Improved accuracy of weak instrument robust GMM statistics through bootstrap and Edgeworth approximations

Frank Kleibergen*

University of Amsterdam

February 2019

Abstract

We construct higher order expressions of (weak instrument robust) generalized method of moment (GMM) test statistics for iid data. We use them to obtain Edgeworth approximations and to reveal sensitivity to instrument quality. The Edgeworth approximations show that usage of bootstrapped critical values reduces the order of the approximation error of the finite sample distribution of the weak instrument robust statistics compared to usage of asymptotic critical values. These results remain to hold when the instruments are weak and extend previous results on the bootstrap and the Edgeworth approximation. We illustrate the resulting reduction of size distortions and conduct a power comparison using a panel autoregressive model of order one.

JEL classification: C11, C20, C30

1 Introduction

Two common approaches for reducing size distortions of test statistics are to Edgeworth-correct asymptotic critical values, see *e.g.* Rothenberg (1984), and usage of bootstrapped critical values, see *e.g.* Horowitz (2001). In many cases, these two approaches remove the approximation error of the finite sample distribution

*Econometrics and Statistics Section, Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018WB Amsterdam, The Netherlands. Email: f.r.kleibergen@uva.nl.

of the statistics up to a higher order in the sample size than the limiting distribution. The size distortions that result when we use such critical values are then less than those that result from using the asymptotic ones. The regularity conditions under which the Edgeworth approximation and the bootstrap improve the approximation of the finite sample distribution request that the hypothesized parameters are well identified, see *e.g.* Bhattacharya and Ghosh (1978) and Horowitz (2001). Parameters that are estimated using the generalized method of moments (GMM) are not well identified when the instruments are weak, see *e.g.* Staiger and Stock (1997) and Stock and Wright (2000). This seems to indicate that the Edgeworth approximation and bootstrap do not improve the approximation of the finite sample distribution of statistics in GMM with weak instruments.

To overcome size distortions in GMM with weak instruments, statistics have been proposed whose limiting distributions are robust to the strength of the instruments, see *e.g.* Stock and Wright (2000), Kleibergen (2005) and Kleibergen and Mavroeidis (2009). The limiting distributions of these statistics apply under more general conditions than those of the traditional GMM statistics. They therefore lead to a better approximation of the finite sample distribution than the approximations that result from the limiting distribution for the traditional GMM statistics. We can further improve upon the approximation of the finite sample distribution of the weak instrument robust GMM statistics by constructing Edgeworth-corrections of the asymptotic critical values or by using bootstrapped critical values. We show that usage of either the Edgeworth-corrected or bootstrapped critical values reduces the order of the approximation error of the finite sample distribution compared to using asymptotic critical values. These improvements remain to hold when the instruments are weak which shows that Edgeworth approximations can be constructed even when the parameters are weakly identified. This is a new result and provides an extension to Bhattacharya and Ghosh (1978).

In a related article Moreira *et. al.* (2009) show the validity of the bootstrap for the weak instrument robust Lagrange multiplier statistic in the linear instrumental variables regression model with one endogenous variable. The bootstrap that they employ resamples the residuals that result after estimating the model. Because of the dependence of these residuals on the involved estimator, the approximation of the finite sample distribution of the Lagrange multiplier statistic by the bootstrap is as accurate as the approximation by the limiting distribution. Hence, there are no higher order improvements that result from the bootstrap. The bootstrap that we employ just resamples the residuals under the null hypothesis so it differs from the one used by Moreira *et. al.* (2010). The higher order improvements that we obtain from our bootstrap do therefore not contradict the results obtained in Moreira *et. al.* (2010).

The paper is organized as follows. In the second section, we introduce GMM and define our GMM statistics of interest: the GMM extension of the Anderson-Rubin statistic (GMM-AR) from Stock and Wright (2000), the GMM Lagrange multiplier (KLM) statistic of Kleibergen (2005), a GMM extension of Moreira's (2003) conditional likelihood ratio (GMM-MLR) statistic and the GMM LM statistic of Newey and West (1987). The second section also states the assumptions under which we derive our results. In the third section, we decompose the statistics in several components that are of a different order in the sample size. The fourth section provides algorithms to bootstrap our statistics of interest and decomposes the bootstrapped statistics into several components that are of a different order in the bootstrap sample size. The fifth section discusses Edgeworth approximations of the finite sample distributions of the orginal statistics and their bootstrapped counterparts. The Edgeworth approximations show the higher order improvement that results from the bootstrap. The sixth section illustrates the theoretical results by conducting a simulation experiment using a panel autoregressive model of order one with Arellano and Bond (1991) moment conditions. It shows that usage of bootstrapped or Edgeworth-corrected critical values reduces size distortion compared to usage of critical values that stem from the limiting distribution. The sixth section also conducts a power comparison. The seventh section briefly discusses further extensions. The eighth section concludes.

Throughout the paper we use the notation: I_m is the $m \times m$ identity matrix, $P_A = A(A'A)^{-1}A'$ for a full rank $n \times m$ matrix A and $M_A = I_n - P_A$. Furthermore, "$\underset{p}{\rightarrow}$" stands for convergence in probability, "$\underset{d}{\rightarrow}$" for convergence in distribution, E is the expectation operator and $\text{vec}(A)$ is the column vectorization of the matrix A.

2 Generalized Method of Moments

We consider the estimation of a scalar parameter θ whose parameter region is $\mathbb{R}$ and for which the $k \times 1$ dimensional moment equation

$$\mathrm{E}(f(\theta, Y_i)) = 0, \qquad i = 1, \ldots, N, \tag{1}$$

holds. We use a scalar parameter instead of a vector of parameters to simplify the analysis. We later show how the results extend to the multiple parameter case. The data vector Y_i is observed for individual/time i. The number of equations k exceeds or is equal to the number of parameters. The $k \times 1$ dimensional vector function f of θ is finite for finite values of θ, continuous and twice continuous differentiable. The unique value of θ, at which (1) holds, is equal to θ_0. To estimate θ in (1), we use Hansen's (1982) GMM.

For a sample of observations $(Y_i,\ i=1,\ldots,N)$, the objective function for the continuous updating estimator (CUE) of Hansen *et. al.* (1996) reads

$$Q(\theta) = Nf_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}f_N(\theta,Y), \tag{2}$$

with $f_N(\theta,Y) = \frac{1}{N}\sum_{i=1}^{N} f_i(\theta),\ f_i(\theta) = f(\theta,Y_i)$. The covariance matrix estimator $\hat{V}_{ff}(\theta)$ that we use in (2) is the Eicker-White covariance matrix estimator, see Eicker (1967) and White (1980). Usage of the Eicker-White covariance matrix estimator implies that we do not allow for dependence between the moments. This is done for expository purposes and we discuss how to deal with dependence between the moments later. We make extensive use of the derivative of the moment functions

$$q_N(\theta,Y) = \tfrac{1}{N}\textstyle\sum_{i=1}^{N} q_i(\theta), \tag{3}$$

with $q_i(\theta) = \frac{\partial}{\partial\theta'} f_i(\theta)$. We use the Eicker-White covariance matrix estimator as well to estimate the covariance between the moments and their derivatives. Thus we employ the covariance matrix estimators:

$$\begin{array}{rl}\hat{V}_{ff}(\theta) = & \frac{1}{N}\sum_{i=1}^{N} f_i(\theta)f_i(\theta)' - f_N(\theta,Y)f_N(\theta,Y)', \\ \hat{V}_{qf}(\theta) = & \frac{1}{N}\sum_{i=1}^{N} q_i(\theta)f_i(\theta)' - q_N(\theta,Y)f_N(\theta,Y)', \\ \hat{V}_{qq}(\theta) = & \frac{1}{N}\sum_{i=1}^{N} q_i(\theta)q_i(\theta)' - q_N(\theta,Y)q_N(\theta,Y)', \end{array} \tag{4}$$

of the covariance matrices $V_{ff}(\theta) = \mathrm{E}(\bar{f}_i(\theta)\bar{f}_i(\theta)')$, $V_{qf}(\theta) = \mathrm{E}(\bar{q}_i(\theta)\bar{f}_i(\theta)')$ and $V_{qq}(\theta) = \mathrm{E}(\bar{q}_i(\theta)\bar{q}_i(\theta)')$, where $\bar{f}_i(\theta) = f_i(\theta) - f_N(\theta,Y)$ and $\bar{q}_i(\theta) = q_i(\theta) - q_N(\theta,Y)$.

We determine the validity of the bootstrap and whether it leads to higher order improvements for four different GMM statistics not all of which are robust to weak instruments.

Definition 1. Four different statistics that test $\mathrm{H}_0 : \theta = \theta_0$ are:

1. The GMM-AR statistic, see Stock and Wright (2000), which is the generalization of the Anderson-Rubin statistic, see Anderson and Rubin (1949), towards GMM,

$$\text{GMM-AR}(\theta_0) = \ Nf_N(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}f_N(\theta_0,Y). \tag{5}$$

2. The KLM-statistic which is a GMM-Lagrange multiplier (LM) statistic based on the CUE, see Kleibergen (2005):

$$\begin{array}{l}\mathrm{KLM}(\theta_0) = Nf_N(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0,Y) \\ \quad \left[\hat{D}_N(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0,Y)\right]^{-1}\hat{D}_N(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}f_N(\theta_0,Y),\end{array} \tag{6}$$

with $\hat{D}_N(\theta_0,Y) = q_N(\theta,Y) - \hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}f_N(\theta_0,Y)$.

3. The GMM-MLR statistic which is Moreira's (2003) conditional likelihood ratio (LR) statistic applied in a GMM-setting, see Kleibergen (2005):

$$\begin{aligned}\text{GMM-MLR}(\theta_0) = \tfrac{1}{2}\Big[&\text{GMM-AR}(\theta_0) - \mathrm{r}(\theta_0) + \\ &\sqrt{(\text{GMM-AR}(\theta_0) + \mathrm{r}(\theta_0))^2 - 4\left[\text{GMM-AR}(\theta_0) - \text{KLM}(\theta_0)\right]\mathrm{r}(\theta_0))}\Big],\end{aligned} \tag{7}$$

with $\mathrm{r}(\theta_0) = N\hat{D}_N(\theta_0, Y)'\left[\hat{V}_{qq}(\theta_0) - \hat{V}_{qf}(\theta_0)\hat{V}_{ff}(\theta_0)^{-1}\hat{V}_{fq}(\theta_0)\right]^{-1}\hat{D}_N(\theta_0, Y)$.

4. The GMM-LM statistic, see Newey and West (1987):

$$\begin{aligned}\text{LM}(\theta_0) = \; & N f_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1}q_N(\theta_0, Y)\left[q_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1}q_N(\theta_0, Y)\right]^{-1} \\ & q_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1}f_N(\theta_0, Y).\end{aligned} \tag{8}$$

The above four statistics are used to test the parameters in models that are estimated by GMM. The GMM-LM statistic is the only one of the above four statistics whose limiting distribution is not robust to weak instruments. We use it to show the issues involved with GMM statistics that are not robust to weak instruments. Since these issues are identical for all non-robust statistics, we just discuss them for the one for which they are the most straightforward to obtain. This explains why we use the GMM-LM statistic instead of the more commonly used Wald statistic.

The limiting distributions of the statistics in Definition 1 result after an assumption on the moment vector and its derivative.

Assumption 1. *Under $H_0 : \theta = \theta_0$, the following assumptions hold jointly:*

a. *The vectors of moments and derivatives $(f_i(\theta_0)' \vdots q_i(\theta_0)')'$ are independent across individuals/time.*

b. *The eighth order moments of $f_i(\theta_0)$ and $q_i(\theta_0)$ are finite.*

Assumption 1a has been mentioned before and justifies usage of the Eicker-White covariance matrix estimator. Assumption 1b implies that the fourth order moment estimator of $f_i(\theta)$ satisfies a central limit theorem. It is somewhat overly restrictive when we just want to use the limiting distributions of the statistics in Definition 1 but we need it for their higher order expansions that we construct lateron.

Corollary 1. *For* $D_N(\theta_0, Y) = \frac{1}{N}\sum_{i=1}^N d_i(\theta_0)$, $d_i(\theta_0) : d_i(\theta_0) = q_i(\theta_0) - V_{qf}(\theta_0)$ $V_{ff}(\theta_0)^{-1} f_i(\theta_0)$, *it holds that under* $H_0 : \theta = \theta_0$ *and Assumption 1,* $\sqrt{N} f_N(\theta_0, Y)$ *and* $\sqrt{N}\left[D_N(\theta_0, Y) - J_\theta(\theta_0)\right]$, *with* $J_\theta(\theta_0)$ =E$[q_i(\theta_0)]$, *have independent normal limiting distributions.*

Proof. Results directly from Assumption 1 and $\mathrm{E}(f_i(\theta_0) d_i(\theta_0)') = 0$. ■

Corollary 1 shows that $D_N(\theta, Y)$, which is an (infeasible) estimator of the derivative of the average moment vector with respect to θ, is independent of the average moment vector in large samples.

Corollary 2. *Under* $H_0 : \theta = \theta_0$ *and Assumption 1, it holds in large samples that*

$$\begin{array}{ll} (i.) & \mathrm{E}\left[\frac{1}{N}\sum_{i=1}^N f_i(\theta_0) f_i(\theta_0)' | D_N(\theta_0, Y) = D\right] = V_{ff}(\theta_0) + O(\frac{1}{\sqrt{N}}) \\ (ii.) & \mathrm{E}\left[\frac{1}{N}\sum_{i=1}^N \left[(f_i(\theta_0) f_i(\theta_0)' \otimes f_i(\theta_0) f_i(\theta_0)'\right] | D_N(\theta_0, Y) = D\right] \\ & \quad = \mathrm{E}\left[f_i(\theta_0) f_i(\theta_0)' \otimes f_i(\theta_0) f_i(\theta_0)'\right] + O(\frac{1}{\sqrt{N}}), \end{array} \tag{9}$$

where $O(\frac{1}{\sqrt{N}})$ *indicates that the (non-stochastic) remainder term is of order* $\frac{1}{\sqrt{N}}$.

Proof. Because of Assumption 1, $D_N(\theta_0, Y)$, $\hat{V}_{ff}(\theta_0)$ and the fourth order moment estimator of $f_i(\theta_0)$ have a joint normal limiting distribution with a convergence rate proportional to $\sqrt{N}$. Hence, the order of covariance between these estimators is proportional to $\frac{1}{\sqrt{N}}$. The conditional expectations of them given one another are then equal to the unconditional expectations plus a term of order $\frac{1}{\sqrt{N}}$.

■

Corollary 2 implies that there is no zero-th order bias in the KLM and GMM-MLR statistics so their (conditional) limiting distributions are valid. We need the expressions of the conditional expectations given $D_N(\theta_0, Y)$ for the conditional higher order expansions given $D_N(\theta_0, Y)$ that we construct lateron. Corollary 2 obviously holds as well when the conditional expectations of $f_i(\theta_0) f_i(\theta_0)'$ and $(f_i(\theta_0) f_i(\theta_0)' \otimes f_i(\theta_0) f_i(\theta_0)')$ given $d_i(\theta)$ are equal to the unconditional expectations but Corollary 2 provides the same result under less restrictive assumptions. [1]

Corollary 1 shows that $D_N(\theta_0, Y)$ is an infeasible estimator of $J_\theta(\theta_0)$ which is in large samples independent of the average moment vector $f_N(\theta_0, Y)$. The KLM and GMM-MLR statistics in Definition 1 therefore use the feasible estimator of $D_N(\theta, Y)$, $\hat{D}_N(\theta, Y)$, that results in their (conditional) limiting distributions not being affected by weak instruments.

[1] Since $f_i(\theta)$ and $d_i(\theta)$ are uncorrelated, this conditional moment assumption is actually less restrictive then perceived at first sight.

Corollary 3. *Under $H_0 : \theta = \theta_0$ and Assumption 1, the limiting distributions of the statistics in Definition 1 are characterized by:*

$$\begin{aligned} \text{GMM-AR}\,(\theta_0) &\overset{d}{\to} \ \psi_1 + \psi_{k-1} \\ \text{KLM}\,(\theta_0) &\overset{d}{\to} \ \psi_1 \\ \text{GMM-MLR}(\theta_0)|\text{r}(\theta_0) = r) &\overset{d}{\to} \ \tfrac{1}{2}\Big[\psi_1 + \psi_{k-1} - r) + \\ &\quad \sqrt{\left(\psi_1 + \psi_{k-1} + r\right)^2 - 4\psi_{k-1}r)}\Big] \\ \text{GMM-LM}(\theta_0) &\overset{d}{\to} \ \psi_1, \ \textit{when } J_\theta(\theta_0) \textit{ does not equal zero,} \end{aligned} \tag{10}$$

with ψ_1 and ψ_{k-1} independent $\chi^2(1)$ and $\chi^2(k-1)$ distributed random variables.

Proof. see Newey and West (1987), Stock and Wright (2000) and Kleibergen (2005). ■

Corollary 3 shows that the (conditional) limiting distributions of the GMM-AR, KLM and GMM-MLR statistics do not depend on nuisance parameters. The bootstrap then typically provides a more accurate approximation of the finite sample distribution than the limiting distribution, see *e.g.* Horowitz (2001). This holds since the bootstrap removes some of the higher order approximation errors of the Edgeworth approximation of the finite sample distribution.

Alongside their usefullness for improving approximations of finite sample distributions of estimators, see *e.g.* Sargan (1976), Edgeworth approximations are used to approximate the finite sample distribution of Wald statistics. They then result from a Taylor approximation of the Wald statistic around the true or expected values of its different elements, see *e.g.* Bhattacharya and Ghosh (1978), Sargan (1980) and Phillips and Park (1988). To construct such a Taylor approximation, the derivatives of the statistic at these values have to be well defined. When we construct such a Taylor approximation for the KLM statistic, which is a function of $f_N(\theta_0, Y)$ and $\hat{D}_N(\theta_0, Y)$, the Taylor approximation uses its derivatives with respect to $f_N(\theta_0, Y)$ and $\hat{D}_N(\theta_0, Y)$ evaluated at their expected values, zero and $J_\theta(\theta_0)$ resp.. The derivative of the KLM statistic with respect to $\hat{D}_N(\theta_0, Y)$ is, however, not well-defined for zero values of $\hat{D}_N(\theta_0, Y)$ so we can not allow for zero values of its expectation, $J_\theta(\theta_0)$. This implies that the resulting Edgeworth approximation does not allow for weak or irrelevant instruments. Hence the Edgeworth approximations in Bhattacharya and Ghosh (1978), Sargan (1980) and Phillips and Park (1988) do not allow for weak or irrelevant instruments and can not be used to show that the bootstrap uniformly improves the approximation of the finite sample distributions of the KLM and GMM-MLR statistics.

The Edgeworth approximations from Bhattacharya and Ghosh (1978), Sargan (1980) and Phillips and Park (1988) are all for Wald statistics and do not exploit the independence between $f_N(\theta_0, Y)$ and $D_N(\theta_0, Y)$ in large samples. Since the

KLM and GMM-MLR statistics are Lagrange Multiplier and (quasi) likelihood-ratio statistics, we construct an alternative Edgeworth approximation that uses the independence between $f_N(\theta_0, Y)$ and $D_N(\theta_0, Y)$ in large samples. We use it to show that the bootstrap uniformly improves the approximation of the finite sample distribution of the KLM and GMM-MLR statistics compared to the (conditional) limiting distribution.

The bootstrap and Edgeworth approximations that we discuss lateron simplify considerably when an additional assumption holds.

Assumption 2. *Under $H_0 : \theta = \theta_0$, it holds that:*

$$\mathrm{E}(\mathrm{vec}(f_i(\theta_0)f_i(\theta_0)')d_i(\theta_0)') = \mathrm{vec}(V_{ff}(\theta_0))\mathrm{E}(d_i(\theta_0))' = \mathrm{vec}(V_{ff}(\theta_0))J(\theta_0)'. \tag{11}$$

Assumption 2 holds if any conditional heteroscedasticity of $f_i(\theta_0)$ does not depend on $d_i(\theta_0)$. It also holds when the conditional moment of $f_i(\theta_0)f_i(\theta_0)'$ given $d_i(\theta_0)$ equals the unconditional moment.

Corollary 4. *When Assumptions 1 and 2 hold, the limiting distributions of $\sqrt{N}(\hat{V}_{ff}(\theta_0) - V_{ff}(\theta_0))$ and $\sqrt{N}D_N(\theta_0, Y)$ are independent and*

$$\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^{N} \mathrm{vec}(f_i(\theta_0)f_i(\theta_0)')d_i(\theta_0)'|D_N(\theta_0, Y) = D) = \mathrm{vec}(V_{ff}(\theta_0))D' + O(\tfrac{1}{\sqrt{N}}). \tag{12}$$

Proof. Assumption 2 implies that the covariance between $\hat{V}_{ff}(\theta_0) - V_{ff}(\theta_0)$ and $D_N(\theta_0, Y) - J_\theta(\theta_0)$ is equal to zero. Since both $\hat{V}_{ff}(\theta_0)$ and $D_N(\theta_0, Y)$ have normal limiting distributions, the zero covariance implies that they are independent. Equation (12) follows along the lines of the proof of Corollary 2. ■

We construct the Edgeworth approximations, that indicate the higher order improvements that result from the bootstrap, using the higher order expressions of the statistics in Definition 1.

3 Higher order expressions

The higher order expressions of Wald statistics result from Taylor approximations. For example, for the Wald statistic that tests the hypothesis $\mathrm{H}_g : g(\theta) = 0$, with $g(\theta)$ a continuous differentiable function of θ:

$$\mathrm{W}(\hat{\theta}) = g(\hat{\theta})'(G(\hat{\theta})\hat{W}_\theta G(\hat{\theta})')^{-1}g(\hat{\theta}), \tag{13}$$

with $G(\theta) = \frac{\partial}{\partial\theta'}g(\theta)$ and $\hat{W}_\theta$ an estimator of the covariance matrix of the asymptotically normal estimator $\hat{\theta}$; these higher order approximations result from Taylor approximations of $g(\hat{\theta})$ and $G(\hat{\theta})$ around $g(\theta_0)$ and $G(\theta_0)$ resp., see *e.g.* Phillips and Park (1988). The higher order derivatives of $g(\theta)$ therefore have to be well defined at θ_0. The statistics in Definition 1 are not Wald statistics so we construct their higher order expressions in a different manner. The resulting expressions do not involve any higher order derivatives whose existence at the true value of the parameters might be questionable.

The higher order expressions of the statistics in Definition 1 result from the large sample behavior of the (covariance) estimators involved in them. We therefore express these estimators as equal to their true value plus an error term. For the GMM-AR statistic, this implies that we use that

$$\hat{V}_{ff}(\theta_0) = V_{ff}(\theta_0) + \hat{V}_{ff}(\theta_0) - V_{ff}(\theta_0), \tag{14}$$

while for the KLM statistic, we use both (14) and

$$\sqrt{N}\hat{D}_N(\theta, Y) = \sqrt{N}D_N(\theta, Y) - \sqrt{N}\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}f_N(\theta, Y) + O_p(\tfrac{1}{N}), \tag{15}$$

with $\hat{V}_{\theta f}(\theta) = \frac{1}{N}\sum_{i=1}^N \bar{d}_i(\theta)\bar{f}_i(\theta)'$, which is an (infeasible) estimator of the (zero) covariance between $d_i(\theta)$ and $f_i(\theta)$, $\bar{d}_i(\theta) = d_i(\theta) - D_N(\theta, Y)$ and $O_p(\frac{1}{N})$ shows that the (random) remainder term is of order $\frac{1}{N}$. These expressions are stated in Lemmas 1-7 in Appendix A.

Theorem 1 states the higher order expressions of the statistics in Definition 1. The order of the different components results from the convergence rate of the expectation of these components given $D_N(\theta_0, Y)$. We use these components to construct the Edgeworth approximation of the conditional distribution of the statistics given $D_N(\theta_0, Y)$. For the non-robust GMM-LM statistic, Theorem 1 just states some of the elements of its higher order expression. These elements suffice to show the dependence of the limiting distribution of the GMM-LM statistic on the strength of the instruments.

Theorem 1. *Under H_0 and Assumption 1, the higher order expressions of the statistics in Definition 1 read:*
1. *For the GMM-AR statistic* (5):

$$\text{GMM-AR}(\theta_0) = GMM\text{-}AR_0 + \tfrac{1}{N}GMM\text{-}AR_1 + O_p(\tfrac{1}{N^2}), \tag{16}$$

where

$$\begin{aligned} GMM\text{-}AR_0 =&\ Nf_N(\theta_0, Y)'V_{ff}(\theta_0)^{-1}f_N(\theta_0, Y) \underset{d}{\rightarrow} \chi^2(k) \\ GMM\text{-}AR_1 =&\ -N^2 f_N(\theta_0, Y)'V_{ff}(\theta_0)^{-1}\left[\hat{V}_{ff}(\theta_0) - V_{ff}(\theta_0)\right]V_{ff}(\theta_0)^{-1}f_N(\theta_0, Y) \\ \mathrm{E}(GMM\text{-}AR_1) =&\ -\tfrac{N-1}{N}\mathrm{vec}(V_{ff}(\theta_0)^{-1})'\mathrm{E}\left[(f_i(\theta_0)f_i(\theta_0)' \otimes f_i(\theta_0)f_i(\theta_0)')\right] \\ &\ \mathrm{vec}(V_{ff}(\theta_0)^{-1}) + \tfrac{N-1}{N}\left[k^2 + 2k\right] + k. \end{aligned} \tag{17}$$

2. *For the KLM-statistic* (6):

$$\text{KLM}(\theta_0) = KLM_0 + \tfrac{1}{N}\left[KLM_1 + KLM_2\right] + O_p(\tfrac{1}{N^2}), \tag{18}$$

where

$$\begin{array}{rl} KLM_0 = & Nf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y) \underset{d}{\to} \chi^2(1) \\ KLM_1 = & -N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\ & V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)- \\ & 2N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\ & V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y) \\ KLM_2 = & -2N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1} \\ & f_N(\theta,Y)\left[D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right]^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y), \end{array} \tag{19}$$

and the conditional expectations of these elements given $D_N(\theta_0,Y)$ *read:*

$$\begin{array}{l} \text{E}\left[KLM_0|D_N(\theta_0,Y)\right] = 1 \\ \text{E}\left[KLM_1|D_N(\theta_0,Y)\right] = -\frac{N-1}{N}\text{vec}(V_{ff}(\theta_0)^{-\frac{1}{2}}P_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}})' \\ \quad \text{E}(f_i(\theta_0)f_i(\theta_0)'\otimes f_i(\theta_0)f_i(\theta_0)')\text{vec}(V_{ff}(\theta_0)^{-1}- \\ \quad V_{ff}(\theta_0)^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}) + \frac{2(N-1)}{N}(k-1)+4-\frac{3}{N} \\ \text{E}\left[KLM_2|D_N(\theta_0,Y)\right] = -\frac{4(N-1)}{N}\text{vec}(M_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)})' \\ \quad \text{E}(\frac{1}{N}\sum_{i=1}^N V_{ff}(\theta_0)^{-\frac{1}{2}}f_i(\theta_0)f_i(\theta_0)'V_{ff}(\theta_0)^{-\frac{1}{2}}\otimes V_{ff}(\theta_0)^{-\frac{1}{2}}\bar{d}_i(\theta_0)|D_N(\theta_0,Y)) \\ \quad V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)\left[D_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}D_N(\theta_0,Y)\right]^{-1} - \frac{2}{N}\text{vec}(M_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)})' \\ \quad \text{E}(\frac{1}{N}\sum_{i=1}^N V_{ff}(\theta_0)^{-\frac{1}{2}}f_i(\theta_0)f_i(\theta_0)'V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'V_{ff}(\theta_0)^{-\frac{1}{2}}\otimes V_{ff}(\theta_0)^{-\frac{1}{2}}\bar{d}_i(\theta_0) \\ \quad |D_N(\theta_0,Y))\text{vec}(V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)\left[D_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}D_N(\theta_0,Y)\right]^{-1}). \end{array} \tag{20}$$

When Assumption 2 holds, the first element of the conditional expectation of KLM_2 *in (20) is equal to zero such that the conditional expectation of* KLM_2 *is of order* $\frac{1}{N}$.

3. *For the GMM-MLR statistic* (7) *given* $\text{r}(\theta_0)$:

$$\begin{array}{l} \text{GMM-MLR}(\theta_0) = \frac{1}{2}\Big[GMM\text{-}AR_0 - \text{r}(\theta_0) + \\ \quad \sqrt{(GMM\text{-}AR_0+\text{r}(\theta_0))^2 - 4\left[GMM\text{-}AR_0 - KLM_0\right]\text{r}(\theta_0))}\Big] + \\ \quad \frac{1}{2N}\left[1+\frac{GMM\text{-}AR_0-\text{r}(\theta_0)}{\sqrt{(GMM\text{-}AR_0+\text{r}(\theta_0))^2-4[GMM\text{-}AR_0-KLM_0]\text{r}(\theta_0))}}\right]GMM\text{-}AR_1 + \\ \quad \frac{1}{N}\frac{\text{r}(\theta_0)}{\sqrt{(GMM\text{-}AR_0+\text{r}(\theta_0))^2-4[GMM\text{-}AR_0-KLM_0]\text{r}(\theta_0))}}(KLM_1+KLM_2)+O_p(\frac{1}{N^2}). \end{array} \tag{21}$$

4. *A part of the higher order expression of the GMM-LM statistic (8) reads:*

$$\begin{array}{c}\mathrm{LM}\,(\theta)=KLM_0+\frac{1}{N}LM_1+\frac{1}{N^2}LM_{D_2}\left(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}+\\ \frac{1}{N}\mathrm{tr}\left(LM_{D_1}\left(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}\right),\end{array} \tag{22}$$

where

$$\begin{array}{l}LM_1=-2N^2f_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y)\\ \quad\left(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)\\ LM_{D_1}=2N\left\{f_N(\theta,Y)f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)-\right.\\ \quad\left.\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}f_N(\theta,Y)f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1}f_N(\theta,Y)\right\}\\ LM_{D_2}=N\left\{f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1}f_N(\theta,Y)\right\}^2\end{array} \tag{23}$$

and the conditional expectations of these elements given $D_N(\theta_0,Y)$ read:

$$\begin{array}{l}\mathrm{E}(LM_1|D_N(\theta_0,Y))=2-2\mathrm{tr}\left\{\mathrm{E}\left[f_i(\theta_0)f_i(\theta_0)'V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0))'\right]V_{ff}(\theta_0)^{-\frac{1}{2}}\right.\\ \quad\left.P_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}\right\}\\ \mathrm{E}(LM_{D_1}|D_N(\theta_0,Y))=\\ \quad 2\mathrm{E}\left[f_i(\theta_0)f_i(\theta_0)'V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}f_i(\theta_0)V_{ff}(\theta_0)^{-1}|D_N(\theta_0,Y)\right]-\\ \quad 2\frac{N-1}{N}\left\{\mathrm{tr}(V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0))\mathrm{E}\left(f_i(\theta_0)'V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'|D_N(\theta_0,Y)\right)+\right.\\ \quad\mathrm{E}\left(f_i(\theta_0)'V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)'V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'|D_N(\theta_0,Y)\right)+\\ \quad\left.\mathrm{E}\left(f_i(\theta_0)'V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'|D_N(\theta_0,Y)\right)\right\}-\\ \quad\frac{2}{N}\mathrm{E}[f_i(\theta_0)'V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'|D_N(\theta_0,Y)]-\\ \quad\frac{2}{N}\mathrm{E}\left(f_i(\theta_0)'V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}f_i(\theta_0)f_i(\theta_0)'|D_N(\theta_0,Y)\right)\\ \mathrm{E}(LM_{D_2}|D_N(\theta_0,Y))=(N-1)\left\{\left[\mathrm{tr}(V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0))\right]^2++\right.\\ \quad\left.\mathrm{tr}(V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)')+\mathrm{tr}(V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0))\right\}+\\ \quad\mathrm{E}\left[f_i(\theta_0)'V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}f(\theta_0))^2\right].\end{array} \tag{24}$$

Proof. see Appendix B. ■

We determined the order of the different elements in the higher order expressions in Theorem 1 using the conditional expectations given $D_N(\theta_0,Y)$. We use these for the Edgeworth approximation of the conditional finite sample distribution of the different statistics given $D_N(\theta_0,Y)$.

