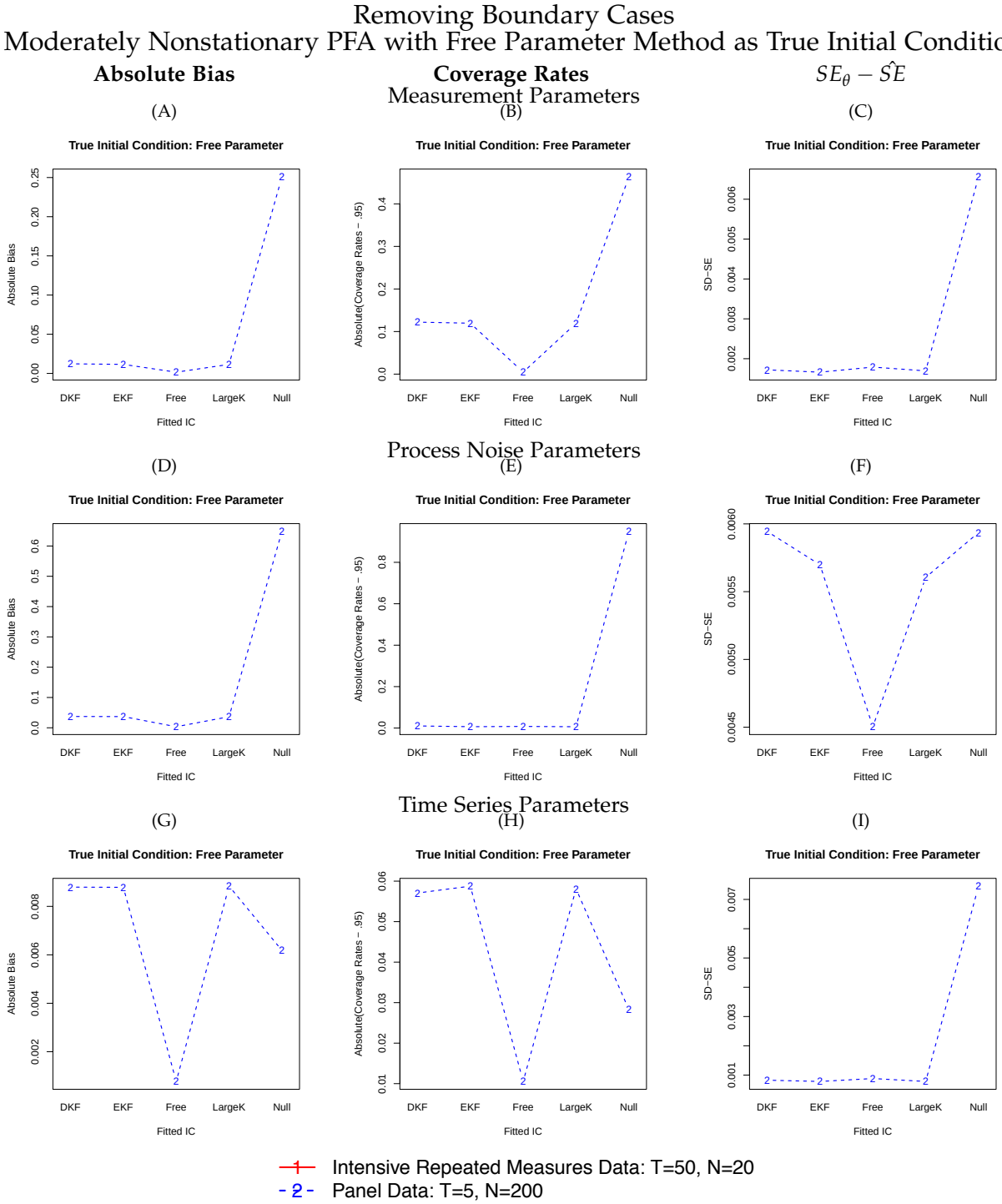

Figure 3.18: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the non-stationary PFA model simulated using the free parameter moments method as the true initial condition specification with moderately nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

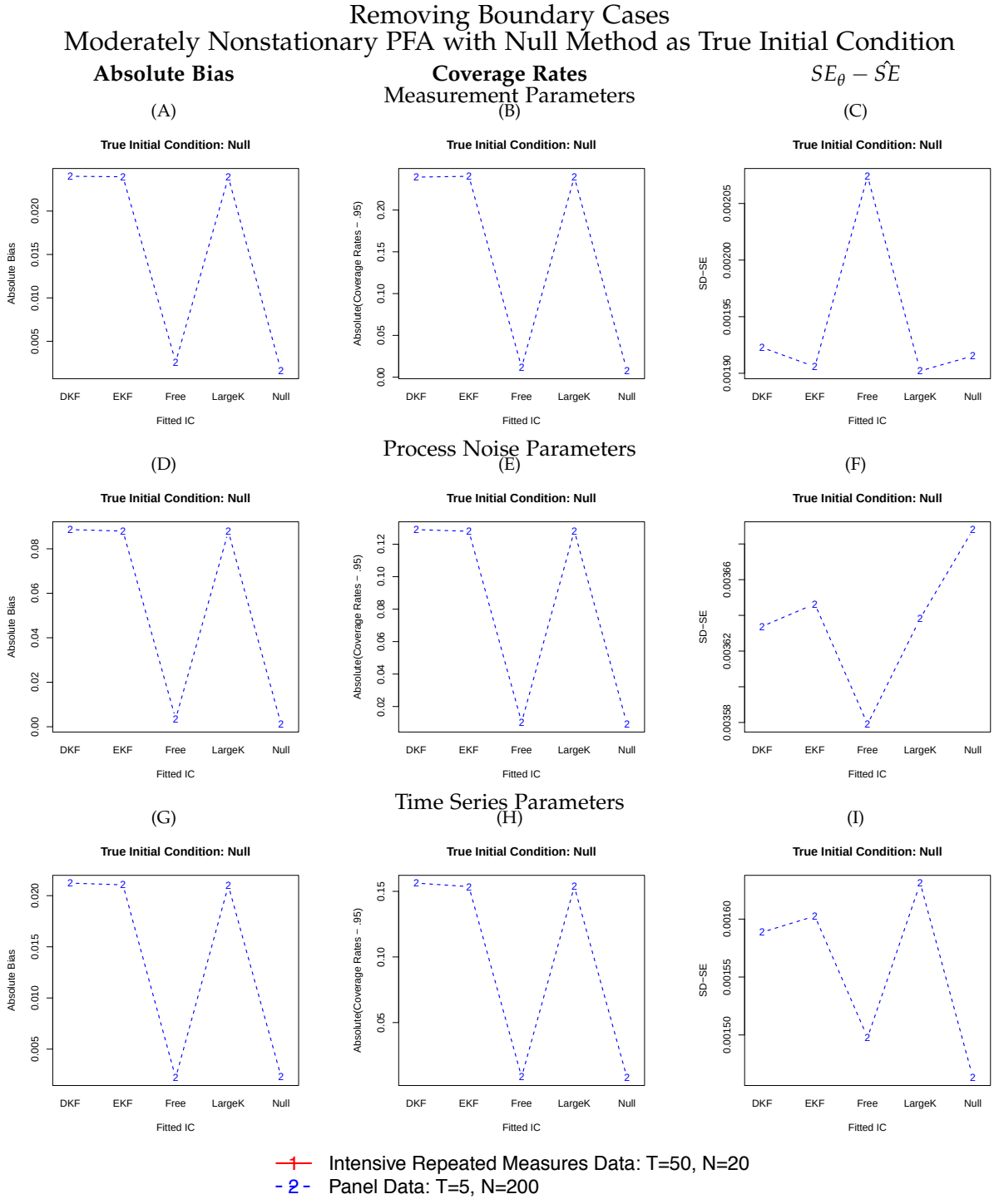

Figure 3.19: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the non-stationary PFA model simulated using the null method as the true initial condition specification with moderately nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

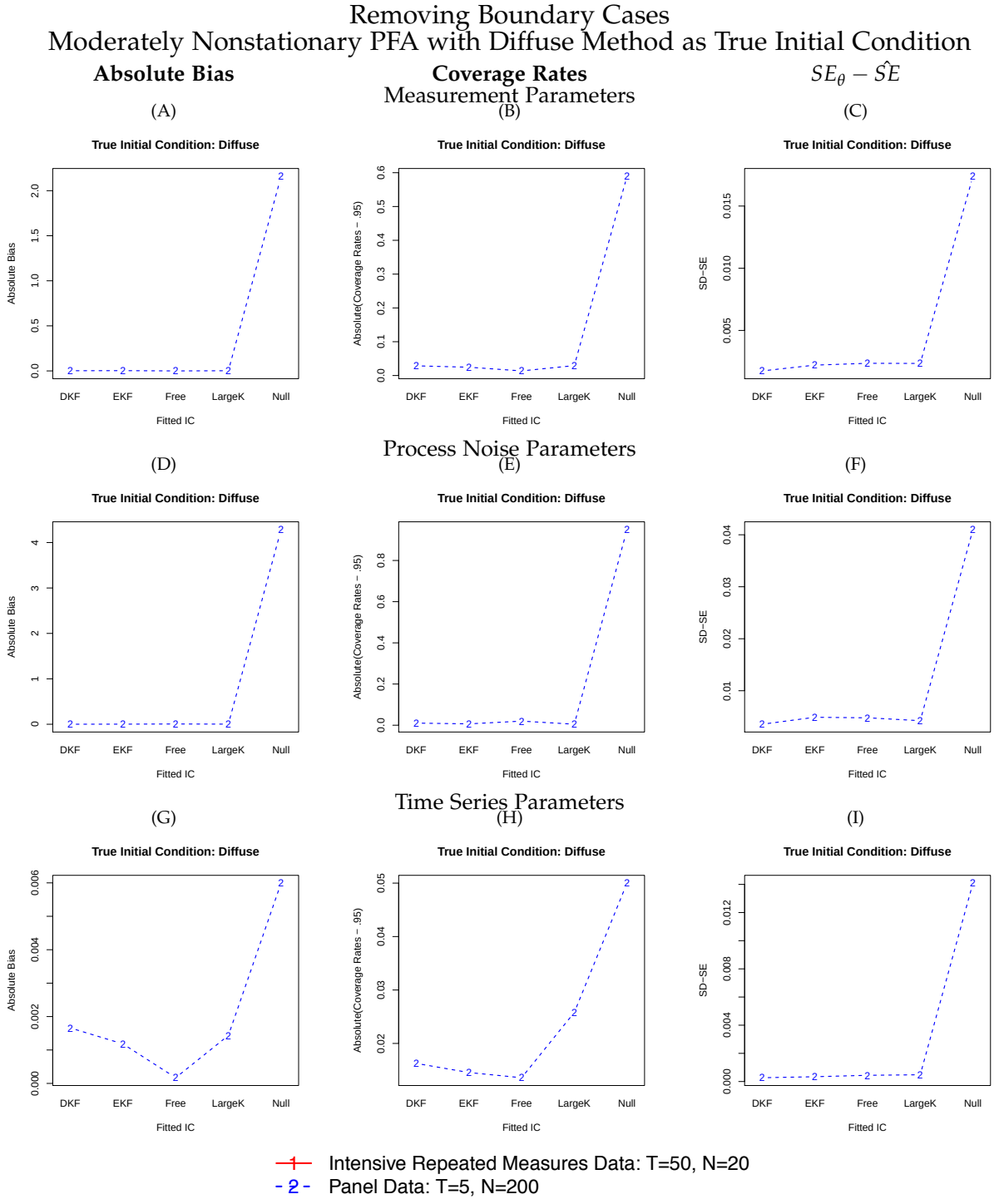

Figure 3.20: Absolute bias, coverage rates, calculated as the absolute difference from .95, absolute empirical standard errors ($\hat{SD}$) minus estimated standard errors (SE_θ), calculated as abs($SE_\theta - \hat{SE}$), for the average of (A)–(B) measurement parameters, (C)–(D) process noise parameters, and (E)–(F) time series parameters for the nonstationary PFA model simulated using the diffuse method as the true initial condition specification with moderately nonstationary population values. DKF=de Jong's diffuse Kalman filter approach, EKF=Koopman's exact initial Kalman filter approach.

3.10 Empirical Example

While the simulation results provide some justification for using a diffuse initial condition, this section illustrates the performance of fitting different initial condition specifications to a real data set. The data used are adopted from the empirical data set used by Du Toit and Browne (2007). From 1974 to 1978, a sample consisting of N=1066 African school children were administered several questionnaires, including the High School Personality Questionnaire (HSPQ; Cattell & Cattell, 1969). This questionnaire seeks to understand dimensions of personality for the purpose of describing individual differences.

Overall, there are fourteen dimensions being tested. Based on results from several exploratory factor analyses, I chose to look more closely at the following dimensions: 1. emotionally stable vs. unstable and easily upset, 2. venturesome and bold vs. shy and timid, 3. conscientious and persevering vs. expedient, 4. outgoing and warmhearted vs. cool and detached, 5. assertive and aggressive vs. obedient and submissive, and 6.lively and enthusiastic vs. sober and serious. The first three dimensions, which serve as manifest variables y_1, y_2, and y_3, loaded onto a factor which I called self-assuredness and the last three dimensions, which serve as manifest variables y_4, y_5, and y_6, loaded onto a factor which I called extraversion. Higher scores on the self-assuredness factor indicated higher levels of self-assuredness while higher scores on the extroversion factor indicated higher levels of extroversion. Summary statistics for the six manifest variables are provided in Table 3.7.

The model estimated is the PFA(1,0) model described in Equation 2.4 in Section 2.1. As the mean values for the manifest variables were all much greater than zero, and the PFA(1,0) model considered does not contain measurement intercepts, the data were de-meaned (i.e., the mean of each observed variable was subtracted from each individuals score for that variable) before estimation so that the observed variables all had a mean of zero. Each initial condition specification, including model-

Time Point	Variable	Mean	Standard Deviation	Skewness	Kurtosis
$T = 1$	y_{11}	9.30	3.17	0.10	-0.02
	y_{21}	10.53	3.38	0.01	-0.05
	y_{31}	7.71	3.16	0.18	-0.40
	y_{41}	9.62	3.21	0.08	-0.04
	y_{51}	12.20	3.24	-0.16	-0.37
	y_{61}	8.80	3.98	0.19	-0.39
$T = 2$	y_{12}	9.56	3.50	0.03	-0.17
	y_{22}	10.80	3.40	-0.09	-0.23
	y_{32}	8.37	3.39	0.18	-0.24
	y_{42}	9.86	3.52	0.11	-0.20
	y_{52}	12.26	3.21	-0.06	-0.27
	y_{62}	8.95	4.26	0.24	-0.51
$T = 3$	y_{13}	9.80	3.61	-0.08	-0.22
	y_{23}	10.93	3.60	0.02	-0.22
	y_{33}	8.79	3.36	0.04	-0.33
	y_{43}	10.07	3.51	0.01	-0.32
	y_{53}	12.21	3.30	-0.04	-0.44
	y_{63}	8.99	4.25	0.19	-0.46
$T = 4$	y_{14}	10.36	3.66	-0.05	-0.32
	y_{24}	11.14	3.51	-0.05	-0.17
	y_{34}	9.01	3.36	0.04	-0.17
	y_{44}	10.26	3.61	0.03	-0.38
	y_{54}	12.49	3.36	-0.23	-0.51
	y_{64}	9.31	4.41	0.10	-0.67
$T = 5$	y_{15}	10.17	3.53	-0.15	-0.30
	y_{25}	11.43	3.42	-0.16	-0.20
	y_{35}	9.23	3.44	-0.01	-0.37
	y_{45}	10.21	3.61	0.01	-0.46
	y_{55}	12.55	3.30	-0.26	-0.14
	y_{65}	9.32	4.33	0.15	-0.46

Table 3.7: Summary Statistics for Manifest Variables of Empirical Example

implied, free-parameter, null, de Jong DKF, Koopman exact initial KF, and large κ, were used in turn to estimate the model using the de-meaned data.

Parameter estimates, associated standard errors, and fit statistics are reported in Table 3.8 and compared across both approaches in Figure 3.21. The fit statistics used include the AIC and BIC, calculated as in the simulation study, and in addition the Root-Mean-Square Error of Approximation (RMSEA; Steiger & Lind, 1980). [1] The model-implied approach failed to converge properly even when given good starting values and relaxing the convergence criteria, thus the estimates are not reported. The de Jong DKF and Koopman exact initial KF produced identical results in terms of parameters estimates, standard errors, fit statistics, and smoothed state estimates to the 7th decimal place, and are therefore grouped together in Figure 3.21.

The large κ approach produced values that were identical to the de Jong and Koopman approaches to the 2nd or 3rd decimal place. However, the log likelihood values, AIC, BIC, and RMSEA as shown in Table 3.8 all indicated that the de Jong and Koopman approaches were best. This is an interesting results because, without other knowledge, a researcher would deem the model with a large κ initial condition as poorly fitting the data, even though the parameter estimates and standard errors

[1] The RMSEA was calculated as follows,

$$\text{RMSEA: } \sqrt{\frac{F_{ML}(N-1)-df}{df(N-1)}}$$

where F_{ML} is the maximum likelihood fit function where the estimated model is compared to a saturated model (Bollen, 1989), df is degrees of freedom calculated as $(((p(p+1))/2)+(ny*nt))-k$ where p is the number of observed variables and k is the number of free parameters being estimated, and N is the total sample size. The log-likelihood values obtained in this manuscript were derived from using full information maximum likelihood estimation, where a saturated model is not calculated. Therefore, the RMSEA was not available for use in the simulation study. However, in the empirical example, the saturated model was calculated as a function of the sample covariance matrix and the maximum likelihood fit function, F_{ML}, was computed as

$$F_{ML} = \frac{-2\log(\boldsymbol{\theta})}{N} - \log(2\pi)(ny*nt) - \log(\det(\text{cov}(S))) - ny*nt \tag{3.7}$$

where S is the sample covariance matrix.

Table 3.8: Parameter Estimates and Standard Errors for PFA(1,0) Empirical Example

	Initial Condition Specification				
Parameter	Null	Free-Parameter	de Jong DKF	Koopman exact initial KF	large κ
Z21	1.34315566 (.03)	1.3857200 (.03)	1.0932273(.02)	1.0932274(.02)	1.0950379(.02)
Z31	0.45139270(.02)	0.4545392 (.02)	0.4103572(.02)	0.4103572(.02)	0.4107510(.02)
Z52	0.88955574(.04)	0.9495546 (.04)	0.5803421(.02)	0.5803421(.02)	0.5822540(.02)
Z62	1.19764514 (.04)	1.2610884 (.04)	0.8129229(.02)	0.8129229(.02)	0.8152876(.02)
V11	2.49249008 (.13)	1.2218748 (.08)	1.7330166(.11)	1.7330165(.11)	1.7281737(.11)
V21	1.08794988 (.08)	0.5297562 (.05)	0.7160510(.07)	0.7160512(.07)	0.7141533(.07)
V22	1.64364074 (.11)	0.4685797 (.06)	0.9547315(.10)	0.9547319(.10)	0.9503385(.10)
T11	0.95612201(.02)	0.9350980 (.01)	0.8516609(.01)	0.8516609(.01)	0.8523320(.01)
T21	-0.02595906(.02)	-0.0239942(.01)	0.0360925(.01)	0.0360925(.01)	0.0361121(.01)
T12	-0.03082725 (.03)	0.0011844 (.02)	0.0263089(.01)	0.0263090(.01)	0.0262464(.01)
T22	0.99170142 (.02)	1.0230739 (.01)	0.8908599(.01)	0.8908599(.01)	0.8917128(.01)
U11	6.89352513 (.16)	5.7343217 (.15)	4.9933912(.16)	4.9933914(.16)	4.9957106(.16)
U22	8.84403143 (.22)	6.0257151 (.22)	7.2919183(.23)	7.2919181(.23)	7.2830298(.23)
U33	9.71798699 (.20)	9.4639671 (.19)	9.2420655(.19)	9.2420655(.19)	9.2429219(.19)
U44	8.63796442 (.21)	8.3367009 (.21)	7.1387172(.24)	7.1387174(.24)	7.1461321(.24)
U55	8.34394523 (.19)	7.6787079 (.18)	8.4980288(.20)	8.4980287(.20)	8.4908715(.20)
U66	7.09426851 (.18)	6.0587904 (.18)	6.9748149(.20)	6.9748144(.20)	6.9648021(.20)
X01	NA	0.0000003 (.08)	NA	NA	NA
X02	NA	-0.0000001(.06)	NA	NA	NA
P011	NA	5.1199089 (.31)	NA	NA	NA
P012	NA	1.8328109 (.17)	NA	NA	NA
P022	NA	2.7515383 (.21)	NA	NA	NA
LL	-82111.13	-80400.53	-75451.28	-75451.28	-84784.65
AIC	164256.3	160845.1	150936.6	150936.6	169603.3
BIC	164340.8	160954.4	151021.1	151021.1	169687.8
RMSEA	0.1329776	0.1053314	0.06367046	0.06367046	0.1678593

are very close to fitted models that indicate good fit. A more detailed examination of the likelihood functions may elucidate this discrepant finding. Reproduced here with the constant term included, the negative two log likelihoods for the de Jong DKF, Koopman exact initial KF, and large κ approaches are,

$$\begin{aligned}
\text{de Jong:} &= (\text{ny} * \text{nt} - \text{nd}) \log(2\pi) + \sum_{i=1}^{N} \sum_{t=1}^{T} \log |\mathbf{F}_{it}| + q_T + \log |\mathbf{S}_T^{-1}| + \mathbf{s}_T' \mathbf{S}_T \mathbf{s}_T \\
\text{Koopman:} &= (\text{ny} * \text{nt} - \text{nd}) \log(2\pi) + \sum_{i=1}^{N} \sum_{t=1}^{T} \log |\mathbf{F}_{*,it}^{-} + \mathbf{F}_{\infty,it}^{-}| + \boldsymbol{\epsilon}_{it}' \mathbf{F}_{*,it}^{-} \boldsymbol{\epsilon}_{it} \\
\text{Large } \kappa\text{:} &= \text{ny} * \text{nt} \log(2\pi) + \text{nd} \log(\kappa) + \log |\mathbf{F}_{it}| + \sum_{i=1}^{N} \sum_{t=1}^{T} \boldsymbol{\epsilon}_{it}' \mathbf{F}_{it}^{-1} \boldsymbol{\epsilon}_{it}
\end{aligned} \tag{3.8}$$

where nd equals the number of diffuse elements. Note that nd appears in the constant term for all three and $\log(\kappa)$ appears in the constant term for the large κ approach. As nd is subtracted from the constant term in the de Jong and Koopman likelihoods, it serves to make the likelihood value smaller in magnitude which will make this model appear to be fitting the data more optimally. Also, the $\log(\kappa)$ is being added to the constant term for the large κ likelihood, which makes this likelihood larger in magnitude and thus less preferred. This may be why the de Jong and Koopman approaches were chosen so frequently in the simulation study and also why they are chosen here over the other methods, even when the large κ method provides very similar point and standard error estimates.

The free-parameter and null conditions yielded similar parameter estimates for the factor loading parameters and time series parameters, as indicated in Figure 3.21A. However, the free-parameter approach yielded process noise parameter estimates that more closely resembled those from the de Jong, Koopman, and large κ approaches. For all methods, the standard errors for the fixed effect parameters are smaller than those for the variance-covariance parameters, as illustrated when comparing 3.21B and 3.21D.

