Testing for Slope Heterogeneity Bias in Panel Data Models∗

Murillo Campello, Antonio F. Galvao, and Ted Juhl

Abstract

Standard econometric methods can overlook individual heterogeneity in empirical work, generating inconsistent parameter estimates in panel data models. We propose the use of methods that allow researchers to easily identify, quantify, and address estimation issues arising from individual slope heterogeneity. We first characterize the bias in the standard fixed effects estimator when the true econometric model allows for heterogeneous slope coefficients. We then introduce a new test to check whether the fixed effects estimation is subject to heterogeneity bias. The procedure tests the population moment conditions required for fixed effects to consistently estimate the relevant parameters in the model. We establish the limiting distribution of the test, and show that it is very simple to implement in practice. We also generalize the test to allow for cross-section dependence in the errors and a form of endogeneity. Examining firm investment models to showcase our approach, we show that heterogeneity bias-robust methods identify cash flow as a more important driver of investment than previously reported. Our study demonstrates analytically, via simulations, and empirically the importance of carefully accounting for individual specific slope heterogeneity in drawing conclusions about economic behavior.

Key Words: Testing, fixed effects, individual heterogeneity, bias, slope heterogeneity

JEL Classification: C12, C23

∗The authors are grateful to Cecilia Bustamante, Zongwu Cai, George Gao, Erasmo Giambona, Hyunseob Kim, Oliver Linton, Whitney Newey, Suyong Song, Albert Wang, Zhijie Xiao, and the participants at seminars at Chapman University, Claremont McKenna College, University of Iowa, University of Kansas, University Wisconsin-Milwaukee, 25th Midwest Econometrics Group, and NY Camp Econometrics VIII for their constructive comments and suggestions. Murillo Campello, 369 Sage Hall, Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853; Tel (607)-255-1282; (campello@cornell.edu). Antonio Galvao, Tippie College of Business, University of Iowa, W284 Pappajohn Business Building, 21 E. Market Street, Iowa City, IA 52242; Tel (319) 335-1810; (antonio-galvao@uiowa.edu). Ted Juhl, Department of Economics, University of Kansas, 415 Snow, Lawrence, KS 66045; Tel (785)-864-2849; (juhl@ku.edu).
1 Introduction

With the increasing availability of large, comprehensive databases, researchers seek to identify and understand richer nuances of microeconomic behavior. Recent computable general equilibrium models introduce behavioral heterogeneity across firms and agents when generating distributions of potential outcomes. To the extent that economic decisions are shaped by these considerations it is natural to consider a methodology that allows for information about the distribution of model inputs (e.g., input variances and covariances) to be incorporated in empirical parameter estimates. This paper shows that estimation methods that ignore relations between model input distribution and parameters can lead to biased inferences about microeconomic behavior. Understanding and addressing these problems is key to advancing empiricists’ ability to inform theory development.

Empirical work in economics often relies on methods that assume a large degree of homogeneity across individuals. One way to account for some degree of heterogeneity in panel data models is to use the ordinary least squares fixed effects (OLS-FE) estimator. When estimating individual-fixed effects models, one often imposes concomitant assumptions of heterogeneous “intercepts” and homogeneous “slope coefficients” across individuals.

This paper characterizes and addresses the issues of estimation and inference of econometric models in the presence of heterogeneity in individual policy responses (“slope coefficients”). We contribute to the literature by proposing the use of methods that allow researchers to easily identify, quantify, and address estimation issues arising from individual heterogeneity. More generally, our analysis warns researchers about imposing arbitrary homogeneity restrictions when studying complex economic behavior.

We start by describing two main results. First, we show analytically how the presence of individual slope heterogeneity may bias the policy estimates obtained under the OLS-FE framework. Second, we discuss alternative methods that account for slope heterogeneity and
produce consistent estimates for the parameters of interest. To do so, we analyze a simple panel data model from which we recover the parameters of interest for each individual. This allows us to study individual slope heterogeneity in detail. To summarize the coefficients of interest, we employ a simple statistic: the mean group (MG) estimator. In its simplest form, the MG estimator is the average of the response (slope) coefficients of each of the individuals used to fit an empirical model (see Pesaran (2006)). This estimator is in the class of minimum distance estimators and has several attractive features.\(^1\) Our analysis shows that the MG method has estimation and inference properties that generally dominate those of the OLS-FE estimator when firms respond differently to innovations to a model driver. At the same time, we provide researchers with diagnostic tests to determine whether the use of the OLS-FE approach is warranted in their particular applications.

Importantly, the main contribution of this paper is to develop a new test to identify whether the presence of slope heterogeneity in the data will cause bias in OLS-FE estimates. The OLS-FE is asymptotically biased when the heterogeneous coefficients are correlated with the variance of the regressors. Accordingly, we propose a test for the null hypothesis that there is no correlation between variances of the data and the heterogeneous parameters. Hence, our procedure tests the population moment conditions required for fixed effects to consistently estimate the relevant parameters in the model. We describe the associated test statistic and derive its limiting distribution. This test is particularly useful for applied researchers in that it follows a standard chi-square distribution, for which critical values are tabulated and widely available.

We also generalize the test to allow for cross-sectional dependence in the errors and in the regressors, and also a form of endogeneity. We show that these extensions can be easily

\(^1\)First, the MG estimator is easy to implement and its computation is no more difficult than that of the OLS-FE. Second, the economic interpretation of the coefficients returned by the MG is similar to that of least squares-based methods (including the OLS-FE). Third, inference procedures using the MG are textbook standard. Fourth, the MG estimator accounts not only for individual-specific slope coefficients, but also for idiosyncratic, individual-fixed intercepts. Finally, the MG estimator can account for time-fixed effects as well.
accommodated in the testing procedure by estimating the general appropriated average of the response (slope) coefficients of each of the individuals and its corresponding variance-covariance matrix. In addition, the liming distribution for such tests remains the same as in the simple case. These are important extensions which broaden the applicability of the proposed tests in practice, since cross-sectional dependence and endogeneity are common concerns in empirical research.

