Pooling is Harmful Sometimes^{*}

Donggyu Sul University of Texas at Dallas

March 29, 2016

Abstract

It is widely believed that the pooled within group (WG) estimator consistently estimates the mean of a cross section of heterogeneous slope coefficients. In this paper we show that the WG estimator is inconsistent when the heterogeneous slope coefficients are correlated with the variances of the regressors. In order to establish whether the WG estimator is consistent in practice we propose a new poolability test based on the difference between the WG estimator and the mean group (MG) estimator. Extensions to instrumental variables methods are explored. By the means of Monte Carlo simulations, we show that the proposed poolability test works well in the finite sample.

Keywords: Poolability Test, Heterogeneous Slope Coefficients, Mean Group Estimator, Panel DID regressions, Trimmed Mean Group Estimator. JEL Classification: C33

^{*}I thank Alexander Chudik and Ryan Greenaway-McGrevy for helpful comments.

1 Introduction

This paper deals with the popular panel fixed effects, or difference-in-difference (DID), regression given by (for example, with one regressor case)

$$y_{it} = a_i + \lambda_t + \beta x_{it} + u_{it}; \ i = 1, ..., n; \ t = 1, ..., T,$$

where a_i and λ_t are individual and time fixed effects, respectively. When the time series observations (T) are much larger than the number of cross sectional units (n), homogeneity of the slope coefficients is testable. As Baltagi, Bresson and Pirotte (2008) point out, the homogeneity restriction, however, is frequently rejected. However, when n is much larger than T, it has been popular practice to pool the cross sectional and time series information as a general homogeneity test is not available.

Under the presence of the heterogeneity of the slope coefficients, pooling the observations is widely-believed to be harmless. In fact, in the treatment literature, the pooled slope coefficient is interpreted as an "average" treatment effect since the individual treatment effects can be heterogeneous. As Baltagi and Griffin (1997) and Woodridge (2005) point out, the standard fixed-effects, or within group (WG), estimator consistently estimates the average of the heterogeneous slope coefficients.

Meanwhile, Pesaran and Smith (1995) and Pesaran, Shin and Smith (1999) proposed the socalled mean group (MG) estimator which is the simple cross sectional average of the time series least squares (LS) estimators. Maddala, Trost, Li and Joutz (1997) considered a model average or shrinkage estimator by utilizing both time series LS estimators and the WG estimator. See Baltagi, Bresson and Pirotte (2002, 2004, 2008) and Baltagi, Griffin and Xiong (2000) for further discussion. More importantly Haque, Pesaran and Sharma (2000) pointed out that the WG estimator becomes inconsistent if the heterogeneous slope coefficients are correlated with the variance of the regressors. If the heterogeneous slope coefficients are correlated with the variances of the regressors, then imposing the homogeneity restriction leads to inconsistent estimation. Hence it is desirable to test the validity of the homogeneity restriction before the WG estimation is used, but there is no test available.

This paper proposes a simple test for the validity of the homogeneity restriction. The MG estimator remains consistent regardless of the potential correlation between the variances of the regressors and heterogeneous slope coefficients. The validity of the homogeneity restriction can be examined by evaluating whether or not the WG estimators are statistically different from the MG estimators. The test tells the practitioner that they should not use the WG estimator when the null hypothesis is rejected. The asymptotic properties of the proposed test are studied under large n with any size of T. As long as $n \to \infty$, the limiting distribution of the proposed test is standard

normal regardless of the size of T.

In the next section, we provide some economic examples demonstrating how the variances of the regressors can be correlated with the slope coefficients. Section 3 provides the asymptotic properties of the suggested test. Section 4 reports the results of Monte Carlo studies and presents one empirical example. Section 5 concludes.

2 Preliminaries and Economic Examples

Throughout this section, we consider a single regressor case for notational simplicity and clear exposition. Consider the following linear regression with a single regressor.

$$y_{it} = a_i + \lambda_t + \beta_i x_{it} + e_{it},\tag{1}$$

Imposing the homogeneity restriction on (1) leads to

$$y_{it} = a_i + \lambda_t + \beta x_{it} + u_{it}, \text{ for } u_{it} = x_{it} \left(\beta_i - \beta\right) + e_{it}.$$
(2)

As long as the slope coefficient, β_i , is not dependent on the variance of x_{it} , the WG estimator becomes consistent even when $\beta_i \neq \beta$ for some *i*.

Next consider the following possibility where the regression coefficient, β_i , becomes correlated with the variance of x_{it} . That is,

$$\phi = Cov\left(\beta_i, \sigma_{x,i}^2\right) \neq 0,\tag{3}$$

where $\sigma_{x,i}^2$ is the variance of the regressor x_{it} , that is $\sigma_{x,i}^2 = \mathbb{E}(T-1)^{-1} \sum_{t=1}^T \left(x_{it} - T^{-1} \sum_{t=1}^T x_{it} \right)^2$.

Here we provide two empirical examples where the regression coefficients may be correlated with the variances of the regressors.

Example 1: Treatment Effects with Missing Doses

In medical science the effectiveness of a new medicine has been evaluated by means of experiments. When a subject (or patient) has not taken the new medicine regularly, the subject may not receive the full advantage of the new medicine. As the number of missing doses increases, the treatment effect of the new medicine may decrease. Let I_{it} be the treatment of the new medicine. If a subject takes the medicine at time t, $I_{it} = 1$. Otherwise $I_{it} = 0$. Then the frequency of missed doses becomes the variance of the regressor, I_{it} . Among many others, see Hernandez-Hernandez, *et al.* (1996) as example, where they investigated the blood pressure effects of missing a dose of a prescribed medicine. In addition, there have been a number of studies regarding the effects of the

frequency of doses. For example, the treatment effect of taking 1,500 mg of antibiotics once a day has been found to differ from that of taking 500 mg of antibiotics three times a day. See Dunn et al. (2005).

Example 2: Income Uncertainty and Precautionary Saving

It is widely well known that an increase in income uncertainty leads to a higher rate of saving. See Canallero (1990) and Guiso, Jappelli and Terlizzese (1992) for more detailed discussions. Suppose that a researcher wants to estimate the marginal propensity to consume (MPC) and runs the following panel regression by imposing the homogeneity restriction on the slope coefficients.

$$c_{it} = a_i + \lambda_t + \beta w_{it} + u_{it},$$

where c_{it} is the log consumption for an individual *i* at time *t*, and w_{it} is the log wealth or income. It is natural to assume that the slope coefficient β is heterogeneous across individuals as some people spend consistently more money than others. Hence the regression error u_{it} may include $(\beta_i - \beta) w_{it}$ as an additive term. If the higher variance or uncertainty of wealth leads to lower MPC (see Kimball (1990) for more additional discussions), then the slope coefficients become dependent on the variances of the independent variables.

When the slope coefficients are correlated with the variances of the regressors, the WG estimators are inconsistent whereas the MG estimators are consistent. Let $\phi \neq 0$ in (3). Further let $\tilde{x}_{it} = x_{it} - T^{-1} \sum_{t=1}^{T} x_{it} - n^{-1} \sum_{i=1}^{n} x_{it} + (nT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} x_{it}$, and similarly define \tilde{e}_{it} . Then the WG estimator, $\hat{\beta}_{wg}$, is given by

$$\hat{\beta}_{wg} - \beta = \frac{\sum_{i=1}^{n} (\beta_i - \beta) \sum_{t=1}^{T} \tilde{x}_{it}^2}{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it}^2} + \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{e}_{it}}{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it}^2}.$$
(4)

Assume that x_{it} is strictly exogenous; $Ex_{it}e_{js} = 0$ for all i, j, t and s. Then it is easy to show that

$$\operatorname{plim}_{n \to \infty} \left(\hat{\beta}_{\mathrm{wg}} - \beta \right) \neq 0 \text{ if } \phi \neq 0,$$
(5)

where β is the mean of β_i .

