

経済統計・政府統計の理論と応用 2017¹

国友直人² & 山本拓³ & 久保川達也⁴

(共編)

2017年2月

¹文部科学省・科学研究費プロジェクト「経済統計・政府統計の応用からの提言」(2015年~2018年)が、2017年2月2日に東京大学経済学研究科において開催した講演をまとめたものである。

²明治大学経済学部

³統計研究会

⁴東京大学経済学研究科

前書き

本報告書は、日本学術振興会・科学研究費プロジェクト「経済統計・政府統計の理論と応用からの提言」（2015年度－2018年度、研究代表者：山本 拓）が、2017年2月2日（木）に東京大学小島ホールにおいて開催した2016年度の研究集会における講演内容をまとめたものである。

本プロジェクトの目的は、経済統計・政府統計における主要な課題の、技術的および制度的問題を、統計学的な立場から理論的・学術的に検討すること、ならびに経済統計・政府統計の応用の際の問題点を明らかにし、それらの解決案を模索・提言することである。

経済統計、とりわけ政府統計は、経済・社会の動向を理解し、政策を実施、評価するためには不可欠な情報であることは言うまでもない。近年はEBPM（Evidence based policy making）ということもよく聞かれ、統計の重要性は一般に広く認識されつつあると思われる。

しかしながら、経済統計・政府統計への信頼性は、近年必ずしも増しているとは言えない状況である。経済社会の急激な変化に伴い、政府統計の質の確保が困難になりつつあると思われる。たとえばマクロ経済統計の側面では、GDP統計や消費統計などに代表されるマクロ公表系列の質と信頼性の問題、信頼性の高い将来人口の推計の問題、地域による経済情勢のばらつきの把握などの問題を挙げることができる。またミクロ経済データにおいては、統計調査をとりまくプライバシー意識の高まりから、調査精度の確保が難しくなりつつあるという問題や、情報開示と秘密保持の両立という匿名化問題などを挙げることができる。

本研究集会の目的は、経済統計・政府統計をめぐる技術的・制度的問題点、あるいは応用にかかわる問題を、また統計学の研究者と実際に経済統計・政府統計に作成者または利用者として携わっている方々との直接的な交流の場を提供しようとするものである。

第1セッションでは、貯蓄率を通じての各種公的消費統計に関する整合性の問題、より正確な消費統計のための需要側と供給側の調整の問題ならび高齢化に伴うと世帯変化の問題、さらに自記式アンケートに伴う標本調査に関する問題が取り上げられた。第2セッションでは、最近大きな話題となっている、GDP統計に関する季節調整法とベンチマークの影響の問題、さらに経済データの応用に関して、GDPと株価収益率の時系列データに基づく因果関係測定にまつわる問題が取り上げられた。第3セッションは統計的手法の問題を取り上げており、GDP統計の需要側と供給側の時系列的側面での調整の方法、ならびの動学的パネルデータ推定の新しい推定法が提案された。

このような研究集会が経済統計・政府統計に様々な形で関わる人々の刺激となり、今後の各種統計の改善の一助になることを期待する次第である。

2017年2月

編者

研究集会・プログラム

科研プロジェクト「経済統計・政府統計の理論と応用」

日程：2017年2月2日(木)

会場：東京大学経済学部小島ホール 1階第2セミナー室

オーガナイザー： 国友直人・山本拓

<挨拶>

13:00～13:05 山本拓（統計研究会）

研究プロジェクトの計画

<セッション I> 政府統計・経済統計の諸問題 1

Chair: 山本拓（統計研究会）

13:05～13:40 *宇南山卓（一橋大学）・大野太郎（信州大学）

世帯類型別の貯蓄率の動向

13:40～14:15 川崎茂（日本大学）

景気指標としての個人消費関連統計について

14:15～14:50 土屋隆裕（統計数理研究所）

自記式調査における複数回答方式をめぐる諸問題

<セッション II> 政府統計・経済統計の諸問題 2

Chair: 星野伸明（金沢大学）

15:00～15:35 長谷川秀司（内閣府）

SNA 統計における時系列データ作成上の課題ーベンチマーキング及び季節調整を中心にー

15:35～16:10 川崎能典（統計数理研究所）

内閣支持率と株価の因果関係

<セッション III> 経済統計の理論と応用

Chair: 国友直人（明治大学）

16:20～16:55 *千木良弘朗（東北大学）・山本拓（統計研究会）

真のデータに線形制約がある際の観測誤差の修正ー国民経済計算の場合

16:55～17:30 Breitung, J. (ケルン大学)・*早川和彦（広島大学）・M. Qi（広島大学）

Double Filter Instrumental Variable Estimation of Panel Data Models with Weakly Exogenous Variables

*は共著の場合の報告者

Theory and Applications of Economic and Government Statistics 2017

Abstract

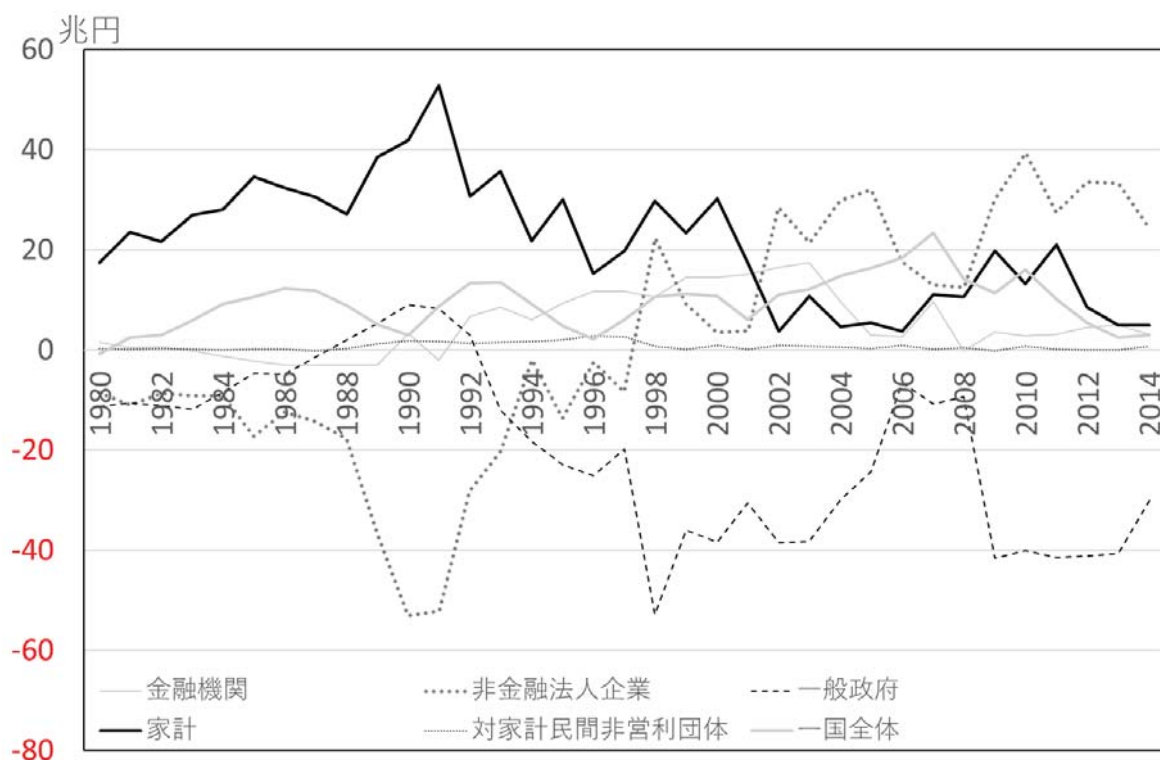
This report is a summary of the unpublished talks and papers presented at the second conference held on February 2, 2016 under the title "Theory and Applications of Economic Statistics and Government Statistics". The topics are on (i) the recent issues of official statistics, (ii) applications of economic statistics, and (iii) statistical methodologies for economic statistics. The papers presented at the second conference have been written in Japanese or English

日本の世帯属性別貯蓄率の動向について

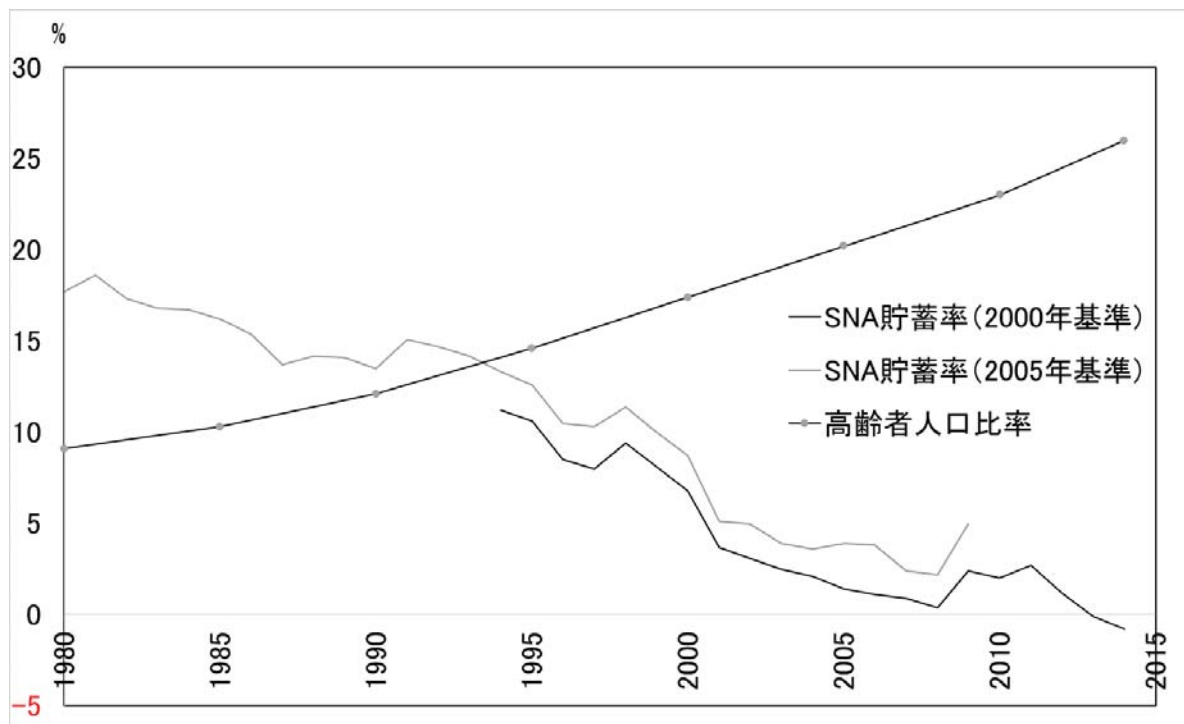
宇南山 卓(一橋大学)

大野太郎(信州大学)

日本の貯蓄投資バランス



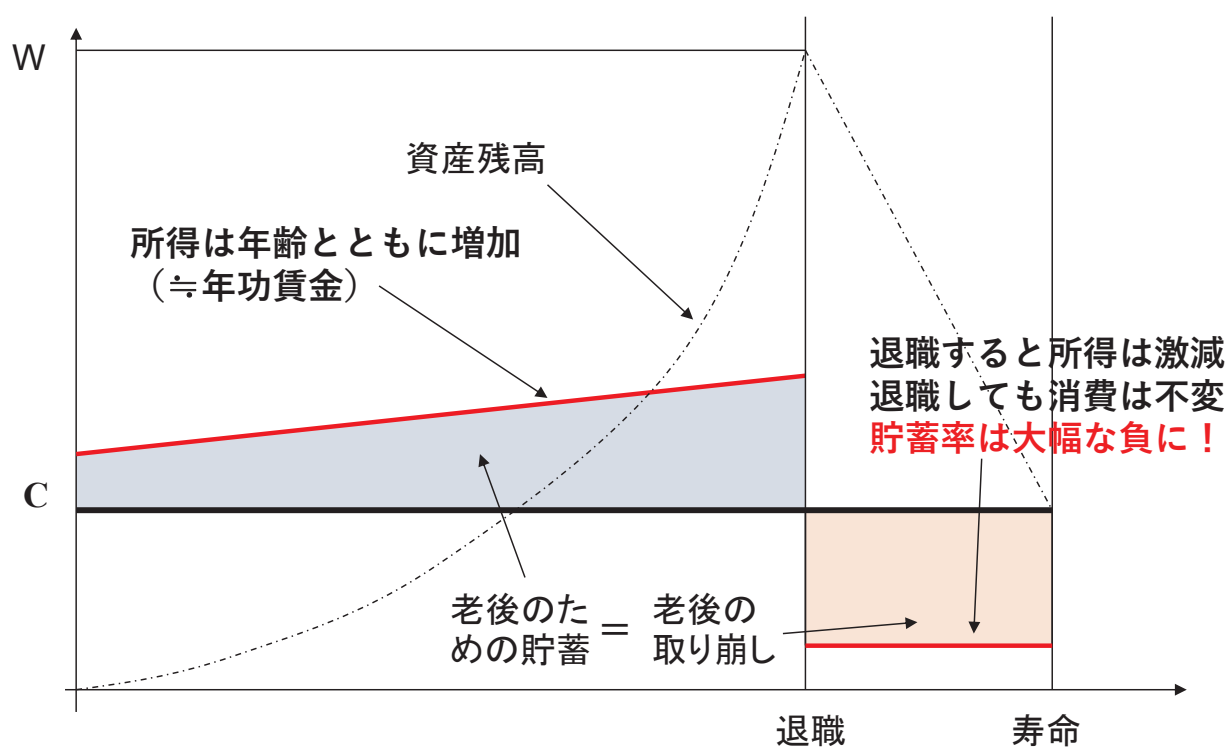
貯蓄率と高齢化



貯蓄率：国民経済計算・制度部門別所得支出勘定・「調整貯蓄率」
 高齢者人口比率：国勢調査・65歳以上人口の比率(2014年のみ人口推計)

3

ライフサイクル仮説



4

貯蓄率の低下と高齢化

- 家計貯蓄率の低下は「高齢化」だけで説明できるのか？
 - 高齢者の割合が増加している
 - ライフサイクル仮説に基づけば高齢者の貯蓄率は低い（マイナスとなる）はず
 - 高齢化はマクロの貯蓄率を引き下げる！
- 年齢別等の世帯属性別の貯蓄率は安定しており、構成だけが変化したのか？
- 世帯属性別の貯蓄率を計算する
 - マクロ統計では不可能！
 - (公式の) ミクロ統計はマクロ統計と大きく乖離

5

ミクロ統計による貯蓄率

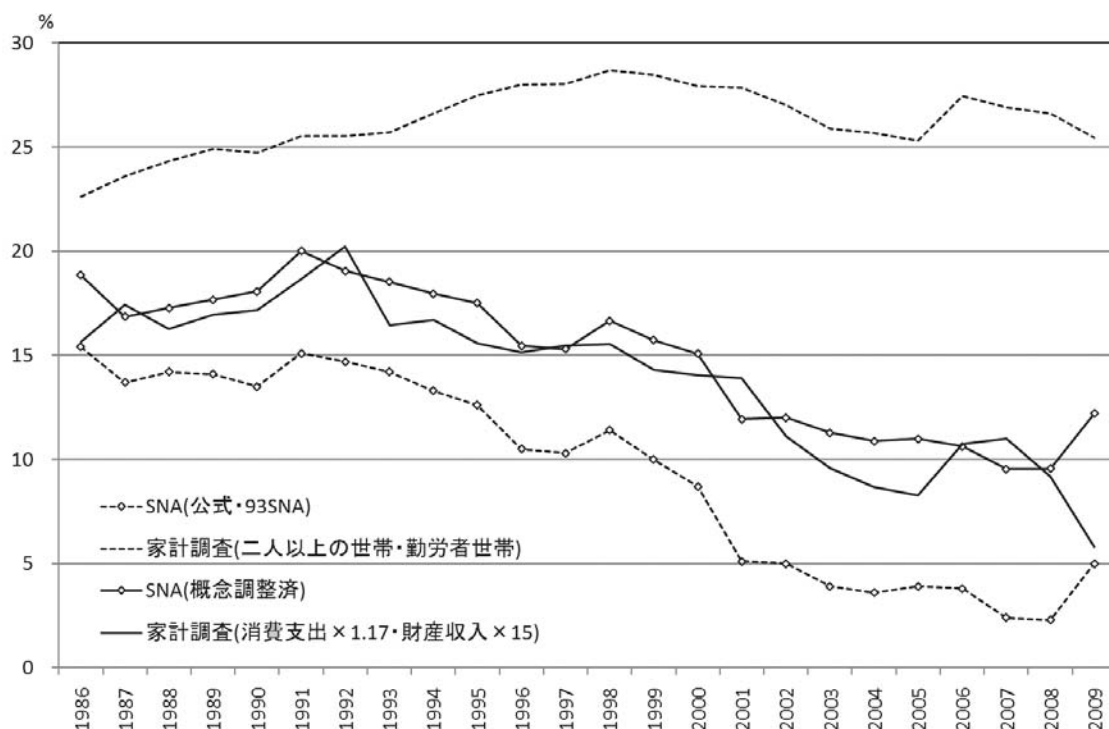
- 日本において個別世帯の貯蓄率を把握可能な統計として家計調査・全国消費実態調査が存在

	家計調査	全国消費実態調査
頻度	月次（各世帯は6か月継続して調査）	5年に一度（9・10・11月の3か月）
サンプルサイズ	毎月約9千世帯	毎回約5万5千世帯
収入	世帯の年間収入 毎月の収入（勤労者・無職世帯のみ）	世帯員ごとの 年間収入 毎月の収入（勤労者・無職世帯のみ）
税・社会保険料	家計簿に記入（勤労者・無職世帯のみ）	家計簿に記入（勤労者・無職世帯のみ）
支出	品目別に詳細に記録	品目別に詳細に記録
資産	金融資産（種類別・ 単身世帯はNA ） 持家に関する詳細な情報（ 帰属家賃はなし ）	金融資産（種類別） 持家に関する詳細な情報 帰属家賃の推計値

- 両者とも全世帯の年間を通じた貯蓄率は計算できない！

6

SNAとミクロ統計の貯蓄率



宇南山(2009)の手法を2009年まで延長したもの (筆者作成)

7

SNAと家計調査の貯蓄率の差

- 貯蓄率の差の原因
 - 家計調査は勤労者世帯のみが対象
 - 引退した高齢者世帯が含まれていない
 - 貯蓄概念に違いがある
 - SNAでは帰属家賃が含まれる
 - 家計調査に正確性に疑義がある
 - 耐久消費財や冠婚葬祭費などの高額消費が過少
 - 財産所得 (金利・配当・賃貸料など) が過少
- 適切な補正をすればマクロとミクロの貯蓄率はほぼ一致する(宇南山2009; 米田 2017)
 - 世帯属性別の補正方法については未検討

8

全国消費実態調査を用いた 世帯属性別貯蓄率の推計

- メリット
 - 家計調査よりもサンプルサイズが大きい
 - 詳細な補正・分析が可能
 - 長期的な動向を見るには頻度が低いことの影響は小さい
 - 全世帯の世帯員レベルでの収入が利用可能
 - 税・社会保険料を推計することで全世帯を対象とできる
 - 年間収入に関してより信頼性が高い(多田・三好 2015)
 - 資産・負債の情報がある
 - 単身世帯についても資産・負債を調査している
 - 自宅以外の不動産なども調査対象
- デメリット
 - 9・10・11月だけを調査対象としている
 - 季節性の強い項目が調査されていない(ボーナスなど)
 - 支出の季節性をコントロールできない
 - 耐久消費財などの高額支出に過少性がある可能性
 - 家計調査と状況は類似

9

貯蓄率の変動の要因分解

- マクロの貯蓄率

$$S_t = \sum_i w_{ti} S_{ti}$$

- w_{ti} : 属性*i*の世帯の構成比
- S_{ti} : 属性*i*の世帯の平均貯蓄率

- マクロの貯蓄率の変化

$$S_t - S_0 = \sum_i \underbrace{s_{0i}(w_{ti} - w_{0i})}_{\text{人口構成変化効果}} + \underbrace{w_{ti}(S_{ti} - S_{0i})}_{\text{貯蓄率変化効果}} \quad (1)$$

世帯属性別貯蓄率の推計手順

- 貯蓄率 = (可処分所得 - 消費支出) / 可処分所得
 - 年間可処分所得 (= 年間収入 - 年間税・社会保険料)
 - 世帯員ごとの年間収入は「年収・貯蓄等調査票」
 - 税・社会保険料は世帯ごとに推計
 - 「世帯票」から得られる世帯構成
 - 税率・保険料率等の制度情報
 - 年間消費支出
 - 耐久財・冠婚葬祭費などは**家計消費状況調査**で推計される年間支出額を世帯属性ごとに代入
 - 同じ世帯属性であれば同じ支出をすると仮定
 - 教育費は**家計消費状況調査**で1人当たりの金額を推計し、通う学校の種類ごとの世帯員数をかけて推計
 - これら以外の品目は、世帯ごとの支出額に、**家計調査**で得られる世帯属性ごとの季節パターンを用いて年間額に換算

11

税・社会保険料の推計

- 年収・貯蓄等調査票で各世帯員の収入を把握
 - 収入の源泉に基づき9つのカテゴリー
- 世帯主や扶養関係をモデル上で再定義
 - 最高所得者を世帯主とする
 - 所得控除等を最大限活用できる扶養関係を想定
- 加入する社会保険制度の推定
 - 雇用者所得・職業の情報に基づき推計
 - 雇用者所得が一定水準（パートタイム労働者等の平均賃金×30時間×52週）より多いか、「勤めか自営かの別」を考慮して健康保険組合、厚生年金の被保険者
 - 医療保険では75歳以上を後期高齢者医療制度、介護保険では65歳以上の者は第1号被保険者
 - それ以外は国民健康保険、国民年金の被保険者に
 - 減免・上限制度は適用可能なものは全て適用
- 所得税率・住民税率・年金・医療・介護・雇用の各保険料率に基づき税額・社会保険料を計算

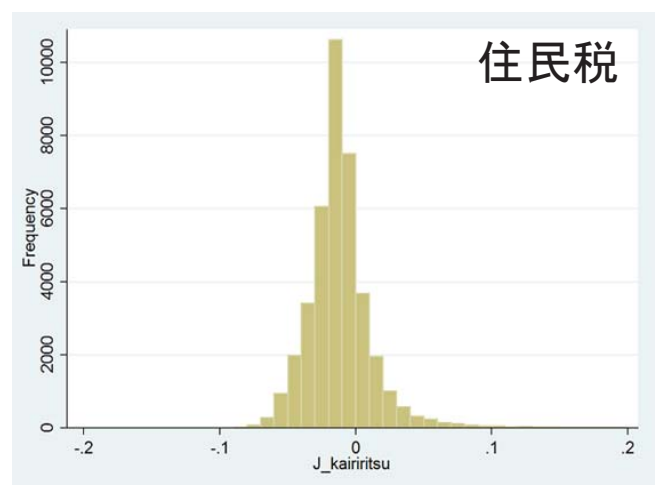
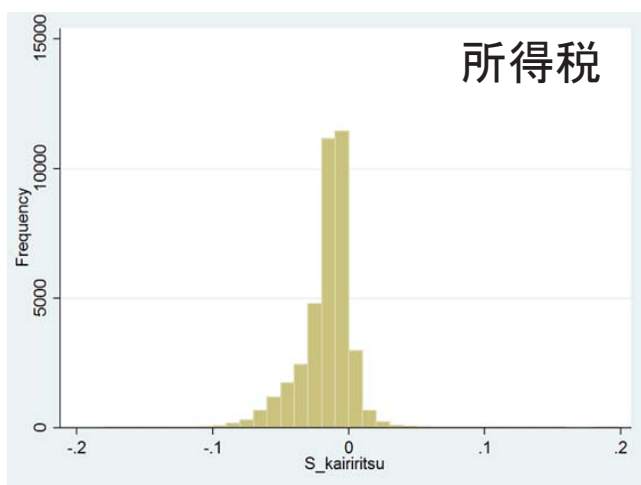
12

保険料の理論値計算

- 医療保険料：
 - 国民健康保険：
 - 所得割率、資産割、均等割、平等割の全国平均で推計
 - 賦課限度額・応益割に対する減額制度も反映
 - 健康保険：
 - 協会けんぽの保険料率を使用し、労使折半と仮定して推計
 - 標準報酬月額と標準賞与額の上限も反映
 - 後期高齢者医療制度：
 - 所得割率、均等割の全国平均を使用して推計
- 賦課限度額、所得割・均等割に対する減額制度も反映年金保険料：
 - 国民年金：定額保険料。免除制度は該当者全員に適用する。
 - 厚生年金：
 - 第1種被保険者の保険料率（労使折半）を使用、標準報酬月額と標準賞与額の上限も反映
- 介護保険料：
 - 第1号被保険者：県別平均保険料基準額を使用して推計
 - 第2号被保険者：（医療保険と同様）
- 雇用保険料：
 - 一般の事業に関する労働者の負担率を使用して推計

13

推計税額の妥当性： 国民生活基礎調査の事例



* 国民生活基礎調査では、調査される税額等は信頼性が高いことは確認されている

(注1) 乖離率 = (記入値 - 理論値) / 世帯所得

(注2) 勤労者世帯の結果

14

季節性の調整と年換算

全国消費実態調査 (YYYY年)			家計調査 (YYYY年)			
		消費支出	世帯属性	年間消費額	9・10・11月の消費額	比率
1						
2			25-29歳 300万円以下 政令指定都市	200万円	40万円	5
3						
4			25-29歳 600万円以下 政令指定都市	300万円	50万円	6
5						
...						
...			

15

耐久財等の高額支出

全国消費実態調査 (YYYY年)			家計消費状況調査 (YYYY年)	
	家計消費状況調査非対象品目	家計消費状況調査対象品目	世帯属性	年間消費額
1		3万円		
2		15万円	25-29歳 300万円以下 政令指定都市	40万円
3		1万円		
4		180万円	25-29歳 600万円以下 政令指定都市	55万円
5				
...				
...		

16

「授業料等」

全国消費実態調査 (YYYY年)				家計消費状況調査 (YYYY年)	
通学中の世帯員1		2	3	世帯属性	一人当たり 年間授業料
学校の種類		小学校 国公立	4 万円
1	小学校	国公立		大学 私立	130 万円
2	大学	私立	
3					
4					
5					
...					
...					

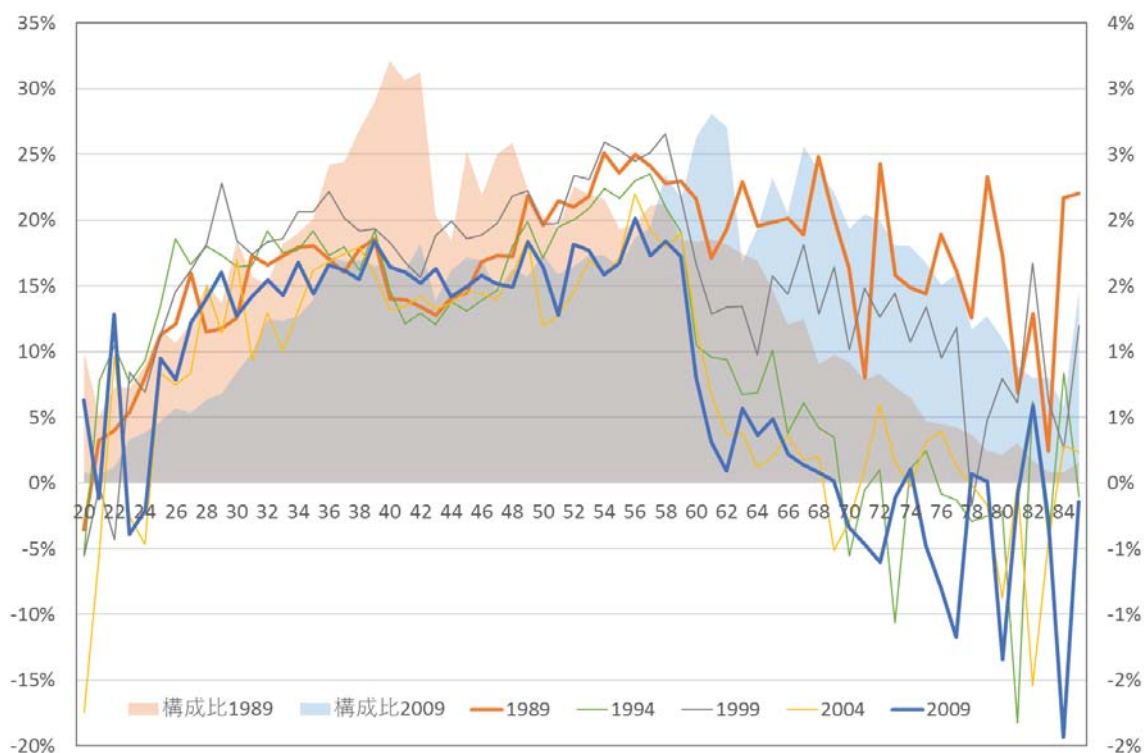
17

推計可処分所得・消費・貯蓄率

		推計結果		全国消費 実態調査 (月平均x12)	家計調査	国民経済 計算
		総世帯	2人以上の世帯 ・勤労者	2人以上の世帯 ・勤労者	2人以上の世帯 ・勤労者	家計 (総貯蓄) (兆円/年)
1989	可処分所得	518.9	577.6	451.2	505.7	262.6
	消費支出	424.7	475.5	377.9	379.8	214.1
	貯蓄率	18.2%	17.7%	16.2%	24.9%	18.5%
1994	可処分所得	566.9	656.8	532.6	580.9	322.4
	消費支出	483.3	544.9	428.0	423.7	263.8
	貯蓄率	14.7%	17.0%	19.6%	27.1%	18.2%
1999	可処分所得	551.2	670.5	545	580.7	334.4
	消費支出	448.9	535.4	423.1	415.4	278.8
	貯蓄率	18.6%	20.2%	22.4%	28.5%	16.6%
2004	可処分所得	493.9	609.2	510.6	535.5	316.7
	消費支出	439.4	519.9	407.1	398.0	283.5
	貯蓄率	11.0%	14.7%	20.3%	25.7%	10.5%
2009	可処分所得	466.2	585.0	472.1	513.5	314.9
	消費支出	419.5	492.9	384.2	382.9	280.5
	貯蓄率	10.0%	15.7%	18.6%	25.4%	10.9%

18

年齢階級別貯蓄率



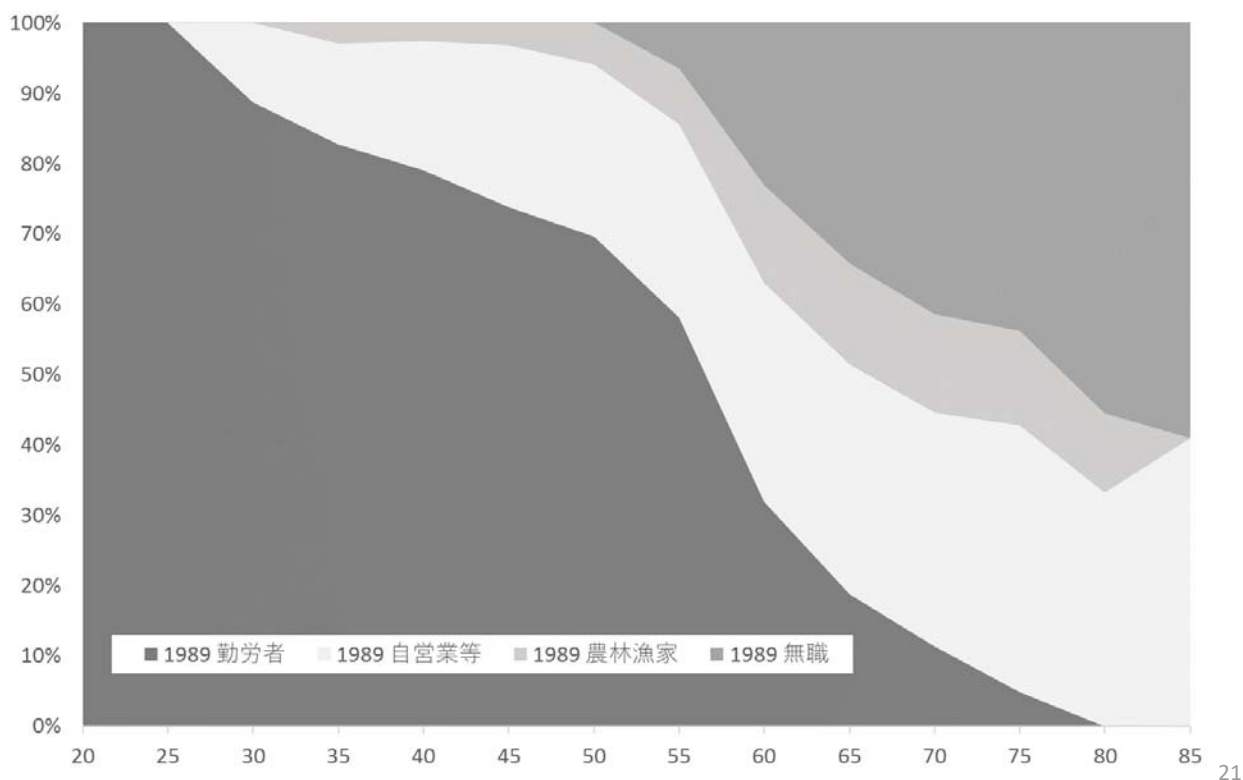
19

高齢化の影響

	貯蓄率	前回調査との差	人口構成変化効果	貯蓄率変化効果	人口構成変化効果の割合
1989	18.2%	-	-	-	
1994	14.7%	-3.4%	-0.2%	-3.2%	
1999	18.6%	3.8%	-0.3%	4.1%	
2004	11.0%	-7.5%	-0.6%	-7.0%	
2009	10.0%	-1.0%	-0.9%	-0.2%	
合計		-8.1%	-1.9%	-6.2%	23.9%
1989 vs 2009			-3.0%	-5.1%	37.3%

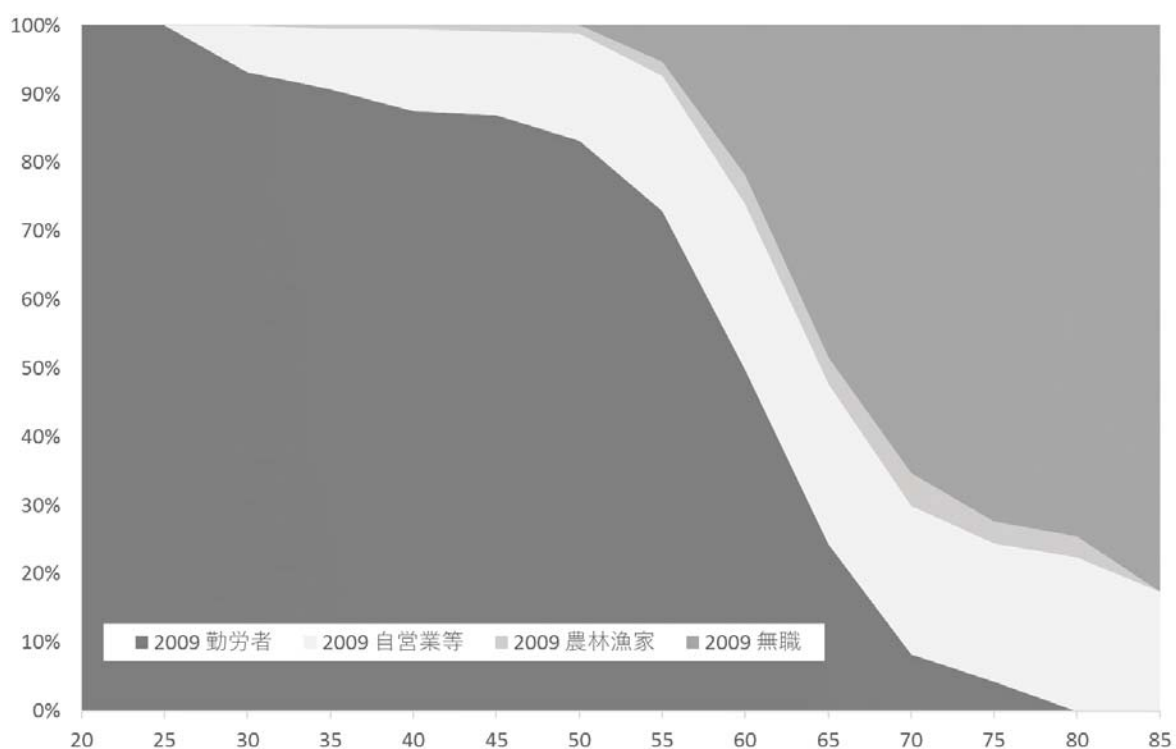
20

職業別構成比：1989年



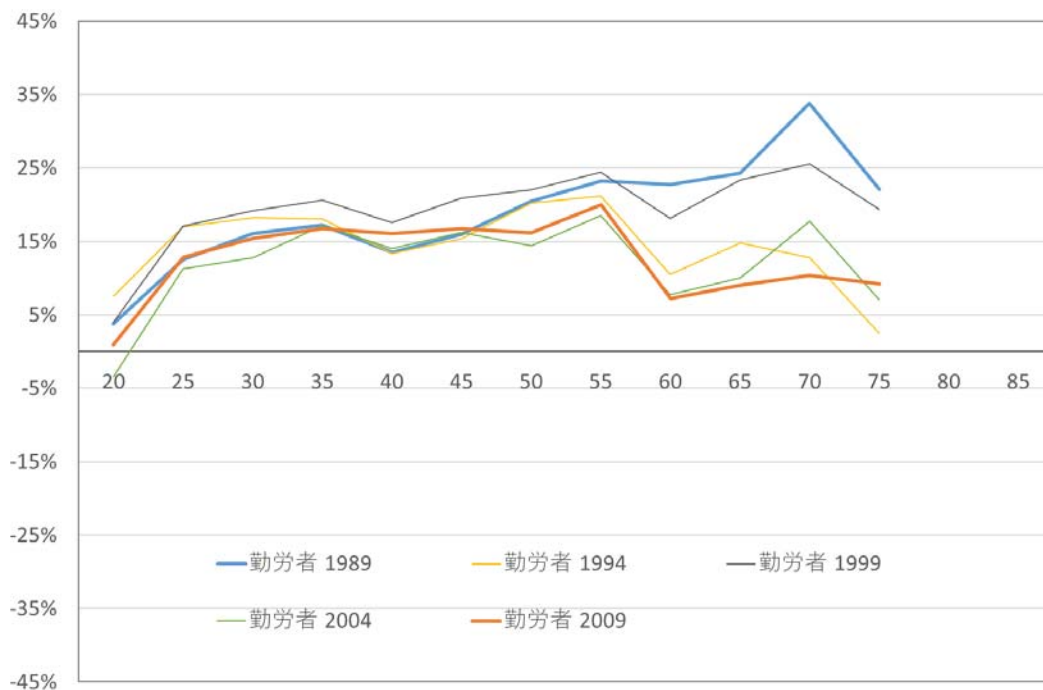
21

職業別構成比：2009年



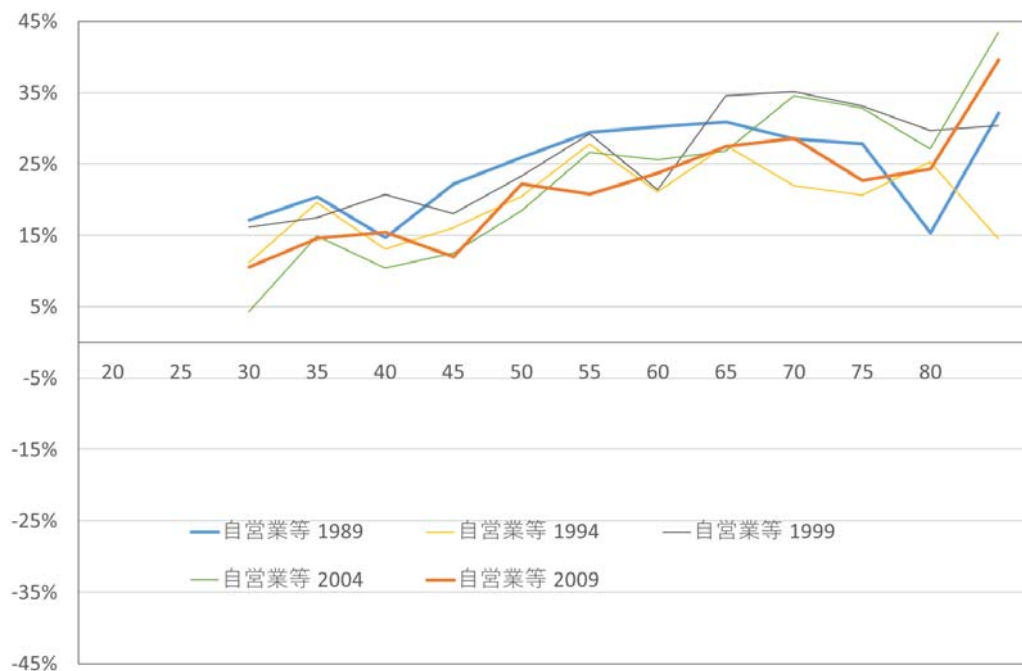
22

年齢別・職業別貯蓄率の推移 勤労者世帯



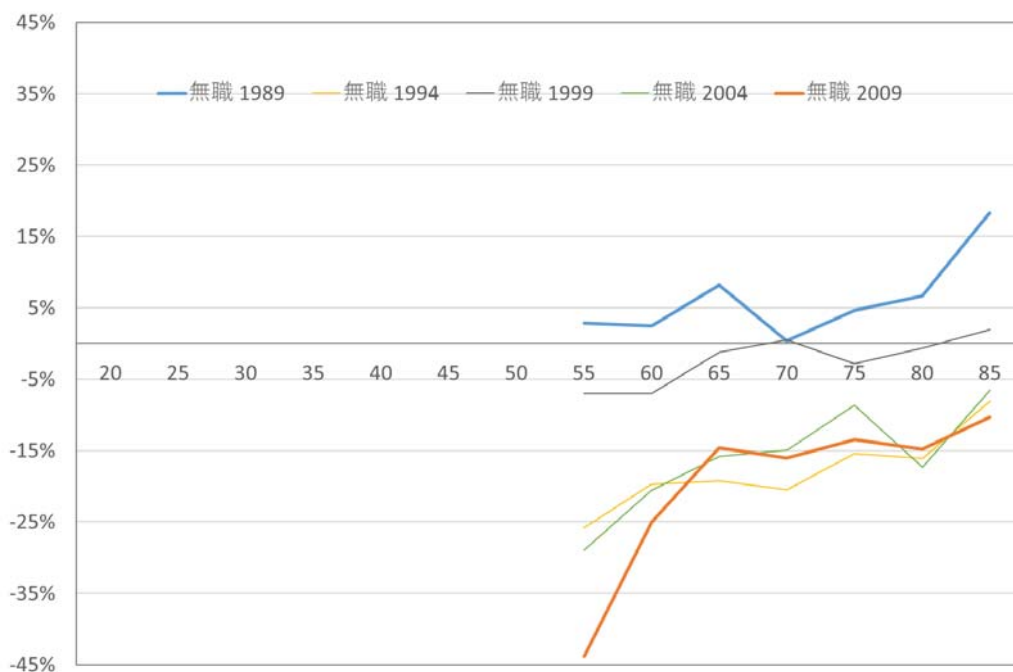
23

年齢別・職業別貯蓄率の推移 自営業等世帯



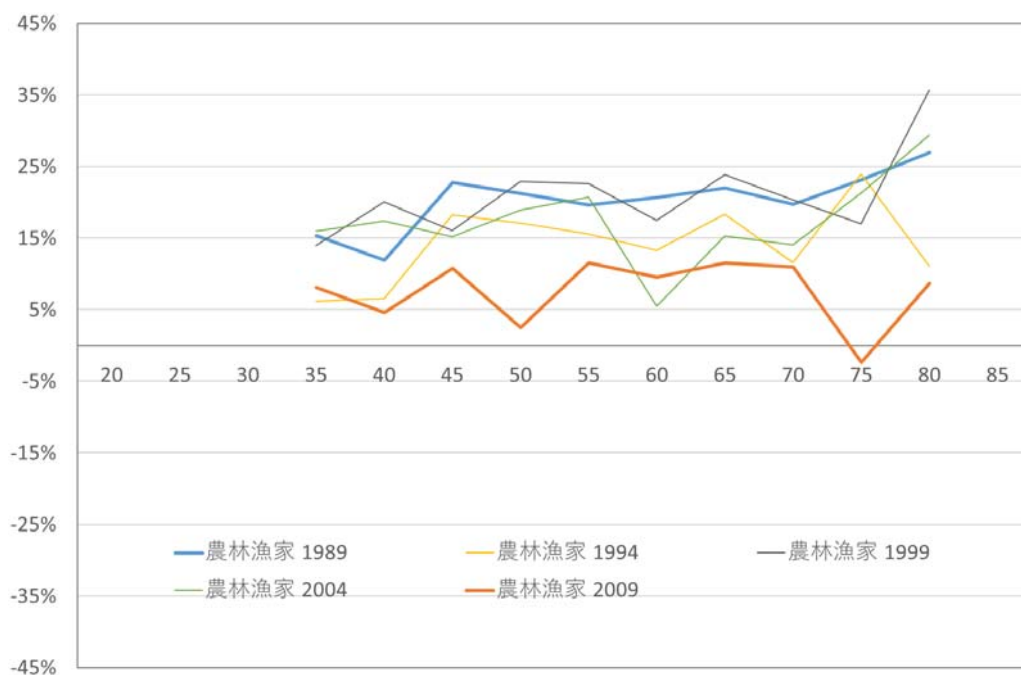
24

年齢別・職業別貯蓄率の推移 無職世帯



25

年齢別・職業別貯蓄率の推移 無職世帯



26

高齢化・サラリーマン化の影響

	貯蓄率	前回調査との差	人口構成変化効果	貯蓄率変化効果	人口構成変化効果の割合
1989	18.2%	-	-	-	
1994	14.7%	-3.4%	-0.1%	-3.3%	
1999	18.6%	3.8%	-0.7%	4.5%	
2004	11.0%	-7.5%	-1.2%	-6.4%	
2009	10.0%	-1.0%	-0.9%	-0.1%	
合計		-8.1%	-2.9%	-5.2%	
1989 vs 2009			-3.5%	-4.7%	37.3%

27

結論

- マクロ統計と統合的なミクロ統計は作成可能
- 1989年から2009年で貯蓄率は8%程度低下
- その大部分は「貯蓄行動の変化」
 - 高齢化という人口構成の変化で説明可能な落ち込みは2-3%程度
 - 就業構造の変化を考慮した人口構成の変化で説明可能な落ち込みは3-3.5%程度
- 引退した高齢無職世帯の貯蓄率の低下が顕著
 - なぜ低下したのかを分析する必要

28

景気指標としての 個人消費関連統計について

2017年2月

科研プロジェクト研究集会

「経済統計・政府統計の理論と応用」

日本大学経済学部 川崎 茂

1

問題意識と背景

個人消費動向への関心の高まり

- 2014年4月の消費税率引き上げ以降、個人消費が低迷

→ 個人消費関連統計への関心の高まり

→ 政府における統計改善の検討

(例 経済財政諮問会議「統計改革の基本方針」(2016年12月)など)

個人消費動向のよりの確な把握に必要なことは？

- 基礎となる個別統計の評価と適切な活用
 - 需要側の統計、供給側の統計の特性の把握
- 個人消費の背景にある長期的傾向の変化に対する理解
 - 高齢化に伴う世帯構成の変化等の消費への影響の把握
- 「景気動向」として注目すべき消費の対象範囲の検討

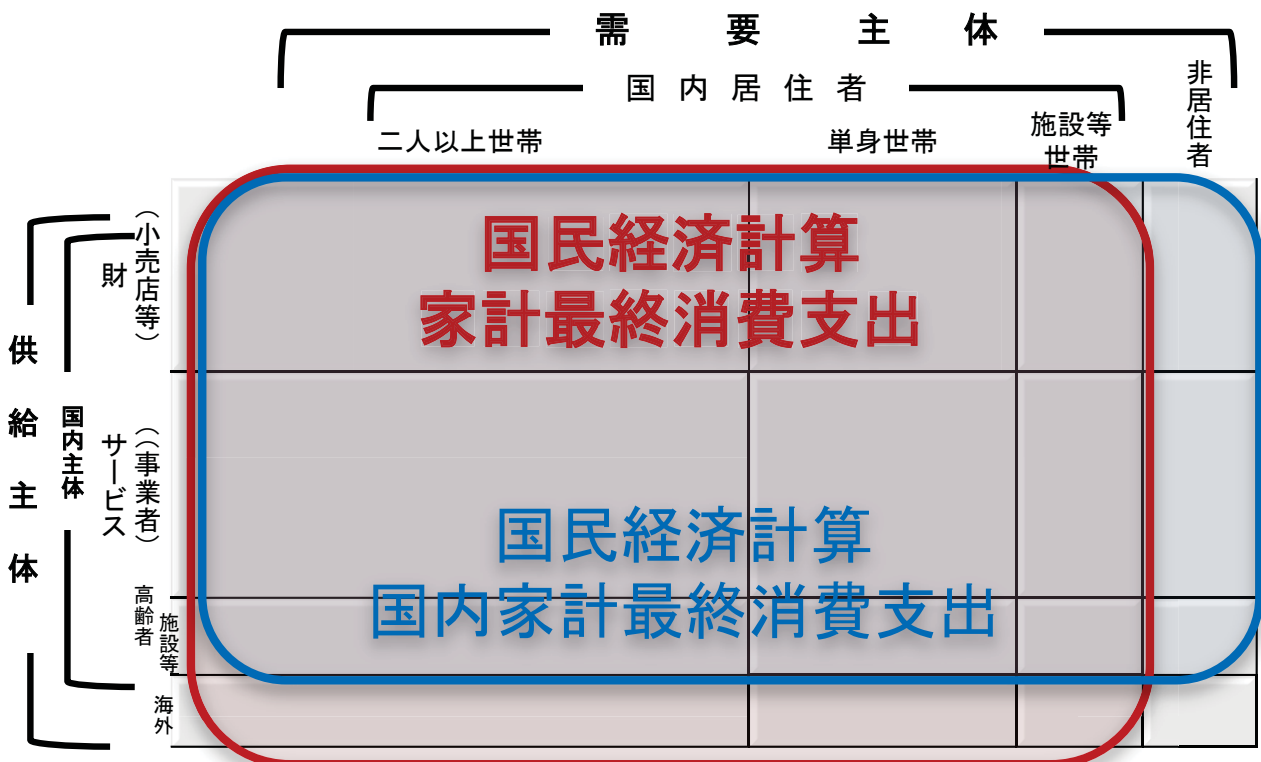
2

目次

1. 個人消費を把握する枠組み
2. 一次統計の状況 — 需要側と供給側
3. 加工統計の状況 — 両面推計、供給側推計
4. 需要側統計の課題
5. 供給側統計の課題
6. まとめ

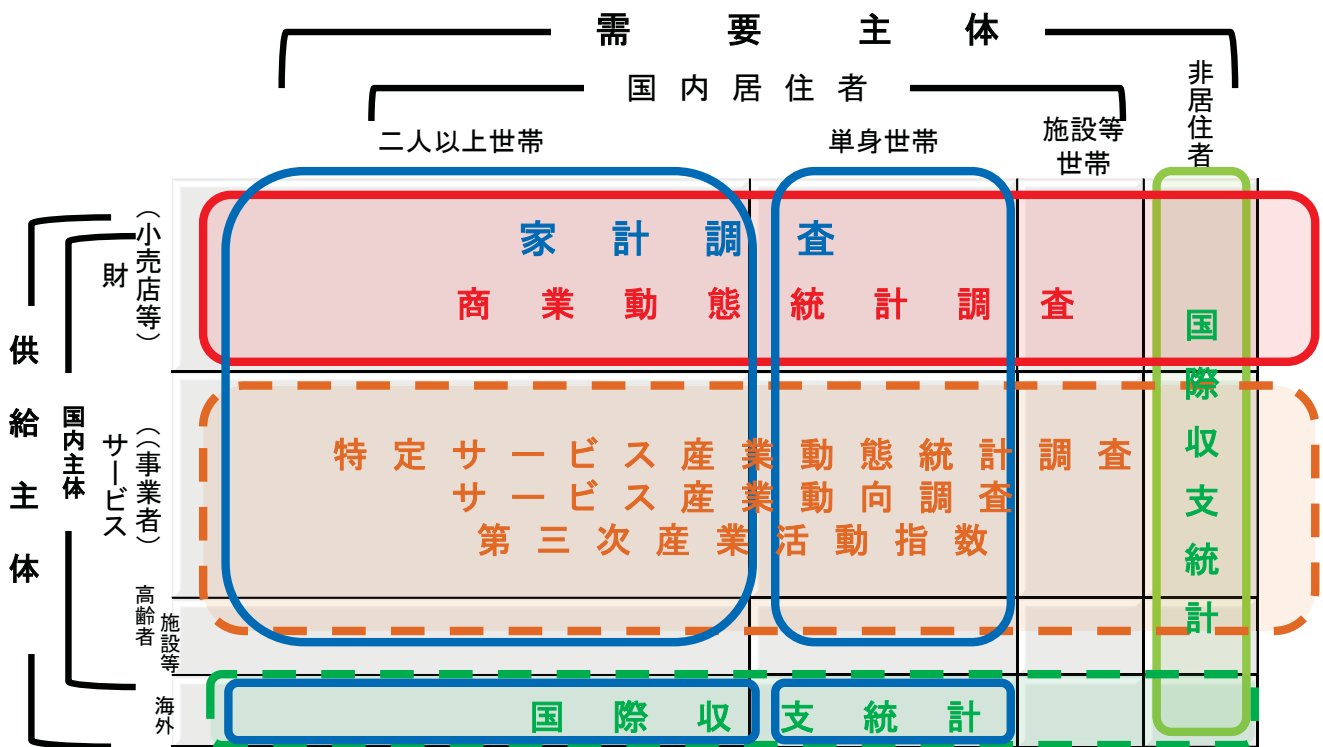
3

1. 個人消費を把握する枠組み



4

個人消費に関する基礎統計



5

主な個人消費関連統計

統計名	作成機関 周期	概要
家計最終消費支出 (国民経済計算)	内閣府 (四半期)	需要側推計, 供給側推計等の結果を統合。需要側では家計調査等を, 供給側では生産動態統計, 特定サービス産業動態統計等を推計に使用。国内家計消費が対象。持家の帰属家賃を含む。
消費支出 (家計調査)	総務省 (月次)	需要側の調査。約9千世帯の標本調査。家計簿支出を集計。1世帯当たり平均値。単身を含む結果もあるが, 二人以上世帯結果が目される。自営業世帯は消費のみ, 勤労者・高齢者世帯は収入も調査。
小売販売額統計 (商業動態統計調査)	経産省 (月次)	供給側の調査。卸売・小売141万事業所を抽出。小売総販売額を推計, 目標精度は5%以下。大規模な百貨店, スーパー, コンビエンスストア等は別に調査し結果を合算。商品等を10区分に分類表章。
消費総合指数	内閣府 (月次)	SNAベースで家計消費(帰属家賃含む)を総合的にとらえる指標として2001年から公表, その後逐次改良。QEの推計方法を簡略化して使用。需要側, 供給側の統計を組み合わせ推計。
消費活動指数	日本銀行 (月次)	SNAの家計消費(帰属家賃除く)の実勢を包括的にとらえる指標として2016年5月から公表, 10月に改良。供給側統計のみ使用。財は主に商業動態統計, サービスは第三次産業活動指数, 業界統計を使用。