The higher order elements of $\mathrm{KLM}(\theta_0)$: KLM_1 and KLM_2 result from the different covariance estimators that are involved in $\mathrm{KLM}(\theta_0)$: KLM_1 results from $\hat{V}_{ff}(\theta)$ while KLM_2 results from (the infeasible covariance matrix estimator) $\hat{V}_{\theta f}(\theta_0)$. KLM_1 is therefore comparable to the $GMM\text{-}AR_1$ higher order element of GMM-AR(θ_0) which also results from $\hat{V}_{ff}(\theta)$. When Assumption 2 holds, the higher order element that results from $\hat{V}_{\theta f}(\theta_0)$, *i.e.* KLM_2, is of a lower order than the one which results from $\hat{V}_{ff}(\theta)$, *i.e.* KLM_1. Assumption 2 namely implies that

$$\begin{array}{l}\mathrm{E}(\frac{1}{N}\sum_{i=1}^{N}V_{ff}(\theta_0)^{-\frac{1}{2}}f_if_i'V_{ff}(\theta_0)^{-\frac{1}{2}}\otimes V_{ff}(\theta_0)^{-\frac{1}{2}}d_i(\theta_0)|D_N(\theta_0,Y))' \\ \mathrm{vec}(M_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)})'=(I_k\otimes V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y))'\mathrm{vec}(M_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)})'=0\end{array} \tag{25}$$

so the first element of the conditional expectation of KLM_2 is equal to zero. Under Assumption 2, the higher order elements that result from $\hat{V}_{\theta f}(\theta_0)$ are therefore of a lower order than those that result from $\hat{V}_{ff}(\theta_0)$.

The higher order expression of the GMM-LM statistic, $\mathrm{LM}(\theta_0)$, in Theorem 1 just states some of its higher order elements. We only want to show the dependence of its higher order elements on $D_N(\theta_0,Y)$ for which we do not need the full higher order expression. The higher order elements of $\mathrm{KLM}(\theta_0)$ also depend on $D_N(\theta_0,Y)$ but that dependence is of order zero so $\mathrm{KLM}(\theta_0)$ is invariant to the length of $D_N(\theta_0,Y)$ and only depends on its direction. The higher order elements of $\mathrm{LM}(\theta_0)$ in Theorem 1 are such that both LM_{D_1} and LM_{D_2} are multiplied by a function that is not invariant with respect to the length of $D_N(\theta_0,Y)$. The conditional expectations show that LM_{D_1} is at most of order zero in the sample size, N. Hence, when multiplied by $(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}D_N(\theta,Y)'$ and divided by N, the full contribution of LM_{D_1} is at most of the order $O(N^{-\frac{1}{2}})$. This occurs when $J_\theta(\theta)$ equals zero in which case $(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}D_N(\theta,Y)'$ is of order $O(N^{\frac{1}{2}})$. LM_{D_1} can therefore not alter the limiting distribution of $\mathrm{LM}(\theta_0)$. The conditional expectation of LM_{D_2} is proportional N. When $J_\theta(\theta_0)$ equals zero, $(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}$ is proportional to N as well. Thus LM_{D_2} alters the limiting distribution of $\mathrm{LM}(\theta)$ when $J_\theta(\theta)$ equals zero since it leads to an element of zero-th order in N in the higher order expression of $\mathrm{LM}(\theta_0)$. This explains why the limiting distribution of $\mathrm{LM}(\theta_0)$ depends on $J_\theta(\theta_0)$. The same result can be shown for other GMM statistics that are not robust to weak instruments. For reasons of brevity, we only did so for the statistic for which it is the least involved to do so and we refrain from showing this for other statistics as well.

The higher order expressions in Theorem 1 also show the sensitivity to the number of instruments. For example, both the GMM-AR statistic and the LM statistics have higher order terms which are proportional to $\frac{k^2}{N}$ while the higher order terms of the KLM statistics are proportional to $\frac{k}{N}$. This shows that the approximation of the finite sample distribution by the limiting distribution is more robust to the number of instruments for the KLM statistic compared to these other two statistics.

The higher order expression of GMM-MLR(θ_0) stated in Theorem 1 is conditional on $\mathrm{r}(\theta_0)$. It is obtained using a Taylor expansion with respect to the other two components of GMM-MLR(θ_0) : GMM-AR(θ_0) and $\mathrm{KLM}(\theta_0)$. Since the limiting distribution of GMM-MLR(θ_0) is conditional on r(θ_0), it is important to determine the highest order of the elements of the higher order expansion of r(θ_0) that are not independent of GMM-AR(θ_0) and KLM(θ_0). We therefore construct the higher order expansion of r(θ_0) which is stated in Theorem 2.

Theorem 2. *Under H_0 and Assumption 1, the higher order expression of the conditioning statistic $r(\theta_0)$ of GMM-MLR(θ_0) (7) reads*

$$\mathrm{r}(\theta_0) = \ \mathrm{r}_0 + \tfrac{1}{N}(\mathrm{r}_1 + \mathrm{r}_2) + \tfrac{1}{N^2}\mathrm{r}_3 + o_p(\tfrac{1}{N^2}), \tag{26}$$

with

$$\begin{aligned}
\mathrm{r}_0 = &\ N D_N(\theta_0, Y)' V_{\theta\theta.f}(\theta_0)^{-1} D_N(\theta_0, Y) \\
\mathrm{r}_1 = &\ 2N^2 f_N(\theta_0, Y)' V_{ff}(\theta_0)^{-1} \hat{V}_{\theta f}(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1} D_N(\theta_0, Y) \\
\mathrm{r}_2 = &\ N^2 D_N(\theta_0, Y)' V_{\theta\theta.f}(\theta_0)^{-1} \left[\hat{V}_{dd}(\theta_0) - V_{\theta\theta.f}(\theta_0)\right] V_{\theta\theta.f}(\theta_0)^{-1} D_N(\theta_0, Y) \\
\mathrm{r}_3 = &\ N^3 f_N(\theta_0, Y)' V_{ff}(\theta_0)^{-1} \hat{V}_{\theta f}(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1} \left[\hat{V}_{dd}(\theta_0) - V_{\theta\theta.f}(\theta_0)\right] \\
&\ V_{\theta\theta.f}(\theta_0)^{-1} \hat{V}_{\theta f}(\theta_0) V_{ff}(\theta_0)^{-1} f_N(\theta_0, Y).
\end{aligned} \tag{27}$$

with $V_{\theta\theta.f}(\theta_0) = V_{qq}(\theta_0) - V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}V_{qf}(\theta_0)'$, $\hat{V}_{dd}(\theta) = \frac{1}{N}\sum_{i=1}^N \bar{d}_i(\theta)\bar{d}_i(\theta)'$ *and*

$$\begin{aligned}
\mathrm{E}[\mathrm{r}_0] = &\ NJ(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1} J(\theta_0) + k \\
\mathrm{E}[\mathrm{r}_1] = &\ 2(N-1)\mathrm{E}\left[f_i(\theta_0)' V_{ff}(\theta_0)^{-1}(d_i(\theta_0) - J_\theta(\theta_0)) f_i(\theta_0)'\right] V_{\theta\theta.f}(\theta_0)^{-1} J_\theta(\theta_0) + \\
&\ 2\tfrac{N-1}{N}\mathrm{E}\left[f_i(\theta_0)' V_{ff}(\theta_0)^{-1}(d_i(\theta_0) - J_\theta(\theta_0)) f_i(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1}(d_i(\theta_0) - J_\theta(\theta_0))\right] + \\
&\ 2\mathrm{E}\left[f_i(\theta_0)' V_{ff}(\theta_0)^{-1} J_\theta(\theta_0) f_i(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1}(d_i(\theta_0) - J_\theta(\theta_0))\right] - 2\tfrac{N-1}{N}k \\
\mathrm{E}[\mathrm{r}_2] = &\ 2(N-1) J_\theta(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1}\mathrm{E}\left[(d_i(\theta_0) - J_\theta(\theta_0))(d_i(\theta_0) - J_\theta(\theta_0))' V_{\theta\theta.f}(\theta_0)^{-1}\right. \\
&\ \left.(d_i(\theta_0) - J_\theta(\theta_0))\right] - k(k+3) + \tfrac{1}{N}(k^2 + 2k) + \tfrac{(N-1)}{N}\mathrm{E}\left[(d_i(\theta_0) - J_\theta(\theta_0))'\right. \\
&\ \left. V_{\theta\theta.f}(\theta_0)^{-1}(d_i(\theta_0) - J_\theta(\theta_0))(d_i(\theta_0) - J_\theta(\theta_0))' V_{\theta\theta.f}(\theta_0)^{-1}(d_i(\theta_0) - J_\theta(\theta_0))\right] \\
\mathrm{E}[\mathrm{r}_3] = &\ O(1),
\end{aligned} \tag{28}$$

such that

$$\begin{aligned}
O(\mathrm{E}(r_1)) = &\ o(1) \quad \textit{when Assumption 2 holds.} \\
O(\mathrm{E}(r_2)) = &\ o(1) \quad \textit{when the third order moment of } d_i(\theta_0) \textit{ equals zero.}
\end{aligned} \tag{29}$$

Proof. see Appendix B. ■

The independence of the limiting distributions of r(θ_0) and the GMM-AR and KLM statistics results from the independence of $f_N(\theta_0, Y)$ and $D_N(\theta_0, Y)$ in large samples. Theorem 2 shows that the independence of r(θ_0) and the GMM-AR and KLM statistics is up to order $\frac{1}{N}$ since the r_1 element of r(θ_0) is not independent of $f_N(\theta_0, Y)$. When Assumption 2 holds, the independence of r(θ_0) and the GMM-AR and KLM statistics is up to including order $\frac{1}{N}$ since r_1 is then of a lower order in the sample size.

4 Bootstraping weak instrument robust statistics

We construct one bootstrap algorithm to resample the GMM-AR and GMM-MLR statistics and two bootstrap algorithms to resample the KLM statistic. One of the bootstrap algorithms for the KLM statistic improves the approximation of its finite sample distribution when both Assumptions 1 and 2 hold while the other improves the approximation when just Assumption 1 holds.

Under Assumption 2, Theorem 2 shows that the independence between $\mathrm{r}(\theta_0)$ and the GMM-AR and KLM statistics is up to including the order $\frac{1}{N}$. This allows us to construct a bootstrap algorithm that resamples the GMM-MLR statistic given $\mathrm{r}(\theta_0)$. It improves the approximation of the conditional finite sample distribution of the GMM-MLR statistic compared to using the conditional limiting distribution.

When Assumption 2 does not hold, the dependence between $\mathrm{r}(\theta_0)$ and the GMM-AR and KLM statistics is of order $\frac{1}{N}$. Unlike when Assumption 2 holds, we then have to construct a formal Edgeworth approximation of the conditional distribution of the GMM-MLR statistic given $r(\theta)$ to establish that the bootstrap leads to higher order improvements. Since it is unclear how to construct such an Edgeworth approximation, because of the non-standard conditional limiting distribution of the GMM-MLR statistic, we refrain from constructing such an Edgeworth approximation and an algorithm to bootstrap the GMM-MLR statistic when just Assumption 1 holds. Assumption 2 is irrelevant for the GMM-AR statistic so its bootstrap algorithm is straightforward.

The first bootstrap algorithm that we propose just resamples $\bar{f}_i(\theta_0)$, $i = 1, \ldots, N$, with replacement.[2] It uses the resampled values of $\bar{f}_i(\theta_0)$ for the bootstrapped moment vector $f_B^*(\theta_0, Y)$ and covariance matrix estimator $V_{ff}^*(\theta_0)$ which are used to construct GMM-AR$^*(\theta)$, KLM$^*(\theta)$ and GMM-MLR$^*(\theta)$. The resampled KLM* and GMM-MLR* statistics are based on the realized sample values of $\hat{D}_N(\theta_0, Y)$ and $\mathrm{r}(\theta_0)$.

The second bootstrap algorithm resamples $(\bar{f}_i(\theta_0)'\ q_i(\theta_0)')'$, $i = 1, \ldots, N$, with replacement. It uses the resampled values of $(\bar{f}_i(\theta_0)'\ q_i(\theta_0)')'$ for the bootstrapped moment and derivative vectors $f_B^*(\theta_0, Y)$ and $q_B^*(\theta_0, Y)$ and the bootstrapped covariance matrix estimators $V_{ff}^*(\theta_0)$ and $V_{qf}^*(\theta_0)$ which are all used to construct KLM$^{**}(\theta)$.

Theorem 1 shows that the limiting distribution of LM(θ) depends on nuisance parameters so we can not construct a bootstrap algorithm that resamples LM(θ) and that approximates the finite sample distribution of LM(θ) for all values of the nuisance parameters. We therefore do not construct a bootstrap algorithm for

[2]We resample $\bar{f}_i(\theta_0)$ instead of $f_i(\theta_0)$ so the (empirical) moment condition holds for the bootstrap population $\bar{f}_i(\theta_0)$, $i = 1, \ldots, N$, see *e.g.* Hall and Horowitz (1996).

LM(θ).

The bootstrap algorithms to approximate the finite sample distributions of the GMM-AR, KLM and GMM-MLR statistics read:

Bootstrap Algorithm 1:

1. Compute $\hat{D}_N(\theta_0, Y)$ and $\mathrm{r}(\theta_0)$ and set bootstrap sample size B and number of simulations S.

2. For $m = 1, \ldots, S$:

   (a) Sample $\{f_j^*(\theta_0),\ j = 1, \ldots, B\}$ independently with replacement from $\{\bar{f}_l(\theta_0),\ l = 1, \ldots, N\}$:

   $$\Pr\left[f_j^*(\theta_0) = \bar{f}_l(\theta_0)\right] = \tfrac{1}{N},\ l = 1, \ldots, N. \tag{30}$$

   (b) Compute:

   $$\begin{aligned} f_B^*(\theta_0, Y)_m = & \ \tfrac{1}{B} \textstyle\sum_{j=1}^B f_j^*(\theta_0) \text{ and} \\ V_{ff}^*(\theta_0)_m = & \ \tfrac{1}{B} \textstyle\sum_{j=1}^B f_j^*(\theta_0) f_j^*(\theta_0)' - f_B^*(\theta_0, Y) f_B^*(\theta_0, Y)', \end{aligned} \tag{31}$$

   from the bootstrap sample $\{f_j^*(\theta_0),\ j = 1, \ldots, B\}$.

   (c) Compute:

   $$\begin{aligned} &\text{GMM-AR}^*(\theta_0)_m = B f_B^*(\theta_0, Y)_m' V_{ff}^*(\theta_0)_m^{-1} f_B^*(\theta_0, Y)_m \\ &\text{KLM}^*(\theta_0)_m = B f_B^*(\theta_0, Y)_m' V_{ff}^*(\theta_0)_m^{-\frac{1}{2}} P_{V_{ff}^*(\theta_0)_m^{-\frac{1}{2}} \hat{D}_N(\theta_0, Y)} V_{ff}^*(\theta_0)_m^{-\frac{1}{2}} f_B^*(\theta_0, Y)_m \\ &\text{GMM-MLR}^*(\theta_0)_m = \tfrac{1}{2} \Big[\text{GMM-AR}^*(\theta_0)_m - \mathrm{r}(\theta_0) + \\ &\sqrt{\left(\text{GMM-AR}^*(\theta_0)_m + \mathrm{r}(\theta_0)\right)^2 - 4\left[\text{GMM-AR}^*(\theta_0)_m - \text{KLM}^*(\theta_0)_m\right] \mathrm{r}(\theta_0))}\Big] \end{aligned} \tag{32}$$

3. Construct the (conditional) bootstrap distributions of GMM-AR(θ_0), KLM(θ_0) and GMM-MLR(θ_0) from the sample $\{$GMM-AR$^*(\theta_0)_m$, KLM$^*(\theta_0)_m$, GMM-MLR$^*(\theta_0)_m$, $m = 1, \ldots, S\}$.

Bootstrap Algorithm 2:

1. Set bootstrap sample size B and number of simulations S.

2. For $m = 1, \ldots, S$:

(a) Sample $\{(f_j^*(\theta_0)'\ q_j^*(\theta_0)')',\ j=1,\ldots,B\}$ independently with replacement from $\{(\bar{f}_l(\theta_0)'\ q_l(\theta_0)')',\ l=1,\ldots,N\}$:

$$\Pr\left[\binom{f_j^*(\theta_0)}{q_j^*(\theta_0)}=\binom{\bar{f}_l(\theta_0)}{q_l(\theta_0)}\right]=\tfrac{1}{N},\ l=1,\ldots,N. \tag{33}$$

(b) Compute:

$$\begin{array}{rl} f_B^*(\theta_0,Y)_m = & \frac{1}{B}\sum_{j=1}^B f_j^*(\theta_0) \\ q_B^*(\theta_0,Y)_m = & \frac{1}{B}\sum_{j=1}^B q_j^*(\theta_0) \\ V_{ff}^*(\theta_0)_m = & \frac{1}{B}\sum_{j=1}^B f_j^*(\theta_0)f_j^*(\theta_0)'-f_B^*(\theta_0,Y)_m f_B^*(\theta_0,Y)_m' \\ V_{qf}^*(\theta_0)_m = & \frac{1}{B}\sum_{j=1}^B q_j^*(\theta_0)f_j^*(\theta_0)'-q_B^*(\theta_0,Y)_m f_B^*(\theta_0,Y)_m' \\ \hat{D}_B^*(\theta_0,Y)_m = & q_B^*(\theta_0,Y)_m - V_{qf}^*(\theta_0)_m V_{ff}^*(\theta_0)_m^{-1} f_B^*(\theta_0,Y)_m \end{array} \tag{34}$$

from the bootstrap sample $\{(f_j^*(\theta_0)'\ q_j^*(\theta_0)')',\ j=1,\ldots,B\}$.

(c) Compute:

$$\begin{array}{rl} \mathrm{KLM}^{**}(\theta_0)_m = & B f_B^*(\theta_0,Y)_m' V_{ff}^*(\theta_0)_m^{-\frac{1}{2}} P_{V_{ff}^*(\theta_0)_m^{-\frac{1}{2}}\hat{D}_B^*(\theta_0,Y)_m} V_{ff}^*(\theta_0)_m^{-\frac{1}{2}} \\ & f_B^*(\theta_0,Y)_m \end{array} \tag{35}$$

3. Construct the bootstrap distribution of $\mathrm{KLM}(\theta_0)$ from the sample $\{\mathrm{KLM}^{**}(\theta_0)_m,\ m=1,\ldots,S\}$.

The bootstrap algorithms are such that $\mathrm{E}^*(f_B^*(\theta_0,Y)_m)=0$, $\mathrm{E}^*(Bf_B^*(\theta_0,Y)_m f_B^*(\theta_0,Y)_m')=\mathrm{E}^*(V_{ff}^*(\theta_0)_m)=\hat{V}_{ff}(\theta)$ and $\mathrm{E}^*(Bq_B^*(\theta_0,Y)_m f_B^*(\theta_0,Y)_m')=\mathrm{E}^*(V_{qf}^*(\theta_0)_m)$ $=\hat{V}_{qf}(\theta)$, where E^* is the expectation operator with respect to the resampling distribution. The above algorithms just bootstrap sample means and covariances so the bootstrap distributions converge to the large sample distributions.

Theorem 3. **a.** *Under H_0, Assumption 1 and when $B=N$, the distribution of $f_B^*(\theta_0,Y)$ that results from Bootstrap Algorithm 1 consistently estimates the (normal) large sample distribution of $f_N(\theta_0,Y)$.*

b. *Under H_0, Assumption 1 and when $B=N$, the distribution of $(f_B^*(\theta_0,Y)'\ q_B^*(\theta_0,Y)')'$ that results from Bootstrap Algorithm 2 consistently estimates the (normal) large sample distribution of $(f_N(\theta_0,Y)'\ q_N(\theta_0,Y)')'$.*

Proof. The bootstrap distributions of $q_B^*(\theta_0,Y)$ and/or $f_B^*(\theta_0,Y)$ that result from Bootstrap algorithms 1 and 2 just bootstrap sample means whose bootstrap distributions converge under Assumption 1 to their large sample (normal) distributions, see Horowitz (2001), Theorem 2.2, and Mammen (1992). ■

Theorem 3 shows that the bootstrap distribution of $f_B^*(\theta_0,Y)$ that results from Bootstrap Algorithm 1 and that of $(f_B^*(\theta_0,Y)'\ q_B^*(\theta_0,Y)')'$ that results from Bootstrap Algorithm 2 converge to normal distributions when $B=N$ goes to infinity.

Corollary 5. *Under H_0, Assumption 1 and when $B = N$, the bootstrap distributions of GMM-AR$^*(\theta_0)$ and KLM$^{**}(\theta_0)$ converge to the (conditional) limiting distributions of GMM-AR(θ_0) and KLM(θ_0).*

Proof. The result for GMM-AR$^*(\theta_0)$ is a direct consequence of Theorem 3a. The covariance estimators $V_{qf}^*(\theta_0)_i$ and $V_{ff}^*(\theta_0)_i$ that are used in Bootstrap Algorithm 2 are consistent estimators of $\hat{V}_{qf}(\theta_0)$ and $\hat{V}_{ff}(\theta_0)$. It therefore holds that $f_B^*(\theta_0, Y)_i$ and $\hat{D}_B^*(\theta_0, Y)_i$ are independently distributed in large samples so the bootstrap distribution of KLM$^{**}(\theta_0)$ that results from Bootstrap Algorithm 2 converges to the limiting distribution of KLM(θ_0). ■

Corollary 6. *Under H_0, Assumptions 1, 2 and when $B = N$, the bootstrap distributions of KLM$^*(\theta_0)$ and GMM-MLR$^*(\theta_0)$ converge to the (conditional) limiting distributions of KLM(θ_0) and GMM-MLR(θ_0).*

Proof. Since $\hat{D}_N(\theta_0, Y)$ is fixed in Bootstrap Algorithm 1 and $V_{ff}^*(\theta_0)$ is a consistent estimator of $\hat{V}_{ff}(\theta_0)$, the bootstrap distributions KLM$^*(\theta_0)$ and GMM-MLR$^*(\theta_0)$ converge to the (conditional) limiting distributions of KLM(θ_0) and GMM-MLR(θ_0). ■

Corollaries 5 and 6 show the validity of the bootstraps proposed in Algorithms 1 and 2. In Moreira *et. al.* (2010), the validity of the bootstrap for the KLM statistic in the linear instrumental variables regression model is shown. Assumption 2 is satisfied for the linear instrumental variables regression model because of the homoscedasticity of the disturbances. Hence bootstrap algorithms 1 and 2 are valid for the linear instrumental regression model. Moreira *et. al.* (2010) do, however, not propose a bootstrap algorithm that resamples the moments under the null hypothesis, like we do, but one that resamples the errors that result from estimating the instrumental variables regression model. Their resampled errors therefore depend on the involved estimator which precludes higher order efficiency gains from the bootstrap. Because our resampled moments do not depend on an estimator, the bootstrap algorithms that we use lead to higher order efficiency gains.

Besides its validity, it is of interest to analyze if the bootstrap leads to a higher order improvement of the approximation of the finite sample distribution of the different statistics compared to the limiting distribution. We therefore construct Edgeworth approximations of the distributions of the bootstrap statistics. These Edgeworth approximations use the higher order expressions of the bootstrap statistics which are stated in Theorem 4.

Theorem 4. **1.** *Under H_0 and Assumption 1, the higher order expression of the bootstrap statistic GMM-AR$^*(\theta_0)$ that results from Bootstrap Algorithm 1 reads:*

$$\text{GMM-AR}^*(\theta_0) = GMM\text{-}AR_0^* + \tfrac{1}{B} GMM\text{-}AR_1^* + O_p(\tfrac{1}{B^2}), \tag{36}$$

where $f_B^*(\theta_0, Y) = \frac{1}{B}\sum_{j=1}^B f_j^*(\theta_0)$,

$$\begin{aligned} GMM\text{-}AR_0^* = & \; B f_B^*(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1} f_B^*(\theta_0,Y) \\ GMM\text{-}AR_1^* = & \; -B^2 f_B^*(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}\left[V_{ff}^*(\theta_0) - \hat{V}_{ff}(\theta_0)\right]\hat{V}_{ff}(\theta_0)^{-1} f_B^*(\theta_0,Y) \end{aligned} \tag{37}$$

and $\mathrm{E}(GMM\text{-}AR_0^*) = k$ *and*

$$E^*(GMM\text{-}AR_1^*) = -\tfrac{B-1}{B}\left\{\tfrac{1}{N}\sum_{i=1}^N \left(\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-1}\bar{f}_i(\theta_0)\right)^2 + [k^2+2k]\right\} + k. \tag{38}$$

2. *Under* H_0 *and Assumptions 1 and 2, the higher order expression of the bootstrap statistic* $\text{KLM}^*(\theta_0)$ *that results from Bootstrap Algorithm 1 reads*:

$$\text{KLM}^*(\theta_0) = KLM_0^* + \tfrac{1}{B} KLM_1^* + O_p(\tfrac{1}{B^2}), \tag{39}$$

where

$$\begin{aligned} KLM_0^* &= B f_B^*(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)} \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} f_B^*(\theta_0,Y) \\ KLM_1^* &= -B^2 f_B^*(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)} \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\left[V_{ff}^*(\theta_0) - \hat{V}_{ff}(\theta_0)\right] \\ &\quad \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)} \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} f_B^*(\theta_0,Y) - 2B^2 f_B^*(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} \\ &\quad P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)} \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\left[V_{ff}^*(\theta_0) - \hat{V}_{ff}(\theta_0)\right]\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} \\ &\quad M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)} \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} f_B^*(\theta_0,Y) \end{aligned} \tag{40}$$

and $\mathrm{E}^*\left[KLM_0^*|\hat{D}_N(\theta_0,Y)\right] = 1$,

$$\begin{aligned} &\mathrm{E}^*\left[KLM_1^*|\hat{D}_N(\theta_0,Y)\right] = \\ &\quad -\tfrac{B-1}{B}\left\{\tfrac{1}{N}\sum_{i=1}^N\left(\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}_i(\theta_0)\right)^2 + \right. \\ &\quad \tfrac{2}{N}\sum_{i=1}^N\left(\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}_i(\theta_0)\right. \\ &\quad \left.\left.\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}_i(\theta_0)\right) - 2(k+1)\right\} + \tfrac{1}{B}. \end{aligned} \tag{41}$$

3. *Under* H_0 *and Assumptions 1 and 2, the higher order expression of the bootstrap statistic* $\text{GMM-MLR}^*(\theta_0)$ *given r(*θ_0*) that results from Bootstrap Algorithm*

1 reads:

$$\text{GMM-MLR}^*(\theta_0) =$$
$$\tfrac{1}{2}\left[GMM\text{-}AR_0^* - \mathrm{r}(\theta_0) + \sqrt{\left(GMM\text{-}AR_0^* + \mathrm{r}(\theta_0)\right)^2 - 4\left[GMM\text{-}AR_0^* - KLM_0^*\right]\mathrm{r}(\theta_0))}\right] +$$
$$\tfrac{1}{2B}\left[1 + \tfrac{GMM\text{-}AR_0^* - \mathrm{r}(\theta_0)}{\sqrt{\left(GMM\text{-}AR_0^* + \mathrm{r}(\theta_0)\right)^2 - 4[GMM\text{-}AR(\theta_0) - KLM(\theta_0)]\mathrm{r}(\theta_0))}}\right] GMM\text{-}AR_1^* +$$
$$\tfrac{1}{B}\tfrac{\mathrm{r}(\theta_0)}{\sqrt{\left(GMM\text{-}AR_0^* + \mathrm{r}(\theta_0)\right)^2 - 4\left[GMM\text{-}AR_0^* - KLM_0^*\right]\mathrm{r}(\theta_0))}} KLM_1^* + O_p(\tfrac{1}{B^2}). \tag{42}$$

4. *Under H_0 and Assumption 1, the higher order expression of the bootstrap statistic* $\text{KLM}^{**}(\theta_0)$ *that results from Bootstrap Algorithm 2 reads*:

$$\text{KLM}^{**}(\theta_0) = KLM_0^* + \tfrac{1}{B}(KLM_1^* + KLM_2^{**}) + O_p(\tfrac{1}{B^2}), \tag{43}$$

where

$$KLM_2^{**} = -2B^2 f_B^*(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)} \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} V_{\theta f}^*(\theta_0)\hat{V}_{ff}(\theta_0)^{-1}$$
$$f_B^*(\theta_0, Y)\left[\hat{D}_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0, Y)\right]^{-1}\hat{D}_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1} f_B^*(\theta_0, Y) \tag{44}$$

and

$$\mathrm{E}\left[KLM_2^*|\hat{D}_N(\theta_0, Y)\right] = -\tfrac{4(B-1)}{B}\mathrm{vec}(M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)})'(\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} \otimes \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}})$$
$$\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N(\bar{f}_i(\theta_0)\bar{f}_i(\theta_0)' \otimes \hat{d}_i(\theta_0))\right]\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0, Y)\left[\hat{D}_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0, Y)\right]^{-1} -$$
$$\tfrac{2}{B}\mathrm{vec}(M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)})'(\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}} \otimes \hat{V}_{ff}(\theta_0)^{-\frac{1}{2}})\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N(\bar{f}_i(\theta_0)\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-1}\right.$$
$$\left.\bar{f}_i(\theta_0)\bar{f}_i(\theta_0)' \otimes \hat{d}_i(\theta_0))\right]\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0, Y)\left[\hat{D}_N(\theta_0, Y)'\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0, Y)\right]^{-1}. \tag{45}$$

Proof. see Appendix B. The results for the GMM-MLR statistic follow from Theorem 1 and Theorems 4.1 and 4.2. ■

The higher order expressions of the bootstrap statistics in Theorem 4 are identical to those of the orginal statistics in Theorem 1. They are also such that the expectations of the higher order elements converge to the same limits when B equals N and N goes to infinity. This holds since

$$\hat{D}_N(\theta_0, Y) = D_N(\theta_0, Y) - \hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1} f_N(\theta_0, Y) + O_p(\tfrac{1}{N}), \tag{46}$$

which is stated in Lemma 2 in Appendix A, and $\hat{V}_{\theta f}(\theta_0)$ converges to zero at rate $\frac{1}{\sqrt{N}}$ such that $\hat{D}_N(\theta_0, Y)$ and $D_N(\theta_0, Y)$ have the same convergence behavior. It suggests that usage of the bootstrap distributions leads to a higher order of precision in terms of the order of the approximation error of the finite sample distribution compared to the limiting distribution. To verify this statement we construct the Edgeworth approximations of the finite sample distributions of the original statistics and of their bootstrapped counterparts.