Meanwhile the LS estimator of β_i for each *i* is consistent as long as x_{it} is strictly exogenous. Let $\hat{\beta}_i$ be the LS estimator for the *i*th time series regression. Then the MG estimator proposed by Pesaran and Smith (1995) is given by

$$\hat{\beta}_{\rm mg} = \frac{1}{n} \sum_{i=1}^{n} \hat{\beta}_i \to^p \beta \text{ as } n \to \infty, \tag{6}$$

under regularity conditions which we will discuss shortly.

It is important to address the fact that the inclusion of x_{it}^2 in the panel fixed effects regression does not solve this problem. To see this, consider the following augmented panel regression with x_{it}^2 .

$$y_{it} = a_i + \lambda_t + \beta x_{it} + \gamma x_{it}^2 + u_{it}, \text{ for } u_{it} = (\beta_i - \beta) x_{it} + e_{it}.$$
(7)

Since $E\sum_{i=1}^{n} \tilde{x}_{it}^3 (\beta_i - \beta) \neq E\sum_{i=1}^{n} \tilde{x}_{it}^2 (\beta_i - \beta)$, the WG estimator for β is still inconsistent if $E(\beta_i - \beta) \sigma_{x,i}^2 \neq 0$. Note that the inclusion of x_{it}^3 does not solve this problem either.

When $\phi \neq 0$, the regression should not be pooled. Hence the poolability can be defined as

$$\mathcal{H}_0$$
: Poolability $\iff \mathcal{H}_0: \phi = 0.$ (8)

Since the MG estimator remains consistent regardless of the values of ϕ but the WG estimator becomes consistent only under the null hypothesis of the poolability, testing for the null of the poolability can be done by measuring the difference between the MG and the WG estimators. To be specific, let

$$S = \frac{\sqrt{n} \left(\hat{\beta}_{\rm wg} - \hat{\beta}_{\rm mg} \right)}{\sqrt{V \left(\hat{\beta}_{\rm wg} - \hat{\beta}_{\rm mg} \right)}}.$$
(9)

Later we will show that under regularity conditions, the limiting distribution of the S-statistic becomes the standard normal distribution as $n \to \infty$. That is,

$$\mathcal{S} \to^d \mathcal{N}(0,1) \text{ as } n \to \infty.$$

3 Asymptotic Properties of the Poolability Test

Here we first consider the asymptotic properties of the poolability test when there is no lagged dependent variable. Later we will consider the dynamic panel regression as a special case.

3.1 Poolability Test with Static Panel Regressions

Rewrite the regression model with multiple regressors as

$$y_{it} = a_i + \lambda_t + \mathbf{x}_{it}\boldsymbol{\beta} + u_{it}, \text{ for } u_{it} = \mathbf{x}_{it} \left(\boldsymbol{\beta}_i - \boldsymbol{\beta}\right) + e_{it},$$
(10)

where \mathbf{x}_{it} is the $(1 \times k)$ vector of regressors and $\boldsymbol{\beta}_i$ is the $(k \times 1)$ vector of the regression coefficients. Define the asymptotic covariance between the regressors and the slope coefficients as $\Sigma_{\boldsymbol{\xi}\boldsymbol{\beta}}$. That is,

$$\Sigma_{\boldsymbol{\xi}\boldsymbol{\beta}} = \operatorname{plim}_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{1}{T} \boldsymbol{\xi}_{iT} - \frac{1}{nT} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right) (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right]$$
(11)

where $\boldsymbol{\xi}_{iT} = \sum_{t=1}^{T} \tilde{\mathbf{x}}'_{it} \tilde{\mathbf{x}}_{it}$. We make the following assumptions.

Assumption

(i) $\operatorname{Ee}_{it}e_{jt} = 0$ if $i \neq j$. (ii) $\operatorname{E}(\beta_{ij} - \beta_j)(\beta_{mj} - \beta_j) = 0$ if $i \neq m$, where $\beta_j = \operatorname{E}(\beta_{ij})$. (iii) $\operatorname{Ex}_{ij,t}e_{ms} = 0$ for all i, j, t, m and s.

We also need other regularity conditions such that all regressors have finite fourth moments and the error terms have zero means with finite second moments. Assumption (i) and (ii) imply cross sectional independence which is the typical assumption in this literature. We can permit weak cross sectional dependence but for the simplicity, we do not consider this case. When Assumption (i) and (ii) are violated, then factor augmented regressors proposed by Pesaran (2006), Bai (2009) and Greenaway-McGrevy, Han and Sul (2012) can be used. See remark 3 for more detailed discussion. Assumption (iii) represents strict exogeneity so that lagged dependent variables are precluded. In fact, Assumption (iii) can be relaxed so that the WG and the MG estimators share the same probability limit under the null of poolability. See Remark 2 for a detailed discussion.

The WG estimator can be written as

$$\hat{\boldsymbol{\beta}}_{wg} - \boldsymbol{\beta} = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) + \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left\{ \sum_{i=1}^{n} \left(\left[\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right] (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right) \right\} + \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right).$$
(12)

Meanwhile the MG estimator is given by

$$\hat{\boldsymbol{\beta}}_{mg} - \boldsymbol{\beta} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\beta}}_{i} - \boldsymbol{\beta} \right) = \frac{1}{n} \sum_{i=1}^{n} \left(\boldsymbol{\beta}_{i} - \boldsymbol{\beta} \right) + \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right).$$
(13)

The difference between the WG and MG estimators becomes

$$\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left\{ \sum_{i=1}^{n} \left(\left[\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right] (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right) \right\} \\ + \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) - \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) \right].$$

Under the null hypothesis of the poolability, $\boldsymbol{\xi}_{iT}$ is not correlated with $\boldsymbol{\beta}_i$. Hence

$$\operatorname{plim}_{n \to \infty} \left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}} - \hat{\boldsymbol{\beta}}_{\mathrm{mg}} \right) = 0.$$
 (14)

Let the covariance matrix of $\left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}} - \hat{\boldsymbol{\beta}}_{\mathrm{mg}} \right)$ be

$$\Omega_{\Delta\beta}^{2} = \mathrm{E}\left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}} - \hat{\boldsymbol{\beta}}_{\mathrm{mg}}\right) \left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}} - \hat{\boldsymbol{\beta}}_{\mathrm{mg}}\right)',\tag{15}$$

and the normalized statistic, \mathcal{S} as

$$S = \sqrt{n} \Omega_{\Delta\beta}^{-1} \left(\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} \right).$$
(16)

Then we have

Theorem 1 (Limiting Distribution under the Null Hypothesis)

Under Assumption A, as
$$n \to \infty$$
 with any fixed T,
(i) $\sqrt{n} \left(\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} \right) \to^{d} \mathcal{N} \left(0, \Omega^{2}_{\Delta \boldsymbol{\beta}} \right),$
(ii) $\mathcal{S} \to^{d} \mathcal{N} \left(0, \mathbf{I}_{k} \right)$

See Appendix for the proof. The covariance matrix of $\Omega^2_{\Delta\beta}$ is unknown but can be estimated consistently. We provide the following lemma.

Lemma 1

Let $\mathbf{z}_{iT} = \left(\boldsymbol{\xi}_{iT} - \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\xi}_{iT}\right)\left(\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\beta}}_{i}\right)$. Then under the null of poolability, the difference between the WG and MG estimators can be expressed as

$$\hat{\boldsymbol{\beta}}_{\rm wg} - \hat{\boldsymbol{\beta}}_{\rm mg} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left[\mathbf{z}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_{iT} \right] + O_p\left(n^{-1}\right).$$
(17)

See the Appendix for the proof of Lemma 1. Let $\eta_{iT} = \frac{1}{T} \left(\mathbf{z}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_{iT} \right)$ and $\Omega_{\eta}^2 = \mathbf{E} \frac{1}{n} \left(\sum_{i=1}^{n} \eta_{iT} \right) \left(\sum_{i=1}^{n} \eta_{iT} \right)'$. Then the probability limit of the sample covariance matrix of η_{iT} becomes

$$\operatorname{plim}_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \eta_{iT} \eta_{iT}' = \Omega_{\eta}^{2}, \tag{18}$$

since η_{iT} is cross sectionally independent. From Lemma 1, the covariance matrix of $\Omega^2_{\Delta\beta}$ can be consistently estimated by

$$\hat{\Omega}_{\Delta\beta}^2 = n \left(\sum_{i=1}^n \boldsymbol{\xi}_{iT}\right)^{-1} \left(\sum_{i=1}^n \eta_{iT} \eta_{iT}'\right) \left(\sum_{i=1}^n \boldsymbol{\xi}_{iT}\right)^{-1} \to^p \Omega_{\Delta\beta}^2.$$
(19)

The S test is similar to Hausman (1978)'s test. The WG estimators can be considered as the restricted estimators with the homogeneity assumption whereas the MG estimators are unrestricted estimators.