There are alternative tests available in the literature for the hypothesis of slope homogeneity across individuals, including, among others, Pesaran, Smith, and Im (1996), Phillips and Sul (2003), Pesaran and Yamagata (2008), Blomquist and Westerlund (2013), and Su and Chen (2013). Pesaran, Smith, and Im (1996) propose an application of the Hausman (1978) testing procedure where the OLS-FE estimator is compared with the MG estimator. Phillips and Sul (2003) suggest a “Hausman-type” test in the context of stationary first-order autoregressive panel models, where the cross-section, \( n \), is fixed as the time-series, \( T \), goes to infinity. Hsiao (2003) describes a variation of the Breusch and Pagan (1979) test for the slope homogeneity, which is valid when both \( n \) and \( T \) dimensions tend to infinity. More recently, Pesaran and Yamagata (2008) propose a dispersion type test based on Swamy (1970) type test. They standardize the Swamy test so that this dispersion test can be applied when both \( n \) and \( T \) are large. Blomquist and Westerlund (2013) propose a test that is robust to general forms of cross-sectional dependence and serial correlation. Su and Chen (2013) develop a test for slope homogeneity in large-dimensional panel models with interactive fixed effects.

Compared to the existing procedures for testing slope homogeneity, our approach has several distinctive advantages. First, the proposed methods test the null hypothesis of lack of correlation between variances of the data and the heterogeneous parameters. This is important because it is possible that the individual heterogeneity is such that there is no bias under OLS-FE, and thus existing test procedures would not be able to detect departures from the null hypothesis. Nevertheless, our proposed tests are able to detect such departures since
the tests are based on the correlation between variances of the data and the heterogeneous parameters. Second, in the simplest case of the model, the tests do not require the time-series to diverge to infinity. Third, the test procedures can be extended to accommodate correlation between errors from different cross-sectional units, as well as a form of endogeneity. This makes the testing procedure beneficial for many empirical applications.

Monte Carlo simulations assess the finite sample properties of the proposed methods. The experiments suggest that the proposed approaches perform very well in finite samples. The bias of fixed effects estimation can be made arbitrarily large by increasing the magnitude of the covariance between the regression slope and the data variance. Critically, the MG estimator we employ is unaffected by the slope heterogeneity bias. Finally, the new tests possess good finite sample performance and have correct empirical size and power to detect precisely the cases where OLS-FE is biased.

To illustrate the performance of the proposed methods in real-world data, we study empirically the investment model of Fazzari et al. (1988), where a firm’s investment is regressed on a proxy for investment demand (Tobin’s $Q$) and cash flow. Using annual COMPUSTAT data covering a four-decade window, we study slope heterogeneity in investment models contrasting estimates from alternative methodologies. The coefficients returned for $Q$ under the methods we consider are fairly comparable in magnitude and significance. Results are very different, however, for the response of investment to cash flow. Indeed, our tests identify pronounced firm response heterogeneity, and the MG-estimated cash flow coefficient is substantially larger than those returned by the other methods. In concrete terms, the estimated coefficient on cash flow increases from 0.057 (0.043) under the OLS (OLS-FE) estimation to 0.291 under the MG estimation; a 6–7-fold increase on the same data. For long panels, the cash flow coefficient is 15 times larger under the MG method. Our results imply that cash flow is a much more relevant driver of investment than previous studies have suggested.

The remainder of the paper is organized as follows. Section 2 describes the biases that
affect OLS-FE in the presence of individual slope heterogeneity, and discusses the MG estimator. Section 3 presents the proposed test for slope heterogeneity and its generalization. Monte Carlo experiments are discussed in Section 4. In Section 5, we estimate a corporate investment model and compare results from different methods. Section 6 concludes.

2 Understanding The Individual Heterogeneity Bias

This section presents the panel data model and characterizes the bias in the OLS-FE estimator when the true econometric model allows for heterogeneous slope coefficients. To guide researchers in their choice of methodology, we also show conditions under which the OLS-FE returns consistent estimates. These conditions are limiting and imply potentially important compromises for researchers’ inferences. We also discuss the mean group estimator.

2.1 Biases In The Individual-Fixed Effects Estimator

When modeling economic behavior, empiricists commonly estimate regression models that allow individuals or firms to have different individual-fixed effects (different “intercepts”), yet impose equal “slope” coefficients across units. As we show next, an incorrect homogenous slope restriction leads to a bias in the policy estimates one obtains under the ordinary least squares fixed effects (OLS-FE) framework.

Assume the following baseline model for the data generating process

\[ y_{it} = \alpha_i + x_{it}^\top \beta_i + u_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \]

where \( y_{it} \) is the dependent variable, \( x_{it} \) is a \( k \times 1 \) vector of exogenous regressors, and \( u_{it} \) is the conditionally mean zero innovation term. The term \( \alpha_i \) captures individual-specific fixed effects, while the slope coefficients \( \beta_i \) may vary across individuals.
In the presence of slope heterogeneity, one would like to recover the slope coefficient of each individual. In most empirical test settings, however, the objective is to report a summary statistic for policy purposes. In that context, a reasonable quantity to estimate is the average individual slope. As such, the empiricist might try to estimate the parameters $E(\beta_i)$, the vector representing the average effect from marginal changes in $x_{it}$.

If all individuals are identical, then the OLS-FE method provides an easy way to estimate the parameters of interest. It is rarely the case, however, that one can \textit{a priori} justify the assumption of homogeneity in individuals’ responses to economic stimuli. Indeed, as discussed above, theoretical modeling and casual observation often suggest otherwise. We now characterize the problems of using the OLS-FE method to estimate the population quantities $E(\beta_i) = \beta$ (the average individual slope coefficients) in the presence of policy heterogeneity.

Consider estimating the following model via OLS-FE

$$y_{it} = \sum_{i=1}^{N} \alpha_i D_i + x_{it}^\top \beta + u_{it}, \quad i = 1, \ldots, N, \; t = 1, \ldots, T,$$

where the slope parameters are forced to be equal across individuals, even though the data generating process is given by equation (1). The variable $D_i$ is one for individual (or firm) $i$ and 0 otherwise. The OLS-FE estimator includes $N$ individual dummies, $D_i$, as a way to account for individual-fixed idiosyncrasy. That is, the only type of heterogeneity allowed in this model concerns intercept terms.

The balanced panel model in matrix form is given as

$$y_i = \alpha_i \iota_T + X_i \beta_i + u_i,$$

for $i = 1, \ldots, N$ where $y_i$ is a $T \times 1$ vector, $y_i = (y_{i1}, \ldots, y_{iT})$, $\iota_T$ is a $T \times 1$ vector of ones, $X_i$
is a $T \times k$ matrix with rows $x_i^T$, $X_i = (x_{i1}, ..., x_{iT})$, and $u_i$ is a vector with the errors. The standard fixed effects estimator is calculated based on the implicit assumption that $\beta_i$ are the same for each $i$. The formula is given as

$$\hat{\beta}_{FE} = \left( \sum_{i=1}^{N} X_i^T M_i X_i \right)^{-1} \sum_{i=1}^{N} X_i^T M_i y_i,$$

where $M_i = I_T - \tau_T (\tau_T^T \tau_T)^{-1} \tau_T^T$, $I_T$ is an identity matrix of order $T$. $M_i$ computes the deviation-from-individual-means.