5 Edgeworth Approximations

Bhattacharya and Ghosh (1978) provide regularity conditions for Edgeworth approximations of finite sample distributions of Wald statistics. Since these Edgeworth approximations are based on higher order expansions that result from Taylor approximations, the regularity conditions request that the expected values of the derivatives are non-zero, *i.e.* $\mathrm{E}(D_N(\theta_0, Y)) = J_\theta(\theta_0) \neq 0$. The Edgeworth approximation that results from Bhattacharya and Ghosh (1978) does therefore not allow for weak instruments. Our Edgeworth approximations allow for a zero value of the Jacobian, $J_\theta(\theta_0)$, since they are not marginal with respect to $D_N(\theta_0, Y)$ but condition on it. We can construct such Edgeworth approximations because the large sample distributions of the weak instrument robust GMM-AR and KLM statistics remain the same irrespective of whether we construct them marginal or conditional with respect to $D_N(\theta_0, Y)$. The (conditional) large sample distribution of the GMM-MLR statistic depends on $D_N(\theta_0, Y)$ so we use a different argument to show the higher order improvements that result from the bootstrap for it.

The Edgeworth approximations are obtained from the characteristic function associated with the (conditional) limiting distribution of $(f_N(\theta_0, Y), \hat{V}_{ff}(\theta_0), \hat{V}_{\theta f}(\theta_0))$. For this characteristic function to exist, Cramèr's condition has to hold with respect to the limiting distribution of $f_N(\theta_0, Y)$.

Assumption 3. *Cramèr condition: for a k-dimensional vector $t \in \mathbb{R}^k$, it holds that*

$$\lim_{\|t\| \to \infty} \sup \left| \mathrm{E}\left[\exp(it'\psi)\right]\right| < 1, \tag{47}$$

where

$$\sqrt{N} f_N(\theta_0, Y) \underset{d}{\rightarrow} \psi. \tag{48}$$

We construct the Edgeworth expansions of the GMM-AR and KLM-statistics and of their bootstrapped counterparts. The Edgeworth expansions are stated in Theorem 5.

Theorem 5. **A.** *Under H_0 and Assumptions 1 and 3, the Edgeworth approximations of the (conditional) finite sample distributions of GMM-AR(θ_0) and KLM(θ_0) read:*

1. *For GMM-AR(θ_0) :*

$$\begin{aligned}\Pr\left[\text{GMM-AR}\left(\theta_0\right) \leq x\right] = &\ \Pr_{\chi^2(k)}(x) - \tfrac{1}{N}\tfrac{\mathrm{E}(GMM\text{-}AR_1)}{k} x p_{\chi^2(k)}(x) + O(N^{-2}) \\ = &\ \Pr_{\chi^2(k)}\left(x - \tfrac{1}{N}\tfrac{\mathrm{E}(GMM\text{-}AR_1)}{k} x\right) + O(N^{-2}),\end{aligned} \tag{49}$$

where $\Pr_{\chi^2(k)}(x)$ *and* $p_{\chi^2(k)}(x)$ *are the distribution and density function of a* $\chi^2(k)$ *distributed random variable evaluated at* x *and* $\mathrm{E}(GMM\text{-}AR_1)$ *is defined in Theorem 1.*

2. *For KLM(θ_0) :*

$$\begin{aligned}&\Pr\left[\text{KLM}\left(\theta_0\right) \leq x | D_N(\theta_0, Y)\right] \\ &= \Pr_{\chi^2(1)}(x) + \tfrac{1}{N}\left(ax + (b+c)\sqrt{2\pi x}\right) p_{\chi^2(1)}(x) + o(N^{-1}) \\ &= \Pr_{\chi^2(1)}\left(x + \tfrac{1}{N}\left(ax + (b+c)\sqrt{2\pi x}\right)\right) + o(N^{-1}),\end{aligned} \tag{50}$$

where $o(N^{-1})$ *indicates that the remaining terms are of a lower order than* N^{-1},

$$\begin{aligned}a = &\ \tfrac{N-1}{N}\left\{\mathrm{E}\left[\left(f_i(\theta_0)' V_{ff}(\theta_0)^{-\frac{1}{2}} P_{V_{ff}(\theta_0)^{-\frac{1}{2}} D_N(\theta_0, Y)} V_{ff}(\theta_0)^{-\frac{1}{2}} f_i(\theta_0)\right)^2\right] - 4\right\} - \tfrac{1}{N} \\ b = &\ \tfrac{N-1}{N}\left\{\mathrm{E}\left[f_i(\theta_0)' V_{ff}(\theta_0)^{-\frac{1}{2}} P_{V_{ff}(\theta_0)^{-\frac{1}{2}} D_N(\theta_0, Y)} V_{ff}(\theta_0)^{-\frac{1}{2}} f_i(\theta_0)\right.\right. \\ &\left.\left. f_i(\theta_0)' V_{ff}(\theta_0)^{-\frac{1}{2}} M_{V_{ff}(\theta_0)^{-\frac{1}{2}} D_N(\theta_0, Y)} V_{ff}(\theta_0)^{-\frac{1}{2}} f_i(\theta_0)\right]\right\} - (k-1)\end{aligned} \tag{51}$$

and c =*E(*$KLM_2 | D_N(\theta_0, Y)$*) which is defined in (19).*

3. *Under* H_0 *and Assumptions 1, 2 and 3, the Edgeworth approximation of the conditional finite sample distributions of KLM(θ_0) reads:*

$$\begin{aligned}\Pr\left[\text{KLM}\left(\theta_0\right) \leq x | D_N(\theta_0, Y)\right] &= \Pr_{\chi^2(1)}(x) + \tfrac{1}{N}\left(ax + b\sqrt{2\pi x}\right) p_{\chi^2(1)}(x) + o(N^{-1}) \\ &= \Pr_{\chi^2(1)}\left(x + \tfrac{1}{N}\left(ax + b\sqrt{2\pi x}\right)\right) + o(N^{-1}),\end{aligned} \tag{52}$$

B. *Under* H_0*, the Edgeworth approximations of the finite sample distributions of bootstrapped GMM-AR(θ_0) and KLM(θ_0) : GMM-AR***(θ_0), KLM***(θ_0) and KLM****(θ_0) read:*

1. *When Assumptions 1 and 3 hold for GMM-AR***(θ_0) :*

$$\Pr\left[\text{GMM-AR}^*\left(\theta_0\right) \leq x\right] = \Pr_{\chi^2(k)}\left(x - \tfrac{1}{B}\tfrac{\mathrm{E}(GMM\text{-}AR_1^*)}{k} x\right) + O(B^{-2}), \tag{53}$$

where $\mathrm{E}(GMM\text{-}AR_1^*)$ *is defined in Theorem 2.*

2. *When Assumption 1 and 3 hold for KLM****(θ_0) :*

$$\begin{aligned}&\Pr\left[\text{KLM}^{**}\left(\theta_0\right) \leq x | \hat{D}_N(\theta_0, Y)\right] \\ &= \Pr_{\chi^2(1)}(x) + \tfrac{1}{B}\left(a^* x + \sqrt{2\pi}(b^* + c^*) x^{\frac{1}{2}}\right) p_{\chi^2(1)}(x) + o(B^{-1}) \\ &= \Pr_{\chi^2(1)}\left(x + \tfrac{1}{B}\left(a^* x + \sqrt{2\pi}(b^* + c^*) x^{\frac{1}{2}}\right)\right) + o(B^{-1}).\end{aligned} \tag{54}$$

where

$$
\begin{aligned}
a^* = & \ \tfrac{B-1}{B}\left\{\tfrac{1}{N}\sum_{i=1}^{N}\left[\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}_i(\theta_0)\right]^2-4\right\}-\tfrac{1}{B}, \\
b^* = & \ \tfrac{B-1}{B}\left\{\tfrac{1}{N}\sum_{i=1}^{N}\left[\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}_i(\theta_0)\right.\right. \\
& \left.\left.\bar{f}_i(\theta_0)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}_i(\theta_0)\right]-(k-1)\right\},
\end{aligned}
\tag{55}
$$

and $c^* = E\left[KLM_2^*|\hat{D}_N(\theta_0,Y)\right]$ *from (45).*

3. *When Assumptions 1, 2 and 3 hold for* $KLM^*(\theta_0)$:

$$
\begin{aligned}
& \Pr\left[\mathrm{KLM}^*(\theta_0)\leq x|\hat{D}_N(\theta_0,Y)\right] \\
& = \Pr_{\chi^2(1)}(x)+\tfrac{1}{B}\left(a^*x+b^*\sqrt{2\pi x}\right)p_{\chi^2(1)}(x)+o(B^{-1}) \\
& = \Pr_{\chi^2(1)}\left(x+\tfrac{1}{B}\left(a^*x+b^*\sqrt{2\pi x}\right)\right)+o(B^{-1}),
\end{aligned}
\tag{56}
$$

where a^* *and* b^* *are defined in (55).*

Proof. see Appendix B. ■

Since $\mathrm{E}(GMM\text{-}AR_1^*)$ converges to $\mathrm{E}(GMM\text{-}AR_1)$ and a^*, b^* and c^* converge to a, b and c when B equals N and N goes to infinity, the Edgeworth approximations of the bootstrapped statistics converge to those of the orginal statistics when B equals N. Thus usage of bootstrapped critical values leads to a higher order efficiency gain, see *e.g.* Horowitz (2001).

Corollary 7. *Under* H_0, *Assumptions 1, 3 and when* B *equals* N, *the bootstrap critical values of GMM-AR*$^*(\theta_0)$ *and KLM*$^{**}(\theta_0)$ *remove the approximation error of the (conditional) finite sample distribution of GMM-AR(*θ*) and KLM(*θ*) up to/including the order* $\frac{1}{N}$.
Under H_0, *Assumptions 1-3 and when* B *equals* N, *the bootstrap critical values of KLM*$^*(\theta_0)$ *remove the approximation error of the (conditional) finite sample distribution of KLM(*θ*) up to/including the order* $\frac{1}{N}$.

Corollary 7 shows that the bootstrap provides higher order efficiency gains even in cases when the parameter of interest is not identified. Corollary 7 therefore extends the previously known results for the bootstrap which only apply to well identified cases.

Theorem 5 only states results for the GMM-AR and KLM statistics and not for the GMM-MLR statistic. Given r(θ), the GMM-MLR statistic is only a function of the GMM-AR and KLM statistics. The GMM-AR and KLM statistic are not asymptotically independent but the GMM-MLR statistic can as well be specified as a function of

$$
\mathrm{JKLM}(\theta_0)=\text{GMM-AR}(\theta_0)-\mathrm{KLM}(\theta_0), \tag{57}
$$

since

$$\text{GMM-MLR}(\theta_0) = \tfrac{1}{2}\Big[\text{KLM}(\theta_0) + \text{JKLM}(\theta_0) - \text{r}(\theta_0) + , \\ \sqrt{(\text{KLM}(\theta_0) + \text{JKLM}(\theta_0) + \text{r}(\theta_0))^2 - 4\text{JKLM}(\theta_0)\text{r}(\theta_0))}\Big]. \tag{58}$$

Under H_0 and Assumption 1, the JKLM statistic converges to a $\chi^2(k-1)$ distributed random variable which is independent of the $\chi^2(1)$ random variable where the KLM statistic converges to, see Kleibergen (2005,2007). The JKLM statistic can easily be incorporated in the bootstrap algorithm in Section 4 which using Theorem 5 can then also be shown to lead to a higher order efficiency gain for approximating the conditional finite sample distribution of $\text{JKLM}(\theta_0)$. Theorem 2 shows that when Assumption 2 holds that the independence between $(\text{KLM}(\theta),\ \text{JKLM}(\theta))$ and $\text{r}(\theta)$ is up to including the order $\frac{1}{N}$. The bootstrapped critical values of GMM-MLR(θ) that result from bootstrapping KLM(θ) and JKLM(θ) therefore also lead to a higher order efficiency gain for the GMM-MLR statistic since this statistic is just a function of KLM(θ) and JKLM(θ), as $\text{r}(\theta)$ is fixed, and the bootstrap leads to a higher order efficiency gain for both of these statistics.

Corollary 8. *Under H_0, Assumptions 1-3 and when B equals N, the bootstrap critical values of GMM-MLR$^*(\theta_0)$ remove the approximation error of the conditional finite sample distribution of GMM-MLR(θ_0) given r(θ_0) up to/including the order $\frac{1}{N}$.*

We did not pursue constructing an Edgeworth approximation of the conditional finite sample distribution of GMM-MLR(θ) given r(θ) since the analytical expression of the conditional characteristic function of GMM-MLR(θ) given r(θ) is unknown. We can therefore not proof the higher order gains from the bootstrap using the Edgeworth approximation for GMM-MLR(θ). We therefore verify the higher order improvement from usage of bootstrap critical values for GMM-MLR(θ) using the argument that the bootstrap leads to a higher order improvement for (KLM(θ), JKLM(θ)) and GMM-MLR(θ) is, given r(θ), just a function of these two statistics.

Besides using the bootstrap to achieve higher order improvements for approximating the (conditional) finite sample distributions of the GMM-AR and KLM statistics, the Edgeworth approximations from Theorem 3 can be used for this purpose as well. It is interesting to note that the Edgeworth corrections of the critical values for the GMM-AR and KLM statistics are almost identical when the $f_i(\theta)$'s are normally distributed. In that case, $\text{E}(GMM\text{-}AR_1) = k$ and $a = -1$ and $b = 0$ so the corrections of the critical value x is $(1 + \frac{1}{N})x$ for both the GMM-AR and the KLM statistic. We note though that x results from a $\chi^2(k)$ distribution for the GMM-AR statistic while it results from a $\chi^2(1)$ distribution for the KLM statistic.

6 Power and size comparison for Panel AR(1)

We illustrate the size improvements from using bootstrap or Edgeworth-corrected critical values for a panel autoregressive model of order one: panel AR(1). An elaborate literature on panel autoregressive models exists, see *e.g.* Anderson and Hsiao (1981), Arellano and Bond (1991) and Arellano and Honoré (2001). For individual i at time t, the panel AR(1) model reads

$$y_{t,i} = c_i + \theta y_{t-1,i} + \varepsilon_{t,i} \qquad t = 1, \ldots, T,\ i = 1, \ldots, N. \tag{59}$$

A sufficient condition for Assumption 1 to hold is that the disturbances $\varepsilon_{t,i}$ are independently distributed with mean zero and finite eighth order moments. We take first differences to remove the individual specific fixed effects c_i, $i = 1, \ldots, N$:

$$\Delta y_{t,i} = \theta \Delta y_{t-1,i} + \Delta\varepsilon_{t,i} \qquad t = 2, \ldots, T,\ i = 1, \ldots, N, \tag{60}$$

with $\Delta y_{t,i} = y_{t,i} - y_{t-1,i}$. Estimation of the parameter θ in (60) by means of least squares leads to a biased estimator in samples with a finite value of T, see *e.g.* Nickel (1981). We therefore estimate it using GMM. We specify the moment equation (1) for the panel AR(1) using all two period and more lagged level values of $y_{t,i}$ as instruments, see Arellano and Bond (1991). The specification of the moment vector $f_i(\theta)$ then reads

$$f_i(\theta) = X_i \varphi_i(\theta) : \tfrac{1}{2}(T-1)(T-2) \times 1 \qquad i = 1, \ldots, N, \tag{61}$$

with $\varphi_i(\theta) = (\Delta y_{3,i} - \theta \Delta y_{2,i} \ldots \Delta y_{T,i} - \theta \Delta y_{T-1,i})'$ and

$$X_i = \begin{pmatrix} y_{1,i} & 0\ldots0 & 0 \\ 0 & \ddots & 0 \\ & & \begin{pmatrix} y_{1,i} \\ \vdots \\ y_{T-2,i} \end{pmatrix} \\ 0 & 0\ldots0 & \end{pmatrix} : \tfrac{1}{2}(T-1)(T-2) \times (T-2). \tag{62}$$

We use the Eicker-White covariance matrix estimator (4) with $q_i(\theta) = \frac{\partial}{\partial\theta} f_i(\theta) = X_i \Delta y_{-1,i}$ for $\Delta y_{-1,i} = (\Delta y_{2,i} \ldots \Delta y_{T-1,i})'$. Because

$$\begin{aligned} \binom{f_i(\theta)}{q_i(\theta)} &= \binom{X_i(\Delta y_i - \theta \Delta y_{-1,i})}{X_i \Delta y_{-1,i}} \\ &= \left[\begin{pmatrix} 1 & -\theta \\ 0 & 1 \end{pmatrix} \otimes I_{\frac{1}{2}(T-1)(T-2)}\right] \binom{X_i \Delta y_i}{X_i \Delta y_{-1,i}}, \end{aligned} \tag{63}$$

θ_0	N	**GMM-AR**			**KLM**				**GMM-MLR**		**LM**
		A	$*$	E	A	$*$	$**$	E	A	$*$	A
0.5	50	22.6	1.1	20.8	15.1	2.4	2.4	14.9	18.8	1.2	18.0
	100	11.3	2.6	10.6	8.0	3.8	3.6	7.9	9.2	3.4	10.5
	250	8.6	6.1	8.5	6.3	5.1	4.9	6.3	6.8	4.7	6.1
0.7	50	22.9	1.2	21.5	16.1	2.1	2.1	15.6	21.5	1.0	20.5
	100	11.3	3.0	11.0	7.7	3.7	3.7	7.6	9.8	3.5	11.3
	250	8.9	5.6	8.9	6.1	5.6	5.7	6.0	6.7	5.5	6.5
0.9	50	22.8	1.0	21.6	15.3	1.8	2.4	15.0	26.0	1.1	37.0
	100	11.3	3.2	10.6	9.6	4.5	3.7	9.3	13.9	2.4	24.6
	250	8.6	6.5	8.6	6.7	4.9	6.2	6.7	9.6	7.3	13.4
0.95	50	22.2	1.0	21.3	14.6	1.6	2.1	14.3	25.3	1.0	44.1
	100	11.0	3.5	10.8	10.5	4.3	5.5	10.2	15.2	3.5	34.0
	250	8.5	5.9	8.5	7.1	5.5	6.2	7.1	10.7	6.3	24.2

Table 1: Size of the different statistics in percentages that test $\mathrm{H}_0 : \theta = \theta_0$ at the 95% significance level. A: asymptotic critical values, $*$ bootstrapped critical values, E Edgeworth-corrected critical values, $**$ bootstrap critical values where also the $q_i(\theta)$'s are resampled.

with $\Delta y_i = (\Delta y_{3,i} \dots \Delta y_{T,i})'$, and X_i (62) consists of lagged values of $y_{t,i}$,

$$\hat{V}(\theta) = \begin{pmatrix} \hat{V}_{ff}(\theta) & \hat{V}_{qf}(\theta)' \\ \hat{V}_{qf}(\theta) & \hat{V}_{qq}(\theta) \end{pmatrix} \tag{64}$$

is singular since some of the elements of $\binom{X_i \Delta y_i}{X_i \Delta y_{-1,i}}$ are identical. We therefore obtain $[\hat{V}_{qq}(\theta) - \hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}\hat{V}_{fq}(\theta)]^{-1}$, that is involved in $\mathrm{r}(\theta)$ and thus in GMM-MLR(θ), from a generalized inverse of $\hat{V}(\theta)$.

The derivative of the moments, $q_i(\theta) = X_i \Delta y_{-1,i}$, is a white noise series when $\theta = 1$. The parameter θ is therefore not identified in the moment equations when it is equal to one. Weakly identified values of θ occur when θ is close to one relative to the sample size, *i.e.* when $\frac{1-\theta}{N}$ is small. This implies that the LM(θ_0) statistic from Definition 1 becomes size distorted when θ_0 is close to one relative to the sample size.

Size results We compute the size of the GMM-AR, KLM and GMM-MLR statistics using asymptotic, bootstrapped and Edgeworth-corrected 95% critical values in a simulation experiment that uses the previously discussed panel AR(1) model. To illustrate the sensitivity of the size of the LM(θ) statistic to the value of θ, we also compute the size of LM(θ) using its asymptotic critical value.

The panel AR(1) model has independent disturbances which are generated from a standard normal distribution so Assumption 1 is satisfied. We use the standard normal distribution since it leads to a straightforward expression for the Edgeworth corrected critical values. The individual specific constants c_i are specified as $c_i = (1-\theta)\mu_i$ where the μ_i's are independent realizations from a $N(0, 2)$ distribution. The initial observations $y_{0,i}$ are simulated such that $y_{0,i} = \mu_i + \varepsilon_{0,i}$ where the $\varepsilon_{0,i}$'s are independent realizations of standard normal random variables. The number of simulated datasets equals one thousand.

We compute the bootstrap critical values that result from GMM-AR$^*(\theta)$, KLM$^*(\theta)$ and KLM$^{**}(\theta)$ using one hundred simulations of bootstrap datasets of N observations. The bootstrap datasets are obtained using the algorithms in Section 4.

The Edgeworth-corrected critical values are computed using the Edgeworth approximations from Theorem 5. Because the disturbances are standard normal, $\mathrm{E}(GMM\text{-}AR_1) = k$, $a = -1$, $b = 0$ and $c = 0$. Hence the Edgeworth corrected critical values are $(1+\frac{1}{N})$ times the asymptotic critical values for both GMM-AR(θ) and KLM(θ).

Table 1 shows the observed size of the statistics when we test at the 95% significance level in a simulation experiment that uses four different values of θ_0 : 0.5, 0.7, 0.9 and 0.95 and three different values of N : 50, 100, 250. The number of time series observations, T, is equal to 6 for all cases.

The observed sizes reported in Table 1 show that usage of the critical values that stem from the asymptotic distributions leads to large size distortions in small samples. Both the Edgeworth correction and the bootstrap decrease these size distortions in all cases. The reduction of the size distortion that results from the bootstraps is, however, much larger than the one that results from the Edgeworth expansion. The reductions result from the higher order improvements that result from using the bootstrap or the Edgeworth corrections of the critical values.

The observed size of KLM(θ) reported in Table 1 that results from using critical values that stem from the bootstrap that resamples the $q_i(\theta)$'s $(**)$, Algorithm 2, is almost identical to the size that results from using the critical values from the bootstrap which uses $\hat{D}_N(\theta, Y)$ $(*)$, Algorithm 1. This holds since resampling $q_i(\theta)$ only effects size distortions which are of order $(N\sqrt{N})^{-1}$ while resampling $f_i(\theta)$ effects size distortions which are of order N^{-1}. The resampling of the $q_i(\theta)$'s is therefore of lesser importance and does not lead to any further size improvements. Thus the size distortions that result from the estimation of the covariance matrix, which is a 10×10 matrix in the simulation experiment, exceed those that result from $\hat{D}_N(\theta, Y)$.

The size distortions of GMM-MLR(θ) can exceed those of GMM-AR(θ) and KLM(θ) when we use the critical values that stem from its conditional limiting distribution. The bootstrap reduces these size distortions which, however, remain larger than those for KLM(θ) and sometimes also larger than those for GMM-AR(θ). This results since for smaller values of θ_0, GMM-MLR(θ) is close to KLM(θ), while it is close to GMM-AR(θ) for values of θ_0 close to one. The size distortions for KLM(θ) are in general the smallest both when using asymptotic and bootstrap critical values.

The size distortions of LM(θ) show the sensitivity of its distribution to the value of θ_0. Table 1 clearly shows that the size distortions increase when θ_0 gets closer to one. An increase of the sample size for the same value of θ_0 decreases the size distortion of LM(θ). The same results can be shown for other statistics whose distributions are sensitive to the value of θ_0, like, for example, the Wald t-statistic. For reasons of brevity, we refrain from showing these results.

Insert Panel 1 around here.

Insert Panel 2 around here.

Insert Panel 3 around here.

Power comparison To further analyse the performance of the different statistics, we compare the power of the different statistics for the three different values of N : 50, 100 and 250, and three of the four different values of θ_0 : 0.5, 0.9 and 0.95, that were used in Table 1. Panels 1-3 in the Figures section show these power curves. All Panels use a value of T equal to 6 while $N = 50$ in Panel 1, $N = 100$ in Panel 2 and $N = 250$ in Panel 3.

All power curves in Panel 1 reveal the large size distortions of the different statistics when we use the asymptotic critical values. Usage of Edgeworth-corrected critical values shifs the power curve downwards but, as already shown in Table 1, not enough to completely remove the size distortions. Usage of bootstrap critical values makes the statistics too conservative but the size distortion is much smaller than when using the Edgeworth-corrected critical values. The power curves show that the statistics have power when using the bootstrap critical values.

The power curves of the KLM statistic in Panel 1 when using the bootstrap with or without resampled $q_i(\theta)$'s are almost indistinguisable. This is in line with Table 1 and shows that there is no need to resample $q_i(\theta)$ in our simulation experiment since Assumption 2 is satisfied.

When $\theta = 1$, the moment conditions do not identify θ and the power and size of the different statistics should coincide. This explains the peculiar shape of the power curves in Panel 1. It is interesting to see that the size distortions at $\theta = \theta_0$ are the same as at $\theta = 1$ of all statistics except LM(θ) which is the only statistic whose distribution depends on the value of θ.

Panel 1 shows that the size distortions of all statistics except LM(θ) just depend on N and not on θ_0. For LM(θ_0), Figures 1.7-1.9 nicely show the increase of the size distortion when we increase θ_0.

In Panel 2, all size distortions have decreased compared to Panel 1. Besides that most general findings from Panel 1 appear in Panel 2 as well: bootstrap power curves lie below the power curves that result from using the asymptotic critical value, resampling $q_i(\theta)$ for the bootstrapped power curve of KLM(θ) does not change it, power and size coincide and $\theta = \theta_0$ and $\theta = 1$ except for LM(θ), size distortions of all statistics except LM(θ) do not depend on θ_0.

Panel 3 shows that the power curves that result from using a different critical value become almost indistinguishable when $N = 250$. Usage of bootstrap critical values still leads to a smaller size distortion but the size distortion from using the asymptotic critical value is now rather small as well. It is interesting to see that the size distortion of LM(θ) is now small for $\theta_0 = 0.5$ but is still very large for $\theta_0 = 0.95$ which shows the sensitivity of the distribution of LM(θ) to the value of θ_0.

7 Extensions

Multiple parameters and subsets Sofar we have just been concerned with testing a parameter vector that consists of only one element. This has been done for expository purposes and the results extend to a vector of multiple parameters as well. This holds true since one of the main results of the paper that the N^{-1} approximation errors just result from the covariance matrix estimator extend towards a multiple parameter setting when Assumption 1b is extended appropriately. We refrain from showing this result since it involves a lot of additional notation as we have to use vectorization results for matrices.

The extension to tests on subsets of the parameters is less straightforward because the limiting distributions of the statistics only extend appropriately when the partialled out parameters are well identified, see Kleibergen (2005). Recently, Kleibergen and Mavroeidis (2010) have, however, shown for the linear instrumental variables regression model that the limiting distributions that result under strong identification assumptions on the partialled out parameters bound the limiting distributions in all other cases. This suggest that the bootstrap might improve on the approximation that results from the bounding limiting distribution.

Dependent observations If the moment equations are dependent, the Eicker-White covariance matrix estimator (4) is not consistent so our results do not apply to that case. It would therefore be of interest to analyze how to extend our results to dependent data.

8 Conclusions

We show that usage of Edgeworth-corrected critical values or bootstrapped critical values leads to a higher order improvement of the approximation of the finite sample distribution of weak instrument robust GMM statistics compared to usage of critical values that stem from their limiting distributions. These improvements remain to hold when the instruments are weak. This extends previous work on the Edgeworth approximation and bootstrap which does not allow for weak instruments. When the moments are homoscedastic, we show that it suffices to resample just the moments under the hypothesis of interest since the covariance matrix estimator is then the sole contributor to the approximation error of the highest order. It is then not necssary to resample anything but the sample moments.

Figures

Panel 1: Power curves for $T = 6,\ N = 50$.

Fig. 1.1-1.3: Power of GMM-AR when using asymptotic (solid), bootstrapped (dashed) and Edgeworth-corrected (dashed-dotted) critical values.

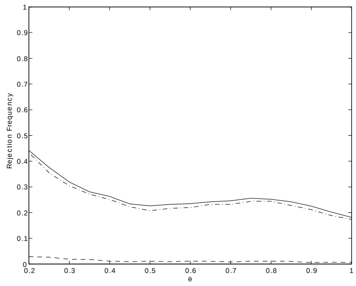

Fig 1.1: $\theta_0 = 0.5$

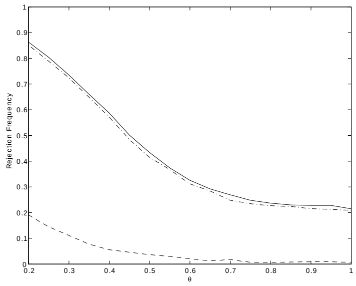

Fig. 1.2: $\theta_0 = 0.9$

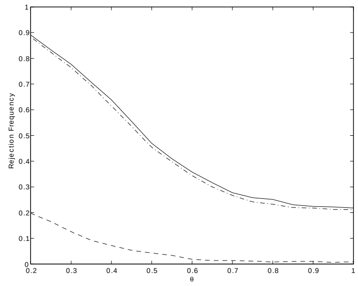

Fig. 1.3: $\theta_0 = 0.95$

Fig. 1.4-1.6: Power of KLM when using asymptotic (solid), bootstrapped (dashed), with resampled $q_i(\theta)$ (dotted) and Edgeworth-corrected (dashed-dotted) critical values.

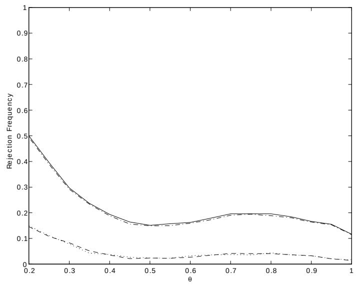

Fig 1.4: $\theta_0 = 0.5$

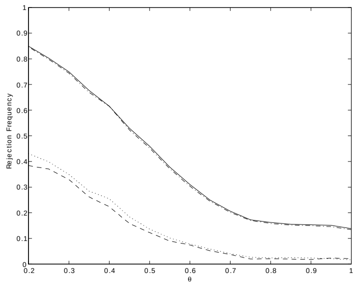

Fig. 1.5: $\theta_0 = 0.9$

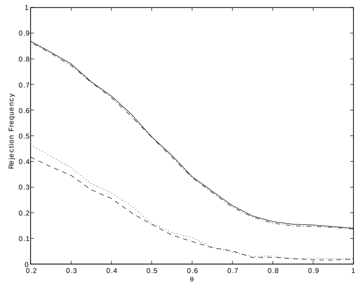

Fig. 1.6: $\theta_0 = 0.95$

Fig. 1.7-1.9: Power of GMM-MLR when using conditional asymptotic (solid) and bootstrapped (dashed) critical values and LM (dashed-dotted).