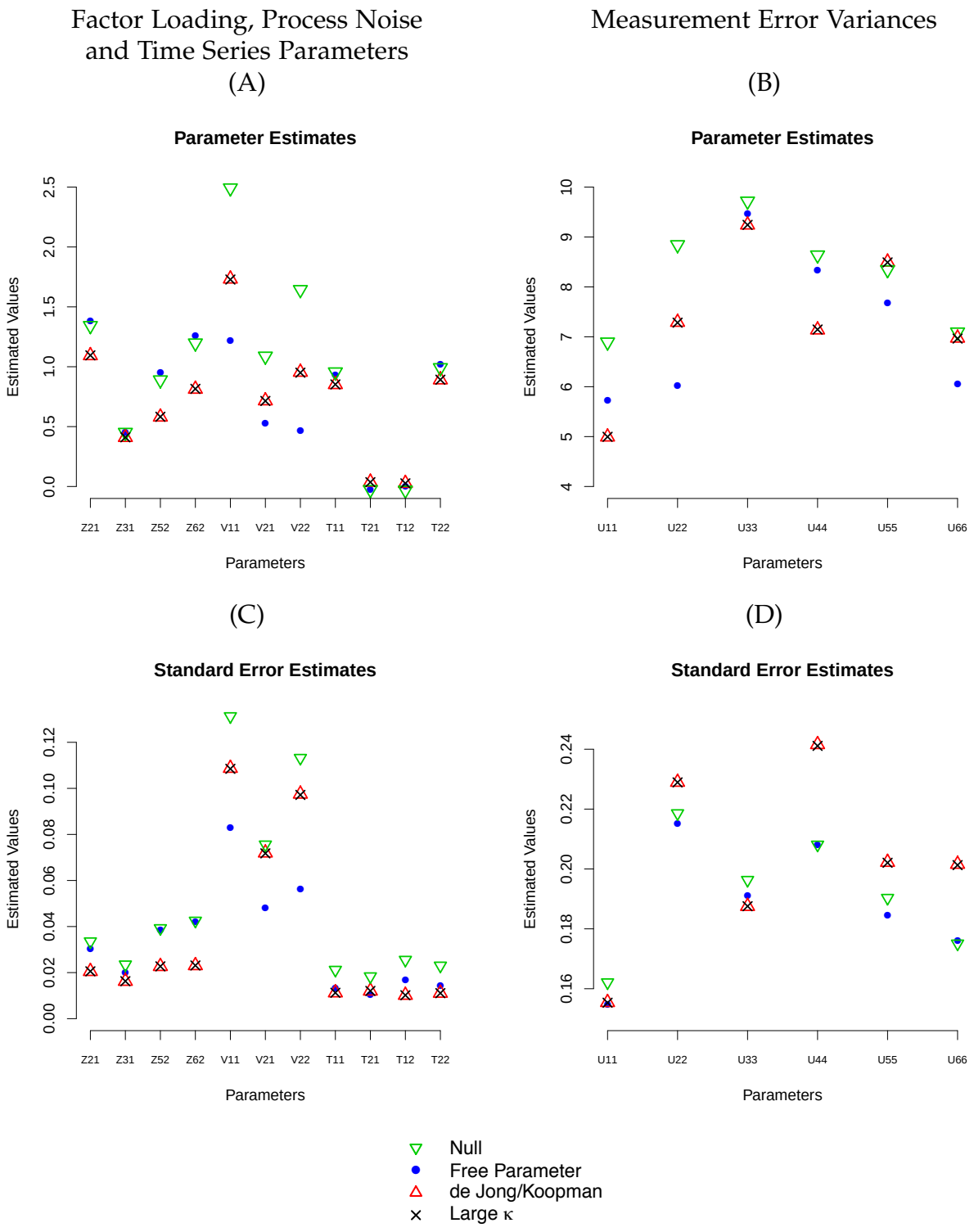

Figure 3.21: Comparison of parameter estimates (A) and standard errors (B) for the factor loading, process noise, and time series parameters, and parameter estimates (C) and standard errors (D) for the measurement error variances, produced from the null, free parameter, de Jong, Koopman, and large κ approaches when applied to the empirical data set. Because the de Jong and Koopman approaches produced identical estimates to the 7th decimal place, they are grouped together

The free-parameter and null approaches estimated the parameters in the transition matrix, i.e., the time series parameters, as being nonstationary. This is verified by calculating the roots of the determinant of Equation 1.7, which were 0.99 and 1.06 for the null approach and 0.98 and 1.07 for the free-parameter approach. As at least one root from each specification has a modulus less than 1, these transition matrices are nonstationary. The diffuse approaches, however, including the de Jong, Koopman and large κ approaches, estimated the time series parameters as being stationary. This is verified as the roots of the determinant of Equation 1.7 are and 1.10 and 1.20, and both have moduli greater than 1.

The cross factor regressions were low in magnitude in the null and free-parameter model, meaning the extroversion factor from the current time point did not have much influence in predicting the self-assuredness factor at the next time point and vice-versa. From a substantive viewpoint, it is interesting that the cross-regression parameters, i.e., $T12$ and $T21$, are significantly different from zero when the de Jong, Koopman, and large κ approaches are implemented but not when the null and free-parameter approaches are implemented.

The null condition estimated larger values for the process noise variables than the other initial condition approaches. The de Jong and Koopman method estimated the process noise parameters as having a larger variance than the free-parameter approach. The free-parameter condition estimated the initial condition matrix with the following values,

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{P}_1 = \begin{bmatrix} 5.12 & 1.83 \\ 1.83 & 2.75 \end{bmatrix}. \tag{3.9}$$

It is also of interest to compare the estimated latent variable scores, or smoothed estimates, obtained from all approaches. Figure 3.22 illustrates the estimated latent variable scores for all approaches, with the de Jong and Koopman approaches again

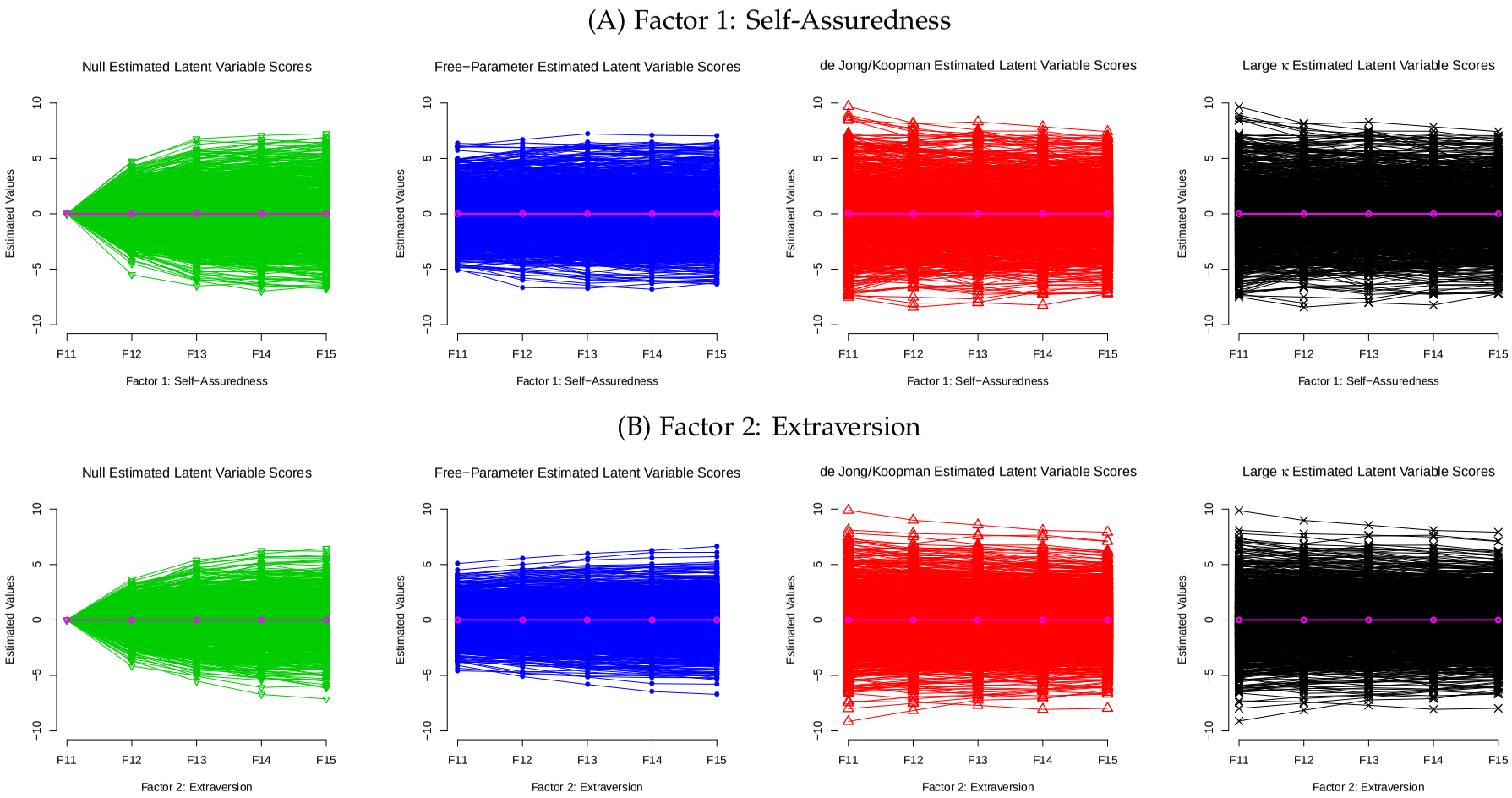

Figure 3.22: Comparison of estimated latent variable scores for the null, free-parameter, de Jong and Koopman, and large κ approaches when applied to the empirical data set for the first factor (A) and second factor (B) from Time 1 to Time 5. For each individual plot, time is on the x-axis and ranges across the five factors. The thick purple line represents the mean estimated latent variable score for that time point.

grouped together. The self-assuredness factor is displayed in the top panel (A) and the extroversion factor in the bottom panel (B). From this plot it is clear that the latent variable scores diverge most at the first time point and converge onto each other as time increases. Also interesting is the how the latent variable scores are clearly restricted at the first time point when the null approach is implemented. The free-parameter approach produced estimated latent variable scores at the first time point between -5 and 5, which coincides with the estimated initial condition mean vector and covariance matrix. The diffuse approaches, however, produced estimated latent variable scores ranging from -10 to 10. Overall, this plot nicely illustrates the idea that the different initial condition approaches will provide different scores at the first time point, but as time increases will provide scores that are more similar to each other. Furthermore, the plot verifies that the diffuse approaches are stationary while the non-diffuse approaches display a nonstationary process. This may be because the diffuse approaches allow for more variability in latent variable scores at the first time point.

CHAPTER 4

Discussion

A specific goal of this thesis was to evaluate two approaches from the time series literature, the de Jong DKF and Koopman exact initial KF, for specifying initial condition in both stationary and nonstationary dynamic factor analysis models estimated in the structural equation modeling framework as equivalent state-space models. Motivation for this research included the fact that data arising from psychological processes has a great potential to be nonstationary and a misspecification of initial condition may be more severe when the number of time points is small, such as in panel data that is frequently collected in psychological studies. Results from this thesis revealed conditions under which using the proper diffuse initial condition specifications may be superior to both an improper diffuse approach and other known stationary approaches. Furthermore, results from this thesis suggest that using the de Jong DKF approach led to fewer numerical difficulties in certain cases than the Koopman exact initial KF approach when estimating PFA(1,0) models with both panel data and intensive repeated measures data.

Given the large number of results presented, it is important to highlight and discuss both interesting and unexpected findings. Overall, results from general model performance emphasized the importance of understanding a model's potential to converge under certain data conditions. Counter to my hypothesis, the model-implied approach did not perform as well as the free parameter approach, even

when the model-implied approach served as the true initial condition specification. One reason for the poor performance of the model-implied approach may be that stationary-constraints were not imposed in the estimation process. Thus, the estimates of the time series parameters may have wandered into an outlying range, which in turn would affect the other estimates. Given this, it is important to complete a follow up simulation study placing stationary constraints on the AR parameters. Also, it would be helpful to simulate models that require less complicated constraints in the covariance initial condition matrix and see if the convergence problems diminish.

Another estimation issue concerned replications that failed when inverting the Hessian matrix, so proceeded to use a generalized inverse. This led to some replications producing parameter estimates that hit a boundary condition for the variance of the process noise or measurement error parameters. Thus, some of these cases did not appear to be representative of the majority of cases. Results with and without cases that hit a boundary condition were presented and compared. While the free-parameter, Koopman exact initial KF, and large κ results significantly improved with boundary cases removed, this does not necessarily indicate that they performed just as well as the de Jong DKF approach. The large discrepancy in cases retained between the de Jong approach and all other approaches suggests that the de Jong approach is better able to estimate the initial condition distribution even when it is diffuse. Furthermore, a sever limitation with the free-parameter approach is that it sometimes soaks up too much variance in the initial condition distribution and subsequently yields process noise variance estimates and measurement error variance estimates that are very close to zero.

Another interesting result was the exceptionally poor performance of the de Jong DKF approach when fitted to a process initialized with a null initial condition specification, particularly in the mildly nonstationary condition using panel data. As

indicated in the results, this may be because there is not enough variability to form accurate maximum likelihood estimates of the diffuse constructs.

Counter to my hypothesis, the de Jong and Koopman approaches did not perform equally well. Across most simulation conditions, the de Jong approach outperformed the Koopman approach on most simulation outcome measures. This is surprising as both approaches should produce identical results. However, the empirical results provide preliminary evidence that the de Jong approach is robust to different starting values, and this may be a reason why it converged more frequently and provided more results that did not produce outlying parameter estimates. Furthermore, when cases hitting a boundary condition were removed, the two approaches produced very similar results. Thus, it may be that the Koopman approach is not as flexible in estimation and resorts to taking a generalized inverse of the Hessian matrix, thereby producing cases that his a boundary condition.

Another interesting result was that the de Jong DKF approach produced very good standard error consistency across most conditions, especially for the process noise parameters. In fact, fitting the de Jong approach resulted in very good point and standard errors estimates for the process noise parameters under all conditions except when the true initial condition was null and panel data was used.

Results concerning the AIC and BIC were particularly interesting and highlight the importance of the constant value in the likelihood function. Even when the de Jong and Koopman approaches did not perform as well with respect to estimation results, the AIC and BIC still chose these models a large percentage of the time, especially in the panel data case. This may be due to the fact that the number of diffuse elements is subtracted out of the constant, thus making the likelihood smaller especially in the panel data case. Also, since $\log(\kappa)$ appears in the likelihood function, this approach was not chosen most of the time.

In line with the goals of this simulation study, it is important to make direct com-

parisons among the initial condition specifications. There were many circumstances under which the free parameter approach yielded the most optimal performance, such as when the process was stationary. The de Jong approach yielded the most optimal results under certain conditions as well, such as when treating the true initial condition as diffuse, especially in the intensive repeated measures condition and in the recover of the process noise parameters.

Because fitting both the de Jong and free parameter approaches produced satisfactory results under different simulation conditions, it is important to understand circumstances under which each is preferred. The results suggest that, as the degree of explosiveness increases (i.e., the process has a greater degree of nonstationarity), the de Jong DKF is more likely to outperform the free parameter approach. Also, as the number of time points increases, the de Jong DKF is more likely to outperform the free parameter approach. Furthermore, developing a sense under which 1. the degree of explosiveness starts to favor using de Jong DKF over free parameter (as the process has a greater degree of nonstationarity), 2. the number of time points used starts to favor using the de Jong DKF over free parameter (as T increases). It may be useful to run a simulation where $T = 14$, which may correspond to a two week period as is commonly used in daily dairy studies (Bolger, Davis, Olchowski, & Rafaeli, 2003).

While a main focus was on comparing fitted initial condition specifications, it is also of interest to discuss differences between data types. For stationary models, the intensive repeated measures data condition was associated with very good simulation results across all true and fitted initial conditions. One reason for this is that as the number of time points increases, the choice of initial condition specification will become less important. As discussed in Section 1.3, the KF becomes independent of the initial state condition after a sufficient number of time points (see Jazwinski, 1970, pp. 239-243; Oud et al., 1990). This means that, as the number of time points

increases, the importance of a correctly specified initial condition is diminished. This makes sense analytically, as when the number of time points is small the diffuse part still overwhelms the likelihood value, affecting two out of five time points instead of two out of 50. Thus, it make sense that, all else held constant, the models estimated under a stationary process, regardless of true or fitted initial condition, would perform well with intensive repeated measures data. With mild nonstationarity the fitted initial condition specifications still performed well with intensive repeated measures. Because the SEM framework is more suited for panel type data, perhaps estimating the free parameter, null, and model-implied approaches in software designed for the SEM framework, such as Mplus, may have produced more optimal results for these fitted initial conditions.

The empirical example illustrated the notion that different initial condition specifications may lead to different point and standard error estimates, and thus different conclusions. This was particularly salient by the fact that the cross-regression parameters were found to be significant when using the diffuse approaches, but not when using the non-diffuse approaches. This is an important distinction as many psychological hypotheses may concern whether or not a given construct significantly predicts future performance on another construct.

4.1 Recommendations for Practice

To make specific recommendations for practice, it is important to consider the circumstances under which the true initial condition specifications used in this manuscript would arise. Therefore, I will next discuss circumstances and constructs in psychological research that would give rise to each of the true initial condition specifications considered and provide suggestions for fitted initial condition specifications based on the results from this thesis.

Three true initial conditions were implemented when the process is stationary: 1. model-implied, 2. free-parameter, and 3. null. Overall, results suggest that if the

true process is indeed stationary, choosing one of the stationary approaches is best. A true initial condition resulting from using a model-implied approach results in a stationary process that has always maintained its stationarity. This may be viable for psychological processes that ebb and flow with time, such as emotions or daily affect (see, for example, Diener, Fujita, & Smith, 1995; Shifren, Hooker, Wood, & Nesselroade, 1997) sampled on a daily basis. A model-implied approach in this case would appear to be the best choice, however this study did not provide supporting evidence as the model-implied approach proved to have severe convergence issues.

Generating a process using the free parameter approach is akin to generating a process where there is only a change in the first occasion, in line with the free-parameter approach discussed by Du Toit and Browne (2007). When the process is stationary, this means that the process before the first occasion was different than the overall process; however, from the first time point onward the process does not display mean trends or changes in variability. This type of process is more likely to arise in psychological data than the null initial condition, although the assumption of stationarity still holds across all time points other than the first. When this is the case, the fitted free-parameter initial condition performs best as it captures the mean and covariance of the process at the first time point, and also the stationary process onward. However, it is important to carefully consider whether this assumption is valid given the type of psychological construct being studied. Still, applying such a specification will provide information regarding the process before the first time point in the form of mean and variance-covariance parameter estimates. There is a cost to this, however, as more degrees of freedom are used. In line with the model selection literature, and a more parsimonious model may be preferred (Preacher, 2006, 2000).

In the stationary case, a true initial condition consisting of null matrices would arise when the process starts exactly at the first time point. Given the types of con-

structs estimated in psychological studies, such as self-assuredness and extroversion in the empirical example considered in this manuscript, this seems like an untenable assumption. Given that the fitted free-parameter method can estimate the value in the initial condition matrices, it seems wise to use this approach over a fitted null approach unless a researcher is certain that the process starts at the first time point, in which case the null model would be a more parsimonious model. The fitted diffuse approaches performed especially poorly in this case, so it would be wise to avoid using them here.

If the process is nonstationary, this thesis considered three true initial condition specifications: 1. free-parameter, 2. null, and 3. diffuse. Like in the nonstationary case, the free-parameter condition indicates that there was indeed a change in process at the first time point, but in this case the process after the first time point is nonstationary. A process like this may occur in an experimental design after a treatment in enacted. Specifically, the first time point serves as a person's process prior to receiving a treatment, and time points thereafter serve as a person's process after receiving a treatment. While the fitted free-parameter approach works well when there are no convergence issues, it may not be the best choice if the process is highly nonstationary, as the variance may be consumed fully by the variance and covariance parameters of the initial condition matrices leading to potential convergence problems, boundary conditions, and parameter estimate outliers.

In the nonstationary case, a true initial condition consisting of null matrices would arise when the process starts exactly at the first time point, but displays a nonstationary process onward. Given psychological constructs, this seems like an untenable assumption. However, there may be cases in experimental designs where, prior to treatment, every person has no proficiency in a given construct. Still, being that the free-parameter approach estimates the process before the first time point, it may be safer to just apply this approach, especially since simulation results suggested

this approach performs well if the true initial condition is null. As in the stationary case, the fitted diffuse approaches did not perform very well when the true initial condition was null.

The final true initial condition considered was that of a diffuse process. In this case, the process began in the distant past, there is no prior information regarding the process before the first time point, and the process after the first time point remains nonstationary. This situation may occur often in psychological processes, as people are constantly changing in complicated ways with respect to many psychological constructs such as personality development, intelligence, and severity of clinical disorders. It is rarely known what a given person's trajectory was before the first time point was collected. It is also unlikely that a person began the process right at the first time point. As the process started in the distant past and there is no prior information about the process, fitting a null initial condition is not a good idea, as confirmed by the simulation results. Fitting the free-parameter approach may work well as it is able to capture some aspects of the process prior to the first observation, but again there are issues of convergence, boundary conditions, and extreme outlying parameter estimates. The de Jong diffuse initial condition specification worked well in this case in terms of both recover point and standard error estimates and not displaying large convergence, boundary, or outlier problems. The other diffuse approaches did not work as well, mainly due to boundary cases and convergence issues. Thus, using the fitted de Jong approach would appear to be the best choice, followed by the free-parameter condition.

4.2 Limitations and Future Directions

Although the results presented in this thesis provide a promising indication that using a diffuse initial condition specification is beneficial when the process is actually diffuse, several limitations are apparent. As with all simulation studies, the findings here are limited to the simulation conditions chosen and may not generalize

to other types of data or models. The conclusions made in this thesis are restricted by the choice of the true initial condition specifications, current modeling choices, and true parameter values considered. The performance of different initial condition specifications needs to be examined in the presence of methodological concerns that commonly arise in psychological data, such as the presence of missing data (both ignorable and non-ignorable) or the use of categorical outcomes (e.g., ordinal, binary, or count data). Also, other forms of misspecification, such as differing degrees of explosiveness and different true parameters values in the fixed initial condition matrices. Furthermore, intercepts were not included in either models nor was a moving average component considered.

The presence of missing data would be an interesting one to consider because, when using the de Jong DKF approach, the number of time points before a collapse to the regular KF may increase with missing data (Harvey, 1991; Doornik, 2007). Thus, the importance of a correctly specified initial condition is magnified as the diffuse part of the likelihood will larger.

The number of replications per condition was 500, which is not very large considering the relatively large number of cases that either did not converge to a proper solution or produced excessive outliers in the parameter estimates. A follow-up simulation study using upwards of 5,000 replications would reveal a more accurate percentage rate of solutions that converged properly and solutions that did not produce extreme outlying parameter estimates. Furthermore, more replications would be available to make more conclusive statements regarding simulation outcome measures such as bias and MSE.

It would also be helpful to analytically determine why the Koopman exact initial KF approach broke down in most simulation conditions. This was certainly an unexpected result and determining the conditions under which both the de Jong and Koopman approaches are likely to fail is important. For example, from this simula-

tion study we can see that the de Jong approach does not converge as often when the a true null initial condition specification is applied and the process is not extremely explosive. Determining the exact analytical conditions under which both approaches fail would be helpful for empirical researchers.

One future consideration is to investigate the merits of and implement a marginal log-likelihood suggested by Francke, Koopman, and deVos (2010). The authors present a modification to the DKF that allows for the marginal log-likelihood to be utilized, which may work better with some state space models. However, it is not clear if this approach would work well when considering panel data. It would also be interesting to consider other types of models, such as models that are partially stationary, models with a time-varying trend, models with covariates, and models with different trajectory shapes. In psychological research there may be processes with both stationary and nonstationary underlying latent processes. For example, daily diary or intensive repeated measures data of positive emotions may produce a process that ebbs and flows around a set point, while ultimately remaining stationary. However, there may also be underlying processes that display a time trend, such as a person's anxiety level over a time period where he or she is attending therapy. Harvey (1991) described a procedure for implementing models where some states are regarded as stationary and some regarded as nonstationary.

These models were estimated using Ox and SSfPack. Given that many psychologists are unfamiliar with these packages and formulating models as state space models in general, it is important to make these statistical methods accessible to empirical researchers. Thus, a future direction for this research is to create R code that automates the process of fitting a model using a given diffuse initial condition specification. Still, this may be difficult as the current version of SSfPack with the canned procedures is expensive. Thus, I will continue to work on the SAS proc IML code and will consider writing the code in R using an existing optimization procedure,

perhaps with openMX.

Despite the above mentioned limitations, the results from this thesis provide promising evidence that using the de Jong DKF may be superior to other methods when the process is truly diffuse. Still, there is much work to be done in determining the exact conditions under which such results will hold true. Overall, the research completed here is a small step forward in the direction of allowing for the feasible estimation of time series models in the SEM framework in the hopes that researchers in the psychological sciences will more readily apply such models.