Next, we consider the asymptotic properties of the poolability test under the alternative. Let

$$\Sigma_{\boldsymbol{\xi}\boldsymbol{\beta}} = \operatorname{plim}_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{1}{T} \boldsymbol{\xi}_{iT} - \frac{1}{nT} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right) (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right] \neq 0,$$

$$\boldsymbol{\varphi}_{nT} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT} (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right).$$

Then its probability limit becomes

$$\boldsymbol{\varphi}_T = \operatorname{plim}_{n \to \infty} \boldsymbol{\varphi}_{nT} = \mathbf{Q}_{xT}^{-1} \Sigma_{\boldsymbol{\xi} \boldsymbol{\beta}}.$$

Under the alternative, the difference between the WG and MG estimators is given by

$$\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} = \boldsymbol{\varphi}_{nT} + \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left\{ \sum_{i=1}^{n} \left(\left[\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right] \left[\boldsymbol{\epsilon}_{i} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\epsilon}_{i} \right] \right) \right\}$$

$$+ \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) - \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) \right].$$

Then we have

Theorem 2 (Limiting Distribution under the Alternative)

Under the Assumption A, as
$$n \to \infty$$
 with any fixed T,
(i) $\sqrt{n} \left(\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} - \boldsymbol{\varphi}_T \right) \to^d \mathcal{N} \left(0, \Omega^2_{\Delta \boldsymbol{\beta}} \right),$
(ii) $\mathcal{S} \to \pm \infty.$

See the Appendix for the proof. When the null of the poolability is rejected, the WG estimator becomes inconsistent. Here we provide three important remarks.

Remark 1: (Homogeneity and Poolability)

If the slope coefficients are homogeneous, then the WG estimators share the same limiting distribution with the MG estimators, so that Theorem 1 still holds. In other words, homogeneity implies poolability. When $\beta_i = \beta$ for all *i*, the difference between the WG and the MG estimators is rewritten as

$$\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right) - \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right)\right].$$

Also, it is straightforward to show that $\hat{\Omega}^2_{\Delta\beta}$ in (19) is still consistent. Note that the poolability in (8) does not imply the homogeneity at all. However when the null of the poolability is rejected, the null of the homogeneity is automatically rejected also.

Even when $\operatorname{plim}_{n\to\infty}\hat{\boldsymbol{\beta}}_{\mathrm{wg}} = \operatorname{plim}_{n\to\infty}\hat{\boldsymbol{\beta}}_{\mathrm{mg}}$ under the null, it is impossible to know whether or not the source of the heterogeneity comes from the heterogeneous inconsistency or slope coefficients. Hence the poolability test becomes meaningful only when both the WG and MG estimators are consistent. The next remark shows how to use the poolability test with instrumental variables (IV).

Remark 2: (IV estimation and Poolability)

Suppose that Assumption (iii) does not hold. In this case, Theorem 1 can be shown to hold for an IV estimator. Let $\zeta_{iT} = \sum_{t=1}^{T} \tilde{\mathbf{w}}'_{it} \tilde{\mathbf{x}}_{it}$ where \mathbf{w}_{it} is the $(1 \times k)$ vector of IVs. Denote $\hat{\boldsymbol{\beta}}^{\text{iv}}_{\text{wg}}$ and $\hat{\boldsymbol{\beta}}^{\text{iv}}_{\text{mg}}$ as the WG-IV and MG-IV estimators, respectively. Then the difference between the WG-IV and MG-IV estimators is rewritten as

$$\hat{\boldsymbol{\beta}}_{wg}^{iv} - \hat{\boldsymbol{\beta}}_{mg}^{iv} = \left(\sum_{i=1}^{n} \boldsymbol{\zeta}_{iT}\right)^{-1} \left\{ \sum_{i=1}^{n} \left(\left[\boldsymbol{\zeta}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\zeta}_{iT} \right] (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right) \right\} \\ + \left(\sum_{i=1}^{n} \boldsymbol{\zeta}_{iT}\right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{w}}_{it}' \tilde{e}_{it} \right) - \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\zeta}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{w}}_{it}' \tilde{e}_{it} \right) \right]$$

Further let $\mathbf{z}_{iT}^{iv} = \left(\boldsymbol{\zeta}_{iT} - \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\zeta}_{iT}\right)\left(\hat{\boldsymbol{\beta}}_{i}^{iv} - \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\beta}}_{i}^{iv}\right)$, and $\eta_{iT}^{iv} = \frac{1}{T}\left(\mathbf{z}_{iT}^{iv} - \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{iT}^{iv}\right)$. Then the covariance matrix of $\hat{\boldsymbol{\beta}}_{wg}^{iv} - \hat{\boldsymbol{\beta}}_{mg}^{iv}$ can be consistently estimated by

$$\hat{\Omega}_{\Delta\beta,\mathrm{iv}}^{2} = n \left(\sum_{i=1}^{n} \boldsymbol{\zeta}_{iT}\right)^{-1} \left(\sum_{i=1}^{n} \eta_{iT}^{\mathrm{iv}} \eta_{iT}^{\mathrm{iv}\prime}\right) \left(\sum_{i=1}^{n} \boldsymbol{\zeta}_{iT}\right)^{-1} \to^{p} \Omega_{\Delta\beta,\mathrm{iv}}^{2},$$

where $\Omega_{\Delta\beta,\mathrm{iv}}^{2} = \mathrm{E}\left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}}^{\mathrm{iv}} - \hat{\boldsymbol{\beta}}_{\mathrm{mg}}^{\mathrm{iv}}\right) \left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}}^{\mathrm{iv}} - \hat{\boldsymbol{\beta}}_{\mathrm{mg}}^{\mathrm{iv}}\right)'.$

Remark 3: (Factor Augmented Regression)

When the regressors and the regression errors are cross sectionally correlated, Assumption (i) does not hold. In this case, the factor augmented estimators suggested by Pesaran (2006), Bai (2009), Greenaway-McGrevy, Han and Sul (2012), Chudik and Pesaran (2015), and Song (2013) can be used for the poolability test. Consider the following regression with cross sectionally dependent errors.

$$y_{it} = a_i + \mathbf{x}_{it}\boldsymbol{\beta}_i + u_{it}, \quad u_{it} = \lambda'_i F_t + u^o_{it}.$$

The factor augmented regression can be written as

$$y_{it} = a_i + \mathbf{x}_{it}\boldsymbol{\beta}_i + \delta'_i \hat{F}_t + u^*_{it}, \tag{20}$$

where \hat{F}_t are the consistent estimators for $H'F_t$ where H is an invertible rotating matrix. Then it is straightforward to show that as long as $n, T \to \infty$ jointly, Theorem 1 holds. We require the condition of $T \to \infty$, otherwise the principal component estimates of the factors, \hat{F}_t , are not consistent. We don't report the finite sample performance of the poolability test in (20) to save the space. But these results are available online.

Even when Assumption (iii) is violated, the poolability test works as long as $\operatorname{plim}_{n\to\infty}\hat{\beta}_{wg} = \operatorname{plim}_{n\to\infty}\hat{\beta}_{mg}$ under the null. The next subsection shows the case where $\operatorname{plim}_{n\to\infty}\hat{\beta}_{wg} \neq \operatorname{plim}_{n\to\infty}\hat{\beta}_{mg}$ with a fixed T but $\operatorname{plim}_{T,n\to\infty}\hat{\beta}_{wg} = \operatorname{plim}_{T,n\to\infty}\hat{\beta}_{mg}$ as $T, n \to \infty$ sequentially.