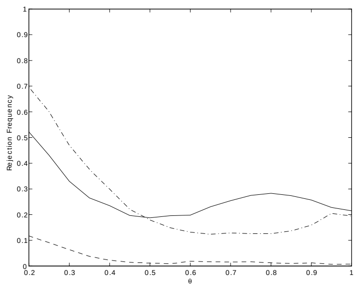

Fig 1.7: $\theta_0 = 0.5$

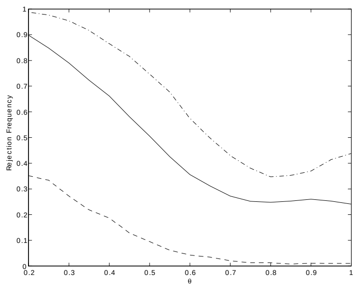

Fig. 1.8: $\theta_0 = 0.9$

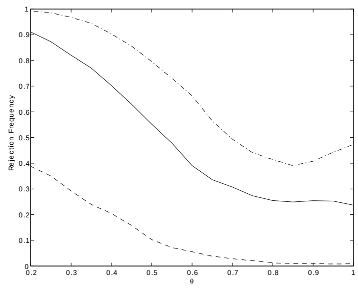

Fig. 1.9: $\theta_0 = 0.95$

Panel 2: Power curves for $T = 6,\ N = 100$.

Fig. 2.1-2.3: Power of GMM-AR when using asymptotic (solid), bootstrapped (dashed) and Edgeworth-corrected (dashed-dotted) critical values.

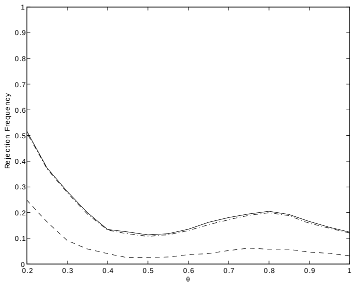

Fig 2.1: $\theta_0 = 0.5$

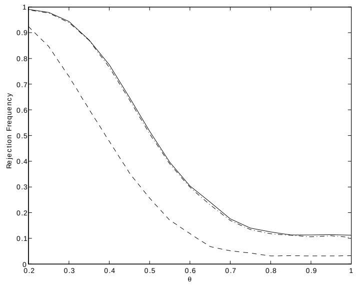

Fig. 2.2: $\theta_0 = 0.9$

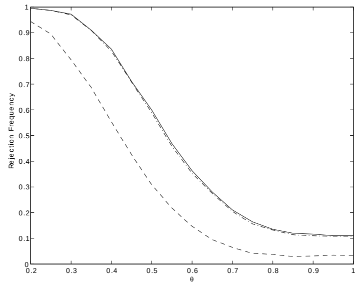

Fig. 2.3: $\theta_0 = 0.95$

Fig. 2.4-2.6: Power of KLM when using asymptotic (solid), bootstrapped (dashed), with resampled $q_i(\theta)$ (dotted) and Edgeworth-corrected (dashed-dotted) critical values.

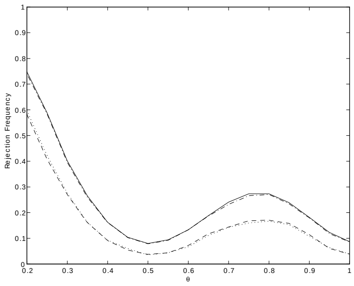

Fig 2.4: $\theta_0 = 0.5$

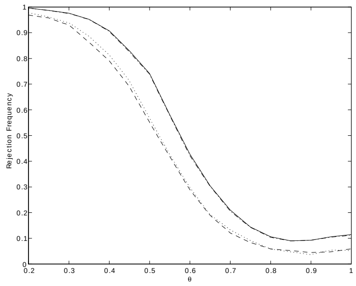

Fig. 2.5: $\theta_0 = 0.9$

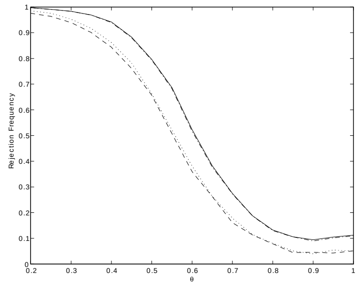

Fig. 2.6: $\theta_0 = 0.95$

Fig. 2.7-2.9: Power of GMM-MLR when using conditional asymptotic (solid) and bootstrapped (dashed) critical values and LM (dashed-dotted).

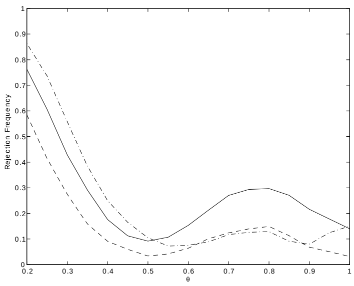

Fig 2.7: $\theta_0 = 0.5$

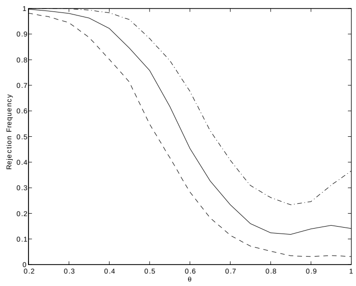

Fig. 2.8: $\theta_0 = 0.9$

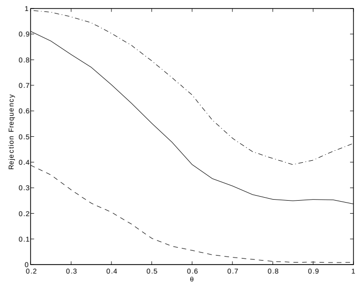

Fig. 2.9: $\theta_0 = 0.95$

Panel 3: Power curves for $T = 6,\ N = 250$.

Fig. 3.1-3.3: Power of GMM-AR when using asymptotic (solid), bootstrapped (dashed) and Edgeworth-corrected (dashed-dotted) critical values.

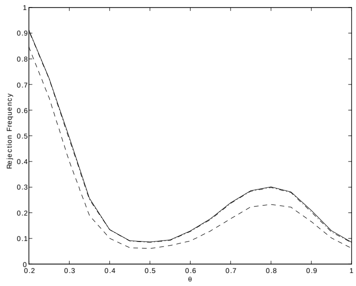

Fig 3.1: $\theta_0 = 0.5$

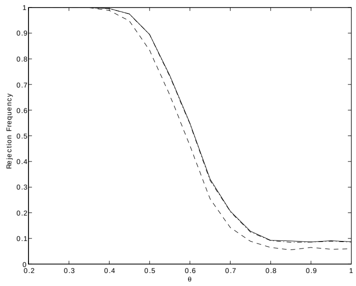

Fig. 3.2: $\theta_0 = 0.9$

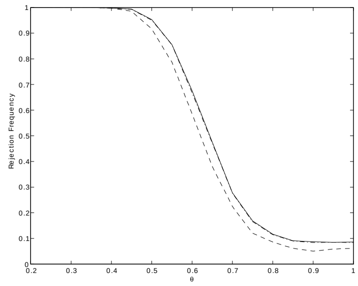

Fig. 3.3: $\theta_0 = 0.95$

Fig. 3.4-3.6: Power of KLM when using asymptotic (solid), bootstrapped (dashed), with resampled $q_i(\theta)$ (dotted) and Edgeworth-corrected (dashed-dotted) critical values.

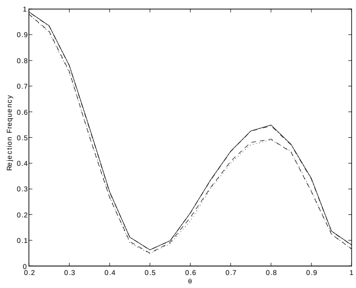

Fig 3.4: $\theta_0 = 0.5$

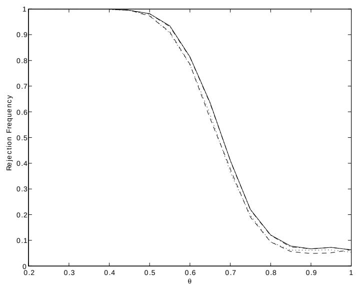

Fig. 3.5: $\theta_0 = 0.9$

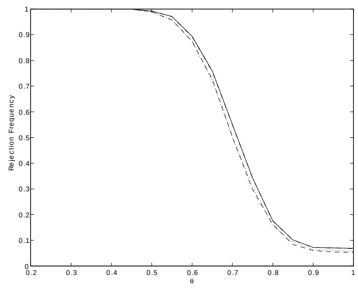

Fig. 3.6: $\theta_0 = 0.95$

Fig. 3.7-3.9: Power of GMM-MLR when using conditional asymptotic (solid) and bootstrapped (dashed) critical values and LM (dashed-dotted).

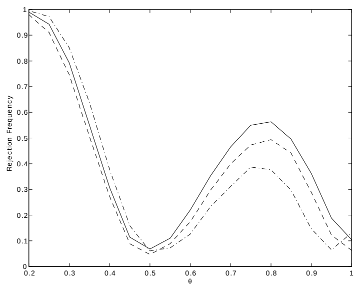

Fig 3.7: $\theta_0 = 0.5$

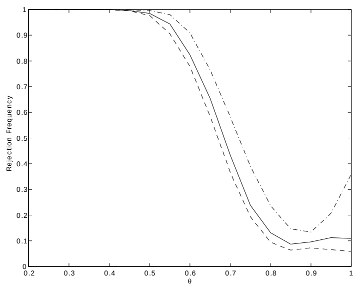

Fig. 3.8: $\theta_0 = 0.9$

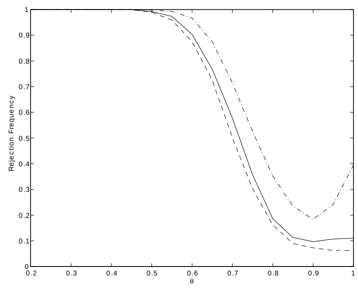

Fig. 3.9: $\theta_0 = 0.95$

Appendix

A. Lemma 1. *When Assumption 1 holds, the higher order decomposition of $\hat{V}_{ff}(\theta)^{-1}$ reads*

$$\hat{V}_{ff}(\theta)^{-1} = V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + \\ V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + o_p(N^{-1})$$

with θ equal to its true value.

Proof. To obtain the higher order specification of $\hat{V}_{ff}(\theta)^{-1}$, we specify it as

$$\begin{aligned}
\hat{V}_{ff}(\theta)^{-1} = &\left[V_{ff}(\theta) + \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)\right]^{-1} \\
= &V_{ff}(\theta)^{-\frac{1}{2}}\left[I_k + V_{ff}(\theta)^{-\frac{1}{2}}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-\frac{1}{2}}\right]^{-1}V_{ff}(\theta)^{-\frac{1}{2}} \\
= &V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + V_{ff}(\theta)^{-1} \\
&\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + o_p(N^{-1}),
\end{aligned}$$

where the $o_p(N^{-1})$ order of the remainder term results from the $\sqrt{N}$ convergence rate of the Eicker-White covariance matrix estimator. ■

Lemma 2. *When Assumption 1 holds, the higher order specification of $\sqrt{N}\hat{D}_N(\theta, Y)$ reads:*

$$\sqrt{N}\hat{D}_N(\theta, Y) = \sqrt{N}D_N(\theta, Y) - \hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) + \\ \hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) + o_p(N^{-1}),$$

where $\hat{V}_{\theta f}(\theta) = \hat{V}_{qf}(\theta) - V_{qf}(\theta)V_{ff}(\theta)^{-1}\hat{V}_{ff}(\theta) = \frac{1}{N}\sum_{i=1}^{N} d_i(\theta)f_i(\theta)' - D_N(\theta, Y)f_N(\theta, Y)'$, $\hat{V}_{\theta f}(\theta) \underset{p}{\rightarrow} 0$, $\sqrt{N}\hat{V}_{\theta f}(\theta) = O_p(1)$ and θ is the true value of θ.

Proof. To obtain the higher order specification, we specify $\hat{D}_N(\theta, Y)$ as

$$\begin{aligned}
&\sqrt{N}\hat{D}_N(\theta, Y) = \sqrt{N}q_N(\theta, Y) - \sqrt{N}\hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}f_N(\theta, Y) \\
= &\sqrt{N}D_N(\theta, Y) - \left[\hat{V}_{qf}(\theta) - V_{qf}(\theta)V_{ff}(\theta)^{-1}\hat{V}_{ff}(\theta)\right]\hat{V}_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) \\
= &\sqrt{N}D_N(\theta, Y) - \left[\hat{V}_{qf}(\theta) - V_{qf}(\theta)V_{ff}(\theta)^{-1}\hat{V}_{ff}(\theta)\right][V_{ff}(\theta)^{-1} - \\
&V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\
&V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big]\left(\sqrt{N}f_N(\theta, Y)\right) + o_p(N^{-1}) \\
= &\sqrt{N}D_N(\theta, Y) - \hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) + \\
&\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) + o_p(\tfrac{1}{N}),
\end{aligned}$$

where we used Lemma 1 for the fourth equation and note that $\sqrt{N}D_N(\theta, Y)$ is at least $O_p(1)$. The order of the remainder terms results from the $\sqrt{N}$ convergence rate of the Eicker-White covariance matrix estimators $\hat{V}_{\theta f}(\theta)$ and $\hat{V}_{ff}(\theta)$. ∎

Lemma 3. *When Assumption 1 holds, the higher order specification of the score vector* $s_N(\theta, Y) = \hat{D}_N(\theta, Y)'\hat{V}_{ff}(\theta_0)^{-1}f_N(\theta, Y)$ *reads*

$$N s_N(\theta, Y) = s_0 + \tfrac{1}{\sqrt{N}} s_1 + \tfrac{1}{N} s_2 + o_p(N^{-1}),$$

with

$$\begin{aligned}
s_0 = & \left[\sqrt{N}D_N(\theta, Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) \\
\sqrt{N}s_1 = & -\left[\sqrt{N}D_N(\theta, Y)\right]' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) - \\
& \left(\sqrt{N}f_N(\theta, Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) \\
N s_2 = & \left[\sqrt{N}D_N(\theta, Y)\right]' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) + \left(\sqrt{N}f_N(\theta, Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right) + \left(\sqrt{N}f_N(\theta, Y)\right)' \\
& V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta, Y)\right),
\end{aligned}$$

θ the true value of θ and s_0, s_1 and s_2 are all at most $O_p(1)$.

Proof. Using Lemmas 1 and 2, the decomposition of the score vector reads

$$
\begin{aligned}
& Ns_N(\theta,Y) \\
= {} & \left(\sqrt{N}\hat{D}_N(\theta,Y)\right)' \hat{V}_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) \\
= {} & \Big[\sqrt{N}D_N(\theta,Y) - \hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + \hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)\Big]' [V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big]\left(\sqrt{N}f_N(\theta,Y)\right) + o_p(N^{-1}) \\
= {} & \left[\sqrt{N}D_N(\theta,Y)\right]'\Big[V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + \\
& V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big]\left(\sqrt{N}f_N(\theta,Y)\right) - \\
& \left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'\left[V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\right] \\
& \left(\sqrt{N}f_N(\theta,Y)\right) + \left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
& \hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + o_P(N^{-1}) \\
= {} & \left[\sqrt{N}D_N(\theta,Y)\right]'\Big[V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} + \\
& V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big]\left(\sqrt{N}f_N(\theta,Y)\right) - \\
& \sqrt{N}f_N(\theta,Y)V_{ff}(\theta)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + \\
& \left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
& \left(\sqrt{N}f_N(\theta,Y)\right) + \left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
& \hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + o_P(N^{-1}),
\end{aligned}
$$

where the order of the remainder term results from the convergence speed of the covariance matrix estimator. ■

Lemma 4. *When Assumption 1 holds, the higher order specification of* $\left[N\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right]^{-1}$ *reads:*

$$\left(N\hat{D}_N(\theta,Y)'V_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right)^{-1} = D_0 + \tfrac{1}{\sqrt{N}}D_1 + \tfrac{1}{N}D_2 + o_p(N^{-1}),$$

with

$$
\begin{aligned}
D_0 = \ & (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
\sqrt{N}D_1 = \ & (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
& \Big[2\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) + \\
& \left(\sqrt{N}D_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\Big] \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
ND_2 = \ & -(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
& \Big[\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + \\
& \left(\sqrt{N}D_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) + 2\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) + 2\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) + 4\left(\sqrt{N}f_N(\theta,Y)\right)' \\
& V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + \\
& \left(\sqrt{N}D_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\Big]\ (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1},
\end{aligned}
$$

θ the true value of θ and D_0, D_1 and D_2 are all at most $O_p(1)$.

Proof. In order to construct the higher order expression of $\left[\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right]^{-1}$, we first use Lemmas 1 and 2 to construct that of $\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)$:

$$
\begin{aligned}
& N\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y) \\
= \; & \left(\sqrt{N}\hat{D}_N(\theta,Y)\right)'\hat{V}_{ff}(\theta)^{-1}\left(\sqrt{N}\hat{D}_N(\theta,Y)\right) \\
= \; & \Big[\sqrt{N}D_N(\theta,Y)-\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)\Big]'\Big[V_{ff}(\theta)^{-1}-V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}+V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big] \\
& \Big[\sqrt{N}D_N(\theta,Y)-\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)\Big]+o_p(N^{-1}) \\
= \; & \left(\sqrt{N}D_N(\theta,Y)\right)'\Big[V_{ff}(\theta)^{-1}-V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}+V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big]\left(\sqrt{N}D_N(\theta,Y)\right)- \\
& 2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'\Big[V_{ff}(\theta)^{-1}-V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big] \\
& \left(\sqrt{N}D_N(\theta,Y)\right)+2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)' \\
& V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+ \\
& \left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+o_p(N^{-1}) \\
= \; & \left(\sqrt{N}D_N(\theta,Y)\right)'\Big[V_{ff}(\theta)^{-1}-V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}+ \\
& V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big]\left(\sqrt{N}D_N(\theta,Y)\right)- \\
& 2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+ \\
& \left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+ \\
& 2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
& \left(\sqrt{N}D_N(\theta,Y)\right)+2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1} \\
& \hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+o_P(N^{-1})
\end{aligned}
$$

so

$$
\begin{aligned}
&\left(N\hat{D}_N(\theta,Y)'V_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right)^{-1} \\
= \; & (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} + (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
& \Big[2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+ \\
& \left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\Big] \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} - (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
& \Big[\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+ \\
& \left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+4\left(\sqrt{N}f_N(\theta,Y)\right)' \\
& V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1} \\
& \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+ \\
& \left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\Big]\,(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}+o_p(N^{-1})
\end{aligned}
$$

where we used that $(A_0+\frac{1}{\sqrt{N}}A_1+\frac{1}{N}A_2+o_p(N^{-1}))^{-1}=A_0^{-1}-\frac{1}{\sqrt{N}}A_0^{-1}A_1A_0^{-1}+\frac{1}{N}A_0^{-1}(A_1A_0^{-1}A_1-A_2)+o_p(N^{-1})$.

Using the higher order decomposition of $\hat{V}_{ff}(\theta)^{-1}$ from Lemma 1, the higher order expression for $\left[N\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right]^{-1}$ results. ∎

Lemma 5. *When Assumption 1 holds, the higher order specification of*

$$N f_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y)$$

reads:

$$
\begin{aligned}
&N f_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y) \\
= \; & N f_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)+N f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)- \\
& N f_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y)+O_p(\tfrac{1}{N}),
\end{aligned}
$$

with θ the true value of θ.

Proof. The higher order decomposition results from the specification of $D_N(\theta,Y)$ and $\hat{V}_{ff}(\theta)^{-1}$:

$$\begin{array}{l} Nf_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y) \\ = \quad Nf_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}D_N(\theta,Y) + Nf_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) \\ = \quad Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y) + Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) - \\ \quad Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y) - Nf_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1} \\ \quad V_{qf}(\theta)V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}f_N(\theta,Y) + O_p(\frac{1}{N}). \end{array}$$

■

Lemma 6. *When Assumption 1 holds, the higher order specification of* $\left[Nq_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y)\right]^{-1}$ *reads:*

$$\begin{array}{l} \left[Nq_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y)\right]^{-1} \\ = \quad (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} - (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\ \quad \{2ND_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) + \\ \quad Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) - ND_N(\theta,Y)'V_{ff}(\theta)^{-1} \\ \quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y) - 2ND_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1} \\ \quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}f_N(\theta,Y) - Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1} \\ \quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)\Big\} \\ \quad (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} + O_p(\frac{1}{N}), \end{array}$$

with θ the true value of θ.

Proof. We obtain the higher order specification of $\left[Nq_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y)\right]^{-1}$ from that of $Nq_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y)$:

$$\begin{array}{l} Nq_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y) \\ = \quad \left[\sqrt{N}D_N(\theta,Y) + \sqrt{N}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)\right]'\{V_{ff}(\theta)^{-1} - V_{ff}(\theta)^{-1} \\ \quad \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\Big\}\left[\sqrt{N}D_N(\theta,Y) + \sqrt{N}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)\right] + O_p(\frac{1}{N}) \\ = \quad ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y) + 2ND_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) + \\ \quad Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) - ND_N(\theta,Y)'V_{ff}(\theta)^{-1} \\ \quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y) - 2ND_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1} \\ \quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}f_N(\theta,Y) - Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1} \\ \quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) + O_p(\frac{1}{N}). \end{array}$$

■

Lemma 7. When Assumption 1 holds, the higher order specification of

$$\left[\hat{V}_{qq}(\theta) - \hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}\hat{V}_{fq}(\theta)\right]^{-1}$$

reads

$$V_{dd}(\theta)^{-1} - V_{dd}(\theta)^{-1}\left[\hat{V}_{dd}(\theta) - V_{dd}(\theta)\right]V_{dd}(\theta)^{-1} + O_p(\tfrac{1}{N}).$$

Proof. We obtain the higher order specification of $\left[\hat{V}_{qq}(\theta) - \hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}\hat{V}_{fq}(\theta)\right]^{-1}$ from that of $\hat{V}_{qq}(\theta) - \hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}\hat{V}_{fq}(\theta)$:

$$\begin{aligned}
&\hat{V}_{qq}(\theta) - \hat{V}_{qf}(\theta)\hat{V}_{ff}(\theta)^{-1}\hat{V}_{fq}(\theta) \\
&= V_{qq}(\theta) + \left[\hat{V}_{qq}(\theta) - V_{qq}(\theta)\right] - \left(V_{qf}(\theta) + \left[\hat{V}_{qf}(\theta) - V_{qf}(\theta)\right]\right) \\
&\quad \left(V_{ff}(\theta) + \left[\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right]\right)^{-1}\left(V_{fq}(\theta) + \left[\hat{V}_{fq}(\theta) - V_{fq}(\theta)\right]\right) \\
&= V_{qq}(\theta) - V_{qf}(\theta)V_{ff}(\theta)^{-1}V_{fq}(\theta) + \left[\hat{V}_{qq}(\theta) - V_{qq}(\theta)\right] - \left[\hat{V}_{qf}(\theta) - V_{qf}(\theta)\right] \\
&\quad V_{ff}(\theta)^{-1}V_{fq}(\theta) - V_{qf}(\theta)V_{ff}(\theta)^{-1}\left[\hat{V}_{fq}(\theta) - V_{fq}(\theta)\right] + V_{qf}(\theta)V_{ff}(\theta)^{-1} \\
&\quad \left[\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}V_{qf}(\theta) + O_p(\tfrac{1}{N}) \\
&= V_{\theta\theta.f}(\theta) + \hat{V}_{dd}(\theta) - V_{\theta\theta.f}(\theta_0) + O_p(\tfrac{1}{N})
\end{aligned}$$

with $V_{\theta\theta.f}(\theta_0) = V_{qq}(\theta_0) - V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}V_{fq}(\theta_0)$, $\hat{V}_{dd}(\theta_0) = \frac{1}{N}\sum_{i=1}^{N}\bar{d}_i(\theta_0)\bar{d}_i(\theta_0)' = \hat{V}_{qq}(\theta_0) - \hat{V}_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}V_{fq}(\theta_0) - V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}\hat{V}_{fq}(\theta_0) + V_{qf}(\theta_0)V_{ff}(\theta_0)^{-1}\hat{V}_{ff}(\theta_0)^{-1}$ $V_{ff}(\theta_0)V_{fq}(\theta_0)$. ■

B. Proof of Theorem 1.

a. GMM–AR statistic: We use the decomposition of $\hat{V}_{ff}(\theta)^{-1}$ from Lemma 1 to obtain the higher order components of the GMM-AR statistic:

$$\begin{aligned}
\text{GMM-AR}(\theta) &= Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y) - Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right] \\
&\quad V_{ff}(\theta)^{-1}f_N(\theta,Y) + Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\
&\quad V_{ff}(\theta)^{-1}f_N(\theta,Y) + O_p(\tfrac{1}{N^2}).
\end{aligned}$$

To determine the order of remainder term and of the different components, we construct their expectations. It is obvious that the expectation of $Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)$ is equal to k since $\text{E}(f_i(\theta)f_j(\theta)') = V_{ff}(\theta)$ when $i = j$ and equals zero when $i \neq j$.

We next construct the expectations of the two additional components in the above expression:

$$
\begin{array}{rl}
\multicolumn{2}{l}{\mathrm{E}\left[Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}f_N(\theta,Y)\right]} \\
= & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_3} f_{i_3}\right] \\
= & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\right]-\mathrm{E}\left[\frac{1}{N}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_4} f_{i_4}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\right] \\
= & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_2}'V_{ff}^{-1}f_{i_1}\right]+\mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_1}\right]- \\
 & \mathrm{E}\left[\frac{1}{N}\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}\right]-\mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_2}'V_{ff}^{-1}f_{i_2}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_1}'V_{ff}^{-1}f_{i_2}\right]-\mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_2}'V_{ff}^{-1}f_{i_1}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_1}\right] \\
= & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_1}\right]+\frac{N-1-N}{N}\mathrm{E}\left[\frac{1}{N}\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_2}'V_{ff}^{-1}f_{i_2}\right]-\mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_1}'V_{ff}^{-1}f_{i_2}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_2}'V_{ff}^{-1}f_{i_1}\right]-\mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_1}\right] \\
= & \frac{N-1}{N^2}\mathrm{vec}(V_{ff}^{-1})'\mathrm{E}\left[(f_if_i'\otimes f_if_i')\right]\mathrm{vec}(V_{ff}^{-1})-\frac{k}{N}-\frac{N-1}{N^2}\left[k^2+2k\right],
\end{array}
$$

where $f_i=f_i(\theta)$, $\bar{f}_i=f_i(\theta)-f_N(\theta,Y)$, $V_{ff}=V_{ff}(\theta)$ and $\sum_{i_1}=\sum_{i_1=1}^N$. The above expression consists of second and fourth order moments of f_i. It uses the independence of f_{i_1} and f_{i_2} for $i_1\neq i_2$ and that $\mathrm{E}(f_i)=0$ which explains why no third order moments are present in the expressions.

$$
\begin{array}{rl}
\multicolumn{2}{l}{Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}f_N(\theta,Y)} \\
= & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] \\
= & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] \\
= & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+ \\
 & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+ \\
 & \mathrm{E}\left[\frac{1}{N^5}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_6} f_{i_6}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
 & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+ \\
 & \mathrm{E}\left[\frac{1}{N}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+\mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]
\end{array}
$$

We next isolate the components that are of order $1/N$ for each of the nine elements in the above expression:

$$
\begin{aligned}
1.\quad & \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}f'_{i_2}V_{ff}^{-1}\sum_{i_3} f_{i_3}f'_{i_3}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_3}V_{ff}^{-1}f_{i_1}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}\right] + O(N^{-2}) \\
= \quad & \tfrac{(N-1)(N-2)}{N^2}k + 3\mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + O(N^{-2}) \\
= \quad & \tfrac{(N-1)(N-2)}{N^2}k + 3\mathrm{E}\left[\tfrac{N-1}{N^3}\textstyle\sum_{i_2} f_{i_2}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + O(N^{-2}) \\
2.\quad & \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}f'_{i_2}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}\right] + O(N^{-2}) \\
= \quad & \tfrac{(N-1)}{N}k + \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}\right] + O(N^{-2})
\end{aligned}
$$

$$
\begin{array}{ll}
3. & \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}f'_{i_2}V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \\
& \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_3}V_{ff}^{-1}f_{i_1}\right] + O(N^{-2}) \\
& \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}f'_{i_3}V_{ff}^{-1}f_{i_3}\right] + \\
& \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_1}V_{ff}^{-1}f_{i_3}\right] + O(N^{-2}) \\
= & \frac{(N-1)(N-2)}{N^3}(k^2+2k) + O(N^{-2}) \\
4. & \mathrm{E}\left[\frac{1}{N^4}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_3} f_{i_3}f'_{i_3}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \\
& \frac{(N-1)(N-2)}{N^3}(k^2+2k) + O(N^{-2}) \\
5. & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \\
& \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_1}V_{ff}^{-1}f_{i_2}\right] + \\
& \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}\right] \\
= & \frac{N-1}{N}(k^2+2k) + O(N^{-2}) \\
6. & \mathrm{E}\left[\frac{1}{N^5}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_6} f'_{i_6}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = O(N^{-2}) \\
7. & \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_3} f_{i_3}f'_{i_3}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \\
& \frac{(N-1)}{N}k + \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}\right] + O(N^{-2}) \\
8. & \mathrm{E}\left[\frac{1}{N}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = k \\
9. & \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] = \\
& \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_1}V_{ff}^{-1}f_{i_2}\right] + \\
& \mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}\right] + O(N^{-2}) \\
= & \frac{N-1}{N^2}(k^2+2k) + O(N^{-2}).
\end{array}
$$

Combining, we get:

$$
\begin{array}{lr}
1-2-3-4+5+6-7+8+9 & = \\
\frac{(N-1)(N-2)}{N^2}k + 3\mathrm{E}\left[\left(\frac{1}{N^2}-\frac{1}{N^3}\right)\sum_{i_2} f_{i_2}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] + & \\
\mathrm{E}\left[\frac{1}{N^3}\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}\right] - \frac{(N-1)}{N}k - & \\
\mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}\right] - 2\frac{(N-1)(N-2)}{N^3}(k^2+2k) + \frac{N-1}{N^2}(k^2+2k) - & \\
\frac{(N-1)}{N}k - \mathrm{E}\left[\frac{1}{N^2}\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}\right] + k + \frac{N-1}{N^2}(k^2+2k) & = \\
-\frac{k}{N} + \frac{1}{N}\mathrm{E}\left[f'_iV_{ff}^{-1}f_if'_iV_{ff}^{-1}f_i\right] + \frac{1}{N}\mathrm{E}\left[f'_iV_{ff}^{-1}f_if'_i\right]V_{ff}^{-1}\mathrm{E}\left[f_if'_iV_{ff}^{-1}f_i\right] + O(N^{-2}). &
\end{array}
$$

We combine this with the results from the first expression so:

$$
\begin{array}{l}
\mathrm{E}(\text{GMM-AR}(\theta)) = \\
k + \frac{k^2+2k}{N} + \frac{1}{N}\mathrm{vec}(V_{ff}^{-1})'E\left[(f_if'_i\otimes f_i)\right]V_{ff}^{-1}E\left[(f_if'_i\otimes f'_i)\right]\mathrm{vec}(V_{ff}^{-1}) + O(N^{-2}).
\end{array}
$$

To determine the order of the remaining components, we analyze the next higher order element

$$\begin{aligned}
&Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
&\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}f_N(\theta,Y) \\
&= \;\; \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1}\sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\right. \\
&\qquad \left.\textstyle\sum_{i_4}\left(\bar{f}_{i_4}\bar{f}_{i_4}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_5} f_{i_5}\right].
\end{aligned}$$

It's highest order will be $O(N^{-2})$ which results from combining two fourth order products with identical indices. The resulting double sum is proportional to N^2 so after dividing by N^4 a $O(N^{-2})$ component remains. Other components of order $O(N^{-2})$ result when combining two third order products and a second order product in deviation of the variance. Despite that this is a triple sum, its order is $O(N^{-2})$ since the deviation of the second order product from the variance is $O(N^{-1})$.