CHAPTER 5

Appendix A

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.205	0.051	0.048	0.005	0.004	0.052	0.946	1
Z31	0.8	0.801	0.049	0.048	0.001	0.002	0.049	0.938	1
Z52	0.9	0.9	0.04	0.04	0	0	0.04	0.958	1
Z62	1.1	1.1	0.038	0.037	0	0	0.038	0.946	1
V11	1	1.002	0.107	0.077	0.002	0.002	0.106	0.943	1
V21	0.4	0.403	0.073	0.045	0.003	0.007	0.073	0.946	1
V22	1	1	0.086	0.075	0	0	0.086	0.943	1
T11	0.5	0.496	0.058	0.033	-0.004	-0.008	0.058	0.941	1
T21	-0.3	-0.306	0.048	0.034	-0.006	0.019	0.049	0.931	1
T12	-0.1	-0.102	0.046	0.028	-0.002	0.022	0.046	0.938	0.946
T22	0.6	0.599	0.047	0.029	-0.001	-0.001	0.047	0.951	0.998
U11	0.8	0.801	0.056	0.054	0.001	0.001	0.056	0.953	1
U22	0.6	0.594	0.068	0.062	-0.006	-0.009	0.068	0.941	0.998
U33	2	1.986	0.095	0.096	-0.014	-0.007	0.096	0.96	1
U44	1	0.996	0.053	0.058	-0.004	-0.004	0.053	0.96	1
U55	1.5	1.495	0.076	0.075	-0.005	-0.003	0.076	0.938	1
U66	0.4	0.399	0.045	0.045	-0.001	-0.003	0.045	0.943	1

Table 5.1: Stationary PFA Results: True IC: Model Implied, Fitted IC: Model Implied, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.202	0.049	0.048	0.002	0.002	0.049	0.947	1
Z31	0.8	0.8	0.048	0.048	0	0	0.048	0.939	1
Z52	0.9	0.9	0.041	0.04	0	0	0.041	0.951	1
Z62	1.1	1.1	0.038	0.037	0	0	0.038	0.949	1
V11	1	1	0.077	0.077	0	0	0.077	0.951	1
V21	0.4	0.399	0.046	0.045	-0.001	-0.002	0.046	0.958	1
V22	1	0.997	0.074	0.075	-0.003	-0.003	0.074	0.941	1
T11	0.5	0.497	0.032	0.033	-0.003	-0.005	0.032	0.947	1
T21	-0.3	-0.304	0.034	0.034	-0.004	0.012	0.034	0.943	1
T12	-0.1	-0.101	0.03	0.029	-0.001	0.015	0.03	0.941	0.932
T22	0.6	0.598	0.029	0.029	-0.002	-0.004	0.029	0.945	1
U11	0.8	0.798	0.055	0.053	-0.002	-0.003	0.055	0.953	1
U22	0.6	0.597	0.062	0.062	-0.003	-0.004	0.062	0.945	1
U33	2	1.992	0.096	0.096	-0.008	-0.004	0.096	0.96	1
U44	1	0.996	0.054	0.058	-0.004	-0.004	0.054	0.958	1
U55	1.5	1.496	0.076	0.075	-0.004	-0.003	0.076	0.939	1
U66	0.4	0.398	0.045	0.045	-0.002	-0.005	0.045	0.943	1
X01	-	0.004	0.282	0.27					
X02	-	-0.004	0.655	0.378					
P011	-	1.274	0.499	0.48					
P012	-	0.133	0.785	0.535					
P022	-	1.591	0.586	1.161					

Table 5.2: Stationary PFA Results: True IC: Model Implied, Fitted IC: Free Parameter, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.171	0.044	0.046	-0.029	-0.024	0.053	0.891	1
Z31	0.8	0.803	0.049	0.049	0.003	0.004	0.049	0.937	1
Z52	0.9	0.9	0.041	0.041	0	0	0.041	0.962	1
Z62	1.1	1.063	0.035	0.036	-0.037	-0.033	0.051	0.828	1
V11	1	1.005	0.078	0.078	0.005	0.005	0.078	0.95	1
V21	0.4	0.406	0.047	0.047	0.006	0.014	0.048	0.956	1
V22	1	1.005	0.075	0.078	0.005	0.005	0.076	0.958	1
T11	0.5	0.505	0.034	0.034	0.005	0.009	0.034	0.937	1
T21	-0.3	-0.308	0.035	0.035	-0.008	0.027	0.036	0.947	1
T12	-0.1	-0.101	0.03	0.029	-0.001	0.011	0.03	0.935	0.935
T22	0.6	0.609	0.029	0.029	0.009	0.015	0.03	0.943	1
U11	0.8	0.814	0.054	0.052	0.014	0.017	0.056	0.941	1
U22	0.6	0.719	0.063	0.056	0.119	0.198	0.134	0.452	1
U33	2	1.996	0.097	0.097	-0.004	-0.002	0.097	0.95	1
U44	1	1.006	0.056	0.058	0.006	0.006	0.056	0.966	1
U55	1.5	1.502	0.077	0.077	0.002	0.001	0.077	0.939	1
U66	0.4	0.54	0.051	0.041	0.14	0.351	0.149	0.111	1

Table 5.3: Stationary PFA Results: True IC: Model Implied, Fitted IC: Null, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.177	0.066	0.047	-0.023	-0.019	0.069	0.9	0.998
Z31	0.8	0.786	0.06	0.047	-0.014	-0.017	0.061	0.93	0.998
Z52	0.9	0.887	0.04	0.039	-0.013	-0.015	0.042	0.952	1
Z62	1.1	1.082	0.036	0.036	-0.018	-0.016	0.04	0.914	1
V11	1	1.021	0.09	0.077	0.021	0.021	0.093	0.954	0.998
V21	0.4	0.408	0.047	0.047	0.008	0.02	0.047	0.956	0.998
V22	1	1.021	0.093	0.076	0.021	0.021	0.095	0.942	1
T11	0.5	0.496	0.034	0.034	-0.004	-0.007	0.034	0.946	0.998
T21	-0.3	-0.295	0.172	0.035	0.005	-0.018	0.172	0.942	1
T12	-0.1	-0.101	0.03	0.029	-0.001	0.007	0.03	0.94	0.932
T22	0.6	0.596	0.042	0.029	-0.004	-0.007	0.042	0.946	0.998
U11	0.8	0.79	0.087	0.053	-0.01	-0.013	0.088	0.94	1
U22	0.6	0.615	0.107	0.061	0.015	0.025	0.108	0.956	1
U33	2	1.993	0.101	0.096	-0.007	-0.004	0.102	0.96	1
U44	1	0.989	0.054	0.058	-0.011	-0.011	0.055	0.95	1
U55	1.5	1.495	0.077	0.075	-0.005	-0.004	0.077	0.938	1
U66	0.4	0.406	0.044	0.044	0.006	0.015	0.045	0.942	1

Table 5.4: Stationary PFA Results: True IC: Model Implied, Fitted IC: deJong DKF, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.179	0.046	0.046	-0.021	-0.018	0.051	0.899	1
Z31	0.8	0.788	0.047	0.047	-0.012	-0.015	0.049	0.93	1
Z52	0.9	0.887	0.039	0.039	-0.013	-0.014	0.041	0.954	1
Z62	1.1	1.083	0.036	0.036	-0.017	-0.016	0.04	0.916	1
V11	1	1.023	0.077	0.077	0.023	0.023	0.081	0.962	1
V21	0.4	0.408	0.046	0.046	0.008	0.021	0.047	0.956	1
V22	1	1.019	0.073	0.076	0.019	0.019	0.076	0.947	1
T11	0.5	0.496	0.033	0.033	-0.004	-0.008	0.033	0.947	1
T21	-0.3	-0.303	0.034	0.034	-0.003	0.008	0.034	0.943	1
T12	-0.1	-0.101	0.03	0.029	-0.001	0.01	0.03	0.939	0.932
T22	0.6	0.597	0.029	0.029	-0.003	-0.005	0.029	0.945	1
U11	0.8	0.786	0.054	0.053	-0.014	-0.017	0.056	0.941	1
U22	0.6	0.612	0.062	0.061	0.012	0.02	0.063	0.956	1
U33	2	1.991	0.096	0.096	-0.009	-0.005	0.096	0.96	1
U44	1	0.988	0.054	0.058	-0.012	-0.012	0.055	0.949	1
U55	1.5	1.495	0.076	0.075	-0.005	-0.003	0.076	0.943	1
U66	0.4	0.406	0.044	0.044	0.006	0.015	0.045	0.941	1

Table 5.5: Stationary PFA Results: True IC: Model Implied, Fitted IC: Koopman exact initial KF, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.179	0.046	0.046	-0.021	-0.018	0.051	0.899	1
Z31	0.8	0.788	0.047	0.047	-0.012	-0.015	0.049	0.93	1
Z52	0.9	0.887	0.039	0.039	-0.013	-0.014	0.041	0.954	1
Z62	1.1	1.083	0.036	0.036	-0.017	-0.016	0.04	0.916	1
V11	1	1.023	0.077	0.077	0.023	0.023	0.081	0.962	1
V21	0.4	0.408	0.046	0.046	0.008	0.021	0.047	0.956	1
V22	1	1.019	0.073	0.076	0.019	0.019	0.076	0.947	1
T11	0.5	0.496	0.033	0.033	-0.004	-0.008	0.033	0.947	1
T21	-0.3	-0.303	0.034	0.034	-0.003	0.008	0.034	0.943	1
T12	-0.1	-0.101	0.03	0.029	-0.001	0.01	0.03	0.939	0.932
T22	0.6	0.597	0.029	0.029	-0.003	-0.005	0.029	0.945	1
U11	0.8	0.786	0.054	0.053	-0.014	-0.017	0.056	0.941	1
U22	0.6	0.612	0.062	0.061	0.012	0.02	0.063	0.956	1
U33	2	1.991	0.096	0.096	-0.009	-0.005	0.096	0.96	1
U44	1	0.988	0.054	0.058	-0.012	-0.012	0.055	0.949	1
U55	1.5	1.495	0.076	0.075	-0.005	-0.003	0.076	0.943	1
U66	0.4	0.406	0.044	0.044	0.006	0.015	0.045	0.941	1

Table 5.6: Stationary PFA Results: True IC: Model Implied, Fitted IC: Large κ, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.168	0.039	0.046	-0.032	-0.027	0.05	0.944	1
Z31	0.8	0.815	0.042	0.049	0.015	0.018	0.044	1	1
Z52	0.9	0.909	0.045	0.041	0.009	0.01	0.045	0.944	1
Z62	1.1	1.072	0.042	0.036	-0.028	-0.025	0.05	0.861	1
V11	1	1.061	0.435	0.103	0.061	0.061	0.433	0.667	1
V21	0.4	0.61	0.345	0.064	0.21	0.526	0.4	0.472	1
V22	1	1.028	0.216	0.097	0.028	0.028	0.215	0.833	1
T11	0.5	0.661	0.515	0.036	0.161	0.322	0.533	0.333	1
T21	-0.3	-0.361	0.414	0.043	-0.061	0.204	0.413	0.389	0.917
T12	-0.1	-0.088	0.373	0.038	0.012	-0.116	0.368	0.361	0.972
T22	0.6	0.853	0.374	0.033	0.253	0.422	0.447	0.417	1
U11	0.8	0.846	0.077	0.056	0.046	0.058	0.089	0.778	1
U22	0.6	0.757	0.122	0.066	0.157	0.261	0.197	0.417	1
U33	2	2.016	0.111	0.098	0.016	0.008	0.111	0.944	1
U44	1	1.008	0.053	0.059	0.008	0.008	0.053	0.972	1
U55	1.5	1.505	0.078	0.077	0.005	0.003	0.077	1	1
U66	0.4	0.512	0.095	0.047	0.112	0.281	0.146	0.389	1

Table 5.7: Simulation Results: True IC: Model Implied, Fitted IC: Model Implied, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.197	0.051	0.049	-0.003	-0.003	0.051	0.949	1
Z31	0.8	0.8	0.049	0.048	0	0	0.049	0.958	1
Z52	0.9	0.906	0.09	0.041	0.006	0.007	0.09	0.941	1
Z62	1.1	1.1	0.049	0.038	0	0	0.049	0.927	1
V11	1	1.003	0.08	0.082	0.003	0.003	0.08	0.954	1
V21	0.4	0.401	0.054	0.05	0.001	0.002	0.054	0.936	1
V22	1	0.994	0.105	0.081	-0.006	-0.006	0.105	0.927	0.998
T11	0.5	0.5	0.037	0.037	0	-0.001	0.037	0.956	1
T21	-0.3	-0.302	0.042	0.038	-0.002	0.006	0.042	0.921	1
T12	-0.1	-0.101	0.032	0.032	-0.001	0.014	0.032	0.954	0.894
T22	0.6	0.6	0.033	0.032	0	0.001	0.033	0.947	1
U11	0.8	0.796	0.061	0.055	-0.004	-0.005	0.061	0.927	1
U22	0.6	0.601	0.07	0.064	0.001	0.002	0.07	0.952	0.998
U33	2	1.996	0.105	0.097	-0.004	-0.002	0.105	0.936	1
U44	1	1.004	0.103	0.059	0.004	0.004	0.103	0.934	1
U55	1.5	1.495	0.095	0.076	-0.005	-0.003	0.095	0.954	1
U66	0.4	0.406	0.059	0.047	0.006	0.014	0.059	0.956	0.998
X01	-	0.008	0.088	0.089	-0.992	-0.992	0.996	0	0.051
X02	-	0.099	2.245	0.22					
P011	-	1.342	0.187	0.172					
P012	-	0.106	3.947	0.338					
P022	-	1.631	0.291	0.474					

Table 5.8: Simulation Results: True IC: Model Implied, Fitted IC: Free Parameter, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.127	0.041	0.059	-0.073	-0.061	0.084	0.811	1
Z31	0.8	0.859	0.056	0.064	0.059	0.074	0.082	0.87	1
Z52	0.9	0.931	0.047	0.054	0.031	0.035	0.057	0.96	1
Z62	1.1	1.019	0.035	0.049	-0.081	-0.074	0.088	0.637	1
V11	1	0.969	0.088	0.096	-0.031	-0.031	0.093	0.95	1
V21	0.4	0.323	0.056	0.057	-0.077	-0.192	0.095	0.717	1
V22	1	1.065	0.099	0.104	0.065	0.065	0.118	0.95	1
T11	0.5	0.54	0.047	0.05	0.04	0.079	0.062	0.88	1
T21	-0.3	-0.315	0.049	0.054	-0.015	0.05	0.051	0.969	1
T12	-0.1	-0.1	0.04	0.043	0	0.001	0.04	0.953	0.663
T22	0.6	0.646	0.041	0.045	0.046	0.076	0.061	0.83	1
U11	0.8	1.155	0.068	0.062	0.355	0.443	0.361	0	1
U22	0.6	1.271	0.082	0.069	0.671	1.119	0.676	0	1
U33	2	2.125	0.112	0.105	0.125	0.063	0.168	0.778	1
U44	1	1.399	0.084	0.073	0.399	0.399	0.408	0	1
U55	1.5	1.756	0.091	0.089	0.256	0.171	0.272	0.163	1
U66	0.4	1.103	0.073	0.058	0.703	1.757	0.707	0	1

Table 5.9: Simulation Results: True IC: Model Implied, Fitted IC: Null, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.027	0.035	0.036	-0.173	-0.144	0.177	0.002	1
Z31	0.8	0.706	0.043	0.041	-0.094	-0.118	0.104	0.385	1
Z52	0.9	0.786	0.246	0.035	-0.114	-0.127	0.271	0.202	1
Z62	1.1	0.941	0.282	0.03	-0.159	-0.144	0.323	0.002	1
V11	1	1.19	0.082	0.088	0.19	0.19	0.206	0.419	1
V21	0.4	0.468	0.075	0.058	0.068	0.17	0.101	0.786	1
V22	1	1.18	0.118	0.089	0.18	0.18	0.215	0.463	1
T11	0.5	0.477	0.035	0.035	-0.023	-0.047	0.042	0.898	1
T21	-0.3	-0.284	0.048	0.036	0.016	-0.053	0.051	0.918	1
T12	-0.1	-0.095	0.038	0.031	0.005	-0.048	0.038	0.936	0.87
T22	0.6	0.581	0.032	0.031	-0.019	-0.032	0.037	0.906	1
U11	0.8	0.713	0.059	0.056	-0.087	-0.109	0.105	0.639	1
U22	0.6	0.716	0.062	0.062	0.116	0.194	0.131	0.531	1
U33	2	2.001	0.103	0.097	0.001	0	0.103	0.94	1
U44	1	0.963	0.252	0.061	-0.037	-0.037	0.255	0.83	1
U55	1.5	1.494	0.075	0.076	-0.006	-0.004	0.075	0.96	1
U66	0.4	0.471	0.045	0.046	0.071	0.177	0.084	0.693	1

Table 5.10: Simulation Results: True IC: Model Implied, Fitted IC: deJong DKF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.027	0.035	0.036	-0.173	-0.144	0.177	0.002	1
Z31	0.8	0.705	0.043	0.041	-0.095	-0.119	0.104	0.381	1
Z52	0.9	0.786	0.226	0.034	-0.114	-0.126	0.253	0.2	1
Z62	1.1	0.941	0.261	0.03	-0.159	-0.145	0.305	0.002	1
V11	1	1.191	0.084	0.088	0.191	0.191	0.208	0.418	1
V21	0.4	0.465	0.088	0.058	0.065	0.163	0.109	0.789	1
V22	1	1.177	0.135	0.089	0.177	0.177	0.222	0.458	0.996
T11	0.5	0.476	0.036	0.036	-0.024	-0.049	0.044	0.893	1
T21	-0.3	-0.284	0.049	0.036	0.016	-0.054	0.051	0.913	0.998
T12	-0.1	-0.096	0.036	0.031	0.004	-0.044	0.037	0.932	0.867
T22	0.6	0.579	0.039	0.031	-0.021	-0.035	0.044	0.9	0.998
U11	0.8	0.712	0.06	0.056	-0.088	-0.11	0.106	0.634	1
U22	0.6	0.717	0.061	0.062	0.117	0.195	0.132	0.527	1
U33	2	2.001	0.104	0.097	0.001	0.001	0.104	0.939	1
U44	1	0.967	0.269	0.061	-0.033	-0.033	0.271	0.824	1
U55	1.5	1.494	0.114	0.076	-0.006	-0.004	0.114	0.954	0.998
U66	0.4	0.472	0.068	0.046	0.072	0.18	0.099	0.684	0.998

Table 5.11: Simulation Results: True IC: Model Implied, Fitted IC: Koopman exact initial KF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.027	0.035	0.036	-0.173	-0.144	0.177	0.002	1
Z31	0.8	0.705	0.043	0.041	-0.095	-0.119	0.104	0.38	1
Z52	0.9	0.787	0.226	0.034	-0.113	-0.126	0.252	0.205	1
Z62	1.1	0.942	0.26	0.03	-0.158	-0.144	0.304	0.002	1
V11	1	1.191	0.083	0.088	0.191	0.191	0.208	0.423	1
V21	0.4	0.465	0.088	0.058	0.065	0.162	0.109	0.79	1
V22	1	1.176	0.134	0.089	0.176	0.176	0.221	0.46	0.996
T11	0.5	0.475	0.036	0.036	-0.025	-0.049	0.044	0.894	1
T21	-0.3	-0.284	0.048	0.036	0.016	-0.054	0.051	0.914	0.998
T12	-0.1	-0.096	0.036	0.031	0.004	-0.043	0.037	0.933	0.868
T22	0.6	0.579	0.039	0.031	-0.021	-0.035	0.044	0.901	0.998
U11	0.8	0.712	0.06	0.056	-0.088	-0.11	0.106	0.633	1
U22	0.6	0.716	0.061	0.062	0.116	0.194	0.132	0.531	1
U33	2	2.001	0.104	0.097	0.001	0.001	0.103	0.94	1
U44	1	0.967	0.268	0.061	-0.033	-0.033	0.27	0.825	1
U55	1.5	1.494	0.114	0.076	-0.006	-0.004	0.114	0.955	0.998
U66	0.4	0.472	0.068	0.046	0.072	0.18	0.099	0.683	0.998

Table 5.12: Simulation Results: True IC: Model Implied, Fitted IC: Large κ, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.199	0.051	0.049	-0.001	-0.001	0.051	0.947	1
Z31	0.8	0.803	0.05	0.049	0.003	0.004	0.05	0.935	1
Z52	0.9	0.899	0.039	0.041	-0.001	-0.001	0.039	0.969	1
Z62	1.1	1.1	0.04	0.038	0	0	0.04	0.94	1
V11	1	0.988	0.129	0.076	-0.012	-0.012	0.13	0.926	1
V21	0.4	0.397	0.085	0.045	-0.003	-0.006	0.085	0.926	1
V22	1	0.988	0.094	0.075	-0.012	-0.012	0.094	0.957	1
T11	0.5	0.486	0.075	0.033	-0.014	-0.028	0.076	0.945	1
T21	-0.3	-0.3	0.058	0.035	0	0	0.058	0.935	1
T12	-0.1	-0.1	0.049	0.029	0	-0.002	0.049	0.94	0.94
T22	0.6	0.586	0.043	0.029	-0.014	-0.023	0.045	0.938	0.998
U11	0.8	0.802	0.055	0.054	0.002	0.003	0.055	0.938	1
U22	0.6	0.601	0.064	0.062	0.001	0.001	0.064	0.94	1
U33	2	1.993	0.096	0.096	-0.007	-0.003	0.096	0.955	1
U44	1	1	0.058	0.058	0	0	0.058	0.933	1
U55	1.5	1.499	0.079	0.076	-0.001	0	0.079	0.943	1
U66	0.4	0.397	0.051	0.045	-0.003	-0.006	0.051	0.926	0.998

Table 5.13: Stationary PFA Results: True IC: Null, Fitted IC: Model Implied, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.199	0.059	0.049	-0.001	-0.001	0.059	0.929	0.998
Z31	0.8	0.8	0.106	0.049	0	0	0.106	0.935	1
Z52	0.9	0.9	0.04	0.041	0	0	0.04	0.961	1
Z62	1.1	1.102	0.039	0.038	0.002	0.001	0.039	0.942	1
V11	1	1.001	0.089	0.077	0.001	0.001	0.089	0.95	0.998
V21	0.4	0.398	0.048	0.045	-0.002	-0.004	0.048	0.944	0.998
V22	1	1.001	0.074	0.076	0.001	0.001	0.074	0.95	1
T11	0.5	0.498	0.036	0.034	-0.002	-0.003	0.036	0.95	0.998
T21	-0.3	-0.301	0.04	0.035	-0.001	0.005	0.04	0.942	1
T12	-0.1	-0.099	0.03	0.029	0.001	-0.013	0.03	0.944	0.931
T22	0.6	0.597	0.029	0.029	-0.003	-0.005	0.029	0.935	1
U11	0.8	0.804	0.084	0.053	0.004	0.005	0.084	0.938	1
U22	0.6	0.601	0.109	0.06	0.001	0.001	0.109	0.931	1
U33	2	1.999	0.105	0.096	-0.001	0	0.105	0.944	1
U44	1	1.002	0.06	0.057	0.002	0.002	0.06	0.931	1
U55	1.5	1.497	0.077	0.075	-0.003	-0.002	0.076	0.946	1
U66	0.4	0.394	0.045	0.044	-0.006	-0.016	0.045	0.942	1
X01	-	-0.006	0.106	0.113					
X02	-	-0.088	2.448	0.129					
P011	-	0.033	0.122	0.04					
P012	-	-0.164	2.335	0.097					
P022	-	0.023	0.042	0.032					

Table 5.14: Stationary PFA Results: True IC: Null, Fitted IC: Free Parameter, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.199	0.049	0.048	-0.001	-0.001	0.049	0.935	1
Z31	0.8	0.804	0.05	0.049	0.004	0.004	0.05	0.937	1
Z52	0.9	0.899	0.039	0.041	-0.001	-0.001	0.038	0.962	1
Z62	1.1	1.092	0.167	0.038	-0.008	-0.007	0.167	0.95	1
V11	1	1.004	0.076	0.077	0.004	0.004	0.076	0.956	1
V21	0.4	0.399	0.055	0.046	-0.001	-0.003	0.054	0.948	1
V22	1	0.999	0.086	0.076	-0.001	-0.001	0.086	0.954	0.998
T11	0.5	0.5	0.034	0.034	0	0	0.034	0.95	1
T21	-0.3	-0.301	0.036	0.035	-0.001	0.002	0.036	0.943	1
T12	-0.1	-0.096	0.07	0.03	0.004	-0.043	0.07	0.945	0.937
T22	0.6	0.595	0.068	0.029	-0.005	-0.008	0.068	0.941	1
U11	0.8	0.801	0.054	0.053	0.001	0.001	0.054	0.939	1
U22	0.6	0.601	0.061	0.059	0.001	0.001	0.061	0.937	1
U33	2	1.997	0.097	0.096	-0.003	-0.002	0.097	0.945	1
U44	1	1.004	0.096	0.058	0.004	0.004	0.096	0.945	1
U55	1.5	1.501	0.108	0.075	0.001	0.001	0.108	0.948	1
U66	0.4	0.404	0.115	0.043	0.004	0.009	0.115	0.941	1