3.2 Poolability Test with Dynamic Panel Regressions

Consider the following simple stationary panel AR(1) model as an example.

$$y_{it} = a_i + \rho_i y_{it-1} + e_{it}, \tag{21}$$

where $|\rho_i| < 1$, and we assume that e_{it} is identically and independently distributed. That is, $e_{it} \sim iid(0, \sigma_e^2)$. In this case, Assumption (iii) is violated since $E(y_{it-1}e_{is}) \neq 0$ for any s < t.

Rewrite (21) after imposing the homogeneity restriction.

$$y_{it} = a_i + \rho y_{it-1} + u_{it}, \tag{22}$$

where $u_{it} = (\rho_i - \rho) y_{it-1} + e_{it}$. Note that

$$E\frac{1}{T-1}\sum_{t=1}^{T}\left(y_{it-1} - \frac{1}{T}\sum_{t=1}^{T}y_{it-1}\right)^2 = \frac{\sigma_e^2}{1-\rho_i^2}$$

It is not hard to show that

$$\mathbf{E}\left[\left(\rho_{i}-\rho\right)\cdot\frac{1}{T-1}\sum_{t=1}^{T}\left(y_{it-1}-\frac{1}{T}\sum_{t=1}^{T}y_{it-1}\right)^{2}\right] = \sigma_{e}^{2}\frac{\rho_{i}-\rho}{1-\rho_{i}^{2}} \neq 0.$$
(23)

The WG estimator in (22) becomes inconsistent even when $n, T \to \infty$. Only when $\rho_i = \rho$ for all *i*, the covariance in (23) becomes zero. However, the poolability test is not equivalent to the homogeneity test. As $n \to \infty$ with a fixed *T*, both the WG and the MG estimators become inconsistent even when $\rho_i = \rho$ for all *i*. With a large *T*, the inconsistency can be written as

$$\operatorname{plim}_{n \to \infty} \left(\hat{\rho}_{wg} - \rho \right) = -\left(\frac{1+\rho}{T} \right) + O\left(T^{-2} \right), \qquad (24)$$

$$\operatorname{plim}_{n \to \infty} \left(\hat{\rho}_{\mathrm{mg}} - \rho \right) = -\left(\frac{1+3\rho}{T} \right) + O\left(T^{-2} \right), \qquad (25)$$

where $\hat{\rho}_{mg} = n^{-1} \sum_{i=1}^{n} \hat{\rho}_i$ and $\hat{\rho}_i$ is the LS estimator of ρ_i .¹ Therefore Theorem 1 holds only when $T \to \infty$ first and then $n \to \infty$ sequentially. See Pesaran, Smith and Im (1996), and Pesaran, Shin and Smith (1999) for further discussions on this issue.

Also it is important noting that Lemma 1 holds only when $T, n \to \infty$ sequentially. As shown in Appendix,

$$\hat{\rho}_{\mathrm{wg}} - \hat{\rho}_{\mathrm{mg}} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left[\mathbf{z}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_{iT}\right] + C,$$

where

$$C = -\left(\frac{1}{nT}\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \left(\boldsymbol{\xi}_{iT} - \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right)\right] \frac{1}{\sqrt{nT}}$$

= $O_p\left(n^{-1/2}T^{-1}\right),$

since

$$\operatorname{plim}_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) \right] = \operatorname{plim}_{n \to \infty} \left(\hat{\rho}_{\mathrm{mg}} - \rho \right) = -\left(\frac{1+3\rho}{T} \right) + O\left(T^{-2} \right).$$

Hence Lemma 1 changes as

$$\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left[\mathbf{z}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_{iT} \right] + O_p \left(n^{-1/2} T^{-1} \right).$$

The point of interest is at least theoretically, then, how the poolability test can be used for testing homogeneity for any fixed T. Without correcting the inconsistency, the difference between $\hat{\rho}_{wg}$ and $\hat{\rho}_{mg}$ still exists so that the test statistic becomes invalid with any fixed T. Hence it is required to correct the bias first. Suppose that the bias functions of $\hat{\rho}_{wg}$ and $\hat{\rho}_i$ are known and are monotonic over ρ_i for any T. Then there exist the unique mean unbiased estimators for $\hat{\rho}_{wg}$ and $\hat{\rho}_i$. Denote them as $\hat{\rho}_{wg,mue}$ and $\hat{\rho}_{i,mue}$. Then the exact mean unbiased MG estimator can be defined as

$$\hat{\rho}_{\mathrm{mg,mue}} = n^{-1} \sum_{i=1}^{n} \hat{\rho}_{i,\mathrm{mue}}.$$

Naturally, we have

$$\operatorname{plim}_{n\to\infty}\hat{\rho}_{\mathrm{wg,mue}} = \operatorname{plim}_{n\to\infty}\hat{\rho}_{\mathrm{mg,mue}} = \rho.$$

$$\mathbf{E}\hat{\rho}_i - \rho = -\left(\frac{1+3\rho}{T}\right) + O\left(T^{-2}\right).$$

By taking the cross sectional mean in the above equation, we can get the result in (25).

¹To get the result in (25), we need to assume that $e_{it} \sim iid\mathcal{N}(0, \sigma_e^2)$. Without this assumption of the normal distribution, the explicit expectation form cannot be obtained. From Kendall (1968) and Tanaka (1983), the LS estimator under the null of the homogeneity can be written as

Hence theoretically for any size of T, the poolability test based on the exact mean unbiased (EMU) estimators becomes the homogeneity test as long as $n \to \infty$.

However, there are two serious drawbacks to use the EMU estimator in practice. First, the EMU estimator for $\hat{\rho}_i$ is depending on the distributional assumption in the finite sample. Unless the distribution of the error term is known, the mean unbiased estimator based on a particular distribution becomes approximated mean unbiased estimator rather than EMU. Second, the distribution of the EMU estimator is usually truncated. Let ρ_1 be the expected value of $\hat{\rho}_i$ when $\rho = 1$. Then when $\hat{\rho}_i \geq \rho_1$, $\hat{\rho}_{i,\text{mue}}$ is usually set to be 1. Even without this truncation rule, the limiting distribution of $\hat{\rho}_{\text{mg,mue}}$ is no longer normal since the distribution of $\hat{\rho}_{i,\text{mue}}$ is not a normal any more when $\hat{\rho}_i \geq \rho_1$. Due to this problem, the variance of $\hat{\rho}_{\text{mg,mue}}$ is usually smaller than the variance of $\hat{\rho}_{\text{mg}}$, which leads to the size distortion.



Figure 1: The size of the poolability test in the AR(1) model

To see this, we perform a Monte Carlo simulation with T = 50. The exact bias function is simulated from the range of $\rho \in [-1, 1.3]$ under the assumption that $u_{it} \sim iid\mathcal{N}(0, 1)$ and $a_i = 0.^2$ The truncation rate – the case of $\hat{\rho}_i \geq \rho_1$ – is recorded. Figure 1 reports the results with the simulation size of 5,000. Evidently, either when $\rho = 0.1$ or $\rho = 0.5$, the poolability test works very well regardless of the size of n because the truncation rate is zero as we discussed before. Meanwhile when $\rho = 0.8$, the truncation rate increases to 1.9%. Even though it is a small fraction, as $n \to \infty$, the size distortion accumulates so that the false rejection rates is increasing over the n/T ratio. Of

²The simulation size is one million.

course, the truncation rate is also depending on T. As T decreases, the truncation rate is increasing as well. Hence with a smaller value of T, the size distortion of the poolability test becomes much more serious.

It is important noting that the exact bias function for any T is not available except for the AR(1) model. Only an approximated bias function is known. See Shaman and Stine (1988) for AR(p) models and Nichollas and Pope (1988) for general VAR models.

For the consistent homogeneity test, the direct homogeneity test is rather desirable. Pesaran and Yamagata (2008) propose a Swamy (1970)'s type homogeneity test, which does not require for the correction of the inconsistency. As long as $n, T \to \infty$ with $n/T \to \kappa$ where $0 \le \kappa < \infty$, Pesaran and Yamagata's test becomes asymptotically valid.

4 Monte Carlo Simulation and Empirical Example

In this section, we report the finite sample performance of the proposed test and provide an empirical example.