The above shows that we can specify GMM-AR(θ) as

$$\text{GMM-AR}(\theta) = GMM\text{-}AR_0 + \tfrac{1}{N}GMM\text{-}AR_1 + O_p(\tfrac{1}{N^2}),$$

with $GMM\text{-}AR_0 = Nf_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y) \underset{d}{\rightarrow} \chi^2(k)$, so $\mathrm{E}(GMM\text{-}AR_0) = k$, and $GMM\text{-}AR_1 = -N^2 f_N(\theta_0,Y)'\,V_{ff}(\theta_0)^{-1}\left[\hat{V}_{ff}(\theta_0)-V_{ff}(\theta_0)\right]V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y) + N^2 f_N(\theta_0,Y)'\,V_{ff}(\theta_0)^{-1}\left[\hat{V}_{ff}(\theta_0)-V_{ff}(\theta_0)\right]V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y)$, $\mathrm{E}(GMM\text{-}AR_1) = k^2+2k+\mathrm{vec}(V_{ff}^{-1})'E\left[(f_if_i'\otimes f_i)\right]V_{ff}^{-1}E\left[(f_if_i'\otimes f_i')\right]\mathrm{vec}(V_{ff}^{-1})$.

b. KLM-statistic: Using the higher order components of $[\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}$ $\hat{D}_N(\theta,Y)]^{-1}$ and $s_N(\theta,Y)$ from Lemmas 3 and 4, the KLM statistic has the higher order specification:

$$\text{KLM}(\theta_0) = \quad KLM_0 + \tfrac{1}{N}\left(KLM_1 + KLM_2\right) + o_p(N^{-1}),$$

with

$$\begin{aligned} KLM_0 &= \ s_0'D_0 s_0 \\ KLM_1 &= \ \sqrt{N}\left(2s_1'D_0s_0 + s_0'D_1s_0\right) \\ KLM_2 &= \ \sqrt{N}\left(2(s_2'D_0s_0 + s_1'D_1s_0) + s_0'D_2s_0 + s_1'D_0s_1\right). \end{aligned}$$

We next determine the conditional expectations given $D_N(\theta_0,Y)$ of each of these components.

KLM$_0$:

$$\begin{aligned} & KLM_0 \\ &= \ s_0'D_0s_0 \\ &= \ \left(\sqrt{N}f_N(\theta,Y)\right)V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left(\sqrt{N}f_N(\theta,Y)\right). \end{aligned}$$

Since $\sqrt{N}f_N(\theta_0,Y)$ and $\sqrt{N}(D_N(\theta_0,Y)-J_\theta(\theta_0))$ are independent in large samples, $\mathrm{E}(Nf_N(\theta_0,Y)f_N(\theta_0,Y)'|D_N(\theta_0,Y) = D)$ $=\mathrm{E}(Nf_N(\theta_0,Y)f_N(\theta_0,Y)') = V_{ff}(\theta_0)$, with $D_N = D_N(\theta_0,Y)$, so $\mathrm{E}(KLM_0|D_N(\theta_0,Y) = D) = 1$.

KLM$_1$:

$$\begin{aligned} & KLM_1 \\ &= \ \sqrt{N}\left(2s_1'D_0s_0 + s_0'D_1s_0\right) \\ &= \ -2N\left(\sqrt{N}f_N(\theta,Y)\right)V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))}V_{ff}(\theta)^{-\frac{1}{2}} \\ & \quad \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) - \\ & \quad 2N\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) \\ & \quad \left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) + \\ & \quad N\left(\sqrt{N}f_N(\theta,Y)\right)V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))}V_{ff}(\theta)^{-\frac{1}{2}}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\ & \quad V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))}V_{ff}(\theta)^{-\frac{1}{2}}\left(\sqrt{N}f_N(\theta,Y)\right) + \\ & \quad 2N\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]\left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1} \\ & \quad \left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\ & \quad \left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right). \end{aligned}$$

$$
\begin{aligned}
KLM_1 = \quad & -2N\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-\frac{1}{2}} M_{V_{ff}(\theta)^{-\frac{1}{2}}\left[\sqrt{N}D_N(\theta,Y)\right]} V_{ff}(\theta)^{-\frac{1}{2}} \hat{V}_{\theta f}(\theta)' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N\right. \\
& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right) - \\
& N\left(\sqrt{N}f_N(\theta,Y)\right) V_{ff}(\theta)^{-\frac{1}{2}} P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))} V_{ff}(\theta)^{-\frac{1}{2}}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-\frac{1}{2}} \\
& M_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))} V_{ff}(\theta)^{-\frac{1}{2}}\left(\sqrt{N}f_N(\theta,Y)\right) - N\left(\sqrt{N}f_N(\theta,Y)\right) V_{ff}(\theta)^{-\frac{1}{2}} P_{V_{ff}(\theta)^{-\frac{1}{2}}} \\
& V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)
\end{aligned}
$$

First component results from higher order dependence between $\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-\frac{1}{2}} M_{V_{ff}(\theta)^{-}}$ and KLM_0 which are first order independent.

next construct the conditional expectations of each of these four components given $D_N(\theta_0, Y)$:

$$\begin{aligned}
& 1. \quad E(N\left(\sqrt{N}f_N(\theta,Y)\right)V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))}V_{ff}(\theta)^{-\frac{1}{2}} \\
& \quad \left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)|\sqrt{N}D_N(\theta_0,Y)=D) \\
= \; & E\left(\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\left[\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right]\right. \\
& \left.V_{ff}^{-1}f_{i_3}|\sqrt{N}D_N(\theta_0,Y)=D\right) \\
= \; & E\left(\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-1}f_{i_3}\right. \\
& -\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_4}'V_{ff}^{-1}f_{i_3} \\
& \left.-\textstyle\sum_{i_1}\sum_{i_3}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|\sqrt{N}D_N(\theta_0,Y)=D\right) \\
= \; & E\left(\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-1}f_{i_1}+\right. \\
& \tfrac{1}{N}\textstyle\sum_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_1}- \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_2}'V_{ff}^{-1}f_{i_2}- \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_1}'V_{ff}^{-1}f_{i_2}- \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-1}f_{i_1}- \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_1}- \\
& \left.\textstyle\sum_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|\sqrt{N}D_N(\theta_0,Y)=D\right) \\
= \; & (N-1)+(1-\tfrac{1}{N})E(f_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_i)-\tfrac{N-1}{N}k- \\
& 2\tfrac{N-1}{N}-N \\
= \; & (1-\tfrac{1}{N})E(f_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_i)-k(1-\tfrac{1}{N})-3
\end{aligned}$$

2. $$\begin{aligned}
& \mathrm{E}\left(N\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}\right. \\
& \left. D_N(\theta,Y))^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)|D_N(\theta,Y)=D\right) \\
= \; & \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}\bar{f}_{i_2}\bar{d}_{i_2}'V_{ff}^{-1}f_{i_3}(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_4}|\right. \\
& \left. D_N(\theta,Y)=D\right) \\
= \; & \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}f_{i_2}d_{i_2}'V_{ff}^{-1}f_{i_3}(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_4}|\right. \\
& \left. D_N(\theta,Y)=D\right)-\mathrm{E}\left(\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} V_{ff}^{-1}f_{i_2}D'V_{ff}^{-1}f_{i_3}\right. \\
& \left.(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_4}|D_N(\theta,Y)=D\right) \\
= \; & \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}f_{i_2}d_{i_2}'V_{ff}^{-1}f_{i_3}(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_4}|\right. \\
& \left. D_N(\theta,Y)=D\right)-\mathrm{E}\left(\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_3}'V_{ff}^{-1}D\right. \\
& \left.(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_4}|D_N(\theta,Y)=D\right) \\
= \; & \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-1}d_{i_1}(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_2}|D_N(\theta,Y)=D\right)+ \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}'V_{ff}^{-1}f_{i_1}f_{i_2}'V_{ff}^{-1}d_{i_1}(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_1}|D_N(\theta,Y)=D\right)+ \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_2}'V_{ff}^{-1}d_{i_1}(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_2}|D_N(\theta,Y)=D\right)- \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right)- \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right)- \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)+o(1) \\
= \; & 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(D'V_{ff}^{-1}D)^{-\frac{1}{2}}d_{i_1}'V_{ff}^{-1}(V_{ff}+(f_{i_1}f_{i_1}'-V_{ff})V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-\frac{1}{2}}|D_N(\theta,Y)=\sum_{i_1}d_{i_1}=\right. \\
& \left.\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(D'V_{ff}^{-1}D)^{-\frac{1}{2}}d_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-\frac{1}{2}}f_{i_1}'V_{ff}^{-1}f_{i_1}|D_N(\theta,Y)=D\right)-\tfrac{N-1}{N^2}(k+2)\right. \\
= \; & 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(D'V_{ff}^{-1}D)^{-\frac{1}{2}}d_{i_1}'V_{ff}^{-1}(V_{ff}+(f_{i_1}f_{i_1}'-V_{ff}))V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-\frac{1}{2}}|D_N(\theta,Y)=\sum_{i_1}d_{i_1}=\right. \\
& \left.\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(D'V_{ff}^{-1}D)^{-\frac{1}{2}}d_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-\frac{1}{2}}(k+(f_{i_1}'V_{ff}^{-1}f_{i_1}-k))|D_N(\theta,Y)=D\right)-\tfrac{N-1}{N^2}(k+\right. \\
= \; & 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(D'V_{ff}^{-1}D)^{-\frac{1}{2}}d_{i_1}'V_{ff}^{-1}(f_{i_1}f_{i_1}'-V_{ff})V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-\frac{1}{2}}|D_N(\theta,Y)=\sum_{i_1}d_{i_1}=D\right)+ \\
& \mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(D'V_{ff}^{-1}D)^{-1}d_{i_1}'V_{ff}^{-1}D(f_{i_1}'V_{ff}^{-1}f_{i_1}-k)|D_N(\theta,Y)=D\right)=\tfrac{N-1}{N}(2\alpha_1+\alpha_2),
\end{aligned}$$

with $E(V_{ff}^{-\frac{1}{2}}(f_if_i'-V_{ff})V_{ff}^{-1}d_i|D)=E((V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}-I_k)V_{ff}^{-\frac{1}{2}}d_i|D)=V_{ff}^{-\frac{1}{2}}D\alpha_1+V_{ff}^{-\frac{1}{2}}D_\perp\beta_1$ and $E((f_i'V_{ff}^{-1}f_i-k)V_{ff}^{-\frac{1}{2}}d_i|D)=V_{ff}^{-\frac{1}{2}}D\alpha_2+V_{ff}^{-\frac{1}{2}}D_\perp\beta_2$ where the α's and β's may depend on D. Note when $E(f_i f_i'|d_i) = V_{ff}$, α and β are zero.

$$
\begin{aligned}
3.\quad & \mathrm{E}\left(N\left(\sqrt{N}f_N(\theta,Y)\right)V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))}V_{ff}(\theta)^{-\frac{1}{2}}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)\right. \\
& \left. V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}(\sqrt{N}D_N(\theta_0,Y))}V_{ff}(\theta)^{-\frac{1}{2}}\left(\sqrt{N}f_N(\theta,Y)\right)|D_N(\theta,Y)=D\right) \\
= \quad & \mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|\right. \\
& \left. D_N(\theta,Y)=D\right) \\
= \quad & \mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|D_N(\theta,Y)=D\right) \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_4}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|\right. \\
& \left. D_N(\theta,Y)=D\right) - \mathrm{E}\left(\textstyle\sum_{i_1}\sum_{i_3} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|D_N(\theta,Y)=D\right) \\
= \quad & \mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)+ \\
& \mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)- \\
& 3\mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right)- \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)- \\
& \mathrm{E}\left(\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right) \\
= \quad & (N-1)-3(\tfrac{N-1}{N})-N \\
& \mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right) \\
= \quad & -4+\tfrac{3}{N}+\mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)
\end{aligned}
$$

$$
\begin{aligned}
4.\quad & \mathrm{E}\left(N\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]\left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}\right. \\
& \left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1} \\
& \left.\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)|D_N(\theta,Y)=D\right) \\
= \quad & \mathrm{E}\Big(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_2}'V_{ff}^{-1}\bar{f}_{i_3}\bar{d}_{i_3}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_4}| \\
& D_N(\theta,Y)=D) \\
= \quad & \mathrm{E}\Big(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_2}'V_{ff}^{-1}f_{i_3}d_{i_3}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_4}| \\
& D_N(\theta,Y)=D)- \\
& \mathrm{E}\Big(\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}\sum_{i_5} f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_2}'V_{ff}^{-1}f_{i_3}d_{i_5}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_4}| \\
& D_N(\theta,Y)=D) \\
= \quad & \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_1}'V_{ff}^{-1}f_{i_1}d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right)+ \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_2}'V_{ff}^{-1}f_{i_1}d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right)+ \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_2}'V_{ff}^{-1}f_{i_1}d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)- \\
& \mathrm{E}\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}f_{i_2}'V_{ff}^{-1}f_{i_3}D'V_{ff}^{-1}f_{i_4}|D_N(\theta,Y)=D\right) \\
= \quad & \mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}d_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)+ \\
& 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1} d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)-\tfrac{(N-1)}{N}(k+2) \\
= \quad & \mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(k+f_{i_1}'V_{ff}^{-1}f_{i_1}-k)d_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)+ \\
& 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1} d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}(V_{ff}+f_{i_1}f_{i_1}'-V_{ff})V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)-\tfrac{(N-1)}{N} \\
= \quad & \mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}(f_{i_1}'V_{ff}^{-1}f_{i_1}-k)d_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)+ \\
& 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1} d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}(f_{i_1}f_{i_1}'-V_{ff})V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right) \\
= \quad & 2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1} d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}(f_{i_1}f_{i_1}'-V_{ff})V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)+\alpha_2
\end{aligned}
$$

The combined effect of the four components is such that

$$
\begin{aligned}
&E(NKLM_1)\\
=\;& -2\left((1-\tfrac{1}{N})E(f_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_i|D_N(\theta,Y)=D)-k(1-\tfrac{1}{N})-3\right)\\
& -2\tfrac{N-1}{N}(2\alpha_1+\alpha_2)+2-4+\tfrac{3}{N}+\\
& \mathrm{E}\left(\tfrac{1}{N}\textstyle\sum_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right)+\\
& 2\left(2\mathrm{E}\left(\tfrac{N-1}{N^2}\textstyle\sum_{i_1}d_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}(f_{i_1}f_{i_1}'-V_{ff})V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)+\alpha_2\right)\\
=\;& -\mathrm{E}\left(f_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_i|D_N(\theta,Y)=D\right)-2(k+2)-\\
& E(f_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_i|D_N(\theta,Y)=D)-\\
& 4\mathrm{E}\left(d_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-1}D(D'V_{ff}^{-1}D)^{-1}|D_N(\theta,Y)=D\right)
\end{aligned}
$$

We can specify $D'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D}d_i$ as

$$
\begin{aligned}
& D'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}d_i=\\
&= D'V_{ff}^{-1}(d_i'\otimes f_if_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D})\mathrm{vec}(V_{ff}^{-\frac{1}{2}})\\
&= \textstyle\sum_{m=1}^k d_{i,m}D'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D}(V_{ff}^{-\frac{1}{2}})_m\\
&= \textstyle\sum_{m=1}^k d_{i,m}D'V_{ff}^{-\frac{1}{2}}g_ig_i'V_{ff}^{\frac{1}{2}}D_\perp(D_\perp'V_{ff}D_\perp)^{-1}D_\perp'V_{ff}^{\frac{1}{2}}(V_{ff}^{-\frac{1}{2}})_m\\
&= \textstyle\sum_{m=1}^k d_{i,m}D'V_{ff}^{-\frac{1}{2}}g_ig_i'V_{ff}^{\frac{1}{2}}D_\perp(D_\perp'V_{ff}D_\perp)^{-1}D_\perp'e_{m,k}
\end{aligned}
$$

with $(V_{ff}^{-\frac{1}{2}})_m$ the m-th column of $V_{ff}^{-\frac{1}{2}}$ and $d_{i,m}$ the m-th element of d_i, $g_i=V_{ff}^{-\frac{1}{2}}f_i$, $M_{V_{ff}^{-\frac{1}{2}}D}=P_{V_{ff}^{\frac{1}{2}}D_\perp}$, $D_\perp$ the $k\times(k-1)$ dimensional orthogonal complement of D, $D_\perp'D\equiv 0$, $e_{m,k}$ the m-th column of I_k. To take the conditional expectation of the above expression given $D_N(\theta,Y)$, we use that we can specify $E(d_{i,m}g_ig_i'|D_N(\theta,Y)=D)=E((d_{i,m}-D_m)(g_ig_i'-I_k)|D_N(\theta,Y)=D)+D_mI_k$

$$
E(d_{i,m}g_ig_i'|D_N(\theta,Y)=D)=(V_{ff}^{-\frac{1}{2}}D\vdots V_{ff}^{\frac{1}{2}}D_\perp)\begin{pmatrix}A_{11,m} & A_{12,m}\\ A_{12,m}' & A_{22,m}\end{pmatrix}(V_{ff}^{-\frac{1}{2}}D\vdots V_{ff}^{\frac{1}{2}}D_\perp)',
$$

so

$$
\begin{aligned}
& E\left((D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}d_i|D_N(\theta,Y)=D\right)=\\
&= \textstyle\sum_{m=1}^k A_{12,m}D_\perp'e_{m,k}.
\end{aligned}
$$

$\mathbf{KLM}_2$: We obtain the conditional expectation of KLM_2 for each one of its four components:

$$KLM_2 = \sqrt{N}\left(2(s_2' D_0 s_0 + s_1' D_1 s_0) + s_0' D_2 s_0 + s_1' D_0 s_1\right).$$

$2\sqrt{N} s_2' D_0 s_0$:

$$\begin{aligned}
&2\sqrt{N} s_2' D_0 s_0 \\
=\ & N\left(\sqrt{N} f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) \\
& V_{ff}(\theta)^{-1}\left[\sqrt{N} D_N(\theta,Y)\right]\left(N D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1} \\
& \left[\sqrt{N} D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N} f_N(\theta,Y)\right) + N\left(\sqrt{N} f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)' V_{ff}(\theta)^{-1} \\
& \left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\left(\sqrt{N} f_N(\theta,Y)\right)\left(N D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1} \\
& \left[\sqrt{N} D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N} f_N(\theta,Y)\right) + N\left(\sqrt{N} f_N(\theta,Y)\right)' \\
& V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right) V_{ff}(\theta)^{-1}\hat{V}_{\theta f}(\theta)' V_{ff}(\theta)^{-1}\left(\sqrt{N} f_N(\theta,Y)\right) \\
& \left(N D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1}\left[\sqrt{N} D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N} f_N(\theta,Y)\right)
\end{aligned}$$

We next construct the conditional expectation of each of these three components:

$$i.\ E\left(N\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)\right.$$
$$V_{ff}(\theta)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]\left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}$$
$$\left.\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)|D_N(\theta,Y)=D\right)$$
$$\begin{aligned}
= \ & E\left(\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}' V_{ff}^{-1}(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff})V_{ff}^{-1}(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff})V_{ff}^{-\frac{1}{2}}\right.\\
& \left.P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_4}|D_N(\theta,Y)=D\right)\\
= \ & \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]-\\
& \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]-\\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f_{i_5}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]-\\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]+\\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]+\\
& \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_6} f_{i_6}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]-\\
& \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]+\\
& \mathrm{E}\left[\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]+\\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f_{i_5}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right]
\end{aligned}$$

We next construct the conditional expectation of each of these nine components:

$$
\begin{aligned}
1.\quad & \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1,i_2,i_3,i_4} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_3}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_4}|D_N(\theta,Y)=D\right] = \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_3}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + O(N^{-2}) \\
= \quad & \tfrac{(N-1)(N-2)}{N} + 3\mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}\right] + O(N^{-1}) \\
= \quad & \tfrac{(N-1)(N-2)}{N} + 3\mathrm{E}\left[\tfrac{N-1}{N^2}\textstyle\sum_{i_2} f_{i_2}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + O(N^{-2}) \\
2.\quad & \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = \\
& \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}|\right. \\
= \quad & (N-1) + \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-1}f_{i_1}|D_N(\theta,Y)=D\right] + O(N^{-2})
\end{aligned}
$$

$$
\begin{aligned}
3.\;\; & \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}f'_{i_2}V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f'_{i_5}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_3}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + O(N^{-2}) \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_1}f'_{i_3}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\sum_{i_3\neq i_1,i_2} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-1}f_{i_3}f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_3}|D_N(\theta,Y)=D\right] + O(N^{-2}) \\
= \;\; & \tfrac{(N-1)(N-2)}{N^2}(k+2) + O(N^{-1}) \\
4.\;\; & \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_3} f_{i_3}f'_{i_3}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] \\
= \;\; & \tfrac{(N-1)(N-2)}{N^2}(k+2) + O(N^{-1}) \\
5.\;\; & \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f'_{i_5}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] \\
= \;\; & \tfrac{N-1}{N}(k+2) + O(N^{-2}) \\
6.\;\; & \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f'_{i_5}V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_6} f'_{i_6}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = O \\
7.\;\; & \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_3} f_{i_3}f'_{i_3}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = \\
& (N-1) + \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + O(N^{-1}) \\
8.\;\; & \mathrm{E}\left[\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = N \\
9.\;\; & \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f'_{i_1}V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f'_{i_5}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}\sum_{i_4} f_{i_4}|D_N(\theta,Y)=D\right] = \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_1}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_1}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N(\theta,Y)=D\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f'_{i_1}V_{ff}^{-1}f_{i_2}f'_{i_2}V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N(\theta,Y)=D\right] + O(N^{-2}) \\
= \;\; & \tfrac{N-1}{N}(k+2) + O(N^{-1}).
\end{aligned}
$$

Combining, we get:

$$
\begin{aligned}
&1-2-3-4+5+6-7+8+9 &&= \\
&\tfrac{(N-1)(N-2)}{N}+3\mathrm{E}\left[\left(\tfrac{1}{N}-\tfrac{1}{N^2}\right)\textstyle\sum_{i_2} f_{i_2}V_{ff}^{-1}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N=D\right]+ \\
&\mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-1}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_2}|D_N=D\right]-(N-1)- \\
&\mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N=D\right]-2\tfrac{(N-1)(N-2)}{N^2}(k+2)+\tfrac{N-1}{N}(k+2)- \\
&(N-1)-\mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_{i_1}|D_N=D\right]+N+\tfrac{N-1}{N}(k+2) &&= \\
&-1+\mathrm{E}\left[f_i'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_i|D_N=D\right]+ \\
&\mathrm{E}\left[f_i'V_{ff}^{-1}f_if_i'|D_N=D\right]V_{ff}^{-1}\mathrm{E}\left[f_if_i'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D}V_{ff}^{-\frac{1}{2}}f_i|D_N=D\right]+O(N^{-1}).
\end{aligned}
$$

We combine this with the results from the first expression so:

$$
\begin{aligned}
&\mathrm{E}(\text{GMM-AR}(\theta))= \\
&k+\tfrac{k^2+2k}{N}+\tfrac{1}{N}\mathrm{vec}(V_{ff}^{-1})'E\left[(f_if_i'\otimes f_i)\right]V_{ff}^{-1}E\left[(f_if_i'\otimes f_i')\right]\mathrm{vec}(V_{ff}^{-1})+O(N^{-2}).
\end{aligned}
$$

To determine the order of the remaining components, we analyze the next higher order element

$$
\begin{aligned}
&Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1} \\
&\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}f_N(\theta,Y) \\
&\quad= \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\right. \\
&\qquad\left.\textstyle\sum_{i_4}\left(\bar{f}_{i_4}\bar{f}_{i_4}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_5} f_{i_5}\right].
\end{aligned}
$$

$$
\begin{aligned}
&Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}f_N(\theta,Y) \\
&= \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] \\
&= \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
&\quad \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3}\left(\bar{f}_{i_3}\bar{f}_{i_3}'-V_{ff}\right)V_{ff}^{-1}\sum_{i_4} f_{i_4}\right] \\
&= \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
&\quad \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}f_{i_2}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
&\quad \mathrm{E}\left[\tfrac{1}{N^4}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N^5}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_2} f_{i_2}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_6} f_{i_6}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]- \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}f_{i_3}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]+\mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-1}\sum_{i_3} f_{i_3}\sum_{i_5} f_{i_5}'V_{ff}^{-1}\sum_{i_4} f_{i_4}\right]
\end{aligned}
$$

$$ii. \quad E\left(N\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\hat{V}_{\theta f}(\theta)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)\right.$$
$$\left.\left(ND_N(\theta,Y)' V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)|D_N(\theta,Y) = D\right)$$
$$\begin{aligned}&= E\left(\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}\sum_{i_5} f'_{i_1}V_{ff}^{-1}\bar{f}_{i_2}\bar{d}'_{i_2}V_{ff}^{-1}(\bar{f}_{i_3}\bar{f}'_{i_3}-V_{ff})V_{ff}^{-1}f_{i_4}\right.\\ &\quad\left.(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_5}|D_N(\theta,Y)=D\right)\\ &= o(1)\end{aligned}$$

$$iii. \quad E\left(N\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\hat{V}_{\theta f}(\theta)' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)\right.$$
$$\left.\left(ND_N(\theta,Y)' V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)|D_N(\theta,Y) = D\right)$$
$$\begin{aligned}&= E\left(\tfrac{1}{N^3}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}\sum_{i_5} f'_{i_1}V_{ff}^{-1}(\bar{f}_{i_2}\bar{f}'_{i_2}-V_{ff})V_{ff}^{-1}\bar{f}_{i_3}\bar{d}'_{i_3}V_{ff}^{-1}f_{i_4}\right.\\ &\quad\left.(D'V_{ff}^{-1}D)^{-1}D'V_{ff}^{-1}f_{i_5}|D_N(\theta,Y)=D\right)\\ &= o(1)\end{aligned}$$

$$\begin{aligned}KLM_0 = &\; Nf_N(\theta_0,Y)' V_{ff}(\theta_0)^{-\frac{1}{2}}P_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta_0,Y)\\ KLM_1 = &\; -N^2 f_N(\theta_0,Y)' V_{ff}(\theta_0)^{-\frac{1}{2}}P_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta_0)-V_{ff}(\theta_0)\right]\\ &\; V_{ff}(\theta_0)^{-\frac{1}{2}}P_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}f_N(\theta_0,Y)-\\ &\; 2N^2 f_N(\theta_0,Y)' V_{ff}(\theta_0)^{-\frac{1}{2}}P_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta_0)-V_{ff}(\theta_0)\right]\\ &\; V_{ff}(\theta_0)^{-\frac{1}{2}}M_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}f_N(\theta_0,Y)\\ KLM_2 = &\; -2N^2 f_N(\theta_0,Y)' V_{ff}(\theta_0)^{-\frac{1}{2}}M_{V_{ff}(\theta_0)^{-\frac{1}{2}}D_N(\theta_0,Y)}V_{ff}(\theta_0)^{-\frac{1}{2}}\hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}\\ &\; f_N(\theta_0,Y)\left[D_N(\theta_0,Y)' V_{ff}(\theta_0)^{-1}D_N(\theta_0,Y)\right]^{-1}D_N(\theta_0,Y)' V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y)\end{aligned}$$

$$
\begin{aligned}
&\left(N\hat{D}_N(\theta,Y)'V_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right)^{-1} \\
= \;& (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} + (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
&\left[2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+\right. \\
&\left.\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\right] \\
&(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} - (ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
&\left[\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+\right. \\
&\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right) \\
&V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1} \\
&\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+2\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1} \\
&\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+4\left(\sqrt{N}f_N(\theta,Y)\right)' \\
&V_{ff}(\theta_0)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\
&(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1} \\
&\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+ \\
&\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right) \\
&(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right) \\
&\left.V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\right](ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}+o_p(N^{-1})
\end{aligned}
$$

$$
\begin{aligned}
&Ns_N(\theta,Y) = \\
&\left[\sqrt{N}D_N(\theta,Y)\right]'\left[V_{ff}(\theta)^{-1}-V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}+\right. \\
&\left.V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\right]\left(\sqrt{N}f_N(\theta,Y)\right)- \\
&\left(\sqrt{N}f_N(\theta,Y)\right)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+\left(\sqrt{N}f_N(\theta,Y)\right)' \\
&V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+\left(\sqrt{N}f_N(\theta,Y)\right)' \\
&V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)+o_P(N^{-1})
\end{aligned}
$$

$$
\begin{aligned}
&\sqrt{N}KLM_1 \\
&= \left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]\left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1} \\
&\quad \left[2\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)+\right. \\
&\quad \left.\left(\sqrt{N}D_N(\theta,Y)\right)' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)V_{ff}(\theta)^{-1}\left(\sqrt{N}D_N(\theta,Y)\right)\right] \\
&\quad \left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta_0)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)- \\
&\quad 2\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\left[\sqrt{N}D_N(\theta,Y)\right]\left(ND_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right)^{-1} \\
&\quad \left\{\left[\sqrt{N}D_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left(\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right)+\right. \\
&\quad \left.\left(\sqrt{N}f_N(\theta,Y)\right)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta)'\right\} V_{ff}(\theta)^{-1}\left(\sqrt{N}f_N(\theta,Y)\right)
\end{aligned}
$$

$D_0+\frac{1}{\sqrt{N}}D_1+\frac{1}{N}D_2+o_p(N^{-1})$
$s_0+\frac{1}{\sqrt{N}}s_1+\frac{1}{N}s_2$
$\mathrm{KLM}_1 = s_0'D_1s_0+2s_1'D_0s_0$

$$
\begin{aligned}
KLM_0 =& \ s_0'D_0s_0 \\
KLM_1 =& \ 2s_1'D_0s_0+s_0'D_1s_0 \\
KLM_2 =& \ 2(s_2'D_0s_0+s_1'D_1s_0)+s_0'D_2s_0+s_1'D_0s_1
\end{aligned}
$$

We determine the order of the four different components of KLM_0, KLM_1 and KLM_2 using their conditional expectations given $D_N(\theta,Y)$. To construct these conditional expectations, we use Lemma 7.
$\mathbf{KLM}_0$: Corollary 1 shows that $\sqrt{N}f_N(\theta_0,Y)$ and $\sqrt{N}(D_N(\theta_0,Y)-J_\theta(\theta_0))$ are independent in large samples. Hence, in large samples $\mathrm{E}(Nf_N(\theta_0,Y)f_N(\theta_0,Y)'|D_N)=$

$\mathrm{E}(Nf_N(\theta_0,Y)f_N(\theta_0,Y)')=V_{ff}(\theta_0)$, with $D_N=D_N(\theta_0,Y)$, so

$$\begin{aligned}
\mathrm{E}(KLM_0) &= \mathrm{E}\left[Nf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)|D_N\right] \\
&= \mathrm{E}\left[\mathrm{tr}\left\{V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}Nf_N(\theta,Y)f_N(\theta,Y)'\right\}|D_N\right] \\
&= \mathrm{tr}\left\{V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\mathrm{E}(\left(\sqrt{N}f_N(\theta,Y)\right)\left(\sqrt{N}f_N(\theta,Y)\right)'|D_N)\right\} \\
&= \mathrm{tr}\left\{V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}V_{ff}(\theta)\right\}=1.
\end{aligned}$$