6

2. 一次統計の状況 — 需要側と供給側

- ・需要側統計は「家計調査」, 「家計消費動向調査」など世帯調査が主体
- ・供給側統計は「商業動態統計調査」, 「特定サービス産業動向調査」など事業所・企業調査が主体

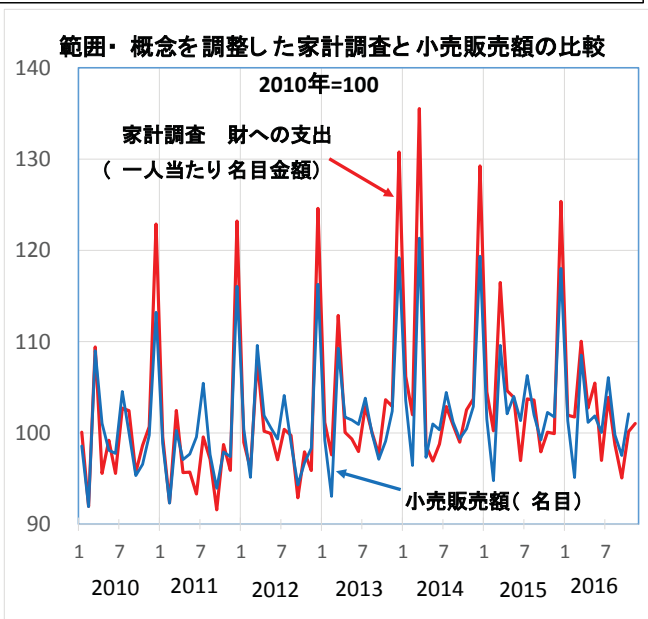
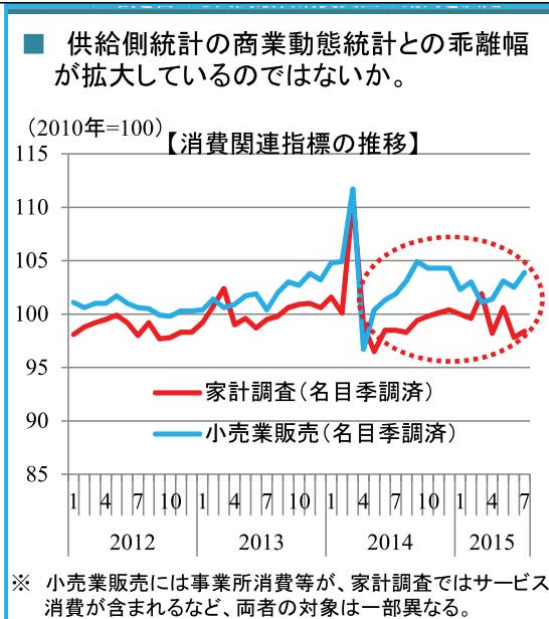
	家計調査	商業動態統計調査
調査方法等	<ul style="list-style-type: none"> ・ 全国の世帯を無作為抽出 ・ 毎月約9千世帯を調査(うち二人以上が約8千世帯) ・ 調査世帯の記入する家計簿のデータを集計 ・ 世帯人員, 世帯主の年齢等世帯属性も調査 	<ul style="list-style-type: none"> ・ 全国の小売業, 卸売業の店舗・企業を調査 ・ 一定規模以上及び特定業態等については悉皆, 小規模は標本調査 ・ 小売販売高の誤差率5%以下が目標 ・ 大型家電専門店, コンビニ, ドラッグストア, ホームセンターは企業調査
統計の概念	<ul style="list-style-type: none"> ・ 消費支出等は1世帯平均 ・ 二人以上世帯の動向が注目され ・ 施設等の居住者は調査の対象外 ・ 中古品購入, 世帯間移転も計上 ・ 発生主義と現金主義が混在(公共料金など支払い時点で計上) 	<ul style="list-style-type: none"> ・ 小売業事業者の販売額の総額 ・ 財(商品)の販売額が中心 ・ 家賃, サービス(電気・ガス・水道代, 学費, 医療費等)は含まれない ・ インバウンド消費を含む ・ 一物品目には企業対象の販売も計上

7

家計調査と商業動態統計のギャップに関する誤解

両統計の概念・範囲の違いを無視して比較しがち。
(例 経済財政諮問会議資料)

統計の概念・範囲を調整(財の購入総額を近似)
→ 数値の動きはかなり一致



(出典) 内閣府 経済財政諮問会議 平成27年10月16日資料4 「企業収益等の動向/基礎統計の更なる充実について」麻生議員提出資料からそのまま転載

(資料)総務省「家計調査」, 経済産業省「商業動態統計」から筆者作成。
(注)この期間, 人口減少率は年率0.1%程度であることから, 家計調査の一人当たり支出金額の動きは総人口ベースの動きに近い。

8

留意点

- 家計調査の結果利用に当たって、標本誤差・非標本誤差に留意が必要。商業動態統計にも同様の課題はあるが、数値は比較的安定。
- 家計調査は1世帯当たり平均、商業動態統計は販売総額、直接比較は不適切
- 家計調査では、店舗形態によらず、財・サービス等すべての購入・支出が対象、商業動態統計では、小売店での販売が対象
 - ➔ 両統計の比較に当たっては、概念を調整して使用することが必要
 - インターネット購入、サービス消費など小売店舗を經由しない消費については、需要側統計からの把握するほうが精度向上の可能性が高い
 - ➔ 需要側、供給側を組み合わせる消費総額を推計するのは合理的

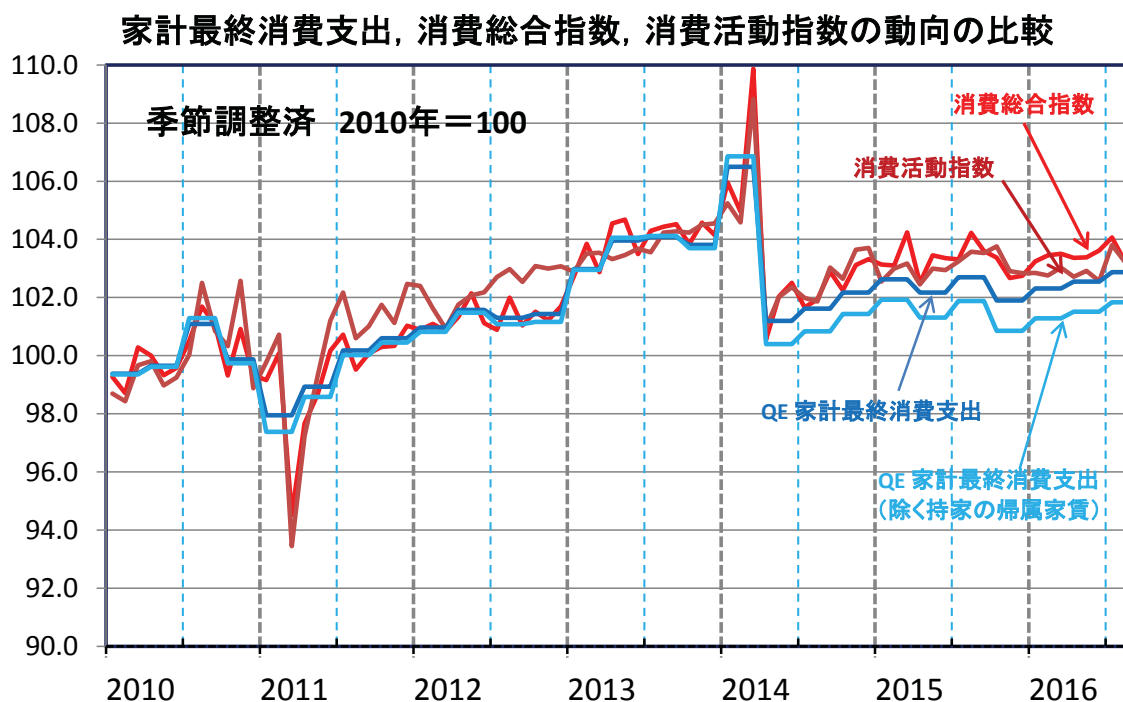
9

3. 加工統計の状況 — 両面推計、供給側推計

	消費総合指数	消費活動指数
作成主体	内閣府 2003年から作成	日本銀行 2016年から作成
公表内容 公表時期	実質季節調整系列(月次, 四半期) 「月例経済報告」と同時に公表, 同報告の検討資料としての位置付け	名目・実質季節調整系列(月次, 四半期) 耐久財、非耐久財、サービス別内訳 毎月第5営業日に公表
対象範囲等	• QE(旧基準)の「国内家計最終消費支出」(帰属家賃, インバウンド消費含む)を近似	• QE(旧基準)の「家計最終消費支出(除く持家の帰属家賃)」を近似 • インバウンド消費を含む・除く二系列
作成方法	• QEと同様, 需要側推計と供給側推計を加重平均して推計 • 需要側は家計調査, 家計消費状況調査等から推計 • 供給側は生産動態統計, 鉱工業出荷指数, 法人企業統計等から推計 • 住宅賃貸料, 医療・介護サービス, 電気水道等については共通の推計値を使用	• 振れが少ないと思われる販売・供給統計のみから推計 • 財は商業動態統計, サービスは特定サービス産業動態統計, 第三次産業活動指数等から推計 • ほかに一部を業界統計で補足 • 各指標を加重平均, ウェイトは国民経済計算, 家計調査等から作成

10

消費総合指数と消費活動指数の動き



(資料) 内閣府「国民経済計算四半期速報」, 内閣府「月例経済報告」, 日本銀行「消費活動指数」に掲載のデータから筆者が作成。家計最終消費支出は四半期系列, その他は月次系列を表示。

11

留意点

- 「消費総合指数」は「月例経済報告」の参考資料として公表。
 - 基礎データはすべて公開情報。ただし, 計算過程の名目系列等は非公表。
 - 推計方法の改良を随時実施。改良後の系列のみ利用可能。
 - 「消費活動指数」は, 耐久財, 非耐久財, サービスの別の指数も公表。
 - 両指数で実質化の方法が異なっており, これも差異が生じる一因。
- ➔ 両指数とも, 推計手法の改定が随時行われており, 今後手法が定着した段階で特性を分析することが必要

12

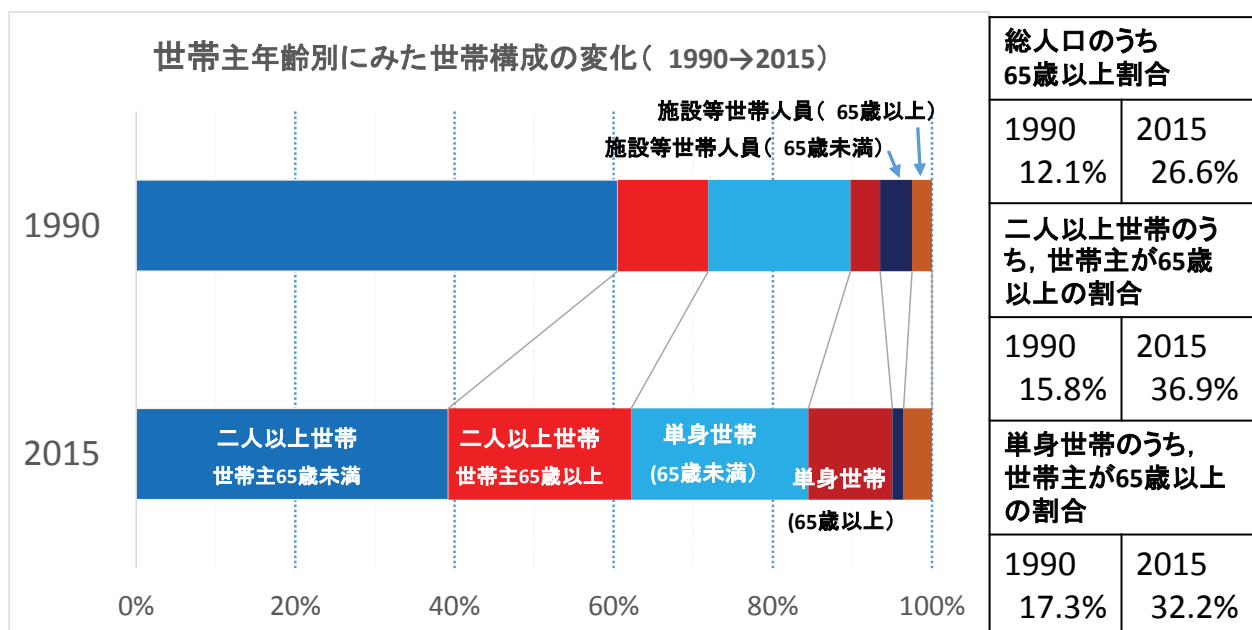
4. 需要側統計の課題

- 家計調査結果に対する主な批判として、①標本誤差、②非標本誤差、③項目別の測定精度、④調査対象世帯のカバレッジ、⑤より詳細なデータ提供 などに関するものがある。
- これに対して、家計調査を担当する総務省では、①3か月移動平均値の公表、②世帯人員・世帯主年齢別分布を調整した結果の公表 などを行っている。
- さらに今後、③オンライン・スマートフォン家計簿の導入、家計簿記入の簡略化、④単身モニター調査の導入 などの可能性が検討中とのことである。
- 以下では、⑤に関連して、需要側統計のより適切な解釈に資するための情報について検討する。

13

調査対象世帯の構成の変化

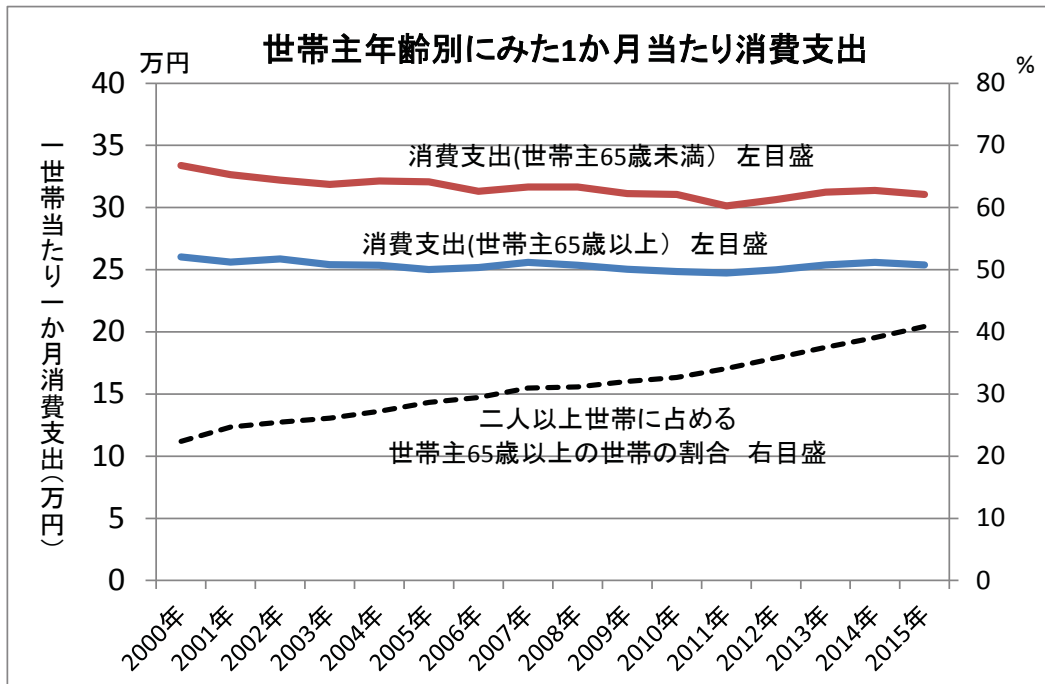
人口の高齢化とともに、世帯主の高齢化も急速に進展。
→ 世帯の構成の変化が一世帯当たり平均値に影響



資料：総務省「国勢調査」

14

世帯主の高齢化は消費支出に影響

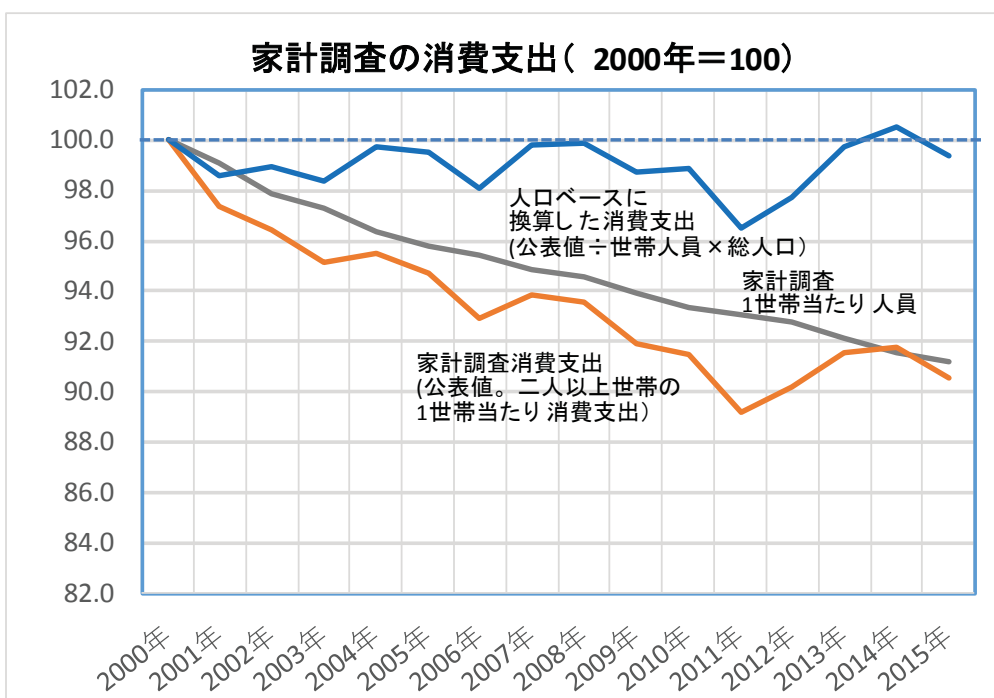


仮に世帯主の年齢構成が2000年のまま不変だった場合、二人以上世帯の平均の消費支出は2000年 2015年 31.7万円→29.8万円となる。実際には、2000年 2015年 31.7万円→28.7万円となった。(年齢構成の変化の影響は▲4.5%)

同様の影響は、所得面でも生じていると考えられる。

資料：総務省「家計調査」(二人以上世帯の結果から作成)

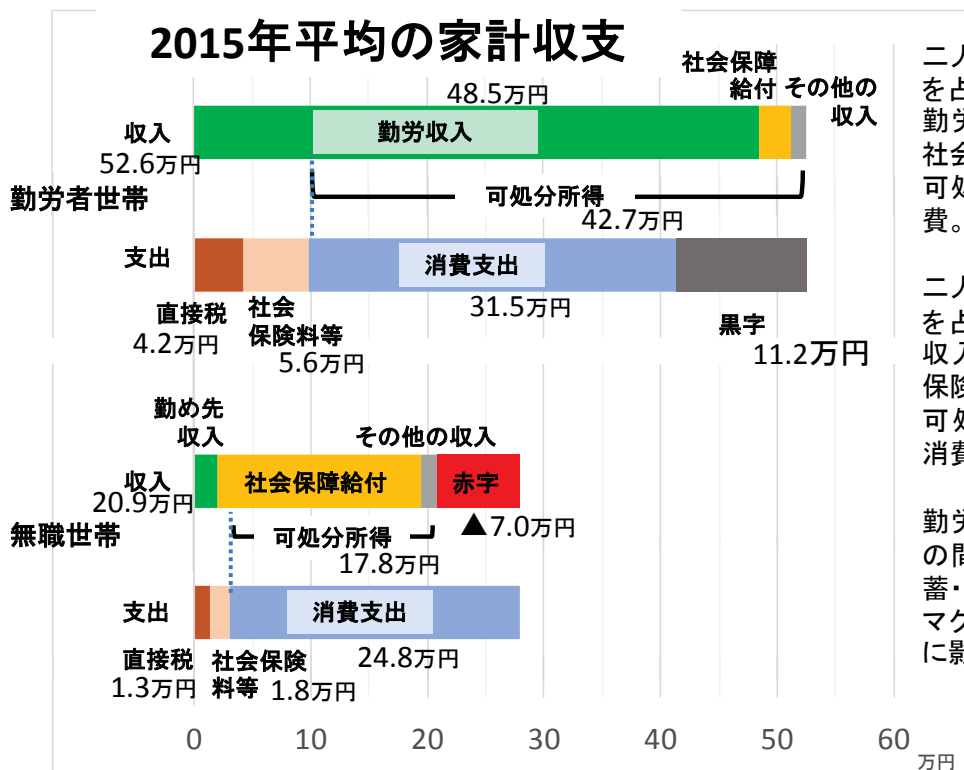
世帯規模の縮小は平均値に影響



家計調査の公表結果(1世帯当り平均)からは、消費が低下傾向にあると理解されがち。この間、1世帯当たり人員が低下傾向にあることに注意が必要。公表結果を人口ベースに換算すると、それほど大幅な低下は見られない。(換算は大雑把な前提に基づいたものであり、解釈には注意が必要)

資料：総務省「家計調査」(家計収支編 二人以上の世帯第1表)及び同「人口推計」(各年10月1日現在)のデータを基に川崎が計算

勤労者と無職者の家計の比較



二人以上世帯の約半数を占める勤労者世帯は、勤労収入の20%を税・社会保険料として負担。可処分所得の74%を消費。

二人以上世帯の約30%を占める無職世帯は、収入の15%を税・社会保険料として負担。可処分所得の139%を消費。

勤労者・無職者の世帯の間の所得再分配、貯蓄・消費性向の差異は、マクロの家計消費動向に影響を与えている。

資料: 総務省「家計調査」家計収支編 二人以上の世帯(勤労者世帯、無職世帯)より作成。
 グラフ中の金額は、1世帯当たり月額である。

5. 供給側統計の課題

- 商品に関する供給側統計として有力なのは、**小売販売額統計(商業動態統計)**
- 広範囲な小売りをカバーするが、販売額(平成27年:140.7兆円)には、事業者向け販売が含まれる、無店舗販売が含まれないなどの課題が存在
- サービスに関する供給側統計としては、**特定サービス産業動態統計調査**、**サービス産業動向調査**、**第3次産業活動指数**、**業種ごとの業務統計**、**業界統計**など
- サービスは、業種が多様である、シェアの高い大規模な特定事業者が存在しない業種が多いなどの理由から、カバレッジの高い統計調査を中心にして、この分野の統計調査の再編を進めるとともに、業務統計、業界統計等の有効活用が必要

特定サービス産業動態統計調査と サービス産業動向調査

- 「特定サービス産業動態統計調査」は、特定のサービス業種の月次動向を把握するため、昭和62年に開始（経済産業省）
- 時期に応じて調査対象の業種を変更し、現在は19業種を調査
- 「サービス産業動向調査」は、サービス業の月次動向を把握するため、平成20年に開始（総務省）
- 産業大分類のG～Rに含まれる大半の業種を対象
- 月次調査では、月間売上高、事業従事者数等を調査
- 年1回、調査対象を拡大し、年間売上高等を追加調査

19

特定サービス産業動態統計調査

- 対象業種は経済産業省所管の主なサービス業
- 調査対象は有意抽出

全国の企業を対象

物品賃貸（リース）業、物品賃貸（レンタル）業、情報サービス業、
広告業、クレジットカード業、エンジニアリング業、パチンコホール、
葬儀業、結婚式場業、外国語会話教室、フィットネスクラブ、学習塾、
インターネット附随サービス業、機械設計業、環境計量証明業、
自動車賃貸業

全国の事業所を対象

遊園地・テーマパーク

特定の地域の事業所を対象

ゴルフ場、ゴルフ練習場、ボウリング場

20

サービス産業動向調査の概要

- 業種に応じて、企業又は事業所を対象として調査
- 調査対象は、事業所・企業母集団データベースから層化無作為抽出
- 月次調査では、約13,000企業、約25,000事業所を調査
- 拡大調査(年次)では、約9,500企業、約69,000事業所を調査
(「G情報通信業」については他の同様の調査が存在するため、調査対象から除外)

21

サービス産業動向調査の対象業種

大分類	中分類		備考
G 情報通信業	37 通信業 38 放送業 39 情報サービス業	40 インターネット附随サービス業 41 映像・音声・文字情報制作業	
H 運輸業、郵便業	42 鉄道業 43 道路旅客運送業 44 道路貨物運送業 45 水運業	46 航空運輸業 47 倉庫業 48 運輸に附帯するサービス業 49 郵便業(信書便事業を含む)	
K 不動産業、物品賃貸業	68 不動産取引業 69 不動産賃貸業・管理業	70 物品賃貸業	
L 学術研究、専門・技術サービス業	72 専門サービス業(他に分類されないもの) 73 広告業 74 技術サービス業(他に分類されないもの)		○ 中分類「71 学術・開発研究機関」、細分類「7282 純粋持株会社」を除く
M 宿泊業、飲食サービス業	75 宿泊業 76 飲食店 77 持ち帰り・配達飲食サービス業		
N 生活関連サービス業、娯楽業	78 洗濯・理容・美容・浴場業 79 その他の生活関連サービス業	80 娯楽業	○ 小分類「792 家事サービス業」を除く
O 教育、学習支援業	82 その他の教育、学習支援業		○ 中分類「81 学校教育」を除く
P 医療、福祉	83 医療業 84 保健衛生 85 社会保険・社会福祉・介護事業		○ 小分類「841 保健所」、「851 社会保険事業団体」及び「852 福祉事務所」を除く
R サービス業(他に分類されないもの)	88 廃棄物処理業 89 自動車整備業 90 機械等修理業(別掲を除く)	91 職業紹介・労働者派遣業 92 その他の事業サービス業 95 その他のサービス業	○ 中分類「93 政治・経済・文化団体」、「94 宗教」、「96 外国公務」を除く

資料:総務省「サービス産業動向調査の概要」

注 赤太文字は、特定サービス産業動態統計調査で調査されている業種(39, 40, 72, 74, 79, 80, 82)

22

6. まとめ

- 需要側統計，供給側統計の利用に当たっては，対象範囲・概念を適切に調整することが必要
- 統合指標については，時系列の蓄積を待って，家計最終消費支出の動向の近似性の良否を評価
- 需要側統計については，世帯の急速な高齢化の進展、世帯規模の縮小、所得の再分配等が所得と消費に影響を与えていることを考慮に入れて解釈・分析を行う必要
- 供給側統計については，サービス産業動向調査を中心に，消費関連のサービス業の動向を一括して月次で把握することが必要
- 供給側統計について，販売先の個人向け，事業者向けの別に関する情報を把握することが必要

23

主要参考文献等

- 宇南山卓 [2011] 「家計調査の課題と改善に向けて」 統計と日本経済
- 宇南山卓 [2015] 「消費関連統計の比較」 フィナンシャルレビュー
- 新家義貴 [2003] 「消費総合指数の改訂とその作成方法について」経済財政分析ディスカッション・ペーパー
- 吉田充 [2016] 「消費総合指数に関する諸課題への検討」経済財政分析ディスカッション・ペーパー
- 中村康治ほか [2016] 「消費活動指数について」 BOJ Reports & Research Papers
- 同 [2016] 「消費活動指数の公表内容の拡充と見直しについて」 同上
- 内閣府経済社会総合研究所 [2016] 「推計手順解説書(四半期別GDP速報(QE編)) 平成23年基準版」
- 同 [2012] 「推計手法解説書(年次推計編) 平成17年基準版」
- 内閣府 平成28年12月13日 「より正確な景気判断のための経済統計の改善に関する研究会 報告」
- 経済財政諮問会議 平成28年12月21日 「統計改革の基本方針」
- 総務省 平成28年7月25日 「家計調査の改善に関するタスクフォース取りまとめ」
- 総務省 「サービス産業動向調査の概要」 <http://www.stat.go.jp/data/mssi/gaiyo.htm#midashi4>
- 経済産業省 [2016] 「商業動態統計調査の概要等について」
- 経済産業省 「特定サービス産業動態統計調査 調査の概要」
- <http://www.meti.go.jp/statistics/tyo/tokusabido/gaiyo.html#menu01>

24

自記式調査における複数回答方式をめぐる諸問題

土屋 隆裕（統計数理研究所）

1

問38 **【リスト19】** つぎのうち、日本人の性質をあらわしていると思うコトバがあったら、いくつでもあげてください？（M.A.）

- 1 (ア) 合理的
- 2 (イ) 勤勉
- 3 (ウ) 自由を尊ぶ
- 4 (エ) 淡泊
- 5 (オ) ねばり強い
- 6 (カ) 親切
- 7 (キ) 独創性にとむ
- 8 (ク) 礼儀正しい
- 9 (ケ) 明朗
- 10 (コ) 理想を求める
- 11 その他〔記入〕
- 12 わからない

問39 **【リスト20】** かりに現在の日本の社会全体を、この表にかいてあるように5つの層に分けるとすれば、お宅は、このどれに入るとお考えですか？

- | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----|-----|-----|-----|-----|------|-------|
| (ア) | (イ) | (ウ) | (エ) | (オ) | | |
| 上 | 中の上 | 中の中 | 中の下 | 下 | その他 | わからない |
| | | | | | 〔記入〕 | 〕 |

問40 あなたは何党を支持していらっしゃいますか？

- | | |
|-------|---------------|
| 1 自民党 | 7 みんなの党 |
| 2 民主党 | 8 生活の党 |
| 3 公明党 | 9 新党改革 |
| 4 共産党 | 10 その他の政党〔記入〕 |
| 5 社民党 | 11 支持政党なし |

2

日本人の長所

日本人の長所は？

複数回答

- 合理的
- 勤勉
- 自由を尊ぶ
- 淡泊
- ねばり強い
- 親切
- 独創性にとむ
- 礼儀正しい
- 明朗
- 理想を求める

強制選択

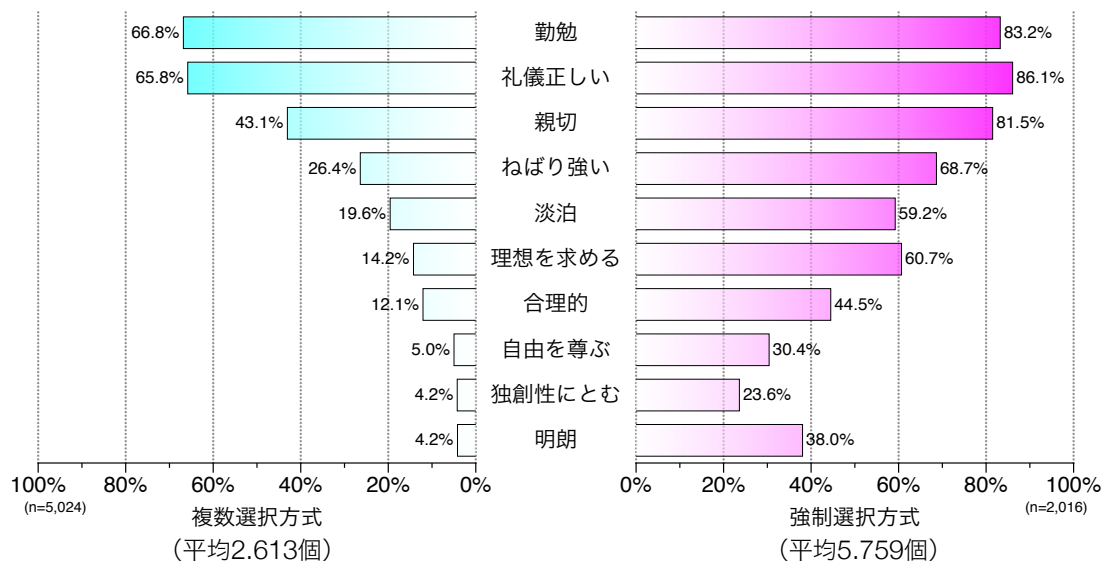
- 当てはまる
当てはまらない
- 合理的 ----- -----
 - 勤勉 ----- -----
 - 自由を尊ぶ ----- -----
 - 淡泊 ----- -----
 - ねばり強い ----- -----
 - 親切 ----- -----
 - 独創性にとむ ----- -----
 - 礼儀正しい ----- -----
 - 明朗 ----- -----
 - 理想を求める ----- -----

3

日本人の長所

日本人の長所は？

(インターネット調査)



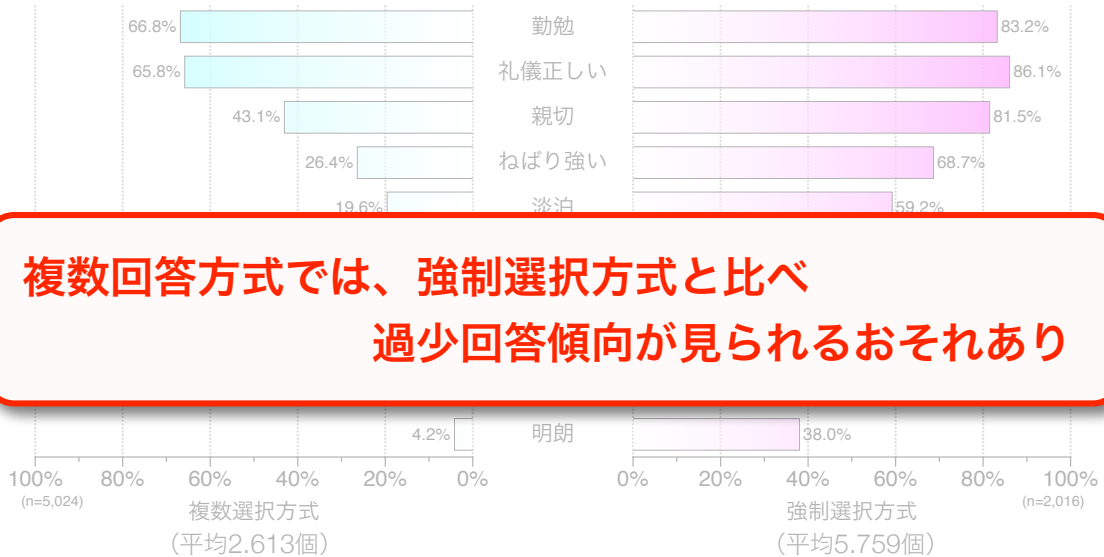
土屋 隆裕 (2014) 事例に見る調査票の設計と回答者の回答行動, マーケティングリサーチ, No.125, 24-32.

4

日本人の長所

日本人の長所は？

(インターネット調査)



**複数回答方式では、強制選択方式と比べ
過少回答傾向が見られるおそれあり**

土屋 隆裕 (2014) 事例に見る調査票の設計と回答者の回答行動, マーケティングリサーチャー, No.125, 24-32.

世論調査 内閣府大臣官庁政府広報室

3 調査票 生涯学習に関する世論調査

平成24年7月

(N=1,956)

1. 生涯学習の現状

Q1 [回答票1] あなたは、生涯学習という言葉から、どのようなイメージを持ちますか。この中から **いくつでもあげてください。** (M. A.)

- (46.2) (ア)幼児期から高齢期まで、生涯を通じて学ぶこと
- (40.6) (イ)趣味・教養を高めること
- (18.1) (ウ)職業上必要な知識・技能を身に付けること
- (34.7) (エ)高齢者の生きがいづくり
- (42.7) (オ)生活を楽しみ、心を豊かにする活動をする
- (21.2) (カ)健康・体力づくりをすること
- (30.9) (キ)公民館や生涯学習センターなど、公の機関での講座や教室における学習活動
- (19.1) (ク)かきかきセンターやスポーツクラブなど、民間の提供する講座や教室における学習活動
- (10.5) (ケ)学校における学習活動
- (0.5) その他
- (6.2) わからない

(M.T.=270.7)

【資料1】を提示して、調査対象者によく読んでもらってから、以下の質問をする。

【資料1】
「生涯学習」とは、人々が、生涯のいつでも、どこでも、自由に行う学習活動のことで、学校教育や、公民館における講座等の社会教育などの学習機会に限らず、自分から進んで行う学習やスポーツ、文化活動、趣味、ボランティア活動などにおけるさまざまな学習活動のことをいいます。

Q2 [回答票2] あなたは、この1年くらいの間に、生涯学習をしたことがありますか。この中から **いくつでもあげてください。** (M. A.)

- (25.7) (ア)趣味的なもの (音楽、美術、華道、舞踊、書道、レクリエーション活動など)
- (11.5) (イ)教養的なもの (文学、歴史、科学、語学など)
- (8.9) (ウ)社会問題に関するもの (社会・時事、国際、環境など)
- (30.4) (エ)健康・スポーツ (健康法、医学、栄養、ジョギング、水泳など)
- (14.1) (オ)家庭生活に役立つ技術 (料理、洋裁、相裁、編み物など)
- (8.9) (カ)育児・教育 (家庭教育、幼児教育、教育問題など)
- (15.2) (キ)職業上必要な知識・技能 (仕事に関係のある知識の習得や資格の取得など)
- (9.9) (ク)情報端末 (コンピュータやタブレット端末など) やインターネットに関すること (プログラムの使用、ホームページの作り方など)
- (8.1) (ケ)ボランティア活動のために必要な知識・技能
- (6.2) (コ)自然体験や生活体験などの体験活動
- (2.8) (サ)学校 (高等・専門・各種学校、大学、大学院など) の正規課程での学習
- (0.4) その他
- (42.5) (シ)生涯学習をしたことがない
- (0.4) わからない

→SQ
bへ
→Q3
へ
(M.T.=184.9)

Q2で(ア)～(サ)、「その他」と答えた方に)

SQ a1 [回答票3] あなたは、何のために生涯学習をしていますか。この中から **いくつでもあげてください。** (M. A.)

(N=1,117)

国民生活基礎調査（厚生労働省）

質問3 あなたはここ数日、病気やけがなどで体の具合の悪いところ（自覚症状）がありますか。

1 ある 2 ない → **質問4へ**

補問3-1 それは、どのような症状ですか。あてはまるすべての症状名の番号に○をつけてください。その中で最も気になる症状名の番号を番号記入欄に記入してください。

全身症状	01 熱がある	呼吸器系	15 せきやたんが出る	筋骨格系	29 肩こり
	02 体がだるい		16 鼻がつまる・鼻汁が出る		30 腰痛
	03 眠れない		17 ゼイゼイする		31 手足の関節が痛む
	04 いらいらしやすい		18 胃のもたれ・むねやけ		32 手足の動きが悪い
	05 ものを忘れる		19 下痢		33 手足のしびれ
眼	06 頭痛	消化器系	20 便秘	手足	34 手足が冷える
	07 めまい		21 食欲不振		35 足のむくみやだるさ
	08 目のかすみ		22 腹痛・胃痛		泌尿生殖器系
09 物を見づらい	23 痔による痛み・出血など	37 頻尿(尿の出る回数が多い)			
耳	10 耳なりがする	歯	24 歯が痛い	損傷	38 尿失禁(尿がもれる)
	11 きこえにくい		25 歯ぐきのはれ・出血		39 月経不順・月経痛
胸部	12 動悸	皮膚	26 かみにくい	損傷	40 骨折・ねんざ・脱ぎゅう
	13 息切れ		27 発疹(じんま疹・できものなど)		41 切り傷・やけどなどのけが
	14 前胸部に痛みがある		28 かゆみ(湿疹・水虫など)		42 その他

最も気になる症状の番号記入欄 → 番

国民生活基礎調査（厚生労働省）

質問4 あなたは**現在**、傷病（病気やけが）で病院や診療所（医院、歯科医院）、あんま・はり・きゅう・柔道整復師（施術所）に通っていますか。（往診、訪問診療、補問3-1の症状で通っているものを含む）

1 通っている 2 通っていない → **質問5へ**

補問4-1 どのような傷病（病気やけが）で通っていますか。あてはまるすべての傷病名の番号に○をつけてください。その中で最も気になる傷病名の番号を番号記入欄に記入してください。

内分泌・代謝障害	01 糖尿病	呼吸器系	15 急性鼻咽喉頭炎(かぜ)	泌尿生殖器系	32 腎臓の病気
	02 肥満症		16 アレルギー性鼻炎		33 前立腺肥大症
	03 脂質異常症 (高コレステロール血症等)		17 慢性閉塞性肺疾患(COPD)		34 閉経期又は閉経後障害 (更年期障害等)
精神・神経	04 甲状腺の病気	消化器系	18 喘息	損傷	35 骨折
	05 うつ病やその他のこころの病気		19 その他の呼吸器系の病気		36 骨折以外のけが・やけど
循環器系	06 認知症	皮膚	20 胃・十二指腸の病気	損傷	37 貧血・血液の病気
	07 パーキンソン病		21 肝臓・胆のうの病気		38 悪性新生物(がん)
	08 その他の神経の病気 (神経痛・麻痺等)		22 その他の消化器系の病気		39 妊娠・産褥 (切迫流産、前置胎盤等)
	09 眼の病気		23 歯の病気		40 不妊症
循環器系	10 耳の病気	筋骨格系	24 アトピー性皮膚炎	損傷	41 その他
	11 高血圧症		25 その他の皮膚の病気		42 不明
	12 脳卒中(脳出血、脳梗塞等)		26 痛風		最も気になる傷病の番号記入欄 → <input type="text"/> 番
	13 狭心症・心筋梗塞		27 関節リウマチ		
	14 その他の循環器系の病気		28 関節症		
			29 肩こり症		
			30 腰痛症		
			31 骨粗しょう症		

国民生活基礎調査

第9表 性・年齢階級・症状（複数回答）別にみた有訴者率（人口千対）

(単位:人口千対)

性 症状	総数	9歳以下	10～19	20～29	30～39	40～49	50～59	60～69	70～79	80歳以上	平成25年	
											(再掲) 65歳以上	(再掲) 75歳以上
男												
有訴者率	276.8	204.7	175.2	168.7	214.4	234.3	271.0	338.5	448.0	528.1	439.9	506.1
熱がある	7.8	28.9	7.8	7.7	7.4	5.4	4.4	3.8	3.9	7.0	4.5	5.4
体がだるい	41.0	5.9	20.9	38.6	47.2	49.0	49.2	43.2	50.4	63.3	50.4	58.7
眠れない	23.0	1.0	4.4	14.4	20.2	23.5	23.0	30.4	45.8	47.8	42.9	47.4
いらいらしやすい	20.2	3.1	10.3	18.2	24.1	23.9	22.1	20.4	28.5	31.7	27.1	30.8
もの忘れする	29.1	0.8	2.5	8.7	10.0	12.8	19.8	34.8	83.8	147.4	84.8	125.6
頭痛	22.0	4.2	17.4	26.3	34.7	30.7	20.0	16.9	20.0	23.6	19.9	22.9
めまい	13.5	0.3	4.3	6.0	10.0	12.4	12.7	15.4	29.1	42.4	27.9	38.7
目のかすみ	36.1	0.7	3.4	10.1	12.8	24.5	36.8	54.0	95.1	114.9	89.4	111.3
物を見づらい	29.3	1.3	4.1	6.6	7.3	21.2	34.2	43.3	72.5	96.6	70.2	89.3
耳なりがする	26.3	0.4	2.4	5.2	7.8	13.9	25.3	50.7	74.5	62.0	66.8	70.5
きこえにくい	29.7	2.5	2.4	4.0	4.9	7.4	16.1	36.6	93.6	175.3	96.6	146.9
動悸	13.9	0.1	0.8	4.8	7.4	9.8	13.7	20.5	32.8	47.5	33.3	43.5
息切れ	18.4	0.3	1.9	3.6	5.0	7.5	14.4	28.2	53.4	77.9	53.0	70.5
前胸部に痛みがある	8.5	0.1	1.1	4.6	6.9	7.0	8.7	10.0	18.6	25.8	17.9	24.3
せきやたんが出る	50.4	71.4	26.4	30.9	37.0	34.7	37.7	55.5	85.8	112.3	84.9	102.7
鼻がつまる・鼻汁が出る	50.9	103.6	61.1	42.2	43.0	31.8	32.5	40.5	64.1	74.2	61.2	73.2
ゼイゼイする	10.8	18.3	5.9	5.1	6.0	5.5	6.7	10.3	20.8	33.4	21.1	28.5
胃のむねやけ	20.8	1.1	2.8	9.0	16.2	20.5	24.2	29.7	41.5	44.7	38.6	45.2
下痢	19.8	13.4	11.6	18.5	21.8	22.1	22.1	21.0	21.4	26.4	22.5	24.7
便秘	26.0	4.6	4.4	8.4	8.8	9.9	14.4	31.2	75.9	128.8	75.6	111.8
食欲不振	8.5	4.1	4.1	6.7	5.4	5.5	7.1	8.3	17.5	29.7	17.6	25.4
腹痛・胃痛	14.9	6.2	11.8	13.3	17.6	16.3	15.4	15.2	18.1	20.0	17.8	19.3
痔による痛み・出血など	9.1	0.5	0.7	6.2	7.8	9.3	10.3	13.4	15.4	19.0	16.1	18.0
歯が痛い	18.0	5.7	6.6	13.6	17.4	16.6	19.9	23.7	30.6	25.6	28.2	28.0
歯ぐきのはれ・出血	16.8	2.3	3.2	7.1	10.4	14.7	24.2	27.6	29.7	28.2	28.7	29.0
かみにくい	18.0	0.5	1.3	2.1	2.3	5.2	13.3	27.8	58.2	78.2	55.0	72.0
発疹(じんま疹・できものなど)	16.1	25.0	14.0	12.8	14.6	12.6	13.1	16.1	20.3	23.3	20.0	22.5
かゆみ(湿疹・水虫など)	39.1	34.0										
肩こり	60.2	0.2										
腰痛	92.2	0.1										
手足の関節が痛む	41.8	1.3										

<http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa13/dl/16.pdf>

目的

- ・（複数回答方式と強制選択方式のどちらが適切な方式か？）
- ・ 複数回答方式の過少回答傾向は**有無が明確な症状等の内容**でも見られる傾向か？
- ・ 複数回答方式の過少回答傾向が生じる**原因**は何か？
- ・ 複数回答方式の過少回答傾向を抑制する**方法**は？

※ 富田尚希医師（東北大学病院）との共同研究

Web実験調査の概要

- **調査票** : 6種類用意
- **調査対象** : 65歳以上個人
 - ネット調査会社の登録パネルから無作為に抽出
 - 各調査票に無作為に割当
- **調査期間** : 2016年2月24日 - 2月25日
- **標本サイズ** : 各調査票 500
 - 目標標本サイズが確保できるまでサイトにアクセス可

11

あなたは、ここ1ヶ月の間に次のようなことがありましたか？

複数回答方式

問1 あなたは、ここ1ヶ月の間に次のようなことがありましたか？(MA)

- 気が散りやすい、ひとつのことに集中しにくい
- よく眠れない、もしくは睡眠薬が欠かせない
- 日付や曜日を思い出すのに時間がかかる
- 調子が良い時と悪い時の違いが大きい
- ここ半年間で2kg以上体重が減った
- わけもなく疲れる、もしくは何もはじめる気がしない
- 目が回る、もしくはふわふわする
- 熱っぽい、もしくは熱が一時的に上がる
- よくつまずく、もしくは転びやすい
- 見えにくい、もしくは見えない
- 聞こえにくい、もしくは聞こえない
- 動作がおそい、ぎこちない
- 腰痛、もしくは肩・膝・その他の関節の痛み
- 背骨あるいは関節の変形
- 皮膚の荒れがある
- トイレが近い、もしくは間に合わない
- 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
- 噛むのに難儀する、もしくは食事中か食後にむせる

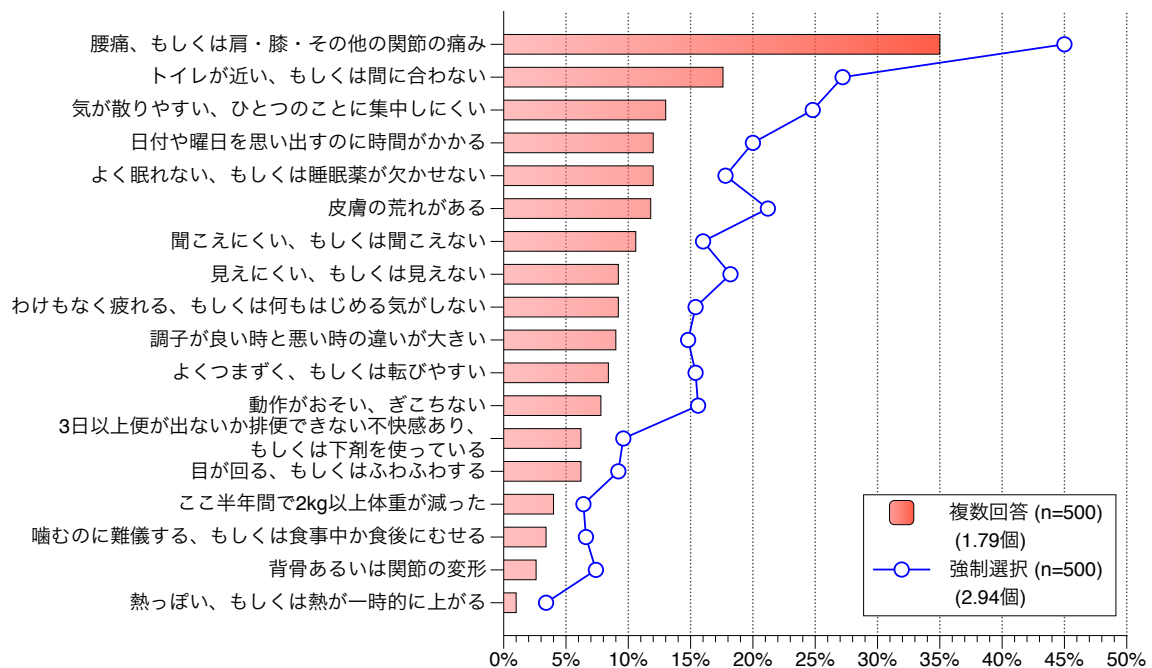
強制選択方式

問1 あなたは、ここ1ヶ月の間に次のようなことがありましたか？(SA)

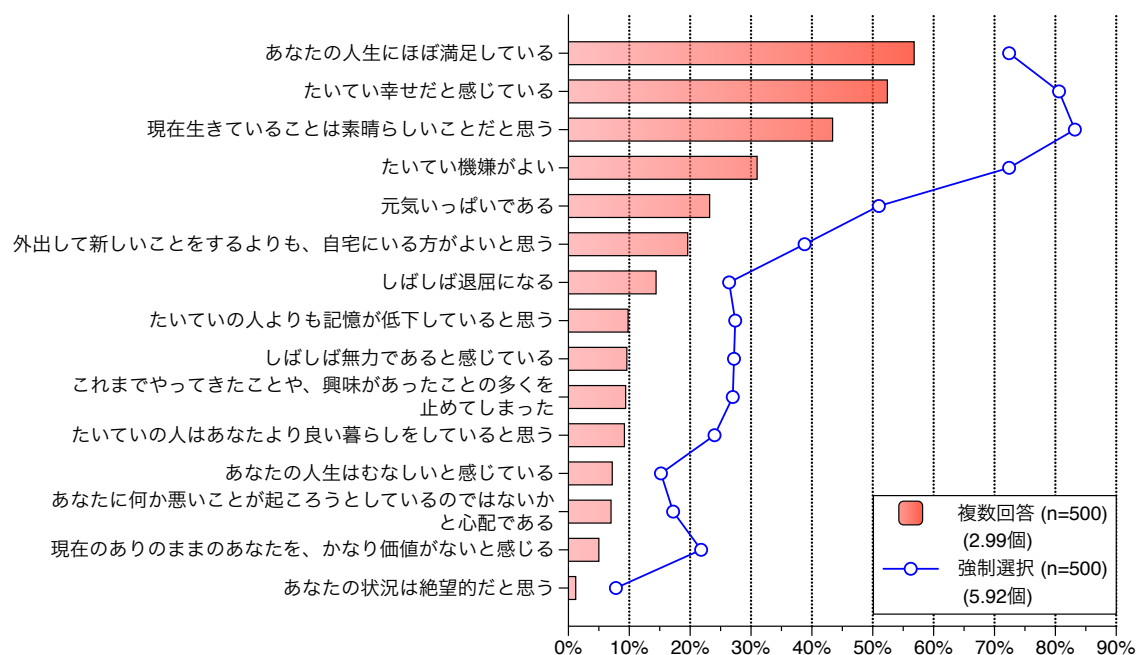
- | | あった | なかった |
|--------------------------------------|-----------------------|-----------------------|
| A 気が散りやすい、ひとつのことに集中しにくい | <input type="radio"/> | <input type="radio"/> |
| B よく眠れない、もしくは睡眠薬が欠かせない | <input type="radio"/> | <input type="radio"/> |
| C 日付や曜日を思い出すのに時間がかかる | <input type="radio"/> | <input type="radio"/> |
| D 調子が良い時と悪い時の違いが大きい | <input type="radio"/> | <input type="radio"/> |
| E ここ半年間で2kg以上体重が減った | <input type="radio"/> | <input type="radio"/> |
| F わけもなく疲れる、もしくは何もはじめる気がしない | <input type="radio"/> | <input type="radio"/> |
| G 目が回る、もしくはふわふわする | <input type="radio"/> | <input type="radio"/> |
| H 熱っぽい、もしくは熱が一時的に上がる | <input type="radio"/> | <input type="radio"/> |
| I よくつまずく、もしくは転びやすい | <input type="radio"/> | <input type="radio"/> |
| J 見えにくい、もしくは見えない | <input type="radio"/> | <input type="radio"/> |
| K 聞こえにくい、もしくは聞こえない | <input type="radio"/> | <input type="radio"/> |
| L 動作がおそい、ぎこちない | <input type="radio"/> | <input type="radio"/> |
| M 腰痛、もしくは肩・膝・その他の関節の痛み | <input type="radio"/> | <input type="radio"/> |
| N 背骨あるいは関節の変形 | <input type="radio"/> | <input type="radio"/> |
| O 皮膚の荒れがある | <input type="radio"/> | <input type="radio"/> |
| P トイレが近い、もしくは間に合わない | <input type="radio"/> | <input type="radio"/> |
| Q 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている | <input type="radio"/> | <input type="radio"/> |
| R 噛むのに難儀する、もしくは食事中か食後にむせる | <input type="radio"/> | <input type="radio"/> |