4.1 Monte Carlo Studies

We consider the following data generating process (DGP) for the case of single regressor. We also considered the case of two regressors but we do not report this case as the results are almost identical to those with single regressor. All results are available on the author's website.

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it},$$

where

$$\begin{aligned} x_{it} &= x_{it}^{o} + \psi w_{it}, \ x_{it}^{o} = \rho x_{it}^{o} + v_{it}^{o}, \ w_{it} = \rho w_{it-1} + v_{it} \\ e_{it} &= e_{it}^{o} + \psi w_{it}, \ e_{it}^{o} = \rho e_{it-1}^{o} + \varepsilon_{it}, \ \beta_{i} = \phi \sigma_{xi}^{2} + \epsilon_{i} \\ v_{it}^{o} &\sim iidN\left(0, \sigma_{xi}^{2}\right), \ v_{it} \sim iidN\left(0, 1\right), \ \varepsilon_{it} \sim iidN\left(0, 1\right), \ \epsilon_{i} \sim iidN\left(0, 1\right). \end{aligned}$$

We set $\phi = 0$ under the null of poolability and $\phi = 0.2$ under the alternative. We set $T \in [5, 10, 25, 50, 100, 200]$, $n \in [100, 200, 500, 1000]$ and $\rho \in [0, 0.9]$. The simulation size is 5,000. We consider the following four cases.

Case 1: No IV $\psi = 0$, $\sigma_{xi}^2 \sim \chi_1^2 + 0.1$, Case 2: No IV $\psi = 0$, $\sigma_{xi}^2 \sim U[0.05, 2.05]$, Case 3: With IV, $\psi = 0.1, 0.2, 0.3, \sigma_{xi}^2 \sim \chi_1^2 + 0.1$, Case 4: With IV, $\psi = 0.1, 0.2, 0.3, \sigma_{xi}^2 \sim \sim U[0.05, 2.05]$. The regressor is strictly exogenous in Cases 1 and 2, while it is endogenous in cases 3 and 4. The distribution of the variance of the regressor is different across cases. In cases 1 and 3, the variances are distributed identically and independently $\chi_1^2 + 0.1$. We add a non-zero mean of 0.1. Without this addition, the variance of the regressor may become zero meaning that the MG estimator is not well defined. In cases 2 and 4, the variance of the regressor is uniformly distributed so that the MG estimator performs relatively better than in the other cases.

Table1 shows the size of the poolability test. The nominal size is 5%. Regardless of the values of ρ , n, and T, the sizes of the test are very good. Even with small T, the tests are only slightly under-sized in cases 1 and 3. However as n increases, the size distortion rapidly dissipates. When the variance of the regressor is distributed evenly (cases 2 and 4), the size of the test becomes very accurate even with small n and T. From these results we can conclude that the poolability test is generally accurate.

Table 2 reports the power of the test at the 5% level. Even though the value of ϕ is the same for all cases, the inconsistent parameter φ is different. In cases 1 and 3, $\varphi = 0.1$ but in cases 2 and 4, $\varphi = 0.042$. Hence the power of the test in cases 1 and 3 is relatively larger than in cases 2 and 4. Interestingly, the power of the test with $\rho = 0$ (no serial dependence) is generally lower than when $\rho = 0.9$ (high serial dependence). However, this finding could be anticipated as the same serial dependent structure is imposed on the regressor and the regression error. Nonetheless, the power of the test in all cases approach unity as n increases. As T increases, the powers of the test improve slightly, but, as we showed in Theorem 1, the powers of the test are mainly dependent on the size of n. From the results in Tables 1 and 2 we can conclude that the proposed test works well.

Table 3 shows the variances of the MG and the WG estimators under the null hypothesis. In all cases, when T is moderately large (for example T > 25), the variances of the MG estimator are smaller than those of the WG estimator. However when T is very small and IV is used, the variances of the MG estimators can be extremely large. For example, regardless of the value of n, the variances of the MG estimator are much larger than those of the WG estimators in cases 3 and 4 when T = 5. It is because the denominator terms, $\sum_{i=1}^{n} \sum_{t=1}^{T} w_{it}x_{it}$, can be very close to zero when T is small. However, as T increases, the variances of the MG estimator become smaller than those of the WG estimator.

Next, we show the effectiveness of the trimmed MG estimation. When the null of the poolability is not rejected, the WG estimator becomes consistent so that the trimmed MG estimation not needed. Only when the null is rejected and the number of time series observations is small, the trimmed MG estimation reduces the variance of the MG estimator. We use the following simple trimmed method. Order the estimates of the slope coefficients from the smallest to the largest, and denote them as $\hat{\beta}_1 < \hat{\beta}_2 < \cdots < \hat{\beta}_n$. Delete the first p% and the last (1-p)% of $\hat{\beta}_i$. Then the trimmed MG estimator can be obtained by taking the averages of the remaining β_i . Next, the trimmed WG estimator can be obtained by running the WG regression only with the cross sectional units used for the trimmed MG estimator.

Table 4 reports the size and power of the poolability test with and without trimmed method for Case 3. We set p = 0.05. See the author's website for more detailed results. The trimmed method reduces the variance of the MG estimator significantly, especially when T is small. In fact, the trimmed MG estimator shows smaller variance than the WG estimator with or without trimming. Also the size of the test becomes more accurate and the power of the test improves when T is small. However, when T is moderately large (for example $T \ge 25$), the trimmed MG estimator does little to reduce the variance and exhibits slightly worse power. Based on these results, we recommend using the trimmed MG estimator only when T is small. Note that when there is more than one regressor, the trimming method can be applied to each regressor individually. For example, consider the case where k = 2. The first trimmed MG estimator for the first regressor can be obtained only with the cross sectional units used in the first trimmed MG estimator. The second trimmed MG and WG estimators are also obtained by discarding the bottom and top p% of $\hat{\beta}_{2i}$. Note that the trimmed MG estimator can be thought as a weighted MG estimator. Finding the optimal weight function is desirable. We will consider this interesting issue in the other work.

4.2 Empirical Example: Food Consumption

This section provides an empirical example of the effects of income elasticity on food consumption. We use PSID data from 1968 to 1972, where the total number of households is 3,577. However, after eliminating missing or zero observations, the total number of households is reduced to 2,952. We run the following simple food consumption expenditure regression.

$$\ln C_{it} = a_i + \lambda_t + \beta_i \ln Y_{it} + e_{it},$$

where C_{it} and Y_{it} are food expenditure and household income, respectively. As we discussed earlier, if the marginal propensity to consume (MPC) for food consumption is dependent on income fluctuations, then the regression coefficient β_i is dependent on the variance of $\ln Y_{it}$. To see this, we re-order $\ln C_{it}$ and $\ln Y_{it}$ based on the sample variances of $\ln Y_{it}$, and estimate $\hat{\beta}_{wg}$ and $\hat{\beta}_{mg}$ with the first k subsamples. We consider the range of $k \in [2,000,2,952]$. Figure 2 displays these recursive WG and MG estimates. When k = 2,000, the panel includes only the first 2,000 samples based on the ranking of the sample variances of $\ln Y_{it}$. Evidently, as we add more volatile series, the point estimate decreases. If β_i is not correlated with the variance of the regressor, these two estimators should be fluctuated over k.



Figure 2: Recursive Within Group and Mean Group Estimators

Table 5 reports the regression results with and without the trimming method. The ordinary MG estimate for the overall MPC becomes 0.1 whereas the WG estimate of 0.07 is slightly smaller than the MG estimate. The S-statistic without trimming is around 1.65 so that the null of the poolability is rejected around at the 10% level. In addition, the trimmed MG and WG estimates show little change relative to their non-trimmed counterparts, but the S-statistic increases to 4.14 so that the null hypothesis is strongly rejected. This evidence suggests that food consumption is dependent on income fluctuations, and the slope coefficients are heterogeneous.