KLM$_1$: KLM_1 consists of two components and we construct the expectation of each one separately. Assumption 1 implies that $\mathrm{E}(\hat{V}_{ff}(\theta_0)|D_N)=V_{ff}(\theta_0)+$

$O(N^{-\frac{1}{2}})$, which we use to construct the expectations of both elements of KLM_1 :

$$
\begin{aligned}
(-)\ & \mathrm{E}\left[N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]\right. \\
& \left. V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)|D_N\right] \\
= \ & \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}'V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}}\sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}'-V_{ff}\right)V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}}\sum_{i_3}f_{i_3}|D_N\right] \\
= \ & \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\left(V_{ff}^{-\frac{1}{2}}\bar{f}_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_3}\otimes V_{ff}^{-\frac{1}{2}}\bar{f}_{i_2}\right)'\right. \\
& \left.\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})|D_N\right]-\mathrm{E}\left[\textstyle\sum_{i_1}\sum_{i_3}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\left(V_{ff}^{-\frac{1}{2}}f_{i_3}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)|D_N\right] \\
= \ & \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_3}\otimes V_{ff}^{-\frac{1}{2}}f_{i_2}\right)'-\right. \\
& \left.\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})|D_N\right]\mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\right. \\
& \left.\left(V_{ff}^{-\frac{1}{2}}f_{i_3}\otimes V_{ff}^{-\frac{1}{2}}f_{i_4}\right)'\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})|D_N\right]-\mathrm{E}\left[\textstyle\sum_{i_1}\sum_{i_3}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\right. \\
& \left.\left(V_{ff}^{-\frac{1}{2}}f_{i_3}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)|D_N\right] \\
= \ & \tfrac{1}{N}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_2}\right)'|D_N\right] \\
& \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})+\tfrac{1}{N}\textstyle\sum_{i_1}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)'|D_N\right] \\
& \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})-\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_2}\right)'\right. \\
& \left.|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})-\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\right. \\
& \left.\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)'|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})-\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})' \\
& \mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_2}\otimes V_{ff}^{-\frac{1}{2}}f_{i_2}\right)'|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})-\tfrac{1}{N^2}\textstyle\sum_{i_1}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})' \\
& \mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)'|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})- \\
& \mathrm{E}\left[\textstyle\sum_{i_1}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)|D_N\right] \\
= \ & \tfrac{1}{N}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}\right)|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})+ \\
& \tfrac{1}{N}\textstyle\sum_{i_1}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}\right)|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})- \\
& \tfrac{2}{N^2}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}f_{i_1}'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}f_{i_2}f_{i_2}'V_{ff}^{-\frac{1}{2}}\right)|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})- \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{\substack{i_2\\ i_2\neq i_1}}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}}f_{i_1}\otimes V_{ff}^{-\frac{1}{2}}f_{i_1}\right)|D_N\right]\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}} f_{i_2} \otimes V_{ff}^{-\frac{1}{2}} f_{i_2}\right) |D_N\right] - \tfrac{1}{N^2}\textstyle\sum_{i_1} \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})' \\
& \mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} \otimes V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}}\right) |D_N\right] \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N}) - \\
& \mathrm{E}\left[\textstyle\sum_{i_1} \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})' \left(V_{ff}^{-\frac{1}{2}} f_{i_1} \otimes V_{ff}^{-\frac{1}{2}} f_{i_1}\right) |D_N\right] \\
= \; & (N-1)\left[1-\tfrac{2}{N}\right] \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})'\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N}) - N\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})'\mathrm{vec}\left(V_{ff}\right) - \\
& \tfrac{N-1}{N}\left[\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})'\mathrm{vec}\left(V_{ff}\right)\right]^2 + \\
& \left(1-\tfrac{1}{N}\right) \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}} \otimes V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}}\right)\right] \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N}) \\
= \; & (N-1)\left[1-\tfrac{2}{N}\right] - N - \tfrac{N-1}{N} + \left(1-\tfrac{1}{N}\right)\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})' \\
& \mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}} \otimes V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}}\right)\right] \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N}) \\
= \; & -4 + \tfrac{3}{N} + \left(1-\tfrac{1}{N}\right) \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})'\mathrm{E}\left[\left(V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}} \otimes V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}}\right)\right] \mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N}),
\end{aligned}
$$

since $\mathrm{vec}(A)'\mathrm{vec}(B)$ $=\mathrm{tr}(AB)$ and $\left(V_{ff}^{-\frac{1}{2}} f_{i_1} \otimes V_{ff}^{-\frac{1}{2}} f_{i_2}\right)'\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N}) = f_{i_2}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} = \left(V_{ff}^{-\frac{1}{2}} f_{i_2} \otimes V_{ff}^{-\frac{1}{2}} f_{i_1}\right)'\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}} D_N})$.

$$
\begin{aligned}
& (-2\times)\ \mathrm{E}\Big[N^2 f_N(\theta,Y)' V_{ff}(\theta)^{-\frac{1}{2}} P_{V_{ff}(\theta)^{-\frac{1}{2}} D_N(\theta,Y)} V_{ff}(\theta)^{-\frac{1}{2}} \left[\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right] \\
& \quad V_{ff}(\theta)^{-\frac{1}{2}} M_{V_{ff}(\theta)^{-\frac{1}{2}} D_N(\theta,Y)} V_{ff}(\theta)^{-\frac{1}{2}} f_N(\theta,Y) |D_N\Big] \\
= \; & \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2}\left(\bar{f}_{i_2}\bar{f}_{i_2}' - V_{ff}\right) V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_3} f_{i_3} |D_N\right] \\
= \; & \mathrm{E}\Big\{\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2} f_{i_2} f_{i_2}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_3} f_{i_3} - \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2} f_{i_2} \sum_{i_4} f_{i_4}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_3} f_{i_3} - \\
& \textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_3} f_{i_3} |D_N\Big\} \\
= \; & \mathrm{E}\Big\{\tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2\neq i_1} f_{i_2} f_{i_2}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} + \\
& \tfrac{1}{N}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} - \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_2}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_2} - \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_2} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_2} - \\
& \tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_2} f_{i_2}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}} D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} -
\end{aligned}
$$

$$
\begin{aligned}
&\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} f_{i_1}|D_N\Big\} \\
= \;& \mathrm{E}\Big\{ -\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} \mathrm{tr}(V_{ff}^{-\frac{1}{2}} P_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_1}')\mathrm{tr}(V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} f_{i_2} f_{i_2}') + \\
&\tfrac{N-1}{N}\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}} \otimes V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}})\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})|D_N\Big\} \\
= \;& \tfrac{N-1}{N}\Big[\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}} \otimes V_{ff}^{-\frac{1}{2}} f_i f_i' V_{ff}^{-\frac{1}{2}})\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N}) - (k-1)\Big]
\end{aligned}
$$

KLM$_2$: KLM_2 consists of one component whose conditional expectation given $D_N(\theta,Y)$ we construct using that $\mathrm{E}(\hat{V}_{ff}(\theta_0)|D_N) = V_{ff}(\theta_0) + O(N^{-\frac{1}{2}})$:

$$
\begin{aligned}
&(-2\times)\ \mathrm{E}\Big[N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}} M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)} V_{ff}(\theta)^{-\frac{1}{2}}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1} \\
&f_N(\theta,Y)\left[D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right]^{-1} D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)|D_N(\theta,Y)\Big] \\
= \;& \mathrm{E}\Big\{\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2} \bar{d}_{i_2}\bar{f}_{i_2}' V_{ff}^{-1}\sum_{i_3} f_{i_3}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} \\
&D_N'V_{ff}^{-1}\textstyle\sum_{i_4} f_{i_4}|D_N\Big\} \\
= \;& \mathrm{E}\Big\{\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2} d_{i_2} f_{i_2}' V_{ff}^{-1}\sum_{i_3} f_{i_3}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} \\
&D_N'V_{ff}^{-1}\textstyle\sum_{i_4} f_{i_4} - \tfrac{1}{N^2}\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} \sum_{i_2} d_{i_2} \\
&\textstyle\sum_{i_5} f_{i_5}' V_{ff}^{-1}\sum_{i_3} f_{i_3}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1}\sum_{i_4} f_{i_4}|D_N\Big\} \\
= \;& \mathrm{E}\Big\{\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_2}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1} f_{i_2} + \\
&\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1} f_{i_2} \\
&\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_2}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1} f_{i_1} + \\
&\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1} f_{i_1}|D_N\Big\} \\
= \;& \mathrm{E}\Big\{\tfrac{N-1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1} + \\
&\tfrac{N-1}{N^2}\textstyle\sum_{i_1} D_N' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} + \\
&\tfrac{N-1}{N^2}\textstyle\sum_{i_1} D_N' V_{ff}^{-1} f_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} + \\
&\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1} f_{i_1}|D_N\Big\} \\
= \;& \mathrm{E}\Big\{\tfrac{2(N-1)}{N^2}\textstyle\sum_{i_1} d_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} f_{i_1} f_{i_1}' V_{ff}^{-1} D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1} + \\
&\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-\frac{1}{2}} M_{V_{ff}^{-\frac{1}{2}}D_N} V_{ff}^{-\frac{1}{2}} d_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1} D_N'V_{ff}^{-1} f_{i_1}|D_N\Big\}
\end{aligned}
$$

$$= \tfrac{2(N-1)}{N}\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}d_i|D_N)V_{ff}^{-\frac{1}{2}}D_N$$
$$\left[D_N'V_{ff}^{-1}D_N\right]^{-1} + \tfrac{1}{N}\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}d_i|D_N)$$
$$\mathrm{vec}(V_{ff}^{-\frac{1}{2}}D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1}).$$

Under Assumption 2, $\mathrm{E}(\frac{1}{N}\sum_{i=1}^N\mathrm{vec}(f_i(\theta_0)f_i(\theta_0)')d_i(\theta_0)'|D_N(\theta_0,Y))=\mathrm{vec}(V_{ff}(\theta_0))$ $D_N(\theta_0,Y)'$ so

$$\begin{aligned}&\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}d_i|D_N)\\ &= (V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}})\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N f_if_i'\otimes d_i|D_N)(V_{ff}^{-\frac{1}{2}}\otimes 1)\\ &= (V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}})(V_{ff}\otimes D_N)(V_{ff}^{-\frac{1}{2}}\otimes 1)\end{aligned}$$

which implies that only the $\frac{1}{N}$ term of KLM_2 remains.

Given $D_N(\theta_0,Y)$, the conditional expectation of the different higher order elements of KLM(θ_0) read

$$\mathrm{E}\left[KLM_0|D_N(\theta_0,Y)\right]=1$$
$$-\mathrm{E}\left[KLM_1|D_N(\theta_0,Y)\right]=\tfrac{N-1}{N}\mathrm{vec}(V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}})'\mathrm{E}(f_if_i'\otimes f_if_i')$$
$$\mathrm{vec}(V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}})-4+\tfrac{3}{N}+$$
$$2\tfrac{N-1}{N}\left[\mathrm{vec}(V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}})'\mathrm{E}(f_if_i'\otimes f_if_i')\mathrm{vec}(V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}})-(k-1)\right]$$
$$=\tfrac{N-1}{N}\mathrm{vec}(V_{ff}^{-\frac{1}{2}}P_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}})'\mathrm{E}(f_if_i'\otimes f_if_i')\mathrm{vec}(V_{ff}^{-1}+V_{ff}^{-\frac{1}{2}}M_{V_{ff}^{-\frac{1}{2}}D_N}V_{ff}^{-\frac{1}{2}})-$$
$$\tfrac{2(N-1)}{N}(k-1)-4+\tfrac{3}{N}$$
$$\mathrm{E}\left[KLM_2|D_N(\theta_0,Y)\right]=-\tfrac{4(N-1)}{N}\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}d_i|D_N)$$
$$V_{ff}^{-\frac{1}{2}}D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1}-\tfrac{2}{N}\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes$$
$$V_{ff}^{-\frac{1}{2}}d_i|D_N)\mathrm{vec}(V_{ff}^{-\frac{1}{2}}D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1}),$$

and the conditional expectation of KLM_2 under Assumption 2 is of order $\frac{1}{N}$.

c. GMM-MLR statistic. To expand the expression of the GMM-MLR statistic, we use a Taylor approximation of GMM-AR(θ) around $GMM\text{-}AR_0$ and KLM(θ_0) around KLM_0 :

$$\mathrm{GMM\text{-}MLR}(\theta_0)=$$
$$\tfrac{1}{2}\left[GMM\text{-}AR_0-\mathrm{r}(\theta_0)+\sqrt{(GMM\text{-}AR_0+\mathrm{r}(\theta_0))^2-4\left[GMM\text{-}AR_0-KLM_0\right]\mathrm{r}(\theta_0))}\right]+$$
$$\tfrac{1}{2N}\left[1+\tfrac{GMM\text{-}AR_0-\mathrm{r}(\theta_0)}{\sqrt{(GMM\text{-}AR_0+\mathrm{r}(\theta_0))^2-4[GMM\text{-}AR(\theta_0)-KLM(\theta_0)]\mathrm{r}(\theta_0))}}\right]GMM\text{-}AR_1+$$
$$\tfrac{1}{N}\tfrac{\mathrm{r}(\theta_0)}{\sqrt{(GMM\text{-}AR_0+\mathrm{r}(\theta_0))^2-4[GMM\text{-}AR(\theta_0)-KLM(\theta_0)]\mathrm{r}(\theta_0))}}(KLM_1+\tfrac{1}{\sqrt{N}}KLM_2)+O_p(\tfrac{1}{N^2}).$$

d. LM statistic. We just show that the higher order components of $\mathrm{LM}(\theta)$ depend on $D_N(\theta,Y)$. We therefore only construct a higher order decomposition for which we uses just the $(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}$-element from the higher order specification of $\left[q_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}q_N(\theta,Y)\right]^{-1}$ in Lemma 6.

$$
\begin{aligned}
\mathrm{LM}(\theta) &= Nf_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}q_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y) \\
&\approx N\left[f_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)+f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)-\right. \\
&\quad \left. f_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y)\right] \\
&\quad (D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}\left[f_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)+\right. \\
&\quad f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)-f_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\
&\quad \left. V_{ff}(\theta)^{-1}D_N(\theta,Y)'\right] \\
&= KLM_0+2Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) \\
&\quad (D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)+ \\
&\quad Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)\,(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
&\quad f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1}f_N(\theta,Y)-2Nf_N(\theta,Y)'V_{ff}(\theta)^{-1} \\
&\quad \left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y)\,(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1} \\
&\quad \left[D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)+f_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)'V_{ff}(\theta)^{-1}f_N(\theta,Y)\right]+ \\
&\quad Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}D_N(\theta,Y) \\
&\quad (D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\
&\quad V_{ff}(\theta)^{-1}f_N(\theta,Y)
\end{aligned}
$$

We use the analog of Lemma 7 to obtain the conditional expectation of $\mathrm{LM}(\theta)$ given $D_N(\theta,Y)$. We therefore first construct the conditional expectation of the first five elements of the approximation of $\mathrm{LM}(\theta)$ given D_N.

$$1.\ \mathrm{E}\left[KLM_0|D_N\right]=1.$$

$$
\begin{aligned}
&2.\ (2\times)\ \mathrm{E}\left[Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}V_{qf}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y)\right. \\
&\quad \left.(D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y))^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)|D_N\right] \\
&= \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}f_{i_1}'V_{ff}^{-1}V_{qf}V_{ff}^{-1}\sum_{i_2}f_{i_2}\left[D_N'V_{ff}^{-1}D_N\right]^{-1}D_N'V_{ff}^{-1}\sum_{i_3}f_{i_3}|D_N\right] \\
&= \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}f_{i_1}'V_{ff}^{-1}V_{qf}V_{ff}^{-1}f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1}D_N'V_{ff}^{-1}f_{i_1}|D_N\right] \\
&= \tfrac{1}{N}\mathrm{tr}\left\{\mathrm{E}\left[f_{i_1}f_{i_1}'V_{ff}^{-1}V_{qf}V_{ff}^{-1}f_{i_1}\left[D_N'V_{ff}^{-1}D_N\right]^{-1}D_N'V_{ff}^{-1}|D_N\right]\right\},
\end{aligned}
$$

where $V_{qf}=V_{qf}(\theta)$.

3. $\mathrm{E}\Big[N f_N(\theta,Y)' V_{ff}(\theta)^{-1} V_{qf}(\theta) V_{ff}(\theta)^{-1} f_N(\theta,Y) \left(D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1}$
$f_N(\theta,Y)' V_{ff}(\theta)^{-1} V_{qf}(\theta)' V_{ff}(\theta)^{-1} f_N(\theta,Y) | D_N]$

$$
\begin{aligned}
&= \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} \sum_{i_2} f_{i_2} \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \sum_{i_3} f_{i_3}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} \sum_{i_4} f_{i_4} | D_N\right] \\
&= \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \sum_{i_3 \neq i_1} f_{i_3}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_{i_3} | D_N\right] + \\
&\quad \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} \sum_{i_2 \neq i_1} f_{i_2} \left[D_N' V_{ff}^{-1} D_N\right]^{-1} f_{i_2}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_{i_1} | D_N\right] + \\
&\quad \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} \sum_{i_2 \neq i_1} f_{i_2} \left[D_N' V_{ff}^{-1} D_N\right]^{-1} f_{i_1}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_{i_2} | D_N\right] + \\
&\quad \mathrm{E}\left[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} \left[D_N' V_{ff}^{-1} D_N\right]^{-1} f_{i_1}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_{i_1} | D_N\right] \\
&= \tfrac{(N-1)}{N^2}\Big\{\left[\mathrm{tr}(V_{ff}^{-1} V_{qf})\right]^2 + \mathrm{tr}(V_{ff}^{-1} V_{qf} V_{ff}^{-1} V_{qf}') + \mathrm{tr}(V_{ff}^{-1} V_{qf} V_{ff}^{-1} V_{qf}) + \\
&\quad \tfrac{1}{N-1}\mathrm{E}\left[(f_i' V_{ff} V_{qf} V_{ff} f_i)^2\right]\Big\} \left[D_N' V_{ff}^{-1} D_N\right]^{-1}
\end{aligned}
$$

4. $(2\times)\ \mathrm{E}\Big[N f_N(\theta,Y)' V_{ff}(\theta)^{-1} \left[\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right] V_{ff}(\theta)^{-1} D_N(\theta,Y)$
$\left(D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1} D_N(\theta,Y)' V_{ff}(\theta)^{-1} f_N(\theta,Y) | D_N\Big]$

$$
\begin{aligned}
&= \mathrm{E}\Big[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} \sum_{i_2} \left(f_{i_2} f_{i_2}' - V_{ff}\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \\
&\quad D_N' V_{ff}^{-1} \textstyle\sum_{i_3} f_{i_3} | D_N\Big] \\
&= \mathrm{E}\Big[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} \sum_{i_2} \left(f_{i_2} f_{i_2}' - V_{ff}\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \\
&\quad D_N' V_{ff}^{-1} f_{i_1} | D_N\Big] + \mathrm{E}\Big[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} \left(f_{i_1} f_{i_1}' - V_{ff}\right) V_{ff}^{-1} D_N \\
&\quad \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1} f_{i_1} | D_N\Big] \\
&= \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1} f_{i_1}' V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1} f_{i_1} | D_N\right] - \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1} f_{i_1} | D_N\right] \\
&= \tfrac{1}{N}\mathrm{tr}\left[\mathrm{E}\left(f_i f_i' V_{ff}^{-1} f_i f_i' | D_N\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1}\right] - \\
&\quad \tfrac{1}{N}\mathrm{tr}\left[\mathrm{E}\left(f_i f_i' | D_N\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1}\right] \\
&= \tfrac{1}{N}\mathrm{tr}\left[\mathrm{E}\left(f_i f_i' V_{ff}^{-1} f_i f_i'\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1}\right] - \tfrac{1}{N}.
\end{aligned}
$$

5. $(2\times)\ \mathrm{E}\Big[N f_N(\theta,Y)' V_{ff}(\theta)^{-1} \left[\hat{V}_{ff}(\theta) - V_{ff}(\theta)\right] V_{ff}(\theta)^{-1} D_N(\theta,Y)$
$\left(D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1} f_N(\theta,Y)' V_{ff}(\theta)^{-1} V_{qf}(\theta)' V_{ff}(\theta)^{-1} f_N(\theta,Y) | D_N\Big]$

$$
\begin{aligned}
&= \mathrm{E}\Big[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} \sum_{i_2} \left(f_{i_2} f_{i_2}' - V_{ff}\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \\
&\quad \textstyle\sum_{i_3} f_{i_3}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} \sum_{i_4} f_{i_4} | d_i\Big] \\
&= \mathrm{E}\Big[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} \sum_{i_2 \neq i_1} \left(f_{i_2} f_{i_2}' - V_{ff}\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \\
&\quad f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} | D_N\Big] + \mathrm{E}\Big[\tfrac{1}{N^3}\textstyle\sum_{i_1} f_{i_1}' V_{ff}^{-1} \left(f_{i_1} f_{i_1}' - V_{ff}\right) V_{ff}^{-1} \\
&\quad D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \textstyle\sum_{i_2 \neq i_1} f_{i_2}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_2} | D_N\Big] +
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{E}\left[\tfrac{1}{N^3}\sum_{i_1} f_{i_1}' V_{ff}^{-1} \sum_{i_2\neq i_1}\left(f_{i_2}f_{i_2}' - V_{ff}\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1}\right. \\
& \left. f_{i_2}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} | D_N\right] + \mathrm{E}\left[\tfrac{1}{N^3}\sum_{i_1} f_{i_1}' V_{ff}^{-1} \sum_{i_2\neq i_1}\left(f_{i_2}f_{i_2}' - V_{ff}\right) V_{ff}^{-1}\right. \\
& \left. D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_2} | D_N\right] + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\sum_{i_1} f_{i_1}' V_{ff}^{-1}\left(f_{i_1}f_{i_1}' - V_{ff}\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1}\right. \\
& \left. f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} | D_N\right] \\
= \; & \tfrac{N-1}{N^2}\left\{\mathrm{E}\left(f_i' V_{ff}^{-1} f_i f_i' | D_N\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \mathrm{E}(f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i | D_N) + \right. \\
& \mathrm{tr}\left[V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \mathrm{E}\left(f_{i_2}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_2} f_{i_2}' | D_N\right)\right] + \\
& \left. \mathrm{tr}\left[V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} \mathrm{E}\left(f_{i_2}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_{i_2} f_{i_2}' | D_N\right)\right]\right\} + \\
& \mathrm{E}\left[\tfrac{1}{N^3}\sum_{i_1} f_{i_1}' V_{ff}^{-1} f_{i_1} f_{i_1}' V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} | D_N\right] - \\
& \mathrm{E}\left[\tfrac{1}{N^3}\sum_{i_1} f_{i_1}' V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} f_{i_1}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_1} | D_N\right] \\
= \; & \tfrac{N-1}{N^2}\left\{\left[\mathrm{tr}(V_{ff}^{-1} V_{qf}) \mathrm{E}\left(f_i' V_{ff}^{-1} f_i f_i'\right) + \mathrm{E}\left(f_{i_2}' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_{i_2} f_{i_2}'\right) + \right.\right. \\
& \left.\left. \mathrm{E}\left(f_{i_2}' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_{i_2} f_{i_2}'\right)\right] V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1}\right\} + \\
& \tfrac{1}{N^2}\mathrm{E}\left[f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i f_i' V_{ff}^{-1} f_i f_i' | D_N\right] V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} + \\
& \tfrac{1}{N^2}\mathrm{E}\left[f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i f_i' | D_N\right] V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1}.
\end{aligned}
$$

The five elements of the approximation of $\mathrm{LM}(\theta)$ are then such that we can specify it as

$$
\begin{aligned}
\mathrm{LM}(\theta) = \; & KLM_0 + \tfrac{1}{N} LM_1 + \tfrac{1}{N^2} LM_{D_2}\left(D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1} + \\
& \mathrm{tr}\left(\tfrac{1}{N} LM_{D_1}\left(D_N(\theta,Y)' V_{ff}(\theta)^{-1} D_N(\theta,Y)\right)^{-1} D_N(\theta,Y)' V_{ff}(\theta)^{-1}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathrm{E}(KLM_0|D_N) = \; & 1 \\
\mathrm{E}(LM_1|D_N) = \; & 2 - 2\mathrm{tr}\left[\mathrm{E}\left(f_i f_i' V_{ff}^{-1} f_i f_i'\right) V_{ff}^{-1} D_N \left[D_N' V_{ff}^{-1} D_N\right]^{-1} D_N' V_{ff}^{-1}\right]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(LM_{D_1}|D_N) = \; & 2\mathrm{E}\left[f_i f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i\right] - 2\tfrac{N-1}{N}\left\{\mathrm{tr}(V_{ff}^{-1} V_{qf}) \mathrm{E}\left(f_i' V_{ff}^{-1} f_i f_i'\right) + \right. \\
& \left. \mathrm{E}\left(f_i' V_{ff}^{-1} V_{qf}' V_{ff}^{-1} f_i f_i'\right) + \mathrm{E}\left(f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i f_i'\right)\right\} - \\
& \tfrac{2}{N}\mathrm{E}\left[f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i f_i' V_{ff}^{-1} f_i f_i'\right] - \tfrac{2}{N}\mathrm{E}\left[f_i' V_{ff}^{-1} V_{qf} V_{ff}^{-1} f_i f_i'\right] \\
\mathrm{E}(LM_{D_2}|D_N) = \; & (N-1)\left\{\left[\mathrm{tr}(V_{ff}^{-1} V_{qf})\right]^2 + \mathrm{tr}(V_{ff}^{-1} V_{qf} V_{ff}^{-1} V_{qf}') + \mathrm{tr}(V_{ff}^{-1} V_{qf} V_{ff}^{-1} V_{qf}) + \right. \\
& \left. \mathrm{tr}(V_{ff}^{-1} V_{qf} V_{ff}^{-1} V_{qf})\right\} + \mathrm{E}\left[(f_i' V_{ff} V_{qf} V_{ff} f_i)^2\right].
\end{aligned}
$$

Proof of Theorem 2.

Using Lemmas 2 and 7, we obtain the higher order specification of $\mathrm{r}(\theta_0)$:

$$\begin{aligned}
\mathrm{r}(\theta_0) = & \; N\hat{D}_N(\theta_0,Y)\left[\hat{V}_{qq}(\theta_0)-\hat{V}_{qf}(\theta_0)\hat{V}_{ff}(\theta_0)^{-1}\hat{V}_{fq}(\theta_0)\right]^{-1}\hat{D}_N(\theta_0,Y) \\
= & \; N\left[D_N(\theta_0,Y)-\hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y)\right]'[V_{\theta\theta.f}(\theta_0)^{-1}-V_{\theta\theta.f}(\theta_0)^{-1} \\
& \left[\hat{V}_{dd}(\theta_0)-V_{\theta\theta.f}(\theta_0)\right]V_{\theta\theta.f}(\theta_0)^{-1}\Big] \\
& \left[D_N(\theta_0,Y)-\hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y)\right]+O_p(\tfrac{1}{N}) \\
= & \; ND_N(\theta_0,Y)'V_{dd}(\theta_0)^{-1}D_N(\theta_0,Y)-N\Big\{2f_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)' \\
& V_{\theta\theta.f}(\theta_0)^{-1}D_N(\theta_0,Y)+D_N(\theta_0,Y)'V_{\theta\theta.f}(\theta_0)^{-1}\left[\hat{V}_{dd}(\theta_0)-V_{\theta\theta.f}(\theta_0)\right]+ \\
& V_{\theta\theta.f}(\theta_0)^{-1}D_N(\theta_0,Y)\}+Nf_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1} \\
& \hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y)+Nf_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1} \\
& \left[\hat{V}_{dd}(\theta_0)-V_{\theta\theta.f}(\theta_0)\right]V_{\theta\theta.f}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y) \\
= & \; \mathrm{r}_0+\tfrac{1}{N}(\mathrm{r}_1+\mathrm{r}_2)+\tfrac{1}{N^2}\mathrm{r}_3+o_p(\tfrac{1}{N^2}),
\end{aligned}$$

with

$$\begin{aligned}
\mathrm{r}_0 = & \; ND_N(\theta_0,Y)'V_{\theta\theta.f}(\theta_0)^{-1}D_N(\theta_0,Y) \\
\mathrm{r}_1 = & \; 2N^2f_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1}D_N(\theta_0,Y) \\
\mathrm{r}_2 = & \; N^2D_N(\theta_0,Y)'V_{\theta\theta.f}(\theta_0)^{-1}\left[\hat{V}_{dd}(\theta_0)-V_{dd}(\theta_0)\right]V_{\theta\theta.f}(\theta_0)^{-1}D_N(\theta_0,Y) \\
\mathrm{r}_3 = & \; N^3f_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1}\left[\hat{V}_{dd}(\theta_0)-V_{\theta\theta.f}(\theta_0)\right] \\
& V_{\theta\theta.f}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y).
\end{aligned}$$

The order of the different elements results from their expected values:
r_0 :

$$\begin{aligned}
\mathrm{E}\left[\mathrm{r}_0\right] = & \; \mathrm{E}\left[ND_N(\theta_0,Y)'V_{\theta\theta.f}(\theta_0)^{-1}D_N(\theta_0,Y)\right] \\
= & \; \mathrm{E}\Bigg[N\left(J(\theta_0)+\tfrac{1}{N}\textstyle\sum_{i=1}^N(d_i(\theta_0)-J(\theta_0))\right)'V_{\theta\theta.f}(\theta_0)^{-1} \\
& \left(J(\theta_0)+\tfrac{1}{N}\textstyle\sum_{i=1}^N(d_i(\theta_0)-J(\theta_0))\right)\Bigg] \\
= & \; NJ(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1}J(\theta_0)+ \\
& \tfrac{1}{N}\mathrm{tr}\left[V_{\theta\theta.f}(\theta_0)^{-1}\mathrm{E}\left(\textstyle\sum_{i=1}^N\sum_{j=1}^N(d_i(\theta_0)-J(\theta_0))(d_j(\theta_0)-J(\theta_0))'\right)\right] \\
= & \; NJ(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1}J(\theta_0)+k.
\end{aligned}$$

so $\mathrm{O}(\mathrm{r}_0)=N$ when $J(\theta_0)\neq 0$ and 1 otherwise.