Table 5.15: Stationary PFA Results: True IC: Null, Fitted IC: Null, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.167	0.132	0.047	-0.033	-0.027	0.135	0.899	1
Z31	0.8	0.788	0.07	0.048	-0.012	-0.016	0.071	0.926	0.996
Z52	0.9	0.886	0.038	0.04	-0.014	-0.016	0.041	0.948	1
Z62	1.1	1.081	0.037	0.037	-0.019	-0.017	0.041	0.901	1
V11	1	1.024	0.1	0.078	0.024	0.024	0.103	0.942	0.996
V21	0.4	0.411	0.046	0.047	0.011	0.026	0.047	0.956	1
V22	1	1.026	0.076	0.077	0.026	0.026	0.081	0.958	1
T11	0.5	0.499	0.038	0.034	-0.001	-0.002	0.038	0.946	1
T21	-0.3	-0.295	0.078	0.035	0.005	-0.015	0.078	0.938	1
T12	-0.1	-0.098	0.029	0.029	0.002	-0.015	0.029	0.956	0.938
T22	0.6	0.596	0.033	0.03	-0.004	-0.007	0.033	0.938	1
U11	0.8	0.795	0.11	0.054	-0.005	-0.006	0.11	0.926	1
U22	0.6	0.624	0.139	0.062	0.024	0.039	0.141	0.938	1
U33	2	1.999	0.113	0.096	-0.001	0	0.113	0.946	1
U44	1	0.991	0.057	0.058	-0.009	-0.009	0.058	0.936	1
U55	1.5	1.498	0.078	0.076	-0.002	-0.002	0.078	0.944	1
U66	0.4	0.406	0.046	0.045	0.006	0.016	0.046	0.942	1

Table 5.16: Stationary PFA Results: True IC: Null, Fitted IC: deJong DKF, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.175	0.048	0.047	-0.025	-0.021	0.054	0.903	1
Z31	0.8	0.791	0.049	0.048	-0.009	-0.012	0.05	0.931	1
Z52	0.9	0.885	0.039	0.04	-0.015	-0.016	0.042	0.943	1
Z62	1.1	1.08	0.048	0.037	-0.02	-0.018	0.052	0.897	1
V11	1	1.029	0.076	0.078	0.029	0.029	0.082	0.945	1
V21	0.4	0.409	0.053	0.047	0.009	0.023	0.054	0.956	1
V22	1	1.023	0.087	0.077	0.023	0.023	0.09	0.96	0.998
T11	0.5	0.497	0.036	0.034	-0.003	-0.005	0.036	0.952	1
T21	-0.3	-0.299	0.035	0.035	0.001	-0.004	0.035	0.939	1
T12	-0.1	-0.099	0.033	0.029	0.001	-0.01	0.033	0.952	0.935
T22	0.6	0.597	0.029	0.029	-0.003	-0.005	0.029	0.939	1
U11	0.8	0.789	0.055	0.054	-0.011	-0.014	0.056	0.931	1
U22	0.6	0.615	0.062	0.062	0.015	0.025	0.064	0.943	1
U33	2	1.996	0.097	0.096	-0.004	-0.002	0.097	0.952	1
U44	1	0.994	0.085	0.058	-0.006	-0.006	0.085	0.931	1
U55	1.5	1.494	0.103	0.075	-0.006	-0.004	0.103	0.947	0.998
U66	0.4	0.408	0.066	0.045	0.008	0.021	0.067	0.937	1

Table 5.17: Stationary PFA Results: True IC: Null, Fitted IC: Koopman exact initial KF, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.175	0.05	0.047	-0.025	-0.021	0.055	0.901	1
Z31	0.8	0.791	0.049	0.048	-0.009	-0.012	0.049	0.931	1
Z52	0.9	0.885	0.04	0.04	-0.015	-0.017	0.042	0.941	1
Z62	1.1	1.079	0.055	0.037	-0.021	-0.019	0.059	0.895	1
V11	1	1.029	0.076	0.078	0.029	0.029	0.082	0.945	1
V21	0.4	0.409	0.054	0.047	0.009	0.022	0.054	0.954	1
V22	1	1.021	0.099	0.077	0.021	0.021	0.101	0.958	0.996
T11	0.5	0.497	0.036	0.034	-0.003	-0.006	0.036	0.95	1
T21	-0.3	-0.299	0.035	0.035	0.001	-0.004	0.035	0.939	1
T12	-0.1	-0.099	0.034	0.029	0.001	-0.013	0.034	0.95	0.933
T22	0.6	0.596	0.034	0.03	-0.004	-0.007	0.034	0.937	1
U11	0.8	0.79	0.057	0.054	-0.01	-0.013	0.057	0.929	1
U22	0.6	0.614	0.067	0.062	0.014	0.024	0.068	0.941	0.998
U33	2	1.997	0.097	0.096	-0.003	-0.002	0.097	0.952	1
U44	1	0.996	0.097	0.058	-0.004	-0.004	0.097	0.929	1
U55	1.5	1.496	0.108	0.075	-0.004	-0.003	0.108	0.945	0.998
U66	0.4	0.407	0.069	0.045	0.007	0.019	0.069	0.935	0.998

Table 5.18: Stationary PFA Results: True IC: Null, Fitted IC: Large κ, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.211	0.054	0.066	0.011	0.009	0.055	1	1
Z31	0.8	0.796	0.055	0.059	-0.004	-0.005	0.055	0.961	1
Z52	0.9	0.89	0.053	0.052	-0.01	-0.011	0.054	0.922	1
Z62	1.1	1.103	0.057	0.053	0.003	0.003	0.056	0.961	1
V11	1	0.802	0.068	0.071	-0.198	-0.198	0.209	0.235	1
V21	0.4	0.306	0.042	0.039	-0.094	-0.234	0.103	0.294	1
V22	1	0.832	0.086	0.072	-0.168	-0.168	0.188	0.392	1
T11	0.5	0.394	0.039	0.039	-0.106	-0.213	0.113	0.235	1
T21	-0.3	-0.237	0.031	0.042	0.063	-0.21	0.07	0.745	1
T12	-0.1	-0.063	0.039	0.037	0.037	-0.37	0.054	0.784	0.49
T22	0.6	0.471	0.033	0.037	-0.129	-0.215	0.133	0.059	1
U11	0.8	0.795	0.047	0.057	-0.005	-0.006	0.047	0.98	1
U22	0.6	0.585	0.056	0.068	-0.015	-0.024	0.057	0.98	1
U33	2	1.989	0.09	0.096	-0.011	-0.005	0.09	0.98	1
U44	1	0.972	0.068	0.06	-0.028	-0.028	0.073	0.902	1
U55	1.5	1.496	0.065	0.076	-0.004	-0.003	0.065	0.98	1
U66	0.4	0.394	0.055	0.05	-0.006	-0.016	0.055	0.941	1

Table 5.19: Stationary PFA Results: True IC: Null, Fitted IC: Model Implied, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.206	0.053	0.051	0.006	0.005	0.054	0.927	1
Z31	0.8	0.799	0.057	0.056	-0.001	-0.001	0.057	0.93	1
Z52	0.9	0.901	0.139	0.05	0.001	0.001	0.138	0.862	1
Z62	1.1	1.097	0.119	0.044	-0.003	-0.003	0.119	0.83	1
V11	1	1.003	0.083	0.084	0.003	0.003	0.083	0.954	1
V21	0.4	0.393	0.062	0.05	-0.007	-0.018	0.062	0.905	1
V22	1	1.004	0.147	0.086	0.004	0.004	0.147	0.847	0.998
T11	0.5	0.488	0.057	0.048	-0.012	-0.023	0.058	0.896	1
T21	-0.3	-0.291	0.06	0.049	0.009	-0.03	0.061	0.915	0.998
T12	-0.1	-0.096	0.069	0.046	0.004	-0.039	0.069	0.93	0.604
T22	0.6	0.58	0.07	0.046	-0.02	-0.034	0.073	0.888	1
U11	0.8	0.794	0.047	0.047	-0.006	-0.008	0.047	0.942	1
U22	0.6	0.586	0.051	0.047	-0.014	-0.024	0.053	0.925	1
U33	2	1.991	0.095	0.094	-0.009	-0.004	0.095	0.937	1
U44	1	1.005	0.125	0.055	0.005	0.005	0.125	0.84	0.998
U55	1.5	1.493	0.142	0.073	-0.007	-0.005	0.142	0.905	0.998
U66	0.4	0.391	0.127	0.036	-0.009	-0.023	0.127	0.82	0.954
X01	-	0.001	0.033	0.034					
X02	-	-0.366	4.929	0.169					
P011	-	0.013	0.023	0.013					
P012	-	-0.183	6.899	0.273					
P022	-	0.007	0.013	0.253					

Table 5.20: Stationary PFA Results: True IC: Null, Fitted IC: Free Parameter, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.051	0.052	0.001	0.001	0.051	0.956	1
Z31	0.8	0.798	0.056	0.057	-0.002	-0.002	0.056	0.947	1
Z52	0.9	0.901	0.053	0.051	0.001	0.001	0.052	0.943	1
Z62	1.1	1.098	0.045	0.045	-0.002	-0.002	0.045	0.958	1
V11	1	1.003	0.083	0.086	0.003	0.003	0.082	0.96	1
V21	0.4	0.4	0.049	0.05	0	0	0.049	0.956	1
V22	1	0.997	0.087	0.088	-0.003	-0.003	0.087	0.952	1
T11	0.5	0.499	0.051	0.048	-0.001	-0.001	0.051	0.936	1
T21	-0.3	-0.3	0.049	0.05	0	0.001	0.049	0.958	1
T12	-0.1	-0.101	0.048	0.046	-0.001	0.013	0.047	0.938	0.608
T22	0.6	0.597	0.045	0.046	-0.003	-0.006	0.045	0.963	1
U11	0.8	0.796	0.047	0.047	-0.004	-0.005	0.047	0.947	1
U22	0.6	0.597	0.045	0.045	-0.003	-0.005	0.045	0.943	1
U33	2	1.995	0.095	0.095	-0.005	-0.003	0.095	0.943	1
U44	1	0.998	0.055	0.054	-0.002	-0.002	0.055	0.936	1
U55	1.5	1.493	0.067	0.073	-0.007	-0.005	0.068	0.965	1
U66	0.4	0.4	0.033	0.032	0	0.001	0.033	0.938	1

Table 5.21: Stationary PFA Results: True IC: Null, Fitted IC: Null, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.926	0.047	0.044	-0.274	-0.228	0.278	0	1
Z31	0.8	0.652	0.05	0.048	-0.148	-0.185	0.156	0.158	1
Z52	0.9	0.749	0.043	0.041	-0.151	-0.168	0.157	0.064	1
Z62	1.1	0.867	0.038	0.036	-0.233	-0.212	0.236	0	1
V11	1	1.319	0.091	0.101	0.319	0.319	0.332	0.068	1
V21	0.4	0.496	0.061	0.065	0.096	0.24	0.114	0.711	1
V22	1	1.329	0.095	0.105	0.329	0.329	0.342	0.056	1
T11	0.5	0.443	0.053	0.048	-0.057	-0.114	0.078	0.772	1
T21	-0.3	-0.26	0.048	0.048	0.04	-0.132	0.062	0.868	0.998
T12	-0.1	-0.076	0.047	0.045	0.024	-0.235	0.053	0.896	0.435
T22	0.6	0.547	0.047	0.045	-0.053	-0.088	0.071	0.79	1
U11	0.8	0.626	0.066	0.065	-0.174	-0.218	0.186	0.23	1
U22	0.6	0.796	0.072	0.068	0.196	0.326	0.209	0.176	1
U33	2	2.02	0.101	0.098	0.02	0.01	0.103	0.942	1
U44	1	0.879	0.066	0.064	-0.121	-0.121	0.137	0.515	1
U55	1.5	1.494	0.072	0.077	-0.006	-0.004	0.072	0.964	1
U66	0.4	0.525	0.053	0.049	0.125	0.313	0.136	0.277	1

Table 5.22: Stationary PFA Results: True IC: Null, Fitted IC: deJong DKF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.927	0.047	0.044	-0.273	-0.228	0.277	0	1
Z31	0.8	0.651	0.05	0.048	-0.149	-0.186	0.157	0.147	1
Z52	0.9	0.747	0.057	0.041	-0.153	-0.17	0.163	0.066	1
Z62	1.1	0.865	0.053	0.036	-0.235	-0.214	0.241	0	1
V11	1	1.321	0.092	0.101	0.321	0.321	0.334	0.066	1
V21	0.4	0.496	0.063	0.065	0.096	0.24	0.115	0.713	1
V22	1	1.321	0.128	0.104	0.321	0.321	0.346	0.061	0.996
T11	0.5	0.443	0.054	0.048	-0.057	-0.114	0.078	0.766	1
T21	-0.3	-0.26	0.049	0.048	0.04	-0.134	0.063	0.873	0.996
T12	-0.1	-0.076	0.052	0.045	0.024	-0.24	0.057	0.895	0.446
T22	0.6	0.545	0.059	0.045	-0.055	-0.091	0.08	0.796	0.996
U11	0.8	0.625	0.066	0.065	-0.175	-0.218	0.187	0.23	1
U22	0.6	0.796	0.072	0.068	0.196	0.326	0.209	0.175	1
U33	2	2.02	0.1	0.098	0.02	0.01	0.102	0.943	1
U44	1	0.882	0.103	0.064	-0.118	-0.118	0.156	0.521	1
U55	1.5	1.497	0.086	0.077	-0.003	-0.002	0.086	0.956	1
U66	0.4	0.524	0.062	0.049	0.124	0.31	0.138	0.282	0.998

Table 5.23: Stationary PFA Results: True IC: Null, Fitted IC: Koopman exact initial KF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.927	0.047	0.044	-0.273	-0.228	0.277	0	1
Z31	0.8	0.651	0.05	0.048	-0.149	-0.186	0.157	0.15	1
Z52	0.9	0.747	0.057	0.041	-0.153	-0.17	0.163	0.065	1
Z62	1.1	0.865	0.053	0.036	-0.235	-0.213	0.241	0	1
V11	1	1.321	0.092	0.101	0.321	0.321	0.333	0.065	1
V21	0.4	0.496	0.063	0.065	0.096	0.24	0.115	0.711	1
V22	1	1.321	0.128	0.104	0.321	0.321	0.346	0.061	0.996
T11	0.5	0.443	0.054	0.048	-0.057	-0.114	0.078	0.766	1
T21	-0.3	-0.26	0.049	0.048	0.04	-0.133	0.063	0.874	0.996
T12	-0.1	-0.076	0.052	0.045	0.024	-0.239	0.057	0.896	0.447
T22	0.6	0.545	0.059	0.045	-0.055	-0.091	0.08	0.794	0.996
U11	0.8	0.625	0.066	0.065	-0.175	-0.218	0.187	0.228	1
U22	0.6	0.796	0.072	0.068	0.196	0.326	0.209	0.174	1
U33	2	2.021	0.1	0.098	0.021	0.01	0.102	0.944	1
U44	1	0.882	0.103	0.064	-0.118	-0.118	0.156	0.518	1
U55	1.5	1.497	0.086	0.077	-0.003	-0.002	0.086	0.957	1
U66	0.4	0.524	0.062	0.049	0.124	0.311	0.139	0.28	0.998

Table 5.24: Stationary PFA Results: True IC: Null, Fitted IC: Large κ, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.199	0.047	0.048	-0.001	-0.001	0.047	0.948	1
Z31	0.8	0.8	0.05	0.048	0	0	0.05	0.946	1
Z52	0.9	0.901	0.042	0.041	0.001	0.001	0.042	0.953	1
Z62	1.1	1.1	0.037	0.038	0	0	0.037	0.946	1
V11	1	1.015	0.077	0.077	0.015	0.015	0.079	0.958	1
V21	0.4	0.403	0.047	0.045	0.003	0.009	0.047	0.931	1
V22	1	0.992	0.077	0.075	-0.008	-0.008	0.077	0.943	1
T11	0.5	0.496	0.033	0.033	-0.004	-0.009	0.033	0.96	1
T21	-0.3	-0.295	0.035	0.034	0.005	-0.017	0.035	0.955	1
T12	-0.1	-0.096	0.03	0.029	0.004	-0.045	0.03	0.938	0.894
T22	0.6	0.593	0.029	0.029	-0.007	-0.012	0.03	0.96	1
U11	0.8	0.797	0.053	0.054	-0.003	-0.004	0.054	0.948	1
U22	0.6	0.597	0.059	0.062	-0.003	-0.005	0.059	0.968	1
U33	2	2	0.096	0.096	0	0	0.096	0.943	1
U44	1	1.002	0.057	0.058	0.002	0.002	0.057	0.955	1
U55	1.5	1.502	0.075	0.076	0.002	0.001	0.075	0.968	1
U66	0.4	0.396	0.044	0.045	-0.004	-0.01	0.044	0.96	1

Table 5.25: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Model Implied, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.2	0.047	0.048	0	0	0.047	0.954	1
Z31	0.8	0.801	0.05	0.048	0.001	0.001	0.05	0.943	1
Z52	0.9	0.902	0.053	0.041	0.002	0.002	0.053	0.954	1
Z62	1.1	1.1	0.038	0.038	0	0	0.038	0.95	1
V11	1	1	0.077	0.076	0	0	0.077	0.95	1
V21	0.4	0.395	0.046	0.045	-0.005	-0.013	0.047	0.929	1
V22	1	0.993	0.089	0.076	-0.007	-0.007	0.089	0.937	0.998
T11	0.5	0.494	0.033	0.033	-0.006	-0.013	0.033	0.956	1
T21	-0.3	-0.302	0.035	0.034	-0.002	0.008	0.035	0.966	1
T12	-0.1	-0.098	0.03	0.029	0.002	-0.024	0.03	0.937	0.914
T22	0.6	0.598	0.031	0.029	-0.002	-0.004	0.031	0.952	1
U11	0.8	0.799	0.053	0.054	-0.001	-0.002	0.053	0.95	1
U22	0.6	0.596	0.06	0.061	-0.004	-0.006	0.06	0.964	1
U33	2	2.001	0.096	0.096	0.001	0.001	0.096	0.943	1
U44	1	1.005	0.092	0.058	0.005	0.005	0.092	0.956	1
U55	1.5	1.499	0.101	0.076	-0.001	-0.001	0.101	0.969	0.998
U66	0.4	0.399	0.068	0.045	-0.001	-0.002	0.068	0.95	1
X01	1	1	0.27	0.261	0	0	0.27	0.939	0.95
X02	0.5	0.498	0.215	0.203	-0.002	-0.004	0.215	0.916	0.677
P011	1.2	1.153	0.443	0.441	-0.047	-0.039	0.445	0.887	0.985
P012	0.3	0.271	0.241	0.248	-0.029	-0.098	0.243	0.943	0.094
P022	0.7	0.637	0.279	0.268	-0.063	-0.09	0.286	0.841	0.87

Table 5.26: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Free Parameter, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_{θ}	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.158	0.042	0.046	-0.042	-0.035	0.059	0.833	1
Z31	0.8	0.807	0.051	0.05	0.007	0.009	0.051	0.95	1
Z52	0.9	0.899	0.058	0.041	-0.001	-0.002	0.058	0.95	0.998
Z62	1.1	1.072	0.07	0.037	-0.028	-0.025	0.075	0.897	1
V11	1	1	0.077	0.079	0	0	0.077	0.967	1
V21	0.4	0.402	0.073	0.047	0.002	0.005	0.073	0.946	1
V22	1	0.994	0.089	0.077	-0.006	-0.006	0.089	0.952	0.998
T11	0.5	0.504	0.037	0.034	0.004	0.009	0.037	0.958	0.998
T21	-0.3	-0.307	0.04	0.036	-0.007	0.023	0.04	0.95	0.998
T12	-0.1	-0.1	0.053	0.03	0	0.005	0.053	0.941	0.916
T22	0.6	0.602	0.086	0.03	0.002	0.003	0.086	0.941	1
U11	0.8	0.836	0.055	0.053	0.036	0.045	0.066	0.912	1
U22	0.6	0.778	0.066	0.056	0.178	0.297	0.19	0.13	1
U33	2	2.013	0.097	0.098	0.013	0.006	0.098	0.956	1
U44	1	1.005	0.106	0.058	0.005	0.005	0.106	0.967	1
U55	1.5	1.504	0.093	0.076	0.004	0.003	0.093	0.971	1
U66	0.4	0.487	0.086	0.041	0.087	0.218	0.122	0.46	1

Table 5.27: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Null, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_{θ}	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.177	0.045	0.046	-0.023	-0.019	0.051	0.905	1
Z31	0.8	0.788	0.048	0.047	-0.012	-0.015	0.049	0.946	1
Z52	0.9	0.887	0.041	0.04	-0.013	-0.014	0.042	0.93	1
Z62	1.1	1.082	0.036	0.036	-0.018	-0.017	0.04	0.93	1
V11	1	1.024	0.076	0.077	0.024	0.024	0.08	0.952	1
V21	0.4	0.404	0.048	0.046	0.004	0.011	0.048	0.942	1
V22	1	1.017	0.077	0.076	0.017	0.017	0.079	0.952	1
T11	0.5	0.492	0.032	0.033	-0.008	-0.015	0.033	0.948	1
T21	-0.3	-0.302	0.034	0.034	-0.002	0.007	0.034	0.966	1
T12	-0.1	-0.098	0.029	0.029	0.002	-0.024	0.029	0.94	0.92
T22	0.6	0.597	0.03	0.029	-0.003	-0.005	0.03	0.95	1
U11	0.8	0.787	0.053	0.053	-0.013	-0.016	0.054	0.936	1
U22	0.6	0.611	0.06	0.061	0.011	0.018	0.061	0.96	1
U33	2	2.004	0.097	0.097	0.004	0.002	0.097	0.942	1
U44	1	0.993	0.058	0.058	-0.007	-0.007	0.058	0.95	1
U55	1.5	1.501	0.073	0.076	0.001	0.001	0.073	0.966	1
U66	0.4	0.405	0.045	0.045	0.005	0.012	0.045	0.952	1

Table 5.28: Stationary PFA Results: True IC: Free Parameter, Fitted IC: deJong DKF, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.176	0.045	0.046	-0.024	-0.02	0.051	0.904	1
Z31	0.8	0.788	0.048	0.047	-0.012	-0.015	0.05	0.945	1
Z52	0.9	0.886	0.044	0.04	-0.014	-0.015	0.046	0.931	1
Z62	1.1	1.08	0.051	0.037	-0.02	-0.018	0.055	0.929	1
V11	1	1.025	0.076	0.077	0.025	0.025	0.08	0.954	1
V21	0.4	0.405	0.047	0.047	0.005	0.011	0.047	0.943	1
V22	1	1.016	0.089	0.076	0.016	0.016	0.091	0.952	0.998
T11	0.5	0.492	0.032	0.033	-0.008	-0.016	0.033	0.945	1
T21	-0.3	-0.301	0.035	0.034	-0.001	0.005	0.035	0.966	1
T12	-0.1	-0.097	0.032	0.029	0.003	-0.029	0.032	0.939	0.92
T22	0.6	0.596	0.039	0.029	-0.004	-0.007	0.04	0.948	0.998
U11	0.8	0.786	0.053	0.053	-0.014	-0.017	0.054	0.937	1
U22	0.6	0.611	0.06	0.061	0.011	0.018	0.061	0.96	1
U33	2	2.001	0.096	0.096	0.001	0.001	0.096	0.941	1
U44	1	0.996	0.085	0.058	-0.004	-0.004	0.085	0.95	1
U55	1.5	1.498	0.101	0.075	-0.002	-0.001	0.1	0.964	0.998
U66	0.4	0.407	0.067	0.045	0.007	0.017	0.068	0.95	1