Table 5: Estimates of MPC of the Food Consumption

	$\hat{\beta}_{ m mg}$	$\hat{eta}_{ m wg}$	S
Without Trimmed	0.1032	0.0675	1.652
With Trimmed at 5%	0.1096	0.0675	4.135

5 Concluding Remarks

Pooling cross sectional and time series information is harmful when the slope coefficients are heterogeneous and are correlated with the variances of the regressors. This paper proposes a simple test to examine whether or not pooling is appropriate. When the null hypothesis of the poolability is rejected, the MG estimator should be used. The suggested test can be used with IV estimators when the regressors are not exogenous. By means of Monte Carlo simulations, we show that the proposed test works well in the finite sample. However, when T is small, the MG estimator becomes inefficient: The variance of the MG estimator is usually much larger than that of the WG estimator. To overcome this efficiency loss, we suggest the trimmed MG estimator when T is small.

References

- [1] Bai, J., 2009, Panel data models with interactive fixed effects, *Econometrica*, 77, 1229-1279.
- [2] Baltagi, B.H. and J.M. Griffin, 1997, Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline, *Journal of Econometrics* 77, 303–327.
- [3] Baltagi, B.H., J.M. Griffin, and W. Xiong, 2000, To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand, *Review of Economics and Statistics* 82, 117–126.
- [4] Baltagi, B.H., G. Bresson, and A. Pirotte, 2002, Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: Some empirical evidence from US electricity and natural-gas consumption, *Economics Letters* 76, 375–382.
- [5] Baltagi, B.H., G. Bresson, and A. Pirotte, 2004, Tobin q: Forecast performance for hierarchical Bayes, heterogeneous and homogeneous panel data estimators, *Empirical Economics* 29, 107– 113.
- [6] Baltagi, B.H., G. Bresson, and A. Pirotte, 2008, To pool or not to pool?, The Econometrics of Panel Data. Springer Berlin Heidelberg, 517-546.
- [7] Caballero, R. J., 1990, Consumption puzzles and precautionary savings, *Journal of Monetary Economics*, 25(1), 113-136.
- [8] Chudik, A. and M.H. Pesaran., 2015, Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors, forthcoming in *Journal of Econometrics*.
- [9] Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G., & Chambliss, H. O., 2005, Exercise treatment for depression: efficacy and dose response. *American Journal of Preventive Medicine*, 28(1), 1-8.
- [10] Greenway-McGrevy, R., C. Han and D. Sul., 2012, Asymptotic distribution of factor augmented estimators for panel regression. *Journal of Econometrics*, 169, 48–53.
- [11] Guiso, L., T. Jappelli and D. Terlizzese. 1992, Earnings uncertainty and precautionary saving, Journal of Monetary Economics, 30(2), 307-337.
- [12] Hasuman, J., 1978, Specification tests in econometrics, *Econometrica*, 46, 1251–1271.

- [13] Haque, N.U., M.H. Pesaran, and S. Sharma. 2000, Neglected heterogeneity and dynamics in cross-country savings regressions, Panel Data Econometrics: Future Directions, Edited by J. Krishmakumar and E. Ronchetti, Elsevier, 53-82.
- [14] Hernandez-Hernandez, R., Armas, D. H. M., Armas-Padilla, M. C., Carvajal, A. R., & Guerrero-Pajuelo, J. (1996). The effects of missing a dose of enalapril versus amlodipine on ambulatory blood pressure. *Blood Pressure Monitoring*, 1(2), 121-126.
- [15] Kendall, M. G., 1954, Note on the bias in the estimation of autocorrelation, *Biometrika*, 41, 403-404.
- [16] Kimball, M.S. 1990, Precautionary saving in the small and in the large, *Econometrica*, 58(1), 53–73.
- [17] Maddala, G.S., R.P. Trost, H. Li, and F. Joutz, 1997, Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators, *Journal of Business* and Economic Statistics 15, 90–100.
- [18] Nichollas, D.F. and A.L. Pope, 1988, Bias in the estimation of multivariate autoregressions, Australian Journal of Statistics, 30, 296–309
- [19] Pesaran, M.H., 2006, Estimation and inference in large heterogenous panels with a multi factor error structure, *Econometrica*, 74, 967–1012.
- [20] Pesaran, M.H. and R. Smith, 1995, Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics* 68, 79–113.
- [21] Pesaran, M.H. and R. Smith, and K.S. Im, 1996. Dynamic linear models for heterogenous panels. In Matyas, L., Sevestre, P. (Eds.), The Econometrics of Panel Data: A Handbook of the Theory with Applications, second revised edition. Kluwer Academic Publishers, Dordrecht, pp. 145-195.
- [22] Pesaran, M.H., Y. Shin, and R. Smith, 1999, Pooled mean group estimation of dynamic heterogeneous panels, *Journal of the American Statistical Association* 94, 621–634.
- [23] Pesaran, M.H. and T. Yamagata. 2008, Testing slope homogeneity in large panels, Journal of Econometrics 142, 50–93.
- [24] Phillips, P.C.B. and D. Sul. 2003, Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal* 6, 217-259.

- [25] Shaman, P. and R.A. Stine. 1988, The bias of autoregressive coefficient estimators. Journal of the American Statistical Association, 83, 842-848.
- [26] Song, M.K. 2013, Asymptotic theory for dynamic heterogeneous panels with cross-sectional dependence and its applications. mimeo, Columbia University.
- [27] Swamy, P.A.V.B. 1970. Efficient inference in a random coefficient regression model. *Econometrica*, 38 (2). 311–323.
- [28] Tanaka, K. 1983, Asymptotic expansion associated with the AR(1) model with unknown mean, *Econometrica*, 51, 1221-1232.
- [29] Wooldridge, J.M., 2005, Fixed effects and related estimators for correlated random-coefficient and treatment effect panel data models, *Review of Economics and Statistics* 87, 385-390.

6 Technical Appendix

Here we rewrite the definition used in the main text. $\boldsymbol{\xi}_{iT} = \sum_{t=1}^{T} \tilde{\mathbf{x}}'_{it} \tilde{\mathbf{x}}_{it}, \mathbf{Q}_{xT} = \lim_{n \to \infty} \mathbf{E} \frac{1}{nT} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT},$ and $\mathbf{z}_{iT} = \left(\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \left(\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\beta}}_{i}\right)$. Under the null, let $\boldsymbol{\beta}_{i} = \boldsymbol{\beta} + \boldsymbol{\epsilon}_{i}$. And denote $\Sigma_{\boldsymbol{\epsilon}}^{2} = \mathbf{I}_{k} \otimes \sigma_{\boldsymbol{\epsilon}}^{2}$ where $\sigma_{\boldsymbol{\epsilon}}^{2} = \left(\sigma_{\boldsymbol{\epsilon},1}^{2}, ..., \sigma_{\boldsymbol{\epsilon},k}^{2}\right)'$. Define $\boldsymbol{\Xi}_{i} = \boldsymbol{\xi}_{i} + \boldsymbol{\xi}_{i} = \mathbf{I}_{i} \sum_{i=1}^{n} \mathbf{\xi}_{i} = \mathbf{I}_{i} \sum_{i=1}^{n} \mathbf{I}_{i} = \mathbf{I}_{i} \sum_{i=1}^{n} \mathbf{I}_{i} = \mathbf{I}_{i}$

$$\boldsymbol{\Xi}_{T} = \mathrm{E}\frac{1}{nT} \sum_{i=1} \left(\boldsymbol{\xi}_{iT} - \mathbf{Q}_{xT} \right) \left(\boldsymbol{\beta}_{i} - \boldsymbol{\beta} \right) \left(\boldsymbol{\beta}_{i} - \boldsymbol{\beta} \right)' \left(\boldsymbol{\xi}_{iT} - \mathbf{Q}_{xT} \right)'.$$

Note that the covariance matrix of Ξ_T is a finite positive definite matrix.