It is instructive to rewrite the fixed effects estimator in two different ways. First, if the corresponding inverses exist, we can write

$$\hat{\beta}_{FE} = \sum_{i=1}^{N} \left( \sum_{i=1}^{N} X_i^T M_i X_i \right)^{-1} (X_i^T M_i X_i)(X_i^T M_i X_i)^{-1} X_i^T M_i y_i$$

$$= \sum_{i=1}^{N} W_i \hat{\beta}_i,$$

where

$$W_i = \left( \sum_{i=1}^{N} X_i^T M_i X_i \right)^{-1} (X_i^T M_i X_i), \quad \text{and} \quad \hat{\beta}_i = (X_i^T M_i X_i)^{-1} X_i^T M_i y_i.$$  

Notice that $\hat{\beta}_i$ is the usual OLS estimator for each $i$.

This first representation in equation (3) shows that fixed effects estimators are a weighted average of the individual OLS estimators for each cross-sectional unit. Moreover, the weights are larger for cross-sectional units with more variation in the $X_i$. Such a weighting scheme has the potential to improve efficiency. If there is no parameter heterogeneity and no heteroskedasticity, the variance of each $\hat{\beta}_i$ is $\sigma^2 (X_i^T M_i X_i)^{-1}$. Therefore, the observations with more variation in $X_i$ are more precisely estimated, and should be given more weight. However, if there is slope heterogeneity so that $\beta_i$ are different, there is a potential for a severe
problem. In particular, if the weights \((W_i)\) and \(\beta_i\) are correlated, we may not consistently estimate any parameter of interest.

The second representation we present illustrates this problem more clearly. We wish to estimate \(E(\beta_i)\), the “average” slope. Then,

\[
\hat{\beta}_{FE} - E(\beta_i) = \left( \frac{1}{N} \sum_{i=1}^{N} X_i^T M_i X_i \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} X_i^T M_i [X_i \beta_i - X_i E(\beta_i) + u_i].
\]

From the representation in equation (4), the term that potentially renders the fixed effects estimator inconsistent is

\[
\frac{1}{N} \sum_{i=1}^{N} X_i^T M_i X_i [\beta_i - E(\beta_i)].
\]

Intuitively, if the slope parameters \(\beta_i\) are correlated with the variation in \(X_i\) in the cross-section, fixed effects estimation will not be consistent.\(^3\)

There are conditions where fixed effects may be appropriate. Indeed, the literature proposes conditions where OLS-FE estimators are robust to heterogeneity in the slope parameters (references include Wooldridge (2003, 2005, 2010)). In the context of our model, the conditions amount to

\[
E(\beta_i | M_i X_i) = E(\beta_i),
\]

which would imply that \(E(X_i^T M_i X_i [\beta_i - E(\beta_i)]) = 0\), and then (5) would converge to zero under suitable regularity conditions. In other words, the assumption such as (6) implies fixed effects estimation is consistent.

Nevertheless, in general, the OLS-FE can be inconsistent. Consider the representation of the bias term. A simple application of the law of large numbers suggests that the asymptotic bias is

\[
E \left( X_i^T M_i X_i \right)^{-1} E \left( X_i^T M_i X_i [\beta_i - E(\beta_i)] \right).\]

\(^3\)Pesaran and Smith (1995) show that in a dynamic panel data model, heterogeneity causes bias in any case.
To illustrate how this might arise in a simple empirical model, consider the standard linear model with two covariates,

\[ y_{it} = \alpha_i + w_{it}\theta_i + z_{it}\gamma_i + u_{it}. \]

The representation of the asymptotic bias for the simple model with two covariates is given by following equation

\[
\begin{pmatrix}
\hat{\theta}_{FE} - E(\theta_i) \\
\hat{\gamma}_{FE} - E(\gamma_i)
\end{pmatrix}
\overset{p}{\rightarrow}
\begin{pmatrix}
E(\sigma_{wi}^2) & E(\sigma_{wzi}) \\
E(\sigma_{wzi}) & E(\sigma_{zi}^2)
\end{pmatrix}
^{-1}
\begin{pmatrix}
E[\sigma_{wi}^2(\theta_i - \theta)] + E[\sigma_{wzi}(\gamma_i - \gamma)] \\
E[\sigma_{wzi}(\theta_i - \theta)] + E[\sigma_{zi}^2(\gamma_i - \gamma)]
\end{pmatrix},
\]

where \(\sigma_{wi}^2\) and \(\sigma_{zi}^2\) represent the individual-specific variance of \(w_{it}\) and \(z_{it}\), respectively, and \(\sigma_{wzi}\) is the individual-specific covariance of \(w_{it}\) and \(z_{it}\). What we see from the representation in (7) is that, in general, the covariance between the variance of the regressors and the parameters causes the bias. For example, if cross-sectional units with a high variance in \(z_{it}\) have a larger response measured by \(\gamma_i\), there will be a positive asymptotic bias in the fixed effects estimator for \(\gamma\). Moreover, these effects can disseminate to the other parameter, depending on the covariance structure. Nevertheless, equation (7) also shows that the OLS-FE can be asymptotically unbiased when there is no covariance between the variance of the regressors and the parameters. In this case, the elements in the vector in the far right-hand-side are equal to zero.

### 2.2 The Mean Group Estimator: Estimation and Inference

Our focus is the estimation of models in which response parameters may vary across individuals observed over time (panel data). If response heterogeneity is part of the structural model, one needs to decide what to report. In principle, one needs an estimator that is useful in summarizing heterogeneity in the coefficients of interest, and that at the same time is easy to compute and interpret. In this context, a measure of centrality of the distribution
of individuals’ responses is arguably a reasonable quantity to estimate.

The mean group (MG) estimator reports the mean of the regression slope parameters as a way to summarize the population.\footnote{This estimator is fully developed in Pesaran (2006) and belongs to the class of minimum distance estimators (see Newey and McFadden (1994)).} For most applications in economics and finance, a viable MG estimator will consist of the average of OLS coefficients from each individual whose data are used to fit an empirical model. This can be represented by

$$\hat{\beta}_{MG} = \frac{1}{N} \sum_{i=1}^{N} \hat{\beta}_i,$$  \hspace{1cm} (8)

where $\hat{\beta}_i$ is the OLS for each sample individual.