Table 5.29: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Koopman exact initial KF, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.176	0.045	0.046	-0.024	-0.02	0.051	0.904	1
Z31	0.8	0.788	0.048	0.047	-0.012	-0.015	0.05	0.945	1
Z52	0.9	0.886	0.044	0.04	-0.014	-0.015	0.046	0.931	1
Z62	1.1	1.08	0.051	0.037	-0.02	-0.018	0.055	0.929	1
V11	1	1.024	0.076	0.077	0.024	0.024	0.08	0.954	1
V21	0.4	0.405	0.047	0.047	0.005	0.011	0.047	0.943	1
V22	1	1.016	0.089	0.076	0.016	0.016	0.091	0.952	0.998
T11	0.5	0.492	0.032	0.033	-0.008	-0.016	0.033	0.945	1
T21	-0.3	-0.301	0.035	0.034	-0.001	0.005	0.035	0.966	1
T12	-0.1	-0.097	0.032	0.029	0.003	-0.029	0.032	0.939	0.92
T22	0.6	0.596	0.039	0.029	-0.004	-0.007	0.04	0.948	0.998
U11	0.8	0.786	0.053	0.053	-0.014	-0.017	0.054	0.937	1
U22	0.6	0.611	0.06	0.061	0.011	0.018	0.061	0.958	1
U33	2	2.001	0.096	0.096	0.001	0.001	0.096	0.941	1
U44	1	0.996	0.085	0.058	-0.004	-0.004	0.085	0.95	1
U55	1.5	1.498	0.101	0.075	-0.002	-0.001	0.1	0.964	0.998
U66	0.4	0.407	0.067	0.045	0.007	0.017	0.068	0.95	1

Table 5.30: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Large κ, N=20, T=50

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.157	0.196	0.047	-0.043	-0.036	0.196	0.957	1
Z31	0.8	0.761	0.138	0.046	-0.039	-0.049	0.141	0.913	0.957
Z52	0.9	0.894	0.038	0.044	-0.006	-0.006	0.038	1	1
Z62	1.1	1.088	0.04	0.043	-0.012	-0.011	0.041	0.957	1
V11	1	1.079	0.252	0.103	0.079	0.079	0.259	0.609	0.957
V21	0.4	0.482	0.055	0.062	0.082	0.205	0.098	0.565	0.957
V22	1	1.013	0.111	0.084	0.013	0.013	0.11	0.913	1
T11	0.5	0.34	0.86	0.046	-0.16	-0.32	0.856	0.87	1
T21	-0.3	-0.255	0.117	0.049	0.045	-0.15	0.123	0.478	1
T12	-0.1	0.014	0.377	0.064	0.114	-1.143	0.386	0.957	0.261
T22	0.6	0.545	0.045	0.044	-0.055	-0.091	0.07	0.652	1
U11	0.8	0.869	0.367	0.063	0.069	0.086	0.366	0.913	1
U22	0.6	0.692	0.418	0.068	0.092	0.153	0.419	0.87	1
U33	2	2.043	0.173	0.098	0.043	0.021	0.174	0.957	1
U44	1	0.986	0.058	0.06	-0.014	-0.014	0.058	1	1
U55	1.5	1.502	0.074	0.076	0.002	0.001	0.072	0.957	1
U66	0.4	0.402	0.043	0.048	0.002	0.005	0.043	1	1

Table 5.31: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Model Implied, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.2	0.047	0.044	0	0	0.047	0.936	1
Z31	0.8	0.798	0.044	0.044	-0.002	-0.002	0.044	0.953	1
Z52	0.9	0.901	0.046	0.045	0.001	0.001	0.046	0.945	1
Z62	1.1	1.102	0.046	0.043	0.002	0.002	0.046	0.949	1
V11	1	1.002	0.078	0.079	0.002	0.002	0.078	0.96	1
V21	0.4	0.397	0.049	0.05	-0.003	-0.007	0.049	0.945	1
V22	1	0.996	0.085	0.084	-0.004	-0.004	0.085	0.94	1
T11	0.5	0.497	0.034	0.034	-0.003	-0.007	0.035	0.947	1
T21	-0.3	-0.302	0.036	0.036	-0.002	0.006	0.036	0.947	1
T12	-0.1	-0.1	0.037	0.038	0	-0.001	0.037	0.953	0.765
T22	0.6	0.598	0.039	0.038	-0.002	-0.003	0.039	0.933	1
U11	0.8	0.803	0.056	0.054	0.003	0.003	0.056	0.942	1
U22	0.6	0.596	0.06	0.062	-0.004	-0.007	0.06	0.96	1
U33	2	1.998	0.097	0.096	-0.002	-0.001	0.096	0.958	1
U44	1	1	0.068	0.058	0	0	0.068	0.929	1
U55	1.5	1.499	0.077	0.076	-0.001	0	0.077	0.947	1
U66	0.4	0.395	0.055	0.046	-0.005	-0.012	0.055	0.94	0.996
X01	1	0.999	0.091	0.088	-0.001	-0.001	0.091	0.938	1
X02	0.5	0.751	3.563	0.068	0.251	0.501	3.568	0.953	1
P011	1.2	1.19	0.139	0.152	-0.01	-0.009	0.139	0.962	1
P012	0.3	0.403	5.975	0.083	0.103	0.343	5.969	0.947	0.965
P022	0.7	0.686	0.133	0.096	-0.014	-0.02	0.134	0.905	0.984

Table 5.32: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Free Parameter, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.127	0.039	0.066	-0.073	-0.061	0.083	0.901	1
Z31	0.8	0.892	0.058	0.07	0.092	0.115	0.109	0.807	1
Z52	0.9	0.919	0.05	0.054	0.019	0.021	0.054	0.953	1
Z62	1.1	1.025	0.035	0.047	-0.075	-0.069	0.083	0.667	1
V11	1	0.907	0.084	0.099	-0.093	-0.093	0.125	0.865	1
V21	0.4	0.364	0.052	0.056	-0.036	-0.09	0.063	0.901	1
V22	1	1.003	0.089	0.096	0.003	0.003	0.089	0.973	1
T11	0.5	0.557	0.047	0.053	0.057	0.114	0.074	0.825	1
T21	-0.3	-0.351	0.054	0.056	-0.051	0.171	0.074	0.852	1
T12	-0.1	-0.1	0.041	0.046	0	-0.003	0.041	0.966	0.611
T22	0.6	0.642	0.043	0.047	0.042	0.07	0.06	0.883	1
U11	0.8	1.409	0.082	0.073	0.609	0.761	0.614	0	1
U22	0.6	1.622	0.102	0.084	1.022	1.703	1.027	0	1
U33	2	2.239	0.108	0.111	0.239	0.119	0.262	0.402	1
U44	1	1.205	0.064	0.064	0.205	0.205	0.215	0.09	1
U55	1.5	1.629	0.078	0.082	0.129	0.086	0.151	0.697	1
U66	0.4	0.841	0.054	0.047	0.441	1.103	0.445	0	1

Table 5.33: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Null, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.056	0.036	0.033	-0.144	-0.12	0.149	0.03	1
Z31	0.8	0.717	0.039	0.038	-0.083	-0.104	0.092	0.394	1
Z52	0.9	0.785	0.036	0.036	-0.115	-0.127	0.12	0.12	1
Z62	1.1	0.934	0.031	0.031	-0.166	-0.151	0.169	0	1
V11	1	1.161	0.081	0.085	0.161	0.161	0.18	0.534	1
V21	0.4	0.469	0.056	0.059	0.069	0.173	0.089	0.824	1
V22	1	1.229	0.087	0.094	0.229	0.229	0.245	0.282	1
T11	0.5	0.477	0.034	0.033	-0.023	-0.046	0.041	0.904	1
T21	-0.3	-0.294	0.035	0.035	0.006	-0.019	0.035	0.94	1
T12	-0.1	-0.087	0.034	0.035	0.013	-0.132	0.037	0.942	0.692
T22	0.6	0.571	0.037	0.037	-0.029	-0.048	0.047	0.892	1
U11	0.8	0.735	0.057	0.054	-0.065	-0.081	0.087	0.768	1
U22	0.6	0.689	0.06	0.061	0.089	0.149	0.108	0.704	1
U33	2	2.002	0.096	0.097	0.002	0.001	0.096	0.96	1
U44	1	0.928	0.064	0.061	-0.072	-0.072	0.096	0.76	1
U55	1.5	1.495	0.075	0.076	-0.005	-0.003	0.075	0.954	1
U66	0.4	0.481	0.048	0.047	0.081	0.204	0.095	0.61	1

Table 5.34: Stationary PFA Results: True IC: Free Parameter, Fitted IC: deJong DKF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.056	0.037	0.033	-0.144	-0.12	0.148	0.036	1
Z31	0.8	0.716	0.037	0.038	-0.084	-0.104	0.091	0.389	1
Z52	0.9	0.784	0.04	0.036	-0.116	-0.129	0.123	0.114	1
Z62	1.1	0.933	0.036	0.031	-0.167	-0.151	0.17	0	1
V11	1	1.162	0.081	0.085	0.162	0.162	0.181	0.528	1
V21	0.4	0.469	0.055	0.059	0.069	0.172	0.088	0.832	1
V22	1	1.224	0.12	0.094	0.224	0.224	0.255	0.277	0.996
T11	0.5	0.476	0.034	0.033	-0.024	-0.048	0.042	0.902	1
T21	-0.3	-0.294	0.036	0.035	0.006	-0.018	0.037	0.937	0.998
T12	-0.1	-0.085	0.037	0.035	0.015	-0.145	0.04	0.94	0.694
T22	0.6	0.57	0.043	0.037	-0.03	-0.05	0.052	0.895	1
U11	0.8	0.737	0.058	0.055	-0.063	-0.079	0.085	0.774	1
U22	0.6	0.688	0.06	0.061	0.088	0.147	0.107	0.716	1
U33	2	2.003	0.097	0.097	0.003	0.002	0.097	0.957	1
U44	1	0.935	0.115	0.061	-0.065	-0.065	0.132	0.756	1
U55	1.5	1.499	0.09	0.076	-0.001	-0.001	0.09	0.953	1
U66	0.4	0.478	0.058	0.047	0.078	0.194	0.097	0.624	0.996

Table 5.35: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Koopman exact initial KF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.056	0.037	0.033	-0.144	-0.12	0.148	0.036	1
Z31	0.8	0.717	0.037	0.038	-0.083	-0.104	0.091	0.393	1
Z52	0.9	0.784	0.04	0.036	-0.116	-0.129	0.123	0.113	1
Z62	1.1	0.933	0.036	0.031	-0.167	-0.151	0.17	0	1
V11	1	1.162	0.081	0.085	0.162	0.162	0.181	0.524	1
V21	0.4	0.469	0.056	0.059	0.069	0.173	0.089	0.827	1
V22	1	1.225	0.12	0.094	0.225	0.225	0.255	0.276	0.996
T11	0.5	0.476	0.034	0.033	-0.024	-0.048	0.041	0.904	1
T21	-0.3	-0.294	0.036	0.035	0.006	-0.019	0.037	0.938	0.998
T12	-0.1	-0.085	0.037	0.035	0.015	-0.145	0.04	0.938	0.696
T22	0.6	0.57	0.043	0.037	-0.03	-0.05	0.052	0.893	1
U11	0.8	0.736	0.058	0.055	-0.064	-0.08	0.086	0.769	1
U22	0.6	0.688	0.06	0.061	0.088	0.147	0.107	0.713	1
U33	2	2.004	0.097	0.097	0.004	0.002	0.096	0.958	1
U44	1	0.934	0.114	0.061	-0.066	-0.066	0.132	0.753	1
U55	1.5	1.498	0.091	0.076	-0.002	-0.001	0.09	0.953	1
U66	0.4	0.478	0.058	0.047	0.078	0.194	0.097	0.624	0.996

Table 5.36: Stationary PFA Results: True IC: Free Parameter, Fitted IC: Large κ, N=200, T=5

CHAPTER 6

Appendix B

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.198	0.018	0.018	-0.002	-0.001	0.018	0.946	1
Z31	0.8	0.799	0.021	0.021	-0.001	-0.001	0.021	0.952	1
Z52	0.9	0.901	0.034	0.019	0.001	0.001	0.034	0.949	1
Z62	1.1	1.103	0.031	0.017	0.003	0.003	0.031	0.952	1
V11	0.05	0.085	0.335	0.011	0.035	0.69	0.336	0.92	1
V21	0.02	0.032	0.137	0.011	0.012	0.596	0.138	0.885	0.713
V22	0.05	0.049	0.012	0.012	-0.001	-0.022	0.012	0.924	0.987
T11	1.2	1.183	0.166	0.008	-0.017	-0.014	0.166	0.949	0.997
T21	0.6	0.595	0.046	0.012	-0.005	-0.008	0.046	0.946	0.99
T12	-0.4	-0.378	0.232	0.01	0.022	-0.055	0.233	0.946	1
T22	0.7	0.701	0.018	0.009	0.001	0.002	0.018	0.962	1
U11	0.8	0.799	0.044	0.038	-0.001	-0.001	0.044	0.939	1
U22	0.6	0.599	0.047	0.03	-0.001	-0.001	0.047	0.917	1
U33	2	1.992	0.092	0.09	-0.008	-0.004	0.092	0.943	1
U44	1	1.062	0.585	0.05	0.062	0.062	0.587	0.93	1
U55	1.5	1.542	0.522	0.071	0.042	0.028	0.523	0.917	0.997
U66	0.4	0.465	0.653	0.025	0.065	0.162	0.655	0.927	1
X01	-	0.001	0.073	0.075					
X02	-	0.043	5.464	0.187					
P011	-	0.065	0.368	0.034					
P012	-	0.992	10.336	0.417					
P022	-	0.009	0.035	0.259					

Table 6.1: Nonstationary PFA Results: True IC: Null, Fitted IC: Free Parameter, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.199	0.018	0.018	-0.001	-0.001	0.018	0.945	1
Z31	0.8	0.8	0.021	0.021	0	0	0.021	0.953	1
Z52	0.9	0.899	0.018	0.017	-0.001	-0.001	0.018	0.945	1
Z62	1.1	1.101	0.019	0.014	0.001	0.001	0.019	0.947	1
V11	0.05	0.067	0.225	0.009	0.017	0.336	0.225	0.947	0.997
V21	0.02	0.025	0.101	0.009	0.005	0.258	0.101	0.895	0.767
V22	0.05	0.051	0.042	0.011	0.001	0.028	0.042	0.925	0.994
T11	1.2	1.191	0.125	0.007	-0.009	-0.008	0.126	0.95	1
T21	0.6	0.598	0.035	0.011	-0.002	-0.004	0.035	0.947	0.994
T12	-0.4	-0.388	0.174	0.008	0.012	-0.031	0.174	0.939	1
T22	0.7	0.699	0.023	0.006	-0.001	-0.002	0.023	0.958	1
U11	0.8	0.795	0.038	0.038	-0.005	-0.007	0.038	0.945	1
U22	0.6	0.601	0.033	0.03	0.001	0.002	0.033	0.931	1
U33	2	1.994	0.091	0.09	-0.006	-0.003	0.091	0.947	1
U44	1	1.035	0.423	0.049	0.035	0.035	0.424	0.953	1
U55	1.5	1.527	0.369	0.07	0.027	0.018	0.37	0.931	1
U66	0.4	0.436	0.456	0.023	0.036	0.09	0.457	0.928	1

Table 6.2: Nonstationary PFA Results: True IC: Null, Fitted IC: Null, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.196	0.018	0.017	-0.004	-0.003	0.018	0.941	1
Z31	0.8	0.798	0.021	0.021	-0.002	-0.003	0.021	0.948	1
Z52	0.9	0.897	0.018	0.017	-0.003	-0.003	0.018	0.955	1
Z62	1.1	1.098	0.015	0.014	-0.002	-0.002	0.015	0.944	1
V11	0.05	0.065	0.207	0.009	0.015	0.294	0.207	0.941	1
V21	0.02	0.026	0.09	0.009	0.006	0.31	0.09	0.918	0.713
V22	0.05	0.05	0.011	0.011	0	-0.007	0.011	0.939	0.995
T11	1.2	1.191	0.115	0.007	-0.009	-0.008	0.115	0.96	1
T21	0.6	0.597	0.032	0.011	-0.003	-0.005	0.032	0.941	0.995
T12	-0.4	-0.389	0.155	0.008	0.011	-0.026	0.155	0.934	1
T22	0.7	0.7	0.009	0.006	0	0	0.009	0.958	1
U11	0.8	0.796	0.038	0.038	-0.004	-0.005	0.038	0.944	1
U22	0.6	0.601	0.034	0.031	0.001	0.001	0.034	0.92	1
U33	2	1.997	0.089	0.091	-0.003	-0.001	0.089	0.955	1
U44	1	1.024	0.374	0.049	0.024	0.024	0.375	0.948	1
U55	1.5	1.519	0.323	0.07	0.019	0.013	0.323	0.929	1
U66	0.4	0.425	0.406	0.023	0.025	0.063	0.406	0.927	1

Table 6.3: Nonstationary PFA Results: True IC: Null, Fitted IC: deJong DKF, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.197	0.021	0.018	-0.003	-0.002	0.022	0.926	1
Z31	0.8	0.797	0.021	0.021	-0.003	-0.003	0.022	0.947	1
Z52	0.9	0.889	0.069	0.019	-0.011	-0.012	0.069	0.937	1
Z62	1.1	1.085	0.094	0.017	-0.015	-0.014	0.095	0.937	1
V11	0.05	0.109	0.367	0.012	0.059	1.174	0.371	0.908	1
V21	0.02	0.044	0.305	0.012	0.024	1.201	0.306	0.881	0.712
V22	0.05	0.048	0.015	0.014	-0.002	-0.041	0.015	0.918	0.958
T11	1.2	1.178	0.174	0.008	-0.022	-0.019	0.175	0.921	1
T21	0.6	0.577	0.255	0.013	-0.023	-0.039	0.256	0.921	0.989
T12	-0.4	-0.37	0.236	0.009	0.03	-0.075	0.238	0.905	1
T22	0.7	0.687	0.119	0.009	-0.013	-0.019	0.119	0.931	1
U11	0.8	0.801	0.059	0.038	0.001	0.001	0.059	0.916	1
U22	0.6	0.59	0.086	0.03	-0.01	-0.017	0.087	0.894	0.995
U33	2	1.997	0.094	0.091	-0.003	-0.002	0.094	0.945	1
U44	1	1.102	0.674	0.053	0.102	0.102	0.681	0.913	1
U55	1.5	1.52	0.583	0.07	0.02	0.013	0.583	0.908	0.979
U66	0.4	0.506	0.733	0.027	0.106	0.266	0.74	0.887	0.997

Table 6.4: Nonstationary PFA Results: True IC: Null, Fitted IC: Koopman exact initial KF, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.197	0.02	0.018	-0.003	-0.002	0.021	0.929	1
Z31	0.8	0.797	0.021	0.021	-0.003	-0.003	0.022	0.947	1
Z52	0.9	0.887	0.08	0.018	-0.013	-0.015	0.081	0.937	1
Z62	1.1	1.082	0.107	0.016	-0.018	-0.017	0.109	0.931	1
V11	0.05	0.101	0.325	0.012	0.051	1.023	0.328	0.91	1
V21	0.02	0.047	0.348	0.012	0.027	1.374	0.349	0.889	0.714
V22	0.05	0.048	0.014	0.014	-0.002	-0.038	0.014	0.926	0.96
T11	1.2	1.182	0.151	0.007	-0.018	-0.015	0.152	0.921	1
T21	0.6	0.568	0.316	0.014	-0.032	-0.054	0.317	0.921	0.992
T12	-0.4	-0.375	0.209	0.009	0.025	-0.063	0.21	0.91	1
T22	0.7	0.682	0.141	0.009	-0.018	-0.025	0.142	0.929	1
U11	0.8	0.8	0.057	0.038	0	0	0.057	0.921	1
U22	0.6	0.589	0.083	0.03	-0.011	-0.018	0.084	0.897	0.995
U33	2	1.997	0.093	0.091	-0.003	-0.002	0.093	0.947	1
U44	1	1.092	0.6	0.053	0.092	0.092	0.606	0.913	1
U55	1.5	1.506	0.519	0.07	0.006	0.004	0.518	0.91	0.979
U66	0.4	0.49	0.646	0.026	0.09	0.224	0.652	0.889	0.997

Table 6.5: Nonstationary PFA Results: True IC: Null, Fitted IC: Large κ, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.2	0.03	0.028	0	0	0.03	0.933	1
Z31	0.8	0.799	0.035	0.033	-0.001	-0.001	0.035	0.935	1
Z52	0.9	0.907	0.064	0.025	0.007	0.007	0.064	0.885	1
Z62	1.1	1.102	0.057	0.021	0.002	0.002	0.057	0.865	1
V11	1	1.009	0.109	0.082	0.009	0.009	0.11	0.904	1
V21	0.4	0.396	0.088	0.055	-0.004	-0.011	0.088	0.93	1
V22	1	0.997	0.132	0.077	-0.003	-0.003	0.131	0.894	0.995
T11	1.2	1.187	0.056	0.038	-0.013	-0.011	0.057	0.882	1
T21	0.6	0.594	0.043	0.039	-0.006	-0.009	0.043	0.928	1
T12	-0.4	-0.397	0.042	0.032	0.003	-0.007	0.042	0.921	0.998
T22	1	0.99	0.055	0.031	-0.01	-0.01	0.056	0.913	1
U11	0.8	0.795	0.043	0.044	-0.005	-0.007	0.043	0.942	1
U22	0.6	0.593	0.047	0.042	-0.007	-0.011	0.047	0.913	1
U33	2	1.995	0.096	0.093	-0.005	-0.003	0.096	0.93	1
U44	1	1.051	0.293	0.053	0.051	0.051	0.297	0.897	1
U55	1.5	1.489	0.142	0.072	-0.011	-0.008	0.142	0.892	0.995
U66	0.4	0.398	0.134	0.031	-0.002	-0.006	0.134	0.894	0.978
X01	-	0	0.032	0.033					
X02	-	0.425	7.684						
P011	-	0.029	0.111						
P012	-	0.416	14.578						
P022	-	0.005	0.01						

Table 6.6: Nonstationary PFA Results: True IC: Null, Fitted IC: Free Parameter, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.198	0.03	0.029	-0.002	-0.001	0.03	0.945	1
Z31	0.8	0.8	0.035	0.034	0	-0.001	0.035	0.934	1
Z52	0.9	0.9	0.025	0.024	0	0	0.025	0.939	1
Z62	1.1	1.101	0.02	0.021	0.001	0.001	0.02	0.961	1
V11	1	0.999	0.083	0.08	-0.001	-0.001	0.083	0.932	1
V21	0.4	0.401	0.055	0.053	0.001	0.003	0.055	0.945	1
V22	1	0.996	0.076	0.075	-0.004	-0.004	0.076	0.945	1
T11	1.2	1.198	0.038	0.038	-0.002	-0.001	0.038	0.936	1
T21	0.6	0.598	0.038	0.039	-0.002	-0.004	0.038	0.957	1
T12	-0.4	-0.402	0.03	0.032	-0.002	0.006	0.03	0.961	1
T22	1	0.997	0.03	0.03	-0.003	-0.003	0.03	0.95	1
U11	0.8	0.796	0.043	0.044	-0.004	-0.005	0.043	0.961	1
U22	0.6	0.601	0.042	0.041	0.001	0.002	0.042	0.955	1
U33	2	1.995	0.094	0.094	-0.005	-0.003	0.094	0.939	1
U44	1	0.996	0.052	0.051	-0.004	-0.004	0.052	0.952	1
U55	1.5	1.497	0.073	0.072	-0.003	-0.002	0.073	0.948	1
U66	0.4	0.399	0.028	0.028	-0.001	-0.001	0.028	0.95	1

Table 6.7: Nonstationary PFA Results: True IC: Null, Fitted IC: Null, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.114	0.089	0.026	-0.086	-0.071	0.123	0.136	1
Z31	0.8	0.746	0.07	0.031	-0.054	-0.068	0.088	0.591	1
Z52	0.9	0.858	0.023	0.023	-0.042	-0.046	0.048	0.55	1
Z62	1.1	1.048	0.019	0.019	-0.052	-0.047	0.055	0.234	1
V11	1	1.137	0.107	0.095	0.137	0.137	0.174	0.706	0.998
V21	0.4	0.433	0.088	0.064	0.033	0.083	0.094	0.916	1
V22	1	1.092	0.092	0.087	0.092	0.092	0.13	0.836	1
T11	1.2	1.126	0.062	0.037	-0.074	-0.061	0.096	0.501	0.998
T21	0.6	0.577	0.047	0.037	-0.023	-0.038	0.052	0.926	0.998
T12	-0.4	-0.37	0.035	0.031	0.03	-0.075	0.046	0.84	0.998
T22	1	0.979	0.031	0.03	-0.021	-0.021	0.037	0.901	1
U11	0.8	0.778	0.16	0.049	-0.022	-0.027	0.161	0.897	1
U22	0.6	0.633	0.134	0.053	0.033	0.054	0.138	0.918	1
U33	2	1.999	0.11	0.095	-0.001	-0.001	0.11	0.943	1
U44	1	0.984	0.056	0.055	-0.016	-0.016	0.058	0.936	1
U55	1.5	1.497	0.075	0.074	-0.003	-0.002	0.074	0.953	1
U66	0.4	0.41	0.037	0.038	0.01	0.025	0.038	0.955	1