12

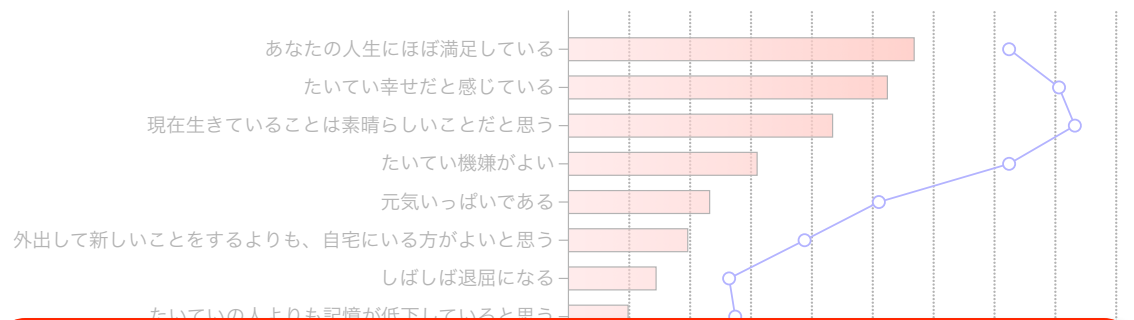
あなたは、ここ1ヶ月の間に次のようなことがありましたか？



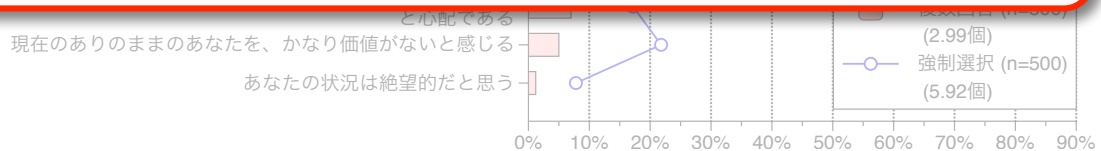
次のうち、あなたがそう思うものをいくつでも選んでください



次のうち、あなたがそう思うものをいくつでも選んでください



**症状等に関しても、複数回答方式では
強制選択方式と比べ、過少回答傾向**



15

なぜ複数回答方式は過少回答となるのか？

・仮説（その1）

複数回答方式では該当が明確な選択肢のみを選ぶため

➡ 強制選択方式を段階評定方式とするとどうなるか？

➡ 該当・非該当が明確であれば過少回答は生じないか？

16

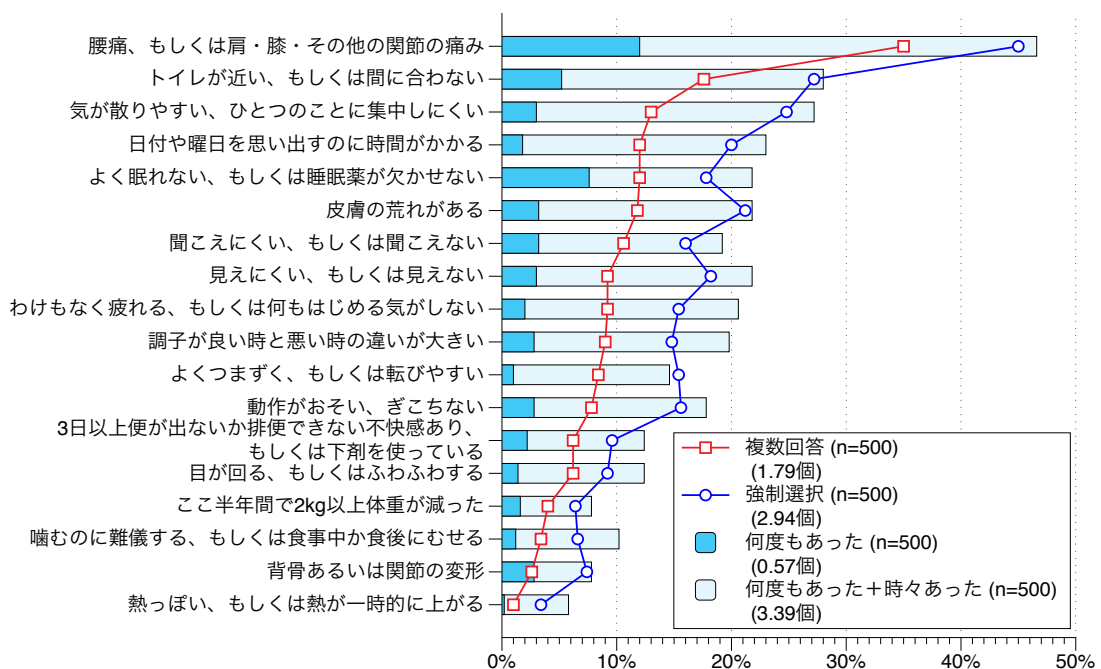
あなたは、ここ1ヶ月の間に次のようなことがありましたか？

段階評定方式

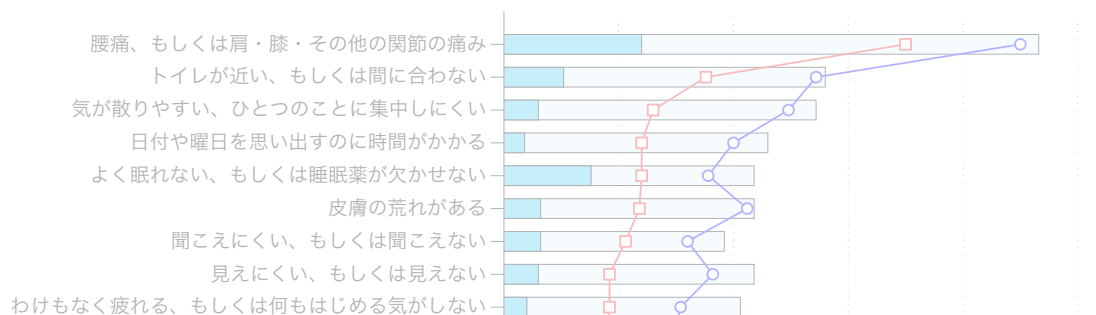
問1 あなたは、ここ1ヶ月の間に次のようなことがありましたか？(SA)

	何度もあった	時々あった	ほとんどなかった	全くなかった
A 気が散りやすい、ひとつのことに集中しにくい	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
B よく眠れない、もしくは睡眠薬が欠かせない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
C 日付や曜日を思い出すのに時間がかかる	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D 調子が良い時と悪い時の違いが大きい	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
E ここ半年間で2kg以上体重が減った	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
F わけもなく疲れる、もしくは何もはじめる気がしない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
G 目が回る、もしくはふわふわする	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
H 熱っぽい、もしくは熱が一時的に上がる	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I よくつまずく、もしくは転びやすい	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J 見えにくい、もしくは見えない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
K 聞こえにくい、もしくは聞こえない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
L 動作がおそい、ぎこちない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
M 腰痛、もしくは肩・膝・その他の関節の痛み	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
N 背骨あるいは関節の変形	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O 皮膚の荒れがある	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
P トイレが近い、もしくは間に合わない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Q 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
R 嘔むのに難儀する、もしくは食事中か食後にむせる	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

あなたは、ここ1ヶ月の間に次のようなことがありましたか？

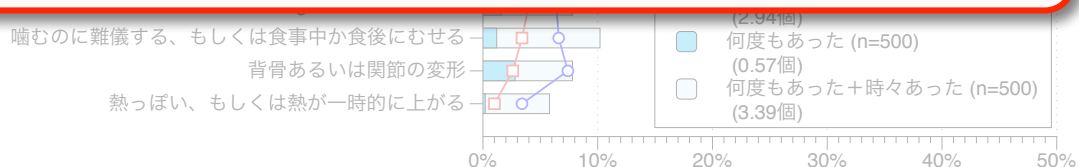


あなたは、ここ1ヶ月の間に次のようなことがありましたか？



「複数回答」 ≠ 「何度もあった」

≠ 「何度もあった」 + 「時々あった」



次のうち、あなたが定期的に使用している薬を選んでください？

複数回答

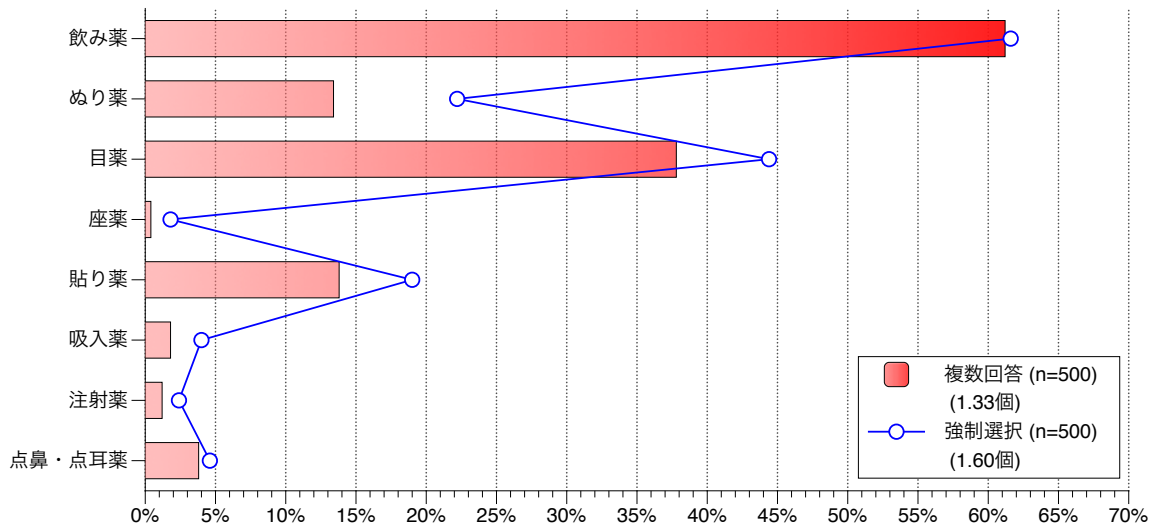
- 飲み薬
- むり薬
- 目薬
- 座薬
- 貼り薬
- 吸入薬
- 注射薬
- 点鼻・点耳薬

強制選択

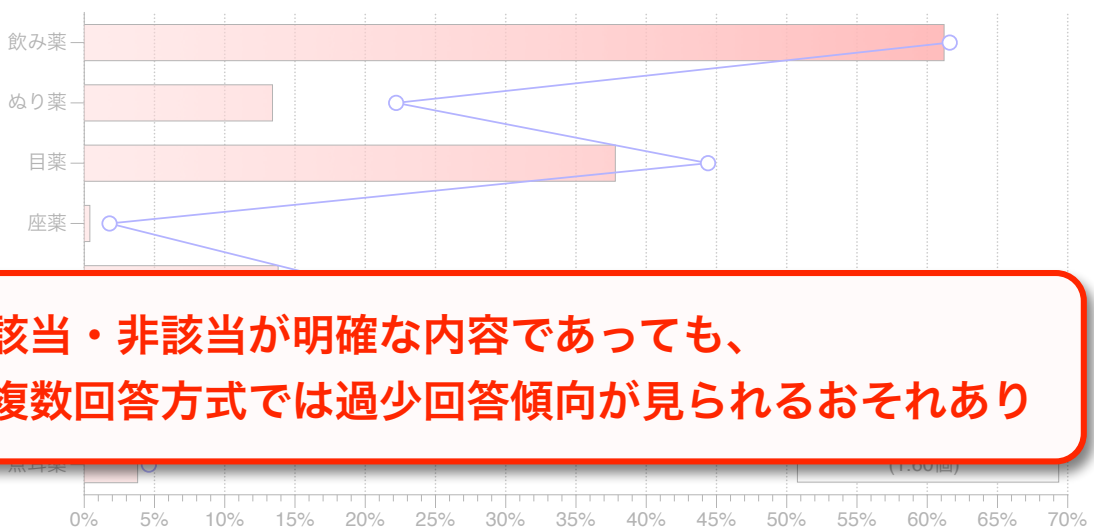
当てはまる 当てはまらない

- 飲み薬 ○ ○
- むり薬 ○ ○
- 目薬 ○ ○
- 座薬 ○ ○
- 貼り薬 ○ ○
- 吸入薬 ○ ○
- 注射薬 ○ ○
- 点鼻・点耳薬 ○ ○

次のうち、あなたが定期的に使用している薬を選んでください？



次のうち、あなたが定期的に使用している薬を選んでください？



該当・非該当が明確な内容であっても、複数回答方式では過少回答傾向が見られるおそれあり

なぜ複数回答方式は過少回答となるのか？

・仮説（その2）

Satisfice（最小限化）傾向のため

➡ 順序効果が生じるか？

➡ 全て該当するような内容でも割合が低くなるか？

23

あなたは、ここ1ヶ月の間に次のようなことがありましたか？

- | | |
|---|---|
| <input type="checkbox"/> 気が散りやすい、ひとつのことに集中しにくい | <input type="checkbox"/> 見えにくい、もしくは見えない |
| <input type="checkbox"/> よく眠れない、もしくは睡眠薬が欠かせない | <input type="checkbox"/> 聞こえにくい、もしくは聞こえない |
| <input type="checkbox"/> 日付や曜日を思い出すのに時間がかかる | <input type="checkbox"/> 動作がおそい、ぎこちない |
| <input type="checkbox"/> 調子が良い時と悪い時の違いが大きい | <input type="checkbox"/> 腰痛、もしくは肩・膝・その他の関節の痛み |
| <input type="checkbox"/> ここ半年間で2kg以上体重が減った | <input type="checkbox"/> 背骨あるいは関節の変形 |
| <input type="checkbox"/> わけもなく疲れる、もしくは何もはじめる気がしない | <input type="checkbox"/> 皮膚の荒れがある |
| <input type="checkbox"/> 目が回る、もしくはふわふわする | <input type="checkbox"/> トイレが近い、もしくは間に合わない |
| <input type="checkbox"/> 熱っぽい、もしくは熱が一時的に上がる | <input type="checkbox"/> 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている |
| <input type="checkbox"/> よくつまずく、もしくは転びやすい | <input type="checkbox"/> 噛むのに難儀する、もしくは食事中か食後にむせる |
| <input type="checkbox"/> 見えにくい、もしくは見えない | <input type="checkbox"/> ここ半年間で2kg以上体重が減った |
| <input type="checkbox"/> 聞こえにくい、もしくは聞こえない | <input type="checkbox"/> わけもなく疲れる、もしくは何もはじめる気がしない |
| <input type="checkbox"/> 動作がおそい、ぎこちない | <input type="checkbox"/> 目が回る、もしくはふわふわする |
| <input type="checkbox"/> 腰痛、もしくは肩・膝・その他の関節の痛み | <input type="checkbox"/> 熱っぽい、もしくは熱が一時的に上がる |
| <input type="checkbox"/> 背骨あるいは関節の変形 | <input type="checkbox"/> よくつまずく、もしくは転びやすい |
| <input type="checkbox"/> 皮膚の荒れがある | <input type="checkbox"/> 気が散りやすい、ひとつのことに集中しにくい |
| <input type="checkbox"/> トイレが近い、もしくは間に合わない | <input type="checkbox"/> よく眠れない、もしくは睡眠薬が欠かせない |
| <input type="checkbox"/> 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている | <input type="checkbox"/> 日付や曜日を思い出すのに時間がかかる |
| <input type="checkbox"/> 噛むのに難儀する、もしくは食事中か食後にむせる | <input type="checkbox"/> 調子が良い時と悪い時の違いが大きい |

24

あなたは、ここ1ヶ月の間に次のようなことがありましたか？

気が散りやすい、ひとつのことに集中しにくい
 よく眠れない、もしくは寝不足
 日付や曜日を思い出すのが難しい
 調子が良い時と悪い時の違いが大きい

**0.448個
(.024)**

- ここ半年間で2kg以上体重が減った
- わけもなく疲れる、もしくは何もはじめる気がしない
- 目が回る、もしくはふわふわする
- 熱っぽい、もしくは熱が一時的に上がる
- よくつまずく、もしくは転びやすい
- 見えにくい、もしくは見えない
- 聞こえにくい、もしくは聞こえない
- 動作がおそい、ぎこちない
- 腰痛、もしくは肩・膝・その他の関節の痛み
- 背骨あるいは関節の変形
- 皮膚の荒れがある
- トイレが近い、もしくは間に合わない
- 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
- 嘔むのに難儀する、もしくは食事中か食後にむせる

- 見えにくい、もしくは見えない
- 聞こえにくい、もしくは聞こえない
- 動作がおそい、ぎこちない
- 腰痛、もしくは肩・膝・その他の関節の痛み
- 背骨あるいは関節の変形
- 皮膚の荒れがある
- トイレが近い、もしくは間に合わない
- 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
- 嘔むのに難儀する、もしくは食事中か食後にむせる
- ここ半年間で2kg以上体重が減った
- わけもなく疲れる、もしくは何もはじめる気がしない
- 目が回る、もしくはふわふわする
- 熱っぽい、もしくは熱が一時的に上がる
- よくつまずく、もしくは転びやすい

気が散りやすい、ひとつのことに集中しにくい
 よく眠れない、もしくは寝不足
 日付や曜日を思い出すのが難しい
 調子が良い時と悪い時の違いが大きい

**0.274個
(.021)**

あなたは、ここ1ヶ月の間に次のようなことがありましたか？

気が散りやすい、ひとつのことに集中しにくい
 よく眠れない、もしくは寝不足
 日付や曜日を思い出すのが難しい
 調子が良い時と悪い時の違いが大きい

**0.448個
(.024)**

- ここ半年間で2kg以上体重が減った
- わけもなく疲れる、もしくは何もはじめる気がしない
- 目が回る、もしくはふわふわする
- 熱っぽい、もしくは熱が一時的に上がる
- よくつまずく、もしくは転びやすい

見えにくい、もしくは見えない
 聞こえにくい、もしくは聞こえない
 動作がおそい、ぎこちない
 腰痛、もしくは肩・膝・その他の関節の痛み
 背骨あるいは関節の変形
 皮膚の荒れがある
 トイレが近い、もしくは間に合わない
 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
 嘔むのに難儀する、もしくは食事中か食後にむせる

**1.017個
(.026)**

見えにくい、もしくは見えない
 聞こえにくい、もしくは聞こえない
 動作がおそい、ぎこちない
 腰痛、もしくは肩・膝・その他の関節の痛み
 背骨あるいは関節の変形
 皮膚の荒れがある
 トイレが近い、もしくは間に合わない
 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
 嘔むのに難儀する、もしくは食事中か食後にむせる

**1.219個
(.026)**

- ここ半年間で2kg以上体重が減った
- わけもなく疲れる、もしくは何もはじめる気がしない
- 目が回る、もしくはふわふわする
- 熱っぽい、もしくは熱が一時的に上がる
- よくつまずく、もしくは転びやすい

気が散りやすい、ひとつのことに集中しにくい
 よく眠れない、もしくは寝不足
 日付や曜日を思い出すのが難しい
 調子が良い時と悪い時の違いが大きい

**0.274個
(.021)**

次のうち、あなたが定期的に使用している薬を選んでください？

- | | |
|---------------------------------|---------------------------------|
| <input type="checkbox"/> 飲み薬 | <input type="checkbox"/> 点鼻・点耳薬 |
| <input type="checkbox"/> めり薬 | <input type="checkbox"/> 注射薬 |
| <input type="checkbox"/> 目薬 | <input type="checkbox"/> 吸入薬 |
| <input type="checkbox"/> 座薬 | <input type="checkbox"/> 貼り薬 |
| <input type="checkbox"/> 貼り薬 | <input type="checkbox"/> 座薬 |
| <input type="checkbox"/> 吸入薬 | <input type="checkbox"/> 目薬 |
| <input type="checkbox"/> 注射薬 | <input type="checkbox"/> めり薬 |
| <input type="checkbox"/> 点鼻・点耳薬 | <input type="checkbox"/> 飲み薬 |

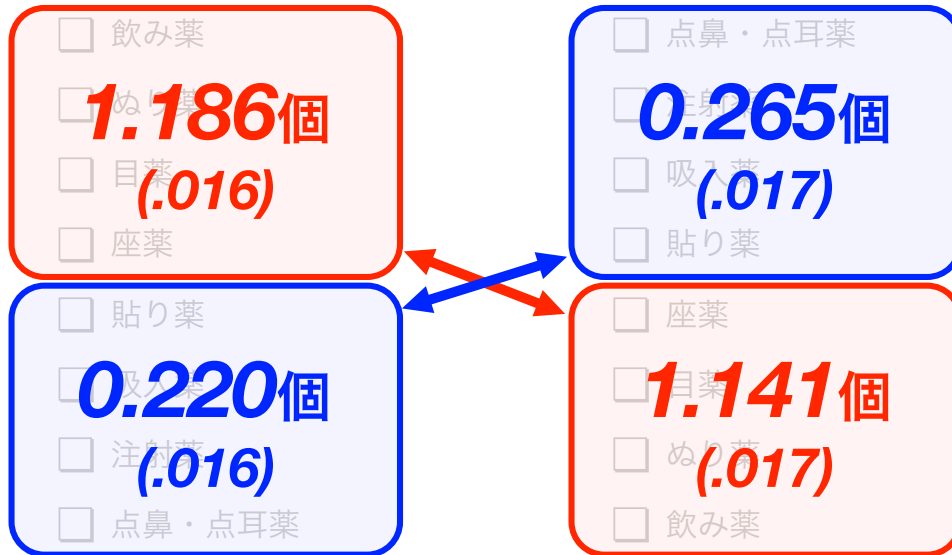
27

次のうち、あなたが定期的に使用している薬を選んでください？

- | | |
|---------------------------------|---------------------------------|
| <input type="checkbox"/> 飲み薬 | <input type="checkbox"/> 点鼻・点耳薬 |
| <input type="checkbox"/> めり薬 | <input type="checkbox"/> 注射薬 |
| <input type="checkbox"/> 目薬 | <input type="checkbox"/> 吸入薬 |
| <input type="checkbox"/> 座薬 | <input type="checkbox"/> 貼り薬 |
| <input type="checkbox"/> 貼り薬 | <input type="checkbox"/> 座薬 |
| <input type="checkbox"/> 吸入薬 | <input type="checkbox"/> 目薬 |
| <input type="checkbox"/> 注射薬 | <input type="checkbox"/> めり薬 |
| <input type="checkbox"/> 点鼻・点耳薬 | <input type="checkbox"/> 飲み薬 |
- 1.186個 (.016)**
- 1.141個 (.017)**

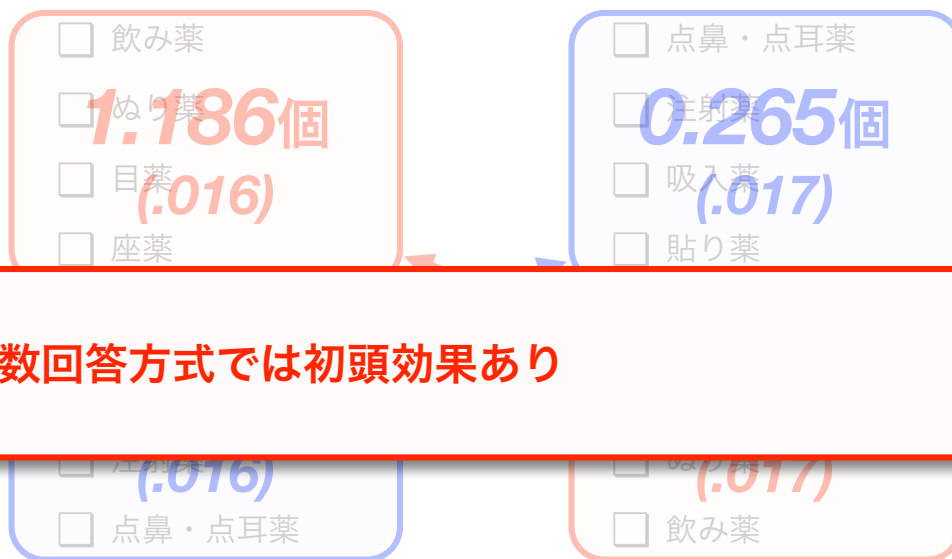
28

次のうち、あなたが定期的に使用している薬を選んでください？



29

次のうち、あなたが定期的に使用している薬を選んでください？



30

郵送調査の概要

- 調査票 : 2種類用意
- 調査対象 : 20歳以上79歳以下立川市民
 - 住民基本台帳から無作為に抽出
 - 各調査票に無作為に割当
- 調査期間 : 2016年1月22日 - 2月29日
- 標本サイズ : 各調査票 2,000 (合計 4,000)
- 回収率 : 75.8%

※ 調査は立川市との連携協定の下、朴堯星助教（統計数理研究所）と実施

31

立川市の魅力

問6 あなたは、立川市の魅力はどのような点にあると思いますか？ (○はいくつでも)

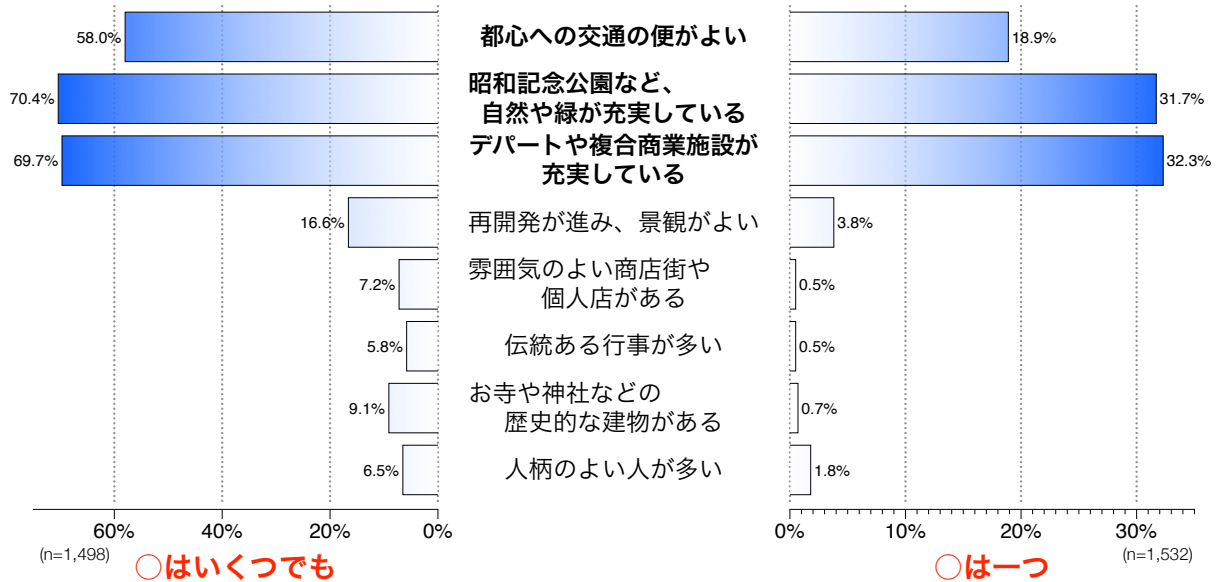
- 1 都心への交通の便がよい
- 2 昭和記念公園など、自然や緑が充実している
- 3 デパートや複合商業施設が充実している
- 4 再開発が進み、景観がよい
- 5 雰囲気の良い商店街や個人店がある
- 6 伝統ある行事が多い
- 7 お寺や神社などの歴史的な建物がある
- 8 人柄のよい人が多い
- 9 その他（具体的に： _____）
- 10 特にない

問6 あなたは、立川市の魅力はどのような点にあると思いますか？ (○は一つ)

- 1 都心への交通の便がよい
- 2 昭和記念公園など、自然や緑が充実している
- 3 デパートや複合商業施設が充実している
- 4 再開発が進み、景観がよい
- 5 雰囲気の良い商店街や個人店がある
- 6 伝統ある行事が多い
- 7 お寺や神社などの歴史的な建物がある
- 8 人柄のよい人が多い
- 9 その他（具体的に： _____）
- 10 特にない

32

立川市の魅力



立川市の重点的取り組み

複数回答方式

問26 あなたは、立川市は今後どのような点に重点的に取り組むべきだと思いますか？ (〇はいくつでも)

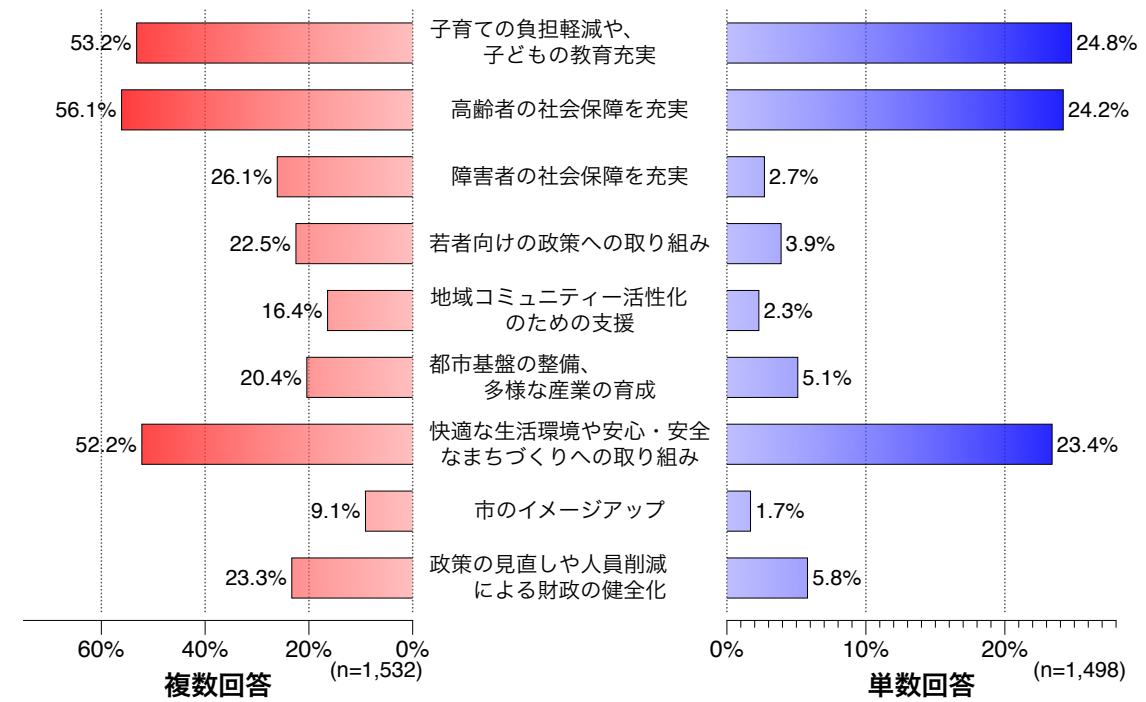
- 1 子育ての負担軽減や、子どもの教育充実
- 2 高齢者の社会保障を充実
- 3 障害者の社会保障を充実
- 4 若者向けの政策への取り組み
- 5 地域コミュニティ活性化のための支援
- 6 都市基盤の整備、多様な産業の育成
- 7 快適な生活環境や安心・安全なまちづくりへの取り組み
- 8 市のイメージアップ
- 9 政策の見直しや人員削減による財政の健全化
- 10 その他 (具体的に:)

単数回答方式

問26 あなたは、立川市は今後どのような点に重点的に取り組むべきだと思いますか？ (〇は一つ)

- 1 子育ての負担軽減や、子どもの教育充実
- 2 高齢者の社会保障を充実
- 3 障害者の社会保障を充実
- 4 若者向けの政策への取り組み
- 5 地域コミュニティ活性化のための支援
- 6 都市基盤の整備、多様な産業の育成
- 7 快適な生活環境や安心・安全なまちづくりへの取り組み
- 8 市のイメージアップ
- 9 政策の見直しや人員削減による財政の健全化
- 10 その他 (具体的に:)

立川市の重点的取り組み

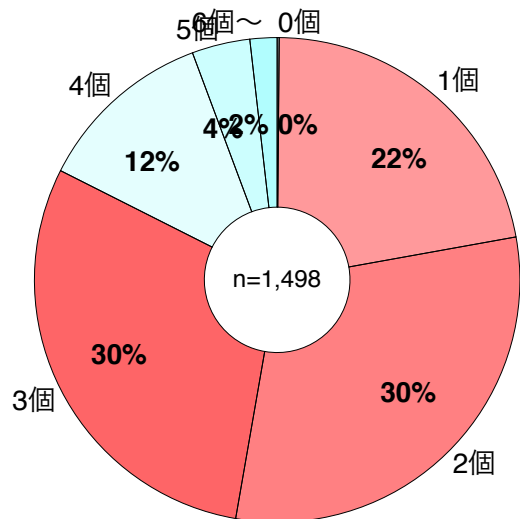


立川市の魅力

複数回答方式

問6 あなたは、立川市の魅力はどのような点にあると思いますか？ (〇はいくつでも)

- 1 都心への交通の便がよい
- 2 昭和記念公園など、自然や緑が充実している
- 3 デパートや複合商業施設が充実している
- 4 再開発が進み、景観がよい
- 5 雰囲気の良い商店街や個人店がある
- 6 伝統ある行事が多い
- 7 お寺や神社などの歴史的な建物がある
- 8 人柄のよい人が多い
- 9 その他 (具体的に:)
- 10 特にない

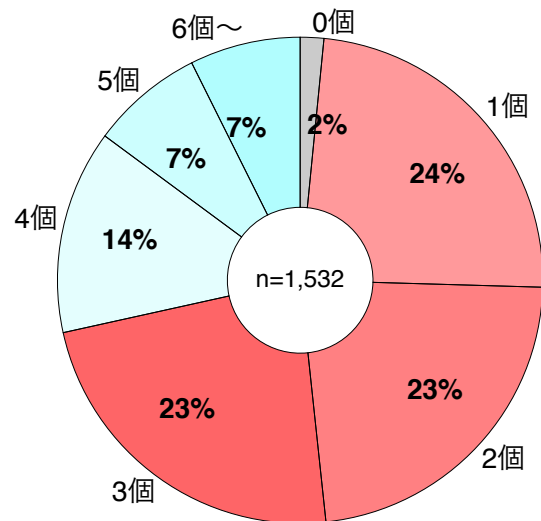


立川市の重点的取り組み

複数回答方式

問26 あなたは、立川市は今後どのような点に重点的に取り組むべきだと思いますか？（〇はいくつでも）

- 1 子育ての負担軽減や、子どもの教育充実
- 2 高齢者の社会保障を充実
- 3 障害者の社会保障を充実
- 4 若者向けの政策への取り組み
- 5 地域コミュニティ活性化のための支援
- 6 都市基盤の整備、多様な産業の育成
- 7 快適な生活環境や安心・安全なまちづくりへの取り組み
- 8 市のイメージアップ
- 9 政策の見直しや人員削減による財政の健全化
- 10 その他（具体的に： ）



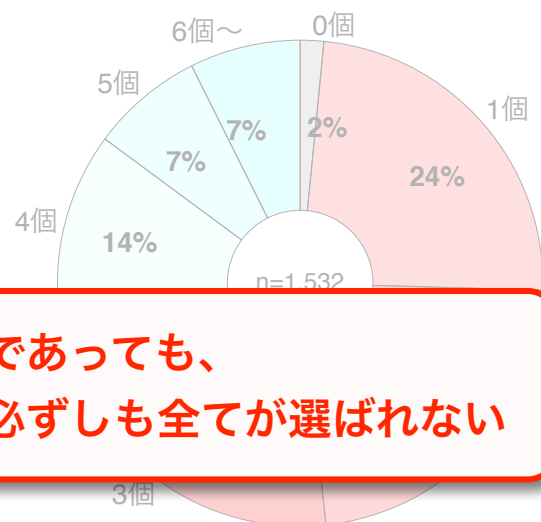
37

立川市の重点的取り組み

複数回答方式

問26 あなたは、立川市は今後どのような点に重点的に取り組むべきだと思いますか？（〇はいくつでも）

- 1 子育ての負担軽減や、子どもの教育充実
- 2 高齢者の社会保障を充実
- 3 障害者の社会保障を充実
- 4 若者向けの政策への取り組み



**多くが該当するような内容であっても、
複数回答方式では必ずしも全てが選ばれない**

38

過少回答を抑制するにはどうすればよいか

・方法（その1）

リストを短くする（選択肢数を減らす）

➡ 短いリストでは割合（選択数）が増えるか？

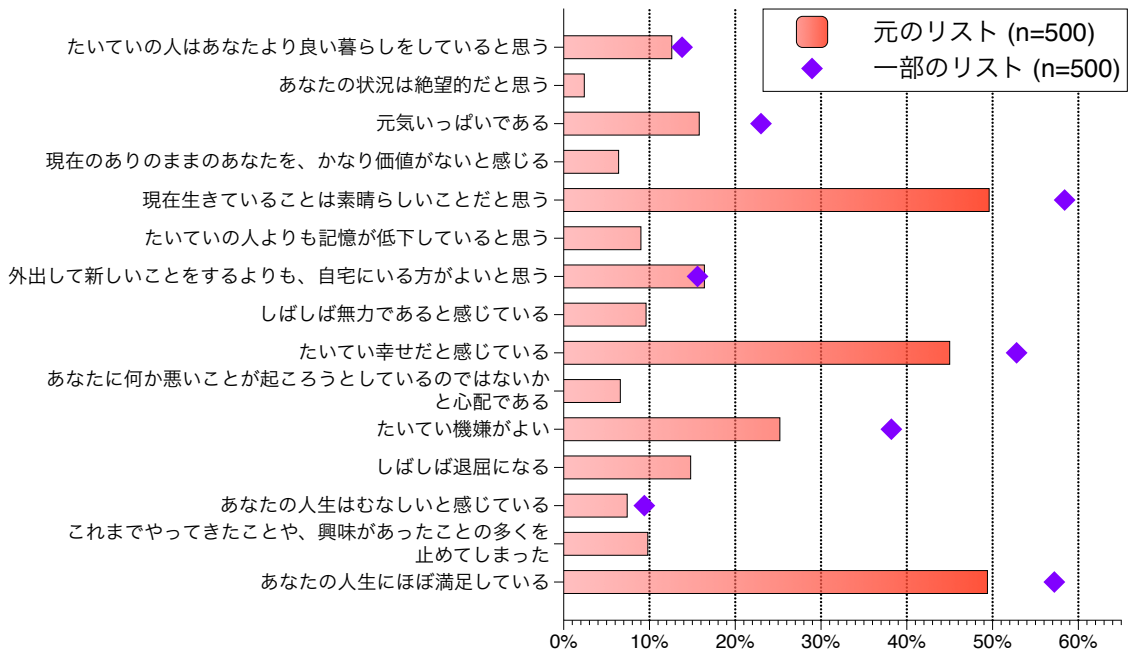
39

次のうち、あなたがそう思うものをいくつでも選んでください。

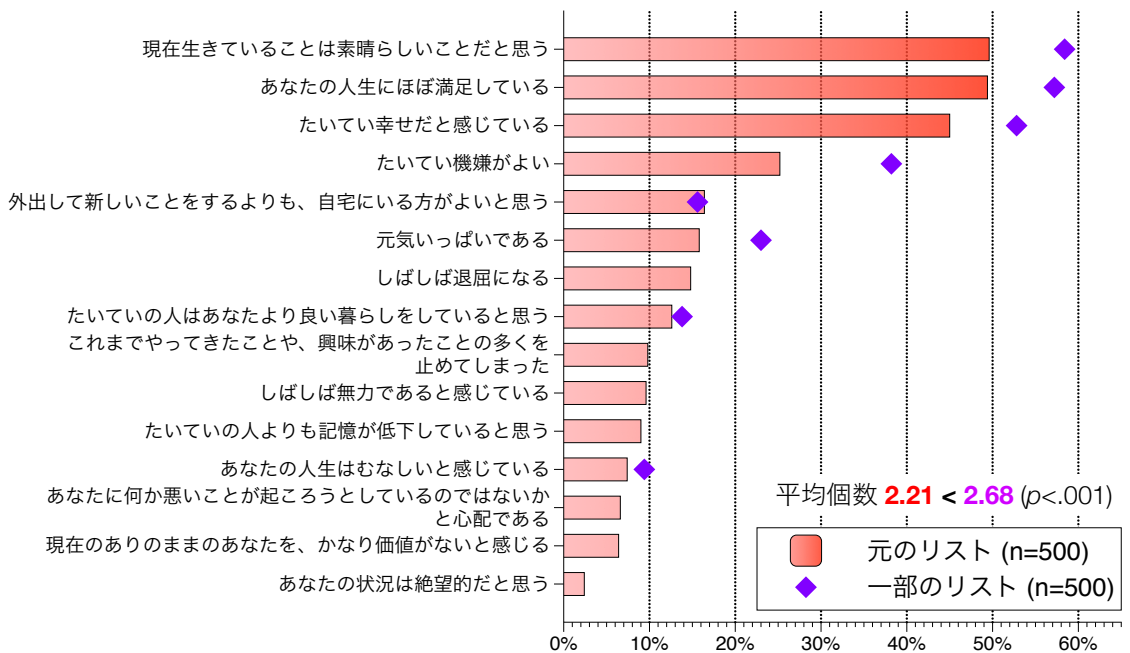
- | | |
|---|---|
| <input type="checkbox"/> たいていの人はあなたより良い暮らしをしていると思う | <input type="checkbox"/> たいていの人はあなたより良い暮らしをしていると思う |
| <input type="checkbox"/> あなたの状況は絶望的だと思う | <input type="checkbox"/> 元気いっぱいである |
| <input type="checkbox"/> 元気いっぱいである | <input type="checkbox"/> 現在生きていることは素晴らしいことだと思う |
| <input type="checkbox"/> 現在のありのままのあなたを、かなり価値がないと感じる | <input type="checkbox"/> 外出して新しいことをするよりも、自宅にいる方がよいと思う |
| <input type="checkbox"/> 現在生きていることは素晴らしいことだと思う | <input type="checkbox"/> たいてい幸せだと感じている |
| <input type="checkbox"/> たいていの人よりも記憶が低下していると思う | <input type="checkbox"/> たいてい機嫌がよい |
| <input type="checkbox"/> 外出して新しいことをするよりも、自宅にいる方がよいと思う | <input type="checkbox"/> あなたの人生はむなしいと感じている |
| <input type="checkbox"/> しばしば無力であると感じている | <input type="checkbox"/> あなたの人生にほぼ満足している |
| <input type="checkbox"/> たいてい幸せだと感じている | |
| <input type="checkbox"/> あなたに何か悪いことが起ころうとしているのではないかと心配である | |
| <input type="checkbox"/> たいてい機嫌がよい | |
| <input type="checkbox"/> しばしば退屈になる | |
| <input type="checkbox"/> あなたの人生はむなしいと感じている | |
| <input type="checkbox"/> これまでやってきたことや、興味があったことの多くを止めてしまった | |
| <input type="checkbox"/> あなたの人生にほぼ満足している | |

40

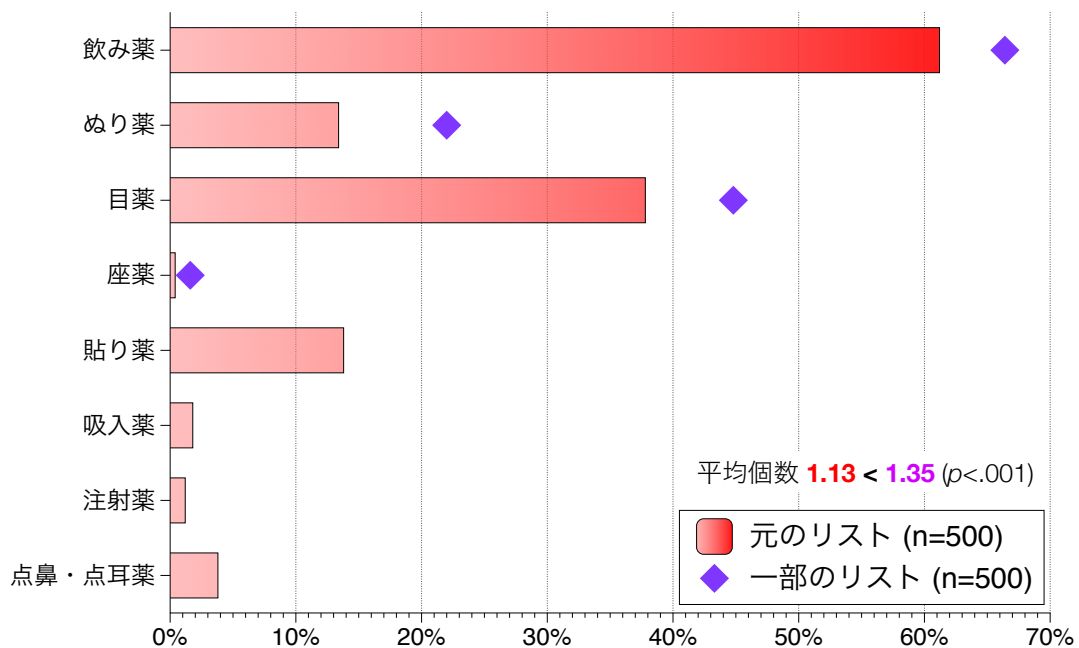
次のうち、あなたがそう思うものをいくつでも選んでください。



次のうち、あなたがそう思うものをいくつでも選んでください。



次のうち、あなたが定期的に使用している薬を選んでください？



43

過少回答を抑制するにはどうすればよいか

・方法 (その2)

リストを分割する

➡ 分割したリストでは割合 (選択数) が増えるか？

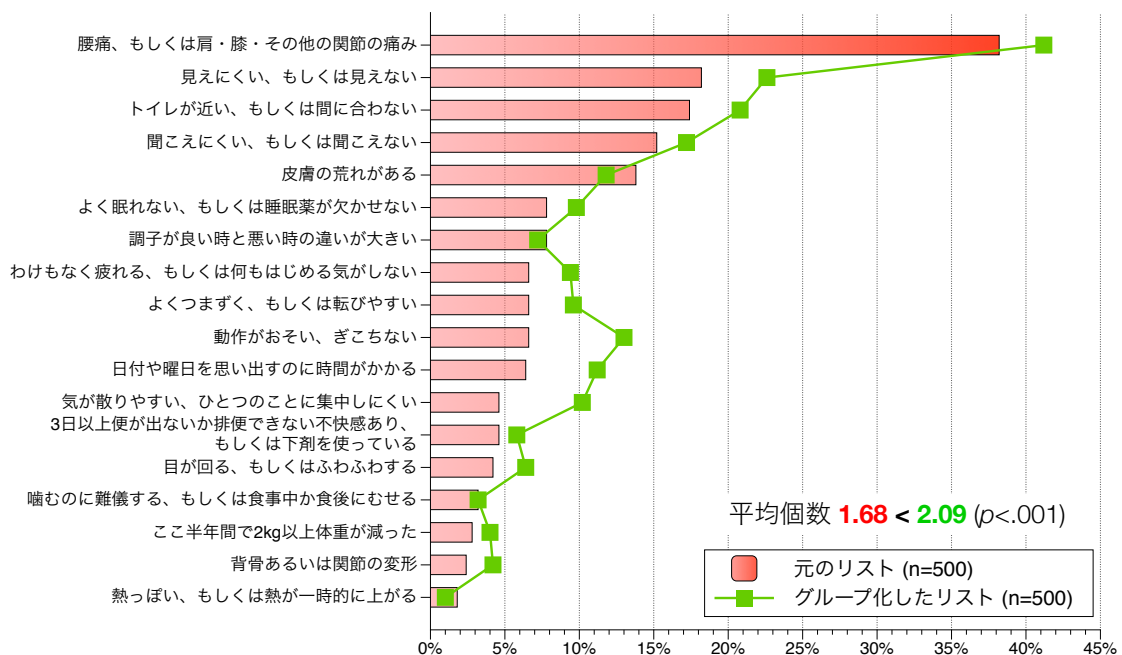
44

あなたは、ここ1ヶ月の間に次のようなことがありましたか？

- 見えにくい、もしくは見えない
 - 聞こえにくい、もしくは聞こえない
 - 動作がおそい、ぎこちない
 - 腰痛、もしくは肩・膝・その他の関節の痛み
 - 背骨あるいは関節の変形
 - 皮膚の荒れがある
 - トイレが近い、もしくは間に合わない
 - 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
 - 嘔むのに難儀する、もしくは食事中か食後にむせる
 - ここ半年間で2kg以上体重が減った
 - わけもなく疲れる、もしくは何もはじめる気がしない
 - 目が回る、もしくはふわふわする
 - 熱っぽい、もしくは熱が一時的に上がる
 - よくつまずく、もしくは転びやすい
 - 気が散りやすい、ひとつのことに集中しにくい
 - よく眠れない、もしくは睡眠薬が欠かせない
 - 日付や曜日を思い出すのに時間がかかる
 - 調子が良い時と悪い時の違いが大きい
- 加齢に伴い増加する身体の状態**
 - 見えにくい、もしくは見えない
 - 聞こえにくい、もしくは聞こえない
 - 動作がおそい、ぎこちない
 - 腰痛、もしくは肩・膝・その他の関節の痛み
 - 背骨あるいは関節の変形
 - 皮膚の荒れがある
 - トイレが近い、もしくは間に合わない
 - 3日以上便が出ないか排便できない不快感あり、もしくは下剤を使っている
 - 嘔むのに難儀する、もしくは食事中か食後にむせる
- 高齢者の疾患罹患時によく見られる身体症状**
 - ここ半年間で2kg以上体重が減った
 - わけもなく疲れる、もしくは何もはじめる気がしない
 - 目が回る、もしくはふわふわする
 - 熱っぽい、もしくは熱が一時的に上がる
 - よくつまずく、もしくは転びやすい
- 高齢者の疾患罹患時によく見られる精神症状**
 - 気が散りやすい、ひとつのことに集中しにくい
 - よく眠れない、もしくは睡眠薬が欠かせない
 - 日付や曜日を思い出すのに時間がかかる
 - 調子が良い時と悪い時の違いが大きい

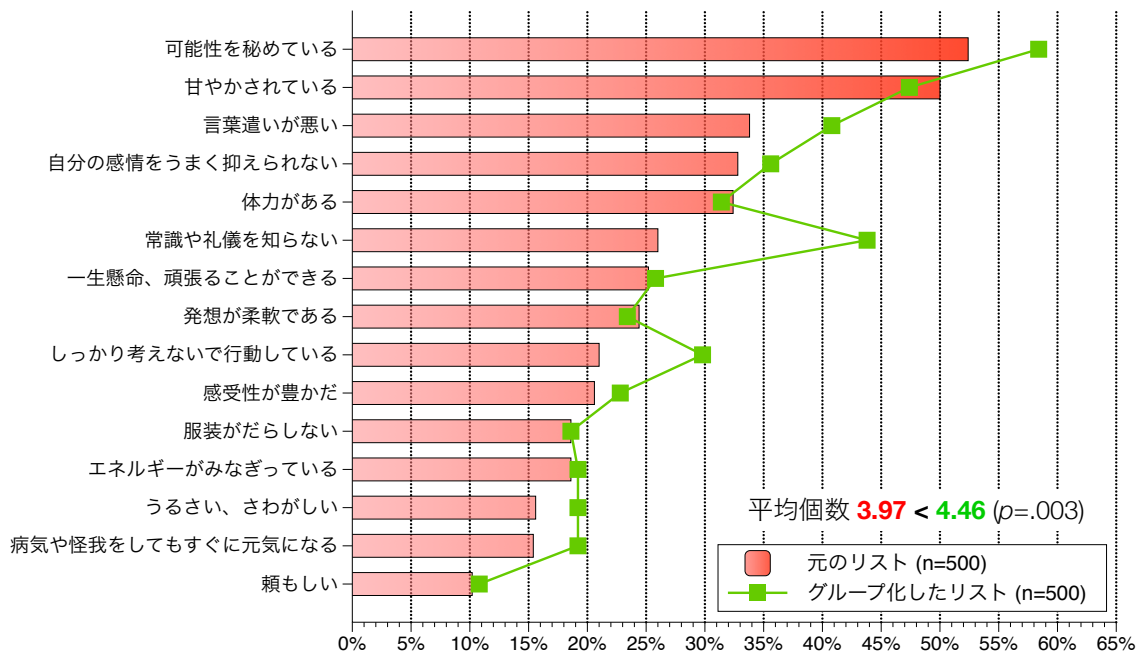
45

あなたは、ここ1ヶ月の間に次のようなことがありましたか？



46

あなたは、今の中学生・高校生を見て次のように
 思いますか？



47

まとめ

- 複数回答方式の過少回答傾向は、内容によって程度の差はあれ、一般に見られる傾向
- 過少回答傾向の原因は、項目の曖昧さだけではない
- 複数回答方式では、選択数を回答者自ら制限か
- 複数回答方式では、選択枝数を減らしてリストを短くしたり、リストを分割することで過少回答傾向を抑制できる可能性

48

SNA統計における時系列データ作成上の課題

ーベンチマーキング及び季節調整を中心にー

平成29年2月2日

長谷川 秀司

目次

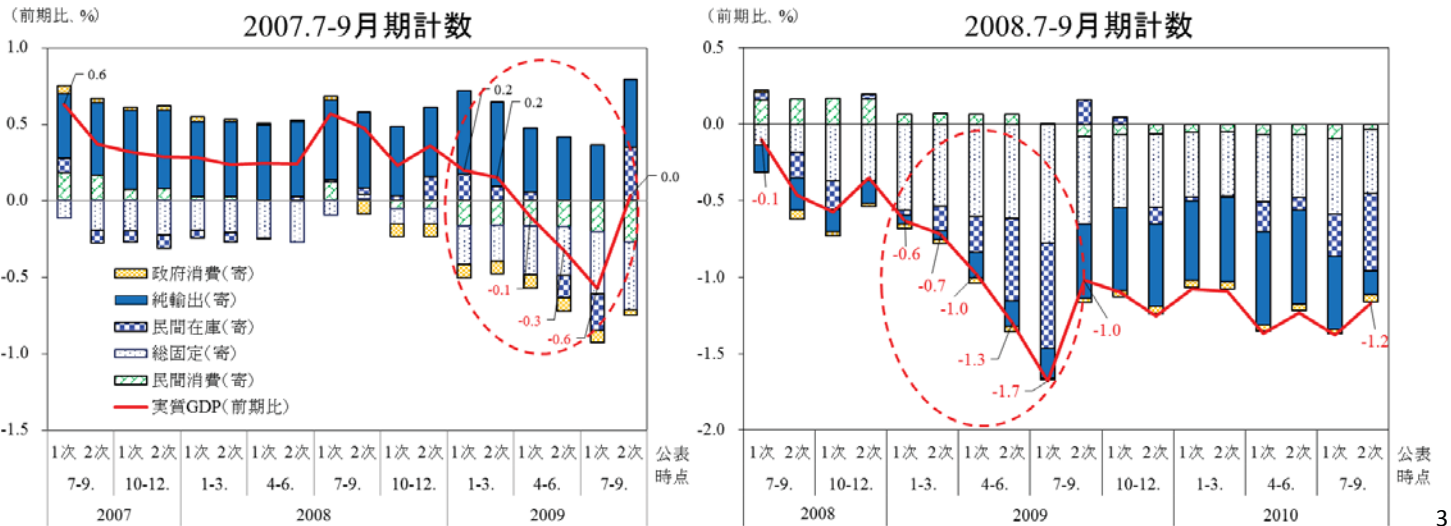
	page
1. 近年のSNA統計を巡る指摘について	3
2. QEの推計と改定	5
3. 予測	7
4. ベンチマーキング	8
5. 季節調整	11

1. 近年のSNA統計を巡る指摘について(1)

① リーマンショック前後のQE(GDP成長率)の連続的な一定方向への改定

2007年7-9月期及び2008年7-9月期の実質GDP成長率(季節調整系列)の改定状況

(リーマンショック(2008年9月)を契機とした世界的景気後退期及び回復期)



3

1. 近年のSNA統計を巡る指摘について(2)

② 1次速報から2次速報への改定

➤2015年7-9月期の成長率の符号が逆転

- 1次 ▲0.2% → 2次 0.3% (改定差0.4%pt)
- 2002年4-6月期以降で符号が逆転したのは4回目 (過去58四半期中)

➤改定幅の絶対値平均は0.17%pt程度 (米0.13%pt程度、英0.03%pt程度)

③ 季節調整における「うるう年」要因の調整

- 2016年1-3月期は季節調整モデルの選定作業で妥当性を検証、検出されず
- 民間エコノミストは、うるう年要因を原系列で調整

4

2. QEの推計と改定(1)

① 改定の考え方

- 利用者にはできるだけタイムリーかつ精度の高いデータを提供するために、改定が行われる
- 基礎データの制約と回答者の負担、利用者の要求
 - ⇒ 【公表の速報性】 vs 【信頼性、精度、包括性】
 - ⇒ より多くのより良い基礎データが得られれば改定

(IMF)

- 将来改定が行われる可能性を踏まえ、直近の公表値を必要以上に重視しないよう注意喚起
- 利用者は、経済動向を慎重に評価することを直近の四半期のみではなく複数の四半期にわたるトレンドを考慮すべき
- 年率換算は、QNAデータの不規則性と不確実性が増幅される

5

2. QEの推計と改定(2)

② 改定要因となる統計処理

データの取り込み、ベンチマーキング、季節調整、概念・推計方法の変更 etc.