The MG estimator allows for both the intercept and slope coefficients to vary across individuals. This happens because, by applying OLS to each individual equation, one simultaneously estimate intercept (standard “fixed effects”) as well as slope coefficients for each individual. Critically, the MG estimator does not suffer from the heterogeneity bias that we characterize in the previous section. That the MG estimator is unbiased stems from the fact that we fit the empirical model onto each individual separately, correctly accounting for individual heterogeneity. The method combines individual estimates equally, averaging over consistent parameters.

The interpretation of the estimator in (8) is straightforward. The population parameter of interest is the mean coefficient over the sample individuals; $\beta_{MG} = E(\beta_i)$. The estimator $\hat{\beta}_{MG}$ is the sample mean of individual slope estimates. In model (1) we can average over the sensitivity of the dependent variable $y_{it}$ with respect to a covariate $w_{it}$, holding $z_{it}$ constant. From equation (1) the interpretation $\theta_i$ is

$$\frac{\partial E(y_{it}|x_{it}, \alpha_i)}{\partial x_{it}} = \beta_i.$$
Similar to standard regression analysis, the MG estimator can be interpreted as the average sensitivity of $y_{it}$ to $x_{it}$, and $E \left[ \frac{\partial E(y_{it}|x_{it},\alpha_i)}{\partial x_{it}} \right] = \hat{\theta}_{MG}$.

Inference for the MG estimator is also straightforward. The asymptotic variance for $\sqrt{N}(\hat{\beta}_{MG} - \beta)$ is estimated via

$$\hat{\Omega}_{MG} = \frac{1}{N-1} \sum_{i=1}^{N} (\hat{\beta}_i - \hat{\beta}_{MG})(\hat{\beta}_i - \hat{\beta}_{MG})^\top. \tag{9}$$

Inference is based on asymptotic normality as the time observations, $T$, and the number of individuals, $N$, increase.

## 3 Testing Slope Heterogeneity

It is important that we present a practical procedure to identify and quantify problem of slope heterogeneity. In this section, we first review a test for the presence of slope heterogeneity building on dispersion tests by Pesaran and Yamagata (2008) and Swamy (1970). Importantly, as noted in Section 2, it is possible that the individual heterogeneity is such that there is no bias under the OLS-FE. Thus, second, we propose a novel test designed to detect precisely when and how slope heterogeneity will cause bias in parameters estimated via OLS-FE. Third, we generalized the proposed test to allow for cross-sectional dependence in the errors and in the regressors, and also a form of endogeneity.

### 3.1 Identifying Slope Heterogeneity

Pesaran and Yamagata (2008) consider the regression panel model with $N$ individuals and $T$ time periods in equation (1). Let $\beta_i$ be the vector of policy parameters in the model, $k \times 1$. 
We wish to test the following null hypothesis of slope homogeneity across individuals

$$H_0 : \beta_i = \beta,$$

for some fixed vector $\beta$ for all $i$, against the alternative

$$H_1 : \beta_i \neq \beta_j \text{ for some } i, j.$$

The strategy is to estimate regression coefficients using the time-series for each individual, then compare the estimates with $\beta$. Under the null, the estimates for each individual should be close to $\beta$. Large differences across these estimates and $\beta$ indicate that the null should be rejected. Since we do not observe the true coefficient $\beta$, we replace it with $\tilde{\beta}_{WFE}$, which is a weighted version of the fixed effects estimator.

The test is defined as

$$\tilde{S} := \sum_{i=1}^{N} (\tilde{\beta}_i - \tilde{\beta}_{WFE})^\top \left( \frac{X_i^\top M_i X_i}{\tilde{\sigma}_i^2} \right)^{-1} \left( \tilde{\beta}_i - \tilde{\beta}_{WFE} \right),$$

where $\tilde{\sigma}_i^2 = \frac{(y_i - X_i \tilde{\beta}_{FE})^\top M_i (y_i - X_i \tilde{\beta}_{FE})}{(T-k-1)}$, $\tilde{\beta}_{FE}$ is the standard OLS-FE, and $\tilde{\beta}_{WFE}$ is defined as $\tilde{\beta}_{WFE} = \left( \sum_{i=1}^{N} \frac{X_i^\top M_i X_i}{\tilde{\sigma}_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{X_i^\top M_i y_i}{\tilde{\sigma}_i^2}$. One can also consider a standardized version of the test as $\tilde{\Delta} = \sqrt{N} \frac{\tilde{S} - k}{\sqrt{2k}}$. Under standard regularity conditions (see Pesaran and Yamagata (2008)):

$$\tilde{S} \xrightarrow{d} \chi^2_{(N-1)k}, \text{ as } T \to \infty;$$

$$\tilde{\Delta} \xrightarrow{d} N(0, 1), \text{ as } (N,T) \to \infty.$$

where $k$ is the number of restriction under the null hypothesis.
3.2 Diagnosing The Bias In The Fixed Effects Estimator

The Pesaran-Yamagata-Swamy (PYS) test presented above provides a way to identify the presence of individual slope heterogeneity in a model. Slope heterogeneity, however, need not cause OLS-FE estimates to be biased as shown in equation (7) (see also Wooldridge (2005, 2010)). Alternatively, the bias could be small enough so as to make the OLS-FE still desirable. We now introduce a novel test measuring the magnitude of the slope heterogeneity bias in OLS-FE estimations. In the next section, we generalize the test to accommodate more general conditions and allow for cross-sectional dependence in the errors and in the regressors, as well as a form of endogeneity.

We wish to test the hypothesis that the heterogeneity across individuals does not induce a bias in the OLS-FE estimator. Equation (7) shows that the OLS-FE estimates are asymptotically biased when the heterogeneous coefficients are related to the variance and covariance of the regressors. Thus, we construct a test based on this result. In particular, we wish to test the null hypothesis that there is no correlation between the coefficients and covariance of the regressors. Formally, the null hypothesis is define as

\[ H_0 : E \left[ X_i^\top M_i X_i (\beta_i - \beta) \right] = 0, \]  

where \( X_i = (x_{i1}, \ldots, x_{iT}) \). The test statistic for the null hypothesis in (10) is based on an estimate of equation (5). Define

\[ \hat{\delta} = \frac{1}{N} \sum_{i=1}^{N} X_i^\top M_i X_i (\hat{\beta}_i - \hat{\beta}_{MG}) \]

\[ = \frac{1}{N} \sum_{i=1}^{N} \hat{\delta}_i. \]

The term \( \hat{\delta} \) is an estimate of the quantity that causes the OLS-FE to be inconsistent. We want to test if that quantity is significantly different from zero. To this end, we need an
estimate of its variance, which is calculated using

$$\hat{\Omega} = \frac{1}{N} \sum_{i=1}^{N} \hat{\delta}_i \hat{\delta}_i^\top.$$ 

The statistic of interest, which we refer to as the heterogeneity bias (HB) test, is given by

$$HB = N \times \hat{\delta}^\top \hat{\Omega}^+ \hat{\delta}, \tag{11}$$

where $\hat{\Omega}^+$ is the Moore-Penrose inverse of $\hat{\Omega}$.