r_1 :

$$
\begin{aligned}
\mathrm{E}\left[\mathrm{r}_1\right]= \;& 2\mathrm{E}\left[N^2 f_N(\theta_0,Y)' V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)' V_{\theta\theta.f}(\theta_0)^{-1} D_N(\theta_0,Y)\right] \\
= \;& 2\mathrm{E}\left[\tfrac{1}{N}\sum_{i_1}\sum_{i_1}\sum_{i_3} f_{i_1}' V_{ff}^{-1}\bar{d}_{i_2}\bar{f}_{i_2}' V_{\theta\theta.f}^{-1} d_{i_3}\right] \\
= \;& 2\mathrm{E}\left[\tfrac{1}{N}\sum_{i_1}\sum_{i_1}\sum_{i_3} f_{i_1}' V_{ff}^{-1} d_{i_2} f_{i_2}' V_{\theta\theta.f}^{-1} d_{i_3}\right] - \\
& 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}' V_{ff}^{-1} d_{i_2} f_{i_3}' V_{\theta\theta.f}^{-1} d_{i_4}\right] \\
= \;& 2\mathrm{E}\left[\tfrac{1}{N}\sum_{i_1} f_{i_1}' V_{ff}^{-1}(J+d_{i_1}-J) f_{i_1}' V_{\theta\theta.f}^{-1}(d_{i_1}-J)\right] + \\
& 2\mathrm{E}\left[\tfrac{1}{N}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}' V_{ff}^{-1}(J+d_{i_2}-J) f_{i_2}' V_{\theta\theta.f}^{-1}(d_{i_1}-J)\right] + \\
& 2\mathrm{E}\left[\sum_{i_1}\sum_{i_2} f_{i_1}' V_{ff}^{-1}(J+d_{i_2}-J) f_{i_2}' V_{\theta\theta.f}^{-1} J\right] - \\
& 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}' V_{ff}^{-1} d_{i_2} f_{i_1}' V_{\theta\theta.f}^{-1} d_{i_3}\right] \\
= \;& 2\mathrm{E}\left[f_i' V_{ff}^{-1} J f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] + 2\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] + \\
& 2J' V_{\theta\theta.f}^{-1} J + 2N\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1} J\right] - \\
& 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}' V_{ff}^{-1}(J+d_{i_2}-J) f_{i_1}' V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] \\
= \;& 2NJ' V_{\theta\theta.f}^{-1} J + 2\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] + 2\mathrm{E}\left[f_i' V_{ff}^{-1} J f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] + \\
& 2N\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1} J\right] - \\
& 2NJ' V_{\theta\theta.f}^{-1} J - 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}' V_{ff}^{-1}(d_{i_2}-J) f_{i_1}' V_{\theta\theta.f}^{-1}(d_{i_3}-J)\right] - \\
& 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}' V_{ff}^{-1} J f_{i_1}' V_{\theta\theta.f}^{-1}(d_{i_3}-J)\right] - \\
& 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3} f_{i_1}' V_{ff}^{-1}(d_{i_2}-J) f_{i_1}' V_{\theta\theta.f}^{-1} J\right] \\
= \;& 2N\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1} J\right] + 2\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] + \\
& 2\mathrm{E}\left[f_i' V_{ff}^{-1} J f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] - 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}' V_{ff}^{-1}(d_{i_2}-J)\right. \\
& \left. f_{i_1}' V_{\theta\theta.f}^{-1}(d_{i_2}-J)\right] - 2\mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1} f_{i_1}' V_{ff}^{-1}(d_{i_1}-J) f_{i_1}' V_{\theta\theta.f}^{-1}(d_{i_1}-J)\right] - \\
& 2\mathrm{E}\left[\tfrac{1}{N}\sum_{i_1} f_{i_1}' V_{ff}^{-1} J f_{i_1}' V_{\theta\theta.f}^{-1}(d_{i_1}-J)\right] - 2\mathrm{E}\left[\tfrac{1}{N}\sum_{i_1} f_{i_1}' V_{ff}^{-1}(d_{i_1}-J) f_{i_1}' V_{\theta\theta.f}^{-1} J\right] \\
= \;& 2(N-1)\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i'\right] V_{\theta\theta.f}^{-1} J + 2\tfrac{N-1}{N}\mathrm{E}\left[f_i' V_{ff}^{-1}(d_i-J) f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] + \\
& 2\mathrm{E}\left[f_i' V_{ff}^{-1} J f_i' V_{\theta\theta.f}^{-1}(d_i-J)\right] - 2\tfrac{N-1}{N} k
\end{aligned}
$$

with $J = J(\theta_0) = \mathrm{E}(d_i(\theta_0))$.

r_2 :

$$
\begin{aligned}
\mathrm{E}\left[\mathrm{r}_2\right]= \;& \mathrm{E}\left[N^2 D_N(\theta_0,Y)' V_{\theta\theta.f}(\theta_0)^{-1}\left[\hat{V}_{dd}(\theta_0)-V_{dd}(\theta_0)\right] V_{\theta\theta.f}(\theta_0)^{-1} D_N(\theta_0,Y)\right] \\
= \;& \mathrm{E}\left[\tfrac{1}{N}\sum_{i_1}\sum_{i_2}\sum_{i_3}(J+d_{i_1}-J)' V_{\theta\theta.f}^{-1}\left[\bar{d}_{i_2}\bar{d}_{i_2}' - V_{\theta\theta.f}\right] V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] \\
= \;& \mathrm{E}\left[\tfrac{1}{N}\sum_{i_1}\sum_{i_2}\sum_{i_3}(J+d_{i_1}-J)' V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)' - \right. \\
& \quad \left. V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] - \mathrm{E}\left[\tfrac{1}{N^2}\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}(J+d_{i_1}-J)' V_{\theta\theta.f}^{-1}\right. \\
& \quad \left.(J+d_{i_2}-J)(J+d_{i_3}-J) V_{\theta\theta.f}^{-1}(J+d_{i_4}-J)\right] - \\
& \mathrm{E}\left[\tfrac{1}{N}\sum_{i_1}\sum_{i_2}\sum_{i_3}(J+d_{i_1}-J)' V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] \\
= \;& N^2 J' V_{\theta\theta.f}^{-1} J (J' V_{\theta\theta.f}^{-1} J + 1) + N\{J' V_{\theta\theta.f}^{-1} J(4+2k) + \tfrac{N-1}{N} k + \\
& 2\tfrac{N+1}{N}\mathrm{E}\left[J' V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)' V_{\theta\theta.f}^{-1}(d_i-J)\right]\} + \\
& \mathrm{E}\left[(d_i-J)' V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)' V_{\theta\theta.f}^{-1}(d_i-J)\right] - \\
& N^2 (J' V_{\theta\theta.f}^{-1} J)^2 - NJ' V_{\theta\theta.f}^{-1} J(4+2k) - 2\tfrac{N-1}{N}(k^2+2k) - \\
& 4\mathrm{E}\left[J' V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)' V_{\theta\theta.f}^{-1}(d_i-J)\right] - \\
& \tfrac{1}{N}\mathrm{E}\left[(d_i-J)' V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)' V_{\theta\theta.f}^{-1}(d_i-J)\right] - N^2 J' V_{\theta\theta.f}^{-1} J - Nk
\end{aligned}
$$

$$
\begin{aligned}
&= 2(N-1)\mathrm{E}\left[J'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]+ \\
&\quad \tfrac{(N-1)}{N}\mathrm{E}\left[(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]- \\
&\quad k(k+3)+\tfrac{1}{N}(k^2+2k)
\end{aligned}
$$

where we used that

$$
\begin{aligned}
&\mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}(J+d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] \\
&= \mathrm{E}\left[\textstyle\sum_{i_2}\sum_{i_3}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] \\
&= N\mathrm{E}\left[\textstyle\sum_{i_2}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}J\right]+ \\
&\quad \mathrm{E}\left[\textstyle\sum_{i_2}\sum_{i_3}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(d_{i_3}-J)\right]+ \\
&\quad \mathrm{E}\left[\textstyle\sum_{i_1}\sum_{i_2}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}J\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(d_{i_3}-J)\right] \\
&= (NJ'V_{\theta\theta.f}^{-1}J)^2+N^2J'V_{\theta\theta.f}^{-1}J+NJ'V_{\theta\theta.f}^{-1}J+ \\
&\quad \mathrm{E}\left[\textstyle\sum_{i_2}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(d_{i_2}-J)\right]+ \\
&\quad \mathrm{E}\left[\textstyle\sum_{i_2}(d_{i_2}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}J\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_2}-J)'V_{\theta\theta.f}^{-1}(d_{i_1}-J)\right] \\
&\quad \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_1}-J)(J+d_{i_1}-J)'V_{\theta\theta.f}^{-1}(d_{i_1}-J)\right] \\
&= (NJ'V_{\theta\theta.f}^{-1}J)^2+N^2J'V_{\theta\theta.f}^{-1}J+3NJ'V_{\theta\theta.f}^{-1}J+2NkJ'V_{\theta\theta.f}^{-1}J+ \\
&\quad 2N\mathrm{E}\left[J'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]+(N-1)J'V_{\theta\theta.f}^{-1}J+(N-1)k+ \\
&\quad \mathrm{E}\left[(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]+J'V_{\theta\theta.f}^{-1}J+ \\
&\quad \mathrm{E}\left[(d_i-J)'V_{\theta\theta.f}^{-1}J(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]+\mathrm{E}\left[(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)J'V_{\theta\theta.f}^{-1}(d_i-J)\right] \\
&= N^2J'V_{\theta\theta.f}^{-1}J(J'V_{\theta\theta.f}^{-1}J+1)+N\{J'V_{\theta\theta.f}^{-1}J(4+2k)+\tfrac{N-1}{N}k+ \\
&\quad \tfrac{N+1}{N}\mathrm{E}\left[J'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]\}+ \\
&\quad \mathrm{E}\left[(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}(J+d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_4}-J)\right] \\
&= \mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_4}-J)\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_4}-J)\right] \\
&= \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}J\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}J'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}(d_{i_4}-J)\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}J\right]+ \\
&\quad \mathrm{E}\left[\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}(d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_2}-J)(J+d_{i_3}-J)'V_{\theta\theta.f}^{-1}(d_{i_4}-J)\right] \\
&= N^2(J'V_{\theta\theta.f}^{-1}J)^2+NJ'V_{\theta\theta.f}^{-1}J+2NkJ'V_{\theta\theta.f}^{-1}J+2NJ'V_{\theta\theta.f}^{-1}J+ \\
&\quad 2\mathrm{E}\left[J'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]+NJ'V_{\theta\theta.f}^{-1}J+\tfrac{N-1}{N}(k^2+2k)+ \\
&\quad 2\mathrm{E}\left[J'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right] \\
&= N^2(J'V_{\theta\theta.f}^{-1}J)^2+NJ'V_{\theta\theta.f}^{-1}J(4+2k)+\tfrac{N-1}{N}(k^2+2k)+ \\
&\quad 4\mathrm{E}\left[J'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right]+ \\
&\quad \tfrac{1}{N}\mathrm{E}\left[(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)(d_i-J)'V_{\theta\theta.f}^{-1}(d_i-J)\right], \\
&\mathrm{E}\left[\tfrac{1}{N}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}(J+d_{i_1}-J)'V_{\theta\theta.f}^{-1}(J+d_{i_3}-J)\right] \\
&= N^2J'V_{\theta\theta.f}^{-1}J+Nk
\end{aligned}
$$

r_3 :

$$\begin{aligned}
\mathrm{E}\left[\mathrm{r}_3\right] = {} & \mathrm{E}\Big\{N^3 f_N(\theta_0,Y)'V_{ff}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)'V_{\theta\theta.f}(\theta_0)^{-1}\left[\hat{V}_{dd}(\theta_0)-V_{\theta\theta.f}(\theta_0)\right] \\
& V_{\theta\theta.f}(\theta_0)^{-1}\hat{V}_{\theta f}(\theta_0)V_{ff}(\theta_0)^{-1}f_N(\theta_0,Y)\Big\} \\
= {} & \mathrm{E}\left\{\tfrac{1}{N^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4}\sum_{i_5} f_{i_1}'V_{ff}^{-1}\bar{d}_{i_2}\bar{f}_{i_2}V_{\theta\theta.f}^{-1}\left[\bar{d}_{i_3}\bar{d}_{i_3}-V_{\theta\theta.f}\right]\right. \\
& \left. V_{\theta\theta.f}^{-1}\bar{f}_{i_4}\bar{d}_{i_4}'V_{ff}^{-1}f_{i_5}\right\},
\end{aligned}$$

which can be shown to be of order $O(1)$.

Proof of Theorem 4.

a. GMM-AR*-statistic: The higher order components of the bootstrapped GMM-AR statistic from Bootstrap Algorithm 1, GMM-AR$^*(\theta)$, read:

$$\begin{aligned}
&\text{GMM-AR}^*(\theta) = Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)- \\
&\quad Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right]\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)+O_p(\tfrac{1}{B^2}).
\end{aligned}$$

The expectation of the first component with respect to the resampling distribution of f_i^*, E^*, results from

$$\begin{aligned}
&\mathrm{E}^*\left[Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)\right] = \mathrm{E}^*\left[\tfrac{1}{B}\mathrm{tr}\left(\hat{V}_{ff}^{-1}\textstyle\sum_{i_1}f_{i_1}^*\sum_{i_2}f_{i_2}^{*\prime}\right)\right] \\
&= \mathrm{E}^*\left[\tfrac{1}{B}\mathrm{tr}\left(\hat{V}_{ff}^{-1}\textstyle\sum_{i_1}f_{i_1}^*f_{i_1}^{*\prime}\right)\right]+\mathrm{E}^*\left[\tfrac{1}{B}\mathrm{tr}\left(\hat{V}_{ff}^{-1}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}^*f_{i_2}^{*\prime}\right)\right] \\
&= \tfrac{1}{B}\mathrm{tr}\left(\hat{V}_{ff}^{-1}\textstyle\sum_{i_1}\mathrm{E}^*(f_{i_1}^*f_{i_1}^{*\prime})\right)+\tfrac{1}{B}\mathrm{tr}\left(\hat{V}_{ff}^{-1}\textstyle\sum_{i_1}\mathrm{E}^*(f_{i_1}^*)\sum_{i_2\neq i_1}\mathrm{E}^*(f_{i_2}^*)'\right) \\
&= \tfrac{1}{B}\mathrm{tr}\left(\hat{V}_{ff}^{-1}\textstyle\sum_{i_1=1}^B\tfrac{1}{N}\sum_{i=1}^N\bar{f}_i\bar{f}_i'\right)=k,
\end{aligned}$$

since $\mathrm{E}^*(f_{i_1}^*)=\frac{1}{N}\sum_i^N\bar{f}_i=0$ and $\mathrm{E}^*(f_{i_1}^*f_{i_1}^{*\prime})=\frac{1}{N}\sum_{j_1=1}^N\bar{f}_i\bar{f}_i'=\hat{V}_{ff}$. The expectation of $Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right]\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)$ results from

$$\begin{aligned}
&\mathrm{E}^*\left[Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right]\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)\right] \\
&= \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}\sum_{i_2}\left(f_{i_2}^*f_{i_2}^{*\prime}-\hat{V}_{ff}\right)\hat{V}_{ff}^{-1}\sum_{i_3}f_{i_3}^*\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^3}\textstyle\sum_{i_1}f_{i_1}^{*\prime}V_{ff}^{-1}\sum_{i_2}f_{i_2}^*\sum_{i_4}f_{i_4}^{*\prime}V_{ff}^{-1}\sum_{i_3}f_{i_3}^*\right] \\
&= \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}\sum_{i_2\neq i_1}\left(f_{i_2}^*f_{i_2}^{*\prime}-\hat{V}_{ff}\right)\hat{V}_{ff}^{-1}f_{i_1}^*\right]+ \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}\left(f_{i_1}^*f_{i_1}^{*\prime}-\hat{V}_{ff}\right)\hat{V}_{ff}^{-1}f_{i_1}^*\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_1}^*f_{i_2}^{*\prime}\hat{V}_{ff}^{-1}f_{i_2}^*\right]-\mathrm{E}^*\left[\tfrac{1}{B^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_2}^*f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_2}^*\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^3}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_2}^*f_{i_2}^{*\prime}\hat{V}_{ff}^{-1}f_{i_1}^*\right]-\mathrm{E}^*\left[\tfrac{1}{B^3}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_1}^*f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_1}^*\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\left(f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_1}^*\right)^2\right] - \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}f_{i_1}^*\right] - \tfrac{B-1}{B^2}\left[k^2+2k\right] - \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\left(f_i^{*\prime}\hat{V}_{ff}^{-1}f_i^*\right)^2\right] \\
&= \tfrac{B-1}{B^2}\mathrm{E}^*\left[\left(f_i^{*\prime}\hat{V}_{ff}^{-1}f_i^*\right)^2\right] - \tfrac{k}{B} - \tfrac{B-1}{B^2}\left[k^2+2k\right] \\
&= \tfrac{B-1}{B^2}\tfrac{1}{N}\textstyle\sum_{i=1}^N\left(\bar{f}_i'\hat{V}_{ff}^{-1}\bar{f}_i\right)^2 - \tfrac{k}{B} - \tfrac{B-1}{B^2}\left[k^2+2k\right].
\end{aligned}
$$

Hence,

$$
\text{GMM-AR}^*(\theta) = GMM\text{-}AR_0^* + \tfrac{1}{B}GMM\text{-}AR_1^* + O_p(\tfrac{1}{B^2}),
$$

where $GMM\text{-}AR_0^* = Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)$, $\mathrm{E}(GMM\text{-}AR_0^*) = k$, and $GMM\text{-}AR_1^* = -Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\,[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)]\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)$, $\mathrm{E}^*(GMM\text{-}AR_1) = -\tfrac{B-1}{B}\tfrac{1}{N}\sum_{i=1}^N\left(\bar{f}_i'\hat{V}_{ff}^{-1}\bar{f}_i\right)^2 + k + \tfrac{B-1}{B}\left[k^2+2k\right]$.

b. KLM*-statistic: The bootstrapped KLM statistic, KLM$^*(\theta)$, from Bootstrap Algorithm 1 has the following higher order components:

$$
\text{KLM}^*(\theta) = KLM_0^* + \tfrac{1}{B}KLM_1^* + O_p(\tfrac{1}{B^2}),
$$

where

$$
\begin{aligned}
&KLM_0^* = Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}f_B^*(\theta,Y) \\
&KLM_1^* = -B^2f_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right] \\
&\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}f_B^*(\theta,Y) - 2B^2f_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}} \\
&P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right]\hat{V}_{ff}(\theta)^{-\frac{1}{2}}M_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}f_B^*(\theta,Y).
\end{aligned}
$$

We determine the order of the these components using their conditional expected values given $\hat{D}_N(\theta_0,Y)$ with respect to the resampling distribution.

$$
\begin{aligned}
&\mathrm{E}^*\left[KLM_0^*\right] = \mathrm{E}^*\left[Bf_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}f_B^*(\theta,Y)|\hat{D}_N\right] \\
&\quad = \mathrm{E}^*\left[\tfrac{1}{B}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_2}f_{i_2}^*|\hat{D}_N\right] \\
&\quad = \mathrm{tr}\left\{\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\tfrac{1}{B}\textstyle\sum_{i_1}\mathrm{E}^*\left(f_{i_1}^*f_{i_1}^{*\prime}|\hat{D}_N\right)\right\} + \\
&\quad\quad \mathrm{tr}\left\{\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\tfrac{1}{B}\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\mathrm{E}^*(f_{i_1}^*f_{i_2}^{*\prime}|\hat{D}_N)\right\} = 1.
\end{aligned}
$$

The conditional expected values of the two components of KLM_1^* given $\hat{D}_N(\theta_0,Y)$

are:

1. $\mathrm{E}^*\left[B^2 f_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right]\right.$
$\left.\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)f_B^*(\theta,Y)|\hat{D}_N\right]$

$$
\begin{aligned}
&= \mathrm{E}^*\left[\tfrac{1}{B}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}\left[\sum_{i_2}\left(f_{i_2}^* f_{i_2}^{*\prime}-\hat{V}_{ff}\right)\right]\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}\sum_{i_3} f_{i_3}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}\left[\sum_{i_2} f_{i_2}^*\sum_{i_3} f_{i_3}^{*\prime}\right]\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}\sum_{i_4} f_{i_4}^*|\hat{D}_N\right] \\
&= \mathrm{E}^*\left[\tfrac{1}{B}\mathrm{tr}\left\{\textstyle\sum_{i_1}\sum_{i_2\neq i_1}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}\left[\left(f_{i_2}^* f_{i_2}^{*\prime}-\hat{V}_{ff}\right)\right]\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^* f_{i_1}^{*\prime}\right\}|\hat{D}_N\right]+ \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B}\mathrm{tr}\left\{\textstyle\sum_{i_1}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}\left[\left(f_{i_1}^* f_{i_1}^{*\prime}-\hat{V}_{ff}\right)\right]\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^* f_{i_1}^{*\prime}\right\}|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2}\sum_{i_3}\sum_{i_4} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_2}^* f_{i_3}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_4}^*|\hat{D}_N\right] \\
&= \mathrm{E}^*\left[\tfrac{1}{B}\left\{\textstyle\sum_{i_1}\left(f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^*\right)^2\right\}|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B}\mathrm{tr}\left\{\textstyle\sum_{i_1}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^* f_{i_1}^{*\prime}\right\}|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_2}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_2}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^* f_{i_2}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_2}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_2}^* f_{i_2}^{*\prime}\hat{V}_{ff}^{-1}A\hat{V}_{ff}^{-1}f_{i_1}^*|\hat{D}_N\right] \\
&= \tfrac{B-1}{B}\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N\left(\bar{f}_i'\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}^{-\frac{1}{2}}\bar{f}_i\right)^2-4\right]-\tfrac{1}{B},
\end{aligned}
$$

with $A=\sum_{i_5}\hat{d}_{i_5}\left[\sum_{i_6}\hat{d}_{i_6}'\hat{V}_{ff}^{-1}\sum_{i_7}\hat{d}_{i_7}\right]^{-1}\sum_{i_8}\hat{d}_{i_8}'$, $\hat{d}_i=q_i(\theta_0)-\hat{V}_{\theta f}(\theta_0)\hat{V}_{ff}(\theta_0)^{-1}f_i(\theta_0)$.

2. $(2\times)$ $\mathrm{E}^*\left[B^2 f_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\left[V_{ff}^*(\theta)-\hat{V}_{ff}(\theta)\right]\right.$
$\left.\hat{V}_{ff}(\theta)^{-\frac{1}{2}}M_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}f_B^*(\theta,Y)|\hat{D}_N\right]$

$$
\begin{aligned}
&= \mathrm{E}^*\left[\tfrac{1}{B}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_2}\left[f_{i_2}^* f_{i_2}^{*\prime}-\hat{V}_{ff}\right]\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_3} f_{i_3}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_2} f_{i_2}^*\sum_{i_4} f_{i_4}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_3} f_{i_3}^*|\hat{D}_N\right] \\
&= \mathrm{E}^*\left[\tfrac{1}{B}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\left[f_{i_2}^* f_{i_2}^{*\prime}-\hat{V}_{ff}\right]\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_1}^*|\hat{D}_N\right]+ \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\left[f_{i_1}^* f_{i_1}^{*\prime}-\hat{V}_{ff}\right]\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_1}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_2}^* f_{i_2}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_1}^*|\hat{D}_N\right]- \\
&\quad \mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_2}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_2}^*|\hat{D}_N\right]-
\end{aligned}
$$

$$\begin{aligned}
&\mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_1}^* f_{i_2}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_2}^*|\hat{D}_N\right]- \\
&\mathrm{E}^*\left[\tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}f_{i_1}^*|\hat{D}_N\right] \\
=\;&\tfrac{B-1}{B}\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N\left(\bar{f}_i'\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\bar{f}_i\bar{f}_i'\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\bar{f}_i\right)-(k-1)\right]
\end{aligned}$$

so $\mathrm{E}^*\left[KLM_0^*\right]=1$ and

$$\begin{aligned}
&\mathrm{E}^*\left[KLM_1^*|\hat{D}_N\right]=-\tfrac{B-1}{B}\left\{\tfrac{1}{N}\textstyle\sum_{i=1}^N\left(\bar{f}_i'\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}^{-\frac{1}{2}}\bar{f}_i\right)^2+\right. \\
&\left.\tfrac{2}{N}\textstyle\sum_{i=1}^N\left(\bar{f}_i'\hat{V}_{ff}^{-\frac{1}{2}}P_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}^{-\frac{1}{2}}\bar{f}_i\bar{f}_i'\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}^{-\frac{1}{2}}\bar{f}_i\right)-2(k+1)\right\}+\tfrac{1}{B}.
\end{aligned}$$

KLM-statistic: The KLM statistic that results from Bootstrap Algorithm 2 has the following higher order components:

$$\mathrm{KLM}^{**}(\theta_0)=\;KLM_0^*+\tfrac{1}{B}\left(KLM_1^*+KLM_2^{**}\right)+O_p(\tfrac{1}{B^2}),$$

with

$$\begin{aligned}
&KLM_2^{**}=-2B^2f_B^*(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}V_{\theta f}^*(\theta_0) \\
&\hat{V}_{ff}(\theta_0)^{-1}f_B^*(\theta_0,Y)\left[\hat{D}_N(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}\hat{D}_N(\theta_0,Y)\right]^{-1}\hat{D}_N(\theta_0,Y)'\hat{V}_{ff}(\theta_0)^{-1}f_B^*(\theta_0,Y)
\end{aligned}$$

with $V_{\theta f}^*(\theta_0)=\frac{1}{B}\sum_{i=1}^B\bar{d}_i^*(\theta_0)_i\bar{f}_i^*(\theta_0)'$, $\bar{d}_i^*(\theta_0)=d_i^*(\theta_0)-D_B^*(\theta_0,Y)$, $d_i^*(\theta_0)=q_i^*(\theta_0)-\hat{V}_{qf}(\theta_0)^{-1}\hat{V}_{ff}(\theta_0)^{-1}f_i^*(\theta_0)$, $D_B^*(\theta_0,Y)=\frac{1}{B}\sum_{i=1}^B d_i^*(\theta_0)=q_B^*(\theta_0,Y)-\hat{V}_{qf}(\theta_0)$ $\hat{V}_{ff}(\theta_0)^{-1}f_B^*(\theta_0,Y)$. We construct the conditional expectation of KLM_2^{**} given $\hat{D}_N(\theta_0,Y)$.

$\mathbf{KLM}_2^{**}$:

$$\begin{aligned}
&-\tfrac{1}{2}\mathrm{E}^*[KLM_2^{**}|\hat{D}_N(\theta,Y)]=\mathrm{E}^*\left[B^2f_B^*(\theta,Y)'\hat{V}_{ff}(\theta)^{-\frac{1}{2}}M_{\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\hat{D}_N(\theta,Y)}\hat{V}_{ff}(\theta)^{-\frac{1}{2}}\right. \\
&V_{\theta f}^*(\theta)\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)\left[\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}\hat{D}_N(\theta,Y)\right]^{-1} \\
&\left.\hat{D}_N(\theta,Y)'\hat{V}_{ff}(\theta)^{-1}f_B^*(\theta,Y)|\hat{D}_N(\theta,Y)\right] \\
&=\;\mathrm{E}^*\left\{\tfrac{1}{B^2}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}V_{ff}^{-\frac{1}{2}}\sum_{i_2}\bar{d}_{i_2}^*\bar{f}_{i_2}^{*\prime}\hat{V}_{ff}^{-1}\sum_{i_3}f_{i_3}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\right. \\
&\quad\left.\hat{D}_N'\hat{V}_{ff}^{-1}\textstyle\sum_{i_4}f_{i_4}^*|\hat{D}_N\right\} \\
&=\;\mathrm{E}^*\left\{\tfrac{1}{B^2}\textstyle\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_2}d_{i_2}^*f_{i_2}^{*\prime}\hat{V}_{ff}^{-1}\sum_{i_3}f_{i_3}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\right. \\
&\quad\hat{D}_N'\hat{V}_{ff}^{-1}\textstyle\sum_{i_4}f_{i_4}^*-\tfrac{1}{B^2}\sum_{i_1}f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}}M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}}\sum_{i_2}d_{i_2}^* \\
&\quad\left.\textstyle\sum_{i_5}f_{i_5}^{*\prime}\hat{V}_{ff}^{-1}\sum_{i_3}f_{i_3}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1}\sum_{i_4}f_{i_4}^*|\hat{D}_N\right\}
\end{aligned}$$

$$
\begin{aligned}
= \; & \mathrm{E}^*\Big\{ \tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_2}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1} f_{i_2}^* + \\
& \tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_1}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1} f_{i_2}^* + \\
& \tfrac{1}{B^2}\textstyle\sum_{i_1}\sum_{i_2\neq i_1} f_{i_2}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_2}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1} f_{i_1}^* + \\
& \tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_1}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1} f_{i_1}^* |\hat{D}_N \Big\} \\
= \; & \mathrm{E}^*\Big\{ \tfrac{B-1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \\
& \tfrac{B-1}{B^2}\textstyle\sum_{i_1} \hat{D}_N'\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_1}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \\
& \tfrac{B-1}{B^2}\textstyle\sum_{i_1} \hat{D}_N'\hat{V}_{ff}^{-1} f_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} \hat{d}_{i_1}\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \\
& \tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_1}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1} f_{i_1}^* |\hat{D}_N \Big\} \\
= \; & \mathrm{E}^*\Big\{ \tfrac{2(B-1)}{B^2}\textstyle\sum_{i_1} d_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} f_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \\
& \tfrac{1}{B^2}\textstyle\sum_{i_1} f_{i_1}^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N}\hat{V}_{ff}^{-\frac{1}{2}} d_{i_1}^* f_{i_1}^{*\prime}\hat{V}_{ff}^{-1} f_{i_1}^*\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}\hat{D}_N'\hat{V}_{ff}^{-1} f_{i_1}^* |\hat{D}_N \Big\} \\
= \; & \tfrac{2(B-1)}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})'\mathrm{E}^*(\tfrac{1}{B}\textstyle\sum_{i=1}^B \hat{V}_{ff}^{-\frac{1}{2}} f_i^* f_i^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}} d_i^* |\hat{D}_N)\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N \\
& \left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \tfrac{1}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})'\mathrm{E}^*(\tfrac{1}{B}\textstyle\sum_{i=1}^B \hat{V}_{ff}^{-\frac{1}{2}} f_i^* f_i^{*\prime}\hat{V}_{ff}^{-1} f_i^* f_i^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} \otimes \\
& \hat{V}_{ff}^{-\frac{1}{2}} d_i^* |\hat{D}_N)\mathrm{vec}(\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}) \\
= \; & \tfrac{2(B-1)}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})'(\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}})\tfrac{1}{B}\textstyle\sum_{i=1}^B \mathrm{E}^*(f_i^* f_i^{*\prime} \otimes d_i^* |\hat{D}_N)\hat{V}_{ff}^{-1}\hat{D}_N \\
& \left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \tfrac{1}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})'(\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}}) \\
& \tfrac{1}{B}\textstyle\sum_{i=1}^B \mathrm{E}^*(f_i^* f_i^{*\prime}\hat{V}_{ff}^{-1} f_i^* f_i^{*\prime} \otimes d_i^* |\hat{D}_N)\hat{V}_{ff}^{-1}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} \\
= \; & \tfrac{2(B-1)}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})'(\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}})\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N (\bar{f}_i\bar{f}_i' \otimes \hat{d}_i)\right]\hat{V}_{ff}^{-1}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} + \\
& \tfrac{1}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})'(\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}})\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N (\bar{f}_i\bar{f}_i'\hat{V}_{ff}^{-1}\bar{f}_i\bar{f}_i' \otimes \hat{d}_i)\right]\hat{V}_{ff}^{-1}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}.
\end{aligned}
$$

Given $\hat{D}_N(\theta_0, Y)$, the conditional expectation of KLM_2^* reads

$$
\begin{aligned}
& \mathrm{E}\left[KLM_2^* |\hat{D}_N(\theta_0, Y)\right] = -\tfrac{4(B-1)}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})' \\
& \mathrm{E}^*(\tfrac{1}{B}\textstyle\sum_{i=1}^B \hat{V}_{ff}^{-\frac{1}{2}} f_i^* f_i^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}} \bar{d}_i^* |\hat{D}_N)\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1} - \tfrac{2}{B}\mathrm{vec}(M_{\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N})' \\
& \mathrm{E}^*(\tfrac{1}{B}\textstyle\sum_{i=1}^B \hat{V}_{ff}^{-\frac{1}{2}} f_i^* f_i^{*\prime}\hat{V}_{ff}^{-1} f_i^* f_i^{*\prime}\hat{V}_{ff}^{-\frac{1}{2}} \otimes \hat{V}_{ff}^{-\frac{1}{2}} \bar{d}_i^* |\hat{D}_N)\mathrm{vec}(\hat{V}_{ff}^{-\frac{1}{2}}\hat{D}_N\left[\hat{D}_N'\hat{V}_{ff}^{-1}\hat{D}_N\right]^{-1}).
\end{aligned}
$$

and the expectation of KLM_2^* under Assumption 2 is such that the conditional expectation of KLM_2^* is of order $\frac{1}{B}$.