③ 改定に関する留意点

- 頻繁に改定される系列は、ほとんど改定されない、または全く改定されない系列と比較して必ずしも正確性に欠けている訳ではない
- 改定がないということは、不十分な一次推計値を改善するために利用可能なより良い情報がないということの意味している可能性

6

3. 予測

- 1次QE段階では、いくつかの系列では、1か月分ないし2か月分のデータのみ利用可能
- 全くデータがない場合も。例えば、1次QEの企業設備投資の推計では、需要側を供給側のトレンド・サイクル成分で外挿

7

4. ベンチマーキング(1)

ベンチマーキング:

ある変数に関する高頻度データ系列(例:四半期データ)と低頻度データ系列(例:年次データ)を一貫した時系列に融合

○プロ・ラータ分割(段差問題step problemの発生)

- 指標の四半期分布に比例的に年次データを按分
- 年をまたがる非連続性(四半期成長率の変更が第1四半期に押し込められる)

8

4. ベンチマーキング(2)

○比例デントン法

→ 年次ベンチマークが与えられる制約条件の下で、隣接四半期に対する相対的な調整の差を最小化

9

4. ベンチマーキング(3)

$$\min_{(X_1, \dots, X_{4\beta}, \dots, X_T)} \sum_{t=2}^T \left[\frac{X_t}{I_t} - \frac{X_{t-1}}{I_{t-1}} \right]^2$$
$$t \in \{1, \dots, (4\beta), \dots, T\}$$

フロー系列においては、制約条件は以下のとおり

$$\sum_{t=2}^T X_t = A_y, \quad y \in \{1, \dots, \beta\}$$

t : 時期 (例 : $t=4y-3$ は y 年第1四半期、 $t=4y$ は y 年第4四半期)

X_t : 四半期 t の求めるQNA推計値

I_t : 四半期 t の指標の水準

A_y : y 年の年次データ

β : 年次ベンチマークが利用可能な最後の年

T : 四半期基礎データが利用可能な最終四半期 10

5. 季節調整(1)

①精度の高い年次推計の四半期データ

- 毎年末、前年年次推計値の10-12月期まで原系列データ(1994年1-3月~)を用いて、AIC最小化によるARIMAモデルの選択(自己回帰や移動平均等の次数、異常値の設定)

11

5. 季節調整(2)

②同時調整

- 一般的に、理論的観点からは、外れ値と原データの改定の影響を除外すれば同時調整(concurrent adjustment)が望ましい。
 - ← 新しいデータは季節パターンの変化についての新しい情報を提供
- 同時調整の潜在利得を左右する要因
 - ・季節成分の安定性
 - ・不規則成分の大きさ
 - ・原データに加えられる改定の大きさ

12

5. 季節調整(3)

③大きな経済変動が生じた場合の対応

➤ 季節調整を通じた過去の計数の改定を抑制する
方策の模索

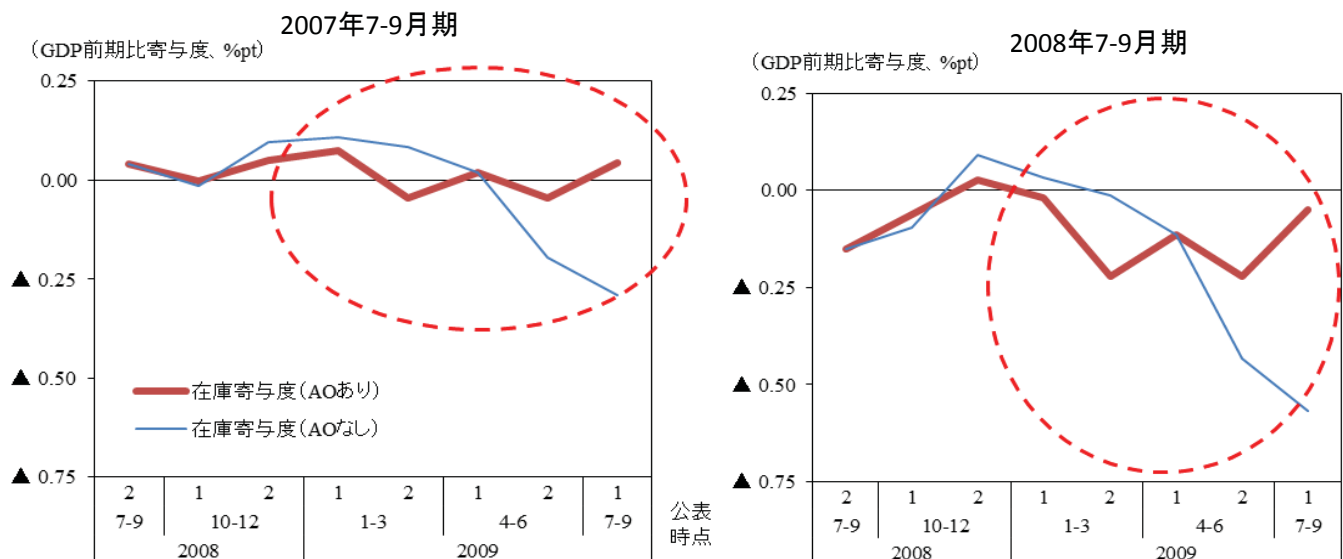
→ リアルタイムにおける異常値処理(AO処理)

(試行)ARIMAモデルにより推計された当期の予定原系列
と実際原系列の乖離の程度で判断

13

5. 季節調整(4)

予定原系列を用いた判定により検出された期における異常値処理設定シミュレーション(民間在庫変動)



14

ご清聴ありがとうございました

(参考)

- 権田[2015](季刊国民経済計算No.158、平成27年11月)
- “Quarterly National Accounts Manual Concepts, Data, Sources, and Compilation” (IMF)2001

内閣支持率と株価の因果関係

川崎能典 (統計数理研究所 モデリング研究系)

概要

内閣支持率と株価の因果関係は、学術的であるに問わずさまざまなレベルで論じられることの多い論題である。本報告の目的は、定常時系列解析における標準的な因果分析法のひとつである、グレンジャーの因果性検定の枠組みに乗せて、議論の出発点程度の認識を確立しておくことにある。内閣支持率は1978年3月から2015年11月までの月次データを取り、それに合わせる形で株価収益率(日経平均月次終値ベース)を使用する。グレンジャーの因果性検定の枠組みに基づけば、内閣支持率から株価収益率への因果性は見られない。一方、株価収益率から内閣支持率への因果性は統計的には有意に検出されるが、支持率の変動の説明という観点からはごく限られた影響にとどまると言うべきである。また、内閣発足直後は概して支持率が高いという「ハネムーン効果」を考慮しても結論は変わらない。

1. データの前処理

まず、分析に用いたデータについて説明する。内閣支持率は、1978年3月から2015年11月までの月次データ(452ヶ月分)を取り、それに合わせる形で株価収益率として日経平均月次終値ベースを使用する。内閣支持率については、読売新聞の調査による数値を使用している。

1.1 欠測値の補間

ここで問題となるのは、月例での調査期間に内閣の交代時期が挟まると、当該月の内閣支持率は欠測データになっていることである。このことは長期の時系列分析を行う際に不都合であるので、今回の分析では便宜的に2階のランダムウォークモデルによる平滑化を行い、欠測時点に関しては平滑化値を代入して時系列の連続性を確保した。すなわち、内閣支持率データを y_t ($t = 1, \dots, T$) とするとき、その背後にトレンド変数 μ_t を想定し、

$$\begin{aligned}\mu_t &= 2\mu_{t-1} - \mu_{t-2} + v_t, & v_t &\sim N(0, \sigma_v^2) \\ y_t &= \mu_t + w_t, & w_t &\sim N(0, \sigma_w^2)\end{aligned}$$

という階層構造モデルをあてはめる。尤度計算はいわゆる予測誤差分解を通じて、アルゴリズムとしてはカルマンフィルタで実現され、パラメータ σ_v^2 , σ_w^2 が最尤推定される。そのパラメータを所与として固定区間平滑化アルゴリズムを適用することで、潜在変数 $\{\mu_t\}$ に関して、データ終端までの情報を利用した平滑化値 $\{\hat{\mu}_{t|T}\}$ ($t = 1, \dots, T$) が得られる。(以上、例えば北川(2005)など標準的な時系列解析のテキストを参照。計算には We Decomp¹ を用いた。) こうして得られた $\{\hat{\mu}_{t|T}\}$ により、欠測部分の補間を行った。

¹<http://ssnt.ism.ac.jp/inets2/title.html>

1.2 定常性の確保

水準 (level) で見た株価時系列はノイズの累積過程として特徴付けられ、データは顕著な平均非定常性を示す。定常性とは、時系列に関する時間方向でのある種の一様性。過去の自分自身との相関関係が、どこで測るかに依存しないことである。後に述べるグレンジャーの因果性検定の枠組みが、定常時系列による多変量自己回帰モデルに基づくため、分析に先立ってデータの定常性はチェック・確保しておかなければならない。

実際、日経平均株価月次系列に t 型 augmented Dickey-Fuller 検定 (以下 ADF 検定と略、Said and Dickey (1984)) を適用すると、水準では定常性の検定をパスしない。具体的には、定数項を含む定式化 (いわゆる type II)

$$\Delta y_t = \alpha + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

で $\gamma = 0$ の t 型検定を行うと、下側 1% 有意点の -3.44 に対し検定統計量は -1.83 となり、「株価が単位根過程である」という帰無仮説は棄却されない²。下側 10% 有意点を使うとしても -2.57 であり、結論は動かない。ちなみに有意性で生き残る差分ラグ項はなく、結果的には差分ラグ項を augment する必要がなかったことがわかる。

そこで株価時系列を対数変換後差分を取り、収益率に変換する。収益率時系列に対し、typeII の定式化で t 型 ADF 検定を適用すると、検定統計量は -20.0 となり、下側 1% 有意点 -3.44 を遙かに超えて左裾に落ちる。従って収益率時系列は定常とみなしてよい。このとき、推定された定数項は有意でない。すなわち、株価収益率の長期的水準はほぼゼロと見なしてよい。そこで念のため、定数項なし (typeI の定式化) で ADF 検定を行っても、左裾 1% 有意点 -2.57 に対して検定統計量は -19.98 となり、単位根仮説が棄却されるという結論には変わりがない。また、収益率でも有意に残る差分ラグ項はない。

一方、内閣支持率は定義から上下に有界なデータであり、平均非定常とは考えられない。実際 typeII の t 型 ADF 検定を適用すると、下側 1% 有意点 -3.44 に対して検定統計量は -5.90 で、単位根仮説は棄却される。すなわち、内閣支持率データは一定水準周りを変動する定常過程と見なして良い。

図 1 に、当該期間の内閣支持率と株価収益率をプロットを示す。一見して両時系列は、観測頻度は共通に揃えているものの、周波数特性は大きく違うとの印象を得る。

2. グレンジャーの因果性検定

時系列 y_t を予測するのに、他の時系列 x_t の過去の値が役立つとき、時系列 x_t から y_t に、「グレンジャーの意味で因果性がある」と言う。時系列 y_t を、自己ラグ y_{t-1}, y_{t-2}, \dots とクロスラグ x_{t-1}, x_{t-2}, \dots に線形回帰し、「 x_{t-1}, x_{t-2}, \dots の係数がゼロ」(つまり不要) という帰無仮説の検定を行う。すなわち、

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + \epsilon_t$$

²以降の解結果析は全て EViews 7.2 を使用して得たものである。

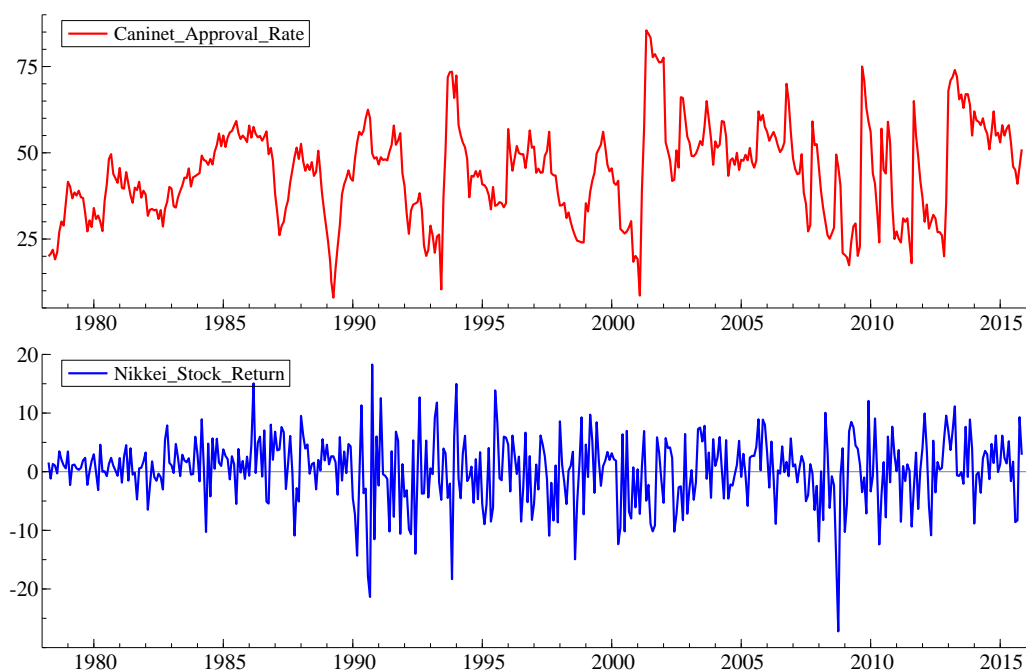


図 1: 内閣支持率 (上段) と株価収益率 (下段)

でよいのか

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + b_1 x_{t-1} + \dots + b_q x_{t-q} + \eta_t$$

であるべきなのかを, $b_1 = \dots = b_q = 0$ を帰無仮説とする F 検定を行って結論する (Granger, 1969). 従って帰無仮説は「(グレンジャーの意味で)³ 因果性がない」である. 仮説が棄却されれば, 時系列 x_t から y_t に「グレンジャーの意味で因果性がある」と言う.

このように自己回帰分布ラグ型モデルに基づくよりも, 多変量自己回帰モデルを推定し, その中でグレンジャー因果性を論じることのほうが多い. すなわち, $Y_t = (y_{1t}, \dots, y_{kt})$ という k 変量時系列ベクトル (平均調整後) に対し,

$$Y_t = \sum_{\tau=1}^P A_\tau Y_{t-\tau} + u_t$$

を推定する. 第 i 時系列 y_{it} から第 j 時系列 y_{jt} に対してグレンジャーの意味で因果性があるとは, $A_\tau(j, i)$ ($\tau = 1, \dots, P$) のうち少なくとも一つは有意にゼロより大きい場合である. 本報告で注目しているのは内閣支持率と株価収益率の関係であるので, 2 変量 VAR に基づいて因果関係を分析することになる.

2.1 内閣支持率から株価収益率へ

³このような断り書きをつけるのは, グレンジャー因果性は予測性を議論しているに過ぎず, 誤解を招くという批判もあるからである.

内閣支持率から株価収益率への因果性は、さまざまにラグの定式化を変えても検出されない。すなわち、内閣支持率が上がったからといって、株価が上がるわけではない。ラグ 11 期 (選択の理由は次の小節に記載) まで取った多変量自己回帰モデルに基づいてグレンジャーの因果性検定を行うと、検定統計量の値は $F = 0.63$ 、その P 値は 0.81 で、「因果性がない」という帰無仮説は棄却されない。

2.2 株価収益率から内閣支持率へ

ラグ構造に関する定式化を探索すると、内閣支持率の自己ラグは 10 期 (y_{t-10}) が有意で、クロスラグ (つまり株価収益率の過去変数) は 11 期 (x_{t-11}) が有意なので、ラグ 11 期までの多変量自己回帰モデルを推定した。検定統計量は $F = 2.14$ 、その P 値は 0.017、すなわち 5% 有意水準で「株価収益率は内閣支持率に影響を与えない」という帰無仮説は棄却される。しかし、問題はその効果の大きさである。

2.3 因果性 (予測性) の大きさ

内閣支持率 y_t は、自身の過去の履歴である y_{t-1} と y_{t-10} (と定数項) があれば、その変動の 73% は説明がつく。これに 2 か月前の株価収益率 (x_{t-2}) をはじめ、 x_{t-6} 、 x_{t-9} 、 x_{t-11} を加えると、少しだけ説明能力は向上して 74% の変動を説明できる。そう観察した上で結論づければ、「株価が上がれば、内閣支持率は上昇する」という主張は、統計的には支持されるが、その影響度はかなり小さいと考えるべきである。

2.4 ハネムーン効果

経験的に観察されることとして、内閣発足当初の「ご祝儀」的期待感を反映して、内閣支持率が高まることがある。これを指して「ハネムーン効果」と呼ぶことがある。ここでは、内閣発足後最初の調査月に 1 を取るダミー時系列をモデルに外生変数として組み込む。

分析の結果、ハネムーン効果は高度に有意 (t 値 10.3) である。因果性検定の結果は変わらず、変動の説明割合は 78% に上昇する。また、このとき内閣支持率のラグ変数は 1ヶ月前 (y_{t-1}) だけが、株価収益率のラグ変数は 11ヶ月前 (x_{t-11}) のみが 5% 有意となる。情報量規準 AIC の値も、この分析を通じて最小 (つまり最良モデル) となった。

3. インパルス応答

インパルス応答関数は一変量の ARMA モデルを考えても出てくる内容であるが、多変量の場合には変数相互の影響をダイナミクスを交えてみるという観点が出てくるので、応用のツールとしてより広がりが出てくる。定常 VAR であれば、(無限次元の) 移動平均表現を与えることができた。

$$y_t = \Phi_0 u_t + \Phi_1 u_{t-1} + \dots \quad (1)$$

ここで u_t の分散共分散行列 Σ_u を $\Sigma_u = PP^T$ と Cholesky 分解し、 $\varepsilon_t = P^{-1}u_t$ 、 $\Psi_s = \Phi_s P$ とすると、

$$y_t = \Psi_0 \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \dots \quad (2)$$

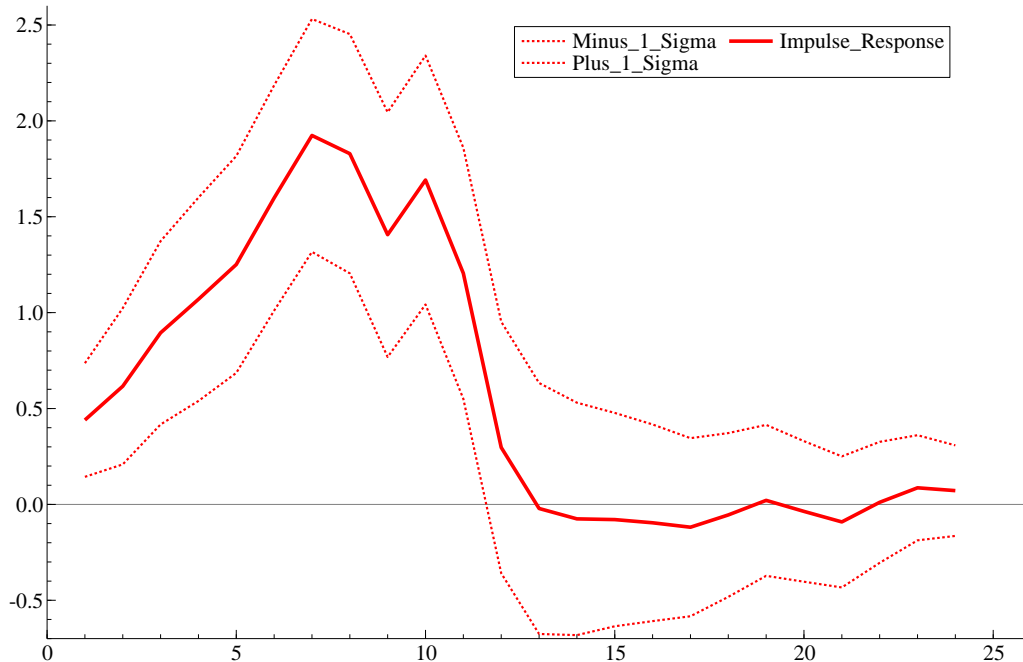


図 2: 株価収益率から内閣支持率へのインパルス応答

となる。 Ψ_s の (i, j) 要素 ψ_{ij}^s は、第 j 変数のイノベーションに 1 単位のショックが加わったときに、 s 期先の第 i 観測時系列 $y_{i,t+s}$ に及ぼす影響度を示す動学乗数と解釈できる。それがインパルス応答である。

$$\frac{\partial y_{j,t+s}}{\partial \varepsilon_{j,t}} = \frac{\partial y_{j,t}}{\partial \varepsilon_{j,t-s}} = \psi_{ij}^s \quad (3)$$

では、前節までに推定したモデルに基づき、株価上昇の経時的効果をインパルス応答で観察しよう。図 2 のグラフの横軸は月単位で 24 ヶ月までを、縦軸は内閣支持率の対する効果 (%) を示している。ある時点で株価が 5.8% (すなわちこれが標準偏差にして 1 単位の変動に相当) 上昇すると、7 ヶ月後に支持率を 1.3% ポイントから 2.5% ポイント (平均的には 1.9% ポイント) 押し上げる効果があると言える。ただし、12 ヶ月後には信頼区間は 0 をはさみ、効果は消失すると考えられる。

4. 文献解題

最後に、内閣支持率と経済指標の関連性を行った文献を、気づいた範囲で挙げておきたい。いずれもモデルの目的変数は内閣支持率である。

Burden (2013) は、時事通信のデータ (1960 年 6 月から 2006 年 4 月の 539 ヶ月) に基づき、消費者物価指数、失業率、日経平均を用いて分析を行っている。消費者物価指数と失業率に関しては、マイナスの効果が有意と報告されている。

前田 (2011) は、読売内閣支持率 (1979 年 1 月から 2008 年 9 月の 357 ヶ月) に時事通信調査の主観的経済指標 (暮らし向き、景況感) を利用した分析を行っており、暮らし向きの

悪化が支持率に対して有意なマイナス効果を持っていることが報告されている。また、内閣発足ダミー (有意) を始め、さまざまなダミー変数を投入している。

中村 (2006) は、長期記憶モデルを用いて内閣支持率の分析を行っている。投入されている予測変数では、暮らし向きの悪化、ハネムーン効果、内閣改造ダミー等が有意と報告されている。形式的なモデルの適用はどんなデータであっても常に可能だが、そもそも内閣支持率における長期記憶性をどう解釈するかは非常に難しいと思われる。

飯田 (2005) も中村 (2006) 同様、政党支持との関連性で内閣支持率を分析しており、長期記憶時系列モデルを推定しながら、与党支持率、消費者物価指数、ハネムーン効果を検証している。

三宅他 (2001) 所収の第 8 章「内閣支持と経済業績評価」においては、時事データに基づき、ハネムーン効果、景況感、暮らし向き等の変数の影響を見る際、自民支持、他党支持、支持なし等の層別での分析結果も報告している。

謝辞 読売新聞世論調査部・福田昌史氏には、データの提供を始め大変お世話になった。ここに謝意を表したい。

参考文献

- [1] 読売新聞世論調査部、『内閣支持率の「通説」検証』、読売新聞 2015 年 12 月 21 日朝刊 9 面「見る」欄, 2015.
- [2] 北川源四郎, 「時系列解析入門」, 岩波書店, 2005.
- [3] Said, S. E. and Dickey, D. A., Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika*, 71, 599-607, 1984.
- [4] Granger, C. W. J., Investigating Causal Relations by Econometric Models and Cross-spectral Methods, *Econometrica*, 37, 424-438, 1969.
- [5] Burden, B. C., Economic accountability and strategic calibration: The case of Japan's Liberal Democratic Party, *Party Politics*, 1-11, 2013.
- [6] 前田幸男, 内閣支持率と与党支持率 (樋渡展洋・斎藤淳編『政党政治の混迷と政権交代』第 10 章), 東京大学出版会, 219-243, 2011.
- [7] 中村悦大, 多変量長期記憶モデルを用いた政党支持と内閣支持の関係性の分析, 選挙学会紀要, 6, 107-126, 2006.
- [8] 飯田健, 政党支持の内閣支持への影響の時間的变化 — ARFIMA モデルと時変パラメータを用いた時系列分析 —, 選挙学会紀要, 4, 41-61, 2005.
- [9] 三宅一郎, 西澤由隆, 河野勝, 「55 年体制下の政治と経済 — 時事世論調査データの分析 —」, 木鐸社, 2001.

真のデータに線形制約がある際の観測誤差の修正-国民経済計算の場合- (未定稿)

千木良 弘朗

山本 拓

東北大学大学院経済学研究科 (一財) 統計研究会

研究集会「経済統計・政府統計の理論と応用」

2017年2月2日 東京大学

- 0/27 -

1 はじめに

□ “観測誤差と線形制約を伴う真のデータ” とは？

例. 国民経済計算における供給表と使用表

		産業		総計
		製造業	サービス業	
供給表:	商品			
	製造品	900	100	1000
	サービス	300	500	800
総計		1200	600	

		産業		消費	総計
		製造業	サービス業		
使用表:	商品				
	製造品	150	50	750	950
	サービス	250	100	250	600
	賃金	400	250		
	営業余剰	350	100		
総計		1150	500		

- 1/27 -

⇒ 、、、 は等しくなければならない

⇒ 真のデータでは等しいが、観測誤差のために等しくなくなっている

- 2/27 -

真のデータの推定

		産業		総計
		製造業	サービス業	
供給表:	商品			
	製造品	880	90	970
	サービス	260	460	720
総計		1140	550	

		産業		消費	総計
		製造業	サービス業		
使用表:	商品				
	製造品	160	40	770	970
	サービス	270	150	300	720
	賃金	410	280		
	営業余剰	300	80		
総計		1140	550		

⇒ 、、、 を等しくするよう観測誤差を修正する

- 3/27 -

⇒ 修正法は国によって違う。アメリカとオランダは統計学の理論に基づいたストーン法 (Stone, Champernowne and Meade (1942)) を採用

⇒ 本稿では、ストーン法も部分的に使うが、全く新しいアプローチを導入する

- 4/27 -

発表の構成

2. ストーン法と先行研究の概観
3. 新しいアプローチ
4. モンテカルロ実験
5. まとめ

- 5/27 -

3 ストーン法と先行研究の概観

□ モデル

$$\underset{(N \times 1)}{X} = X^* + u, u \sim (0, \Omega_u)$$

$$\underset{(R \times N)}{A'} X^* = h$$

$\left\{ \begin{array}{l} X: \text{観測されたデータ} \\ X^*: \text{真のデータ} \\ u: \text{観測誤差} \\ A, h: X^* \text{が満たすべき線形制約} \end{array} \right.$

- 6/27 -

例. 国民経済計算における供給表と使用表

		産業		総計
		製造業	サービス業	
供給表:	商品			
	製造品	x_1	x_2	
	サービス	x_3	x_4	
	総計			

		産業		消費	総計
		製造業	サービス業		
使用表:	商品				
	製造品	x_5	x_6	x_7	
	サービス	x_8	x_9	x_{10}	
	賃金	x_{11}	x_{12}		
	営業余剰	x_{13}	x_{14}		
	総計				

→ $A'X^* = h$ は、

$$\begin{bmatrix} 1 & 1 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_{14}^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- 7/27 -

□ ストーン法

$$\min(X - X^*)' \Omega_u^{-1} (X - X^*), \text{ s.t. } A' X^* = h$$

$$\rightarrow \widehat{X}^* = X - \Omega_u A (A' \Omega_u A)^{-1} (A' X - h)$$

別の解釈:

簡単化のため $h = 0$ とすると制約条件は $A' X^* = 0$ で

$$\begin{aligned} \widehat{X}^* &= X - \Omega_u A (A' \Omega_u A)^{-1} A' X \\ &= (I - \tilde{A} (A' \tilde{A})^{-1} A') X \quad (\tilde{A} = \Omega_u A) \\ &= A_{\perp} (\tilde{A}'_{\perp} A_{\perp})^{-1} \tilde{A}'_{\perp} X \quad (\text{“}_{\perp}\text{” は直行補空間}) \end{aligned}$$

→ X^* が A_{\perp} に属している（フルランク N ではなくランク落ちして $N - R$ ）ため、 X も A_{\perp} に属するよう変換しているのがストーン法

- 8/27 -

□ Ω_u の設定について

ストーン法は Ω_u が解らないと実行不能

→ 統計理論に基づいた解決法はそう多くなく、ある程度恣意的に決められることが多い

	業種						
	農林水産業	鉱業	製造業	建設業	電気・ガス・水道業	卸売・小売業	金融・保険業
農林水産業	1,981.2	1.5	8,412.4	250.1	1.7	1,605.1	0.0
鉱業	0.3	8.4	8,506.8	1,037.0	2,336.4	3.5	0.0
製造業	3,510.3	281.2	138,525.2	28,727.9	1,669.4	5,702.9	1,517.6
建設業	85.5	9.7	1,410.0	208.1	1,131.2	554.8	160.1
電気・ガス・水道業	106.8	43.3	6,522.2	500.4	1,335.1	1,128.4	220.5
卸売・小売業	4.0	0.0	0.0	0.0	0.0	672.0	0.0
金融・保険業	137.3	42.1	1,419.1	480.4	224.5	1,723.5	1,150.0

野木森 稔, 「加重最小二乗法を利用したバランシング・モデル-SUT バランスシステム開発に向けた一考察」, 季刊国民経済計算, 147, 69-89, 2012年
より抜粋

- 9/27 -

○ 統計理論に基づいた解決法

Weale (1992) 等のアイデア:

$$X_t = X_t^* + u_t, u_t \sim (0, \Omega_u), t = 1, \dots, T$$

$$\begin{cases} X_t, X_t^*, u_t \text{ は定常} \\ \text{Cov}(X_t^*, u_t) = 0 \end{cases}$$

とすると、簡単化のため $E(X_t^*) = 0$ として

$$\begin{aligned} \widehat{A'\Omega_u} &= \frac{1}{T} \sum_{t=1}^T A' X_t X_t' \\ &= \frac{1}{T} \sum_{t=1}^T A' u_t (X_t^* + u_t)' \xrightarrow{p} A'\Omega_u \end{aligned}$$

⇒ Ω_u ではなく $A'\Omega_u$ の一致推定でストーン法を実行可能にしている

- 10/27 -

□ 先行研究の問題点と本稿での新しいアプローチ

問題点:

$$\begin{cases} \widehat{A'\Omega_u} \text{ は } N \ll T \text{ でないとパフォーマンスが悪い} \\ X_t^* \text{ に定常性を仮定するのは非現実的} \\ \text{ストーン法では } u_t \text{ を一致推定できない} \\ \Omega_u \text{ は一致推定できない} \end{cases}$$

新しいアプローチ:

$$\Rightarrow \begin{cases} \text{大きな } N \text{ に対して良好なパフォーマンス} \\ X_t^* \text{ に定常性を仮定しない} \\ u_t \text{ (の一部) を一致推定する} \\ \Omega_u \text{ を一致推定する} \end{cases}$$

- 11/27 -

3 新しいアプローチ

□ モデル

観測誤差に**ファクターモデル**を適用する

$$X_t = X_t^* + u_t = X_t^* + \underset{(N \times q)}{C} \underset{(q \times 1)}{F_t} + \varepsilon_t, t = 1, \dots, T$$

$$\begin{cases} C \text{ はファクター負荷} \\ F_t \text{ はファクター} \\ q \text{ はファクター数} \end{cases}$$

⇒ CF_t を一致推定して観測誤差を修正する。 ε_t についてはストーン法で修正

(簡単化のため、以下では $h = 0$ とする)

- 12/27 -

□ モデルの推定の直感的な概観

$$X_t = X_t^* + CF_t + \varepsilon_t, \varepsilon_t \sim (0, \Omega), \Omega = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \dots & \\ 0 & & \sigma_N^2 \end{bmatrix}$$

↓

$$A'_N X_t = A'_N C F_t + A'_N \varepsilon_t, A_N = A \Upsilon_N^{-1}, \Upsilon_N = \begin{bmatrix} n_1 & & 0 \\ & \dots & \\ 0 & & n_R \end{bmatrix},$$

n_r は A の第 r 列の non-zero 要素の数

↓

この式に Bai (2003) の方法を適用して $A'_N \widehat{CH}'^{-1}$ と $\widehat{H}' F_t$ を得る

- 13/27 -



$Cov(X_t^*, \varepsilon_t) = 0$ 、 $Cov(F_t, \varepsilon_t) = 0$ を仮定すると

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \Upsilon_N (A'_N X_t - A'_N \widehat{C} \widehat{H}^{-1} \widehat{H}' F_t) X_t' &\approx \frac{1}{T} \sum_{t=1}^T \widehat{A}' \varepsilon_t X_t' \\ &\approx \frac{1}{T} \sum_{t=1}^T \widehat{A}' \varepsilon_t \varepsilon_t' \\ &\approx A' \Omega \end{aligned}$$



この式を積率条件としてGMMで $\hat{\sigma}_1^2, \dots, \hat{\sigma}_N^2$ を得る

- 14/27 -



$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T A'_N X_t X_t' &\approx A'_N C \Omega_F C' + A'_N \Omega \quad (Cov(X_t^*, F_t) = 0 \text{を追加}) \\ &= A'_N C H^{-1} H' \Omega_F H H^{-1} C' + A'_N \Omega \end{aligned}$$

において、

$$\underbrace{\frac{1}{T} \sum_{t=1}^T A'_N X_t X_t'}_{\text{計算可能}} = \underbrace{A'_N C H^{-1} H' \Omega_F H}_{A'_N \widehat{C} \widehat{H}^{-1} \text{と } \widehat{H}' F_t \text{で推定可能}} \underbrace{H^{-1} C'}_{\text{未知パラメーター}} + \underbrace{A'_N \Omega}_{\hat{\sigma}_1^2, \dots, \hat{\sigma}_N^2 \text{で推定可能}}$$



この式から $H^{-1} C'$ を逆算して $\widehat{H}^{-1} C'$ を得る

- 15/27 -

仮定:

1. $E(F_t) = 0$ 、 $F_t, \varepsilon_t \sim I(0)$

2. A の第 r 列と第 s 列で共通する項目の数 c_{rs} が十分小さい
(全ての r について $\sum_{s=1}^R c_{rs}$ が $R \rightarrow \infty$ としても発散しない)

3. $q < R$

4. A の全行には少なくとも1つ non-zero 要素がある

5. $X_t^* \sim I(0)$

6. $N, R, T \rightarrow \infty$ with $\frac{N^K}{\sqrt{T}R^{1-W}} \rightarrow 0$

($\sum_{r=1}^R n_R = O(N^K)$, $1 \leq K \leq 2$ 、
 n_1, \dots, n_R の内で $O(1)$ なものの個数を $O(R^W)$, $0 \leq W \leq 1$)

- 16/27 -

$$\widehat{CH}'^{-1}\widehat{H}'F_t \xrightarrow{p} CF_t \quad (\text{観測誤差 (の一部) の一致推定})$$
$$\hat{\sigma}_1^2, \dots, \hat{\sigma}_N^2 \xrightarrow{p} \sigma_1^2, \dots, \sigma_N^2 \quad (\text{ストーン法の改善})$$

□ 実際の procedure

ファクター-ストーン法:

$$\widetilde{X}_t^* = X_t - \widehat{CH}'^{-1}\widehat{H}'F_t \text{ に、 } \hat{\Omega} = \begin{bmatrix} \hat{\sigma}_1^2 & & 0 \\ & \dots & \\ 0 & & \hat{\sigma}_N^2 \end{bmatrix} \text{ を使ってス}$$

トーン法を適用

$$\rightarrow \widehat{X}_t^* = \widetilde{X}_t^* - \hat{\Omega}A(A'\hat{\Omega}A)^{-1}A'\widetilde{X}_t^*$$

- 17/27 -

□ $X_t^* \sim I(0)$ を緩めるには?

Nelson and Plosser (1982): GNPは $I(1)$

→ $X_t^* \sim I(1)$ とする

⇒ $Cov(X_t^*, \varepsilon_t) = 0 \cdot Cov(X_t^*, F_t) = 0$ を仮定しても
 $\frac{1}{T} \sum_{t=1}^T X_t^* \varepsilon_t \approx 0 \cdot \frac{1}{T} \sum_{t=1}^T X_t^* F_t \approx 0$ とならず問題

⇒ $\Delta X_t = \Delta X_t^* + C \Delta F_t + \Delta \varepsilon_t$ だと厳しい仮定が必要

Perron (1989): GNPは構造変化を伴うトレンド定常

Diebold and Rudebusch (1989): GNPは長期記憶過程

⇒ エルゴード性が成立する範囲まで X_t^* に構造変化付きトレンドや長期記憶過程を入れて一般化する

- 18/27 -

4 モンテカルロ実験

□ DGPと実験の設定

$$X_t = X_t^* + u_t = X_t^* + C F_t + \varepsilon_t$$

$$\begin{cases} X_t^* \sim \text{トレンド定常 VAR}(1) \\ F_t \sim \text{期待値0の定常 VAR}(1) & (C \Omega_F C' / \text{Var}(X_t^*) \approx 5\%) \\ \varepsilon_t \sim \text{期待値0の定常 VAR}(1) & (\Omega / \text{Var}(X_t^*) \approx 5\%) \\ X_t^*, F_t, \varepsilon_t \text{は全て正規乱数で互いに独立} \end{cases}$$

⇒ $\text{Var}(X_t)$ の内、 $\text{Var}(X_t^*)$ が約90%で Ω_u が約10%

$\text{Var}(X_t^*)$ 、 $C \Omega_F C'$ 、 Ω 全てにおいて最小の分散と最大の分散には約500~1000倍の差をつけ、不均一性を表現

- 19/27 -

- ・ X_t^* の VAR(1) パラメーターと定数・トレンド項は乱数で生成して固定
- ・ F_t と ε_t の VAR(1) パラメーターは乱数で生成して固定
- ・ C は乱数で生成して固定
- ・ A は $-1, 0, 1, 2$ しかとらない sparse 行列として乱数で生成して固定
- ・ $q = 2$ で既知
- ・ $N = 25, 100, 400$ 、 $R = N/4$ 、 \sqrt{N} 、 $T = 25, 50, 800$
- ・ 繰り返し回数 1000 回

□ 各修正法の具体的な計算

- 原始的なストーン法 (Stone1)

$$\Omega_u = I_N \text{ として } \widehat{X}_t^* = X_t - A(A'A)^{-1}A'X_t$$

- 20/27 -

- Weale (1992) の方法 (Weale)

$$\text{demean, detrend した } X_t^d \text{ より } \widehat{A'\Omega_u} = \frac{1}{T} \sum_{t=1}^T A'X_tX_t^{d'} \text{ と}$$

$$\text{して } \widehat{X}_t^* = X_t - \widehat{\Omega_u}A(\widehat{A'\Omega_u}A)^{-1}A'X_t$$

- ファクター-ストーン法 (F-Stone)

先述の通りだが、

$$\begin{cases} \frac{1}{T} \sum_{t=1}^T \Upsilon_N(A'_N X_t - A'_N \widehat{CH}^{-1} \widehat{H}' F_t) X_t^{d'} \approx A' \Omega \\ \frac{1}{T} \sum_{t=1}^T A'_N X_t X_t^{d'} \approx A'_N C H^{-1} H' \Omega_F H H^{-1} C' + A'_N \Omega \end{cases}$$

のように X_t^d で計算する。また、 $\hat{\sigma}_i^2 \leq 0$ だった時は $\hat{\sigma}_i^2 = 0.01 \times \frac{1}{T} \sum_{t=1}^T (x_{it}^d)^2$ とする

- 21/27 -

□ 推定精度の尺度

MSE = $E\{(X_t^* - \widehat{X}_t^*)'(X_t^* - \widehat{X}_t^*)\}$ だと、

$\left\{ \begin{array}{l} \text{変動の大きい } x_{it}^* - \hat{x}_{it}^* \text{ だけに dominate されてしまい不公平} \\ N \text{ や } R \text{ を変えると DGP のパラメーターの値の変化を反映して} \\ \text{しまい、純粋な sample size の効果を捉えられない} \end{array} \right.$

基準化MSE (sMSE) :

$$\Rightarrow \text{sMSE} = \frac{1}{N} E \left\{ (X_t^* - \widehat{X}_t^*)' \begin{bmatrix} \text{Var}(x_1^*) & & 0 \\ & \dots & \\ 0 & & \text{Var}(x_N^*) \end{bmatrix}^{-1} (X_t^* - \widehat{X}_t^*) \right\}$$

$$= \frac{1}{N} E \left\{ \sum_{i=1}^N \frac{(x_{it}^* - \hat{x}_{it}^*)^2}{\text{Var}(x_i^*)} \right\}$$

- 22/27 -

$$\Rightarrow \left\{ \begin{array}{l} x_{it}^* = \mu_i + \delta_i t + z_{it}, E(z_{it}) = 0, z_{it} \sim I(0) \\ \hat{x}_{it}^* = \mu_i + \delta_i t \quad (\text{確定項だけ捉えている“最低限”の推定量}) \end{array} \right.$$

とすると、 $E \left\{ \frac{(x_{it}^* - \hat{x}_{it}^*)^2}{\text{Var}(x_i^*)} \right\} = E \left\{ \frac{z_{it}^2}{\text{Var}(x_i^*)} \right\} = 1$

$$\Rightarrow \left\{ \begin{array}{l} \text{sMSE} < 1 \quad \dots \text{最低限の規準はクリア} \\ \text{sMSE} \geq 1 \quad \dots \text{推定精度に問題あり} \end{array} \right.$$

(実際はsMSEが $t = 1, \dots, T$ について得られるので $\frac{1}{T} \sum_{t=1}^T \text{sMSE}_t$ を計算)

- 23/27 -

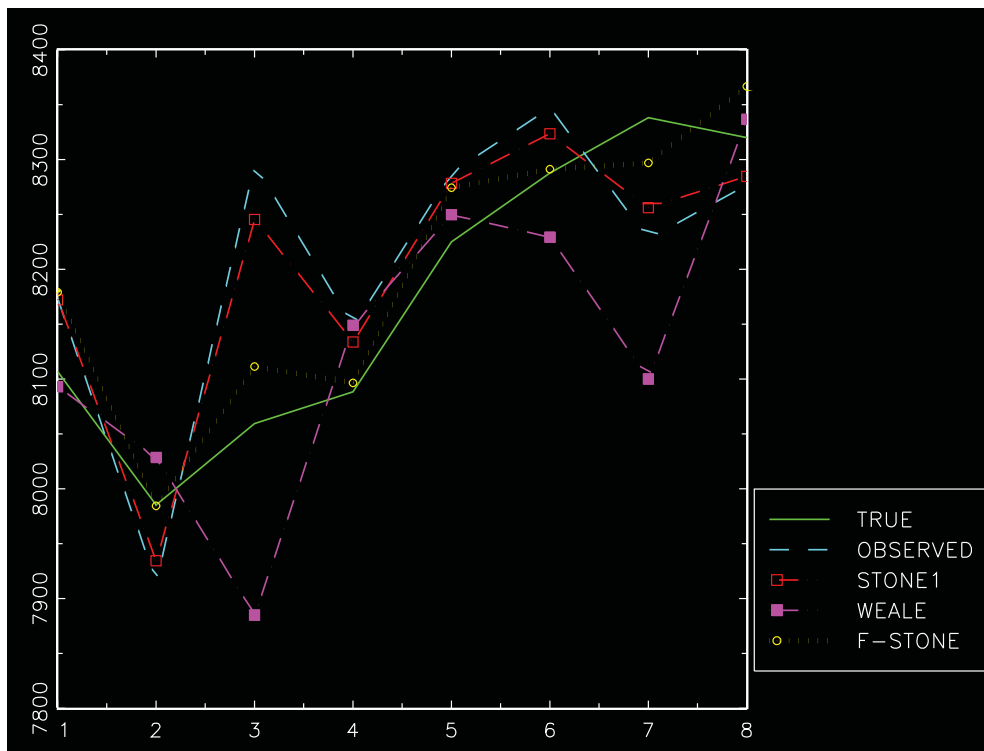
□ sMSE

N	R	Stone1			Weale			F-Stone		
		T			T			T		
		25	50	800	25	50	800	25	50	800
25	5	0.137	0.136	0.136	0.294	0.178	0.062	0.144	0.109	0.064
	7	0.152	0.152	0.151	0.398	0.222	0.060	0.143	0.106	0.057
100	10	0.138	0.139	0.139	0.599	0.307	0.068	0.144	0.110	0.059
	25	0.192	0.192	0.193	.	0.632	0.081	0.147	0.106	0.052
400	20	0.116	0.115	0.115	2.005	0.524	0.081	0.135	0.105	0.053
	100	0.137	0.138	0.138	.	.	0.195	0.139	0.104	0.045

- 24/27 -

□ 実験データの時系列プロット

$N = 100$ 、 $R = 25$ 、 $T = 50$



- 25/27 -

5 まとめ

□ 本稿の貢献

観測誤差を伴う真のデータの推定について、

{ 観測誤差（の一部）の一致推定による直接的な推定
（それ以外の観測誤差に関する）修正精度の改善

を提案した

□ 推定精度

sMSE という尺度では、概ね $F\text{-Stone} > \text{Stone1} > \text{Weale}$

– 26/27 –

□ 今後の課題

$X_t^* \sim I(0)$ という仮定をどこまで緩められるか？

– 27/27 –

Double Filter Instrumental Variable Estimation of Panel Data Models with Weakly Exogenous Variables*

Kazuhiko Hayakawa
Department of Economics,
Hiroshima University

Meng Qi
Department of Economics,
Hiroshima University

Jörg Breitung
Center of Econometrics and Statistics
University of Cologne

January 16, 2017

Abstract

In this paper, we propose instrumental variables (IV) and generalized method of moments (GMM) estimators for panel data models with weakly exogenous variables. The model is allowed to include heterogeneous time trends besides the standard fixed effects. The proposed IV and GMM estimators are obtained by applying a forward filter to the model and a backward filter to the instruments in order to remove fixed effects, thereby called the double filter IV and GMM estimators. We derive the asymptotic properties of the proposed estimators under fixed T and large N , and large T and large N asymptotics where N and T denote the dimensions of cross section and time series, respectively. It is shown that the proposed IV estimator has the same asymptotic distribution as the bias corrected fixed effects estimator when both N and T are large. Monte Carlo simulation results reveal that the proposed estimator performs well in finite samples and outperforms the conventional IV/GMM estimators using instruments in levels in many cases.

*The authors are grateful to Graham Elliott, Jan Kiviet, Koichi Maekawa, Peter Phillips, Hiroshi Yamada, Jun Yu and the participants of SETA 2015 at Hitotsubashi University, CFE 2015 at University of London, ESEM at Geneva, 2016 Japan-Korea Allied Conference in Econometrics at Hitotsubashi University and seminars at Singapore Management University and Nanyang Technological University for helpful comments. This research is supported by KAKENHI(25780153, 15H01943, 16H03606).

1 Introduction

Using panel data in empirical studies has become much more popular than before since many panel data sets are available in these days. Accordingly, many types of panel data models and estimation procedures have been proposed. Among them, the most basic approach is the fixed effects (FE) regression model where unobserved individual specific effects are allowed to be correlated with regressors. However, consistency of the fixed effects estimator relies on the strict exogeneity assumption, i.e., the regressors and idiosyncratic errors are uncorrelated for all periods when the time series dimension, denoted by T , is small. Unfortunately, there are many cases in which the strict exogeneity assumption is violated. A leading example is a dynamic panel model. Regardless of whether the regressors besides the lagged dependent variables are strictly or weakly exogenous, or endogenous, the lagged dependent variable is correlated with the idiosyncratic errors by construction, and hence the fixed effect estimator is inconsistent when T is small (cf. Nickell, 1981). To address this problem, estimation procedures using instrumental variables (IV) have been extensively considered since the work of Anderson and Hsiao (1981). These include, among others, Holtz-Eakin, Newey and Rosen (1988), Arellano and Bond (1991), Arellano and Bover (1995), Ahn and Schmidt (1995) and Blundell and Bond (1998) etc. While most of these studies focus on short panels, there are cases where long panel data are available, typically in macro panels. Inspired by the availability of long panel data, several papers study large N and large T asymptotic properties of aforementioned estimators where N is the number of cross-sectional units. Earlier papers that considered large N and large T dynamic panels are Hahn and Kuersteiner (2002) and Alvarez and Arellano (2003). Hahn and Kuersteiner (2002) and Alvarez and Arellano (2003) demonstrate that, when T and N are large, the fixed effect estimator is consistent but its asymptotic distribution is not centered around the true value in the context of (vector) autoregressive models. To correct for the bias, Hahn and Kuersteiner (2002) also proposes a bias-corrected fixed effects estimator.

More recently, an alternative IV estimator has been proposed in the literature, where a forward demeaning (detrending) is applied to the model while backward demeaning (detrending) is applied to the instruments. We call that IV estimator the double filter IV (DFIV) estimator since unobserved heterogeneity is filtered out forward and backward. Perhaps, the first paper that considers the DFIV estimator is Moon and Phillips (2000) where a near integrated autoregressive panel data model is studied. However, they did not provide theoretical discussion on the DFIV estimator. Hayakawa (2009) considers the DFIV estimator in a stationary panel autoregressive model and derives the asymptotic properties when both N and T are large. A novel feature of the DFIV estimator is that it has the same asymptotic distribution as the bias-corrected FE estimator when T and N are large. Since the DFIV estimator simply uses variables derived from past means as instruments, as opposed to the commonly used level variables, it is quite easy to use in practice¹. From the theoretical point of view, the DFIV estimator has addressed the trade-off problem of using many instruments. Although many instruments are required to improve efficiency, the DFIV estimator becomes efficient despite the same number of instruments as the parameters is used. Hence, the DFIV estimator becomes efficient with the minimal number of instruments. This property has the advantage that it does not cause a large finite sample bias induced by using many instruments. Thereby, the trade-off between bias and efficiency of the generalized method of moments (GMM) estimator is addressed: both the bias and variance of the DFIV estimator become small simultaneously.

¹In the Stata command `xtabond2` by David Roodman, which is routinely used in empirical studies, we can estimate dynamic panel data models by that instruments.

The DFIV estimator is also extended to a panel VAR model (Hayakawa, 2016), dynamic panel simultaneous equation model (Akashi and Kunitomo, 2015) and an infinite order panel autoregressive model (Lee, Okui and Shintani, 2016). However, while there are several nice features as above, unfortunately, the asymptotic equivalence between the DFIV and bias-corrected FE estimators are only proved in the context of (V)AR models, which are somewhat restrictive in practice. One of the purposes of this paper is to demonstrate that this equivalence result holds for more general case with additional regressors. Specifically, we demonstrate that the asymptotic distributions of the DFIV and bias-corrected fixed effect estimator with large N and T are identical for linear panel data models including dynamic models as well as static panel data models with weakly exogenous regressors. Moreover, we demonstrate that this equivalence result holds even when the errors are heteroskedastic and heterogeneous time trends are included in the model, which are not allowed in Hayakawa (2009, 2016). We also investigate the efficiency property of the DFIV and related GMM estimators when T is small and N is large. We conduct Monte Carlo simulation to investigate the finite sample behavior of estimators. Consequently, we find that the DFIV/DFGMM estimators tend to outperform the fixed effects estimator and IV/GMM estimators using instruments in levels.

The rest of this paper is organized as follows. In Section 2, we introduce the models and estimators. In Section 3, the large N and T asymptotic properties of estimators introduced in Section 2 are derived. In Section 4, we carry out Monte Carlo simulation to investigate the finite sample behavior of estimators, and in Section 5, we conclude.