Rewrite (12) as

$$\begin{split} \sqrt{n} \left(\hat{\boldsymbol{\beta}}_{wg} - \boldsymbol{\beta} \right) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\boldsymbol{\beta}_{i} - \boldsymbol{\beta} \right) + \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right)^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left[\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right] \left(\boldsymbol{\beta}_{i} - \boldsymbol{\beta} \right) \right) \right\} \\ &+ \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right). \end{split}$$

Meanwhile the OLS estimator is given by

$$\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i = \boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^T \tilde{\mathbf{x}}'_{it} \tilde{e}_{it} \right),$$

and the mean group estimator is defined as

$$\hat{\boldsymbol{\beta}}_{mg} = \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\beta}}_{i} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\beta}_{i} + \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) \right].$$

Proof of Lemma 1 Note that

$$\begin{aligned} \boldsymbol{\beta}_{i} - \frac{1}{n} \sum^{n} \boldsymbol{\beta}_{i} &= \left(\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n} \sum^{n} \hat{\boldsymbol{\beta}}_{i} \right) - \left(\hat{\boldsymbol{\beta}}_{i} - \boldsymbol{\beta}_{i} \right) + \frac{1}{n} \sum^{n} \left(\hat{\boldsymbol{\beta}}_{i} - \boldsymbol{\beta}_{i} \right) \\ &= \left(\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n} \sum^{n} \hat{\boldsymbol{\beta}}_{i} \right) - \boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) + \frac{1}{n} \sum^{n} \left(\hat{\boldsymbol{\beta}}_{i} - \boldsymbol{\beta}_{i} \right), \end{aligned}$$

so that

$$\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n} \sum^{n} \hat{\boldsymbol{\beta}}_{i} = \left(\boldsymbol{\beta}_{i} - \frac{1}{n} \sum^{n} \boldsymbol{\beta}_{i}\right) + \boldsymbol{\xi}_{iT}^{-1} \left(\sum^{T}_{t=1} \tilde{\mathbf{x}}_{it}' e_{it}\right) - \frac{1}{n} \sum^{n} \left(\hat{\boldsymbol{\beta}}_{i} - \boldsymbol{\beta}_{i}\right).$$

Under the null hypothesis, note that $E_n^{\frac{1}{n}} \sum \mathbf{z}_{iT} = 0$ since $\boldsymbol{\xi}_{iT}$ is independent from $\hat{\boldsymbol{\beta}}_i$. Hence we consider only \mathbf{z}_{iT} . Next observe this. For a large n, we have

$$\hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left\{ \sum_{i=1}^{n} \left(\left[\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT} \right] (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right) \right\} \\ + \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) - \frac{1}{n} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} \right) \right].$$

$$\left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \left(\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\beta}}_{i}\right)$$

$$= \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \left[\left(\boldsymbol{\beta}_{i} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\beta}_{i}\right) + \boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{\boldsymbol{e}}_{it}\right) - \frac{1}{n} \sum_{i=1}^{n} \left(\hat{\boldsymbol{\beta}}_{i} - \boldsymbol{\beta}_{i}\right)\right]$$

$$= A + B + C,$$

where

$$A = \left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) \left(\beta_{i} - \frac{1}{n} \sum_{i=1}^{n} \beta_{i}\right),$$

$$B = \left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left[\left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) \xi_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right)\right],$$

$$C = -\left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) \frac{1}{n} \sum_{i=1}^{n} \left[\xi_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right)\right].$$

The ${\cal A}$ term is given by

$$A = \left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) \left(\beta_{i} - \frac{1}{n} \sum_{i=1}^{n} \beta_{i}\right)$$

$$= \left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) \left(\beta_{i} - \beta - \frac{1}{n} \sum_{i=1}^{n} \epsilon_{i}\right)$$

$$= \left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) (\beta_{i} - \beta)$$

$$- \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \xi_{iT}\right)^{-1} \left[\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right)\right] \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_{i}\right) \frac{1}{n\sqrt{T}}$$

$$= \left(\sum_{i=1}^{n} \xi_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\xi_{iT} - \frac{1}{n} \sum_{i=1}^{n} \xi_{iT}\right) (\beta_{i} - \beta) + O_{p} (n^{-1})$$

for a fixed T but a large n.

The B term can be rewritten as

$$B = \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left[\left(\mathbf{I} - \left(\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \boldsymbol{\xi}_{iT}^{-1} \right) \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right) \right]$$

$$= \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} - \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right)$$

$$= \left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right).$$

Meanwhile the C term can be expressed as

$$C = -\left(\frac{1}{nT}\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{1}{T}\boldsymbol{\xi}_{iT} - \frac{1}{nT}\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\boldsymbol{\xi}_{iT}^{-1} \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{e}_{it}\right)\right] \frac{1}{n}$$

= $O_p(n^{-1}).$

Hence

$$\left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \sum_{i=1}^{n} \left(\boldsymbol{\xi}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right) \left(\hat{\boldsymbol{\beta}}_{i} - \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\beta}}_{i}\right)$$
$$= \hat{\boldsymbol{\beta}}_{wg} - \hat{\boldsymbol{\beta}}_{mg} + O_{p} \left(n^{-1}\right).$$

Q.E.D. \Box

Proof of Theorem 1 Note that we defined $\eta_{iT} = \frac{1}{T} \left(\mathbf{z}_{iT} - \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_{iT} \right)$ and $\Omega_{\eta}^{2} = \mathbb{E} \frac{1}{n} \left(\sum_{i=1}^{n} \eta_{iT} \right) \left(\sum_{i=1}^{n} \eta_{iT} \right)'$. Since both $\boldsymbol{\xi}_{iT}$ and $\hat{\boldsymbol{\beta}}_{i}$ have the finite variances, it is easy to show that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_{iT}\to^{d}\mathcal{N}\left(0,\Omega_{\eta}^{2}\right),$$

so that

$$\left(\sum_{i=1}^{n} \boldsymbol{\xi}_{iT}\right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_{iT}\right) \to^{d} \mathcal{N}\left(0, \Omega_{\Delta\beta}^{2}\right).$$

Q.E.D. \Box

Proof of Theorem 2 Under the alternative, we have

$$\operatorname{plim}_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{1}{T} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{\mathbf{x}}_{it} - \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{\mathbf{x}}_{it} \right) (\boldsymbol{\beta}_{i} - \boldsymbol{\beta}) \right] = \Sigma_{\boldsymbol{\xi}\boldsymbol{\beta}} \neq 0.$$

Define

$$\varphi_{nT} = \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}'_{it} \tilde{\mathbf{x}}_{it}\right)^{-1} \left(\sum_{i=1}^{n} \left[\sum_{t=1}^{T} \tilde{\mathbf{x}}'_{it} \tilde{\mathbf{x}}_{it}\right] (\boldsymbol{\beta}_{i} - \boldsymbol{\beta})\right).$$

Then its probability limit becomes

$$\boldsymbol{\varphi}_T = \operatorname{plim}_{n \to \infty} \boldsymbol{\varphi}_{nT} = \mathbf{Q}_{xT}^{-1} \Sigma_{\boldsymbol{\xi} \boldsymbol{\beta}}.$$

Meanwhile the limiting distribution of the MG estimator does not change under the alternative. Hence it is straightforward to show that

$$\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\mathrm{wg}}-\hat{\boldsymbol{\beta}}_{\mathrm{mg}}-\boldsymbol{\varphi}_{T}\right)\rightarrow^{d}\mathcal{N}\left(0,\Omega_{\Delta\boldsymbol{\beta}}^{2}\right).$$