Proof of Theorem 5.

A.1. For the Edgeworth approximation of the distribution of the GMM-AR statistic, we construct the characteristic function of GMM-AR(θ_0) for which we use the joint large sample density of its' two components, $f_N(\theta_0, Y)$ and $\hat{V}_{ff}(\theta_0)$:

$$\begin{aligned}
&\mathrm{cf}_{\text{GMM-AR}(\theta)}(t) = \iint \exp\left(it\text{GMM-AR}(\theta)\right) p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta))}(u,v)dudv \\
&= \iint \exp\left(it\left[GMM\text{-}AR_0 + \tfrac{1}{N}GMM\text{-}AR_1 + \tfrac{1}{N^2}GMM\text{-}AR_2\right]\right) \\
&\qquad p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta))}(u,v)dudv + o(N^{-2}) \\
&= \iint \exp\left(itGMM\text{-}AR_0\right)\exp\left(\tfrac{it}{N}GMM\text{-}AR_1 + \tfrac{it}{N^2}GMM\text{-}AR_2\right) \\
&\qquad p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta))}(u,v)dudv + o(N^{-2}) \\
&= \iint \left[1 + \tfrac{it}{N}GMM\text{-}AR_1\right]\exp\left(itGMM\text{-}AR_0\right) \\
&\qquad p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta))}(u,v)dudv + O(N^{-2}) \\
&= \int \left[1 + \tfrac{it}{N}\int GMM\text{-}AR_1 p_{\hat{V}_{ff}(\theta)|\sqrt{N}f_N(\theta_0,Y)}(v|u)dv\right]\exp\left(itGMM\text{-}AR_0\right) \\
&\qquad p_{\sqrt{N}f_N(\theta,Y)}(u)du + O(N^{-2})
\end{aligned}$$

where we used that

$$p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta))}(u,v) = p_{\hat{V}_{ff}(\theta)|\sqrt{N}f_N(\theta,Y)}(v|u)p_{\sqrt{N}f_N(\theta,Y)}(u)$$

is the joint density of $f_N(\theta,Y)$ and $\hat{V}_{ff}(\theta)$ for a sample of size N and $p_{\hat{V}_{ff}(\theta)|\sqrt{N}f_N(\theta,Y)}(v|u)$, $p_{\sqrt{N}f_N(\theta,Y)}(v|u)$ are the conditional density of $\hat{V}_{ff}(\theta)$ given $\sqrt{N}f_N(\theta,Y)$ and the marginal density of $\sqrt{N}f_N(\theta,Y)$. The second last equation results from a Taylor approximation of $\exp\left(\frac{it}{N}GMM\text{-}AR_1 + \frac{it}{N^2}GMM\text{-}AR_2\right)$. The order of the error term results from the Mean Value Theorem and the convergence rates of $\left(\frac{it}{N}GMM\text{-}AR_1\right)^2$ and $\frac{it}{N^2}GMM\text{-}AR_2$.

We use the limiting distribution of $\sqrt{N}f_N(\theta,Y)$, which is $N(0,V_{ff}(\theta_0))$, to construct the expression of the characteristic function. To account for $\exp(itGMM\text{-}AR_0)$, we conduct a transformation of random variables from $\sqrt{N}f_N(\theta,Y)$ to $\sqrt{N(1-2it)}f_N(\theta,Y)$. The Jacobian of this transformation equals $(1-2it)^{-\frac{1}{2}k}$:

$$\begin{aligned}
&\mathrm{cf}_{\text{GMM-AR}(\theta)}(t) \\
&= \int \left[1 + \tfrac{it}{N}\int GMM\text{-}AR_1 p_{\hat{V}_{ff}(\theta)|f_N(\theta_0,Y)}(v|u)dv\right]\exp\left(itGMM\text{-}AR_0\right) \\
&\qquad p_{f_N(\theta,Y)}(u)du + O(N^{-2}) \\
&= (1-2it)^{-\frac{1}{2}k}\int \left[1 + \tfrac{it}{N}\int GMM\text{-}AR_1 p_{\hat{V}_{ff}(\theta)|\sqrt{N}f_N(\theta,Y)}(v|\tfrac{1}{\sqrt{1-2it}}u)dv\right] \\
&\qquad p_{\sqrt{N(1-2it)}f_N(\theta,Y)}(u)du + O(N^{-2}),
\end{aligned}$$

where the density function of $\sqrt{N(1-2it)}f_N(\theta,Y)$ is a normal one with mean zero and covariance matrix $V_{ff}(\theta_0)$. If we now specify

$$\begin{aligned} GMM\text{-}AR_1 = & -\tfrac{N}{1-2it}\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\ & V_{ff}(\theta)^{-1}\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right], \end{aligned}$$

we can use the elements of the expression for $\mathrm{E}(GMM\text{-}AR_1)$ in Theorem 1 to obtain the characteristic function since the transformation from $\sqrt{N}f_N(\theta,Y)$ to $\sqrt{N(1-2it)}f_N(\theta,Y)$ is such that

$$\begin{aligned} & -\int\left\{\int\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right]' V_{ff}(\theta)^{-1}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-1}\right. \\ & \quad \left.\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right] p_{\hat{V}_{ff}(\theta)|\sqrt{N}f_N(\theta,Y)}(v|\tfrac{1}{\sqrt{1-2it}}u)dv\right\} p_{\sqrt{N(1-2it)}f_N(\theta,Y)}(u)du \\ & = \mathrm{E}(GMM\text{-}AR_1). \end{aligned}$$

Hence,

$$\begin{aligned} & \mathrm{cf}_{\text{GMM-AR}(\theta)}(t) \\ & = (1-2it)^{-\frac{1}{2}k}\int\left[1+\tfrac{it}{N}\int GMM\text{-}AR_1 p_{\hat{V}_{ff}(\theta)|\sqrt{N(1-2it)}f_N(\theta,Y)}(v|u)dv\right] \\ & \quad p_{\sqrt{N(1-2it)}f_N(\theta,Y)}(u)du+O(N^{-2}) \\ & = (1-2it)^{-\frac{1}{2}k}\left[1+\tfrac{it}{N(1-2it)}\mathrm{E}(GMM\text{-}AR_1)\right]+O(N^{-2}), \end{aligned}$$

which is such that $(-i\times\frac{\partial cf_{\text{GMM-AR}(\theta)}(0)}{\partial t}) = k+\frac{1}{N}\mathrm{E}(GMM\text{-}AR_1)$. This characteristic function results from the density function:

$$p_{\text{GMM-AR}(\theta)}(u) = p_{\chi^2(k)}(u) - \tfrac{\mathrm{E}(GMM\text{-}AR_1)}{N}\left(\tfrac{\partial}{\partial u}p_{\chi^2(k+2)}(u)\right),$$

where $p_{\chi^2(k)}(u)$ is the density function of a $\chi^2(k)$ distributed random variable, since $\mathrm{cf}_{\text{GMM-AR}(\theta)}(t)=\int\exp(itu)p_{\text{GMM-AR}(\theta)}(u)du$, such that the Edgeworth approximation of the distribution function of the GMM-AR statistic reads

$$\begin{aligned} \Pr\left[\text{GMM-AR}\left(\theta_0\right)\leq x\right] = & \Pr{}_{\chi^2(k)}(x)-\tfrac{\mathrm{E}(GMM\text{-}AR_1)}{Nk}xp_{\chi^2(k)}(x)+O(N^{-2}) \\ = & \Pr{}_{\chi^2(k)}\left(x(1-\tfrac{\mathrm{E}(GMM\text{-}AR_1)}{Nk})\right)+O(N^{-2}), \end{aligned}$$

which uses that $p_{\chi^2(k+2)}(u) = p_{\chi^2(k)}(u)\frac{u}{k}$. Thus the Edgeworth approximation amounts to multiplying a standard $\chi^2(k)$ critical value x by $1+\frac{\mathrm{E}(GMM\text{-}AR_1)}{Nk}$.

A.2. For the Edgeworth approximation of the distribution of KLM(θ), we construct the conditional characteristic function of KLM(θ) given $D_N(\theta,Y)$:

$$\begin{array}{rl}
\multicolumn{2}{l}{\mathrm{cf}_{\mathrm{KLM}(\theta)}(t|D_N(\theta,Y))} \\
= & \iiint \exp\left(it\mathrm{KLM}(\theta)\right) p_{(\sqrt{N}f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|D_N(\theta,Y))}(u,v,w|D_N(\theta,Y))dudvdw \\
= & \iiint \exp\left(it\left[KLM_0+\frac{1}{N}\left(KLM_1+KLM_2\right)\right]\right) \\
& \quad p_{(\sqrt{N}f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|D_N(\theta,Y))}(u,v,w|D_N(\theta,Y))dudvdw+o(\frac{1}{N}) \\
= & \iiint \exp\left(itKLM_0\right)\exp\left(\frac{it}{N}\left(KLM_1+KLM_2\right)\right) \\
& \quad p_{(\sqrt{N}f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|D_N(\theta,Y))}(u,v,w|D_N(\theta,Y))dudvdw+o(\frac{1}{N}) \\
= & \iiint \left[1+\frac{it}{N}\left(KLM_1+KLM_2\right)\right]\exp\left(itKLM_0\right) \\
& \quad p_{(\sqrt{N}f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|D_N(\theta,Y))}(u,v,w|D_N(\theta,Y))dudvdw+o(\frac{1}{N}), \\
= & \int \left[1+\frac{it}{N}\iint\left[KLM_1+KLM_2\right]p_{\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|\sqrt{N}f_N(\theta,Y),D_N(\theta,Y)}(v,w|u,D_N(\theta,Y))dvdw\right] \\
& \quad \exp\left(itKLM_0\right)p_{\sqrt{N}f_N(\theta,Y)|D_N(\theta,Y)}(u|D_N(\theta,Y))du+o(\frac{1}{N}),
\end{array}$$

where

$$\begin{array}{l}
p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|D_N(\theta,Y))}(u,v,w|D_N(\theta,Y))= \\
p_{\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|\sqrt{N}f_N(\theta,Y),D_N(\theta,Y))}(u,v,w|D_N(\theta,Y))p_{\sqrt{N}f_N(\theta,Y)|D_N(\theta,Y)}(u|D_N(\theta,Y))
\end{array}$$

is the conditional density of $(f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta))$ given $D_N(\theta,Y)$ and $p_{\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta)|f_N(\theta,Y),D_N(\theta,Y))}(v,w|u,D_N(\theta,Y))$ and $p_{\sqrt{N}f_N(\theta,Y)|D_N(\theta,Y)}(u|D_N(\theta,Y))$ are the conditional densities of $(\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta))$ given $(f_N(\theta,Y),D_N(\theta,Y))$ and of $\sqrt{N}f_N(\theta,Y)$ given $D_N(\theta,Y)$.

We use the limiting distributions to construct the expectations. Because of the independence of $f_N(\theta,Y)$ and $D_N(\theta,Y)$ in large samples, the conditional limiting distribution of $\sqrt{N}f_N(\theta,Y)$ given $D_N(\theta,Y)$ is $N(0,V_{ff}(\theta_0)$. We combine the exponent term of the density function of this limiting distribution with $\exp(itKLM_0)$:

$$\begin{array}{rl}
\multicolumn{2}{l}{\exp\left[itKLM_0-\frac{1}{2}Nf_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)\right]} \\
= & \exp\left[itNf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)-\right. \\
& \left.\frac{1}{2}Nf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}\left(P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}+M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}\right)V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)\right] \\
= & \exp\left[-\frac{1}{2}\left(1-2it\right)Nf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)-\right. \\
& \left.\frac{1}{2}Nf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)\right] \\
= & \exp\left[-\frac{1}{2}\left(1-2it\right)KLM_0-\right. \\
& \left.\frac{1}{2}Nf_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)\right].
\end{array}$$

We now conduct a transformation of random variables from $V_{ff}(\theta)^{-\frac{1}{2}}\sqrt{N}f_N(\theta,Y)$

to

$$\begin{pmatrix} (1-2it)^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y) & 0 \\ 0 & D_N(\theta,Y)_{\perp}'V_{ff}(\theta)D_N(\theta,Y)_{\perp} \end{pmatrix}^{-\frac{1}{2}}$$
$$\begin{pmatrix} D_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}} \\ D_N(\theta,Y)_{\perp}'V_{ff}(\theta)^{\frac{1}{2}} \end{pmatrix} V_{ff}(\theta)^{-\frac{1}{2}}\sqrt{N}f_N(\theta,Y),$$

with $D_N(\theta,Y)_{\perp}$: $(k-1)\times k$, $D_N(\theta,Y)_{\perp}'D_N(\theta,Y) \equiv 0$ and which uses that $V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}} = P_{V_{ff}(\theta)^{\frac{1}{2}}D_N(\theta,Y)_{\perp}}$. The Jacobian of this transformation is equal to $(1-2it)^{-\frac{1}{2}}$. The KLM_1 and KLM_2 higher order components of KLM(θ) contain both elements that result from the transformation and we therefore specify them as:

$$\begin{aligned}
&KLM_1 = -\tfrac{N}{1-2it}\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right]' V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}} \\
&\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right]V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right] - \\
&2\tfrac{N}{\sqrt{(1-2it)}}\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right]' V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\
&V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\sqrt{N}f_N(\theta,Y)\right] \\
&KLM_2 = -2\tfrac{N}{\sqrt{(1-2it)}}\left[\sqrt{N}f_N(\theta,Y)\right]' V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1} \\
&f_N(\theta,Y)\left[D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right]^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}\left[\sqrt{N(1-2it)}f_N(\theta,Y)\right].
\end{aligned}$$

We now use the results from Theorem 1 that:

$$\begin{aligned}
a = {} & \mathrm{E}\Big[N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\
& V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)|D_N\Big] \\
= {} & \tfrac{N-1}{N}\left\{\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left(V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\right)\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})-4\right\}-\tfrac{1}{N} \\
b = {} & \mathrm{E}\Big[N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}P_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\left[\hat{V}_{ff}(\theta)-V_{ff}(\theta)\right] \\
& V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}f_N(\theta,Y)|D_N\Big] \\
= {} & \tfrac{N-1}{N}\left\{\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}\left(V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\right)\mathrm{vec}(P_{V_{ff}^{-\frac{1}{2}}D_N})-(k-1)\right\} \\
c = {} & \mathrm{E}\Big[N^2 f_N(\theta,Y)'V_{ff}(\theta)^{-\frac{1}{2}}M_{V_{ff}(\theta)^{-\frac{1}{2}}D_N(\theta,Y)}V_{ff}(\theta)^{-\frac{1}{2}}\hat{V}_{\theta f}(\theta)V_{ff}(\theta)^{-1}f_N(\theta,Y) \\
& \left[D_N(\theta,Y)'V_{ff}(\theta)^{-1}D_N(\theta,Y)\right]^{-1}D_N(\theta,Y)'V_{ff}(\theta)^{-1}f_N(\theta,Y)|D_N\Big] \\
= {} & \tfrac{2(N-1)}{N}\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}d_i|D_N)V_{ff}^{-\frac{1}{2}}D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1}+ \\
& \tfrac{1}{N}\mathrm{vec}(M_{V_{ff}^{-\frac{1}{2}}D_N})'\mathrm{E}(\tfrac{1}{N}\textstyle\sum_{i=1}^N V_{ff}^{-\frac{1}{2}}f_if_i'V_{ff}^{-1}f_if_i'V_{ff}^{-\frac{1}{2}}\otimes V_{ff}^{-\frac{1}{2}}d_i|D_N) \\
& \mathrm{vec}(V_{ff}^{-\frac{1}{2}}D_N\left[D_N'V_{ff}^{-1}D_N\right]^{-1}),
\end{aligned}$$

to obtain the characteristic function:

$$\begin{array}{rl} \mathrm{cf}_{\mathrm{KLM}(\theta)}(t|D_N(\theta_0,Y)) = & \int \left[1+\frac{it}{N}KLM_1+\frac{it}{N}KLM_2\right]\exp\left(itKLM_0\right) \\ & p_{(f_N(\theta,Y),\hat{V}_{ff}(\theta),\hat{V}_{\theta f}(\theta))}(u,v,w|D_N(\theta,Y))dudvdw+o(\frac{1}{N}) \\ = & (1-2it)^{-\frac{1}{2}}\left[1-\frac{1}{N}\frac{it}{(1-2it)}a-\frac{2}{N}\frac{it}{\sqrt{1-2it}}(b+c)+o(\frac{1}{N})\right], \end{array}$$

where $(1-2it)^{-\frac{1}{2}}$ results from the Jacobian of the transformation from KLM_0 to $(1-2it)KLM_0$. The density function that leads to the above characteristic function reads

$$p_{\mathrm{KLM}(\theta)}(u) = \quad p_{\chi^2(1)}(u)+\frac{1}{N}\left[(b+c)\left(\frac{\partial}{\partial u}p_{\chi^2(2)}(u)\right)+a\left(\frac{\partial}{\partial u}p_{\chi^2(3)}(u)\right)\right]+o(\frac{1}{N}),$$

so the resulting Edgeworth approximation reads:

$$\begin{array}{l} \Pr\left[\mathrm{KLM}\left(\theta_0\right)\leq x|D_N(\theta_0,Y)\right] = \Pr_{\chi^2(1)}(x)+ \\ \frac{1}{N}\left((b+c)p_{\chi^2(2)}(x)+ap_{\chi^2(3)}(x)\right)+o(N^{-1}) \\ = \Pr_{\chi^2(1)}(x)+\frac{1}{N}\left(ax+\sqrt{2\pi}(b+c)x^{\frac{1}{2}}\right)p_{\chi^2(1)}(x)+o(N^{-1}) \\ = \Pr_{\chi^2(1)}\left(x+\frac{1}{N}\left(ax+\sqrt{2\pi}(b+c)x^{\frac{1}{2}}\right)\right)+o(N^{-1}), \end{array}$$

which uses that $p_{\chi^2(2)}(x)=\sqrt{2\pi}\sqrt{x}p_{\chi^2(1)}(x)$ and $p_{\chi^2(3)}(x)=xp_{\chi^2(1)}(x)$. Thus the Edgeworth approximation alters a $\chi^2(1)$ critical value x towards $x-\frac{1}{N}\left(ax+\sqrt{2\pi}(b+c)\sqrt{x}\right)$.
B.1. For the Edgeworth approximation of the distribution of GMM-AR$^*(\theta)$, we construct the characteristic function of GMM-AR$^*(\theta_0)$:

$$\begin{array}{rl} \mathrm{cf}_{\text{GMM-AR}^*(\theta)}(t) = & \int \exp\left(it\text{GMM-AR}^*(\theta)\right)p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv \\ = & \int \exp\left(it\left[GMM\text{-}AR_0^*+\frac{1}{B}GMM\text{-}AR_1^*\right]\right)p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv+o(B^{-1}) \\ = & \int\left[1+\frac{it}{B}GMM\text{-}AR_1^*\right]\exp\left(itGMM\text{-}AR_0^*\right)p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv+o(B^{-1}), \end{array}$$

where $p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)$ is the joint density of $f_B^*(\theta,Y)$ and $V_{ff}^*(\theta)$ for a bootstrap sample of size B. The order of the error term results from the Mean Value Theorem and the convergence rates of $\left(\frac{it}{B}GMM\text{-}AR_1^*\right)^2$. The addition of $\exp(itGMM\text{-}AR_0^*)$ alters the variance of the limiting distribution of $\sqrt{B}f_B^*(\theta,Y)$ so it becomes $N(0,\frac{1}{1-2it}\hat{V}_{ff}(\theta))$. The characteristic function then becomes

$$\mathrm{cf}_{\text{GMM-AR}^*(\theta)}(t) = \quad (1-2it)^{-\frac{1}{2}k}\left[1+\frac{1}{B}\frac{it}{(1-2it)}\mathrm{E}(GMM\text{-}AR_1^*)+o(B^{-1})\right],$$

where $\mathrm{E}(GMM\text{-}AR_1^*)$ results from the higher order specification of GMM-AR$^*(\theta)$ in Theorem 2, such that the Edgeworth approximation reads:

$$\Pr\left[GMM\text{-}AR^*\left(\theta_0\right)\leq x\right] = \quad \Pr_{\chi^2(k)}\left(x-\frac{1}{B}\frac{\mathrm{E}(GMM\text{-}AR_1^*)}{k}x\right)+o(B^{-1}).$$

B.2. For the Edgeworth approximation of the distribution of KLM$^*(\theta)$, we construct the characteristic function of KLM$^*(\theta_0)$ given $\hat{D}_N(\theta_0,Y)$:

$$\begin{array}{l} \mathrm{cf}_{\mathrm{KLM}^*(\theta)}(t|\hat{D}_N(\theta_0,Y)) \\ \quad = \int \exp\left(it\mathrm{KLM}^*(\theta)\right) p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv \\ \quad = \int \exp\left(it\left[KLM_0^* + \frac{1}{B}KLM_1^*\right]\right) p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv + o(B^{-1}) \\ \quad = \int \left[1+\frac{it}{B}KLM_1^*\right]\exp\left(itKLM_0^*\right) p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv + o(B^{-1}). \end{array}$$

Identical to KLM(θ), the exponent term $\exp(itKLM_0^*)$ alters the exponent term of the density function of the limiting distribution such that the characteristic function of KLM$^*(\theta)$ reads

$$\begin{array}{l} \mathrm{cf}_{\mathrm{KLM}^*(\theta)}(t|\hat{D}_N(\theta_0,Y)) \\ \quad = \int \left[1+\frac{it}{B}KLM_1^*\right]\exp\left(itKLM_0^*\right) p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv + o(B^{-1}) \\ \quad = (1-2it)^{-\frac{1}{2}}\left[1-\frac{1}{B}\frac{it}{(1-2it)}\mathrm{E}(a^*) - \frac{2}{B}\frac{it}{\sqrt{1-2it}}\mathrm{E}(b^*) + o(B^{-1})\right], \end{array}$$

where

$$\begin{array}{rcl} a^* &=& \frac{B-1}{B}\left\{\frac{1}{N}\sum_{i=1}^N\left[\bar{f}(\theta_0)_i'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}(\theta_0)\right]^2 - 4\right\} - \frac{1}{B}, \\ b^* &=& \frac{B-1}{B}\left\{\frac{1}{N}\sum_{i=1}^N\left[\bar{f}(\theta_0)_i'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}(\theta_0)\right.\right. \\ && \left.\left.\bar{f}(\theta_0)_i'\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}M_{\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\hat{D}_N(\theta_0,Y)}\hat{V}_{ff}(\theta_0)^{-\frac{1}{2}}\bar{f}(\theta_0)\right] - (k-1)\right\}. \end{array}$$

The Edgeworth approximation of KLM$^*(\theta_0)$ then reads

$$\begin{array}{l} \Pr\left[\mathrm{KLM}^*\left(\theta_0\right) \leq x|\hat{D}_N(\theta_0,Y)\right] = \Pr_{\chi^2(1)}(x) + \\ \frac{1}{B}\left(a^*x+\sqrt{2\pi}b^*x^{\frac{1}{2}}\right)p_{\chi^2(1)}(x) + o(B^{-1}) \\ = \Pr_{\chi^2(1)}\left(x+\frac{1}{B}\left(a^*x+\sqrt{2\pi}b^*x^{\frac{1}{2}}\right)\right) + o(B^{-1}). \end{array}$$

B.3. For the Edgeworth approximation of the distribution of KLM$^{**}(\theta)$, we construct the characteristic function of KLM$^{**}(\theta_0)$ given $\hat{D}_N(\theta_0,Y)$:

$$\begin{array}{l} \mathrm{cf}_{\mathrm{KLM}^{**}(\theta)}(t|\hat{D}_N(\theta,Y)) \\ \quad = \iiint \exp\left(it\mathrm{KLM}^{**}(\theta)\right) p_{(\sqrt{B}f_B^*(\theta,Y),V_{ff}^*(\theta),V_{\theta f}^*(\theta)|\hat{D}_N(\theta,Y))}(u,v,w|\hat{D}_N(\theta,Y))dudvdw \\ \quad = \iiint \exp\left(it\left[KLM_0^*+\frac{1}{B}\left(KLM_1^*+KLM_2^{**}\right)\right]\right) \\ \qquad\quad p_{(\sqrt{B}f_N(\theta,Y),V_{ff}^*(\theta),V_{\theta f}^*(\theta)|\hat{D}_N(\theta,Y))}(u,v,w|\hat{D}_N(\theta,Y))dudvdw + o(\frac{1}{B}) \\ \quad = \int\left[1+\frac{it}{B}\iint\left[KLM_1^*+KLM_2^{**}\right]p_{V_{ff}^*(\theta),V_{\theta f}^*(\theta)|\sqrt{B}f_B^*(\theta,Y),\hat{D}_N(\theta,Y)}(v,w|u,\hat{D}_N(\theta,Y))dvdw\right] \\ \qquad\quad \exp\left(itKLM_0\right)p_{\sqrt{B}f_B^*(\theta,Y)|\hat{D}_N(\theta,Y)}(u|D_N(\theta,Y))du + o(\frac{1}{B}), \end{array}$$

The characteristic function of KLM$^{**}(\theta)$ reads

$$\begin{aligned}
&\mathrm{cf}_{\mathrm{KLM}^{**}(\theta)}(t|\hat{D}_N(\theta_0, Y)) \\
&= \int \left[1 + \tfrac{it}{B} KLM_1^*\right] \exp\left(itKLM_0^*\right) p_{(f_B^*(\theta,Y),V_{ff}^*(\theta))}(u,v)dudv + o(B^{-1}) \\
&= (1-2it)^{-\frac{1}{2}} \left[1 - \tfrac{1}{B}\tfrac{it}{(1-2it)}\mathrm{E}(a^*) - \tfrac{2}{B}\tfrac{it}{\sqrt{1-2it}}\left(\mathrm{E}(b^*) + \mathrm{E}(c^*)\right) + o(B^{-1})\right],
\end{aligned}$$

where a^* and b^* are defined in B.2 and c^* =E$\Big[KLM_2^{**}|\hat{D}_N(\theta_0, Y)\Big]$ which is stated in (45).The Edgeworth approximation of KLM$^{**}(\theta_0)$ then reads

$$\begin{aligned}
&\Pr\left[\mathrm{KLM}^{**}\left(\theta_0\right) \leq x|\hat{D}_N(\theta_0, Y)\right] = \Pr_{\chi^2(1)}(x) + \\
&\tfrac{1}{B}\left(a^*x + \sqrt{2\pi}b^*x^{\frac{1}{2}}\right)p_{\chi^2(1)}(x) + o(B^{-1}) \\
&= \Pr_{\chi^2(1)}\left(x + \tfrac{1}{B}\left(a^*x + \sqrt{2\pi}(b^* + c^*)x^{\frac{1}{2}}\right)\right) + o(B^{-1}).
\end{aligned}$$

References

[1] Anderson, T.W. and C. Hsiao. Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*, **76**:598–606, 1981.

[2] Anderson, T.W. and H. Rubin. Estimation of the Parameters of a Single Equation in a Complete Set of Stochastic Equations. *The Annals of Mathematical Statistics*, **21**:570–582, 1949.

[3] Arellano, M. and B. Honoré. Panel Data Models: Some Recent Developments. In J.L. Heckman and E. Leamer, editors, *Handbook of Econometrics*, volume **5**, chapter 53. Elsevier North-Holland, 2001.

[4] Arellano, M. and S. Bond. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, **58**:277–297, 1991.

[5] Bhattacharya, R.N. and J.K. Ghosh. On the Validity of the Formal Edgeworth Expansion. *Annals of Statistics*, **6**:434–451, 1978.

[6] Eicker, F. Limit Theorems for Regressions with Unequal and Dependent Errors. In L.M. LeCam and J. Neyman, editor, *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics*. Berkeley: University of California Press, 1967.

[7] Hansen, L.P. Large Sample Properties of Generalized Method Moments Estimators. *Econometrica*, **50**:1029–1054, 1982.

[8] Hansen, L.P., J. Heaton and A. Yaron. Finite Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics*, **14**:262–280, 1996.

[9] Horowitz, J.L. The Bootstrap. In J.L. Heckman and E. Leamer, editors, *Handbook of Econometrics*, volume 5, chapter 52, pages 3159–3228. Elsevier Science B.V., 2002.

[10] Kleibergen, F. Testing Parameters in GMM without assuming that they are identified. *Econometrica*, **73**:1103–1124, 2005.

[11] Kleibergen, F. Generalizing weak instrument robust IV statistics towards multiple parameters, unrestricted covariance matrices and identification statistics. *Journal of Econometrics*, **139**:181–216, 2007.

[12] Kleibergen, F. and S. Mavroeidis. Weak instrument robust tests in GMM and the new Keynesian Phillips curve. *Journal of Business and Economic Statistics*, **27**:293–311, 2009.

[13] Kleibergen, F. and S. Mavroeidis. Inference on subsets of parameters in linear IV without assuming identification. 2010. Working Paper, Brown University.

[14] Mammen, E. *When Does the Bootstrap Work? Asymptotic Results and Simulations*. Springer, New York, 1992.

[15] Moreira, M.J.,. A Conditional Likelihood Ratio Test for Structural Models. *Econometrica*, **71**:1027–1048, 2003.

[16] Moreira, M.J., J.R. Porter and G. Suarez. Bootstrap validity of the score test when instruments may be weak. *Journal of Econometrics*, **149**:52–64, 2009.

[17] Newey, W.K. and K.D. West. Hypothesis Testing with Efficient Method of Moments Estimation. *International Economic Review*, **28**:777–787, 1987.

[18] Nickell, S.J. Biases in dynamic models with fixed effects. *Econometrica*, **49**:1417–1426, 1981.

[19] Phillips, P.C.B. and J.Y. Park. On the Formulation of Wald Tests of Nonlinear Restrictions. *Econometrica*, **56**:1065–1083, 1988.

[20] Rothenberg, T.J. Approximating the Distributions of Econometric Estimators and Test Statistics. In Z. Griliches and M.D. Intrilligator, editor, *Handbook of Econometrics, Volume 2*, chapter 15, pages 881–935. Elsevier Science B.V., 1984.

[21] Sargan, J.D. Econometric Estimators and the Edgeworth Approximation. *Econometrica*, **44**:421–448, 1976.

[22] Sargan, J.D. Some Approxmations to the Distributions of Econometric Criteria which are Asymptotically distributed as Chi-Squared. *Econometrica*, **48**:1107–1138, 1980.

[23] Staiger, D. and J.H. Stock. Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65**:557–586, 1997.

[24] Stock, J.H. and J.H. Wright. GMM with Weak Identification. *Econometrica*, **68**:1055–1096, 2000.

[25] White, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, **48**:817–838, 1980.