The null hypothesis is that heterogeneity in slope parameters does not cause OLS-FE to be inconsistent. Rejection of the null implies that one should use the MG estimator instead.

**Remark 3.1** Hsiao and Pesaran (2008) propose an alternative Hausman test for slope heterogeneity. In particular, they test the difference between the Mean Group estimator and a pooled estimator (like fixed effects). However, in their framework, under the null of no heterogeneity in slopes (nor in error variances), fixed effects is an efficient estimator. Hence, they are able to apply the usual Hausman type variance estimator (the difference between the consistent estimator and efficient estimator variances) to standardize the statistic. In our setup, we do not impose lack of heterogeneity under the null, only a population moment condition as in (10) where fixed effects is consistent. Rejection of the Hausman test might be due to slope heterogeneity, slope heterogeneity bias of fixed effects, or lack of efficiency of the fixed effects estimator for a given empirical situation. Our test is designed to target only the potential bias of fixed effects estimation.

Next, we derive the limiting distribution of the proposed HB test. To this end, we consider the following set of assumptions.

**Assumption 1** Let the matrix $X_i$ have rows $x_{it}^\top$. Suppose that for all $i$, $T^{-1}(X_i M_i X_i)$ is invertible.
Assumption 2 The terms $T^{-1}(X_i^\top M_i X_i)(\beta_i - \beta)$ are independent across $i$ with finite variance.

Assumption 3 Errors $u_i$ are independent across $i$ with $E(u_i|X_j, \alpha_j) = 0$ for all $i$ and $j$, and $u_j$ is independent of $\beta_i$ for all $i$ and $j$.

Assumption 1 allows us to estimate each individual slope parameter, $\beta_i$. This assumption could be violated for cases with a limited number of time-series observations, $T$, for example. It is important to notice that, Condition 1 does not explicitly require $T \to \infty$, which is a common requirement in the literature (see, e.g., Phillips and Sul (2003), Pesaran and Yamagata (2008), Su and Chen (2013)), but it requires enough variability in the time-series dimension to be satisfied. Assumption 2 is related to the usual assumption in random coefficient models where the parameters are all assumed to be independent of each other as well as independent of the data (see, e.g., Hsiao (2003)) for a discussion. Assumption 3 is similar to the standard assumption restricting the errors on the cross-section dimension when $T$ is finite (such as Wooldridge (2010)). Note that the assumptions allow for serially correlated errors as well as heterosekasticity. Assumptions 1–3 are very mild and standard in the literature. Nevertheless, below we will relax these assumptions and modify our test to accommodate the more general conditions.

Now we present the asymptotic distributions of the HB test.

Theorem 1 Suppose that Assumptions 1-3 hold and that $H_0 : E[X_i^\top M_i X_i(\beta_i - \beta)] = 0$, then as $N \to \infty$ with $T$ finite,

$$N\hat{\delta}^\top \hat{\Omega}^+ \hat{\delta} \overset{d}{\to} \chi_k^2,$$

where $k$ is the number of regressors in the model.

Proof. In Appendix A. □
Theorem 1 provides the asymptotic distribution of new test. Notably, implementation of the proposed heterogeneity test is straightforward. One simply: (I) computes the test statistic, $H_B$, using in equation (11); (II) sets the level of significance; and (III) finds the critical values from standard distribution tables. Since $H_B$ is asymptotically bounded by a chi-square distribution, critical values are tabulated and widely available. The null hypothesis is rejected if the value of the test is outside the interval defined by the critical value of choice.

It is important to highlight the differences between the $H_B$ and PYS tests. The $H_B$ test is designed to test a completely different null hypothesis. In particular, the null hypothesis associated with the $H_B$ test is the lack of correlation between variances of the data and the heterogeneous parameters. The PYS test takes the null hypothesis of no parameter heterogeneity, and detects heterogeneity. Therefore, it is possible to reject the null of parameter heterogeneity with the PYS test, yet fail to reject the null of no bias in the fixed effects estimator. The tests are very much complementary.

3.3 Cross-Sectional Dependence

The results presented in the paper can be extended to more general cases. In particular, we can allow for correlation between errors from different cross-sectional units, and a form of endogeneity. In this subsection we show that these extensions can be easily accommodated in the testing procedure by estimating the general appropriated average of the response (slope) coefficients of each of the individuals and its corresponding variance-covariance matrix.

In particular, consider the following model

$$y_{it} = \alpha_i + x_{it}^T \beta_i + u_{it},$$
$$u_{it} = \phi_i^T f_t + \epsilon_{it},$$

where each $\phi_i$ is $1 \times j$ and $f_t$ is a $j \times 1$ vector of unobserved factors.
In the above model, the presence of \( f_t \) allow us to consider cross-section dependence in the errors and in the regressors, and also a form of endogeneity. First, the existence of \( f_t \) in each of the cross-sectional equations allows for cross-sectional dependence in the \( u_{it} \) error terms. Such dependence is modeled in Bai (2009) and Pesaran (2006), among others. In addition to allowing for cross-sectional dependence between the errors, explicit dependence between the factors and the covariates is allowed via

\[
x_{it} = \eta_i + \Lambda_i^T f_t + v_{it}.
\]

Hence, second, \( f_t \) allows for dependence between cross-sectional covariates. That is, \( x_{it} \) and \( x_{jt} \) are dependent through the mutual presence of \( f_t \). Third, the more general model uses \( f_t \) in the dual roles of cross-sectional dependence of error terms, as well as endogeneity of \( x_{it} \) through the effects of \( f_t \) on both \( u_{it} \) and \( x_{it} \).