With regard to the notation, we define $T_j = T - j$. For a matrix $\mathbf{A} = \{a_{ij}\}$, a_{ij} denotes the (i, j) element of \mathbf{A} . $\|\mathbf{A}\|^2 = \text{tr}(\mathbf{A}'\mathbf{A}) = \sum_{ij} a_{ij}^2$ denotes the Euclidean norm of a matrix \mathbf{A} .

2 Model and estimators

In this section, we introduce models and estimators. We first consider a model with fixed effects and then consider a model with heterogeneous time trends.

2.1 Fixed effects model

Consider a panel data model with fixed effects, given by

$$y_{it} = \mathbf{w}'_{it}\boldsymbol{\delta} + \eta_i + v_{it}, \quad (i = 1, \dots, N; t = 1, \dots, T) \quad (1)$$

where \mathbf{w}_{it} and $\boldsymbol{\delta}$ are $k \times 1$ vectors. We assume that the error term v_{it} is serially and cross-sectionally uncorrelated. The fixed effect η_i may be correlated with the regressor \mathbf{w}_{it} . Also, we assume that the regressor is weakly exogenous in the sense that $E(\mathbf{w}_{it}v_{is}) = \mathbf{0}$ for $t \leq s$ and $E(\mathbf{w}_{it}v_{is}) \neq \mathbf{0}$ for $t > s$.

This model includes several models as special cases:

Static model: $y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_i + v_{it}$

where $\mathbf{w}_{it} = \mathbf{x}_{it}$, $\boldsymbol{\delta} = \boldsymbol{\beta}$, and \mathbf{x}_{it} and $\boldsymbol{\beta}$ are $q \times 1$ vectors with $k = q$.

AR(p) model: $y_{it} = \alpha_1 y_{i,t-1} + \dots + \alpha_p y_{i,t-p} + \eta_i + v_{it}$,

where $\mathbf{w}_{it} = (y_{i,t-1}, \dots, y_{i,t-p})'$, $\boldsymbol{\delta} = (\alpha_1, \dots, \alpha_p)'$, and \mathbf{w}_{it} and $\boldsymbol{\delta}$ are $p \times 1$ vectors with $k = p$.

ARX(p) model: $y_{it} = \alpha_1 y_{i,t-1} + \dots + \alpha_p y_{i,t-p} + \mathbf{x}'_{it}\boldsymbol{\beta} + \eta_i + v_{it}$

where $\mathbf{w}_{it} = (y_{i,t-1}, \dots, y_{i,t-p}, \mathbf{x}'_{it})'$, $\boldsymbol{\delta} = (\alpha_1, \dots, \alpha_p, \boldsymbol{\beta}')'$, and \mathbf{w}_{it} and $\boldsymbol{\delta}$ are $(p+q) \times 1$ vectors with $k = p+q$.

In a matrix form, the model (1) can be written as

$$\mathbf{y}_i = \mathbf{W}_i \boldsymbol{\delta} + \eta_i \boldsymbol{\nu}_T + \mathbf{v}_i, \quad (i = 1, \dots, N) \quad (2)$$

where $\mathbf{y}_i = (y_{i1}, \dots, y_{iT})'$, $\mathbf{W}_i = (\mathbf{w}_{i1}, \dots, \mathbf{w}_{iT})'$, $\boldsymbol{\nu}_T = (1, \dots, 1)'$ and $\mathbf{v}_i = (v_{i1}, \dots, v_{iT})'$. Define the following matrix that can be used to remove fixed effects:

$$\mathbf{F}_T^t = \text{diag}(c_1^t, c_2^t, \dots, c_{T_1}^t) \begin{bmatrix} 1 & \frac{-1}{T-1} & \frac{-1}{T-1} & \frac{-1}{T-1} & \cdots & \frac{-1}{T-1} & \frac{-1}{T-1} \\ 0 & 1 & \frac{-1}{T-2} & \frac{-1}{T-2} & \cdots & \frac{-1}{T-2} & \frac{-1}{T-2} \\ 0 & 0 & 1 & \frac{-1}{T-3} & \cdots & \frac{-1}{T-3} & \frac{-1}{T-3} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & 1 & \frac{-1}{2} & \frac{-1}{2} \\ 0 & 0 & \cdots & \cdots & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} F_{11}^t & \mathbf{F}_{12}^t & F_{13}^t \\ \mathbf{0}_{T_2 \times 1} & \mathbf{F}_{22}^t & \mathbf{F}_{23}^t \end{bmatrix} \quad (3)$$

$$= \{f_{st}^t\} = \begin{cases} 0 & \text{if } s > t \\ c_t^t = 1 + O\left(\frac{1}{T-t}\right) & \text{if } s = t \\ \frac{-c_s^t}{T-s} = O\left(\frac{1}{T-s}\right) & \text{if } s < t \end{cases}$$

where $c_t^t = \sqrt{(T-t)/(T-t+1)}$, F_{11}^t and F_{13}^t are scalars, \mathbf{F}_{12}^t is $1 \times T_2$, \mathbf{F}_{22}^t is $T_2 \times T_2$, and \mathbf{F}_{23}^t is $T_2 \times 1$.

Multiplying (2) by \mathbf{F}_T^t , the model to be estimated becomes

$$\dot{\mathbf{y}}_i^t = \dot{\mathbf{W}}_i^t \boldsymbol{\delta} + \dot{\mathbf{v}}_i^t, \quad (i = 1, \dots, N) \quad (4)$$

where $\dot{\mathbf{y}}_i^t = \mathbf{F}_T^t \mathbf{y}_i = (\dot{y}_{i1}^t, \dots, \dot{y}_{iT_1}^t)'$, $\dot{\mathbf{W}}_i^t = \mathbf{F}_T^t \mathbf{W}_i = (\dot{\mathbf{w}}_{i1}^t, \dots, \dot{\mathbf{w}}_{iT_1}^t)'$ and $\dot{\mathbf{v}}_i^t = \mathbf{F}_T^t \mathbf{v}_i = (\dot{v}_{i1}^t, \dots, \dot{v}_{iT_1}^t)'$ with $\dot{y}_{it}^t = c_t^t [y_{it} - (y_{i,t+1} + \dots + y_{iT})/(T-t)]$, $\dot{\mathbf{w}}_{it}^t = c_t^t [\mathbf{w}_{it} - (\mathbf{w}_{i,t+1} + \dots + \mathbf{w}_{iT})/(T-t)]$ and $\dot{v}_{it}^t = c_t^t [v_{it} - (v_{i,t+1} + \dots + v_{iT})/(T-t)]$ for $t = 1, \dots, T_1$. Note that the fixed effects η_i is removed by taking a deviation from future means. The t th row of (4) can be written as

$$\dot{y}_{it}^t = \dot{\mathbf{w}}_{it}^t \boldsymbol{\delta} + \dot{v}_{it}^t, \quad (t = 1, \dots, T_1; i = 1, \dots, N) \quad (5)$$

This is the model in forward orthogonal deviations (FOD).

Next, we introduce an instrumental variable. In empirical studies, (a subset of) lagged level variables $\mathbf{w}_{i1}, \dots, \mathbf{w}_{it}$ are commonly used as instruments. Instead of using variables in levels, Hayakawa (2009, 2016) suggest to use variables deviated from past means. To introduce variables deviated from past means, let us define

$$\mathbf{B}_T^t = \text{diag}(c_{T_1}^t, \dots, c_2^t, c_1^t) \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \frac{-1}{2} & \frac{-1}{2} & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \ddots & & \vdots \\ \frac{-1}{T-3} & \frac{-1}{T-3} & \cdots & \frac{-1}{T-3} & 1 & 0 & 0 \\ \frac{-1}{T-2} & \frac{-1}{T-2} & \cdots & \frac{-1}{T-2} & \frac{-1}{T-2} & 1 & 0 \\ \frac{-1}{T-1} & \frac{-1}{T-1} & \cdots & \frac{-1}{T-1} & \frac{-1}{T-1} & \frac{-1}{T-1} & 1 \end{bmatrix} = \{b_{st}^t\}. \quad (6)$$

\mathbf{B}_T^t is obtained by rotating \mathbf{F}_T^t . The mathematical relationship between \mathbf{F}_T^t and \mathbf{B}_T^t is given in (70) in the appendix. Using this, we define an instrumental variable $\ddot{\mathbf{W}}_i^t = \mathbf{B}_T^t \mathbf{W}_i =$

$(\ddot{\mathbf{w}}_{i2}^t, \dots, \ddot{\mathbf{w}}_{iT}^t)'$ where²

$$\ddot{\mathbf{w}}_{it}^t = c_{T-t+1}^t \left[\mathbf{w}_{it} - \frac{\mathbf{w}_{i,t-1} + \dots + \mathbf{w}_{i1}}{t-1} \right], \quad (i = 1, \dots, N; t = 2, \dots, T). \quad (7)$$

Note that the first period is lost due to the difference property of the transformation matrix (6). The transformation that induces $\ddot{\mathbf{w}}_{it}^t$ is called the backward orthogonal deviation (BOD) transformation as opposed to FOD transformation.

Since $E(\ddot{\mathbf{w}}_{is}^t v_{it}^t) = \mathbf{0}$ for $2 \leq s \leq t \leq T_1$ holds, we can construct moment conditions from this. Specifically, we consider the moment conditions $E\left(\sum_{t=2}^{T_1} \ddot{\mathbf{w}}_{it}^t v_{it}^t\right) = \mathbf{0}$. The corresponding instrumental variable estimator is given by

$$\widehat{\boldsymbol{\delta}}_{IV}^t = \left(\sum_{i=1}^N \sum_{t=2}^{T_1} \ddot{\mathbf{w}}_{it}^t \ddot{\mathbf{w}}_{it}^{t'} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=2}^{T_1} \ddot{\mathbf{w}}_{it}^t \dot{y}_{it}^t \right). \quad (8)$$

We call $\widehat{\boldsymbol{\delta}}_{IV}^t$ the *double filter instrumental variable* (DFIV) estimator since it is based on forward and backward filtering.

How are the moment conditions derived?

In Hayakawa (2009), it is shown in the context of AR(p) models that $\ddot{\mathbf{w}}_{it}^t$ has the same structure as the infeasible optimal instruments which leads to efficient estimation. Here, we provide an alternative explanation how the moment conditions $E(\ddot{\mathbf{w}}_{it}^t \dot{v}_{it}^t) = \mathbf{0}$ are derived. For this, let us define two variables r_{it}^b and r_{it}^f for some r_{it} such that $r_{it}^b = r_{it} - (r_{i,t-1} + \dots + r_{i1})/(t-1)$ and $r_{it}^f = r_{it} - (r_{i,t+1} + \dots + r_{iT})/(T-t)$. Note that r_{it}^b is a variable deviated from backward means while r_{it}^f is a variable deviated from forward means. Hence, when $r_{it} = v_{it}$, r_{it}^f and \dot{v}_{it}^t are related such that $r_{it}^f = \dot{v}_{it}^t / c_t^t$. We demonstrate that the moment conditions $E(\ddot{\mathbf{w}}_{it}^t \dot{v}_{it}^t) = \mathbf{0}$ can be obtained from the fixed effects model:

$$(y_{it} - \bar{y}_i) = (\mathbf{w}_{it} - \bar{\mathbf{w}}_i)' \boldsymbol{\delta} + (v_{it} - \bar{v}_i), \quad (i = 1, \dots, N; t = 1, \dots, T) \quad (9)$$

where $\bar{y}_i = T^{-1} \sum_{t=1}^T y_{it}$, $\bar{\mathbf{w}}_i = T^{-1} \sum_{t=1}^T \mathbf{w}_{it}$ and $\bar{v}_i = T^{-1} \sum_{t=1}^T v_{it}$. Note that, after some algebra, $(\mathbf{w}_{it} - \bar{\mathbf{w}}_i)$ and $(v_{it} - \bar{v}_i)$ can be written as

$$\begin{aligned} \mathbf{w}_{it} - \bar{\mathbf{w}}_i &= \frac{t-1}{T} \mathbf{w}_{it}^b + \frac{T-t}{T} \mathbf{w}_{it}^f, \\ v_{it} - \bar{v}_i &= \frac{t-1}{T} v_{it}^b + \frac{T-t}{T} v_{it}^f. \end{aligned}$$

Hence, the covariance between the regressors and error term in (9) becomes

$$\begin{aligned} E[(\mathbf{w}_{it} - \bar{\mathbf{w}}_i)(v_{it} - \bar{v}_i)] &= \frac{(t-1)^2}{T^2} E(\mathbf{w}_{it}^b v_{it}^b) + \frac{(T-t)(t-1)}{T^2} E(\mathbf{w}_{it}^f v_{it}^b) \\ &\quad + \frac{(T-t)(t-1)}{T^2} E(\mathbf{w}_{it}^b v_{it}^f) + \frac{(T-t)^2}{T^2} E(\mathbf{w}_{it}^f v_{it}^f) \neq \mathbf{0}. \end{aligned}$$

This non-zero correlation is the reason why the fixed effect estimator is inconsistent when T is small. However, among the four terms, the third term has zero mean $E(\mathbf{w}_{it}^b v_{it}^f) = \mathbf{0}$, which can be used to consistently estimate $\boldsymbol{\delta}$ even when T is small. Multiplying $c_{T-t+1}^t c_t^t$ to this moment condition in order to account for time series heteroskedasticity, we obtain $c_{T-t+1}^t c_t^t E(\mathbf{w}_{it}^b v_{it}^f) = E(\ddot{\mathbf{w}}_{it}^t \dot{v}_{it}^t) = \mathbf{0}$. This indicates that the proposed moment conditions are derived from the valid part (i.e., no correlation) of the moment conditions of the fixed effects estimator.

²Compared with the form given in Hayakawa (2009), the coefficient c_{T-t+1} is slightly different. However, this is inconsequential and does not affect the main result of this paper. Also, although it is possible to remove c_{T-t+1}^t from $\ddot{\mathbf{w}}_{it}^t$, we keep it to simplify the mathematical relationship between \mathbf{F}_T^t and \mathbf{B}_T^t (see (70)).

2.2 Trend model

Next, we consider a panel data model with usual fixed effects and heterogeneous time trends, given by³

$$y_{it} = \mathbf{w}'_{it}\boldsymbol{\delta} + \eta_i + \lambda_i t + v_{it}, \quad (i = 1, \dots, N; t = 1, \dots, T). \quad (10)$$

In this model, since both η_i and λ_i can be correlated with \mathbf{w}_{it} , the FE estimator augmented with time trend is inconsistent when T is small. In a matrix form, this model can be written as

$$\mathbf{y}_i = \mathbf{W}_i \boldsymbol{\delta} + \eta_i \boldsymbol{\iota}_T + \lambda_i \boldsymbol{\tau}_T + \mathbf{v}_i, \quad (i = 1, \dots, N) \quad (11)$$

where $\boldsymbol{\tau}_T = (1, 2, \dots, T)'$. To remove both η_i and λ_i , we need to multiply a matrix that is orthogonal to both $\boldsymbol{\iota}_T$ and $\boldsymbol{\tau}_T$. While there are several matrices that achieves this (e.g. the second differences), we consider the following matrix:

$$\begin{aligned} \mathbf{F}_T^\tau &= \mathbf{F}_T^{\tau 1} \begin{bmatrix} 1 & \frac{2(-2T_2)}{T_1 T_2} & \frac{2(-2T_2+3)}{T_1 T_2} & \frac{2(-2T_2+6)}{T_1 T_2} & \dots & \dots & \frac{2(-2T_2+3T_3)}{T_1 T_2} & \frac{2(-2T_2+3T_2)}{T_1 T_2} \\ 0 & 1 & \frac{2(-2T_3)}{T_2 T_3} & \frac{2(-2T_3+3)}{T_2 T_3} & \dots & \dots & \frac{2(-2T_3+3T_4)}{T_2 T_3} & \frac{2(-2T_3+3T_3)}{T_2 T_3} \\ 0 & 0 & 1 & \frac{2(-2T_4)}{T_3 T_4} & \dots & \dots & \frac{2(-2T_4+3T_5)}{T_3 T_4} & \frac{2(-2T_4+3T_4)}{T_3 T_4} \\ \vdots & \vdots & 0 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & 1 & \frac{2(-4)}{3 \cdot 2} & \frac{2(-4+3)}{3 \cdot 2} & \frac{2(-4+6)}{3 \cdot 2} \\ 0 & 0 & \dots & \dots & 0 & 1 & \frac{2(-2)}{2 \cdot 1} & \frac{2(-2+3)}{2 \cdot 1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{F}_{11}^\tau & \mathbf{F}_{12}^\tau & \mathbf{F}_{13}^\tau \\ \mathbf{0}_{T_4 \times 2} & \mathbf{F}_{22}^\tau & \mathbf{F}_{23}^\tau \end{bmatrix} \end{aligned} \quad (12)$$

$$= \{f_{st}^\tau\} = \begin{cases} 0 & \text{if } s > t \\ c_t^\tau = 1 + O\left(\frac{1}{T-t}\right) & \text{if } s = t \\ \frac{2c_s^\tau[-2(T-s-1)+3(t-s-1)]}{(T-s)(T-s-1)} = O\left(\frac{1}{T-s}\right) + O\left(\frac{t-s}{(T-s)^2}\right) & \text{if } s < t \end{cases} \quad (13)$$

where $\mathbf{F}_T^{\tau 1} = \text{diag}(c_1^\tau, c_2^\tau, \dots, c_{T_2}^\tau)$, $c_t^\tau = \sqrt{(T-t)(T-t-1)/(T-t+1)(T-t+2)}$, \mathbf{F}_{11}^τ and \mathbf{F}_{13}^τ are 2×2 , \mathbf{F}_{12}^τ is $2 \times T_4$, \mathbf{F}_{22}^τ is $T_4 \times T_4$, and \mathbf{F}_{23}^τ is $T_4 \times 2$. The matrix \mathbf{F}_T^τ is obtained as a GLS transformation of second differences. The formal derivation of \mathbf{F}_T^τ is provided in appendix.

Multiplying (12) to (11), we have the following transformed model

$$\dot{\mathbf{y}}_i^\tau = \dot{\mathbf{W}}_i^\tau \boldsymbol{\delta} + \dot{\mathbf{v}}_i^\tau, \quad (i = 1, \dots, N) \quad (14)$$

where $\dot{\mathbf{y}}_i^\tau = \mathbf{F}_T^\tau \mathbf{y}_i = (\dot{y}_{i1}^\tau, \dots, \dot{y}_{iT_2}^\tau)'$, $\dot{\mathbf{W}}_i^\tau = \mathbf{F}_T^\tau \mathbf{W}_i = (\dot{\mathbf{w}}_{i1}^\tau, \dots, \dot{\mathbf{w}}_{iT_2}^\tau)'$, and $\dot{\mathbf{v}}_i^\tau = \mathbf{F}_T^\tau \mathbf{v}_i = (\dot{v}_{i1}^\tau, \dots, \dot{v}_{iT_2}^\tau)'$. The t th row of (14) can be written as

$$\dot{y}_{it}^\tau = \dot{\mathbf{w}}_{it}^{\tau'} \boldsymbol{\delta} + \dot{v}_{it}^\tau, \quad (t = 1, \dots, T_2; i = 1, \dots, N) \quad (15)$$

where

$$\begin{aligned} \dot{y}_{it}^\tau &= f_{tt}^\tau y_{it} + f_{t,t+1}^\tau y_{i,t+1} + \dots + f_{tT}^\tau y_{iT}, & \dot{\mathbf{w}}_{it}^\tau &= f_{tt}^\tau \mathbf{w}_{it} + f_{t,t+1}^\tau \mathbf{w}_{i,t+1} + \dots + f_{tT}^\tau \mathbf{w}_{iT}, \\ \dot{v}_{it}^\tau &= f_{tt}^\tau v_{it} + f_{t,t+1}^\tau v_{i,t+1} + \dots + f_{tT}^\tau v_{iT}, \end{aligned}$$

and f_{st}^τ is defined in (13).

³Panel data models with heterogeneous time trends are studied by, say, Wansbeek and Knaap (1999), and Phillips and Sul (2007) etc..

Next, to introduce an instrumental variables, we define

$$\mathbf{B}_T^\tau = \mathbf{B}_T^{\tau 1} \begin{bmatrix} \frac{2(-2+3)}{2 \cdot 1} & \frac{2(-2)}{2 \cdot 1} & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \frac{2(-4+6)}{3 \cdot 2} & \frac{2(-4+3)}{3 \cdot 2} & \frac{2(-4)}{3 \cdot 2} & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \frac{2(-2T_4+3T_4)}{T_3 T_4} & \frac{2(-2T_4+3T_5)}{T_3 T_4} & \cdots & \cdots & \frac{2(-2T_4)}{T_3 T_4} & 1 & \cdots & \vdots \\ \frac{2(-2T_3+3T_3)}{T_2 T_3} & \frac{2(-2T_3+3T_4)}{T_2 T_3} & \cdots & \cdots & \frac{2(-2T_3+3)}{T_2 T_3} & \frac{2(-2T_3)}{T_2 T_3} & 1 & 0 \\ \frac{2(-2T_2+3T_2)}{T_1 T_2} & \frac{2(-2T_2+3T_3)}{T_1 T_2} & \cdots & \cdots & \frac{2(-2T_2+6)}{T_1 T_2} & \frac{2(-2T_2+3)}{T_1 T_2} & \frac{2(-2T_2)}{T_1 T_2} & 1 \end{bmatrix} = \{b_{st}^\tau\} \quad (16)$$

where $\mathbf{B}_T^{\tau 1} = \text{diag}(c_{T_2}^\tau, \dots, c_2^\tau, c_1^\tau)$. Note that \mathbf{B}_T^τ can be obtained by rotating \mathbf{F}_T^τ (see (70) in the appendix). Using this, we define an instrumental variable $\ddot{\mathbf{W}}_i^\tau = \mathbf{B}_T^\tau \mathbf{W}_i = (\mathbf{z}_{i3}^\tau, \dots, \mathbf{z}_{iT}^\tau)'$ where its t th row is given by

$$\ddot{w}_{it}^\tau = b_{t-2,t}^\tau \mathbf{w}_{it} + b_{t-2,t-1}^\tau \mathbf{w}_{i,t-1} + \cdots + b_{t-2,1}^\tau \mathbf{w}_{i1}, \quad (i = 1, \dots, N; t = 3, \dots, T) \quad (17)$$

with b_{st}^τ being defined in (16). Note that the first two periods are lost due to the difference property of the transformation matrix \mathbf{B}_T^τ .

Since $E(\ddot{w}_{is}^\tau \dot{v}_{it}^\tau) = \mathbf{0}$, ($3 \leq s \leq t \leq T_2$) holds, we can construct moment conditions from them. Specifically, we consider the moment conditions $E\left(\sum_{t=3}^{T_2} \ddot{w}_{it}^\tau \dot{v}_{it}^\tau\right) = \mathbf{0}$. The corresponding instrumental variable estimator is given by

$$\hat{\delta}_{IV}^\tau = \left(\sum_{i=1}^N \sum_{t=3}^{T_2} \ddot{w}_{it}^\tau \ddot{w}_{it}^{\tau'} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=3}^{T_2} \ddot{w}_{it}^\tau \dot{y}_{it}^\tau \right). \quad (18)$$

This is the DFIV estimator for the model with fixed effects and time trend given by (10).

2.3 Unified model

To derive the asymptotic properties of the proposed IV estimators (8) and (18) in the next section, we formulate the above two models (1) and (11) in a unified framework. For this, let us define a variable d such that $d = 1$ corresponds to the FE model while $d = 2$ corresponds to the trend model. Also, let us define \mathbf{C}_T and \mathbf{F}_T such that $\mathbf{C}_T = \boldsymbol{\nu}_T$ and $\mathbf{F}_T = \mathbf{F}_T^l$ for the FE model, and $\mathbf{C}_T = (\boldsymbol{\nu}_T, \boldsymbol{\tau}_T)$ and $\mathbf{F}_T = \mathbf{F}_T^r$ for the trend model. Thereby, the case $(d, T_d, \mathbf{F}_T, \mathbf{C}_T) = (1, T_1, \mathbf{F}_T^l, \boldsymbol{\nu}_T)$ corresponds to the FE model while $(d, T_d, \mathbf{F}_T, \mathbf{C}_T) = (2, T_2, \mathbf{F}_T^r, (\boldsymbol{\nu}_T, \boldsymbol{\tau}_T))$ corresponds to the trend model. Note that \mathbf{F}_T has the properties such that $\mathbf{F}_T \mathbf{C}_T = \mathbf{0}$, $\mathbf{F}_T \mathbf{F}_T' = \mathbf{I}_{T_d}$ and

$$\mathbf{F}_T' \mathbf{F}_T = \mathbf{Q}_T = \mathbf{I}_T - \mathbf{C}_T (\mathbf{C}_T' \mathbf{C}_T)^{-1} \mathbf{C}_T' = \mathbf{I}_T - \mathbf{R}_T, \quad (19)$$

$$\mathbf{R}_T = \begin{cases} \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' & \text{FE model} \\ \frac{2(2T+1)}{T(T-1)} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' + \frac{12}{T(T-1)(T+1)} \boldsymbol{\tau}_T \boldsymbol{\tau}_T' - \frac{6}{T(T-1)} (\boldsymbol{\nu}_T \boldsymbol{\tau}_T' + \boldsymbol{\tau}_T \boldsymbol{\nu}_T') & \text{trend model} \end{cases} \quad (20)$$

Using these, the models (1) and (11) can be written as

$$\mathbf{y}_i = \mathbf{W}_i \boldsymbol{\delta} + \mathbf{C}_T \boldsymbol{\eta}_i + \mathbf{v}_i, \quad (i = 1, \dots, N). \quad (21)$$

Multiplying (21) by \mathbf{F}_T we obtain the following transformed model:

$$\mathbf{y}_i^* = \mathbf{W}_i^* \boldsymbol{\delta} + \mathbf{v}_i^* \quad (22)$$

where $\mathbf{y}_i^* = \mathbf{F}_T \mathbf{y}_i = (y_{i1}^*, \dots, y_{iT_d}^*)'$. $\mathbf{W}_i^* = \mathbf{F}_T \mathbf{W}_i = (\mathbf{w}_{i1}^*, \dots, \mathbf{w}_{iT_d}^*)'$ and $\mathbf{v}_i^* = \mathbf{F}_T \mathbf{v}_i = (v_{i1}^*, \dots, v_{iT_d}^*)'$. The t th row of (22) can be written as

$$y_{it}^* = \mathbf{w}_{it}^{*'} \boldsymbol{\delta} + v_{it}^*, \quad (i = 1, \dots, N; t = 1, \dots, T_d) \quad (23)$$

Note that the models (5) and (15) are the special cases of (23).

Similarly, let \mathbf{w}_{it}^{**} , ($i = 1, \dots, N$; $t = d+1, \dots, T$) denote $\check{\mathbf{w}}_{it}^t$ for FE model given by (7), and $\check{\mathbf{w}}_{it}^T$ for the trend model given by (17), respectively. Since $E\left(\sum_{t=d+1}^{T_d} \mathbf{w}_{it}^{**} v_{it}^*\right) = \mathbf{0}$, we have the following instrumental variable estimator

$$\hat{\boldsymbol{\delta}}_{IV}^B = \left(\sum_{i=1}^N \sum_{t=d+1}^{T_d} \mathbf{w}_{it}^{**} \mathbf{w}_{it}^{*'} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=d+1}^{T_d} \mathbf{w}_{it}^{**} y_{it}^* \right). \quad (24)$$

This is the DFIV estimator for the unified model given by (23). Note that the previous two estimators (8) and (18) are the special case of (24).

Alternatively, we can consider the GMM estimators that are more efficient than IV estimators especially when T is small. Since the first and last d periods are lost due to difference properties of \mathbf{F}_T and \mathbf{B}_T , the middle T_{2d} periods are used in estimation. Hence, the model in a matrix form becomes

$$\check{\mathbf{y}}_i^* = \check{\mathbf{W}}_i^{*'} \boldsymbol{\delta} + \check{\mathbf{v}}_i^*, \quad (i = 1, \dots, N) \quad (25)$$

where $\check{\mathbf{y}}_i^* = (y_{i,d+1}^*, \dots, y_{iT_d}^*)'$, $\check{\mathbf{W}}_i^* = (\mathbf{w}_{i,d+1}^*, \dots, \mathbf{w}_{iT_d}^*)'$ and $\check{\mathbf{v}}_i^* = (v_{i,d+1}^*, \dots, v_{iT_d}^*)'$. For this model, we consider the moment conditions given by $E(\mathbf{Z}_i^{B'} \check{\mathbf{v}}_i^*) = \mathbf{0}$ where $\mathbf{Z}_i^B = \text{diag}(\mathbf{z}_{i,d_0+1}^{B'}, \dots, \mathbf{z}_{iT_d}^{B'})$, $\mathbf{z}_{it}^B = (\mathbf{w}_{i,t-\ell+1}^{*'}, \dots, \mathbf{w}_{it}^{*'})'$, ($1 \leq \ell \leq t-d$) and $\mathbf{v}_i^* = (v_{i,d_0+1}^*, \dots, v_{iT_d}^*)'$. The corresponding one-step GMM estimator is given by⁴

$$\begin{aligned} \hat{\boldsymbol{\delta}}_{GMM}^B &= \left[\left(\sum_{i=1}^N \check{\mathbf{W}}_i^{*'} \mathbf{Z}_i^B \right) \left(\sum_{i=1}^N \mathbf{Z}_i^{B'} \mathbf{Z}_i^B \right)^{-1} \left(\sum_{i=1}^N \mathbf{Z}_i^{B'} \check{\mathbf{W}}_i^* \right) \right]^{-1} \\ &\quad \times \left[\left(\sum_{i=1}^N \check{\mathbf{W}}_i^{*'} \mathbf{Z}_i^B \right) \left(\sum_{i=1}^N \mathbf{Z}_i^{B'} \mathbf{Z}_i^B \right)^{-1} \left(\sum_{i=1}^N \mathbf{Z}_i^{B'} \check{\mathbf{y}}_i^* \right) \right] \\ &= \left[\sum_{t=d_0+1}^{T_d} \underline{\mathbf{W}}_t^{*'} \underline{\mathbf{Z}}_t^B (\underline{\mathbf{Z}}_t^{B'} \underline{\mathbf{Z}}_t^B)^{-1} \underline{\mathbf{Z}}_t^{B'} \underline{\mathbf{W}}_t^* \right]^{-1} \left[\sum_{t=d_0+1}^{T_d} \underline{\mathbf{W}}_t^{*'} \underline{\mathbf{Z}}_t^B (\underline{\mathbf{Z}}_t^{B'} \underline{\mathbf{Z}}_t^B)^{-1} \underline{\mathbf{Z}}_t^{B'} \underline{\mathbf{y}}_t^* \right] \end{aligned} \quad (26)$$

where $\underline{\mathbf{W}}_t^* = (\mathbf{w}_{1t}^*, \dots, \mathbf{w}_{Nt}^*)'$, $\underline{\mathbf{Z}}_t^B = (\mathbf{z}_{1t}^B, \dots, \mathbf{z}_{Nt}^B)'$ and $\underline{\mathbf{y}}_t^* = (y_{1t}^*, \dots, y_{Nt}^*)'$. We call $\hat{\boldsymbol{\delta}}_{GMM}^B$ the *double filter GMM* (DFGMM) estimator by the same reason as the DFIV estimator.

In order to compare the IV estimator with the FE estimator in the next section, we further reformulate (25) in terms of \mathbf{y}_i , \mathbf{W}_i and \mathbf{v}_i . For this, let us define $\mathbf{L}_T = (\mathbf{0}_{T_d \times d}, \mathbf{I}_{T_d})$. Then, by noting that $\mathbf{L}_T \mathbf{y}_i = (y_{i,d+1}, \dots, y_{iT})'$ and $\mathbf{L}_T \mathbf{W}_i = (\mathbf{w}_{i,d+1}, \dots, \mathbf{w}_{iT})'$, the model (25) can be written as

$$\mathbf{F}_{T_d} \mathbf{L}_T \mathbf{y}_i = \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{W}_i \boldsymbol{\delta} + \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{v}_i, \quad (i = 1, \dots, N). \quad (27)$$

⁴The two-step GMM estimator is not considered in this paper since it requires a computation of large dimensional weighting matrix. Indeed, if the number of moment conditions $m = \sum_{t=d_0+1}^{T_d} m_t$ exceeds the sample size where m_t denotes the number of instruments used at period t , the optimal weighting matrix cannot be computed. However, the one-step GMM estimator (26) can be computed as long as $m_t < N$ for all t even when $m > N$.

Similarly, by using $\mathbf{K}_T = (\mathbf{I}_{T_d}, \mathbf{0}_{T_d \times d})$ and $\mathbf{K}_T \mathbf{W}_i = (\mathbf{w}_{i1}, \dots, \mathbf{w}_{iT_d})'$, we have $\mathbf{W}_i^{**} = (\mathbf{w}_{i,d+1}^{**}, \dots, \mathbf{w}_{iT_d}^{**})' = \mathbf{B}_{T_d} \mathbf{K}_T \mathbf{W}_i$ where \mathbf{B}_T denotes \mathbf{B}_T^L for the FE model and \mathbf{B}_T^r for trend model. Using these, the moment conditions $E\left(\sum_{t=d+1}^{T_d} \mathbf{w}_{it}^{**} v_{it}^*\right) = \mathbf{0}$ can be written as $E(\mathbf{W}_i' \mathbf{K}_T' \mathbf{B}_{T_d}' \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{v}_i) = \mathbf{0}$, and the IV estimator (24) can be written as

$$\widehat{\boldsymbol{\delta}}_{IV}^B = \left(\sum_{i=1}^N \mathbf{W}_i^{**'} \mathbf{W}_i^* \right)^{-1} \left(\sum_{i=1}^N \mathbf{W}_i^{**'} \mathbf{y}_i^* \right) = \left[\sum_{i=1}^N \mathbf{W}_i' \mathbf{K}_T' \mathbf{B}_{T_d}' \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{W}_i \right]^{-1} \sum_{i=1}^N \mathbf{W}_i' \mathbf{K}_T' \mathbf{B}_{T_d}' \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{y}_i. \quad (28)$$

In the next section, we compare the asymptotic properties of this IV estimator and the FE estimator given by

$$\widehat{\boldsymbol{\delta}}_{FE} = \left[\sum_{i=1}^N \mathbf{W}_i' \mathbf{Q}_T \mathbf{W}_i \right]^{-1} \sum_{i=1}^N \mathbf{W}_i' \mathbf{Q}_T \mathbf{y}_i \quad (29)$$

where \mathbf{Q}_T is defined in (19).

Also, for later use, we define IV and GMM estimators using instrument in levels. The IV estimator for model (23) using instruments \mathbf{w}_{it} is given by

$$\widehat{\boldsymbol{\delta}}_{IV}^L = \left(\sum_{i=1}^N \sum_{t=1}^{T_d} \mathbf{w}_{it} \mathbf{w}_{it}^{*'} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=1}^{T_d} \mathbf{w}_{it} y_{it}^* \right). \quad (30)$$

The GMM estimator based on the moment condition $E(\mathbf{Z}_i^{L'} \mathbf{v}_i^*) = \mathbf{0}$ where $\mathbf{Z}_i^L = \text{diag}(\mathbf{z}_{i1}^{L'}, \dots, \mathbf{z}_{iT_d}^{L'})$, $\mathbf{z}_{it}^L = (\mathbf{w}'_{i,t-\ell+1}, \dots, \mathbf{w}'_{it})'$, ($1 \leq \ell \leq t$) is given by

$$\widehat{\boldsymbol{\delta}}_{GMM}^L = \left[\sum_{t=1}^{T_d} \mathbf{W}_t^{*'} \mathbf{Z}_t^L (\mathbf{Z}_t^L \mathbf{Z}_t^L)^{-1} \mathbf{Z}_t^L \mathbf{W}_t^* \right]^{-1} \left[\sum_{t=1}^{T_d} \mathbf{W}_t^{*'} \mathbf{Z}_t^L (\mathbf{Z}_t^L \mathbf{Z}_t^L)^{-1} \mathbf{Z}_t^L \mathbf{y}_t^* \right] \quad (31)$$

where $\mathbf{Z}_t^L = (\mathbf{z}_{1t}^L, \dots, \mathbf{z}_{Nt}^L)'$.

3 Asymptotic properties

In this section, we derive the asymptotic properties of the IV and GMM estimators introduced in the previous section. Specifically, we consider two asymptotic schemes: fixed T and large N asymptotics and large N and large T asymptotics.

We first consider the case with small T and large N , and then consider large N and large T case.

3.1 Fixed T and large N case

Fixed T and large N asymptotic properties of IV and GMM estimators are well established in the literature. Under suitable conditions, the IV and GMM estimators are consistent and asymptotically normally distributed. Moreover, since the GMM estimator exploits more moment conditions than the IV estimator, the GMM estimator is more efficient than the IV estimator. To investigate the efficiency associated with different form of instruments in detail, let us consider the simple AR(1) model with fixed effects:

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it}, \quad (t = 1, \dots, T; i = 1, \dots, N). \quad (32)$$

We make the following assumption

Assumption 1. We assume that (a) $|\alpha| < 1$, (b) the idiosyncratic error term v_{it} is iid with $E(v_{it}) = 0$ and $\text{Var}(v_{it}) = \sigma_v^2$, (c) the unobserved individual effects η_i is iid with $E(\eta_i) = 0$ and $\text{Var}(\eta_i) = \sigma_\eta^2$, (d) the initial conditions follow the stationary distribution:

$$y_{i0} = \frac{\eta_i}{1 - \alpha} + e_{i0} \quad (33)$$

where $e_{i0} \sim \text{iid}(0, \sigma_v^2/(1 - \alpha^2))$.

Most of these assumptions are made just to simplify the theoretical consideration. Indeed, for consistency of IV and GMM estimator, the idiosyncratic term can be heteroskedastic and initial conditions do not need to follow the stationary distribution.

In practice, researchers do not use all past variables as instruments since it causes many instruments problem and resulting GMM estimator is biased. Instead, they used only a few lagged variable in each period. Okui (2009) proposed a statistical procedure to select the number of instruments so that mean-squared error of the GMM estimator is minimized under large T and large N framework. Here, we investigate the effect of lag length of instruments used in each period in terms of efficiency. Since one of the main reasons to use many instruments is to improve efficiency, it would be of interest how efficiency changes depending on the lag length of instruments. To the best of authors' knowledge, such an analysis has not been conducted in the literature even in the simple AR(1) model.

The model after the FOD transformation is given by

$$y_{it}^* = \alpha y_{i,t-1}^* + v_{it}^*, \quad (t = 1, \dots, T - 1; i = 1, \dots, N).$$

Let $\mathbf{z}_{it}^{L(\ell)} = (y_{i,t-\ell}, \dots, y_{i,t-1})'$ and $\mathbf{z}_{it}^{B(\ell)} = (y_{i,t-\ell}^{**}, \dots, y_{i,t-1}^{**})'$ be an $m_t^L \times 1$ and $m_t^B \times 1$ vectors of instruments, respectively, where ℓ denotes the maximum length of instruments used in each period and $y_{i,t-k}^{**} = c_{T-t+k+1}^L [y_{i,t-k} - (y_{i,t-k-1} + \dots + y_{i0})/(t-k)]$, ($1 \leq k \leq \ell$). Note that $m_t^L, m_t^B = t$ for $t < \ell$ and $m_t^L, m_t^B = \ell$ for $t \geq \ell$.

Specifically, the GMM estimator with $\mathbf{z}_{it}^{L(\ell)}$ and $\mathbf{z}_{it}^{B(\ell)}$ as instruments are respectively given as

$$\hat{\alpha}_{GMM}^{L(\ell)} = \left(\sum_{t=1}^{T-1} \mathbf{y}_{t-1}^{*'} \mathbf{P}_t^{L(\ell)} \mathbf{y}_{t-1}^* \right)^{-1} \left(\sum_{t=1}^{T-1} \mathbf{y}_{t-1}^{*'} \mathbf{P}_t^{L(\ell)} \mathbf{y}_t^* \right), \quad (34)$$

$$\hat{\alpha}_{GMM}^{B(\ell)} = \left(\sum_{t=2}^{T-1} \mathbf{y}_{t-1}^{*'} \mathbf{P}_t^{B(\ell)} \mathbf{y}_{t-1}^* \right)^{-1} \left(\sum_{t=2}^{T-1} \mathbf{y}_{t-1}^{*'} \mathbf{P}_t^{B(\ell)} \mathbf{y}_t^* \right) \quad (35)$$

where $\mathbf{y}_{t-1}^* = (y_{1,t-1}^*, \dots, y_{N,t-1}^*)'$, $\mathbf{y}_t^* = (y_{1t}^*, \dots, y_{Nt}^*)'$, $\mathbf{P}_t^{L(\ell)} = \mathbf{z}_t^{L(\ell)} (\mathbf{z}_t^{L(\ell)'} \mathbf{z}_t^{L(\ell)})^{-1} \mathbf{z}_t^{L(\ell)'}$, $\mathbf{P}_t^{B(\ell)} = \mathbf{z}_t^{B(\ell)} (\mathbf{z}_t^{B(\ell)'} \mathbf{z}_t^{B(\ell)})^{-1} \mathbf{z}_t^{B(\ell)'}$, $\mathbf{z}_t^{L(\ell)} = (\mathbf{z}_{1t}^{L(\ell)}, \dots, \mathbf{z}_{Nt}^{L(\ell)})'$ and $\mathbf{z}_t^{B(\ell)} = (\mathbf{z}_{1t}^{B(\ell)}, \dots, \mathbf{z}_{Nt}^{B(\ell)})'$.

The asymptotic variances of $\hat{\alpha}_{GMM}^{L(\ell)}$ and $\hat{\alpha}_{GMM}^{B(\ell)}$ with fixed T and large N are given by

$$\text{Avar} \left(\hat{\alpha}_{GMM}^{L(\ell)} \right) = \sigma_v^2 \left[\sum_{t=1}^{T-1} E \left(y_{i,t-1}^* \mathbf{z}_{it}^{L(\ell)'} \right) \left[E \left(\mathbf{z}_{it}^{L(\ell)} \mathbf{z}_{it}^{L(\ell)'} \right) \right]^{-1} E \left(\mathbf{z}_{it}^{L(\ell)} y_{i,t-1}^* \right) \right]^{-1}, \quad (36)$$

$$\text{Avar} \left(\hat{\alpha}_{GMM}^{B(\ell)} \right) = \sigma_v^2 \left[\sum_{t=1}^{T-1} E \left(y_{i,t-1}^* \mathbf{z}_{it}^{B(\ell)'} \right) \left[E \left(\mathbf{z}_{it}^{B(\ell)} \mathbf{z}_{it}^{B(\ell)'} \right) \right]^{-1} E \left(\mathbf{z}_{it}^{B(\ell)} y_{i,t-1}^* \right) \right]^{-1}. \quad (37)$$

Note that the asymptotic variance $\text{Avar} \left(\hat{\alpha}_{GMM}^{L(\ell)} \right)$ is a function of α and the variance ratio $r = \sigma_\eta^2/\sigma_v^2$ whereas $\text{Avar} \left(\hat{\alpha}_{GMM}^{B(\ell)} \right)$ is a function of α only since a covariance-stationarity is

imposed in Assumption 1 (see Theorem 1). Figure 1 shows the asymptotic variances of $\hat{\alpha}_{GMM}^{L(\ell)}$ and $\hat{\alpha}_{GMM}^{B(\ell)}$ based on (36) and (37) with various lag length of instruments for the cases with $\alpha = 0.3, 0.6, 0.9$, $r = 0.2, 1, 5$ and $T = 10$. For the detail of computation of (36) and (37), see appendix. From the figure, it is found that when $r = 0.2$, the asymptotic variances of the two GMM estimators are very similar regardless of lag length of instruments ℓ . This implies that many lags are not required to improve efficiency. Also it is found that $\hat{\alpha}_{GMM}^L$ tends to be more efficient than $\hat{\alpha}_{GMM}^B$. This is because one estimation period is lost in $\hat{\alpha}_{GMM}^B$ compared to $\hat{\alpha}_{GMM}^L$. However, when $r = 1$ and $r = 5$, the result dramatically changes. From the figure, it is found that $\hat{\alpha}_{GMM}^B$ is little affected by lag length ℓ and hence, we do not need to use many lags; in view of the figure, one lagged instrument is sufficient to obtain nearly efficient estimator. However, this is not the case for $\hat{\alpha}_{GMM}^L$. When $r = 1$ and $r = 5$, the lag length of instruments substantially affects the asymptotic variance. When r is large, $\hat{\alpha}_{GMM}^L$ with one or two lagged instruments are far less efficient than $\hat{\alpha}_{GMM}^B$ despite the same number of instruments is used. But as the lag length increases, efficiency improve. What is striking is that efficiency gain when lag length ℓ is increased from one to two or two from three is substantial. If lag length is more than four, the asymptotic variances of $\hat{\alpha}_{GMM}^L$ and $\hat{\alpha}_{GMM}^B$ are very similar. This implies that although many instruments leads to efficiency gain, in the current case, three or four lags are sufficient to obtain reasonably efficient $\hat{\alpha}_{GMM}^L$. In other words, using higher order lags does not contribute to efficiency gain so much. Thus, this result supports the use of a few lagged variable as instruments. Based on this results, in the following, we mainly consider $\hat{\alpha}_{GMM}^L$ with $\ell = 3$ and $\hat{\alpha}_{GMM}^B$ with $\ell = 1$. From the figure, we also find an interesting relationship between the variance ratio r and instruments lag length ℓ . For example, consider the case with $(\alpha, r) = (0.6, 0.2)$ and $(\alpha, r) = (0.6, 5)$. Comparing these two cases, we find that $\hat{\alpha}_{GMM}^L$ becomes substantially less efficient when ℓ is small if r is increased from 0.2 to 5, while efficiency loss is not so evident when ℓ is large. For instance, the variance of $\hat{\alpha}_{GMM}^L$ with $r = 5$ is about 11.7 times larger than that with $r = 0.2$ when $\ell = 1$. However, when $\ell = 5$, the variance of $\hat{\alpha}_{GMM}^L$ with $r = 5$ is only 1.8 times larger than that with $r = 0.2$. This indicates that the variance inflation of $\hat{\alpha}_{GMM}^L$ caused by a large r is more evident when the instruments lag length is small compared with a large ℓ ; in other words, the extent of reactivity of the variance of $\hat{\alpha}_{GMM}^L$ to a large r can be reduced by using many instruments, i.e., a large ℓ .

Next, we investigate the effect of time length T on the efficiency of IV and GMM estimators with level and BOD instruments. In addition to the GMM estimators (34) and (35), we consider IV estimators given by

$$\begin{aligned}\hat{\alpha}_{IV}^L &= \left(\sum_{i=1}^N \sum_{t=1}^{T-1} y_{i,t-1} y_{i,t-1}^* \right)^{-1} \left(\sum_{i=1}^N \sum_{t=1}^{T-1} y_{i,t-1} y_{it}^* \right), \\ \hat{\alpha}_{IV}^B &= \left(\sum_{i=1}^N \sum_{t=2}^{T-1} y_{i,t-1} y_{i,t-1}^{**} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=2}^{T-1} y_{i,t-1} y_{it}^{**} \right).\end{aligned}$$

Theorem 1. *Asymptotic variances of IV and GMM estimators with fixed T and large N asymptotics are given by*

$$Avar \left(\hat{\alpha}_{GMM}^{L(t)} \right) = (1 - \alpha^2) \left[\sum_{t=1}^{T-1} \psi_t^2 \left(1 - \frac{r \left(\frac{1+\alpha}{1-\alpha} \right)}{1 + r \left\{ \left(\frac{1+\alpha}{1-\alpha} \right) + (t-1) \right\}} \right) \right]^{-1}, \quad (38)$$

$$Avar \left(\hat{\alpha}_{GMM}^{L(3)} \right) = (1 - \alpha^2) \left[\sum_{t=1}^{T-1} \psi_t^2 \left(1 - \frac{r \left(\frac{1+\alpha}{1-\alpha} \right)}{1 + r \left\{ \left(\frac{1+\alpha}{1-\alpha} \right) + m_t^L \right\}} \right) \right]^{-1}, \quad (39)$$

$$Avar\left(\widehat{\alpha}_{GMM}^{L(1)}\right) = (1 - \alpha^2) \left[\sum_{t=1}^{T-1} \psi_t^2 \left(\frac{1}{1 + r \left(\frac{1+\alpha}{1-\alpha} \right)} \right) \right]^{-1}, \quad (40)$$

$$Avar\left(\widehat{\alpha}_{IV}^L\right) = (1 - \alpha^2) \left[1 + r \left(\frac{1+\alpha}{1-\alpha} \right) \right] \left[\sum_{t=1}^{T-1} \psi_t \right]^{-2}, \quad (41)$$

$$Avar\left(\widehat{\alpha}_{GMM}^{B(1)}\right) = (1 - \alpha^2) \left[\sum_{t=2}^{T-1} \psi_t^2 \left(1 - \frac{\alpha\phi_{t-1}}{t-1} \right)^2 A_t^{-1} \right]^{-1}, \quad (42)$$

$$Avar\left(\widehat{\alpha}_{IV}^B\right) = (1 - \alpha^2) \left(\sum_{t=2}^{T-1} c_{T-t+1}^2 A_t \right) \left[\sum_{t=2}^{T-1} \psi_t c_{T-t+1}^2 \left(1 - \frac{\alpha\phi_{t-1}}{t-1} \right)^2 \right]^{-2} \quad (43)$$

where

$$\psi_t = c_t^2 \left[1 - \frac{\alpha\phi_{T-t}}{T-t} \right], \quad (44)$$

$$\phi_j = \frac{1 - \alpha^j}{1 - \alpha} = 1 + \alpha + \dots + \alpha^{j-1}, \quad (45)$$

$$A_t = \left[1 - \frac{2\alpha\phi_{t-1}}{t-1} + \frac{1}{(t-1)^2} \left\{ \frac{(t-1)(1+\alpha)}{1-\alpha} - \frac{2\alpha(1-\alpha^{t-1})}{(1-\alpha)^2} \right\} \right]. \quad (46)$$

Figure 2 depicts $Avar\left(\widehat{\alpha}_{GMM}^{L(3)}\right)$, $Avar\left(\widehat{\alpha}_{IV}^L\right)$, $Avar\left(\widehat{\alpha}_{GMM}^{B(1)}\right)$ and $Avar\left(\widehat{\alpha}_{IV}^B\right)$ for $\alpha = 0.3, 0.6, 0.9$, $r = 0.2, 1, 5$ and $T = 5, 6, \dots, 20$ based on Theorem 1⁵. From the figure, it is found that the efficiency of $\widehat{\alpha}_{IV}^L$ is substantially affected by the variance ratio r . When r is large, $\widehat{\alpha}_{IV}^L$ is much less efficient than other estimators. Also, it is found that the variances of $\widehat{\alpha}_{GMM}^{B(1)}$ and $\widehat{\alpha}_{IV}^B$ are almost identical in all cases. With regard to the effect of T , we find that the difference in efficiency between GMM estimators using instruments in levels and new instruments are not small when T is less than 10 and $r = 0.2$. However, that difference becomes smaller as r increases. Indeed, when r is larger than 1 and T is larger than 10, $\widehat{\alpha}_{GMM}^{L(3)}$, $\widehat{\alpha}_{GMM}^{B(1)}$ and $\widehat{\alpha}_{IV}^B$ have a very similar efficiency property. However, it should be noted that $\widehat{\alpha}_{GMM}^{B(1)}$ and $\widehat{\alpha}_{IV}^B$ use less instruments than $\widehat{\alpha}_{GMM}^{L(3)}$.

Also, from Theorem 1, we heuristically find that $T^{-1}Avar\left(\widehat{\alpha}_{GMM}^{L(t)}\right)$, $T^{-1}Avar\left(\widehat{\alpha}_{GMM}^{B(1)}\right)$ and $T^{-1}Avar\left(\widehat{\alpha}_{IV}^B\right)$ tend to $(1 - \alpha^2)$, which coincides with the asymptotic variance under large T and large N , whereas it is not the case for $Avar\left(\widehat{\alpha}_{GMM}^{L(3)}\right)$, $Avar\left(\widehat{\alpha}_{GMM}^{L(1)}\right)$ and $Avar\left(\widehat{\alpha}_{IV}^L\right)$. A formal discussion under large N and large T asymptotics is given next.

3.2 Large T and large N case

In order to derive the asymptotic property for large N and large T , we first make the following assumptions.

Assumption 2. *The error term v_{it} are serially and cross-sectionally uncorrelated and satisfy*

$$E(v_{it} | \mathbf{w}_{it}, \dots, \mathbf{w}_{i1}, \eta_i) = 0. \quad (47)$$

⁵Since $Avar\left(\widehat{\alpha}_{GMM}^{L(\ell=3)}\right)$ and $Avar\left(\widehat{\alpha}_{GMM}^{L(\ell=t)}\right)$, and $Avar\left(\widehat{\alpha}_{GMM}^{L(\ell=1)}\right)$ and $Avar\left(\widehat{\alpha}_{IV}^L\right)$ are very similar, $Avar\left(\widehat{\alpha}_{GMM}^{L(\ell=t)}\right)$ and $Avar\left(\widehat{\alpha}_{GMM}^{L(\ell=1)}\right)$ are excluded in the figure.

Assumption 3. The regressor \mathbf{w}_{it} follows the process:

$$\mathbf{w}_{it} = \begin{cases} \boldsymbol{\mu}_i + \boldsymbol{\xi}_{it} & \text{FE model} \\ \boldsymbol{\mu}_i + \boldsymbol{\kappa}_{it} + \boldsymbol{\xi}_{it} & \text{trend model} \end{cases} \quad (48)$$

where $E(\boldsymbol{\xi}_{it}) = \mathbf{0}$, $E(\boldsymbol{\xi}_{it}\boldsymbol{\xi}'_{i,t+s}) = \boldsymbol{\Gamma}_{i,s}$ and $\sum_{l=-\infty}^{\infty} \|\boldsymbol{\Gamma}_{i,l}\| < \infty$ for all i . Also, for all i , assume that $E(\boldsymbol{\xi}_{it}v_{is}) = \mathbf{0}$ for $t \leq s$ and $E(\boldsymbol{\xi}_{it}v_{is}) = \boldsymbol{\phi}_{i,t-s} \neq \mathbf{0}$ for $t > s$ where $\sum_{l=1}^{\infty} \|\boldsymbol{\phi}_{i,l}\| < \infty$. $\boldsymbol{\mu}_i$ and $\boldsymbol{\kappa}_i$ are uncorrelated with v_{it} for all i and t , but can be correlated with η_i and λ_i in an unrestricted manner.