Table 6.8: Nonstationary PFA Results: True IC: Null, Fitted IC: deJong DKF, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.115	0.092	0.026	-0.085	-0.071	0.126	0.144	1
Z31	0.8	0.746	0.072	0.031	-0.054	-0.068	0.09	0.59	1
Z52	0.9	0.854	0.038	0.023	-0.046	-0.051	0.059	0.539	1
Z62	1.1	1.039	0.07	0.02	-0.061	-0.055	0.092	0.228	1
V11	1	1.147	0.138	0.096	0.147	0.147	0.201	0.694	0.998
V21	0.4	0.438	0.102	0.066	0.038	0.096	0.109	0.914	1
V22	1	1.073	0.171	0.089	0.073	0.073	0.185	0.816	0.982
T11	1.2	1.121	0.076	0.037	-0.079	-0.066	0.11	0.499	0.998
T21	0.6	0.578	0.049	0.037	-0.022	-0.037	0.054	0.922	0.998
T12	-0.4	-0.363	0.064	0.031	0.037	-0.094	0.074	0.827	0.987
T22	1	0.971	0.072	0.031	-0.029	-0.029	0.078	0.887	1
U11	0.8	0.78	0.166	0.05	-0.02	-0.025	0.167	0.894	1
U22	0.6	0.63	0.143	0.053	0.03	0.051	0.146	0.909	0.998
U33	2	1.999	0.111	0.095	-0.001	-0.001	0.11	0.945	1
U44	1	1.034	0.369	0.057	0.034	0.034	0.37	0.922	1
U55	1.5	1.471	0.211	0.073	-0.029	-0.02	0.213	0.942	0.982
U66	0.4	0.437	0.204	0.039	0.037	0.093	0.207	0.936	1

Table 6.9: Nonstationary PFA Results: True IC: Null, Fitted IC: Koopman exact initial KF, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.115	0.092	0.026	-0.085	-0.071	0.125	0.144	1
Z31	0.8	0.746	0.072	0.031	-0.054	-0.068	0.09	0.591	1
Z52	0.9	0.855	0.036	0.023	-0.045	-0.05	0.058	0.542	1
Z62	1.1	1.039	0.07	0.02	-0.061	-0.055	0.093	0.232	1
V11	1	1.147	0.137	0.096	0.147	0.147	0.201	0.692	0.998
V21	0.4	0.438	0.102	0.066	0.038	0.096	0.109	0.912	1
V22	1	1.072	0.171	0.089	0.072	0.072	0.185	0.819	0.982
T11	1.2	1.121	0.075	0.037	-0.079	-0.066	0.109	0.496	0.998
T21	0.6	0.578	0.048	0.037	-0.022	-0.037	0.053	0.927	0.998
T12	-0.4	-0.363	0.064	0.031	0.037	-0.094	0.074	0.827	0.987
T22	1	0.971	0.072	0.031	-0.029	-0.029	0.078	0.887	1
U11	0.8	0.78	0.166	0.05	-0.02	-0.025	0.167	0.894	1
U22	0.6	0.63	0.143	0.053	0.03	0.05	0.146	0.909	0.998
U33	2	1.999	0.111	0.095	-0.001	-0.001	0.111	0.942	1
U44	1	1.031	0.355	0.057	0.031	0.031	0.356	0.923	1
U55	1.5	1.476	0.205	0.073	-0.024	-0.016	0.206	0.942	0.985
U66	0.4	0.433	0.194	0.038	0.033	0.082	0.196	0.936	0.998

Table 6.10: Nonstationary PFA Results: True IC: Null, Fitted IC: Large κ, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.235	0.222	0.177	0.035	0.029	0.225	0.91	1
Z31	0.8	0.801	0.206	0.193	0.001	0.001	0.205	0.92	1
Z52	0.9	0.739	0.363	0.143	-0.161	-0.179	0.396	0.757	0.931
Z62	1.1	0.912	0.422	0.123	-0.188	-0.171	0.462	0.774	0.958
V11	0.05	0.064	0.04	0.019	0.014	0.289	0.043	0.74	0.861
V21	0.02	-0.135	0.483	0.075	-0.155	-7.756	0.507	0.833	0.465
V22	0.05	0.141	0.394	0.157	0.091	1.823	0.403	0.83	0.819
T11	1.2	1.021	0.437	0.121	-0.179	-0.149	0.471	0.681	0.972
T21	0.6	0.259	1.011	0.25	-0.341	-0.569	1.065	0.778	0.927
T12	-0.4	-0.333	0.224	0.117	0.067	-0.167	0.234	0.712	0.753
T22	0.7	0.586	0.321	0.154	-0.114	-0.163	0.34	0.823	0.875
U11	0.8	0.795	0.039	0.039	-0.005	-0.006	0.039	0.931	1
U22	0.6	0.59	0.039	0.034	-0.01	-0.016	0.04	0.924	1
U33	2	1.997	0.089	0.09	-0.003	-0.001	0.089	0.944	1
U44	1	0.962	0.361	0.07	-0.038	-0.038	0.363	0.861	0.951
U55	1.5	1.512	0.075	0.069	0.012	0.008	0.076	0.938	1
U66	0.4	0.416	0.049	0.023	0.016	0.04	0.051	0.757	1
X01	-	0	0.023	0.021	-1	-1	1	0	0.042
X02	-	-0.049	0.971	0.094					
P011	-	0.006	0.01	0.006					
P012	-	0.083	2.383	0.075					
P022	-	0.011	0.081	0.006					

Table 6.11: Nonstationary PFA Results: True IC: Null, Fitted IC: Free Parameter, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.229	0.199	0.178	0.029	0.024	0.201	0.934	1
Z31	0.8	0.8	0.209	0.199	0	0	0.209	0.941	0.997
Z52	0.9	0.687	0.392	0.138	-0.213	-0.236	0.445	0.716	0.848
Z62	1.1	0.842	0.466	0.121	-0.258	-0.235	0.532	0.716	0.868
V11	0.05	0.067	0.033	0.021	0.017	0.34	0.037	0.848	0.997
V21	0.02	-0.28	0.567	0.251	-0.3	-15.012	0.64	0.908	0.459
V22	0.05	0.071	0.264	1.007	0.021	0.411	0.264	0.97	0.726
T11	1.2	0.951	0.465	0.152	-0.249	-0.208	0.527	0.729	0.855
T21	0.6	-0.101	1.299	0.376	-0.701	-1.168	1.474	0.71	0.901
T12	-0.4	-0.319	0.197	0.126	0.081	-0.202	0.212	0.716	0.723
T22	0.7	0.539	0.399	0.225	-0.161	-0.23	0.429	0.924	0.809
U11	0.8	0.797	0.041	0.04	-0.003	-0.003	0.041	0.931	1
U22	0.6	0.594	0.031	0.033	-0.006	-0.009	0.032	0.954	1
U33	2	1.992	0.088	0.091	-0.008	-0.004	0.089	0.947	1
U44	1	1.01	0.074	0.064	0.01	0.01	0.075	0.937	1
U55	1.5	1.514	0.072	0.07	0.014	0.01	0.073	0.954	1
U66	0.4	0.418	0.044	0.023	0.018	0.045	0.047	0.703	1

Table 6.12: Nonstationary PFA Results: True IC: Null, Fitted IC: Null, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.197	0.059	0.056	-1.003	-0.836	1.004	0	0.975
Z31	0.8	0.134	0.067	0.07	-0.666	-0.832	0.669	0	0.456
Z52	0.9	0.18	0.08	0.064	-0.72	-0.8	0.725	0	0.785
Z62	1.1	0.188	0.056	0.046	-0.912	-0.829	0.914	0	1
V11	0.05	0.733	0.15	0.213	0.683	13.652	0.699	0.025	0.987
V21	0.02	0.098	0.036	0.038	0.078	3.902	0.086	0.456	0.747
V22	0.05	0.797	0.197	0.232	0.747	14.942	0.772	0	1
T11	1.2	0.138	0.057	0.059	-1.062	-0.885	1.063	0	0.684
T21	0.6	0.151	0.053	0.061	-0.449	-0.748	0.452	0	0.785
T12	-0.4	0.03	0.048	0.044	0.43	-1.075	0.432	0	0.127
T22	0.7	0.161	0.073	0.062	-0.539	-0.77	0.544	0	0.684
U11	0.8	0.208	0.151	0.211	-0.592	-0.74	0.611	0.025	0.241
U22	0.6	0.732	0.034	0.035	0.132	0.22	0.137	0	1
U33	2	2.052	0.08	0.092	0.052	0.026	0.095	0.962	1
U44	1	0.342	0.2	0.228	-0.658	-0.658	0.687	0.038	0.418
U55	1.5	1.572	0.075	0.072	0.072	0.048	0.104	0.861	1
U66	0.4	0.525	0.027	0.026	0.125	0.313	0.128	0	1

Table 6.13: Nonstationary PFA Results: True IC: Null, Fitted IC: deJong DKF, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.16	0.041	0.049	-1.04	-0.866	1.04	0	0.993
Z31	0.8	0.107	0.054	0.061	-0.693	-0.866	0.695	0	0.368
Z52	0.9	0.125	0.061	0.056	-0.775	-0.861	0.777	0	0.599
Z62	1.1	0.145	0.045	0.042	-0.955	-0.868	0.956	0	0.971
V11	0.05	0.879	0.115	0.269	0.829	16.583	0.837	0.02	0.987
V21	0.02	0.056	0.226	0.057	0.036	1.799	0.228	0.684	0.508
V22	0.05	0.976	0.234	0.36	0.926	18.516	0.955	0.052	0.954
T11	1.2	0.098	0.046	0.051	-1.102	-0.918	1.103	0	0.456
T21	0.6	0.091	0.097	0.059	-0.509	-0.849	0.518	0.01	0.472
T12	-0.4	0.014	0.041	0.039	0.414	-1.036	0.416	0	0.065
T22	0.7	0.101	0.099	0.061	-0.599	-0.856	0.607	0.01	0.375
U11	0.8	0.054	0.111	0.267	-0.746	-0.932	0.754	0.039	0.039
U22	0.6	0.744	0.034	0.034	0.144	0.241	0.148	0	1
U33	2	2.055	0.092	0.092	0.055	0.027	0.107	0.912	1
U44	1	0.165	0.256	0.334	-0.835	-0.835	0.873	0.078	0.156
U55	1.5	1.59	0.07	0.072	0.09	0.06	0.114	0.801	1
U66	0.4	0.532	0.026	0.025	0.132	0.33	0.135	0	1

Table 6.14: Nonstationary PFA Results: True IC: Null, Fitted IC: Koopman exact initial KF, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.161	0.042	0.049	-1.039	-0.866	1.04	0	0.994
Z31	0.8	0.108	0.056	0.061	-0.692	-0.865	0.694	0	0.368
Z52	0.9	0.123	0.064	0.056	-0.777	-0.864	0.78	0	0.568
Z62	1.1	0.143	0.046	0.043	-0.957	-0.87	0.958	0	0.946
V11	0.05	0.876	0.122	0.268	0.826	16.519	0.835	0.019	0.99
V21	0.02	0.048	0.234	0.062	0.028	1.383	0.235	0.676	0.505
V22	0.05	0.964	0.263	0.38	0.914	18.29	0.951	0.079	0.93
T11	1.2	0.098	0.047	0.051	-1.102	-0.918	1.103	0	0.454
T21	0.6	0.086	0.104	0.061	-0.514	-0.856	0.524	0.01	0.463
T12	-0.4	0.014	0.042	0.039	0.414	-1.034	0.416	0	0.063
T22	0.7	0.097	0.102	0.065	-0.603	-0.861	0.611	0.013	0.365
U11	0.8	0.057	0.119	0.266	-0.743	-0.928	0.752	0.054	0.054
U22	0.6	0.744	0.034	0.034	0.144	0.24	0.148	0	1
U33	2	2.056	0.093	0.092	0.056	0.028	0.108	0.908	1
U44	1	0.171	0.262	0.343	-0.829	-0.829	0.869	0.102	0.159
U55	1.5	1.589	0.071	0.072	0.089	0.059	0.114	0.787	1
U66	0.4	0.531	0.027	0.025	0.131	0.328	0.134	0	1

Table 6.15: Nonstationary PFA Results: True IC: Null, Fitted IC: Large κ, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.01	0.009	0.001	0.001	0.01	0.944	1
Z31	0.8	0.8	0.011	0.011	0	0	0.011	0.96	1
Z52	0.9	0.933	0.197	0.015	0.033	0.037	0.2	0.916	0.996
Z62	1.1	1.151	0.235	0.021	0.051	0.046	0.24	0.888	0.996
V11	0.05	0.114	0.233	0.012	0.064	1.272	0.241	0.88	1
V21	0.02	-0.006	0.208	0.011	-0.026	-1.287	0.209	0.848	0.776
V22	0.05	0.046	0.017	0.013	-0.004	-0.072	0.017	0.9	0.928
T11	1.2	1.198	0.009	0.004	-0.002	-0.001	0.009	0.88	1
T21	0.6	0.6	0.008	0.006	0	-0.001	0.008	0.916	1
T12	-0.4	-0.414	0.063	0.009	-0.014	0.035	0.065	0.912	0.996
T22	0.7	0.732	0.162	0.012	0.032	0.046	0.165	0.888	0.996
U11	0.8	0.804	0.062	0.039	0.004	0.006	0.062	0.884	1
U22	0.6	0.569	0.11	0.03	-0.031	-0.052	0.114	0.904	0.992
U33	2	2.006	0.102	0.091	0.006	0.003	0.101	0.94	1
U44	1	1.098	0.446	0.051	0.098	0.098	0.456	0.864	1
U55	1.5	1.453	0.296	0.068	-0.047	-0.031	0.299	0.868	0.964
U66	0.4	0.462	0.411	0.025	0.062	0.154	0.415	0.864	0.964
X01	1	0.996	0.247	0.244	-0.004	-0.004	0.246	0.944	0.976
X02	0.5	0.489	0.483	0.253	-0.011	-0.022	0.482	0.932	0.68
P011	1.2	1.182	0.654	0.395	-0.018	-0.015	0.653	0.812	0.996
P012	0.3	0.313	0.924	0.314	0.013	0.042	0.922	0.876	0.148
P022	0.7	0.642	0.293	0.318	-0.058	-0.083	0.298	0.824	0.952

Table 6.16: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Free Parameter, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.203	0.042	0.01	0.003	0.002	0.042	0.936	1
Z31	0.8	0.801	0.02	0.011	0.001	0.001	0.02	0.968	1
Z52	0.9	0.896	0.054	0.008	-0.004	-0.004	0.054	0.928	1
Z62	1.1	1.095	0.065	0.007	-0.005	-0.004	0.065	0.924	1
V11	0.05	0.156	0.114	0.014	0.106	2.111	0.155	0	1
V21	0.02	0.096	0.23	0.012	0.076	3.784	0.241	0	1
V22	0.05	0.13	0.024	0.019	0.08	1.592	0.083	0.004	0.996
T11	1.2	1.2	0.01	0.004	0	0	0.01	0.98	1
T21	0.6	0.579	0.346	0.008	-0.021	-0.035	0.346	0.952	1
T12	-0.4	-0.399	0.016	0.005	0.001	-0.003	0.016	0.972	1
T22	0.7	0.697	0.032	0.004	-0.003	-0.004	0.033	0.976	1
U11	0.8	0.871	0.547	0.042	0.071	0.089	0.55	0.843	1
U22	0.6	0.679	0.268	0.035	0.079	0.132	0.279	0.594	1
U33	2	2.043	0.246	0.093	0.043	0.022	0.25	0.948	1
U44	1	1.047	0.462	0.055	0.047	0.047	0.463	0.928	1
U55	1.5	1.513	0.072	0.07	0.013	0.009	0.073	0.956	1
U66	0.4	0.431	0.026	0.024	0.031	0.077	0.04	0.763	1

Table 6.17: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Null, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.199	0.01	0.009	-0.001	-0.001	0.01	0.946	1
Z31	0.8	0.799	0.011	0.011	-0.001	-0.001	0.011	0.953	1
Z52	0.9	0.899	0.009	0.008	-0.001	-0.001	0.009	0.941	1
Z62	1.1	1.099	0.007	0.007	-0.001	-0.001	0.007	0.928	1
V11	0.05	0.05	0.008	0.008	0	-0.008	0.008	0.946	1
V21	0.02	0.02	0.009	0.008	0	0.008	0.009	0.917	0.747
V22	0.05	0.05	0.012	0.01	0	0.007	0.012	0.928	0.99
T11	1.2	1.2	0.003	0.003	0	0	0.003	0.938	1
T21	0.6	0.6	0.006	0.005	0	0	0.006	0.943	1
T12	-0.4	-0.4	0.004	0.004	0	0	0.004	0.951	1
T22	0.7	0.7	0.003	0.003	0	-0.001	0.003	0.941	1
U11	0.8	0.799	0.038	0.038	-0.001	-0.001	0.038	0.941	1
U22	0.6	0.596	0.03	0.03	-0.004	-0.007	0.03	0.948	1
U33	2	2.005	0.096	0.091	0.005	0.002	0.096	0.943	1
U44	1	1	0.048	0.047	0	0	0.048	0.941	1
U55	1.5	1.495	0.072	0.069	-0.005	-0.004	0.072	0.946	1
U66	0.4	0.399	0.023	0.022	-0.001	-0.002	0.023	0.933	1

Table 6.18: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: deJong DKF, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.01	0.009	0.001	0.001	0.01	0.916	1
Z31	0.8	0.799	0.011	0.011	-0.001	-0.001	0.011	0.947	1
Z52	0.9	0.876	0.133	0.014	-0.024	-0.027	0.135	0.877	1
Z62	1.1	1.057	0.169	0.017	-0.043	-0.039	0.174	0.825	1
V11	0.05	0.175	0.314	0.016	0.125	2.492	0.338	0.818	1
V21	0.02	-0.039	0.381	0.015	-0.059	-2.952	0.385	0.789	0.796
V22	0.05	0.043	0.02	0.016	-0.007	-0.148	0.021	0.895	0.86
T11	1.2	1.196	0.011	0.004	-0.004	-0.003	0.012	0.839	1
T21	0.6	0.586	0.227	0.007	-0.014	-0.024	0.227	0.912	1
T12	-0.4	-0.386	0.052	0.007	0.014	-0.035	0.054	0.86	0.996
T22	0.7	0.667	0.143	0.01	-0.033	-0.047	0.146	0.849	1
U11	0.8	0.809	0.083	0.039	0.009	0.012	0.084	0.853	1
U22	0.6	0.545	0.143	0.03	-0.055	-0.092	0.153	0.835	0.979
U33	2	2.012	0.111	0.091	0.012	0.006	0.112	0.912	1
U44	1	1.228	0.678	0.056	0.228	0.228	0.714	0.804	1
U55	1.5	1.375	0.448	0.064	-0.125	-0.083	0.464	0.811	0.909
U66	0.4	0.578	0.63	0.03	0.178	0.444	0.654	0.804	0.954

Table 6.19: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Koopman exact initial KF, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.011	0.009	0.001	0.001	0.011	0.91	1
Z31	0.8	0.799	0.011	0.011	-0.001	-0.001	0.011	0.95	1
Z52	0.9	0.878	0.111	0.016	-0.022	-0.024	0.113	0.878	1
Z62	1.1	1.062	0.143	0.018	-0.038	-0.035	0.148	0.839	1
V11	0.05	0.18	0.325	0.017	0.13	2.595	0.349	0.806	1
V21	0.02	-0.062	0.31	0.014	-0.082	-4.123	0.32	0.781	0.792
V22	0.05	0.042	0.02	0.015	-0.008	-0.151	0.021	0.892	0.857
T11	1.2	1.196	0.012	0.004	-0.004	-0.003	0.013	0.835	1
T21	0.6	0.599	0.011	0.007	-0.001	-0.002	0.011	0.921	1
T12	-0.4	-0.389	0.038	0.008	0.011	-0.029	0.04	0.867	1
T22	0.7	0.673	0.103	0.011	-0.027	-0.039	0.106	0.853	1
U11	0.8	0.811	0.09	0.039	0.011	0.013	0.09	0.849	1
U22	0.6	0.543	0.147	0.03	-0.057	-0.095	0.157	0.832	0.975
U33	2	2.012	0.112	0.091	0.012	0.006	0.113	0.91	1
U44	1	1.244	0.696	0.057	0.244	0.244	0.737	0.803	1
U55	1.5	1.349	0.479	0.063	-0.151	-0.101	0.501	0.814	0.892
U66	0.4	0.618	0.68	0.031	0.218	0.544	0.713	0.803	0.964

Table 6.20: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Large κ, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.202	0.021	0.021	0.002	0.002	0.021	0.95	1
Z31	0.8	0.8	0.025	0.024	0	0	0.025	0.95	1
Z52	0.9	0.915	0.157	0.019	0.015	0.017	0.158	0.935	1
Z62	1.1	1.114	0.149	0.016	0.014	0.013	0.15	0.916	1
V11	1	1	0.105	0.082	0	0	0.105	0.956	1
V21	0.4	0.395	0.097	0.057	-0.005	-0.012	0.097	0.95	1
V22	1	0.992	0.121	0.076	-0.008	-0.008	0.121	0.937	0.992
T11	1.2	1.197	0.032	0.025	-0.003	-0.002	0.032	0.927	1
T21	0.6	0.602	0.029	0.026	0.002	0.003	0.029	0.948	1
T12	-0.4	-0.402	0.039	0.02	-0.002	0.005	0.039	0.953	1
T22	1	1.01	0.114	0.021	0.01	0.01	0.114	0.958	1
U11	0.8	0.798	0.053	0.047	-0.002	-0.003	0.053	0.943	1
U22	0.6	0.596	0.064	0.049	-0.004	-0.007	0.064	0.94	0.997
U33	2	1.994	0.098	0.095	-0.006	-0.003	0.098	0.945	1
U44	1	1.034	0.282	0.055	0.034	0.034	0.284	0.932	1
U55	1.5	1.482	0.156	0.073	-0.018	-0.012	0.157	0.932	0.992
U66	0.4	0.415	0.164	0.035	0.015	0.036	0.165	0.943	0.997
X01	1	0.997	0.082	0.084	-0.003	-0.003	0.082	0.961	1
X02	0.5	0.426	1.655	0.131	-0.074	-0.149	1.655	0.93	0.995
P011	1.2	1.215	0.321	0.144	0.015	0.013	0.321	0.93	1
P012	0.3	0.378	8.382	0.34	0.078	0.261	8.371	0.914	0.971
P022	0.7	0.658	0.171	0.577	-0.042	-0.06	0.176	0.911	0.958

Table 6.21: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Free Parameter, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.185	0.021	0.03	-0.015	-0.012	0.026	0.979	1
Z31	0.8	0.817	0.026	0.029	0.017	0.022	0.031	0.948	1
Z52	0.9	0.901	0.029	0.016	0.001	0.001	0.029	0.966	1
Z62	1.1	1.092	0.033	0.015	-0.008	-0.007	0.034	0.953	1
V11	1	1.604	0.125	0.123	0.604	0.604	0.617	0.003	1
V21	0.4	0.986	0.26	0.091	0.586	1.466	0.641	0	1
V22	1	1.737	0.142	0.117	0.737	0.737	0.75	0.003	0.997
T11	1.2	1.229	0.03	0.038	0.029	0.024	0.042	0.94	1
T21	0.6	0.598	0.079	0.039	-0.002	-0.004	0.079	0.99	1
T12	-0.4	-0.412	0.019	0.025	-0.012	0.031	0.022	0.977	1
T22	1	1.002	0.081	0.024	0.002	0.002	0.081	0.984	1
U11	0.8	1.363	0.079	0.07	0.563	0.703	0.568	0	1
U22	0.6	1.57	0.102	0.081	0.97	1.616	0.975	0	1
U33	2	2.248	0.116	0.108	0.248	0.124	0.274	0.369	1
U44	1	1.201	0.065	0.062	0.201	0.201	0.211	0.081	1
U55	1.5	1.629	0.085	0.08	0.129	0.086	0.154	0.673	1
U66	0.4	0.801	0.047	0.044	0.401	1.001	0.403	0	1