Pesaran (2006) shows that the average parameter \( E(\beta_i) \) can be consistently estimated by including the cross-sectional averages at each time period, \( \bar{y}_t \) and \( \bar{x}_t \), in each unit regression, and combining via the sample average. To see this, write the model as a system,

\[
\begin{pmatrix}
y_{it} \\
x_{it}
\end{pmatrix} = \begin{pmatrix}
\alpha_i + \beta_i^T \eta_i \\
\eta_i
\end{pmatrix} + \begin{pmatrix}
\beta_i^T \Lambda_i^T + \phi_i^T \\
\Lambda_i^T
\end{pmatrix} f_t + \begin{pmatrix}
\beta_i^T v_{it} + \epsilon_{it} \\
v_{it}
\end{pmatrix}
= B_i + C_i^T f_t + \xi_{it}.
\]
The factors, $f_t$ can be estimated using this model. Define

$$\bar{y}_t = \frac{1}{N} \sum_{i=1}^{N} y_{it}, \quad \bar{x}_t = \frac{1}{N} \sum_{i=1}^{N} x_{it},$$

$$\bar{\xi}_t = \frac{1}{N} \sum_{i=1}^{N} \xi_{it}, \quad B = \frac{1}{N} \sum_{i=1}^{N} B_i,$$

$$\bar{C} = \frac{1}{N} \sum_{i=1}^{N} C_i.$$

If the rank of $\bar{C}$ is full for all $N$, then we can solve for $f_t$ as

$$f_t = (\bar{C} \bar{C}^\top)^{-1} \bar{C} \begin{pmatrix} \bar{y}_t \\ \bar{x}_t \\ \bar{\xi}_t \end{pmatrix} - B - \bar{\xi}_t.$$

The model for $y_{it}$ is written as

$$y_{it} = \alpha_i + x_{it}^\top \beta_i + \phi_i^\top (\bar{C} \bar{C}^\top)^{-1} \bar{C} \begin{pmatrix} \bar{y}_t \\ \bar{x}_t \end{pmatrix} - B - \bar{\xi}_t + \epsilon_{it}.$$

Under regularity conditions, $\bar{\xi}_t$ will converge to zero, and hence we can use $\bar{y}_t$ and $\bar{x}_t$ in the model to account for $f_t$. For each cross-sectional unit, we can estimate $\beta_i$, and then consider the mean group estimator accounting for cross-sectional correlation. To this end, define the $T \times (k + 2)$ matrix

$$Q = \begin{pmatrix} 1 & \bar{y}_1 & \bar{x}_1^\top \\ \vdots & \vdots & \vdots \\ 1 & \bar{y}_T & \bar{x}_T^\top \end{pmatrix},$$
and let $M_Q = I - Q(Q^T Q)^+ Q^T$, so that

$$\hat{\beta}_{Ci} = (X_i^T M_Q X_i)^{-1} X_i^T M_Q y_i.$$ 

The goal of using the cross-sectional averages as regressors is to approximate the factors, $f_t$. If the factors were actually observed, we could replace $M_Q$ with $M_G$ where we define

$$G = \begin{pmatrix} 1 & f_1^T \\ \vdots & \vdots \\ 1 & f_T^T \end{pmatrix}.$$ 

This matrix will be important for the asymptotic results, and we will include an assumption about this matrix in what follows.

The analogue of the fixed effects estimator when allowing for cross-sectional dependence and endogeneity is given by the pooled estimator of Pesaran (2006)

$$\hat{\beta}_P = \left( \sum_{i=1}^N X_i^T M_Q X_i \right)^{-1} \sum_{i=1}^N X_i^T M_Q y_i,$$

where efficiency gains are possible if there is no heterogeneity of the $\beta_i$ parameters. Therefore, the analogue of the Mean Group estimator, which we label Mean Group Correlated (MGC) is given as

$$\hat{\beta}_{MGC} = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_{Ci}.$$ 

The $\hat{\beta}_{MGC}$ is used in the construction of the general version of the $HB$ test. Similar to
the construction of the test above, we have

\[ \hat{\delta}_C = \frac{1}{N} \sum_{i=1}^{N} X_i^\top M_Q X_i (\hat{\beta}_{Ci} - \hat{\beta}_{MGC}) \]  

\[ = \frac{1}{N} \sum_{i=1}^{N} \hat{\delta}_{ci}. \]

Finally, to obtain the statistic of test we need the corresponding variance-covariance matrix, which is computed as

\[ \hat{\Omega}_C = \frac{1}{N - 1} \sum_{i=1}^{N} \hat{\delta}_{Ci} \hat{\delta}_{Ci}^\top. \]

Therefore, analogously to the previous test case, we wish to test the hypothesis that the heterogeneity across individuals does not induce a bias in the pooled estimator (\(\hat{\beta}_P\)). These estimates are asymptotically biased when the heterogeneous coefficients are related to the variance and covariance of the regressors. As in previous section, we wish to test the null hypothesis that there is no correlation between the coefficients and covariance of the regressors. The null hypothesis is define as

\[ H_0 : E \left[ (X_i^\top M_G X_i)(\beta_i - \beta) \right] = 0. \]

The test statistic for the null hypothesis in (13) will be based on an estimate of equation (12).

We refer to the modified test as the Heterogeneity Bias Cross test (HBC) as it allows for the cross-sectional dependence and endogeneity through the factor structure. The test statistic is defined as following

\[ HBC = N \hat{\delta}_C^\top \hat{\Omega}_C^{-1} \hat{\delta}_C. \]

The presence of the factors \(f_t\) changes the assumptions needed to obtain the same limiting distribution, and are similar to those of Pesaran (2006). We consider the following set of
Assumption 4 The common factors \( f_t \) are covariance stationary with absolutely summable autocovariances. Moreover, \( f_t \) is independent of \( \epsilon_{it'} \) and \( v_{it'} \) for all \( i, t, t' \).

Assumption 5 The variables \( \epsilon_{it} \) and \( v_{it} \) are independently distributed for all \( i, j, t, t' \). Both \( \epsilon_{it} \) and \( v_{it} \) are linear stationary processes with absolutely summable autocovariances, where the error process has finite fourth order cumulants.

Assumption 6 The factor loadings \( \phi_i \) and \( \lambda_i \) are iid and independent of all \( f_t, \epsilon_{jt}, v_{jt} \) for all \( i, j, t \). \( \phi_i \) and \( \lambda_i \) have finite second moment matrices.

Assumption 7 The \( \beta_i \) are iid and independent of all \( f_t, \epsilon_{jt}, \phi_j \) and \( \lambda_j \) for all \( i, j, t \). The terms \( \beta_i^\top v_{it} \) are independent across \( i \) with mean zero and finite fourth order cumulants.

Assumption 8 The matrices \( T^{-1}(X_{i}^\top M_Q X_i) \) and \( T^{-1}(X_{i}^\top M_G X_i) \) are non-singular for all \( i \), and their respective inverses have finite second-order moment matrices for all \( i \).