Assumption 4. As $N, T \rightarrow \infty$,

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \boldsymbol{\xi}_{it}\boldsymbol{\xi}'_{it} \xrightarrow{p} \boldsymbol{\Gamma}_0, \quad (49)$$

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \boldsymbol{\xi}_{it}v_{it} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \quad (50)$$

where $\boldsymbol{\Gamma}_0 = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \boldsymbol{\Gamma}_{i0}$, $\boldsymbol{\Omega} = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N E(v_{it}^2 \boldsymbol{\xi}_{it}\boldsymbol{\xi}'_{it})$ and both $\boldsymbol{\Gamma}_0$ and $\boldsymbol{\Omega}$ are positive definite.

Assumption 2 indicates that the regressor \mathbf{w}_{it} is weakly exogenous. The correlation structure between regressors and errors are specified in Assumption 3. Assumption 4 is a high-level assumption that can be used to derive the large N and T asymptotic properties. More primitive assumptions can be found in Phillips and Moon (1999).

The following Lemma 1 is useful to understand the relationship between $\widehat{\boldsymbol{\delta}}_{FE}$ and $\widehat{\boldsymbol{\delta}}_{IV}^B$.

Lemma 1. Let Assumptions 2 and 3 hold. Then,

$$(a) \quad \frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{W}_i = \frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{W}_i + O_p\left(\frac{\log T}{T}\right), \quad (51)$$

$$(b) \quad \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{v}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \boldsymbol{\xi}_{it}v_{it} - O_p\left(\sqrt{\frac{N}{T}}\right), \quad (52)$$

$$(c) \quad \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}'_i \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{v}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \boldsymbol{\xi}_{it}v_{it} + o_p(1). \quad (53)$$

Lemma 1(a) is a result for the denominator of the FE and DFIV estimators, and Lemma 1(b) and (c) are the results for the numerator of the FE and DFIV estimators, respectively. Lemma 1(a) indicates that the denominators of $\widehat{\boldsymbol{\delta}}_{FE}$ and $\widehat{\boldsymbol{\delta}}_{IV}^B$ are asymptotically equivalent when T is large. Also, comparing (b) and (c), we find that the second term of the right-hand side makes a significant difference in FE and DFIV estimators. When N/T converges to a non-zero constant, the second term in (b) becomes $O_p(1)$, and because of this, the asymptotic distribution of $\widehat{\boldsymbol{\delta}}_{FE}$ is not centered around the true value as shown in Theorem 2 below. This bias is due the incidental parameter problem. Contrary to the FE estimator, the second term of (c) vanishes asymptotically. Hence, as shown in Theorem 4 below, the asymptotic distribution of $\widehat{\boldsymbol{\delta}}_{IV}^B$ is centered around the true value.

Specifically, the asymptotic distributions of $\widehat{\boldsymbol{\delta}}_{FE}$ and $\widehat{\boldsymbol{\delta}}_{IV}^B$ are given in the following theorems.

Theorem 2. *Let Assumptions 2, 3 and 4 hold. Also assume that $N/T \rightarrow \kappa$, ($0 < \kappa < \infty$). Then the asymptotic distribution of $\widehat{\boldsymbol{\delta}}_{FE}$ as $N, T \rightarrow \infty$ is given by*

$$\sqrt{NT}(\widehat{\boldsymbol{\delta}}_{FE} - \boldsymbol{\delta}) \xrightarrow{d} \mathcal{N}(\mathbf{b}, \boldsymbol{\Gamma}_0^{-1} \boldsymbol{\Omega} \boldsymbol{\Gamma}_0^{-1})$$

where $\mathbf{b} = \sqrt{\kappa} \boldsymbol{\Gamma}_0^{-1} \bar{\mathbf{h}}$ and $\bar{\mathbf{h}} = \text{plim}_{N, T \rightarrow \infty} N^{-1} \sum_{i=1}^N E(\mathbf{W}_i' \mathbf{Q}_T \mathbf{v}_i)$.

This result implies that the asymptotic distribution of the FE estimator is not centered around the true value due to the bias caused by the incidental parameter problem. To correct for this bias, we consider a bias-corrected FE estimator:

$$\widehat{\boldsymbol{\delta}}_{BCFE} = \widehat{\boldsymbol{\delta}}_{FE} - \frac{1}{T} \widehat{\boldsymbol{\Gamma}}_0^{-1} \widehat{\mathbf{h}} \quad (54)$$

where $\widehat{\boldsymbol{\Gamma}}_0 = \frac{1}{NT} \sum_{i=1}^N \mathbf{W}_i' \mathbf{Q}_T \mathbf{W}_i$ and $\widehat{\mathbf{h}}$ is a consistent estimator of $\bar{\mathbf{h}}$. Note that this bias correction is not always possible in practice and feasibility depends on the model specification. For instance, if the model is assumed to be AR(1), then, it is possible to correct the bias as proposed in Hahn and Kuersteiner (2002). However, for other cases, say, for a model with weakly exogenous regressors, bias-correction is infeasible unless a specific form is assumed for the regressors, which is undesirable in practice, since the form of bias depends on the correlation structure between the regressors and errors. Apart from the feasibility, the asymptotic distribution of bias-corrected FE estimator is given in the following theorem.

Theorem 3. *Let Assumptions 2, 3 and 4 hold. Then the asymptotic distribution of $\widehat{\boldsymbol{\delta}}_{BCFE}$ as $N, T \rightarrow \infty$ is given by*

$$\sqrt{NT}(\widehat{\boldsymbol{\delta}}_{BCFE} - \boldsymbol{\delta}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_0^{-1} \boldsymbol{\Omega} \boldsymbol{\Gamma}_0^{-1}).$$

Finally, the asymptotic distribution of the DFIV estimator is given in the following theorem.

Theorem 4. *Let Assumptions 2, 3 and 4 hold. Then, the asymptotic distribution of $\widehat{\boldsymbol{\delta}}_{IV}^B$ as $N, T \rightarrow \infty$ is given by*

$$\sqrt{NT}(\widehat{\boldsymbol{\delta}}_{IV}^B - \boldsymbol{\delta}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_0^{-1} \boldsymbol{\Omega} \boldsymbol{\Gamma}_0^{-1}).$$

This theorem implies that $\widehat{\boldsymbol{\delta}}_{IV}^B$ and $\widehat{\boldsymbol{\delta}}_{BCFE}$ have the same asymptotic distribution when both N and T are large. This is because $\widehat{\boldsymbol{\delta}}_{IV}^B$ does not suffer from the incidental parameter problem nor many instruments problem.

For the GMM estimator, since it is quite involved to derive the large T and large N asymptotic properties for the general model (21), we instead consider a simple AR(1) model given by (32). The asymptotic distributions of IV/GMM estimators for the AR(1) model (32) under large N and large T asymptotics is given in the following theorem.

Theorem 5. *The asymptotic distributions of the FE, IV and GMM estimators under large N and large T are given as follows:*

- (a) $\sqrt{NT} \left[\widehat{\alpha}_{FE} - \left(\alpha - \frac{1}{T}(1 + \alpha) \right) \right] \xrightarrow{d} \mathcal{N}(0, 1 - \alpha^2), \quad \text{if } T^3/N \rightarrow 0,$
- (b) $\sqrt{NT} \left[\widehat{\alpha}_{GMM}^{L(t)} - \left(\alpha - \frac{1}{N}(1 + \alpha) \right) \right] \xrightarrow{d} \mathcal{N}(0, 1 - \alpha^2),$
if $(\log T)^2/N \rightarrow \infty$ and $T/N \rightarrow c$, ($0 \leq c \leq 1$),
- (c) $\sqrt{NT} \left(\widehat{\alpha}_{GMM}^{L(1)} - \alpha \right) \xrightarrow{d} \mathcal{N} \left(0, (1 - \alpha^2) \left(1 + \frac{r(1 + \alpha)}{1 - \alpha} \right) \right), \quad \text{if } T/N \rightarrow c, (0 \leq c \leq 1),$

- (d) $\sqrt{NT}(\hat{\alpha}_{IV}^L - \alpha) \xrightarrow{d} \mathcal{N}\left(0, (1 - \alpha^2) \left(1 + \frac{r(1 + \alpha)}{1 - \alpha}\right)\right)$,
- (e) $\sqrt{NT}(\hat{\alpha}_{GMM}^{B(1)} - \alpha) \xrightarrow{d} \mathcal{N}(0, 1 - \alpha^2)$, if $T/N \rightarrow c, (0 \leq c \leq 1)$,
- (f) $\sqrt{NT}(\hat{\alpha}_{IV}^B - \alpha) \xrightarrow{d} \mathcal{N}(0, 1 - \alpha^2)$.

These results are already derived in the literature. (a) is due to Hahn and Kuersteiner (2002) and Alvarez and Arellano (2003), and (b) is due to Alvarez and Arellano (2003), (c) is derived by Hayakawa (2006) and Hsiao and Zhou (2015), (d) is derived by Hsiao and Zhou (2015), (e) is derived by Hayakawa (2006), and (f) is derived by Hayakawa (2009).

Note that while the GMM estimators require conditions on the relative speed of N and T , such a condition is not required for the IV estimators $\hat{\alpha}_{IV}^L$ and $\hat{\alpha}_{IV}^B$. From Theorem 5, we find that the GMM estimator using instruments in levels becomes efficient (Hahn and Kuersteiner, 2002) if all past variables are used as instruments in each period. However, if only one lagged variable is used as an instrument, the GMM estimator is not efficient and also it has the same asymptotic distribution as IV estimator $\hat{\alpha}_{IV}^L$. However, for IV and GMM estimators using BOD filtered instruments, both $\hat{\alpha}_{GMM}^{B(1)}$ and $\hat{\alpha}_{IV}^B$ become efficient, which implies that we do not need to use many instruments to enhance efficiency.

In the AR(1) model given by (32), Hahn and Kuersteiner (2002) demonstrate that $\mathcal{N}(0, 1 - \alpha^2)$ is the minimal asymptotic distribution under the normality assumption on v_{it} . Hence, the bias-corrected FE estimator given by $\hat{\alpha}_{BCFE} = [(T + 1)/T]\hat{\alpha}_{FE} + (1/T)$, the GMM estimators using all instruments in levels and using one lagged new instrument, and the IV estimator using new instruments are asymptotically efficient. Given this, it is conjectured that $\hat{\delta}_{BCFE}^B$ and $\hat{\delta}_{IV}^B$ are asymptotically efficient under certain conditions. However, a formal discussion is beyond the scope of the present paper and left as a future topic.

4 Monte Carlo simulation

In this section, we investigate the finite sample properties of the proposed estimators in the context of dynamic panel data models with/without time trends.

4.1 Design

The data are generated as

$$\begin{aligned} y_{it} &= \alpha y_{i,t-1} + \beta x_{it} + \eta_i + \varphi \lambda_i t + v_{it}, \\ x_{it} &= \rho x_{i,t-1} + \tau_\eta \eta_i + \varphi \tau_\lambda \lambda_i t + \theta v_{i,t-1} + e_{it}. \end{aligned}$$

Note that the case with $\varphi = 0$ corresponds to the FE model while that with $\varphi = 1$ corresponds to the trend model. In a matrix form, this can be written as

$$\begin{pmatrix} y_{it} \\ x_{it} \end{pmatrix} = \begin{pmatrix} \alpha & \beta\rho \\ 0 & \rho \end{pmatrix} \begin{pmatrix} y_{i,t-1} \\ x_{i,t-1} \end{pmatrix} + \begin{pmatrix} (1 + \beta\tau_\eta) \\ \tau_\eta \end{pmatrix} \eta_i + \varphi \begin{pmatrix} (1 + \beta\tau_\lambda) \\ \tau_\lambda \end{pmatrix} \lambda_i t + \begin{pmatrix} v_{it} + \beta\theta v_{i,t-1} + \beta e_{it} \\ \theta v_{i,t-1} + e_{it} \end{pmatrix}$$

or

$$\mathbf{p}_{it} = \mathbf{\Phi} \mathbf{p}_{i,t-1} + \mathbf{c}_\eta \eta_i + \varphi \mathbf{c}_\lambda \lambda_i t + \boldsymbol{\varepsilon}_{it} \quad (55)$$

where $\mathbf{p}_{it} = (y_{it}, x_{it})'$, $\mathbf{c}_\eta = (1 + \beta\tau_\eta, \tau_\eta)'$, $\mathbf{c}_\lambda = (1 + \beta\tau_\lambda, \tau_\lambda)'$, $\boldsymbol{\varepsilon}_{it} = (v_{it} + \beta\theta v_{i,t-1} + \beta e_{it}, \theta v_{i,t-1} + e_{it})'$ and

$$\boldsymbol{\Phi} = \begin{pmatrix} \alpha & \beta\rho \\ 0 & \rho \end{pmatrix}. \quad (56)$$

Alternatively, \mathbf{p}_{it} can be written as in a component form:

$$\begin{aligned} \mathbf{p}_{it} &= \mathbf{a}_i + \varphi \mathbf{b}_i t + \boldsymbol{\zeta}_{it}, \\ \boldsymbol{\zeta}_{it} &= \boldsymbol{\Phi} \boldsymbol{\zeta}_{i,t-1} + \boldsymbol{\varepsilon}_{it} \end{aligned} \quad (57)$$

where

$$\begin{aligned} \mathbf{a}_i &= (\mathbf{I} - \boldsymbol{\Phi})^{-1} \mathbf{c}_\eta \eta_i - (\mathbf{I} - \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi} (\mathbf{I} - \boldsymbol{\Phi})^{-1} \mathbf{c}_\lambda \lambda_i, \\ \mathbf{b}_i &= (\mathbf{I} - \boldsymbol{\Phi})^{-1} \mathbf{c}_\lambda \lambda_i, \\ \text{Var}(\boldsymbol{\varepsilon}_{it}) &= \begin{bmatrix} (1 + \beta^2\theta^2) \sigma_v^2 + \beta^2 \sigma_e^2 & \beta\theta^2 \sigma_v^2 + \beta \sigma_e^2 \\ \beta\theta^2 \sigma_v^2 + \beta \sigma_e^2 & \theta^2 \sigma_v^2 + \sigma_e^2 \end{bmatrix}. \end{aligned}$$

Data for y_{it} and x_{it} are generated from (57). For the sample size, we consider $T = 10, 25, 50, 100$ and $N = 50, 100, 250$. For parameter values, we consider $\alpha = 0.4, 0.8$, $\beta = 1$, $\rho = 0.5$, $\theta = -0.2$, $\tau_\eta = 0.5$, $\tau_\lambda = 0.5$. v_{it} , e_{it} , η_i and λ_i are independently generated as $v_{it} \sim \mathcal{N}(0, \sigma_v^2)$, $e_{it} \sim \mathcal{N}(0, \sigma_e^2)$, $\eta_i \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\lambda_i \sim \mathcal{N}(0, \sigma_\lambda^2)$ with $\sigma_v^2 = 1$, $\sigma_e^2 = 0.16$, $\sigma_\eta^2 = 1, 5$ and $\sigma_\lambda^2 = 1^6$. We report the median bias, interquartile range (IQR), median absolute error (MAE) and empirical size with 5% significance level based on 2,000 replications.

4.2 Estimators to be compared

We consider seven estimators. The first is the FE estimator $\widehat{\boldsymbol{\delta}}_{FE}$ given in (29).⁷ The second is the IV estimator $\widehat{\boldsymbol{\delta}}_{IV}^L$ given in (30) where instruments in levels are used. The third and fourth are the GMM estimator $\widehat{\boldsymbol{\delta}}_{GMM}^L$ given in (31) where instruments in levels are used. For the choice of lag length of instruments, we consider $\ell = 1$ and 3. The corresponding GMM estimators are denoted as “LEV1” and “LEV3”, respectively. The fifth estimator is the DFIV estimator $\widehat{\boldsymbol{\delta}}_{IV}^B$ defined in (24). The last two estimators are the GMM estimator $\widehat{\boldsymbol{\delta}}_{GMM}^B$ defined in (26) where backward filtered instruments are used. For the choice of lag length of instruments, we consider $\ell = 1$ and 3. The corresponding GMM estimators are denoted as “BOD1” and “BOD3”, respectively. For the computation of standard errors, we use those obtained under large N and fixed T since they are more accurate than those obtained under large N and large T (see Hayakawa, 2015).

4.3 Results

Simulation results are provided in Tables 1-4. We first consider the model with fixed effects only. From Tables 1 and 2, we find that the FE estimator for α is severely biased when $T = 10$. However, as T gets larger, the bias becomes smaller as expected since the FE estimator is consistent when T is large. However, in terms of accuracy of inference, the sizes are severely

⁶Although we tried the cases with $(\sigma_\eta^2, \sigma_\lambda^2) = (1, 5), (5, 5)$, the results are very similar to those with $(\sigma_\eta^2, \sigma_\lambda^2) = (5, 1)$. Hence the results of these cases are not reported to save space.

⁷A bias corrected FE estimator is not compared since it is not available in the current case where the regressors is weakly exogenous.

distorted even when T is large, say, $T = 100$. This is because the asymptotic distribution of the FE estimator is not centered around the true value due to the incidental parameter problem. Also, note that increase in N does not reduce the bias since the bias of FE estimator does not depend on N . With regard to the FE estimator of β , the performance is better than those of α . However, it still shows some bias and size distortions. This result implies that the FE estimator does not work even when T is large. Also, note that a widely acceptable bias-correction method is not available since the regressor is weakly exogenous⁸. With regard to the IV and GMM estimators, in terms of MAE, the IV estimators using instruments in levels perform worst among the four estimators mainly due to the large dispersions. With regard to the remaining three estimators, they perform very similarly in terms of MAE when $T = 10$. However, as T gets larger, IV and GMM estimators using new instruments outperform the GMM estimator using instruments in levels. With regard to the choice of IV or GMM estimators using new instruments, it is observed that the GMM estimator tends to have slightly smaller MAEs than IV estimator. In terms of accuracy of inference, IV and GMM estimators using new instruments have almost correct empirical sizes in all cases while the GMM estimator using instruments in levels have large size distortions especially when $T = 10$ and $\alpha = 0.8$. We find that the efficiency of GMM estimator using instruments in levels depends substantially on the lag length ℓ . Comparing the IQRs with $\ell = 1$ and 3, the reduction of dispersion with $\ell = 3$ is substantial though it induces many instruments. Contrary to IV/GMM with instruments in levels, the effects of ℓ on the performance of the GMM with new instruments are minor and the IQRs are relatively smaller than those of GMM with instruments in levels. This result is consistent with the theoretical implication that using new instruments leads to efficient estimation. Considering overall performance, we may conclude that the IV or GMM estimator using new instruments with $\ell = 1$ tend to perform best in many cases.

Next, we consider the models with both fixed effects and heterogeneous time trends. The results are provided in Tables 3 and 4. Compared with the models with fixed effects only, the FE estimator is severely biased when T is small in this model too, and the magnitude of bias is larger. This also can be seen in the substantial size distortions even for a large $T = 100$. This implies that the FE estimator deteriorates further if time trends are included in the model. With regard to the IV and GMM estimators, the IV estimator using instruments in levels perform poorly compared with other estimators. However, contrary to the previous model, the other three IV and GMM estimators perform poorly when $T = 10$. Compared with the previous model with fixed effects, the dispersion is much larger when $T = 10$. However, the performance of these estimators improves as T gets larger. When $T = 25$ or larger, three estimators perform reasonably well when $\alpha = 0.4$ while more than $T = 50$ is required when $\alpha = 0.8$. For the relative performance among the three estimators, we find that the GMM estimator using instruments in levels perform best when $T = 10$. However, for all other cases, the GMM estimator using new instruments perform best.

5 Conclusion

In this paper, we have proposed a new instrumental variable estimator for panel data models including static and dynamic models with weakly exogenous variables and with fixed effects and/or heterogeneous time trends. We showed that the new IV estimator called the DFIV

⁸If the regressors are strictly exogenous, bias-corrected FE estimators such as Bun and Carree (2005) can be used.

estimator is consistent and has the same asymptotic distribution as the bias-corrected fixed effects estimator, which is sometimes infeasible, when both N and T are large. This implies that the DFIV estimator is as efficient as the fixed effects estimator. Monte Carlo simulation results revealed that the DFIV and DFGMM estimators tend to perform better than the conventional IV/GMM estimators using instruments in levels in almost all cases.

Appendix

Derivation of \mathbf{F}_T

We derive the form of \mathbf{F}_T^l and \mathbf{F}_T^r . Although a brief derivation of \mathbf{F}_T^l is given in Arellano (2003), a complete derivation is not provided. Hence, we fill that gap. Let us define the following $T_1 \times T$ matrix that takes the first difference:

$$\mathbf{D}_T = \begin{bmatrix} -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & \\ & & & \ddots & \ddots \\ 0 & & & & -1 & 1 \end{bmatrix}.$$

Multiplying \mathbf{D}_T by (1) and noting that $\mathbf{D}_T \boldsymbol{\nu}_T = \mathbf{0}$, we have

$$\mathbf{D}_T \mathbf{y}_i = \mathbf{D}_T \mathbf{W}_i + \mathbf{D}_T \mathbf{v}_i$$

where it is simply assumed that $\text{Var}(\mathbf{v}_i) = \sigma_v^2 \mathbf{I}_T$. Since $\text{Var}(\mathbf{D}_T \mathbf{v}_i) = \sigma_v^2 \mathbf{D}_T \mathbf{D}_T'$, the transformed error is serially correlated. To correct for the serial correlation, we use the following transformation matrix, which is a GLS transformation:

$$\mathbf{F}_T^l = (\mathbf{D}_T \mathbf{D}_T')^{-1/2} \mathbf{D}_T,$$

where $(\mathbf{D}_T \mathbf{D}_T')^{-1/2}$ is the *upper* triangular Cholesky factorization of $(\mathbf{D}_T \mathbf{D}_T')^{-1}$ with⁹

$$\mathbf{D}_T \mathbf{D}_T' = \begin{matrix} (T_1 \times T_1) \\ \begin{bmatrix} 2 & -1 & 0 & & 0 \\ -1 & 2 & -1 & & \\ 0 & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots & 0 \\ & & & -1 & 2 & -1 \\ 0 & & & 0 & -1 & 2 \end{bmatrix} \end{matrix}.$$

To compute $(\mathbf{D}_T \mathbf{D}_T')^{-1/2}$, we need to derive the inverse matrix $\mathbf{H}^l = (\mathbf{D}_T \mathbf{D}_T')^{-1} = \{h_{st}^l\}$.¹⁰ Using the results by El-Mikkawy and Karawia (2006), we have

$$h_{st}^l = \begin{cases} \frac{T_1}{T_1+1} & \text{if } s = t = 1 \text{ or } s = t = T_1 \\ \frac{s(T_1-s+1)}{T_1+1} & \text{if } s = t < T_1 \\ \frac{s(T_1-t+1)}{T_1+1} & \text{if } s < t \\ \frac{t(T_1-s+1)}{T_1+1} & \text{if } s > t \end{cases}$$

Next, we need to compute the Cholesky factorization to \mathbf{H}^l . For a $K \times K$ matrix $\mathbf{A} = \{a_{ij}\}$, its Cholesky factorization is given by

$$\mathbf{A} = \mathbf{L}\mathbf{L}'$$

⁹A matrix with this structure is called *tridiagonal* matrix.

¹⁰Arellano (2003) does not provide the details how the upper triangular Cholesky factorization can be computed.

where $\mathbf{L} = (\ell_{ij})$ is the lower triangular matrix. Then using ℓ_{ij} , we can write the elements of \mathbf{A} as follows:

$$\begin{aligned} a_{11} &= \ell_{11}^2, \\ a_{21} &= \ell_{21}\ell_{11}, & a_{22} &= \ell_{21}^2 + \ell_{22}^2, \\ a_{31} &= \ell_{31}\ell_{11}, & a_{32} &= \ell_{31}\ell_{21} + \ell_{32}\ell_{22}, & \ell_{33} &= \ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2, \\ &\vdots & &\vdots & &\ddots \\ a_{K1} &= \ell_{K1}\ell_{11}, & a_{K2} &= \ell_{K1}\ell_{21} + \ell_{K2}\ell_{22}, & \cdots & , a_{KK} = \ell_{K1}^2 + \cdots + \ell_{KK}^2. \end{aligned}$$

ℓ_{ij} can be solved sequentially as follows:

$$\begin{aligned} \ell_{11} &= \sqrt{a_{11}}, \\ \ell_{21} &= a_{21}/\ell_{11}, & \ell_{22} &= \sqrt{a_{22} - \ell_{21}^2}, \\ \ell_{31} &= a_{31}/\ell_{11}, & \ell_{32} &= (a_{32} - \ell_{31}\ell_{21})/\ell_{22}, & \ell_{33} &= \sqrt{a_{33} - \ell_{31}^2 - \ell_{32}^2}, \\ &\vdots & &\vdots & &\ddots \\ \ell_{K1} &= a_{K1}/\ell_{11}, & \ell_{K2} &= (a_{K2} - \ell_{K1}\ell_{21})/\ell_{22}, & \cdots & \ell_{KK} = \sqrt{a_{KK} - \ell_{K1}^2 - \cdots - \ell_{K,K-1}^2}. \end{aligned}$$

The explicit form of \mathbf{F}_T^l is obtained by letting $\mathbf{A} = \mathbf{H}^l$.

Next, we consider a model with individual effects and heterogeneous time trends given by (11). To remove both η_i and λ_i from the model, we take second differences. In terms of a model in matrix, this corresponds to multiplying by $\mathbf{D}_{T_1}\mathbf{D}_T$, ($T_2 \times T$), we have

$$\mathbf{D}_{T_1}\mathbf{D}_T\mathbf{y}_i = \mathbf{D}_{T_1}\mathbf{D}_T\mathbf{W}_i\boldsymbol{\delta} + \mathbf{D}_{T_1}\mathbf{D}_T\mathbf{v}_i.$$

Since the transformed error is serially correlated, we consider the following GLS-type transformation matrix:

$$\mathbf{F}_T^\tau = (\mathbf{D}_{T_1}\mathbf{D}_T\mathbf{D}'_T\mathbf{D}'_{T_1})^{-1/2}\mathbf{D}_{T_1}\mathbf{D}_T,$$

where $(\mathbf{D}_{T_1}\mathbf{D}_T\mathbf{D}'_T\mathbf{D}'_{T_1})^{-1/2}$ is the *upper* triangular Cholesky factorization of $(\mathbf{D}_{T_1}\mathbf{D}_T\mathbf{D}'_T\mathbf{D}'_{T_1})^{-1}$. To compute \mathbf{F}_T^τ , we need to derive the inverse matrix $\mathbf{H}^\tau = (\mathbf{D}_{T_1}\mathbf{D}_T\mathbf{D}'_T\mathbf{D}'_{T_1})^{-1} = \{h_{st}^\tau\}$ with¹¹

$$\mathbf{D}_{T_1}\mathbf{D}_T\mathbf{D}'_T\mathbf{D}'_{T_1} = \begin{matrix} (T_2 \times T_2) \\ \begin{bmatrix} 6 & -4 & 1 & & 0 \\ -4 & 6 & -4 & & \\ 1 & -4 & 6 & -4 & \\ & & \ddots & \ddots & \ddots & 1 \\ & & & -4 & 6 & -4 \\ 0 & & & 1 & -4 & 6 \end{bmatrix} \end{matrix}.$$

Using the results by Dow (2003), we have¹²

$$h_{st}^\tau = \begin{cases} a_{t0}s^3 + a_{t1}s^2 + a_{t2}s, & s \leq t + 1 \\ b_{t0}s^3 + b_{t1}s^2 + b_{t2}s + b_{t3}, & s \geq t + 1 \end{cases}$$

where

$$a_{t0} = -(3 + 2t + T_2)d_t/c, \quad a_{t1} = 3t(1 + T_2)d_t/c, \quad a_{t2} = (3 + 5t + T_2 + 3tT_2)d_t/c,$$

¹¹A matrix with this structure is called *pentadiagonal* matrix.

¹²See also Chen (2013) for an alternative expression.

$$\begin{aligned}
b_{t0} &= (5 - 2t + 3T_2)e_t/c, & b_{t1} &= -3(1 + T_2)(4 - t + 2T_2)e_t/c, \\
b_{t2} &= (1 + 5t + 12T_2 + 3tT_2 + 12T_2^2 + 3T_2^3)e_t/c, & b_{t3} &= (1 - t)e_t/6, \\
d_t &= (T_2 - t + 1)(T_2 - t + 2), & e_t &= t(t + 1), & c &= 6(T_2 + 1)(T_2 + 2)(T_2 + 3).
\end{aligned}$$

Using these results and applying the algorithm of Cholesky factorization introduced above where $\mathbf{A} = \mathbf{H}^\tau$, after a lengthy calculation, we obtain the explicit expression of \mathbf{F}_T^τ as in (12).

Derivation of asymptotic variances (36) and (37)

Let us define $\mathbf{y}_{i,-1} = (y_{i0}, \dots, y_{i,T-1})'$ and $\text{Var}(\mathbf{y}_{i,-1}) = \mathbf{V}_T = \sigma_\mu^2 \boldsymbol{\nu}_T \boldsymbol{\nu}_T' + \sigma_v^2 \boldsymbol{\Phi}_T$ where $\sigma_\mu^2 = \sigma_\eta^2 / (1 - \alpha)^2$ and $\boldsymbol{\Phi}_T = \{\phi_{st}\} = \alpha^{|s-t|} / (1 - \alpha^2)$. Also, let \mathbf{f}_t' be the t th row of \mathbf{F}_T^τ , $\mathbf{L}_{(s:t)}$ be the s th to t th rows of \mathbf{I}_T and $\mathbf{B}_{(s:t)}$ be the s th to t th rows of \mathbf{B}_T^τ . Then, we obtain $y_{i,t-1}^* = \mathbf{f}_t' \mathbf{y}_{i,-1}$ and

$$\mathbf{z}_{it}^L = (y_{i,t-\ell}, \dots, y_{i,t-1})' = \mathbf{L}_{(t-\ell+1:t)} \mathbf{y}_{i,-1}, \quad \mathbf{z}_{it}^B = (y_{i,t-\ell}^*, \dots, y_{i,t-1}^*)' = \mathbf{B}_{(t-\ell+1:t)} \mathbf{y}_{i,-1}.$$

Using this, we have

$$\begin{aligned}
E\left(\mathbf{z}_{it}^{L(\ell)} y_{i,t-1}^*\right) &= \mathbf{L}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{f}_t', & E\left(\mathbf{z}_{it}^{B(\ell)} y_{i,t-1}^*\right) &= \mathbf{B}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{f}_t', \\
E\left(\mathbf{z}_{it}^{L(\ell)} \mathbf{z}_{it}^{L'(\ell)}\right) &= \mathbf{L}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{L}_{(t-\ell+1:t)}', & E\left(\mathbf{z}_{it}^{B(\ell)} \mathbf{z}_{it}^{B'(\ell)}\right) &= \mathbf{B}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{B}_{(t-\ell+1:t)}'.
\end{aligned}$$

Hence, (36) and (37) can be written as

$$\begin{aligned}
Avar(\widehat{\alpha}_{GMM}^L) &= \sigma_v^2 \left[\sum_{t=1}^{T-1} \mathbf{f}_t' \mathbf{V}_T \mathbf{L}_{(t-\ell+1:t)}' \left[\mathbf{L}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{L}_{(t-\ell+1:t)}' \right]^{-1} \mathbf{L}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{f}_t \right]^{-1} \quad (58) \\
Avar(\widehat{\alpha}_{GMM}^B) &= \sigma_v^2 \left[\sum_{t=2}^{T-1} \mathbf{f}_t' \mathbf{V}_T \mathbf{B}_{(t-\ell+1:t)}' \left[\mathbf{B}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{B}_{(t-\ell+1:t)}' \right]^{-1} \mathbf{B}_{(t-\ell+1:t)} \mathbf{V}_T \mathbf{f}_t \right]^{-1} \quad (59)
\end{aligned}$$

Figure 1 is described based on (58) and (59) numerically without deriving the explicit form of expectations such as $E\left(\mathbf{z}_{it}^{L(\ell)} y_{i,t-1}^*\right)$.

Proof of Theorem 1

We derive the explicit formula for asymptotic variances of IV and GMM estimators. First, consider $Avar(\widehat{\alpha}_{GMM}^{L(t)})$. Note that under Assumption 1, $y_{i,t-1}$ can be written as

$$y_{i,t-1} = \frac{\eta_i}{1 - \alpha} + \xi_{i,t-1} \quad (60)$$

where $\xi_{i,t-1} = \sum_{j=0}^{\infty} \alpha^j v_{i,t-1-j}$. Also, from (A43) of Alvarez and Arellano (2003), we have

$$y_{i,t-1}^* = \psi_t \xi_{i,t-1} - c_t \left(\frac{\phi_{T-t} v_{it} + \dots + \phi_1 v_{i,T-1}}{T - t} \right). \quad (61)$$

where ϕ is defined in (45). Using (60) and (61), and under Assumption 1, we have

$$E(\mathbf{z}_{it}^L y_{i,t-1}^*) = \psi_t \left(\frac{\sigma_v^2}{1 - \alpha^2} \right) (\alpha^{\ell-1}, \dots, 1)', \quad (62)$$

$$[E(\mathbf{z}_{it}^L \mathbf{z}_{it}^{L'})]^{-1} = \mathbf{V}_\ell^{-1} = \frac{1}{\sigma_v^2} \left[(\sqrt{\lambda} \boldsymbol{\nu}_\ell) (\sqrt{\lambda} \boldsymbol{\nu}_\ell)' + \boldsymbol{\Phi}_\ell \right]^{-1} \quad (63)$$

where $\lambda = \sigma_\mu^2/\sigma_v^2 = r/(1 - \alpha)^2$, $\boldsymbol{\iota}_\ell$ is an ℓ dimensional column vector of ones, and \mathbf{V}_ℓ is the upper-left $\ell \times \ell$ matrix of \mathbf{V}_T . The explicit expression of (63) is obtained as follows. By using the Sherman-Morrison-Woodbury inversion formula

$$[\mathbf{A} + \mathbf{b}\mathbf{b}']^{-1} = \mathbf{A}^{-1} - \left[\frac{1}{1 + \mathbf{b}'\mathbf{A}^{-1}\mathbf{b}} \right] \mathbf{A}^{-1}\mathbf{b}\mathbf{b}'\mathbf{A}^{-1}$$

and the decomposition of \mathbf{V}_ℓ^{-1} ¹³

$$\mathbf{V}_\ell^{-1} = \mathbf{C}'\mathbf{C}$$

where

$$\mathbf{C} = \begin{bmatrix} \sqrt{1 - \alpha^2} & 0 & 0 & \cdots & 0 & 0 \\ -\alpha & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\alpha & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\alpha & 1 \end{bmatrix}$$

we obtain

$$[E(\mathbf{z}_{it}^L \mathbf{z}_{it}^{L'})]^{-1} = \sigma_v^{-2} \left[\mathbf{C}'\mathbf{C} - \frac{\lambda}{1 + \lambda \boldsymbol{\iota}_\ell' \mathbf{C}' \mathbf{C} \boldsymbol{\iota}_\ell} \mathbf{C}' \mathbf{C} \boldsymbol{\iota}_\ell \boldsymbol{\iota}_\ell' \mathbf{C}' \mathbf{C} \right]. \quad (64)$$

By substituting (62) and (64) into (36), we obtain (38). The results (39) and (40) are obtained from (38). Next, we consider $Var(\hat{\alpha}_{GMM}^{B(1)})$. First, note that $y_{i,t-1}^{**}$ can be written as

$$y_{i,t-1}^{**} = c_{T-t+1}^L \left[\xi_{i,t-1} - \frac{\xi_{i,t-2} + \cdots + \xi_{i0}}{t-1} \right]. \quad (65)$$

Then, using (61) and (65), we obtain

$$E(y_{i,t-1}^{**} y_{i,t-1}^*) = \left(\frac{\sigma_v^2}{1 - \alpha^2} \right) \psi_t c_{T-t+1}^L \left(1 - \frac{\phi_{t-1}}{t-1} \right). \quad (66)$$

Also, from (60), we obtain

$$\begin{aligned} E[(y_{i,t-1}^{**})^2] &= c_{T-t+1}^{L2} E \left[\xi_{i,t-1} - \frac{1}{t-1} (\xi_{i,0} + \cdots + \xi_{i,t-2}) \right]^2 \\ &= c_{T-t+1}^{L2} \left[\frac{\sigma_v^2}{1 - \alpha^2} \left(1 - \frac{2\alpha\phi_{t-1}}{t-1} \right) + \frac{1}{(t-1)^2} E(\xi_{i0} + \cdots + \xi_{i,t-1})^2 \right]. \end{aligned} \quad (67)$$

Using the result of (A8) in Alvarez and Arellano (2003), we have

$$E(\xi_{i0} + \cdots + \xi_{i,t-1})^2 = \frac{\sigma_v^2}{1 - \alpha^2} \left[\frac{(t-1)(1 + \alpha)}{1 - \alpha} - \frac{2\alpha(1 - \alpha^{t-1})}{(1 - \alpha)^2} \right]. \quad (68)$$

By substituting this into (67), we get

$$E[(y_{i,t-1}^{**})^2] = \left(\frac{\sigma_v^2}{1 - \alpha^2} \right) c_{T-t+1}^{L2} A_t \quad (69)$$

where A_t is defined in (46). The result (42) is obtained by substituting (66) and (69) into (36).

¹³See Amemiya (1985, p.164), Hamilton (1994, p.120) and Greene (2001, p.822).

Next, we derive the asymptotic variances of $\widehat{\alpha}_{IV}^L$ and $\widehat{\alpha}_{IV}^B$. Using (60) and (61) and the fact that v_{it}^* is serially uncorrelated, we obtain

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{T-1} E(y_{i,t-1} y_{i,t-1}^*) &= \left(\frac{\sigma_v^2}{1-\alpha^2} \right) \sum_{t=1}^{T-1} \psi_t, \\ \text{Var} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \sum_{t=1}^{T-1} y_{i,t-1} v_{it}^* \right) &= \text{Var} \left(\sum_{t=1}^{T-1} y_{i,t-1} v_{it}^* \right) = \sigma_v^2 \sum_{t=1}^{T-1} E(y_{i,t-1}^2) = \sigma_v^2 \sigma_\mu^2 + \frac{\sigma_v^4}{1-\alpha^2}. \end{aligned}$$

Using these, we obtain the asymptotic variance of $\widehat{\alpha}_{IV}^L$ as in (41). The asymptotic variance of $\widehat{\alpha}_{IV}^B$ can be derived similarly. Using (66) and (69), and under Assumption 1, we obtain

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sum_{t=2}^{T-1} E(y_{i,t-1}^{**} y_{i,t-1}^*) &= \left(\frac{\sigma_v^2}{1-\alpha^2} \right) \sum_{t=2}^{T-1} \psi_t c_{T-t+1}^t \left(1 - \frac{\phi_{t-1}}{t-1} \right), \\ \text{Var} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \sum_{t=2}^{T-1} y_{i,t-1}^{**} v_{it}^* \right) &= \text{Var} \left(\sum_{t=2}^{T-1} y_{i,t-1}^{**} v_{it}^* \right) = \sigma_v^2 \sum_{t=2}^{T-1} E[(y_{i,t-1}^{**})^2] = \left(\frac{\sigma_v^4}{1-\alpha^2} \right) \sum_{t=2}^{T-1} c_{T-t+1}^{2t} A_t. \end{aligned}$$

From these, the asymptotic variance of $\widehat{\alpha}_{IV}^L$ is obtained as (43).

Proof of Lemma 1

First, we decompose $T_d \times T$ matrix \mathbf{F}_T as

$$\mathbf{F}_T = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} & \mathbf{F}_{13} \\ \mathbf{0}_{T_{2d} \times d} & \mathbf{F}_{22} & \mathbf{F}_{23} \end{bmatrix} = \{f_{st}\}, (s = 1, \dots, T_d; t = 1, \dots, T)$$

where \mathbf{F}_{11} is $d \times d$, \mathbf{F}_{12} is $d \times T_{2d}$, \mathbf{F}_{13} is $d \times d$, \mathbf{F}_{22} is $T_{2d} \times T_{2d}$, and \mathbf{F}_{23} is $T_{2d} \times d$. Note that \mathbf{B}_T and \mathbf{F}_T have the following relationship

$$\mathbf{B}_T = \mathcal{I}_{T_d} \mathbf{F}_T \mathcal{I}_T \tag{70}$$

where

$$\mathcal{I}_T = \begin{bmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{bmatrix}$$

and $\mathcal{I}_T^2 = \mathcal{I}_T' \mathcal{I}_T = \mathbf{I}_T$. Furthermore, using (48), \mathbf{W}_i can be written as

$$\mathbf{W}_i = \nu_T \boldsymbol{\mu}'_i + \tau_T \boldsymbol{\kappa}'_i + \boldsymbol{\Xi}_i = \mathbf{C}_T \boldsymbol{\Psi}_i + \boldsymbol{\Xi}_i$$

where $\boldsymbol{\Xi}_i = (\boldsymbol{\xi}'_{i1}, \dots, \boldsymbol{\xi}'_{iT})'$ and $\boldsymbol{\Psi}_i = (\boldsymbol{\mu}_i, \boldsymbol{\kappa}_i)'$.

Proof of (a): Note the following decomposition:

$$\frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{W}_i = \frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{W}_i + \frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i (\mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T - \mathbf{Q}_T) \mathbf{W}_i. \tag{71}$$

Using $\mathbf{F}_{T_d} \mathbf{L}_T \mathbf{C}_T = \mathbf{B}_{T_d} \mathbf{K}_T \mathbf{C}_T = \mathbf{Q}_T \mathbf{C}_T = \mathbf{0}$ and (19), the second term of (71) can be further decomposed as

$$\frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i (\mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T - \mathbf{Q}_T) \mathbf{W}_i = \frac{1}{NT} \sum_{i=1}^N \boldsymbol{\Xi}'_i (\mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T - \mathbf{I}_T) \boldsymbol{\Xi}_i + \frac{1}{NT} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{R}_T \boldsymbol{\Xi}_i. \quad (72)$$

To consider the first term of right-hand side of (72), we derive the explicit form of $\mathbf{A}_T = \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T - \mathbf{I}_T$. Using (70), $\mathbf{F}_{T_d} = \mathbf{L}_{T_d} \mathbf{F}_T \mathbf{L}'_T$ and

$$\begin{aligned} \mathbf{K}'_T \mathcal{I}_{T_d} \mathbf{L}_T &= \begin{bmatrix} \mathbf{0}_{d \times d} & \mathbf{0}_{d \times T_{2d}} & \mathcal{I}_d \\ \mathbf{0}_{T_{2d} \times d} & \mathcal{I}_{T_{2d}} & \mathbf{0}_{T_{2d} \times d} \\ \mathbf{0}_{d \times d} & \mathbf{0}_{d \times T_{2d}} & \mathbf{0}_{d \times d} \end{bmatrix}, & \mathbf{L}'_{T_d} \mathcal{I}_{T_{2d}} \mathbf{L}_{T_d} &= \begin{bmatrix} \mathbf{0}_{d \times d} & \mathbf{0}_{d \times T_{2d}} \\ \mathbf{0}_{T_{2d} \times d} & \mathcal{I}_{T_{2d}} \end{bmatrix}, \\ \mathbf{L}'_T \mathbf{L}_T &= \begin{bmatrix} \mathbf{0}_{d \times d} & \mathbf{0}_{d \times T_{2d}} & \mathbf{0}_{d \times d} \\ \mathbf{0}_{T_{2d} \times d} & \mathbf{I}_{T_{2d}} & \mathbf{0}_{T_{2d} \times d} \\ \mathbf{0}_{d \times d} & \mathbf{0}_{d \times T_{2d}} & \mathbf{I}_d \end{bmatrix}, \end{aligned}$$

we have

$$\begin{aligned} \mathbf{A}_T &= \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{L}_{T_d} \mathbf{F}_T \mathbf{L}'_T \mathbf{L}_T - \mathbf{I}_T = \mathbf{K}'_T \mathcal{I}_{T_d} \mathbf{F}'_{T_d} \mathcal{I}_{T_{2d}} \mathbf{L}_{T_d} \mathbf{F}_T \mathbf{L}'_T \mathbf{L}_T - \mathbf{I}_T \\ &= (\mathbf{K}'_T \mathcal{I}_{T_d} \mathbf{L}_T) \mathbf{F}'_T (\mathbf{L}'_{T_d} \mathcal{I}_{T_{2d}} \mathbf{L}_{T_d}) \mathbf{F}_T (\mathbf{L}'_T \mathbf{L}_T) - \mathbf{I}_T \\ &= \begin{bmatrix} -\mathbf{I}_d & \mathcal{I}_d \mathbf{F}'_{23} \mathcal{I}_{T_{2d}} \mathbf{F}_{22} & \mathcal{I}_d \mathbf{F}'_{23} \mathcal{I}_{T_{2d}} \mathbf{F}_{23} \\ \mathbf{0}_{T_{2d} \times d} & \mathcal{I}_{T_{2d}} \mathbf{F}'_{22} \mathcal{I}_{T_{2d}} \mathbf{F}_{22} - \mathbf{I}_{T_{2d}} & \mathcal{I}_{T_{2d}} \mathbf{F}'_{22} \mathcal{I}_{T_{2d}} \mathbf{F}_{23} \\ \mathbf{0}_{d \times d} & \mathbf{0}_{d \times T_{2d}} & -\mathbf{I}_d \end{bmatrix} \\ &= \{\mathbf{A}_{ij}\}, (i, j = 1, 2, 3). \end{aligned} \quad (73)$$

Next, we derive the form of each \mathbf{A}_{ij} . Using

$$\begin{aligned} \mathcal{I}_{T_{2d}} \mathbf{F}_{22} &= \begin{bmatrix} 0 & \cdots & 0 & f_{T_d T_d} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & f_{d+2, d+2} & \cdots & f_{d+2, T_d} \\ f_{d+1, d+1} & f_{d+1, d+2} & \cdots & f_{d+1, T_d} \end{bmatrix}, & \mathcal{I}_{T_{2d}} \mathbf{F}'_{22} &= \begin{bmatrix} f_{d+1, T_d} & f_{d+2, T_d} & \cdots & f_{T_d T_d} \\ \vdots & \vdots & \ddots & 0 \\ f_{d+1, d+2} & f_{d+2, d+2} & & \vdots \\ f_{d+1, d+1} & 0 & \cdots & 0 \end{bmatrix}, \\ \mathcal{I}_{T_{2d}} \mathbf{F}_{23} &= \begin{bmatrix} f_{T_d, T_d+1} & f_{T_d, T} \\ \vdots & \vdots \\ f_{d+2, T_d+1} & f_{d+2, T} \\ f_{d+1, T_d+1} & f_{d+1, T} \end{bmatrix}, & \mathcal{I}_d \mathbf{F}'_{23} &= \begin{bmatrix} f_{d+1, T} & f_{d+2, T} & \cdots & f_{T_d, T} \\ f_{d+1, T_d+1} & f_{d+2, T_d+1} & \cdots & f_{T_d, T_d+1} \end{bmatrix} \end{aligned}$$

we have

$$\begin{aligned} \mathbf{A}_{12} &= \mathcal{I}_d \mathbf{F}'_{23} \mathcal{I}_{T_{2d}} \mathbf{F}_{22} \\ &= \begin{bmatrix} f_{T_d, T} f_{d+1, d+1} & f_{T_d-1, T} f_{d+2, d+2} + f_{T_d, T} f_{d+1, d+2} & \cdots & \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T} f_{T_d-\ell+1, T_d} \\ f_{T_d, T_d+1} f_{d+1, d+1} & f_{T_d-1, T_d+1} f_{d+2, d+2} + f_{T_d, T_d+1} f_{d+1, d+2} & \cdots & \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T_d+1} f_{T_d-\ell+1, T_d} \end{bmatrix} \\ &= \left\{ a_{12}^{jk} \right\} = \sum_{\ell=1}^k f_{T_d-\ell+1, T-j+1} f_{d+\ell, d+k} = \sum_{\ell=1}^{k-1} f_{T_d-\ell+1, T-j+1} f_{d+\ell, d+k} + f_{T_d-k+1, T-j+1} f_{d+k, d+k} \end{aligned}$$

$$\begin{aligned}
&= \begin{cases} \sum_{\ell=1}^{k-1} \frac{c_{T-\ell}^c c_{\ell+1}^c}{\ell(T-\ell-1)} - \frac{c_{T-k}^c c_{k+1}^c}{k} < \sum_{\ell=1}^{T_3} \frac{c_{T-\ell}^c c_{\ell+1}^c}{\ell(T-\ell-1)} & \text{for FE model} \\ \sum_{\ell=1}^{k-1} \frac{4c_{\ell+2}^r c_{T-\ell-1}^r (3j-\ell-3)(2T-3k+\ell-3)}{\ell(\ell+1)(T-\ell-2)(T-\ell-3)} - \frac{2c_{k+2}^r c_{T-k-1}^r (3j-k-3)}{k(k+1)} & \text{for trend model} \end{cases} \\
&\quad (j = 1; k = 1, \dots, T_{2d}) \\
&\quad (j = 1, 2; k = 1, \dots, T_{2d}) \\
&= O\left(\frac{\log T}{T}\right),
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{13} &= \mathcal{I}_d \mathbf{F}'_{23} \mathcal{I}_{T_{2d}} \mathbf{F}_{23} = \begin{bmatrix} \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T} f_{T_d-\ell+1, T_d+1} & \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T} f_{T_d-\ell+1, T} \\ \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T_d+1} f_{T_d-\ell+1, T_d+1} & \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T_d+1} f_{T_d-\ell+1, T} \end{bmatrix} \\
&= \left\{ a_{13}^{jk} \right\} = \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T-j+1} f_{T_d-\ell+1, T_d+k} \\
&= \begin{cases} \sum_{\ell=1}^{T-2} \frac{c_{\ell+1}^c c_{T-\ell}^c}{(T-\ell-1)\ell}, & (j = 1, k = 1) & \text{for FE model} \\ \sum_{\ell=1}^{T_{2d}} \frac{4c_{\ell+2}^r c_{T-\ell-1}^r (3k+\ell-6)(T-3j-\ell)}{\ell(\ell+1)(T-\ell-2)(T-\ell-3)}, & (j = 1, 2; k = 1, 2) & \text{for trend model} \end{cases} \\
&= O\left(\frac{\log T}{T}\right),
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{22} &= \mathcal{I}_{T_{2d}} \mathbf{F}'_{22} \mathcal{I}_{T_{2d}} \mathbf{F}_{22} - \mathbf{I}_{T_{2d}} = \begin{bmatrix} a_{22}^{11} & a_{22}^{12} & \cdots & a_{22}^{1, T_{2d}} \\ 0 & a_{22}^{22} & \cdots & a_{22}^{2, T_{2d}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{22}^{T_{2d}, T_{2d}} \end{bmatrix} \\
&= \left\{ a_{22}^{jk} \right\} = \begin{cases} 0 & \text{if } j > k \\ f_{T_d-j+1, T_d-j+1} f_{d+j, d+j} - 1 & \text{if } j = k \\ \sum_{\ell=1}^k f_{T_d-\ell+1, T_d-j+1} f_{d+\ell, d+k} & \text{if } j < k \end{cases} \\
&= \begin{cases} \begin{cases} 0 & \text{if } j > k \\ c_{T-j}^c c_{j+1}^c - 1 = O\left(\frac{1}{j}\right) + O\left(\frac{1}{(T-j)}\right) + O\left(\frac{1}{j^2} + \frac{1}{(T-j)^2}\right) & \text{if } j = k \\ \sum_{\ell=1}^{k-j+1} \frac{c_{\ell+1}^c c_{T-\ell}^c}{(T-\ell-1)\ell} = O\left(\frac{\log T}{T}\right) & \text{if } j < k \end{cases} & \text{for FE model} \\ \begin{cases} 0 & \text{if } j > k \\ c_{T-j-1}^r c_{j+2}^r - 1 = O\left(\frac{1}{j}\right) + O\left(\frac{1}{(T-j)}\right) & \text{if } j = k \\ \sum_{\ell=1, \ell > j, \ell < k}^k \frac{4c_{\ell+2}^r c_{T-\ell-1}^r (3j-\ell+3)(2T-3k+\ell-3)}{\ell(\ell+1)(T-\ell-2)(T-\ell-3)} \\ - \frac{2c_{j+2}^r c_{T-j-1}^r (2T+j-3k-3)}{(T-j-2)(T-j-3)} - \frac{2c_{k+2}^r c_{T-k-1}^r (3j-k+3)}{k(k+1)} = O\left(\frac{\log T}{T}\right) & \text{if } j < k \end{cases} & \text{for trend model} \end{cases} \\
&\quad (j, k = 1, \dots, T_{2d})
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{23} &= \mathcal{I}_{T_{2d}} \mathbf{F}'_{22} \mathcal{I}_{T_{2d}} \mathbf{F}_{23} = \begin{bmatrix} \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T_d} f_{T_d-\ell+1, T_d+1} & \sum_{\ell=1}^{T_{2d}} f_{d+\ell, T_d} f_{T_d-\ell+1, T} \\ \vdots & \vdots \\ \sum_{\ell=1}^2 f_{d+\ell, d+2} f_{T_d-\ell+1, T_d+1} & \sum_{\ell=1}^2 f_{d+\ell, d+2} f_{T_d-\ell+1, T} \\ \sum_{\ell=1}^1 f_{d+\ell, d+1} f_{T_d-\ell+1, T_d+1} & \sum_{\ell=1}^1 f_{d+\ell, d+1} f_{T_d-\ell+1, T} \end{bmatrix} \\
&= \left\{ a_{23}^{jk} \right\} = \sum_{\ell=1}^{T_{2d}-j+1} f_{d+\ell, T_d-j+1} f_{T_d-\ell+1, T_d+k}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell=1}^{T_{2d-j}} f_{d+\ell, T_d-j+1} f_{T_d-\ell+1, T_d+k} + f_{T_d-j+1, T_d-j+1} f_{d+j, T_d+k} \\
&= \begin{cases} \sum_{\ell=1}^{T_{2d-j}} \frac{c_{\ell+1}^{\ell} c_{T-\ell}^{\ell} - c_{T_d-j+1}^{\ell} c_{d+j}^{\ell}}{(T-\ell-1)\ell - (T_d-j)\ell}, & (j=1, \dots, T_{2d}; k=1) & \text{for FE model} \\ \sum_{\ell=1}^{T_{2d-j}} \frac{4c_{\ell+2}^{\ell} c_{T-\ell-1}^{\ell} (3k+\ell-6)(T-3j-\ell-6)}{\ell(\ell+1)(T-\ell-2)(T-\ell-3)} + \frac{2c_{j+2}^{\ell} c_{T-j-1}^{\ell} (T-j+3k-9)}{(T-j-2)(T-j-3)} & & \text{for trend model} \end{cases} \\
&= O\left(\frac{\log T}{T}\right).
\end{aligned}$$

We now assess the first term of (72). Using $\Xi_i = (\Xi_{1i}, \Xi_{2i}, \Xi_{3i})'$ where Ξ_{1i} is $d \times k$, Ξ_{2i} is $T_{2d} \times k$, and Ξ_{3i} is $d \times k$, we have

$$\mathbf{S}_i = \Xi_i' \mathbf{A}_T \Xi_i = \mathbf{S}_{1i} + \mathbf{S}_{2i} + \mathbf{S}_{3i} + \mathbf{S}_{4i} + \mathbf{S}_{5i} + \mathbf{S}_{6i}$$

where

$$\begin{aligned}
\mathbf{S}_{1i} &= -\Xi_{1i}' \Xi_{1i}, & \mathbf{S}_{2i} &= \Xi_{1i}' \mathbf{A}_{12} \Xi_{2i}, & \mathbf{S}_{3i} &= \Xi_{1i}' \mathbf{A}_{13} \Xi_{3i}, & \mathbf{S}_{4i} &= \Xi_{2i}' \mathbf{A}_{22} \Xi_{2i}, \\
\mathbf{S}_{5i} &= \Xi_{2i}' \mathbf{A}_{23} \Xi_{3i}, & \mathbf{S}_{6i} &= -\Xi_{3i}' \Xi_{3i}.
\end{aligned}$$

We now evaluate each term. We consider the FE model and trend model separately below.