Q.E.D. \Box

=				ρ :	=0		$\rho = 0.9$			
	n	T	Case1	Case2	Case3	Case4	Case1	Case2	Case3	Case4
-	100	5	0.035	0.047	0.037	0.048	0.032	0.048	0.035	0.045
	100	10	0.040	0.051	0.035	0.044	0.036	0.045	0.036	0.048
	100	25	0.041	0.052	0.047	0.053	0.037	0.052	0.040	0.050
	100	50	0.043	0.047	0.041	0.051	0.041	0.044	0.043	0.047
	100	100	0.043	0.052	0.042	0.052	0.039	0.053	0.040	0.046
	100	200	0.044	0.050	0.040	0.048	0.037	0.049	0.039	0.055
	200	5	0.041	0.050	0.041	0.048	0.036	0.045	0.037	0.042
	200	10	0.048	0.050	0.045	0.057	0.044	0.050	0.043	0.046
	200	25	0.043	0.044	0.041	0.049	0.043	0.051	0.040	0.043
	200	50	0.045	0.049	0.042	0.046	0.037	0.054	0.040	0.050
	200	100	0.047	0.056	0.045	0.051	0.044	0.050	0.046	0.052
	200	200	0.045	0.053	0.049	0.055	0.040	0.049	0.049	0.053
	500	5	0.047	0.053	0.042	0.051	0.044	0.051	0.041	0.047
	500	10	0.047	0.049	0.051	0.050	0.044	0.050	0.047	0.048
	500	25	0.045	0.050	0.042	0.052	0.046	0.049	0.047	0.051
	500	50	0.047	0.056	0.045	0.052	0.049	0.048	0.047	0.051
	500	100	0.044	0.047	0.042	0.047	0.046	0.044	0.046	0.046
	500	200	0.042	0.048	0.046	0.051	0.047	0.049	0.047	0.051
	1000	5	0.052	0.047	0.050	0.049	0.045	0.042	0.041	0.046
	1000	10	0.051	0.052	0.047	0.045	0.045	0.050	0.049	0.052
	1000	25	0.049	0.048	0.048	0.050	0.048	0.045	0.050	0.046
	1000	50	0.054	0.046	0.045	0.049	0.047	0.047	0.048	0.053
	1000	100	0.051	0.052	0.049	0.046	0.050	0.048	0.053	0.049
-	1000	200	0.049	0.048	0.043	0.049	0.050	0.049	0.041	0.054

Table 1: The Size of the Poolability Test (Nominal size = 5%)

			ho =	= 0		ho=0.9			
n	T	Case1	Case2	Case3	Case4	Case1	Case2	Case3	Case4
100	5	0.238	0.075	0.243	0.080	0.179	0.061	0.171	0.069
100	10	0.415	0.114	0.417	0.112	0.247	0.077	0.260	0.077
100	25	0.520	0.148	0.536	0.162	0.329	0.089	0.342	0.090
100	50	0.560	0.176	0.557	0.178	0.384	0.102	0.393	0.101
100	100	0.578	0.188	0.574	0.182	0.459	0.127	0.465	0.123
100	200	0.582	0.190	0.593	0.197	0.512	0.145	0.520	0.157
200	5	0.548	0.105	0.527	0.118	0.405	0.082	0.373	0.087
200	10	0.753	0.176	0.777	0.184	0.525	0.105	0.565	0.111
200	25	0.875	0.250	0.869	0.273	0.656	0.129	0.653	0.139
200	50	0.898	0.300	0.902	0.294	0.743	0.168	0.748	0.162
200	100	0.905	0.333	0.907	0.332	0.815	0.202	0.818	0.207
200	200	0.916	0.343	0.911	0.364	0.866	0.258	0.861	0.265
500	5	0.933	0.210	0.874	0.217	0.833	0.143	0.721	0.158
500	10	0.994	0.381	0.996	0.388	0.933	0.197	0.929	0.212
500	25	0.998	0.559	0.999	0.565	0.975	0.262	0.975	0.278
500	50	1.000	0.643	1.000	0.636	0.989	0.340	0.989	0.345
500	100	1.000	0.678	1.000	0.690	0.997	0.442	0.997	0.453
500	200	1.000	0.713	0.999	0.703	0.999	0.546	0.998	0.553
1000	5	0.998	0.361	0.934	0.356	0.989	0.241	0.842	0.240
1000	10	1.000	0.642	1.000	0.663	0.997	0.347	0.992	0.378
1000	25	1.000	0.841	1.000	0.859	1.000	0.442	0.999	0.492
1000	50	1.000	0.906	1.000	0.906	1.000	0.574	1.000	0.603
1000	100	1.000	0.931	1.000	0.934	1.000	0.720	1.000	0.741
1000	200	1.000	0.945	1.000	0.944	1.000	0.834	1.000	0.837

Table 2: The Power of the Poolability Test (Nominal Size = 5%)

			Μ	[G			W	/G	
n	T	Case1	Case2	Case3	Case4	Case1	Case2	Case3	Case4
100	5	3.423	2.479	< 100	46.76	4.981	2.859	4.776	2.886
100	10	2.180	1.694	8.105	4.077	4.447	2.515	4.359	2.535
100	25	1.722	1.413	1.730	67.19	3.729	2.142	3.891	2.123
100	50	1.471	1.266	1.481	1.270	3.398	1.827	3.505	1.862
100	100	1.242	1.170	1.222	1.160	3.093	1.631	2.967	1.630
100	200	1.072	1.064	1.201	1.050	2.758	1.431	2.912	1.447
200	5	1.647	1.212	< 100	17.04	2.430	1.375	2.546	1.420
200	10	1.048	0.853	10.44	1.544	2.266	1.231	2.248	1.200
200	25	0.858	0.706	0.915	0.719	2.025	1.090	2.002	1.057
200	50	0.738	0.615	0.746	0.660	1.740	0.919	1.753	0.930
200	100	0.642	0.585	0.637	0.572	1.604	0.804	1.537	0.811
200	200	0.566	0.538	0.585	0.554	1.419	0.746	1.486	0.729
500	5	0.662	0.487	< 100	45.80	1.014	0.565	1.042	0.577
500	10	0.428	0.338	0.853	1.185	0.928	0.487	0.953	0.488
500	25	0.351	0.281	0.363	0.924	0.787	0.426	0.815	0.430
500	50	0.302	0.256	0.295	0.265	0.715	0.370	0.711	0.385
500	100	0.256	0.236	0.252	0.234	0.608	0.321	0.612	0.324
500	200	0.223	0.219	0.240	0.220	0.559	0.298	0.581	0.299
1000	5	0.337	0.240	< 100	52.28	0.513	0.279	0.542	0.279
1000	10	0.213	0.169	16.24	5.338	0.470	0.251	0.456	0.244
1000	25	0.168	0.143	0.184	0.143	0.412	0.216	0.421	0.206
1000	50	0.151	0.128	0.150	0.123	0.362	0.180	0.348	0.178
1000	100	0.126	0.116	0.126	0.113	0.322	0.165	0.326	0.160
1000	200	0.112	0.107	0.118	0.113	0.300	0.148	0.287	0.149

Table 3: Variances of the MG and WG Estimators under the Null Hypothesis $(\rho=0.9)$

			Without Trimmed			With Trimmed		
			5%	$Variance \times 10^2$		5%	Varian	$ce \times 10^2$
	n	T	MG	MG	WG	MG	MG	WG
Under the Null	100	5	0.035	< 100	4.776	0.058	2.755	4.709
	100	10	0.036	8.105	4.359	0.041	2.108	4.166
	100	25	0.040	1.730	3.891	0.045	1.651	3.546
	100	50	0.043	1.481	3.505	0.043	1.466	3.143
	100	100	0.040	1.222	2.967	0.041	1.233	2.604
	100	200	0.039	1.201	2.912	0.042	1.224	2.567
	1000	5	0.041	< 100	0.542	0.054	0.263	0.531
	1000	10	0.049	16.24	0.456	0.054	0.204	0.436
	1000	25	0.050	0.184	0.421	0.054	0.173	0.384
	1000	50	0.048	0.150	0.348	0.050	0.148	0.319
	1000	100	0.053	0.126	0.326	0.050	0.128	0.274
	1000	200	0.041	0.118	0.287	0.045	0.121	0.250
Under the Alternative	100	5	0.171	< 100	6.947	0.241	2.938	5.921
	100	10	0.260	7.517	6.185	0.253	2.246	4.906
	100	25	0.342	1.857	5.809	0.288	1.788	4.159
	100	50	0.393	1.611	5.289	0.307	1.592	3.525
	100	100	0.465	1.327	4.543	0.324	1.335	2.967
	100	200	0.520	1.309	4.561	0.363	1.326	2.835
	1000	5	0.842	< 100	0.807	0.993	0.275	0.643
	1000	10	0.992	16.80	0.693	0.992	0.215	0.519
	1000	25	0.999	0.192	0.628	0.995	0.182	0.438
	1000	50	1.000	0.158	0.542	0.998	0.156	0.356
	1000	100	1.000	0.134	0.498	0.999	0.136	0.314
	1000	200	1.000	0.127	0.456	1.000	0.129	0.282

Table 4: Efficiency of Trimmed Estimation