Assumption 9 \( T^{-1}(X_{i}^\top M_G X_i)(\beta_i - \beta) \) are independent across \( i \) with finite second-order moment matrices.

Assumption 10 The matrix \( \bar{C} \) is full rank \( j \leq (k + 1) \) for all \( N \).

These assumptions are general and allow for a wide variety of data generating processes. Assumption 4 restricts the dependence between the cross-sectional \( x_{it} \) variables to be solely a function of the factors, while Assumption 6 ensures that the factors and the other error processes do not influence the loadings. Assumption 5 is a standard condition on stationarity of the innovation terms only requiring the corresponding fourth comments to be finite. Assumption 7 restricts the dependence between \( \beta_i \) and the other variables.\(^5\) Assumption 8

\(^5\)This assumption is a modification of Assumption 4 of Pesaran (2006) in that we allow for \( \beta_i \) to be correlated with the variance of \( v_{it} \), as will happen when fixed effects is not consistent.
is an identification condition for each of the cross-sectional units. The independence condition from Assumption 9 is more general than it might first appear. The $X_i$ matrices might themselves be dependent across $i$ due to the factor structure. However, $M_G$ matrix will eliminate the factors, so that we are essentially restricting the dependence of $\beta_i$ between cross-sectional units. Assumption 10 is the rank condition for identification of $f_t$. Pesaran (2006) shows that the mean group estimator is still consistent if this assumption does not hold. We conjecture that our results would also hold if this assumption does not hold, but with a modification of Assumption 9.

The next result presents the asymptotic distribution of the $HBC$ test.

**Theorem 2** Suppose that Assumptions 4-10 hold and that $H_0 : E \left[ T^{-1} (X_i^\top M_G X_i) (\beta_i - \beta) \right] = 0$. Then as $(N, T) \to \infty$,

$$ N \hat{\delta}_C^\top \hat{\Omega}_C \hat{\delta}_C \to \chi^2_q, $$

where $q \leq k$, with $q$ the rank of $E \left[ T^{-2} (X_i^\top M_G X_i) (\beta_i - \beta)(\beta_i - \beta)^\top (X_i^\top M_G X_i) \right]$.

**Proof.** In Appendix B. ■

Notice that if we fail to reject the null hypothesis, then it may be appropriate to use the pooled estimator. The limiting distribution of the test in Theorem 2 is different from than in Theorem 1 for two reasons. First, for the $HBC$ we require $T \to \infty$ to apply the central limit theorem to $\hat{\delta}_C$. Second, $\hat{\Omega}_C$ converges to

$$ E \left[ T^{-2} (X_i^\top M_G X_i)(\beta_i - \beta)(\beta_i - \beta)^\top (X_i^\top M_G X_i) \right]. $$

The rank of this matrix may vary from 0 to $k$. For example, in the extreme case where there is no heterogeneity, $\beta_i = \beta$ for each $i$, and the rank is 0. The statistic would converge to 0, so that critical values for a $\chi^2_k$ would be conservative. Of course, we would correctly fail to reject that null hypothesis that heterogeneity does not cause bias, as there is no heterogeneity.
To summarize, practical implementation of the test is straightforward. The test is computed in a similar way to the $HB$ test, but with the addition of cross-sectional averages in the regressions for each unit. One compares the $HBC$ statistic to a $\chi^2_k$ distribution. The test will reject the null as there is more heterogeneity bias. In addition, the test correctly becomes more conservative if there are fewer parameters in the model that exhibit heterogeneity.

4 Monte Carlo

We use Monte Carlo simulations to assess the performance of the methods discussed in the previous sections. Our simulations allow for varying degrees of importance assigned to individual slope heterogeneity. The main objective is to assess the finite sample performance of the $HB$ and $HBC$ tests. We evaluate them in terms of size and power. We also compare the standard OLS (without fixed effects), OLS-FE, MG, and MGC estimators in terms of inferential bias and efficiency. Our goal is to illustrate the importance of the individual heterogeneity bias, and its impact on different estimation methods under various data scenarios. In doing so, we restrict our attention to the following cases: (1) no heterogeneity bias; (2) heterogeneity affecting one regressor; (3) heterogeneity affecting both regressors. The first case illustrates the case in which OLS-FE is unbiased and gives the size of the tests. The second case shows that one coefficient can be biased while the other remains unbiased. The last case is more general, as both coefficients can be biased. The last two cases provide the power of the tests.

4.1 Monte Carlo Designs

We consider a simple data generating process (DGP). The variable $y_{it}$ is generated by the model

$$y_{it} = \alpha_i + \theta_i w_{it} + \gamma_i z_{it} + u_{it},$$

(15)
where $\alpha_i$ is the individual-specific intercept term, while $\theta_i$ and $\gamma_i$ are individual-specific slope coefficients associated with exogenous variables $w_{it}$ and $z_{it}$, respectively. The regression error term, $u_{it}$, is normally distributed with mean zero and variance one.

As discussed in Section 2, fixed effects estimation of the population averages $E(\theta_i)$ and $E(\gamma_i)$ will be biased if there is a non-zero covariance between $\theta_i$ and the variance of $w_{it}$, or if there is non-zero covariance between $\gamma_i$ and the variance of $z_{it}$. We model the dependence in $z_{it}$ and $\gamma_i$, and the heterogeneity in $\gamma_i$ as follows:

$$z_{it} = \alpha_i + \varepsilon_{it}^z \quad \text{and} \quad w_{it} = 1 + \varepsilon_{it}^w,$$

where

$$\varepsilon_{it}^z \sim N(0, \sigma_{z_i}^2), \quad \varepsilon_{it}^w \sim N(0, \sigma_{w_i}^2); \quad \sigma_{z_i}^2 = 1 + v_{zi}, \quad \sigma_{w_i}^2 = 1 + v_{wi}; \quad v_{zi} \sim \chi^2_1, \quad v_{wi} \sim \chi^2_1;$$

$$\alpha_i \sim N(0, 1), \quad u_{it} \sim N(0, 1).$$

We divide our experiments into three parts. In the first two parts main objective is to investigate the finite sample performance of the $HB$ test, and in the third part we assess the $HBC$ test.\footnote{We conducted other experiments where the errors were serially correlated. The results were similar and are available upon request.}

In the first part, we set $\theta_i = 1$ for all individuals in order to isolate the effect of just one of the variables having a heterogeneous effect on $y_{it}$. The parameter $\gamma_i$, representing the individual-specific slope, is generated as

$$\gamma_i = 1 + 2c - c\sigma_{z_i}^2 + d\varepsilon_{i}^\gamma,$$

where $\varepsilon_{i}^\gamma$ is $N(0, 1)$ and $c = 0, \pm 0.5, \pm 1$. The parameter $d$ is given by $d = (2 - 2c^2)^{1/2}$ so
that the variance of $\gamma_i$ equals 2 regardless of the correlation between $\sigma_{zi}^2$ and $\gamma_i$. 