Table 6.22: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Null, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.158	0.019	0.019	-0.042	-0.035	0.046	0.43	1
Z31	0.8	0.772	0.022	0.023	-0.028	-0.035	0.036	0.777	1
Z52	0.9	0.883	0.014	0.015	-0.017	-0.019	0.022	0.799	1
Z62	1.1	1.08	0.012	0.012	-0.02	-0.018	0.023	0.634	1
V11	1	1.063	0.078	0.086	0.063	0.063	0.1	0.927	1
V21	0.4	0.417	0.056	0.059	0.017	0.042	0.058	0.955	1
V22	1	1.032	0.074	0.079	0.032	0.032	0.081	0.953	1
T11	1.2	1.168	0.025	0.024	-0.032	-0.027	0.04	0.764	1
T21	0.6	0.59	0.024	0.025	-0.01	-0.017	0.026	0.944	1
T12	-0.4	-0.389	0.018	0.018	0.011	-0.028	0.021	0.919	1
T22	1	0.996	0.018	0.017	-0.004	-0.004	0.018	0.944	1
U11	0.8	0.785	0.048	0.048	-0.015	-0.019	0.05	0.938	1
U22	0.6	0.613	0.051	0.051	0.013	0.022	0.053	0.942	1
U33	2	1.995	0.096	0.095	-0.005	-0.003	0.097	0.936	1
U44	1	0.992	0.051	0.054	-0.008	-0.008	0.052	0.955	1
U55	1.5	1.495	0.074	0.074	-0.005	-0.003	0.074	0.946	1
U66	0.4	0.403	0.037	0.037	0.003	0.007	0.037	0.931	1

Table 6.23: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: deJong DKF, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.16	0.022	0.019	-0.04	-0.033	0.045	0.443	1
Z31	0.8	0.772	0.023	0.023	-0.028	-0.035	0.036	0.759	1
Z52	0.9	0.876	0.045	0.015	-0.024	-0.027	0.051	0.778	1
Z62	1.1	1.066	0.076	0.013	-0.034	-0.031	0.083	0.616	1
V11	1	1.102	0.222	0.088	0.102	0.102	0.244	0.895	1
V21	0.4	0.424	0.218	0.062	0.024	0.06	0.219	0.914	0.992
V22	1	0.988	0.223	0.08	-0.012	-0.012	0.223	0.908	0.957
T11	1.2	1.16	0.045	0.025	-0.04	-0.033	0.06	0.735	1
T21	0.6	0.592	0.03	0.025	-0.008	-0.014	0.031	0.916	1
T12	-0.4	-0.378	0.055	0.019	0.022	-0.054	0.059	0.881	0.995
T22	1	0.982	0.077	0.018	-0.018	-0.018	0.079	0.911	1
U11	0.8	0.793	0.068	0.048	-0.007	-0.008	0.068	0.924	1
U22	0.6	0.599	0.099	0.051	-0.001	-0.002	0.099	0.905	0.986
U33	2	2.001	0.098	0.095	0.001	0.001	0.097	0.93	1
U44	1	1.116	0.62	0.059	0.116	0.116	0.63	0.908	1
U55	1.5	1.449	0.295	0.071	-0.051	-0.034	0.299	0.919	0.965
U66	0.4	0.468	0.345	0.038	0.068	0.171	0.351	0.9	0.997

Table 6.24: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Koopman exact initial KF, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.16	0.022	0.019	-0.04	-0.033	0.045	0.446	1
Z31	0.8	0.772	0.022	0.023	-0.028	-0.035	0.036	0.761	1
Z52	0.9	0.876	0.049	0.016	-0.024	-0.026	0.054	0.777	1
Z62	1.1	1.066	0.077	0.013	-0.034	-0.031	0.084	0.621	1
V11	1	1.106	0.248	0.088	0.106	0.106	0.27	0.892	1
V21	0.4	0.43	0.295	0.062	0.03	0.075	0.296	0.911	0.995
V22	1	0.986	0.229	0.081	-0.014	-0.014	0.229	0.906	0.954
T11	1.2	1.16	0.044	0.025	-0.04	-0.033	0.06	0.734	1
T21	0.6	0.591	0.031	0.026	-0.009	-0.015	0.032	0.914	1
T12	-0.4	-0.379	0.054	0.019	0.021	-0.053	0.058	0.882	0.995
T22	1	0.982	0.079	0.018	-0.018	-0.018	0.081	0.914	1
U11	0.8	0.791	0.065	0.048	-0.009	-0.011	0.065	0.927	1
U22	0.6	0.601	0.094	0.051	0.001	0.001	0.094	0.906	0.989
U33	2	2.001	0.098	0.095	0.001	0.001	0.098	0.933	1
U44	1	1.133	0.69	0.06	0.133	0.133	0.702	0.906	1
U55	1.5	1.45	0.294	0.072	-0.05	-0.034	0.298	0.917	0.965
U66	0.4	0.468	0.347	0.039	0.068	0.17	0.353	0.901	0.997

Table 6.25: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Large κ, N=200, T=5, Moderate Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.235	0.222	0.177	0.035	0.029	0.225	0.91	1
Z31	0.8	0.801	0.206	0.193	0.001	0.001	0.205	0.92	1
Z52	0.9	0.739	0.363	0.143	-0.161	-0.179	0.396	0.757	0.931
Z62	1.1	0.912	0.422	0.123	-0.188	-0.171	0.462	0.774	0.958
V11	0.05	0.064	0.04	0.019	0.014	0.289	0.043	0.74	0.861
V21	0.02	-0.135	0.483	0.075	-0.155	-7.756	0.507	0.833	0.465
V22	0.05	0.141	0.394	0.157	0.091	1.823	0.403	0.83	0.819
T11	1.2	1.021	0.437	0.121	-0.179	-0.149	0.471	0.681	0.972
T21	0.6	0.259	1.011	0.25	-0.341	-0.569	1.065	0.778	0.927
T12	-0.4	-0.333	0.224	0.117	0.067	-0.167	0.234	0.712	0.753
T22	0.7	0.586	0.321	0.154	-0.114	-0.163	0.34	0.823	0.875
U11	0.8	0.795	0.039	0.039	-0.005	-0.006	0.039	0.931	1
U22	0.6	0.59	0.039	0.034	-0.01	-0.016	0.04	0.924	1
U33	2	1.997	0.089	0.09	-0.003	-0.001	0.089	0.944	1
U44	1	0.962	0.361	0.07	-0.038	-0.038	0.363	0.861	0.951
U55	1.5	1.512	0.075	0.069	0.012	0.008	0.076	0.938	1
U66	0.4	0.416	0.049	0.023	0.016	0.04	0.051	0.757	1
X01	1	0	0.023	0.021	-1	-1	1	0	0.042
X02	0.5	-0.049	0.971	0.094	-0.549	-1.099	1.114	0.038	0.066
P011	1.2	0.006	0.01	0.006	-1.194	-0.995	1.194	0	0.028
P012	0.3	0.083	2.383	0.075	-0.217	-0.723	2.389	0.028	0.115
P022	0.7	0.011	0.081	0.006	-0.689	-0.984	0.694	0	0.031

Table 6.26: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Free Parameter, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.229	0.199	0.178	0.029	0.024	0.201	0.934	1
Z31	0.8	0.8	0.209	0.199	0	0	0.209	0.941	0.997
Z52	0.9	0.687	0.392	0.138	-0.213	-0.236	0.445	0.716	0.848
Z62	1.1	0.842	0.466	0.121	-0.258	-0.235	0.532	0.716	0.868
V11	0.05	0.067	0.033	0.021	0.017	0.34	0.037	0.848	0.997
V21	0.02	-0.28	0.567	0.251	-0.3	-15.012	0.64	0.908	0.459
V22	0.05	0.071	0.264	1.007	0.021	0.411	0.264	0.97	0.726
T11	1.2	0.951	0.465	0.152	-0.249	-0.208	0.527	0.729	0.855
T21	0.6	-0.101	1.299	0.376	-0.701	-1.168	1.474	0.71	0.901
T12	-0.4	-0.319	0.197	0.126	0.081	-0.202	0.212	0.716	0.723
T22	0.7	0.539	0.399	0.225	-0.161	-0.23	0.429	0.924	0.809
U11	0.8	0.797	0.041	0.04	-0.003	-0.003	0.041	0.931	1
U22	0.6	0.594	0.031	0.033	-0.006	-0.009	0.032	0.954	1
U33	2	1.992	0.088	0.091	-0.008	-0.004	0.089	0.947	1
U44	1	1.01	0.074	0.064	0.01	0.01	0.075	0.937	1
U55	1.5	1.514	0.072	0.07	0.014	0.01	0.073	0.954	1
U66	0.4	0.418	0.044	0.023	0.018	0.045	0.047	0.703	1

Table 6.27: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Null, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.197	0.059	0.056	-1.003	-0.836	1.004	0	0.975
Z31	0.8	0.134	0.067	0.07	-0.666	-0.832	0.669	0	0.456
Z52	0.9	0.18	0.08	0.064	-0.72	-0.8	0.725	0	0.785
Z62	1.1	0.188	0.056	0.046	-0.912	-0.829	0.914	0	1
V11	0.05	0.733	0.15	0.213	0.683	13.652	0.699	0.025	0.987
V21	0.02	0.098	0.036	0.038	0.078	3.902	0.086	0.456	0.747
V22	0.05	0.797	0.197	0.232	0.747	14.942	0.772	0	1
T11	1.2	0.138	0.057	0.059	-1.062	-0.885	1.063	0	0.684
T21	0.6	0.151	0.053	0.061	-0.449	-0.748	0.452	0	0.785
T12	-0.4	0.03	0.048	0.044	0.43	-1.075	0.432	0	0.127
T22	0.7	0.161	0.073	0.062	-0.539	-0.77	0.544	0	0.684
U11	0.8	0.208	0.151	0.211	-0.592	-0.74	0.611	0.025	0.241
U22	0.6	0.732	0.034	0.035	0.132	0.22	0.137	0	1
U33	2	2.052	0.08	0.092	0.052	0.026	0.095	0.962	1
U44	1	0.342	0.2	0.228	-0.658	-0.658	0.687	0.038	0.418
U55	1.5	1.572	0.075	0.072	0.072	0.048	0.104	0.861	1
U66	0.4	0.525	0.027	0.026	0.125	0.313	0.128	0	1

Table 6.28: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: deJong DKF, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.16	0.041	0.049	-1.04	-0.866	1.04	0	0.993
Z31	0.8	0.107	0.054	0.061	-0.693	-0.866	0.695	0	0.368
Z52	0.9	0.125	0.061	0.056	-0.775	-0.861	0.777	0	0.599
Z62	1.1	0.145	0.045	0.042	-0.955	-0.868	0.956	0	0.971
V11	0.05	0.879	0.115	0.269	0.829	16.583	0.837	0.02	0.987
V21	0.02	0.056	0.226	0.057	0.036	1.799	0.228	0.684	0.508
V22	0.05	0.976	0.234	0.36	0.926	18.516	0.955	0.052	0.954
T11	1.2	0.098	0.046	0.051	-1.102	-0.918	1.103	0	0.456
T21	0.6	0.091	0.097	0.059	-0.509	-0.849	0.518	0.01	0.472
T12	-0.4	0.014	0.041	0.039	0.414	-1.036	0.416	0	0.065
T22	0.7	0.101	0.099	0.061	-0.599	-0.856	0.607	0.01	0.375
U11	0.8	0.054	0.111	0.267	-0.746	-0.932	0.754	0.039	0.039
U22	0.6	0.744	0.034	0.034	0.144	0.241	0.148	0	1
U33	2	2.055	0.092	0.092	0.055	0.027	0.107	0.912	1
U44	1	0.165	0.256	0.334	-0.835	-0.835	0.873	0.078	0.156
U55	1.5	1.59	0.07	0.072	0.09	0.06	0.114	0.801	1
U66	0.4	0.532	0.026	0.025	0.132	0.33	0.135	0	1

Table 6.29: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Koopman exact initial KF, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	0.161	0.042	0.049	-1.039	-0.866	1.04	0	0.994
Z31	0.8	0.108	0.056	0.061	-0.692	-0.865	0.694	0	0.368
Z52	0.9	0.123	0.064	0.056	-0.777	-0.864	0.78	0	0.568
Z62	1.1	0.143	0.046	0.043	-0.957	-0.87	0.958	0	0.946
V11	0.05	0.876	0.122	0.268	0.826	16.519	0.835	0.019	0.99
V21	0.02	0.048	0.234	0.062	0.028	1.383	0.235	0.676	0.505
V22	0.05	0.964	0.263	0.38	0.914	18.29	0.951	0.079	0.93
T11	1.2	0.098	0.047	0.051	-1.102	-0.918	1.103	0	0.454
T21	0.6	0.086	0.104	0.061	-0.514	-0.856	0.524	0.01	0.463
T12	-0.4	0.014	0.042	0.039	0.414	-1.034	0.416	0	0.063
T22	0.7	0.097	0.102	0.065	-0.603	-0.861	0.611	0.013	0.365
U11	0.8	0.057	0.119	0.266	-0.743	-0.928	0.752	0.054	0.054
U22	0.6	0.744	0.034	0.034	0.144	0.24	0.148	0	1
U33	2	2.056	0.093	0.092	0.056	0.028	0.108	0.908	1
U44	1	0.171	0.262	0.343	-0.829	-0.829	0.869	0.102	0.159
U55	1.5	1.589	0.071	0.072	0.089	0.059	0.114	0.787	1
U66	0.4	0.531	0.027	0.025	0.131	0.328	0.134	0	1

Table 6.30: Nonstationary PFA Results: True IC: Free Parameter, Fitted IC: Large κ, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.006	0.006	0.001	0.001	0.007	0.953	1
Z31	0.8	0.8	0.008	0.008	0	0.001	0.008	0.953	1
Z52	0.9	0.989	0.263	0.028	0.089	0.099	0.277	0.853	1
Z62	1.1	1.206	0.302	0.033	0.106	0.096	0.32	0.816	1
V11	0.05	0.189	0.33	0.02	0.139	2.78	0.357	0.805	0.995
V21	0.02	-0.104	0.358	0.016	-0.124	-6.205	0.378	0.795	0.779
V22	0.05	0.04	0.021	0.019	-0.01	-0.193	0.023	0.905	0.837
T11	1.2	1.198	0.006	0.003	-0.002	-0.002	0.006	0.879	1
T21	0.6	0.599	0.009	0.005	-0.001	-0.002	0.009	0.921	1
T12	-0.4	-0.421	0.058	0.01	-0.021	0.051	0.061	0.837	1
T22	0.7	0.775	0.217	0.022	0.075	0.107	0.229	0.858	1
U11	0.8	0.812	0.096	0.04	0.012	0.015	0.097	0.847	1
U22	0.6	0.544	0.151	0.03	-0.056	-0.093	0.161	0.774	0.974
U33	2	2.01	0.097	0.091	0.01	0.005	0.097	0.921	1
U44	1	1.335	0.882	0.06	0.335	0.335	0.941	0.768	1
U55	1.5	1.308	0.529	0.063	-0.192	-0.128	0.562	0.795	0.863
U66	0.4	0.686	0.75	0.034	0.286	0.714	0.801	0.763	0.974
X01	-	0.013	0.43						
X02	-	0.025	0.439						
P011	-	3.6	0.848						
P012	-	1.745	0.698						
P022	-	3.17	1.076						

Table 6.31: Nonstationary PFA Results: True IC: Diffuse, Fitted IC: Free Parameter, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.2	0.006	0.007	0	0	0.006	0.99	1
Z31	0.8	0.801	0.008	0.008	0.001	0.001	0.008	0.953	1
Z52	0.9	0.9	0.006	0.006	0	0	0.006	0.964	1
Z62	1.1	1.1	0.005	0.005	0	0	0.005	0.943	1
V11	0.05	0.209	0.025	0.018	0.159	3.184	0.161	0	1
V21	0.02	0.142	0.025	0.015	0.122	6.112	0.125	0	1
V22	0.05	0.205	0.031	0.018	0.155	3.099	0.158	0	1
T11	1.2	1.199	0.002	0.003	-0.001	-0.001	0.002	1	1
T21	0.6	0.6	0.004	0.005	0	-0.001	0.004	0.974	1
T12	-0.4	-0.4	0.003	0.004	0	-0.001	0.003	0.99	1
T22	0.7	0.7	0.002	0.003	0	0	0.002	0.99	1
U11	0.8	0.869	0.046	0.043	0.069	0.086	0.083	0.677	1
U22	0.6	0.725	0.049	0.039	0.125	0.208	0.134	0.141	1
U33	2	2.034	0.097	0.094	0.034	0.017	0.103	0.927	1
U44	1	1.071	0.053	0.052	0.071	0.071	0.089	0.745	1
U55	1.5	1.562	0.072	0.073	0.062	0.042	0.096	0.896	1
U66	0.4	0.527	0.034	0.029	0.127	0.318	0.132	0.016	1

Table 6.32: Nonstationary PFA Results: True IC: Diffuse, Fitted IC: Null, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.2	0.006	0.007	0	0	0.006	0.976	1
Z31	0.8	0.8	0.008	0.008	0	0	0.008	0.946	1
Z52	0.9	0.9	0.006	0.006	0	0	0.006	0.961	1
Z62	1.1	1.1	0.005	0.005	0	0	0.005	0.94	1
V11	0.05	0.049	0.008	0.008	-0.001	-0.015	0.008	0.943	1
V21	0.02	0.02	0.008	0.008	0	-0.002	0.008	0.946	0.752
V22	0.05	0.049	0.011	0.01	-0.001	-0.011	0.011	0.94	0.997
T11	1.2	1.2	0.002	0.002	0	0	0.002	0.943	1
T21	0.6	0.6	0.004	0.004	0	0	0.004	0.943	1
T12	-0.4	-0.4	0.003	0.003	0	-0.001	0.003	0.946	1
T22	0.7	0.7	0.002	0.002	0	0	0.002	0.961	1
U11	0.8	0.798	0.039	0.038	-0.002	-0.003	0.04	0.925	1
U22	0.6	0.6	0.031	0.03	0	0	0.031	0.931	1
U33	2	2.001	0.085	0.091	0.001	0	0.085	0.961	1
U44	1	0.999	0.047	0.047	-0.001	-0.001	0.046	0.943	1
U55	1.5	1.504	0.065	0.069	0.004	0.002	0.065	0.958	1
U66	0.4	0.4	0.023	0.022	0	0	0.023	0.922	1

Table 6.33: Nonstationary PFA Results: True IC: Diffuse, Fitted IC: deJong DKF, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.2	0.007	0.006	0	0	0.007	0.946	1
Z31	0.8	0.8	0.008	0.008	0	0	0.008	0.955	1
Z52	0.9	0.886	0.08	0.014	-0.014	-0.016	0.081	0.897	1
Z62	1.1	1.074	0.103	0.018	-0.026	-0.024	0.106	0.848	1
V11	0.05	0.211	0.35	0.019	0.161	3.211	0.385	0.781	1
V21	0.02	-0.067	0.416	0.017	-0.087	-4.34	0.424	0.772	0.795
V22	0.05	0.052	0.179	0.019	0.002	0.046	0.179	0.888	0.821
T11	1.2	1.197	0.007	0.003	-0.003	-0.002	0.007	0.862	1
T21	0.6	0.583	0.25	0.006	-0.017	-0.029	0.25	0.933	1
T12	-0.4	-0.391	0.041	0.006	0.009	-0.022	0.042	0.848	1
T22	0.7	0.678	0.116	0.011	-0.022	-0.031	0.117	0.875	1
U11	0.8	0.819	0.111	0.04	0.019	0.024	0.112	0.817	1
U22	0.6	0.529	0.165	0.03	-0.071	-0.119	0.179	0.763	0.964
U33	2	2.008	0.102	0.091	0.008	0.004	0.102	0.915	1
U44	1	1.303	0.77	0.059	0.303	0.303	0.826	0.754	1
U55	1.5	1.33	0.513	0.063	-0.17	-0.114	0.54	0.799	0.875
U66	0.4	0.65	0.724	0.033	0.25	0.624	0.765	0.741	0.951

Table 6.34: Nonstationary PFA Results: True IC: Diffuse, Fitted IC: Koopman exact initial KF, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\widehat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.007	0.007	0.001	0.001	0.007	0.951	1
Z31	0.8	0.8	0.008	0.008	0	0	0.008	0.938	1
Z52	0.9	0.885	0.08	0.015	-0.015	-0.016	0.081	0.898	1
Z62	1.1	1.06	0.207	0.02	-0.04	-0.036	0.21	0.841	1
V11	0.05	0.23	0.372	0.021	0.18	3.604	0.413	0.765	0.996
V21	0.02	-0.083	0.429	0.018	-0.103	-5.139	0.44	0.757	0.783
V22	0.05	0.039	0.027	0.019	-0.011	-0.211	0.029	0.885	0.796
T11	1.2	1.197	0.007	0.003	-0.003	-0.003	0.008	0.854	1
T21	0.6	0.582	0.25	0.006	-0.018	-0.03	0.25	0.938	1
T12	-0.4	-0.392	0.047	0.007	0.008	-0.019	0.047	0.841	1
T22	0.7	0.669	0.161	0.013	-0.031	-0.044	0.163	0.872	1
U11	0.8	0.825	0.124	0.041	0.025	0.032	0.126	0.801	1
U22	0.6	0.519	0.179	0.03	-0.081	-0.136	0.196	0.748	0.947
U33	2	2.013	0.109	0.092	0.013	0.007	0.109	0.903	1
U44	1	1.337	0.806	0.061	0.337	0.337	0.872	0.743	1
U55	1.5	1.321	0.561	0.063	-0.179	-0.119	0.587	0.779	0.863
U66	0.4	0.677	0.76	0.034	0.277	0.693	0.808	0.726	0.947

Table 6.35: Nonstationary PFA Results: True IC: Diffuse, Fitted IC: Large κ, N=20, T=50, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.201	0.009	0.008	0.001	0.001	0.009	0.92	1
Z31	0.8	0.8	0.01	0.01	0	0	0.01	0.962	1
Z52	0.9	0.978	0.261	0.022	0.078	0.087	0.272	0.803	1
Z62	1.1	1.189	0.307	0.024	0.089	0.081	0.319	0.864	1
V11	1	1.188	0.447	0.088	0.188	0.188	0.484	0.789	1
V21	0.4	0.306	0.683	0.066	-0.094	-0.234	0.687	0.798	0.981
V22	1	0.818	0.383	0.082	-0.182	-0.182	0.424	0.742	0.826
T11	1.2	1.196	0.013	0.009	-0.004	-0.003	0.013	0.869	1
T21	0.6	0.598	0.015	0.011	-0.002	-0.004	0.015	0.911	1
T12	-0.4	-0.425	0.104	0.013	-0.025	0.063	0.107	0.854	1
T22	1	1.08	0.277	0.024	0.08	0.08	0.288	0.859	1
U11	0.8	0.83	0.096	0.049	0.03	0.038	0.101	0.911	1
U22	0.6	0.543	0.161	0.048	-0.057	-0.094	0.171	0.812	0.962
U33	2	2.008	0.109	0.095	0.008	0.004	0.109	0.939	1
U44	1	1.724	1.727	0.083	0.724	0.724	1.869	0.798	1
U55	1.5	1.313	0.538	0.066	-0.187	-0.124	0.569	0.789	0.869
U66	0.4	0.676	0.74	0.047	0.276	0.689	0.788	0.808	0.981
X01	-	0.036	0.278	0.26					
X02	-	0.026	0.953	0.255					
P011	-	13.31	0.913	1.354					
P012	-	4.02	6.454	0.966					
P022	-	12.872	3.606	1.353					