FE model From the definition of Ξ_{1i} and Assumption 3, we have

$$E(\mathbf{S}_{1i}) = -E(\xi_{i1} \xi_{i1}') = -\Gamma_{i0} = O(1).$$

Using $a_{12}^{1,t-1} = O\left(\frac{\log T}{T}\right)$ for all t , $a_{13}^{11} = O\left(\frac{\log T}{T}\right)$ and Assumption 3, we have

$$E(\mathbf{S}_{2i}) = \sum_{t=2}^{T_d} a_{12}^{1,t-1} E(\xi_{i1} \xi_{it}') = O\left(\frac{\log T}{T}\right) \sum_{t=2}^{T_1} \Gamma_{i,t-1} = O\left(\frac{\log T}{T}\right),$$

$$E(\mathbf{S}_{3i}) = a_{13}^{1,1} E(\xi_{i1} \xi_{iT}') = O\left(\frac{\log T}{T}\right) \Gamma_{i,T-1} = O\left(\frac{\log T}{T}\right).$$

Similarly, using $a_{22}^{t-1,t-1} = O(1/(t+1)) + O(1/(T-t))$ and $a_{22}^{s-1,t-1} = O(\log T/T)$ for all $s \neq t$, we have

$$\begin{aligned}
E(\mathbf{S}_{4i}) &= \sum_{s=2}^{T_d} \sum_{t=2}^{T_d} a_{22}^{s-1,t-1} E(\xi_{is} \xi_{it}') = \sum_{t=2}^{T_d} a_{22}^{t-1,t-1} E(\xi_{it} \xi_{it}') + \sum_{s=2}^{T_d-1} \sum_{t=s+1}^{T_d} a_{22}^{s-1,t-1} E(\xi_{is} \xi_{it}') \\
&= \sum_{t=2}^{T_d} \left[O\left(\frac{1}{t+1}\right) + O\left(\frac{1}{T-t}\right) \right] E(\xi_{it} \xi_{it}') + O\left(\frac{\log T}{T}\right) \sum_{s=2}^{T_d-1} \sum_{t=s+1}^{T_d} \Gamma_{t-s,i} \\
&= O(\log T).
\end{aligned}$$

Finally, using $a_{23}^{t-1,1} = O(\log T/T)$ for all t , and the definition of Ξ_{3i} , we have

$$E(\mathbf{S}_{5i}) = \sum_{t=2}^{T_d} a_{23}^{t-1,1} E(\xi_{it} \xi_{iT}') = \sum_{t=2}^{T_d} O\left(\frac{\log T}{T}\right) \Gamma_{i,T-t} = O\left(\frac{\log T}{T}\right),$$

$$E(\mathbf{S}_{6i}) = -E(\xi_{iT} \xi_{iT}') = -\Gamma_{i0} = O(1).$$

Thus, for the FE model, we have $\mathbf{S}_i = \sum_{l=1}^6 \mathbf{S}_{li} = O(\log T)$ for all i and obtain

$$\frac{1}{NT} \sum_{i=1}^N \Xi_i' \mathbf{A}_T \Xi_i = O_p\left(\frac{\log T}{T}\right). \quad (74)$$

Trend model From the definition of Ξ_{1i} , we have

$$E(\mathbf{S}_{1i}) = -E(\xi_{i1}\xi'_{i1}) - E(\xi_{i2}\xi'_{i2}) = -2\Gamma_{i0} = O(1).$$

Since $a_{12}^{1,t-1}$, $a_{12}^{2,t-1}$, a_{13}^{1k} and a_{13}^{2k} are $O\left(\frac{\log T}{T}\right)$ for all t and k , using Assumption 3, we have

$$\begin{aligned} E(\mathbf{S}_{2i}) &= \sum_{t=2}^{T_d} \left[a_{12}^{1,t-1} E(\xi_{i1}\xi'_{it}) + a_{12}^{2,t-1} E(\xi_{i2}\xi'_{it}) \right] \\ &= O\left(\frac{\log T}{T}\right) \sum_{t=2}^{T_1} \Gamma_{i,t-1} + O\left(\frac{\log T}{T}\right) \sum_{t=2}^{T_1} \Gamma_{i,t-2} = O\left(\frac{\log T}{T}\right), \end{aligned}$$

$$\begin{aligned} E(\mathbf{S}_{3i}) &= \sum_{t=T-1}^T \left[\left(a_{13}^{1,1} + a_{13}^{1,2} \right) E(\xi_{i1}\xi'_{it}) + \left(a_{13}^{2,1} + a_{13}^{2,2} \right) E(\xi_{i2}\xi'_{it}) \right] \\ &= O\left(\frac{\log T}{T}\right) (\Gamma_{i,T-1} + 2\Gamma_{i,T-2} + \Gamma_{i,T-3}) = O\left(\frac{\log T}{T}\right). \end{aligned}$$

Similarly, using $a_{22}^{t-1,t-1} = O(1/(t+1)) + O(1/(T-t))$ and $a_{22}^{s-1,t-1} = O(\log T/T)$ for all $s \neq t$, we have

$$\begin{aligned} E(\mathbf{S}_{4i}) &= \sum_{s=3}^{T_d} \sum_{t=3}^{T_d} a_{22}^{s-1,t-1} E(\xi_{is}\xi'_{it}) = \sum_{t=3}^{T_d} a_{22}^{t-1,t-1} E(\xi_{it}\xi'_{it}) + \sum_{s=3}^{T_d-1} \sum_{t=s+1}^{T_d} a_{22}^{s-1,t-1} E(\xi_{is}\xi'_{it}) \\ &= \sum_{t=3}^{T_d} \left[O\left(\frac{1}{t+1}\right) + O\left(\frac{1}{T-t}\right) \right] E(\xi_{it}\xi'_{it}) + O\left(\frac{\log T}{T}\right) \sum_{s=3}^{T_d-1} \sum_{t=s+1}^{T_d} \Gamma_{t-s,i} \\ &= O(\log T). \end{aligned}$$

Finally, using $a_{23}^{t-1,1} = O(\log T/T)$ and $a_{23}^{t-1,2} = O(\log T/T)$ for all t , and the definition of Ξ_{3i} , we have

$$\begin{aligned} E(\mathbf{S}_{5i}) &= \sum_{t=2}^{T_d} a_{23}^{t-1,1} E(\xi_{it}\xi'_{iT_1}) + \sum_{t=2}^{T_d} a_{23}^{t-1,2} E(\xi_{it}\xi'_{iT}) \\ &= \sum_{t=2}^{T_d} O\left(\frac{\log T}{T}\right) (\Gamma_{i,T-t-1} + \Gamma_{i,T-t}) = O\left(\frac{\log T}{T}\right), \end{aligned}$$

$$E(\mathbf{S}_{6i}) = -E(\xi_{iT}\xi'_{iT}) - E(\xi_{iT_1}\xi'_{iT_1}) = -2\Gamma_{i0} = O(1).$$

Thus, for the trend model, we have $\mathbf{S}_i = \sum_{l=1}^6 \mathbf{S}_{li} = O(\log T)$ for all i and obtain

$$\frac{1}{NT} \sum_{i=1}^N \Xi'_i \mathbf{A}_T \Xi_i = O_p\left(\frac{\log T}{T}\right). \quad (75)$$

Next, we consider the second term of (72). Let us define $\mathbf{H}_i = \Xi'_i \mathbf{R}_T \Xi_i$. Then, for the FE model, using (20), we have

$$\begin{aligned} E(\mathbf{H}_i) &= \frac{1}{T} E(\Xi'_i \mathbf{u}_T \mathbf{u}'_T \Xi_i) = \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T E(\xi_{it}\xi'_{is}) \\ &= \Gamma_{i0} + \frac{1}{T} \sum_{s=1}^{T-1} \sum_{t=s+1}^T (\Gamma_{i,t-s} + \Gamma'_{i,t-s}) = O(1) \end{aligned} \quad (76)$$

where we used $\|\sum_{t=s+1}^T \mathbf{\Gamma}_{i,t-s}\| = \|\sum_{l=1}^{T-s} \mathbf{\Gamma}_{i,l}\| \leq \sum_{l=1}^{T-s} \|\mathbf{\Gamma}_{i,l}\| < \sum_{l=1}^{\infty} \|\mathbf{\Gamma}_{i,l}\| < \infty$. For the trend model, using (20), we have

$$\begin{aligned} \mathbf{H}_i &= \frac{2(2T+1) \mathbf{\Xi}'_i \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \mathbf{\Xi}_i}{T(T-1)} + \frac{12 \mathbf{\Xi}'_i \boldsymbol{\tau}_T \boldsymbol{\tau}'_T \mathbf{\Xi}_i}{T(T-1)(T+1)} - \frac{6(\mathbf{\Xi}'_i \boldsymbol{\nu}_T \boldsymbol{\tau}'_T \mathbf{\Xi}_i + \mathbf{\Xi}'_i \boldsymbol{\tau}_T \boldsymbol{\nu}'_T \mathbf{\Xi}_i)}{T(T-1)} \\ &= \mathbf{H}_{1i} + \mathbf{H}_{2i} + \mathbf{H}_{3i}. \end{aligned}$$

Using Assumption 3 and (76), we have

$$\begin{aligned} E(\mathbf{H}_{1i}) &= \frac{2(2T+1)}{T(T-1)} E \left[\left(\sum_{t=1}^T \boldsymbol{\xi}_{it} \right) \left(\sum_{s=1}^T \boldsymbol{\xi}'_{is} \right) \right] = O(1), \\ E(\mathbf{H}_{2i}) &= \frac{12}{T(T-1)(T+1)} E \left(\left(\sum_{t=1}^T t \boldsymbol{\xi}_{it} \right) \left(\sum_{s=1}^T s \boldsymbol{\xi}'_{is} \right) \right) \\ &= \frac{12T}{(T-1)(T+1)} \sum_{s=1}^T \sum_{t=1}^T \left(\frac{t}{T} \right) \left(\frac{s}{T} \right) E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{is}) = O(1), \\ E(\mathbf{H}_{3i}) &= \frac{6}{T(T-1)} \left(\sum_{t=1}^T \sum_{s=1}^T s E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{is}) + \sum_{t=1}^T \sum_{s=1}^T t E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{is}) \right) \\ &= \frac{6}{(T-1)} \left(\sum_{s=1}^T \sum_{t=1}^T \left(\frac{s}{T} \right) E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{is}) + \sum_{s=1}^T \sum_{t=1}^T \left(\frac{t}{T} \right) E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{is}) \right) = O(1) \end{aligned}$$

where we used $0 < t/T \leq 1$ and $0 < s/T \leq 1$ for all s and t . Hence, for each i , we have $E(\mathbf{H}_i) = O(1)$ for both FE and trend models, and we obtain

$$\frac{1}{NT} \sum_{i=1}^N \mathbf{\Xi}'_i \mathbf{R}_T \mathbf{\Xi}_i = O_p \left(\frac{1}{T} \right). \quad (77)$$

By combining (74), (75), and (77), we obtain

$$\frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{W}_i = \frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{W}_i + O_p \left(\frac{\log T}{T} \right).$$

Proof of (b): Using $\mathbf{Q}_T(\boldsymbol{\nu}_T, \boldsymbol{\tau}_T) = \mathbf{0}$, we have

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{v}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{\Xi}'_i \mathbf{Q}_T \mathbf{v}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{\Xi}'_i \mathbf{v}_i - \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{\Xi}'_i \mathbf{R}_T \mathbf{v}_i.$$

The first term converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ by Assumption 4. To assess the second term, let us define $\mathbf{h}_i = -\mathbf{\Xi}'_i \mathbf{R}_T \mathbf{v}_i$. Then, for the case of FE model, using Assumption 3, we have

$$\begin{aligned} E(\mathbf{h}_i) &= \frac{-1}{T} E \left[\left(\sum_{t=1}^T \boldsymbol{\xi}_{it} \right) \left(\sum_{s=1}^T v_{is} \right) \right] = \frac{-1}{T} \sum_{s=1}^{T-1} \sum_{t=s+1}^T E(\boldsymbol{\xi}_{it} v_{is}) + \frac{-1}{T} \sum_{s,t=1, t \leq s}^T E(\boldsymbol{\xi}_{it} v_{is}) \\ &= \frac{-1}{T} \sum_{s=1}^{T-1} \sum_{t=s+1}^T \phi_{i,t-s} = O(1) \end{aligned} \quad (78)$$

where we used $\|\sum_{t=s+1}^T \phi_{i,t-s}\| = \|\sum_{l=1}^{T-s} \phi_{i,l}\| \leq \sum_{l=1}^{T-s} \|\phi_{i,l}\| < \sum_{l=1}^{\infty} \|\phi_{i,l}\| < \infty$. For trend model, we have

$$\mathbf{h}_i = -\frac{2(2T+1) \mathbf{\Xi}'_i \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \mathbf{v}_i}{T(T-1)} - \frac{12 \mathbf{\Xi}'_i \boldsymbol{\tau}_T \boldsymbol{\tau}'_T \mathbf{v}_i}{T(T-1)(T+1)} + \frac{6(\mathbf{\Xi}'_i \boldsymbol{\nu}_T \boldsymbol{\tau}'_T \mathbf{v}_i + \mathbf{\Xi}'_i \boldsymbol{\tau}_T \boldsymbol{\nu}'_T \mathbf{v}_i)}{T(T-1)}$$

$$= -\mathbf{h}_{1i} - \mathbf{h}_{2i} + \mathbf{h}_{3i}.$$

Using Assumption 3 and (78), we have

$$\begin{aligned} E(\mathbf{h}_{1i}) &= \frac{2(2T+1)}{T(T-1)} E \left[\left(\sum_{t=1}^T \boldsymbol{\xi}_{it} \right) \left(\sum_{s=1}^T v_{is} \right) \right] = \frac{2(2T+1)}{T(T-1)} \sum_{s=1}^{T-1} \sum_{t=s+1}^T \phi_{i,t-s} = O(1), \\ E(\mathbf{h}_{2i}) &= \frac{12}{T(T-1)(T+1)} E \left(\left(\sum_{t=1}^T t \boldsymbol{\xi}_{it} \right) \left(\sum_{s=1}^T s v_{is} \right) \right) \\ &= \frac{12}{(T+1)} \left(\sum_{s=1}^{T-1} \sum_{t=s+1}^T \left(\frac{t}{T} \right) \left(\frac{s}{T-1} \right) \phi_{i,t-s} \right) = O(1), \\ E(\mathbf{h}_{3i}) &= \frac{6}{T(T-1)} \left(\sum_{t=1}^T \sum_{s=1}^T s E(\boldsymbol{\xi}_{it} v_{is}) + \sum_{t=1}^T \sum_{s=1}^T t E(\boldsymbol{\xi}_{it} v_{is}) \right) \\ &= \frac{6}{T} \sum_{s=1}^{T-1} \sum_{t=s+1}^T \left(\frac{s}{T-1} \right) \phi_{i,t-s} + \frac{6}{(T-1)} \sum_{s=1}^{T-1} \sum_{t=s+1}^T \left(\frac{t}{T} \right) \phi_{i,t-s} = O(1) \end{aligned}$$

where we used $0 < t/T \leq 1$ and $0 < s/(T-1) \leq 1$ for all s and t . Thus, both for FE and trend models, $E(\mathbf{h}_i) = O(1)$ and obtain

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{R}_T \mathbf{v}_i = \sqrt{\frac{N}{T}} \bar{\mathbf{h}}_N = O_p \left(\sqrt{\frac{N}{T}} \right)$$

where $\bar{\mathbf{h}}_N = N^{-1} \sum_{i=1}^N \mathbf{h}_i$.

Proof of (c): Noting $\mathbf{B}_{T_d} \mathbf{K}_T \mathbf{C}_T = \mathbf{0}$, we have the following decomposition:

$$\begin{aligned} \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}'_i \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{v}_i &= \frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{v}_i \\ &= \frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{v}_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{A}_T \mathbf{v}_i. \end{aligned}$$

The first term converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ by Assumption 4. To derive the order of the second term, let us define $\mathbf{s}_i = \boldsymbol{\Xi}'_i \mathbf{A}_T \mathbf{v}_i$. Then, using (73), and $\mathbf{v}_i = (\mathbf{v}'_{i1}, \mathbf{v}'_{i2}, \mathbf{v}'_{i3})'$ where \mathbf{v}_{i1} and \mathbf{v}_{i3} are $d \times 1$ and \mathbf{v}_{i2} is $T_{2d} \times 1$, \mathbf{s}_i can be decomposed as

$$\mathbf{s}_i = \mathbf{s}_{1i} + \mathbf{s}_{2i} + \mathbf{s}_{3i} + \mathbf{s}_{4i} + \mathbf{s}_{5i} + \mathbf{s}_{6i}$$

where

$$\begin{aligned} \mathbf{s}_{1i} &= -\boldsymbol{\Xi}'_{1i} \mathbf{v}_{1i}, & \mathbf{s}_{2i} &= \boldsymbol{\Xi}'_{1i} \mathbf{A}_{12} \mathbf{v}_{2i}, & \mathbf{s}_{3i} &= \boldsymbol{\Xi}'_{1i} \mathbf{A}_{13} \mathbf{v}_{3i}, & \mathbf{s}_{4i} &= \boldsymbol{\Xi}'_{2i} \mathbf{A}_{22} \mathbf{v}_{2i}, \\ \mathbf{s}_{5i} &= \boldsymbol{\Xi}'_{2i} \mathbf{A}_{23} \mathbf{v}_{3i}, & \mathbf{s}_{6i} &= -\boldsymbol{\Xi}'_{3i} \mathbf{v}_{3i}. \end{aligned}$$

To derive the variance of \mathbf{s}_i , we need to calculate $Var(\mathbf{s}_{ki})$ and $Cov(\mathbf{s}_{ki}, \mathbf{s}_{li}), (k \neq l)$ for $k, l = 1, \dots, 6$. We consider the FE and trend models separately.

FE model Since $d = 1$ in the FE model, we have

$$\text{Var}(\mathbf{s}_{1i}) = \text{Var}(\boldsymbol{\xi}_{i1}v_{i1}) = O(1).$$

Using $a_{12}^{1,t-1} = O\left(\frac{\log T}{T}\right)$ for all t , $a_{13}^{11} = O\left(\frac{\log T}{T}\right)$, we have

$$\text{Var}(\mathbf{s}_{2i}) = \text{Var}\left(\sum_{t=2}^{T_d} a_{12}^{1,t-1} \boldsymbol{\xi}_{i1}v_{it}\right) = \sum_{t=2}^{T_d} \left(a_{12}^{1,t-1}\right)^2 \text{Var}(\boldsymbol{\xi}_{i1}v_{it}) = O\left(\frac{(\log T)^2}{T}\right),$$

$$\text{Var}(\mathbf{s}_{3i}) = \text{Var}(a_{13}^{11} \boldsymbol{\xi}_{i1}v_{iT}) = (a_{13}^{11})^2 \text{Var}(\boldsymbol{\xi}_{i1}v_{iT}) = O\left(\frac{(\log T)^2}{T^2}\right).$$

Similarly, using $a_{22}^{t-1,t-1} = O(1/(t+1)) + O(1/(T-t))$ and $a_{22}^{s-1,t-1} = O(\log T/T)$ for all $s \neq t$, we have

$$\begin{aligned} \text{Var}(\mathbf{s}_{4i}) &= \text{Var}\left[\sum_{s=2}^{T_d-1} \sum_{t=s}^{T_d} a_{22}^{s-1,t-1} \boldsymbol{\xi}_{is}v_{it}\right] = \sum_{s=2}^{T_d-1} \sum_{t=s}^{T_d} (a_{22}^{s-1,t-1})^2 \text{Var}(\boldsymbol{\xi}_{is}v_{it}) \\ &= \sum_{t=2}^{T_d} (a_{22}^{t-1,t-1})^2 \text{Var}(\boldsymbol{\xi}_{it}v_{it}) + \sum_{s=2}^{T_d-1} \sum_{t=s+1}^{T_d} (a_{22}^{s-1,t-1})^2 \text{Var}(\boldsymbol{\xi}_{is}v_{it}) \\ &= O(1) + O((\log T)^2). \end{aligned}$$

Finally, using $a_{23}^{t-1,1} = O(\log T/T)$ for all t , and the definition of $\boldsymbol{\Xi}_{3i}$, we have

$$\text{Var}(\mathbf{s}_{5i}) = \text{Var}\left(\sum_{t=2}^{T_d} a_{23}^{t-1,1} \boldsymbol{\xi}_{it}v_{iT}\right) = \sum_{t=2}^{T_d} (a_{23}^{t-1,1})^2 \text{Var}(\boldsymbol{\xi}_{it}v_{iT}) = O\left(\frac{(\log T)^2}{T}\right),$$

$$\text{Var}(\mathbf{s}_{6i}) = \text{Var}(\boldsymbol{\xi}_{iT}v_{iT}) = O(1).$$

For the covariances, we have

$$\begin{aligned} \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{2i}) &= \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{3i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{4i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{5i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{6i}) = \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{3i}) \\ &= \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{5i}) = \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{6i}) = \text{Cov}(\mathbf{s}_{3i}, \mathbf{s}_{4i}) = \text{Cov}(\mathbf{s}_{4i}, \mathbf{s}_{5i}) = \text{Cov}(\mathbf{s}_{4i}, \mathbf{s}_{6i}) = \mathbf{0}, \end{aligned}$$

$$\begin{aligned} \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{4i}) &= E\left[\left(\sum_{t=2}^{T_d} a_{12}^{1,t-1} \boldsymbol{\xi}_{i1}v_{it}\right) \left(\sum_{t=2}^{T_d} a_{22}^{t-1,t-1} \boldsymbol{\xi}_{it}v_{it} + \sum_{s=2}^{T_d-1} \sum_{t=s+1}^{T_d} a_{22}^{s-1,t-1} \boldsymbol{\xi}_{is}v_{it}\right)'\right] \\ &= \sum_{t_1=2}^{T_d} \sum_{t_2=2}^{T_d} a_{12}^{1,t_1-1} a_{22}^{t_2-1,t_2-1} E(\boldsymbol{\xi}_{i1} \boldsymbol{\xi}'_{it_2} v_{it_1} v_{it_2}) + \sum_{t_1=2}^{T_d} \sum_{s=2}^{T_d-1} \sum_{t_2=s+1}^{T_d} a_{12}^{1,t_1-1} a_{22}^{s-1,t_2-1} E(\boldsymbol{\xi}_{i1} \boldsymbol{\xi}'_{is} v_{it_1} v_{it_2}) \\ &= \sum_{t=2}^{T_d} a_{12}^{1,t-1} a_{22}^{t-1,t-1} E(\boldsymbol{\xi}_{i1} \boldsymbol{\xi}'_{it} v_{it}^2) + \sum_{s=2}^{T_d-1} \sum_{t=s+1}^{T_d} a_{12}^{1,t-1} a_{22}^{s-1,t-1} E(\boldsymbol{\xi}_{i1} \boldsymbol{\xi}'_{is} v_{it}^2) \\ &= O\left(\frac{(\log T)^2}{T}\right) + O((\log T)^2), \end{aligned}$$

$$\text{Cov}(\mathbf{s}_{3i}, \mathbf{s}_{5i}) = \sum_{t=2}^{T_d} a_{13}^{11} a_{23}^{t-1,1} E(\boldsymbol{\xi}_{i1} \boldsymbol{\xi}'_{it} v_{iT}^2) = O\left(\frac{(\log T)^2}{T}\right),$$

$$\text{Cov}(\mathbf{s}_{3i}, \mathbf{s}_{6i}) = a_{13}^{11} E(\boldsymbol{\xi}_{i1} \boldsymbol{\xi}'_{iT} v_{iT}^2) = O\left(\frac{\log T}{T}\right),$$

$$\text{Cov}(\mathbf{s}_{5i}, \mathbf{s}_{6i}) = \sum_{t=2}^{T_d} a_{23}^{t-1,1} E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{iT} v_{iT}^2) = O(\log T).$$

Therefore, for FE model, we have $\text{Var}(\mathbf{s}_i) = O((\log T)^2)$, and

$$\text{Var}\left(\frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{A}_T \mathbf{v}_i\right) = \frac{1}{NT} \sum_{i=1}^N \text{Var}(\mathbf{s}_i) = O\left(\frac{(\log T)^2}{T}\right).$$

Hence, it follows that $\frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{A}_T \mathbf{v}_i = O_p(\log T/\sqrt{T}) = o_p(1)$.

Trend model

$$\text{Var}(\mathbf{s}_{1i}) = \text{Var}\left(\sum_{s=1}^2 \sum_{t=1}^2 \boldsymbol{\xi}_{it} v_{is}\right) = \text{Var}(\boldsymbol{\xi}_{i1} v_{i1}) + \text{Var}(\boldsymbol{\xi}_{i2} v_{i2}) = O(1).$$

Using $a_{12}^{s,t-1} = O\left(\frac{\log T}{T}\right)$, $s = 2$ for all t and $a_{13}^{jk} = O\left(\frac{\log T}{T}\right)$, we have

$$\text{Var}(\mathbf{s}_{2i}) = \text{Var}\left(\sum_{s=1}^2 \sum_{t=3}^{T_d} a_{12}^{j,t-1} \boldsymbol{\xi}_{i1} v_{it}\right) = O\left(\frac{(\log T)^2}{T}\right),$$

$$\text{Var}(\mathbf{s}_{3i}) = \text{Var}\left(\sum_{s=1}^2 \sum_{t=1}^2 a_{13}^{st} \boldsymbol{\xi}_{it} v_{iT_{s+1}}\right) = \sum_{s=1}^2 \sum_{t=1}^2 (a_{13}^{st})^2 \text{Var}(\boldsymbol{\xi}_{it} v_{iT_{s+1}}) = O\left(\frac{(\log T)^2}{T^2}\right).$$

Similarly, using $a_{22}^{t-1,t-1} = O(1/(t+1)) + O(1/(T-t))$ and $a_{22}^{s-1,t-1} = O(\log T/T)$ for all $s \neq t$, we have

$$\begin{aligned} \text{Var}(\mathbf{s}_{4i}) &= \text{Var}\left[\sum_{s=3}^{T_d-1} \sum_{t=s}^{T_d} a_{22}^{s-1,t-1} \boldsymbol{\xi}_{is} v_{it}\right] = \sum_{s=3}^{T_d-1} \sum_{t=s}^{T_d} (a_{22}^{s-1,t-1})^2 \text{Var}(\boldsymbol{\xi}_{is} v_{it}) \\ &= \sum_{t=3}^{T_d} (a_{22}^{t-1,t-1})^2 \text{Var}(\boldsymbol{\xi}_{it} v_{it}) + \sum_{s=3}^{T_d-1} \sum_{t=s+1}^{T_d} (a_{22}^{s-1,t-1})^2 \text{Var}(\boldsymbol{\xi}_{is} v_{it}) \\ &= O(1) + O((\log T)^2). \end{aligned}$$

Finally, using $a_{23}^{t-1,s} = O(\log T/T)$, $s = 1, 2$ for all t , and the definition of $\boldsymbol{\Xi}_{3i}$, we have

$$\begin{aligned} \text{Var}(\mathbf{s}_{5i}) &= \text{Var}\left(\sum_{s=1}^2 \sum_{t=3}^{T_d} a_{23}^{t-1,s} \boldsymbol{\xi}_{it} v_{iT_{s+1}}\right) \\ &= \sum_{s=1}^2 \sum_{t=2}^{T_d} (a_{23}^{t-1,s})^2 \text{Var}(\boldsymbol{\xi}_{it} v_{iT_{s+1}}) = O\left(\frac{(\log T)^2}{T}\right), \end{aligned}$$

$$\text{Var}(\mathbf{s}_{6i}) = \text{Var}\left(\sum_{s=1}^2 \sum_{t=1}^2 \boldsymbol{\xi}_{iT_t} v_{iT_s}\right) = O(1).$$

For the covariances of trend model,

$$\begin{aligned} &\text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{2i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{3i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{4i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{5i}) = \text{Cov}(\mathbf{s}_{1i}, \mathbf{s}_{6i}) = \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{3i}) \\ &= \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{5i}) = \text{Cov}(\mathbf{s}_{2i}, \mathbf{s}_{6i}) = \text{Cov}(\mathbf{s}_{3i}, \mathbf{s}_{4i}) = \text{Cov}(\mathbf{s}_{4i}, \mathbf{s}_{5i}) = \text{Cov}(\mathbf{s}_{4i}, \mathbf{s}_{6i}) = \mathbf{0}, \end{aligned}$$

$$\begin{aligned}
Cov(\mathbf{s}_{2i}, \mathbf{s}_{4i}) &= E \left[\left(\sum_{s=1}^2 \sum_{t=2}^{T_d} a_{12}^{s,t-1} \boldsymbol{\xi}_{is} v_{it} \right) \left(\sum_{t=3}^{T_d} a_{22}^{t-1,t-1} \boldsymbol{\xi}_{it} v_{it} + \sum_{s=3}^{T_d-1} \sum_{t=s+1}^{T_d} a_{22}^{s-1,t-1} \boldsymbol{\xi}_{is} v_{it} \right) \right] \\
&= \sum_{s=1}^2 \sum_{t_1=3}^{T_d} \sum_{t_2=3}^{T_d} a_{12}^{1,t_1-1} a_{22}^{t_2-1,t_2-1} E(\boldsymbol{\xi}_{is} \boldsymbol{\xi}'_{it_2} v_{it_1} v_{it_2}) \\
&+ \sum_{s_1=1}^2 \sum_{t_1=3}^{T_d} \sum_{s_2=3}^{T_d-1} \sum_{t_2=s_2+1}^{T_d} a_{12}^{1,t_1-1} a_{22}^{s_2-1,t_2-1} E(\boldsymbol{\xi}_{is_1} \boldsymbol{\xi}'_{is_2} v_{it_1} v_{it_2}) \\
&= \sum_{s=1}^2 \sum_{t=3}^{T_d} a_{12}^{s,t-1} a_{22}^{t-1,t-1} E(\boldsymbol{\xi}_{is} \boldsymbol{\xi}'_{it} v_{it}^2) + \sum_{s_1=1}^2 \sum_{s_2=2}^{T_d-1} \sum_{t=s_2+1}^{T_d} a_{12}^{s_1,t-1} a_{22}^{s_2-1,t-1} E(\boldsymbol{\xi}_{is_1} \boldsymbol{\xi}'_{is_2} v_{it}^2) \\
&= O\left(\frac{(\log T)^2}{T}\right) + O((\log T)^2),
\end{aligned}$$

$$Cov(\mathbf{s}_{3i}, \mathbf{s}_{5i}) = \sum_{s_1=1}^2 \sum_{s_2=1}^2 \sum_{t=3}^{T_d} a_{13}^{s_1 s_2 1} a_{23}^{t-1, s_2} E(\boldsymbol{\xi}_{is_1} \boldsymbol{\xi}'_{it} v_{iT_{s_1+1}}^2) = O\left(\frac{(\log T)^2}{T}\right),$$

$$Cov(\mathbf{s}_{3i}, \mathbf{s}_{6i}) = \sum_{s_1=1}^2 \sum_{s_2=1}^2 a_{13}^{s_1 s_2} E(\boldsymbol{\xi}_{is_1} \boldsymbol{\xi}'_{iT_{s_2+1}} v_{iT_{s_1+1}}^2) = O\left(\frac{\log T}{T}\right),$$

$$Cov(\mathbf{s}_{5i}, \mathbf{s}_{6i}) = \sum_{s=1}^2 \sum_{t=3}^{T_d} a_{23}^{t-1, s} E(\boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{iT_{s+1}} v_{iT_s+1}^2) = O(\log T).$$

Therefore, for trend model, we have $Var(\mathbf{s}_i) = O((\log T)^2)$, and

$$Var\left(\frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{A}_T \mathbf{v}_i\right) = \frac{1}{NT} \sum_{i=1}^N Var(\mathbf{s}_i) = O\left(\frac{(\log T)^2}{T}\right).$$

Hence, it follows that $\frac{1}{\sqrt{NT}} \sum_{i=1}^N \boldsymbol{\Xi}'_i \mathbf{A}_T \mathbf{v}_i = O_p(\log T / \sqrt{T}) = o_p(1)$.

Proof of Theorems 2 and 3

We first provide a proof of Theorem 2. Using (77) and Assumption 4, we have

$$\frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{W}_i = \frac{1}{NT} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \boldsymbol{\xi}_{it} \boldsymbol{\xi}'_{it} + O_p\left(\frac{1}{T}\right) \xrightarrow{P} \boldsymbol{\Gamma}_0. \quad (79)$$

Next, we have the following decomposition

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{v}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \boldsymbol{\xi}_{it} v_{it} + \sqrt{\frac{N}{T}} \bar{\mathbf{h}}_N.$$

Hence, using Assumption 4, as $N, T \rightarrow \infty$ with $N/T \rightarrow \kappa$, ($0 < \kappa < \infty$), we obtain

$$\sqrt{NT}(\hat{\boldsymbol{\delta}}_{FE} - \boldsymbol{\delta}) = \left(\frac{1}{NT} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{W}_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}'_i \mathbf{Q}_T \mathbf{v}_i$$

$$\begin{aligned}
&= \left(\frac{1}{NT} \sum_{i=1}^N \mathbf{W}_i' \mathbf{Q}_T \mathbf{W}_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \xi_{it} v_{it} + \left(\frac{1}{NT} \sum_{i=1}^N \mathbf{W}_i' \mathbf{Q}_T \mathbf{W}_i \right)^{-1} \sqrt{\frac{N}{T}} \bar{\mathbf{h}}_N \\
&\xrightarrow{d} \mathcal{N}(\sqrt{\kappa} \mathbf{\Gamma}_0^{-1} \bar{\mathbf{h}}, \mathbf{\Gamma}_0^{-1} \mathbf{\Omega} \mathbf{\Gamma}_0^{-1})
\end{aligned}$$

where $\bar{\mathbf{h}} = \text{plim}_{N,T \rightarrow \infty} \bar{\mathbf{h}}_N$. Theorem 3 can be proved by noting that $\widehat{\mathbf{\Gamma}}_0$ and $\widehat{\mathbf{h}}$ are consistent estimators of $\mathbf{\Gamma}_0$ and $\bar{\mathbf{h}}$ with large N and T .

Proof of Theorem 4

Using Lemma 1(a), (79) and Assumption 4, we have

$$\frac{1}{NT} \sum_{i=1}^N \mathbf{W}_i' \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{W}_i = \frac{1}{NT} \sum_{i=1}^N \mathbf{W}_i' \mathbf{Q}_T \mathbf{W}_i + o_p \left(\frac{\log T}{T} \right) \rightarrow^p \mathbf{\Gamma}_0. \quad (80)$$

Also, using Lemma 1(c) and Assumption 4, we have

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{W}_i' \mathbf{K}'_T \mathbf{B}'_{T_d} \mathbf{F}_{T_d} \mathbf{L}_T \mathbf{v}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=d+1}^{T_d} \xi_{it} v_{it} + o_p(1) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}). \quad (81)$$

Combining (80) and (81), we obtain the result.

References

- Ahn, S. C. and P. Schmidt (1995) “Efficient Estimation of Models for Dynamic Panel Data,” *Journal of Econometrics*, 68, 5-27.
- Akashi, K. and N. Kunitomo (2015) “The limited information maximum likelihood approach to dynamic panel structural equation models,” *Annals of the Institute of Statistical Mathematics*, 67, 39–73.
- Alvarez, J. and M. Arellano (2003) “The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators,” *Econometrica*, 71, 1121-1159.
- Anderson, T. W. and C. Hsiao (1981) “Estimation of Dynamic Models with Error Components,” *Journal of the American Statistical Association*, 76, 598-606.
- Arellano, M. (2003) *Panel Data Econometrics*, Oxford: Oxford University Press.
- Arellano, M. and S. Bond (1991) “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *Review of Economic Studies*, 58, 277-297.
- Arellano, M. and O. Bover (1995) “Another Look at the Instrumental Variable Estimation of Error-Components Models,” *Journal of Econometrics*, 68, 29-51.
- Blundell, R. and S. Bond (1998) “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics*, 87, 115-143.
- Bun, M. J. G. and M. A. Carree (2005) “Bias-Corrected Estimation in Dynamic Panel Data Models,” *Journal of Business and Economic Statistics*, 23, 200-210.

- Chen, Y. (2013) “A New Algorithm for Computing the Determinant and the Inverse of a Pentadiagonal Toeplitz Matrix,” *Engineering*, 5, p. 25.
- Dow, M. (2003) “Explicit inverses of Toeplitz and associated matrices,” *ANZIAM Journal*, 44, E185–E215.
- El-Mikkawy, M. and A. Karawia (2006) “Inversion of General Tridiagonal Matrices,” *Applied Mathematics Letters*, 19, 712–720.
- Hahn, J. and G. Kuersteiner (2002) “Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T Are Large,” *Econometrica*, 70, 1639-1657.
- Hayakawa, K. (2006) “Efficient GMM estimation of dynamic panel data models where large heterogeneity may be present.”
- Hayakawa, K. (2009) “A Simple Efficient Instrumental Variable Estimator in Panel AR(p) Models When Both N and T Are Large,” *Econometric Theory*, 25, 873-890.
- Hayakawa, K. (2015) “The Asymptotic Properties of the System GMM Estimator in Dynamic Panel Data Models When Both N and T Are Large,” *Econometric Theory*, 31, 647-667.
- Hayakawa, K. (2016) “Improved GMM Estimation of Panel VAR Models,” *Computational Statistics & Data Analysis*.
- Holtz-Eakin, D., W. K. Newey, and H. S. Rosen (1988) “Estimating Vector Autoregressions with Panel Data,” *Econometrica*, 56, 1371-1395.
- Hsiao, C. and Q. Zhou (2015) “First Difference or Forward Demeaning: Implications for the Method of Moments Estimators.” Working paper.
- Lee, Y.J., R. Okui, and M. Shintani (2016) “Asymptotic inference for dynamic panel estimators of infinite order autoregressive processes.” Working paper.
- Moon, H. R. and P. C. B. Phillips (2000) “Estimation of Autoregressive Roots Near Unity Using Panel Data,” *Econometric Theory*, 16, 927-997.
- Nickell, S. J. (1981) “Biases in Dynamic Models with Fixed Effects,” *Econometrica*, 49, 1417-1426.
- Okui, R. (2009) “The Optimal Choice of Moments in Dynamic Panel Data Models,” *Journal of Econometrics*, 151, 1-16.
- Phillips, P. C. B. and H. R. Moon (1999) “Linear Regression Limit Theory for Nonstationary Panel Data,” *Econometrica*, 67, 1057-1111.
- Phillips, P. C. B. and D. Sul (2007) “Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence,” *Journal of Econometrics*, 127, 162-188.
- Wansbeek, T. J. and T. Knaap (1999) “Estimating a Dynamic Panel Data Model with Heterogeneous Trends,” *Annales d’Economie et de Statistique*, 55-56, 331-349.

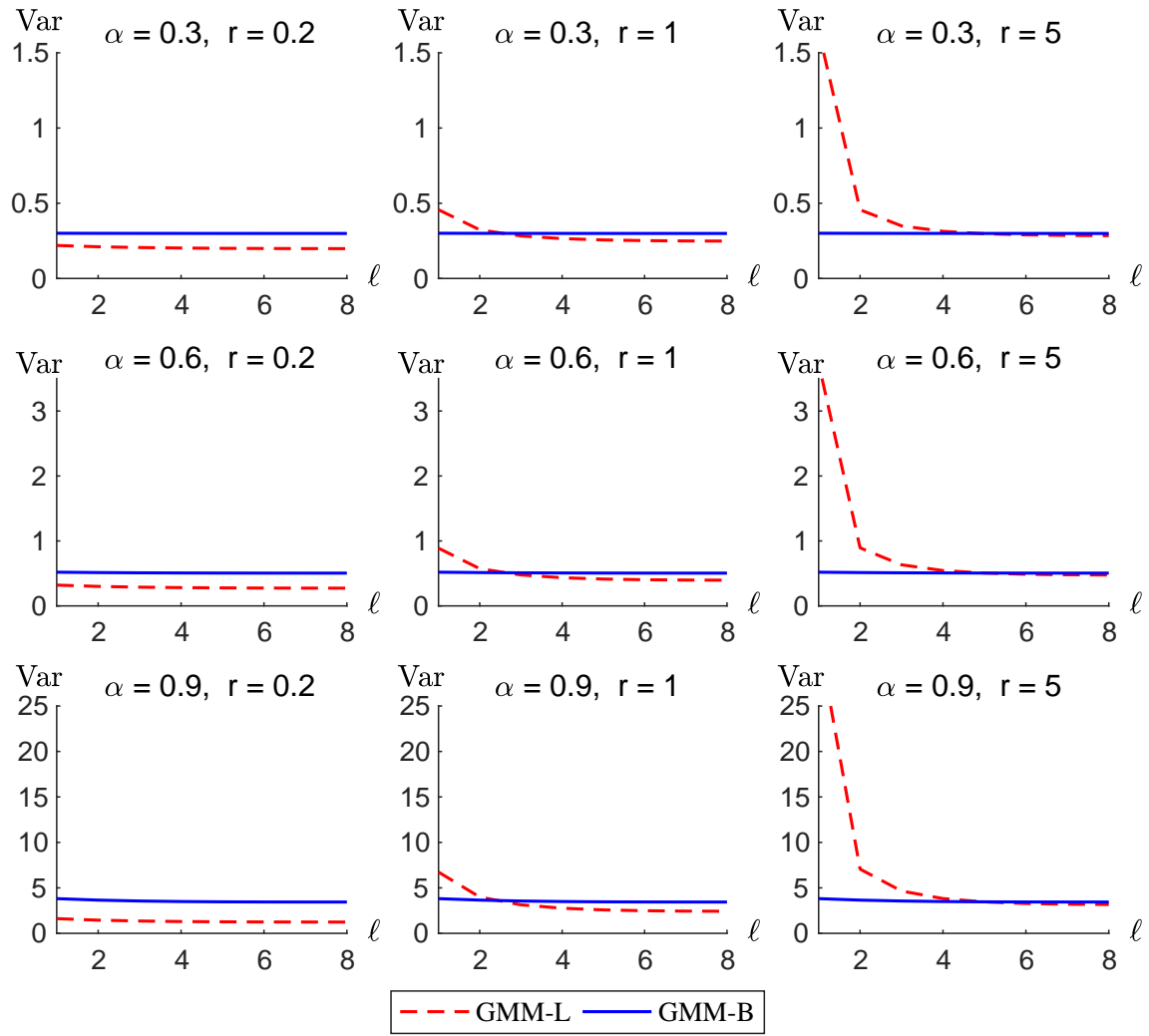


Figure 1: Asymptotic variance of GMM estimators with various instruments lag length ℓ ($T = 10$).

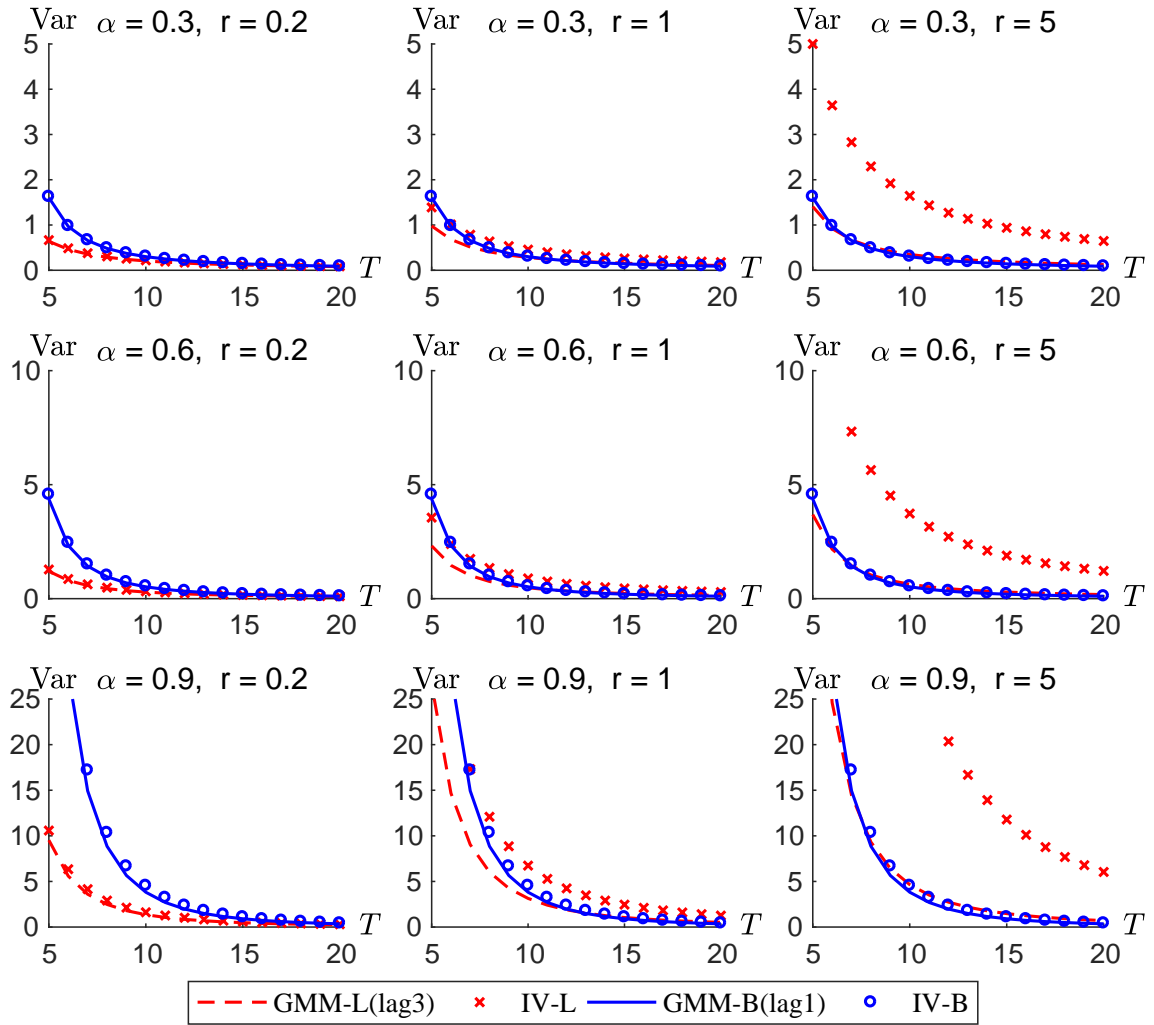


Figure 2: Asymptotic variance of IV/GMM estimators with various T .