Table 6.36: Simulation Results: True IC: Diffuse, Fitted IC: Free Parameter, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.196	0.009	0.023	-0.004	-0.003	0.009	1	1
Z31	0.8	0.81	0.01	0.017	0.01	0.012	0.014	0.994	1
Z52	0.9	0.903	0.007	0.014	0.003	0.003	0.008	1	1
Z62	1.1	1.097	0.006	0.016	-0.003	-0.003	0.006	1	1
V11	1	5.762	0.333	0.381	4.762	4.762	4.774	0	1
V21	0.4	2.461	0.282	0.293	2.061	5.154	2.081	0	1
V22	1	7.032	0.376	0.438	6.032	6.032	6.044	0	1
T11	1.2	1.219	0.009	0.025	0.019	0.015	0.021	1	1
T21	0.6	0.606	0.011	0.028	0.006	0.01	0.012	1	1
T12	-0.4	-0.407	0.008	0.019	-0.007	0.017	0.01	1	1
T22	1	1.006	0.008	0.02	0.006	0.006	0.01	1	1
U11	0.8	4.283	0.196	0.205	3.483	4.353	3.488	0	1
U22	0.6	5.827	0.252	0.278	5.227	8.712	5.233	0	1
U33	2	3.866	0.19	0.186	1.866	0.933	1.875	0	1
U44	1	4.621	0.184	0.221	3.621	3.621	3.625	0	1
U55	1.5	4.221	0.195	0.202	2.721	1.814	2.728	0	1
U66	0.4	5.075	0.194	0.241	4.675	11.687	4.679	0	1

Table 6.37: Simulation Results: True IC: Diffuse, Fitted IC: Null, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.193	0.008	0.008	-0.007	-0.006	0.011	0.851	1
Z31	0.8	0.795	0.009	0.01	-0.005	-0.006	0.011	0.911	1
Z52	0.9	0.896	0.007	0.007	-0.004	-0.004	0.008	0.909	1
Z62	1.1	1.095	0.005	0.006	-0.005	-0.004	0.007	0.875	1
V11	1	1.004	0.079	0.079	0.004	0.004	0.079	0.962	1
V21	0.4	0.395	0.06	0.056	-0.005	-0.012	0.06	0.952	1
V22	1	1	0.093	0.075	0	0	0.093	0.933	0.998
T11	1.2	1.195	0.009	0.009	-0.005	-0.004	0.01	0.916	1
T21	0.6	0.598	0.01	0.009	-0.002	-0.004	0.01	0.928	1
T12	-0.4	-0.398	0.007	0.007	0.002	-0.004	0.007	0.947	1
T22	1	0.999	0.007	0.007	-0.001	-0.001	0.007	0.94	1
U11	0.8	0.801	0.046	0.048	0.001	0.002	0.046	0.954	1
U22	0.6	0.605	0.055	0.05	0.005	0.008	0.055	0.947	1
U33	2	1.991	0.096	0.095	-0.009	-0.004	0.097	0.952	1
U44	1	1.002	0.053	0.054	0.002	0.002	0.053	0.957	1
U55	1.5	1.502	0.074	0.074	0.002	0.001	0.074	0.962	1
U66	0.4	0.402	0.044	0.036	0.002	0.006	0.044	0.95	1

Table 6.38: Simulation Results: True IC: Diffuse, Fitted IC: deJong DKF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.194	0.009	0.008	-0.006	-0.005	0.011	0.848	1
Z31	0.8	0.795	0.009	0.01	-0.005	-0.007	0.011	0.924	1
Z52	0.9	0.782	0.579	0.014	-0.118	-0.131	0.59	0.777	1
Z62	1.1	0.969	0.593	0.014	-0.131	-0.119	0.605	0.711	1
V11	1	1.215	0.437	0.085	0.215	0.215	0.486	0.766	1
V21	0.4	0.219	0.578	0.064	-0.181	-0.453	0.604	0.787	0.98
V22	1	0.805	0.407	0.077	-0.195	-0.195	0.45	0.761	0.807
T11	1.2	1.191	0.012	0.009	-0.009	-0.007	0.015	0.843	1
T21	0.6	0.597	0.013	0.011	-0.003	-0.005	0.013	0.898	1
T12	-0.4	-0.359	0.188	0.01	0.041	-0.103	0.192	0.761	1
T22	1	0.879	0.555	0.015	-0.121	-0.121	0.567	0.766	1
U11	0.8	0.837	0.119	0.048	0.037	0.046	0.124	0.848	1
U22	0.6	0.526	0.187	0.047	-0.074	-0.124	0.201	0.766	0.934
U33	2	2.01	0.114	0.093	0.01	0.005	0.114	0.909	1
U44	1	1.727	1.729	0.082	0.727	0.727	1.872	0.777	1
U55	1.5	1.313	0.579	0.067	-0.187	-0.125	0.607	0.766	0.868
U66	0.4	0.709	0.775	0.046	0.309	0.771	0.832	0.756	0.964

Table 6.39: Simulation Results: True IC: Diffuse, Fitted IC: Koopman exact initial KF, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.194	0.009	0.008	-0.006	-0.005	0.011	0.84	1
Z31	0.8	0.795	0.009	0.01	-0.005	-0.007	0.011	0.912	1
Z52	0.9	0.746	0.667	0.017	-0.154	-0.171	0.683	0.753	1
Z62	1.1	0.933	0.685	0.016	-0.167	-0.152	0.703	0.722	1
V11	1	1.227	0.455	0.087	0.227	0.227	0.507	0.763	1
V21	0.4	0.178	0.656	0.069	-0.222	-0.556	0.692	0.778	0.974
V22	1	0.8	0.408	0.08	-0.2	-0.2	0.453	0.753	0.799
T11	1.2	1.191	0.014	0.009	-0.009	-0.008	0.017	0.835	1
T21	0.6	0.593	0.045	0.012	-0.007	-0.011	0.045	0.902	0.995
T12	-0.4	-0.351	0.207	0.01	0.049	-0.124	0.213	0.753	1
T22	1	0.846	0.647	0.018	-0.154	-0.154	0.663	0.758	1
U11	0.8	0.829	0.124	0.049	0.029	0.036	0.127	0.861	0.995
U22	0.6	0.539	0.193	0.049	-0.061	-0.101	0.201	0.763	0.948
U33	2	2.011	0.122	0.093	0.011	0.006	0.122	0.897	1
U44	1	1.816	1.954	0.087	0.816	0.816	2.113	0.778	0.995
U55	1.5	1.307	0.571	0.07	-0.193	-0.129	0.602	0.789	0.871
U66	0.4	0.718	0.762	0.048	0.318	0.796	0.824	0.758	0.974

Table 6.40: Simulation Results: True IC: Diffuse, Fitted IC: Large κ, N=200, T=5

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.202	0.028	0.022	0.002	0.001	0.028	0.931	1
Z31	0.8	0.797	0.027	0.026	-0.003	-0.004	0.028	0.961	1
Z52	0.9	0.915	0.164	0.023	0.015	0.017	0.164	0.897	1
Z62	1.1	1.106	0.127	0.02	0.006	0.006	0.127	0.856	1
V11	0.05	0.084	0.15	0.02	0.034	0.67	0.154	0.867	0.842
V21	0.02	-0.058	0.477	0.022	-0.078	-3.898	0.483	0.875	0.403
V22	0.05	0.046	0.02	0.035	-0.004	-0.087	0.021	0.914	0.817
T11	1.2	1.191	0.041	0.012	-0.009	-0.008	0.042	0.878	1
T21	0.6	0.536	0.459	0.024	-0.064	-0.106	0.462	0.886	1
T12	-0.4	-0.399	0.024	0.012	0.001	-0.003	0.024	0.886	1
T22	0.7	0.674	0.247	0.016	-0.026	-0.037	0.248	0.911	1
U11	0.8	0.807	0.082	0.04	0.007	0.009	0.083	0.928	1
U22	0.6	0.587	0.107	0.034	-0.013	-0.021	0.108	0.869	0.992
U33	2	2.013	0.113	0.092	0.013	0.006	0.114	0.936	1
U44	1	1.068	0.335	0.052	0.068	0.068	0.341	0.892	1
U55	1.5	1.468	0.26	0.069	-0.032	-0.021	0.262	0.875	0.975
U66	0.4	0.444	0.28	0.027	0.044	0.111	0.283	0.914	0.986
X01	-	0.003	0.14	0.137					
X02	-	0.075	0.82	0.166					
P011	-	3.696	0.279	0.391					
P012	-	2.518	9.308	0.565					
P022	-	3.591	0.723	0.725					

Table 6.41: Simulation Results: True IC: Diffuse, Fitted IC: Free Parameter, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.18	0.028	0.041	-0.02	-0.017	0.034	0.985	1
Z31	0.8	0.825	0.035	0.037	0.025	0.031	0.043	0.946	1
Z52	0.9	0.906	0.039	0.028	0.006	0.007	0.039	0.98	1
Z62	1.1	1.085	0.083	0.028	-0.015	-0.014	0.084	0.964	1
V11	0.05	1.291	0.126	0.107	1.241	24.829	1.248	0.005	0.995
V21	0.02	0.891	0.224	0.08	0.871	43.527	0.899	0	1
V22	0.05	1.348	0.139	0.109	1.298	25.963	1.306	0.005	0.995
T11	1.2	1.21	0.117	0.045	0.01	0.008	0.117	0.992	1
T21	0.6	0.58	0.175	0.047	-0.02	-0.033	0.176	0.992	1
T12	-0.4	-0.413	0.061	0.036	-0.013	0.033	0.062	0.997	1
T22	0.7	0.71	0.07	0.034	0.01	0.014	0.07	0.995	0.997
U11	0.8	1.728	0.163	0.086	0.928	1.16	0.942	0	1
U22	0.6	2.076	0.2	0.103	1.476	2.461	1.49	0	1
U33	2	2.455	0.147	0.118	0.455	0.228	0.478	0.015	1
U44	1	1.892	0.211	0.093	0.892	0.892	0.917	0	1
U55	1.5	2.137	0.185	0.103	0.637	0.425	0.664	0	1
U66	0.4	1.664	0.162	0.082	1.264	3.159	1.274	0	1

Table 6.42: Simulation Results: True IC: Diffuse, Fitted IC: Null, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.152	0.02	0.02	-0.048	-0.04	0.052	0.34	1
Z31	0.8	0.766	0.025	0.025	-0.034	-0.043	0.042	0.722	1
Z52	0.9	0.871	0.019	0.019	-0.029	-0.032	0.035	0.644	1
Z62	1.1	1.065	0.016	0.016	-0.035	-0.031	0.038	0.37	1
V11	0.05	0.062	0.019	0.018	0.012	0.233	0.022	0.917	0.949
V21	0.02	0.024	0.015	0.013	0.004	0.179	0.015	0.924	0.428
V22	0.05	0.054	0.018	0.017	0.004	0.077	0.019	0.931	0.883
T11	1.2	1.17	0.011	0.011	-0.03	-0.025	0.032	0.193	1
T21	0.6	0.588	0.015	0.014	-0.012	-0.021	0.019	0.809	1
T12	-0.4	-0.385	0.011	0.011	0.015	-0.037	0.019	0.685	1
T22	0.7	0.7	0.009	0.009	0	0.001	0.009	0.947	1
U11	0.8	0.801	0.04	0.04	0.001	0.001	0.04	0.949	1
U22	0.6	0.6	0.035	0.034	0	0	0.035	0.933	1
U33	2	2.01	0.097	0.092	0.01	0.005	0.098	0.949	1
U44	1	1.001	0.049	0.049	0.001	0.001	0.049	0.947	1
U55	1.5	1.499	0.071	0.07	-0.001	-0.001	0.07	0.933	1
U66	0.4	0.403	0.024	0.025	0.003	0.007	0.024	0.963	1

Table 6.43: Simulation Results: True IC: Diffuse, Fitted IC: deJong DKF, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.155	0.03	0.02	-0.045	-0.037	0.054	0.36	1
Z31	0.8	0.765	0.027	0.025	-0.035	-0.043	0.044	0.715	1
Z52	0.9	0.84	0.288	0.021	-0.06	-0.067	0.294	0.627	1
Z62	1.1	1.029	0.187	0.017	-0.071	-0.064	0.199	0.345	1
V11	0.05	0.116	0.211	0.022	0.066	1.324	0.221	0.842	0.938
V21	0.02	-0.074	0.619	0.027	-0.094	-4.681	0.625	0.855	0.461
V22	0.05	0.049	0.022	0.046	-0.001	-0.021	0.022	0.92	0.813
T11	1.2	1.159	0.044	0.011	-0.041	-0.034	0.06	0.181	1
T21	0.6	0.489	0.618	0.027	-0.111	-0.186	0.627	0.782	1
T12	-0.4	-0.378	0.04	0.012	0.022	-0.055	0.045	0.648	1
T22	0.7	0.636	0.313	0.015	-0.064	-0.091	0.319	0.881	1
U11	0.8	0.809	0.072	0.041	0.009	0.012	0.073	0.917	1
U22	0.6	0.577	0.127	0.034	-0.023	-0.039	0.129	0.845	0.984
U33	2	2.011	0.102	0.092	0.011	0.006	0.103	0.93	1
U44	1	1.096	0.394	0.056	0.096	0.096	0.405	0.863	1
U55	1.5	1.481	0.306	0.07	-0.019	-0.013	0.307	0.883	0.979
U66	0.4	0.462	0.341	0.028	0.062	0.156	0.347	0.894	0.984

Table 6.44: Simulation Results: True IC: Diffuse, Fitted IC: Koopman exact initial KF, N=200, T=5, Mild Nonstationarity

θ	True Value	$\hat{\theta}$	SE_θ	$\hat{SE}$	Bias	RB	RMSE	cov	pow
Z21	1.2	1.156	0.025	0.02	-0.044	-0.037	0.051	0.377	1
Z31	0.8	0.766	0.025	0.024	-0.034	-0.042	0.042	0.722	1
Z52	0.9	0.839	0.285	0.02	-0.061	-0.067	0.291	0.615	1
Z62	1.1	1.022	0.27	0.018	-0.078	-0.071	0.281	0.352	1
V11	0.05	0.113	0.206	0.021	0.063	1.269	0.215	0.848	0.944
V21	0.02	-0.074	0.603	0.026	-0.094	-4.69	0.609	0.863	0.451
V22	0.05	0.049	0.022	0.045	-0.001	-0.025	0.022	0.927	0.815
T11	1.2	1.159	0.043	0.011	-0.041	-0.034	0.059	0.18	1
T21	0.6	0.493	0.599	0.027	-0.107	-0.178	0.607	0.782	1
T12	-0.4	-0.378	0.038	0.012	0.022	-0.054	0.044	0.661	1
T22	0.7	0.635	0.316	0.015	-0.065	-0.093	0.322	0.881	1
U11	0.8	0.806	0.065	0.041	0.006	0.008	0.065	0.916	1
U22	0.6	0.576	0.101	0.034	-0.024	-0.04	0.104	0.861	0.985
U33	2	2.009	0.098	0.092	0.009	0.005	0.099	0.934	1
U44	1	1.091	0.384	0.055	0.091	0.091	0.394	0.873	1
U55	1.5	1.483	0.31	0.07	-0.017	-0.011	0.31	0.889	0.98
U66	0.4	0.458	0.315	0.028	0.058	0.144	0.32	0.909	0.987

Table 6.45: Simulation Results: True IC: Diffuse, Fitted IC: Large κ, N=200, T=5, Mild Nonstationarity

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Ansley, C. F., & Kohn, R. (1985). Estimation, filtering and smoothing in state space models with incompletely specified initial conditions. *Annals of Statistics*, *13*(4), 1286–1316.

Ayers, S. (1997). The application of chaos theory to psychology. *Theory and Psychology*, *7*(3), 373–398.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Bolger, N., Davis, A., Olchowski, A. E., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579–616.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (alt) models: A synthesis of two traditions. *Sociological Methods and Research*, *32*, 336–383.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.

Bowlby, J. (1982). *Attachment and loss: Vol. 1. attachment* (2nd ed.). New York: Basic Books.

Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of arma time series models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Advances in psychometrics: A festschrift for roderick p. mcdonald* (pp. 415–452). Mahwah, NJ: Erlbaum.

Browne, M. W., & Zhang, G. (2005). DyFA: Dynamic Factor Analysis of Lagged Correlation Matrices, Version 2.03 [Computer software manual]. Available from `http://quantrm2.psy.ohio-state.edu/browne/`

Caspi, A., Moffitt, T. E., Newman, D. L., & Silva, P. A. (1996). Behavioral observations at age 3 years predict adult psychiatric disorders: Longitudinal evidence from a birth cohort. *Psychological Bulletin*, *53*, 1033–1039.

Cattell, R. B., & Cattell, M. D. L. (1969). *Jr.-sr. high school personality questionnaire*. Institute for Personality and Ability Testing.

Chatfield, C. (2004). *The analysis of time series: An introduction* (6th ed.). Boca Raton, FL: Chapman & Hall/CRC.

Chow, S.-M., Ho., M. H. R., Hamaker, E., & Dolan, C. (2010).

Structural Equation Modeling.

Cicchetti, D., & Tucker, D. (1994). Development of self-regulatory structures of the mind. *Developmental Psychopathology*, *6*, 533–549.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

De Jong, P. (1988). The likelihood for a state space model. *Biometrika*, *75*, 165–169.

De Jong, P. (1991). The diffuse kalman filter. *The Annals of Statistics*, *19*(2), 1073–1083.

De Jong, P. (2003). Smoothing with an unknown initial condition. *Journal of time series analysis*, *24*(2), 141–148.

De Jong, P., & Chu-Chun Lin, S. (1994). Stationary and non-stationary state space models. *Journal of Time Series Analysis*, *15*, 151–166.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistics Society, Series B*, *39*, 1–38.

Diener, E., Fujita, F., & Smith, H. (1995). The personality structure of affect. *Journal of Personality and Social Psychology*, *69*(1), 130–141.

Dolan, C. V., & Molenaar, P. C. M. (1991). A note on the calculation of latent trajectories in the quasi markov simplex model by means of regression method and the discrete kalman filter. *Kwantitatieve Methoden*, *38*, 29–44.

Doornik, J. A. (2007). *Object-oriented matrix programming using ox* (3rd ed.). London: Timberlake Consultants Press and Oxford.

Du Toit, S. H. C., & Browne, M. W. (2001). The covariance structure of a vector time series. In R. Cudeck, S. H. C. D. Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future a festschrift for karl jreskog* (pp. 279–314). Chicago: Scientific Software International Inc.

Du Toit, S. H. C., & Browne, M. W. (2007). Structural equation modeling of multivariate time series. *Multivariate Behavioral Research*, *42(1)*, 67–101.

Enders, C. K. (2001). A primer on the use of maximum likelihod algorithms available for use with missing data. *Structural Equation Modeling*, *8*, 128–141.

Francke, M. K., Koopman, S. J., & deVos, A. F. (2010). Likelihood functions for state space models with diffuse initial conditions. *Journal of Time Series Analysis*, *31*.

Gottschalk, A., Bauer, M. S., & Whybrow, P. C. (1995). Evidence of chaotic mood variation in bipolar disorder. *Archive of General Psychiatry*, *52*, 947–959.

Greenberg, M. T., Kusche, C. A., & Spelz, M. (1991). Emotion regularion, self-control, and psychopathology: The role of relationships in early childhood. In D. Cicchetti & S. L. Toth (Eds.), *Rocheester symposium on developmental psychopathology: Vol. 2. internalizing and externalizing expressions of dysfunction* (pp. 21–55). Rochester, NY: University of Rochester Press.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*, 430–450.

Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods and Research*, *33*, 404–416.

Hamaker, E. L., & Dolan, C. V. (2009). Idiographic data analysis: Quantitative methods from simple to advanced. In J. Valsiner, P. C. M. Molenaar, M. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 191–216). New York: Springer-Verlag.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2003). Arma-based sem when the number of time points t exceeds the number of cases n: Raw data maximum likelihood. *Structural Equation Modeling*, *10(3)*, 352–379.

Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Harvey, A. C. (1991). *Forecasting, structural time series models and the kalman filter*. Cambridge: Cambridge University Press.

Harvey, A. C., & Phillips, G. D. A. (1979). Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika*, *66*, 49–58.

Jessor, R., Donovan, J. E., & Costa, F. M. (1991). *Beyond adolescence: Problem behavior and young adult development* (2nd ed.). New York: Cambridge University Press.

Jöreskog, K. G., & Sörbom, D. (2001). *LISREL user's guide*. Chicago, IL: Scientific Software International.

Kitagawa, G. (1981). A nonstationary time series model and its fitting by a recursive filter. *Journal of Time Series Analysis*, *2*, 103–116.

Koopman, S. J. (1997). Exact initial kalman filtering and smoothing for non-stationary time series models. *Journal of American Statistical Association*, *92*, 1630–1638.

Koopman, S. J., & Durbin, J. (2000). Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis*, *21*, 281–296.

Koopman, S. J., & Durbin, J. (2003). Filtering and smoothing of state vector for diffuse state-space models. *Journal of Time Series Analysis*, *24*, 85–98.

Koopman, S. J., Shephard, N., & Doornik, J. A. (2008). *Statistical algorithms for models in state space form: Ssfpack 3.0*. London, UK: Timberlake Consultants.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworths.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York, N.Y.: Wiley.

Ltkepohl, H. (1993). *Introduction to multiple time series analysis, second edition*. Berlin: Springer-Verlag.

MacCallum, R., & Ashby, F. G. (1986). Relationship between linear systems theory and covariance structure modeling. *Journal of Mathematical Psychology*, *1*, 1–27.

Maughan, A., & Cicchetti, D. (2002). Impact of child maltreatment and interadult violence on children's emotion regulation abilities and socioemotional adjustment. *Child Development*, *73*(5), 1525–1542.

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analysis with incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 139–175). Washington, DC: American Psychological Association.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*(2), 181–202.

Molenaar, P. C. M., & Nesselroade, J. R. (1998). A comparison of pseudo-maximum likelihood and asymptotically distribution-free dynamic factor analysis parameter estimation in fitting covariance-structure models to block-toeplitz matrices representing single-subject multivariate time series. *Multivariate Behavioral Research*, *33*, 313–342.

Muthén, L. K., & Muthén, B. O. (1998–2004). *Mplus technical appendices* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998–2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. (2001). Dynamic factor analysis modles for representing process in multivariate time-series. In D. M. Moskowitz & S. L. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 235–265). Mahwah, NJ: Lawrence Erlbaum.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time series for dynamic factor analysis. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 223–250). Thousand Oaks, CA: Sage Publications.

Oud, J. H. L., Bercken, J. H. van den, & Essers, R. J. (1990). Longitudinal factor score estimation using the kalman filter. *Applied psychological measurement*, *14*(4), 395–418.

Oud, J. H. L., & Jansen, R. A. R. G. (1996). Nonstationary longitudinal lisrel model estimation from incomplete panel data using em and the kalman smoother. In U. Engel & R. Reinecke (Eds.), *Analysis of change: Advanced techniques in panel data analysis* (pp. 135–159). Berlin: Walter de Gruyter.

Preacher, K. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, *44*.

Preacher, K. J. (2006). Quantifying parsimony in structural equation modeling. *Multivariate Behavioral Research*, *41*(3), 227–259.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Robertson, R., & Combs, A. (1995). *Chaos theory in psychology and the life sciences*. Hillsdale, N.J: Lawrence Erlbaum Associates.

Schweppe, F. C. (1965). Evaluation of likelihood functions for gaussian signals. *IEEE Trans. Inform. Theory*, *IT-11*, 61–70.

Schweppe, F. C. (1973). *Uncertain dynamic systems*. Englewood Cliffs, N.J.: Prentice-Hall.

Shafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.

Shifren, K., Hooker, K. A., Wood, P. K., & Nesselroade, J. R. (1997). Structure and variation of mood in individuals with Parkinson's disease: A dynamic factor analysis. *Psychology and Aging*, *12*, 328–339.

Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications*. New York, NY: Springer Texts in Statistics.

Skrondal, A. (2000). Design and analysis of monte carlo experiments: Attacking the conventional wisdom. *Multivariate Behavioral Research*, *35*(2), 137–167.

Steiger, J. H., & Lind, J. (1980). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City.

Sternberg, R. J. (1999). Successful intelligence: Finding a balance. *Trends in Cognitive Sciences*, *3*, 436–442.

Van Vuuren, D. P., De Beer, J. J., & Du Toit, S. H. C. (1982). *Television needs and personality needs of South African adolescents.* Unpublished report, Pretoria: Human Sciences Research Council.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wei, W. W. S. (1990). *Time series analysis*. Redwood City, CA: Addison–Wesley.

Zucker, R. A., Chermack, S. T., & Curran, G. M. (2000). Alcoholism: A life span perspective on etiology and course. In A. J. Sameroff, M. Lewis, & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (2nd ed., pp. 569–587). Dordrecht, Netherlands: Kluwer Academic Publishers.