Table 1: Fixed effects model: $\alpha = 0.4$, $\beta = 1.0$

$N = 50, \sigma_\eta^2 = 1$														
	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.113	0.003	-0.042	-0.036	0.000	-0.013	-0.035	0.034	0.004	-0.085	-0.039	0.000	-0.014	-0.025
IQR	0.054	0.254	0.156	0.086	0.100	0.097	0.088	0.140	0.579	0.361	0.223	0.302	0.275	0.252
MAE	0.113	0.126	0.080	0.052	0.050	0.048	0.050	0.075	0.285	0.183	0.112	0.150	0.137	0.129
Size	0.818	0.021	0.057	0.086	0.055	0.061	0.090	0.072	0.017	0.041	0.064	0.055	0.056	0.058
$T = 25$														
Bias	-0.045	0.001	-0.016	-0.011	0.000	-0.004	-0.009	0.021	0.007	-0.023	-0.008	-0.001	-0.003	-0.001
IQR	0.031	0.131	0.090	0.045	0.043	0.041	0.039	0.084	0.243	0.168	0.111	0.116	0.110	0.105
MAE	0.045	0.065	0.045	0.024	0.021	0.020	0.020	0.044	0.120	0.086	0.056	0.058	0.055	0.052
Size	0.490	0.045	0.062	0.070	0.063	0.057	0.066	0.077	0.035	0.055	0.066	0.064	0.067	0.062
$T = 50$														
Bias	-0.022	-0.002	-0.010	-0.004	0.000	-0.001	-0.003	0.014	0.000	-0.012	0.001	0.002	0.002	0.002
IQR	0.022	0.083	0.063	0.031	0.026	0.026	0.025	0.054	0.144	0.105	0.068	0.068	0.066	0.064
MAE	0.022	0.042	0.032	0.016	0.013	0.013	0.013	0.029	0.071	0.053	0.034	0.034	0.033	0.033
Size	0.281	0.045	0.057	0.075	0.066	0.062	0.063	0.073	0.044	0.056	0.056	0.062	0.061	0.056
$T = 100$														
Bias	-0.011	0.001	-0.004	-0.002	0.001	0.000	-0.001	0.006	0.001	-0.003	0.000	0.001	0.002	0.001
IQR	0.015	0.058	0.043	0.020	0.017	0.017	0.017	0.038	0.090	0.072	0.045	0.043	0.043	0.042
MAE	0.011	0.029	0.022	0.010	0.008	0.008	0.008	0.020	0.045	0.035	0.023	0.022	0.021	0.021
Size	0.167	0.053	0.058	0.060	0.056	0.055	0.059	0.070	0.049	0.050	0.056	0.057	0.058	0.055

$N = 100, \sigma_\eta^2 = 1$														
	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.114	0.001	-0.027	-0.021	0.001	-0.005	-0.017	0.038	0.009	-0.053	-0.021	0.005	0.000	-0.014
IQR	0.039	0.177	0.130	0.070	0.070	0.069	0.066	0.100	0.402	0.302	0.168	0.213	0.207	0.192
MAE	0.114	0.089	0.066	0.037	0.035	0.035	0.036	0.055	0.201	0.157	0.086	0.109	0.104	0.097
Size	0.987	0.027	0.055	0.073	0.061	0.061	0.073	0.094	0.026	0.045	0.062	0.055	0.060	0.058
$T = 25$														
Bias	-0.044	0.002	-0.011	-0.005	0.000	-0.001	-0.004	0.022	0.003	-0.018	-0.003	0.001	0.001	0.002
IQR	0.022	0.090	0.073	0.033	0.030	0.029	0.029	0.057	0.169	0.135	0.077	0.079	0.076	0.075
MAE	0.044	0.045	0.036	0.017	0.015	0.015	0.014	0.034	0.085	0.068	0.038	0.039	0.039	0.037
Size	0.749	0.038	0.058	0.055	0.046	0.049	0.051	0.083	0.040	0.055	0.053	0.060	0.060	0.059
$T = 50$														
Bias	-0.022	-0.002	-0.006	-0.003	0.000	-0.001	-0.002	0.013	-0.001	-0.007	0.000	0.001	0.001	0.002
IQR	0.016	0.062	0.051	0.024	0.018	0.018	0.018	0.037	0.104	0.084	0.049	0.047	0.046	0.045
MAE	0.022	0.031	0.025	0.012	0.009	0.009	0.009	0.021	0.051	0.041	0.025	0.023	0.023	0.023
Size	0.461	0.050	0.060	0.058	0.063	0.062	0.064	0.072	0.045	0.055	0.052	0.049	0.051	0.050
$T = 100$														
Bias	-0.011	0.001	-0.003	-0.001	0.000	0.000	-0.001	0.006	0.001	-0.002	0.000	0.000	0.000	0.000
IQR	0.011	0.041	0.034	0.016	0.012	0.012	0.012	0.028	0.064	0.056	0.034	0.031	0.031	0.031
MAE	0.011	0.021	0.017	0.008	0.006	0.006	0.006	0.015	0.032	0.027	0.017	0.016	0.016	0.015
Size	0.266	0.056	0.054	0.053	0.050	0.051	0.053	0.061	0.051	0.049	0.049	0.048	0.051	0.052

$N = 250, \sigma_\eta^2 = 1$														
	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.115	0.001	-0.017	-0.009	0.001	-0.003	-0.008	0.035	0.003	-0.028	-0.011	0.003	0.000	-0.007
IQR	0.024	0.111	0.092	0.045	0.046	0.045	0.044	0.061	0.249	0.209	0.102	0.130	0.124	0.117
MAE	0.115	0.055	0.047	0.023	0.023	0.022	0.023	0.040	0.124	0.107	0.052	0.065	0.062	0.059
Size	1.000	0.035	0.040	0.058	0.057	0.058	0.060	0.106	0.039	0.037	0.047	0.045	0.043	0.047
$T = 25$														
Bias	-0.044	0.000	-0.006	-0.003	0.000	-0.001	-0.002	0.022	-0.001	-0.008	-0.001	-0.001	-0.001	0.000
IQR	0.014	0.057	0.048	0.022	0.018	0.018	0.018	0.036	0.107	0.092	0.050	0.048	0.048	0.048
MAE	0.044	0.028	0.025	0.011	0.009	0.009	0.009	0.024	0.053	0.047	0.025	0.024	0.024	0.024
Size	0.982	0.053	0.053	0.055	0.048	0.051	0.055	0.129	0.056	0.054	0.054	0.057	0.050	0.047
$T = 50$														
Bias	-0.022	-0.001	-0.003	-0.002	0.000	-0.001	-0.001	0.012	-0.002	-0.004	-0.001	0.000	0.000	0.001
IQR	0.010	0.039	0.035	0.015	0.012	0.012	0.012	0.025	0.065	0.058	0.032	0.029	0.029	0.029
MAE	0.022	0.019	0.017	0.008	0.006	0.006	0.006	0.015	0.033	0.030	0.016	0.015	0.015	0.014
Size	0.840	0.046	0.053	0.056	0.060	0.055	0.055	0.105	0.050	0.054	0.050	0.052	0.050	0.053
$T = 100$														
Bias	-0.011	0.000	-0.001	0.000	0.000	0.000	0.000	0.006	0.000	-0.001	0.000	0.000	0.000	0.000
IQR	0.007	0.025	0.023	0.010	0.008	0.007	0.007	0.018	0.042	0.037	0.020	0.019	0.018	0.018
MAE	0.011	0.012	0.011	0.005	0.004	0.004	0.004	0.010	0.021	0.019	0.010	0.009	0.009	0.009
Size	0.549	0.050	0.050	0.052	0.059	0.063	0.061	0.076	0.049	0.046	0.055	0.052	0.051	0.051

Table 1(cont.): Fixed effects model: $\alpha = 0.4$, $\beta = 1.0$

$N = 50, \sigma_{\eta}^2 = 5$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.113	0.010	-0.054	-0.036	0.000	-0.013	-0.035	0.034	0.039	-0.112	-0.044	0.000	-0.014	-0.025
IQR	0.054	0.389	0.172	0.089	0.100	0.097	0.088	0.140	0.929	0.439	0.228	0.302	0.275	0.252
MAE	0.113	0.197	0.091	0.052	0.050	0.048	0.050	0.075	0.467	0.223	0.118	0.150	0.137	0.129
Size	0.818	0.012	0.054	0.086	0.055	0.061	0.090	0.072	0.009	0.044	0.068	0.055	0.056	0.058
$T = 25$														
Bias	-0.045	0.007	-0.022	-0.011	0.000	-0.004	-0.009	0.021	0.014	-0.032	-0.010	-0.001	-0.003	-0.001
IQR	0.031	0.217	0.103	0.046	0.043	0.041	0.039	0.084	0.417	0.201	0.114	0.116	0.110	0.105
MAE	0.045	0.108	0.053	0.024	0.021	0.020	0.020	0.044	0.205	0.104	0.057	0.058	0.055	0.052
Size	0.490	0.035	0.066	0.068	0.063	0.057	0.066	0.077	0.027	0.061	0.070	0.064	0.067	0.062
$T = 50$														
Bias	-0.022	-0.004	-0.014	-0.004	0.000	-0.001	-0.003	0.014	-0.003	-0.020	0.000	0.002	0.002	0.002
IQR	0.022	0.137	0.074	0.031	0.026	0.026	0.025	0.054	0.241	0.123	0.068	0.068	0.066	0.064
MAE	0.022	0.067	0.037	0.016	0.013	0.013	0.013	0.029	0.117	0.064	0.034	0.034	0.033	0.033
Size	0.281	0.039	0.063	0.074	0.066	0.062	0.063	0.073	0.037	0.068	0.061	0.062	0.061	0.056
$T = 100$														
Bias	-0.011	0.002	-0.004	-0.002	0.001	0.000	-0.001	0.006	0.003	-0.005	0.000	0.001	0.002	0.001
IQR	0.015	0.091	0.050	0.020	0.017	0.017	0.017	0.038	0.149	0.084	0.047	0.043	0.043	0.042
MAE	0.011	0.045	0.024	0.010	0.008	0.008	0.008	0.020	0.075	0.042	0.023	0.022	0.021	0.021
Size	0.167	0.045	0.062	0.060	0.056	0.055	0.059	0.070	0.042	0.051	0.057	0.057	0.058	0.055

$N = 100, \sigma_{\eta}^2 = 5$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.114	0.012	-0.040	-0.022	0.001	-0.005	-0.017	0.038	0.030	-0.079	-0.025	0.005	0.000	-0.014
IQR	0.039	0.291	0.164	0.072	0.070	0.069	0.066	0.100	0.698	0.391	0.178	0.213	0.207	0.192
MAE	0.114	0.145	0.082	0.037	0.035	0.035	0.036	0.055	0.351	0.199	0.090	0.109	0.104	0.097
Size	0.987	0.016	0.050	0.073	0.061	0.061	0.073	0.094	0.017	0.047	0.068	0.055	0.060	0.058
$T = 25$														
Bias	-0.044	0.004	-0.016	-0.005	0.000	-0.001	-0.004	0.022	0.013	-0.025	-0.003	0.001	0.001	0.002
IQR	0.022	0.149	0.091	0.034	0.030	0.029	0.029	0.057	0.286	0.177	0.079	0.079	0.076	0.075
MAE	0.044	0.073	0.045	0.018	0.015	0.015	0.014	0.034	0.141	0.090	0.039	0.039	0.039	0.037
Size	0.749	0.038	0.052	0.056	0.046	0.049	0.051	0.083	0.034	0.057	0.052	0.060	0.060	0.059
$T = 50$														
Bias	-0.022	-0.002	-0.007	-0.003	0.000	-0.001	-0.002	0.013	-0.003	-0.013	0.000	0.001	0.001	0.002
IQR	0.016	0.096	0.064	0.024	0.018	0.018	0.018	0.037	0.167	0.108	0.051	0.047	0.046	0.045
MAE	0.022	0.048	0.031	0.012	0.009	0.009	0.009	0.021	0.084	0.054	0.025	0.023	0.023	0.023
Size	0.461	0.048	0.054	0.055	0.063	0.062	0.064	0.072	0.044	0.049	0.050	0.049	0.051	0.050
$T = 100$														
Bias	-0.011	0.001	-0.003	-0.001	0.000	0.000	-0.001	0.006	0.001	-0.006	-0.001	0.000	0.000	0.000
IQR	0.011	0.066	0.044	0.016	0.012	0.012	0.012	0.028	0.107	0.072	0.035	0.031	0.031	0.031
MAE	0.011	0.033	0.022	0.008	0.006	0.006	0.006	0.015	0.052	0.036	0.017	0.016	0.016	0.015
Size	0.266	0.053	0.061	0.053	0.050	0.051	0.053	0.061	0.056	0.053	0.049	0.048	0.051	0.052

$N = 250, \sigma_{\eta}^2 = 5$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.115	0.006	-0.023	-0.009	0.001	-0.003	-0.008	0.035	0.010	-0.048	-0.014	0.003	0.000	-0.007
IQR	0.024	0.175	0.121	0.046	0.046	0.045	0.044	0.061	0.436	0.288	0.112	0.130	0.124	0.117
MAE	0.115	0.088	0.061	0.023	0.023	0.022	0.023	0.040	0.219	0.147	0.056	0.065	0.062	0.059
Size	1.000	0.019	0.038	0.057	0.057	0.058	0.060	0.106	0.019	0.032	0.047	0.045	0.043	0.047
$T = 25$														
Bias	-0.044	0.000	-0.008	-0.003	0.000	-0.001	-0.002	0.022	-0.002	-0.014	-0.002	-0.001	-0.001	0.000
IQR	0.014	0.094	0.066	0.022	0.018	0.018	0.018	0.036	0.180	0.131	0.052	0.048	0.048	0.048
MAE	0.044	0.047	0.034	0.011	0.009	0.009	0.009	0.024	0.090	0.065	0.026	0.024	0.024	0.024
Size	0.982	0.042	0.054	0.054	0.048	0.051	0.055	0.129	0.042	0.058	0.057	0.057	0.050	0.047
$T = 50$														
Bias	-0.022	-0.001	-0.006	-0.002	0.000	-0.001	-0.001	0.012	-0.005	-0.008	-0.001	0.000	0.000	0.001
IQR	0.010	0.059	0.049	0.015	0.012	0.012	0.012	0.025	0.106	0.082	0.033	0.029	0.029	0.029
MAE	0.022	0.030	0.024	0.008	0.006	0.006	0.006	0.015	0.052	0.042	0.016	0.015	0.015	0.014
Size	0.840	0.042	0.053	0.055	0.060	0.055	0.055	0.105	0.039	0.060	0.047	0.052	0.050	0.053
$T = 100$														
Bias	-0.011	0.001	-0.002	0.000	0.000	0.000	0.000	0.006	0.001	-0.002	0.000	0.000	0.000	0.000
IQR	0.007	0.040	0.032	0.010	0.008	0.007	0.007	0.018	0.066	0.054	0.021	0.019	0.018	0.018
MAE	0.011	0.020	0.016	0.005	0.004	0.004	0.004	0.010	0.033	0.026	0.011	0.009	0.009	0.009
Size	0.549	0.051	0.058	0.051	0.059	0.063	0.061	0.076	0.048	0.050	0.049	0.052	0.051	0.051

Table 2: Fixed effects model: $\alpha = 0.8$, $\beta = 1.0$

$N = 50, \sigma_{\eta}^2 = 1$															
	α							β							
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	
$T = 10$															
Bias	-0.150	0.036	-0.189	-0.109	0.004	-0.053	-0.101	0.006	0.034	-0.164	-0.111	0.013	-0.074	-0.122	
IQR	0.049	0.641	0.243	0.112	0.188	0.143	0.112	0.137	0.745	0.407	0.254	0.429	0.335	0.279	
MAE	0.150	0.320	0.191	0.110	0.094	0.079	0.102	0.068	0.370	0.224	0.151	0.210	0.168	0.159	
Size	0.995	0.001	0.106	0.243	0.029	0.085	0.215	0.062	0.005	0.038	0.097	0.031	0.057	0.094	
$T = 25$															
Bias	-0.054	0.062	-0.072	-0.032	-0.002	-0.010	-0.022	0.021	0.046	-0.030	-0.023	0.000	-0.008	-0.016	
IQR	0.024	0.387	0.115	0.048	0.049	0.045	0.040	0.084	0.293	0.184	0.123	0.122	0.116	0.112	
MAE	0.054	0.205	0.075	0.035	0.025	0.024	0.025	0.043	0.150	0.095	0.062	0.061	0.058	0.057	
Size	0.908	0.012	0.067	0.142	0.057	0.064	0.124	0.081	0.014	0.054	0.068	0.052	0.061	0.064	
$T = 50$															
Bias	-0.026	0.061	-0.034	-0.012	-0.001	-0.003	-0.007	0.017	0.012	-0.004	-0.003	0.002	0.001	0.001	
IQR	0.015	0.289	0.075	0.029	0.023	0.022	0.021	0.055	0.156	0.113	0.075	0.066	0.065	0.065	
MAE	0.026	0.157	0.042	0.017	0.012	0.011	0.011	0.029	0.078	0.056	0.037	0.033	0.033	0.032	
Size	0.664	0.030	0.059	0.086	0.056	0.063	0.080	0.079	0.016	0.060	0.059	0.063	0.061	0.062	
$T = 100$															
Bias	-0.012	0.044	-0.015	-0.005	0.000	-0.001	-0.002	0.009	-0.001	0.005	0.000	0.001	0.001	0.002	
IQR	0.010	0.207	0.048	0.019	0.013	0.012	0.012	0.039	0.089	0.072	0.049	0.042	0.041	0.040	
MAE	0.012	0.113	0.024	0.010	0.006	0.006	0.006	0.020	0.045	0.036	0.025	0.021	0.020	0.020	
Size	0.400	0.072	0.063	0.083	0.056	0.059	0.064	0.071	0.012	0.044	0.048	0.057	0.059	0.056	

$N = 100, \sigma_{\eta}^2 = 1$															
	α							β							
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	
$T = 10$															
Bias	-0.150	0.041	-0.174	-0.077	0.001	-0.033	-0.068	0.006	0.046	-0.177	-0.097	0.009	-0.041	-0.093	
IQR	0.034	0.662	0.257	0.097	0.134	0.110	0.093	0.098	0.758	0.374	0.210	0.299	0.254	0.222	
MAE	0.150	0.323	0.178	0.078	0.067	0.060	0.072	0.050	0.381	0.204	0.126	0.151	0.126	0.126	
Size	1.000	0.001	0.088	0.186	0.044	0.068	0.165	0.054	0.004	0.036	0.086	0.035	0.060	0.090	
$T = 25$															
Bias	-0.054	0.057	-0.064	-0.020	-0.002	-0.006	-0.013	0.022	0.038	-0.026	-0.011	0.002	-0.001	-0.008	
IQR	0.016	0.342	0.114	0.039	0.033	0.031	0.029	0.055	0.235	0.143	0.087	0.089	0.084	0.078	
MAE	0.054	0.178	0.069	0.024	0.017	0.016	0.017	0.033	0.124	0.073	0.045	0.044	0.041	0.041	
Size	0.995	0.010	0.062	0.096	0.044	0.051	0.082	0.083	0.018	0.041	0.065	0.057	0.056	0.059	
$T = 50$															
Bias	-0.025	0.040	-0.029	-0.008	0.000	-0.001	-0.004	0.016	0.009	-0.004	-0.002	0.001	0.001	0.000	
IQR	0.011	0.237	0.067	0.023	0.017	0.016	0.016	0.038	0.105	0.078	0.053	0.046	0.045	0.045	
MAE	0.025	0.126	0.037	0.013	0.008	0.008	0.008	0.022	0.053	0.039	0.026	0.023	0.023	0.022	
Size	0.910	0.033	0.050	0.082	0.060	0.063	0.075	0.085	0.021	0.046	0.058	0.052	0.055	0.052	
$T = 100$															
Bias	-0.012	0.020	-0.015	-0.003	0.000	-0.001	-0.001	0.009	0.000	0.002	-0.001	0.000	0.000	0.000	
IQR	0.007	0.163	0.046	0.015	0.009	0.009	0.009	0.028	0.058	0.052	0.037	0.031	0.031	0.031	
MAE	0.012	0.083	0.024	0.008	0.004	0.004	0.004	0.015	0.029	0.026	0.018	0.015	0.015	0.015	
Size	0.668	0.060	0.059	0.069	0.050	0.046	0.051	0.075	0.021	0.054	0.050	0.051	0.052	0.052	

$N = 250, \sigma_{\eta}^2 = 1$															
	α							β							
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	
$T = 10$															
Bias	-0.150	0.038	-0.157	-0.044	0.002	-0.016	-0.037	0.004	0.047	-0.161	-0.054	0.002	-0.024	-0.054	
IQR	0.022	0.572	0.217	0.068	0.084	0.077	0.069	0.060	0.654	0.271	0.145	0.179	0.168	0.155	
MAE	0.150	0.286	0.166	0.049	0.042	0.040	0.044	0.031	0.327	0.176	0.081	0.090	0.086	0.086	
Size	1.000	0.001	0.073	0.118	0.056	0.056	0.108	0.041	0.003	0.038	0.074	0.050	0.052	0.065	
$T = 25$															
Bias	-0.054	0.029	-0.051	-0.011	0.000	-0.002	-0.005	0.021	0.019	-0.025	-0.009	-0.001	-0.003	-0.005	
IQR	0.011	0.274	0.099	0.028	0.021	0.020	0.020	0.035	0.176	0.096	0.058	0.055	0.051	0.050	
MAE	0.054	0.140	0.060	0.015	0.011	0.010	0.010	0.024	0.092	0.048	0.030	0.028	0.026	0.026	
Size	1.000	0.014	0.043	0.082	0.049	0.048	0.062	0.130	0.012	0.038	0.062	0.052	0.049	0.051	
$T = 50$															
Bias	-0.025	0.014	-0.023	-0.004	0.000	-0.001	-0.002	0.015	0.001	-0.005	-0.002	0.000	0.000	0.000	
IQR	0.007	0.182	0.062	0.017	0.011	0.010	0.010	0.025	0.067	0.051	0.035	0.030	0.029	0.029	
MAE	0.025	0.089	0.035	0.009	0.005	0.005	0.005	0.016	0.034	0.026	0.018	0.015	0.015	0.015	
Size	1.000	0.031	0.052	0.060	0.048	0.060	0.057	0.140	0.023	0.044	0.051	0.051	0.052	0.052	
$T = 100$															
Bias	-0.012	0.004	-0.013	-0.002	0.000	0.000	-0.001	0.009	0.001	0.001	0.001	0.000	0.001	0.000	
IQR	0.004	0.112	0.042	0.010	0.006	0.006	0.006	0.017	0.036	0.035	0.023	0.019	0.018	0.018	
MAE	0.012	0.053	0.022	0.005	0.003	0.003	0.003	0.011	0.018	0.017	0.011	0.009	0.009	0.009	
Size	0.958	0.059	0.059	0.058	0.054	0.048	0.051	0.099	0.025	0.044	0.050	0.056	0.055	0.056	

Table 2(cont.): Fixed effects model: $\alpha = 0.8$, $\beta = 1.0$ $N = 50, \sigma_{\eta}^2 = 5$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.150	0.054	-0.199	-0.109	0.004	-0.053	-0.101	0.006	0.037	-0.120	-0.115	0.013	-0.074	-0.122
IQR	0.049	0.651	0.260	0.112	0.188	0.143	0.112	0.137	1.000	0.597	0.269	0.429	0.335	0.279
MAE	0.150	0.325	0.202	0.110	0.094	0.079	0.102	0.068	0.500	0.299	0.161	0.210	0.168	0.159
Size	0.995	0.000	0.079	0.244	0.029	0.085	0.215	0.062	0.005	0.037	0.104	0.031	0.057	0.094
$T = 25$														
Bias	-0.054	0.067	-0.077	-0.032	-0.002	-0.010	-0.022	0.021	0.045	-0.032	-0.025	0.000	-0.008	-0.016
IQR	0.024	0.380	0.121	0.049	0.049	0.045	0.040	0.084	0.459	0.287	0.131	0.122	0.116	0.112
MAE	0.054	0.199	0.081	0.034	0.025	0.024	0.025	0.043	0.237	0.143	0.067	0.061	0.058	0.057
Size	0.908	0.015	0.056	0.140	0.057	0.064	0.124	0.081	0.007	0.053	0.075	0.052	0.061	0.064
$T = 50$														
Bias	-0.026	0.062	-0.036	-0.012	-0.001	-0.003	-0.007	0.017	0.019	-0.003	-0.004	0.002	0.001	0.001
IQR	0.015	0.289	0.075	0.029	0.023	0.022	0.021	0.055	0.267	0.184	0.077	0.066	0.065	0.065
MAE	0.026	0.160	0.044	0.017	0.012	0.011	0.011	0.029	0.134	0.093	0.039	0.033	0.033	0.032
Size	0.664	0.030	0.057	0.089	0.056	0.063	0.080	0.079	0.012	0.060	0.061	0.063	0.061	0.062
$T = 100$														
Bias	-0.012	0.051	-0.015	-0.005	0.000	-0.001	-0.002	0.009	0.010	0.004	0.000	0.001	0.001	0.002
IQR	0.010	0.218	0.048	0.019	0.013	0.012	0.012	0.039	0.163	0.113	0.053	0.042	0.041	0.040
MAE	0.012	0.117	0.024	0.010	0.006	0.006	0.006	0.020	0.081	0.057	0.026	0.021	0.020	0.020
Size	0.400	0.065	0.058	0.074	0.056	0.059	0.064	0.071	0.012	0.047	0.050	0.057	0.059	0.056

 $N = 100, \sigma_{\eta}^2 = 5$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.150	0.031	-0.189	-0.078	0.001	-0.033	-0.068	0.006	0.038	-0.169	-0.110	0.009	-0.041	-0.093
IQR	0.034	0.671	0.271	0.097	0.134	0.110	0.093	0.098	0.907	0.515	0.229	0.299	0.254	0.222
MAE	0.150	0.334	0.195	0.080	0.067	0.060	0.072	0.050	0.453	0.270	0.140	0.151	0.126	0.126
Size	1.000	0.001	0.085	0.188	0.044	0.068	0.165	0.054	0.006	0.034	0.091	0.035	0.060	0.090
$T = 25$														
Bias	-0.054	0.063	-0.071	-0.021	-0.002	-0.006	-0.013	0.022	0.051	-0.030	-0.017	0.002	-0.001	-0.008
IQR	0.016	0.344	0.122	0.039	0.033	0.031	0.029	0.055	0.366	0.237	0.094	0.089	0.084	0.078
MAE	0.054	0.184	0.075	0.025	0.017	0.016	0.017	0.033	0.185	0.119	0.046	0.044	0.041	0.041
Size	0.995	0.012	0.054	0.099	0.044	0.051	0.082	0.083	0.013	0.038	0.067	0.057	0.056	0.059
$T = 50$														
Bias	-0.025	0.046	-0.034	-0.008	0.000	-0.001	-0.004	0.016	0.022	-0.003	-0.004	0.001	0.001	0.000
IQR	0.011	0.245	0.072	0.024	0.017	0.016	0.016	0.038	0.187	0.138	0.057	0.046	0.045	0.045
MAE	0.025	0.133	0.041	0.013	0.008	0.008	0.008	0.022	0.097	0.068	0.029	0.023	0.023	0.022
Size	0.910	0.032	0.043	0.080	0.060	0.063	0.075	0.085	0.017	0.046	0.053	0.052	0.055	0.052
$T = 100$														
Bias	-0.012	0.024	-0.016	-0.003	0.000	-0.001	-0.001	0.009	0.006	0.000	-0.001	0.000	0.000	0.000
IQR	0.007	0.177	0.048	0.015	0.009	0.009	0.009	0.028	0.111	0.086	0.039	0.031	0.031	0.031
MAE	0.012	0.091	0.025	0.008	0.004	0.004	0.004	0.015	0.056	0.043	0.020	0.015	0.015	0.015
Size	0.668	0.057	0.050	0.069	0.050	0.046	0.051	0.075	0.019	0.054	0.056	0.051	0.052	0.052

 $N = 250, \sigma_{\eta}^2 = 5$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.150	0.039	-0.192	-0.047	0.002	-0.016	-0.037	0.004	0.065	-0.192	-0.064	0.002	-0.024	-0.054
IQR	0.022	0.580	0.248	0.069	0.084	0.077	0.069	0.060	0.742	0.395	0.166	0.179	0.168	0.155
MAE	0.150	0.286	0.197	0.051	0.042	0.040	0.044	0.031	0.378	0.230	0.091	0.090	0.086	0.086
Size	1.000	0.001	0.065	0.125	0.056	0.056	0.108	0.041	0.006	0.032	0.080	0.050	0.052	0.065
$T = 25$														
Bias	-0.054	0.034	-0.068	-0.011	0.000	-0.002	-0.005	0.021	0.030	-0.036	-0.010	-0.001	-0.003	-0.005
IQR	0.011	0.286	0.117	0.028	0.021	0.020	0.020	0.035	0.253	0.162	0.064	0.055	0.051	0.050
MAE	0.054	0.151	0.075	0.016	0.011	0.010	0.010	0.024	0.133	0.084	0.032	0.028	0.026	0.026
Size	1.000	0.013	0.040	0.086	0.049	0.048	0.062	0.130	0.010	0.031	0.063	0.052	0.049	0.051
$T = 50$														
Bias	-0.025	0.017	-0.030	-0.004	0.000	-0.001	-0.002	0.015	0.006	-0.010	-0.002	0.000	0.000	0.000
IQR	0.007	0.190	0.072	0.017	0.011	0.010	0.010	0.025	0.123	0.090	0.039	0.030	0.029	0.029
MAE	0.025	0.096	0.039	0.009	0.005	0.005	0.005	0.016	0.062	0.045	0.020	0.015	0.015	0.015
Size	1.000	0.031	0.045	0.060	0.048	0.060	0.057	0.140	0.021	0.035	0.059	0.051	0.052	0.052
$T = 100$														
Bias	-0.012	0.006	-0.016	-0.001	0.000	0.000	-0.001	0.009	0.004	0.001	0.000	0.000	0.001	0.000
IQR	0.004	0.118	0.045	0.010	0.006	0.006	0.006	0.017	0.064	0.061	0.025	0.019	0.018	0.018
MAE	0.012	0.057	0.025	0.005	0.003	0.003	0.003	0.011	0.032	0.031	0.012	0.009	0.009	0.009
Size	0.958	0.057	0.058	0.060	0.054	0.048	0.051	0.099	0.026	0.045	0.052	0.056	0.055	0.056

Table 3: Trend model: $\alpha = 0.4$, $\beta = 1.0$, $N = 50$, $\sigma_\eta^2 = 1$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.247	-0.025	-0.071	-0.113	-0.024	-0.129	-0.231	0.009	-0.034	-0.175	-0.178	-0.139	-0.167	-0.170
IQR	0.062	0.636	0.173	0.123	0.510	0.261	0.193	0.160	3.425	0.757	0.465	2.720	0.963	0.613
MAE	0.247	0.321	0.099	0.115	0.259	0.160	0.231	0.080	1.724	0.386	0.260	1.389	0.500	0.331
Size	1.000	0.007	0.083	0.242	0.014	0.117	0.369	0.063	0.003	0.054	0.079	0.003	0.038	0.066
$T = 25$														
Bias	-0.092	0.022	-0.031	-0.035	-0.002	-0.020	-0.041	0.034	0.075	-0.079	-0.052	0.010	-0.022	-0.031
IQR	0.032	0.737	0.096	0.057	0.092	0.076	0.063	0.089	2.306	0.316	0.182	0.336	0.266	0.201
MAE	0.092	0.367	0.053	0.039	0.046	0.041	0.045	0.049	1.162	0.165	0.096	0.168	0.133	0.102
Size	0.958	0.001	0.067	0.123	0.052	0.073	0.140	0.094	0.001	0.051	0.064	0.050	0.056	0.059
$T = 50$														
Bias	-0.045	0.046	-0.017	-0.014	0.001	-0.005	-0.012	0.024	0.118	-0.035	-0.012	0.001	-0.002	-0.003
IQR	0.022	0.740	0.072	0.037	0.040	0.038	0.035	0.057	1.811	0.180	0.094	0.113	0.099	0.091
MAE	0.045	0.377	0.037	0.020	0.020	0.019	0.019	0.033	0.907	0.089	0.047	0.057	0.050	0.046
Size	0.758	0.001	0.078	0.090	0.051	0.055	0.086	0.095	0.003	0.065	0.055	0.060	0.056	0.058
$T = 100$														
Bias	-0.021	0.077	-0.008	-0.006	0.000	-0.002	-0.005	0.012	0.169	-0.016	-0.003	0.001	0.001	0.001
IQR	0.015	0.666	0.052	0.022	0.022	0.021	0.021	0.040	1.359	0.112	0.058	0.059	0.056	0.054
MAE	0.021	0.344	0.026	0.012	0.011	0.010	0.011	0.021	0.717	0.057	0.029	0.029	0.028	0.027
Size	0.456	0.007	0.060	0.071	0.055	0.055	0.063	0.082	0.007	0.060	0.050	0.056	0.050	0.056

 $N = 100$, $\sigma_\eta^2 = 1$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.245	-0.038	-0.054	-0.077	-0.009	-0.087	-0.168	0.013	-0.145	-0.182	-0.190	-0.024	-0.179	-0.198
IQR	0.045	0.605	0.141	0.104	0.373	0.214	0.160	0.120	3.100	0.674	0.455	2.103	0.855	0.565
MAE	0.245	0.303	0.079	0.081	0.186	0.124	0.169	0.061	1.551	0.351	0.253	1.056	0.457	0.321
Size	1.000	0.004	0.061	0.165	0.013	0.077	0.275	0.056	0.003	0.040	0.079	0.004	0.040	0.070
$T = 25$														
Bias	-0.090	0.025	-0.026	-0.023	0.001	-0.010	-0.024	0.035	0.129	-0.069	-0.040	0.010	-0.008	-0.026
IQR	0.023	0.707	0.085	0.047	0.063	0.056	0.050	0.060	2.173	0.284	0.151	0.235	0.195	0.163
MAE	0.090	0.344	0.047	0.029	0.032	0.029	0.031	0.040	1.076	0.154	0.081	0.118	0.099	0.081
Size	0.999	0.002	0.058	0.083	0.051	0.053	0.095	0.114	0.001	0.055	0.059	0.044	0.052	0.057
$T = 50$														
Bias	-0.044	0.061	-0.014	-0.008	0.001	-0.002	-0.006	0.023	0.144	-0.033	-0.010	0.000	-0.002	-0.002
IQR	0.016	0.689	0.060	0.029	0.029	0.027	0.026	0.038	1.690	0.152	0.075	0.083	0.077	0.073
MAE	0.044	0.357	0.034	0.016	0.014	0.013	0.014	0.026	0.870	0.084	0.038	0.042	0.039	0.036
Size	0.956	0.001	0.050	0.083	0.058	0.060	0.074	0.113	0.001	0.046	0.052	0.060	0.052	0.054
$T = 100$														
Bias	-0.022	0.085	-0.008	-0.003	0.000	-0.001	-0.002	0.012	0.187	-0.014	-0.003	0.000	0.000	0.000
IQR	0.011	0.668	0.048	0.018	0.015	0.015	0.014	0.029	1.393	0.102	0.043	0.042	0.039	0.039
MAE	0.022	0.344	0.024	0.009	0.008	0.007	0.007	0.017	0.707	0.050	0.021	0.021	0.019	0.019
Size	0.750	0.003	0.059	0.061	0.058	0.055	0.064	0.097	0.002	0.054	0.052	0.058	0.055	0.054

 $N = 250$, $\sigma_\eta^2 = 1$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.247	-0.019	-0.035	-0.048	-0.002	-0.045	-0.099	0.011	-0.059	-0.144	-0.167	-0.060	-0.150	-0.188
IQR	0.027	0.488	0.115	0.084	0.225	0.161	0.134	0.073	2.367	0.577	0.417	1.323	0.756	0.531
MAE	0.247	0.245	0.063	0.055	0.113	0.083	0.104	0.037	1.179	0.298	0.237	0.664	0.388	0.296
Size	1.000	0.003	0.040	0.113	0.020	0.065	0.198	0.050	0.001	0.035	0.069	0.010	0.025	0.059
$T = 25$														
Bias	-0.090	0.015	-0.017	-0.013	0.001	-0.004	-0.011	0.034	0.073	-0.053	-0.028	0.006	-0.002	-0.014
IQR	0.014	0.571	0.071	0.037	0.039	0.036	0.035	0.039	1.775	0.247	0.123	0.149	0.129	0.119
MAE	0.090	0.288	0.037	0.020	0.020	0.019	0.020	0.034	0.896	0.125	0.064	0.075	0.064	0.061
Size	1.000	0.001	0.060	0.091	0.050	0.054	0.077	0.221	0.001	0.053	0.066	0.051	0.052	0.053
$T = 50$														
Bias	-0.044	0.041	-0.010	-0.005	-0.001	-0.002	-0.004	0.022	0.106	-0.024	-0.006	0.000	-0.001	-0.002
IQR	0.010	0.645	0.054	0.019	0.018	0.018	0.017	0.026	1.544	0.140	0.053	0.052	0.050	0.047
MAE	0.044	0.322	0.027	0.010	0.009	0.009	0.009	0.022	0.789	0.069	0.027	0.026	0.025	0.024
Size	1.000	0.001	0.059	0.058	0.052	0.050	0.053	0.217	0.001	0.057	0.051	0.046	0.050	0.049
$T = 100$														
Bias	-0.021	0.071	-0.007	-0.001	0.000	0.000	-0.001	0.012	0.150	-0.013	-0.001	0.001	0.000	0.001
IQR	0.007	0.675	0.040	0.012	0.010	0.009	0.009	0.018	1.404	0.085	0.028	0.025	0.024	0.025
MAE	0.021	0.341	0.021	0.006	0.005	0.005	0.005	0.013	0.700	0.045	0.014	0.013	0.012	0.012
Size	0.987	0.003	0.063	0.063	0.052	0.054	0.058	0.152	0.003	0.058	0.050	0.051	0.048	0.048

Table 3(cont.): Trend model: $\alpha = 0.4$, $\beta = 1.0$, $N = 50$, $\sigma_{\eta}^2 = 5$, $\sigma_{\lambda}^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.247	-0.007	-0.106	-0.115	-0.024	-0.129	-0.231	0.009	0.054	-0.196	-0.190	-0.139	-0.167	-0.170
IQR	0.062	0.632	0.207	0.125	0.510	0.261	0.193	0.160	3.398	0.782	0.472	2.720	0.963	0.613
MAE	0.247	0.320	0.124	0.116	0.259	0.160	0.231	0.080	1.665	0.417	0.269	1.389	0.500	0.331
Size	1.000	0.005	0.084	0.249	0.014	0.117	0.369	0.063	0.002	0.054	0.077	0.003	0.038	0.066
$T = 25$														
Bias	-0.092	0.025	-0.033	-0.036	-0.002	-0.020	-0.041	0.034	0.084	-0.079	-0.055	0.010	-0.022	-0.031
IQR	0.032	0.717	0.103	0.057	0.092	0.076	0.063	0.089	2.242	0.330	0.191	0.336	0.266	0.201
MAE	0.092	0.362	0.055	0.040	0.046	0.041	0.045	0.049	1.120	0.170	0.102	0.168	0.133	0.102
Size	0.958	0.001	0.067	0.120	0.052	0.073	0.140	0.094	0.001	0.053	0.067	0.050	0.056	0.059
$T = 50$														
Bias	-0.045	0.046	-0.017	-0.014	0.001	-0.005	-0.012	0.024	0.128	-0.039	-0.014	0.001	-0.002	-0.003
IQR	0.022	0.733	0.075	0.038	0.040	0.038	0.035	0.057	1.801	0.180	0.098	0.113	0.099	0.091
MAE	0.045	0.372	0.038	0.021	0.020	0.019	0.019	0.033	0.910	0.092	0.050	0.057	0.050	0.046
Size	0.758	0.002	0.081	0.094	0.051	0.055	0.086	0.095	0.002	0.066	0.055	0.060	0.056	0.058
$T = 100$														
Bias	-0.021	0.078	-0.009	-0.006	0.000	-0.002	-0.005	0.012	0.168	-0.018	-0.004	0.001	0.001	0.001
IQR	0.015	0.656	0.053	0.023	0.022	0.021	0.021	0.040	1.349	0.111	0.058	0.059	0.056	0.054
MAE	0.021	0.343	0.026	0.012	0.011	0.010	0.011	0.021	0.706	0.058	0.030	0.029	0.028	0.027
Size	0.456	0.007	0.058	0.072	0.055	0.055	0.063	0.082	0.007	0.062	0.054	0.056	0.050	0.056

 $N = 100$, $\sigma_{\eta}^2 = 5$, $\sigma_{\lambda}^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.245	-0.023	-0.078	-0.081	-0.009	-0.087	-0.168	0.013	-0.101	-0.228	-0.195	-0.024	-0.179	-0.198
IQR	0.045	0.632	0.171	0.108	0.373	0.214	0.160	0.120	3.055	0.716	0.466	2.103	0.855	0.565
MAE	0.245	0.320	0.100	0.083	0.186	0.124	0.169	0.061	1.500	0.385	0.261	1.056	0.457	0.321
Size	1.000	0.002	0.067	0.173	0.013	0.077	0.275	0.056	0.003	0.038	0.090	0.004	0.040	0.070
$T = 25$														
Bias	-0.090	0.029	-0.030	-0.025	0.001	-0.010	-0.024	0.035	0.111	-0.079	-0.050	0.010	-0.008	-0.026
IQR	0.023	0.705	0.095	0.049	0.063	0.056	0.050	0.060	2.207	0.295	0.167	0.235	0.195	0.163
MAE	0.090	0.354	0.052	0.031	0.032	0.029	0.031	0.040	1.086	0.164	0.089	0.118	0.099	0.081
Size	0.999	0.002	0.062	0.092	0.051	0.053	0.095	0.114	0.001	0.057	0.061	0.044	0.052	0.057
$T = 50$														
Bias	-0.044	0.060	-0.016	-0.009	0.001	-0.002	-0.006	0.023	0.151	-0.035	-0.012	0.000	-0.002	-0.002
IQR	0.016	0.682	0.064	0.031	0.029	0.027	0.026	0.038	1.682	0.157	0.082	0.083	0.077	0.073
MAE	0.044	0.350	0.035	0.016	0.014	0.013	0.014	0.026	0.844	0.083	0.041	0.042	0.039	0.036
Size	0.956	0.002	0.050	0.079	0.058	0.060	0.074	0.113	0.001	0.045	0.055	0.060	0.052	0.054
$T = 100$														
Bias	-0.022	0.083	-0.009	-0.003	0.000	-0.001	-0.002	0.012	0.182	-0.016	-0.003	0.000	0.000	0.000
IQR	0.011	0.676	0.049	0.018	0.015	0.015	0.014	0.029	1.394	0.102	0.044	0.042	0.039	0.039
MAE	0.022	0.344	0.025	0.009	0.008	0.007	0.007	0.017	0.703	0.051	0.022	0.021	0.019	0.019
Size	0.750	0.003	0.062	0.060	0.058	0.055	0.064	0.097	0.002	0.052	0.051	0.058	0.055	0.054

 $N = 250$, $\sigma_{\eta}^2 = 5$, $\sigma_{\lambda}^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.247	-0.009	-0.057	-0.053	-0.002	-0.045	-0.099	0.011	0.012	-0.205	-0.199	-0.060	-0.150	-0.188
IQR	0.027	0.534	0.145	0.090	0.225	0.161	0.134	0.073	2.544	0.670	0.423	1.323	0.756	0.531
MAE	0.247	0.267	0.084	0.059	0.113	0.083	0.104	0.037	1.266	0.354	0.261	0.664	0.388	0.296
Size	1.000	0.001	0.048	0.123	0.020	0.065	0.198	0.050	0.001	0.035	0.075	0.010	0.025	0.059
$T = 25$														
Bias	-0.090	0.024	-0.023	-0.015	0.001	-0.004	-0.011	0.034	0.094	-0.072	-0.038	0.006	-0.002	-0.014
IQR	0.014	0.572	0.083	0.039	0.039	0.036	0.035	0.039	1.775	0.265	0.142	0.149	0.129	0.119
MAE	0.090	0.288	0.043	0.022	0.020	0.019	0.020	0.034	0.889	0.143	0.076	0.075	0.064	0.061
Size	1.000	0.001	0.063	0.096	0.050	0.054	0.077	0.221	0.001	0.053	0.067	0.051	0.052	0.053
$T = 50$														
Bias	-0.044	0.043	-0.012	-0.005	-0.001	-0.002	-0.004	0.022	0.112	-0.030	-0.008	0.000	-0.001	-0.002
IQR	0.010	0.643	0.058	0.021	0.018	0.018	0.017	0.026	1.536	0.147	0.058	0.052	0.050	0.047
MAE	0.044	0.320	0.031	0.011	0.009	0.009	0.009	0.022	0.778	0.077	0.030	0.026	0.025	0.024
Size	1.000	0.001	0.060	0.061	0.052	0.050	0.053	0.217	0.002	0.056	0.058	0.046	0.050	0.049
$T = 100$														
Bias	-0.021	0.073	-0.008	-0.001	0.000	0.000	-0.001	0.012	0.147	-0.016	-0.001	0.001	0.000	0.001
IQR	0.007	0.682	0.043	0.012	0.010	0.009	0.009	0.018	1.412	0.091	0.030	0.025	0.024	0.025
MAE	0.021	0.342	0.022	0.006	0.005	0.005	0.005	0.013	0.702	0.047	0.015	0.013	0.012	0.012
Size	0.987	0.003	0.059	0.063	0.052	0.054	0.058	0.152	0.003	0.056	0.052	0.051	0.048	0.048

Table 4: Trend model: $\alpha = 0.8$, $\beta = 1.0$, $N = 50$, $\sigma_\eta^2 = 1$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.365	-0.091	-0.390	-0.301	-0.184	-0.370	-0.488	-0.099	0.125	-0.234	-0.379	-0.559	-0.406	-0.355
IQR	0.063	1.181	0.322	0.171	1.251	0.369	0.249	0.153	3.577	0.805	0.467	3.756	0.885	0.583
MAE	0.365	0.569	0.390	0.301	0.671	0.373	0.488	0.109	1.801	0.434	0.386	1.968	0.529	0.405
Size	1.000	0.004	0.275	0.682	0.010	0.303	0.782	0.162	0.002	0.050	0.207	0.004	0.067	0.138
$T = 25$														
Bias	-0.122	0.008	-0.084	-0.085	0.001	-0.086	-0.106	0.013	0.077	-0.131	-0.133	0.028	-0.156	-0.131
IQR	0.029	0.662	0.110	0.065	0.399	0.113	0.077	0.086	2.043	0.320	0.191	1.011	0.316	0.222
MAE	0.122	0.330	0.085	0.085	0.198	0.091	0.106	0.044	1.024	0.179	0.142	0.507	0.195	0.145
Size	1.000	0.001	0.120	0.422	0.002	0.168	0.492	0.071	0.002	0.064	0.159	0.004	0.107	0.135
$T = 50$														
Bias	-0.054	0.022	-0.037	-0.036	0.002	-0.025	-0.036	0.024	0.115	-0.070	-0.055	0.002	-0.037	-0.039
IQR	0.017	0.577	0.068	0.039	0.080	0.051	0.039	0.057	1.129	0.165	0.108	0.187	0.132	0.106
MAE	0.054	0.288	0.041	0.036	0.040	0.032	0.037	0.032	0.582	0.091	0.066	0.094	0.070	0.060
Size	0.995	0.001	0.111	0.272	0.034	0.093	0.256	0.094	0.007	0.079	0.122	0.036	0.065	0.090
$T = 100$														
Bias	-0.025	0.059	-0.019	-0.015	0.000	-0.006	-0.011	0.016	0.113	-0.029	-0.018	0.000	-0.004	-0.006
IQR	0.010	0.447	0.044	0.023	0.027	0.023	0.020	0.039	0.699	0.090	0.059	0.065	0.057	0.056
MAE	0.025	0.225	0.025	0.017	0.013	0.012	0.013	0.023	0.372	0.048	0.032	0.033	0.029	0.028
Size	0.918	0.006	0.088	0.157	0.049	0.072	0.138	0.092	0.007	0.074	0.075	0.057	0.067	0.067

 $N = 100$, $\sigma_\eta^2 = 1$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.362	-0.091	-0.352	-0.250	-0.187	-0.315	-0.428	-0.098	-0.002	-0.318	-0.446	-0.577	-0.498	-0.435
IQR	0.046	1.174	0.316	0.166	1.203	0.324	0.227	0.118	3.279	0.756	0.444	3.752	0.847	0.570
MAE	0.362	0.567	0.352	0.250	0.623	0.320	0.428	0.100	1.640	0.442	0.446	1.928	0.578	0.451
Size	1.000	0.000	0.236	0.572	0.007	0.254	0.735	0.245	0.000	0.052	0.275	0.001	0.086	0.172
$T = 25$														
Bias	-0.122	-0.009	-0.073	-0.075	0.003	-0.073	-0.091	0.014	0.109	-0.141	-0.149	0.025	-0.148	-0.156
IQR	0.020	0.735	0.103	0.060	0.280	0.113	0.074	0.060	1.923	0.304	0.172	0.755	0.295	0.195
MAE	0.122	0.367	0.076	0.075	0.140	0.079	0.091	0.032	0.971	0.174	0.151	0.375	0.181	0.159
Size	1.000	0.000	0.103	0.350	0.004	0.149	0.413	0.065	0.003	0.074	0.195	0.004	0.099	0.169
$T = 50$														
Bias	-0.054	0.042	-0.034	-0.032	0.000	-0.015	-0.028	0.023	0.122	-0.072	-0.059	0.004	-0.022	-0.039
IQR	0.012	0.592	0.060	0.036	0.060	0.043	0.034	0.039	1.145	0.145	0.096	0.140	0.108	0.091
MAE	0.054	0.299	0.039	0.032	0.030	0.024	0.028	0.026	0.575	0.090	0.066	0.070	0.056	0.052
Size	1.000	0.001	0.077	0.231	0.050	0.071	0.210	0.116	0.004	0.073	0.137	0.052	0.057	0.101
$T = 100$														
Bias	-0.025	0.055	-0.016	-0.011	0.000	-0.003	-0.007	0.015	0.104	-0.026	-0.016	-0.001	-0.003	-0.006
IQR	0.007	0.431	0.041	0.020	0.018	0.017	0.015	0.029	0.651	0.080	0.046	0.045	0.042	0.040
MAE	0.025	0.215	0.022	0.013	0.009	0.009	0.009	0.018	0.348	0.042	0.026	0.023	0.021	0.020
Size	0.998	0.003	0.064	0.114	0.051	0.063	0.092	0.126	0.009	0.069	0.074	0.062	0.062	0.062

 $N = 250$, $\sigma_\eta^2 = 1$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.365	-0.070	-0.322	-0.211	-0.159	-0.258	-0.354	-0.100	-0.032	-0.405	-0.484	-0.504	-0.566	-0.535
IQR	0.029	0.952	0.296	0.154	1.178	0.317	0.212	0.070	2.496	0.659	0.426	3.560	0.817	0.531
MAE	0.365	0.475	0.322	0.211	0.607	0.270	0.354	0.100	1.249	0.463	0.484	1.858	0.607	0.536
Size	1.000	0.002	0.235	0.499	0.007	0.211	0.620	0.483	0.002	0.071	0.340	0.002	0.105	0.259
$T = 25$														
Bias	-0.121	-0.011	-0.059	-0.062	-0.004	-0.059	-0.076	0.013	0.076	-0.144	-0.154	0.001	-0.137	-0.159
IQR	0.013	0.713	0.093	0.059	0.176	0.099	0.066	0.038	1.746	0.258	0.176	0.469	0.257	0.174
MAE	0.121	0.355	0.063	0.062	0.088	0.065	0.076	0.021	0.866	0.165	0.155	0.235	0.161	0.160
Size	1.000	0.001	0.088	0.307	0.015	0.128	0.353	0.071	0.003	0.071	0.241	0.015	0.097	0.224
$T = 50$														
Bias	-0.054	0.034	-0.027	-0.024	-0.001	-0.009	-0.018	0.022	0.095	-0.056	-0.051	-0.001	-0.017	-0.031
IQR	0.008	0.570	0.055	0.030	0.035	0.030	0.026	0.025	1.094	0.129	0.081	0.085	0.074	0.064
MAE	0.054	0.288	0.033	0.025	0.017	0.016	0.019	0.022	0.561	0.073	0.056	0.043	0.038	0.038
Size	1.000	0.001	0.079	0.167	0.043	0.060	0.146	0.215	0.005	0.067	0.139	0.046	0.055	0.090
$T = 100$														
Bias	-0.025	0.051	-0.014	-0.007	0.000	-0.001	-0.003	0.015	0.096	-0.024	-0.011	0.001	-0.001	-0.003
IQR	0.005	0.458	0.035	0.015	0.012	0.010	0.010	0.018	0.735	0.067	0.035	0.029	0.028	0.027
MAE	0.025	0.240	0.020	0.009	0.006	0.005	0.006	0.015	0.388	0.038	0.019	0.015	0.014	0.014
Size	1.000	0.001	0.064	0.092	0.059	0.050	0.071	0.217	0.007	0.060	0.065	0.050	0.048	0.054

Table 4(cont.): Trend model: $\alpha = 0.8$, $\beta = 1.0$, $N = 50$, $\sigma_\eta^2 = 5$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.365	-0.057	-0.407	-0.302	-0.184	-0.370	-0.488	-0.099	0.106	-0.146	-0.376	-0.559	-0.406	-0.355
IQR	0.063	1.170	0.335	0.173	1.251	0.369	0.249	0.153	3.291	0.832	0.475	3.756	0.885	0.583
MAE	0.365	0.580	0.408	0.302	0.671	0.373	0.488	0.109	1.639	0.421	0.383	1.968	0.529	0.405
Size	1.000	0.004	0.280	0.683	0.010	0.303	0.782	0.162	0.001	0.042	0.198	0.004	0.067	0.138
$T = 25$														
Bias	-0.122	0.006	-0.103	-0.087	0.001	-0.086	-0.106	0.013	0.097	-0.100	-0.137	0.028	-0.156	-0.131
IQR	0.029	0.644	0.122	0.065	0.399	0.113	0.077	0.086	1.972	0.363	0.191	1.011	0.316	0.222
MAE	0.122	0.323	0.103	0.087	0.198	0.091	0.106	0.044	0.985	0.192	0.146	0.507	0.195	0.145
Size	1.000	0.001	0.134	0.431	0.002	0.168	0.492	0.071	0.002	0.050	0.165	0.004	0.107	0.135
$T = 50$														
Bias	-0.054	0.022	-0.044	-0.037	0.002	-0.025	-0.036	0.024	0.112	-0.075	-0.061	0.002	-0.037	-0.039
IQR	0.017	0.568	0.070	0.040	0.080	0.051	0.039	0.057	1.165	0.177	0.109	0.187	0.132	0.106
MAE	0.054	0.284	0.046	0.037	0.040	0.032	0.037	0.032	0.593	0.102	0.072	0.094	0.070	0.060
Size	0.995	0.001	0.112	0.279	0.034	0.093	0.256	0.094	0.006	0.073	0.115	0.036	0.065	0.090
$T = 100$														
Bias	-0.025	0.063	-0.020	-0.015	0.000	-0.006	-0.011	0.016	0.115	-0.032	-0.021	0.000	-0.004	-0.006
IQR	0.010	0.440	0.046	0.023	0.027	0.023	0.020	0.039	0.689	0.093	0.061	0.065	0.057	0.056
MAE	0.025	0.224	0.026	0.017	0.013	0.012	0.013	0.023	0.361	0.051	0.034	0.033	0.029	0.028
Size	0.918	0.006	0.095	0.155	0.049	0.072	0.138	0.092	0.008	0.074	0.073	0.057	0.067	0.067

 $N = 100$, $\sigma_\eta^2 = 5$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.362	-0.079	-0.380	-0.249	-0.187	-0.315	-0.428	-0.098	-0.064	-0.212	-0.441	-0.577	-0.498	-0.435
IQR	0.046	1.105	0.332	0.164	1.203	0.324	0.227	0.118	3.218	0.821	0.449	3.752	0.847	0.570
MAE	0.362	0.567	0.380	0.249	0.623	0.320	0.428	0.100	1.608	0.432	0.444	1.928	0.578	0.451
Size	1.000	0.000	0.237	0.569	0.007	0.254	0.735	0.245	0.000	0.032	0.275	0.001	0.086	0.172
$T = 25$														
Bias	-0.122	-0.007	-0.093	-0.077	0.003	-0.073	-0.091	0.014	0.100	-0.135	-0.159	0.025	-0.148	-0.156
IQR	0.020	0.707	0.118	0.062	0.280	0.113	0.074	0.060	1.870	0.337	0.180	0.755	0.295	0.195
MAE	0.122	0.353	0.094	0.077	0.140	0.079	0.091	0.032	0.955	0.188	0.161	0.375	0.181	0.159
Size	1.000	0.000	0.118	0.372	0.004	0.149	0.413	0.065	0.002	0.062	0.203	0.004	0.099	0.169
$T = 50$														
Bias	-0.054	0.043	-0.040	-0.033	0.000	-0.015	-0.028	0.023	0.117	-0.081	-0.068	0.004	-0.022	-0.039
IQR	0.012	0.583	0.065	0.037	0.060	0.043	0.034	0.039	1.156	0.163	0.104	0.140	0.108	0.091
MAE	0.054	0.296	0.043	0.034	0.030	0.024	0.028	0.026	0.582	0.099	0.073	0.070	0.056	0.052
Size	1.000	0.001	0.085	0.250	0.050	0.071	0.210	0.116	0.004	0.069	0.153	0.052	0.057	0.101
$T = 100$														
Bias	-0.025	0.054	-0.018	-0.012	0.000	-0.003	-0.007	0.015	0.106	-0.030	-0.020	-0.001	-0.003	-0.006
IQR	0.007	0.430	0.043	0.021	0.018	0.017	0.015	0.029	0.653	0.084	0.050	0.045	0.042	0.040
MAE	0.025	0.219	0.024	0.013	0.009	0.009	0.009	0.018	0.347	0.045	0.029	0.023	0.021	0.020
Size	0.998	0.003	0.068	0.115	0.051	0.063	0.092	0.126	0.009	0.064	0.077	0.062	0.062	0.062

 $N = 250$, $\sigma_\eta^2 = 5$, $\sigma_\lambda^2 = 1$

	α							β						
	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3	FE	IV LEV	GMM LEV1	GMM LEV3	IV BOD	GMM BOD1	GMM BOD3
$T = 10$														
Bias	-0.365	-0.065	-0.361	-0.213	-0.159	-0.258	-0.354	-0.100	-0.028	-0.311	-0.508	-0.504	-0.566	-0.535
IQR	0.029	0.949	0.318	0.153	1.178	0.317	0.212	0.070	2.545	0.770	0.426	3.560	0.817	0.531
MAE	0.365	0.469	0.361	0.213	0.607	0.270	0.354	0.100	1.269	0.455	0.509	1.858	0.607	0.536
Size	1.000	0.002	0.228	0.500	0.007	0.211	0.620	0.483	0.001	0.042	0.351	0.002	0.105	0.259
$T = 25$														
Bias	-0.121	0.003	-0.080	-0.068	-0.004	-0.059	-0.076	0.013	0.088	-0.175	-0.175	0.001	-0.137	-0.159
IQR	0.013	0.692	0.111	0.061	0.176	0.099	0.066	0.038	1.698	0.314	0.180	0.469	0.257	0.174
MAE	0.121	0.346	0.082	0.068	0.088	0.065	0.076	0.021	0.845	0.197	0.175	0.235	0.161	0.160
Size	1.000	0.000	0.082	0.335	0.015	0.128	0.353	0.071	0.002	0.070	0.261	0.015	0.097	0.224
$T = 50$														
Bias	-0.054	0.039	-0.037	-0.028	-0.001	-0.009	-0.018	0.022	0.099	-0.079	-0.064	-0.001	-0.017	-0.031
IQR	0.008	0.572	0.065	0.033	0.035	0.030	0.026	0.025	1.121	0.152	0.092	0.085	0.074	0.064
MAE	0.054	0.287	0.042	0.028	0.017	0.016	0.019	0.022	0.564	0.091	0.067	0.043	0.038	0.038
Size	1.000	0.001	0.079	0.185	0.043	0.060	0.146	0.215	0.005	0.066	0.151	0.046	0.055	0.090
$T = 100$														
Bias	-0.025	0.050	-0.018	-0.008	0.000	-0.001	-0.003	0.015	0.096	-0.030	-0.013	0.001	-0.001	-0.003
IQR	0.005	0.457	0.040	0.017	0.012	0.010	0.010	0.018	0.739	0.075	0.042	0.029	0.028	0.027
MAE	0.025	0.239	0.024	0.010	0.006	0.005	0.006	0.015	0.386	0.044	0.023	0.015	0.014	0.014
Size	1.000	0.001	0.061	0.104	0.059	0.050	0.071	0.217	0.007	0.058	0.076	0.050	0.048	0.054