The experiments use the parameter $c$ to modulate the importance of individual slope heterogeneity. When $c = 0$, there is heterogeneity in $\gamma_i$ coming from $\epsilon_i^\gamma$. However, since $c = 0$, the heterogeneity in the slope coefficient $\gamma_i$ is not correlated with the variance in the regressor $z_{it}$, and the fixed effects estimator provides an unbiased estimate of $E(\gamma_i)$. As $c$ increases in magnitude, the covariance between $\gamma_i$ and $\sigma_{zi}^2$ increases, leading to biases in the fixed effects estimation of $E(\gamma_i)$. For example, when $c = 1$, the covariance between $\gamma_i$ and $\sigma_{zi}^2$ is $-2$, resulting in negative bias of the fixed effects estimator. Intuitively, the fixed effects estimator assigns more weight to individuals that have smaller slope coefficients $\gamma_i$. When $c = -1$, we have the opposite effect, and the fixed effects estimator will be positively biased. Therefore, the parameter $c$ is very important. Under the null hypothesis of slope homogeneity $c = 0$, and we obtain the size of the $HB$ test. Then, $c \neq 0$ we have slope heterogeneity that causes bias in OLS-FE, and we are able to examine the power of the proposed $HB$ test.

In the second part of the experiment, we also allow $\theta_i$ to be heterogeneous, and generated by $\theta_i = 1 + 2c - c\sigma_{\theta_i}^2 + d\epsilon_i^\theta$, where $\epsilon_i^\theta \sim N(0, 1)$. In this set of simulations, by controlling the constant $c$, we are examining the size and power of the tests. Again, when $c = 0$ we recover size, when $c \neq 0$ we have power of the tests.

Finally, in the third part, we explore size and power of the $HBC$ test by allowing for a factor structure to generate cross-sectional dependence in the errors and endogeneity. In particular, we use the DGP model as in equation (15). In this case, we maintain the same format as in the first case and set $\theta_i = 1$ for all individuals. The covariate $z_{it}$ has $z_{it} = \alpha_i + \epsilon_{it}^z$ with $\epsilon_{it}^z \sim N(0, \sigma_{zi}^2), \sigma_{zi}^2 = 1 + \nu_{zi}, \nu_{zi} \sim \chi_1^2$. Nevertheless we introduce the structure of $w_{it}$ generates cross-sectional dependence in the errors and endogeneity as follows

$$ w_{it} = 1 + \lambda_i^T f_t + \epsilon_{it}^w, \text{ and } u_{it} = \phi_i^T f_t + \epsilon_{it}, $$

$$ \lambda_i = (1, 1)^T, \text{ and } \phi_i = (1, 1)^T + \xi_i, $$

25
where $f_t$ is a $2 \times 1$ vector of iid standard normal variables, and $\xi_t$ is a $2 \times 1$ vector of iid standard normal variables. The presence of the factor in the variable $w_{it}$ causes endogeneity between $u_{it}$ and $w_{it}$. Finally, as in the first case, the parameter $\gamma_i$, representing the individual-specific slope, is generated as

$$
\gamma_i = 1 + 2c - c\sigma^2_{zi} + d\epsilon_i^\gamma,
$$

where $\epsilon_i^\gamma \sim N(0, 1)$ and $c = 0, \pm 0.5, \pm 1$, and the parameter $d$ is given by $d = (2 - 2c^2)^{1/2}$. Once again, the parameter $c$ controls the size and the power of the test.

To benchmark our results, we estimate the model using traditional OLS, where all individuals are (incorrectly) assumed to have the same intercept and slopes. Under the data generating process above, the OLS estimate of $E(\gamma_i)$ will be biased since the individual-specific intercept terms are correlated with $z_{it}$, and OLS is unable to account for individual effects and cross-sectional slope heterogeneity. We also report results for the OLS-FE model that allows for individual intercepts, but as is standard, assumes that the slopes are homogeneous. Finally, we include the mean group estimators. The MG is presented for all comparisons, it allows for individual-specific intercepts as well as slope coefficients. The MGC is presented for the third part of the exercise, and it allows for the cross-sectional dependence.

### 4.2 Monte Carlo Results

We examine the finite sample properties of the methods considering different DGP. We use $N = 500$ for the number of individuals and set the number of time observations, $T$, alternatively, to 5, 10, 20, and 30. The number of replications in each experiment is 5,000.7 We first report the results for evaluation of the estimators. Then, we describe the results for the tests.

7Since the data used in many empirical applications face limitations on the times series dimension, $T$, our main presentation focuses on variations along this dimension. However, in unreported tables we also experiment with variations in the number of individuals, $N$ (e.g., $N = 100$ and $N = 1,000$). These alternative experiments lead to similar inferences and are readily available from the authors.
4.2.1 Results For Bias And RMSE

Results For Experimental Design 1

Table 1 reports the bias and RMSE associated with each of the estimators considered in the first experimental design. In the absence of estimation biases, we would expect to find $E(\theta_i) = 1$ and $E(\gamma_i) = 0$. Deviations from these benchmarks measure the degree of bias in the estimated coefficient. Estimators with better inference properties should present low RMSEs.

Table 1 About Here

As predicted, the OLS estimator is biased even when $c = 0$, and they produce similar results. For example, with $T = 5$ the bias for $E(\gamma_i)$ under OLS is 0.334. As the time dimension, $T$, increases, the OLS bias does not decline. In sharp contrast, the bias under OLS-FE and MG is virtually zero; less than $-0.001$ in both cases. While the OLS-FE and MG estimators have similar, negligible biases, the MG has the smallest RMSE. The reason that OLS-FE does not have the smallest RMSE for the case of $c = 0$ is that it is not efficient even though it is unbiased. The intuition is that if OLS is appropriate in a model with heterogeneity only in intercepts, the random effects model is more efficient. Even though OLS-FE is consistent (unbiased) when $c = 0$, it is not efficient. Moreover, MG has smaller variance for this case even though its use is not necessary.

Now we allow for $c \neq 0$. When $c = -0.5$, the covariance between $\gamma_i$ and $\sigma^2_{z_i}$ is positive and we should see a positive bias in the OLS-FE estimator. This is what we see in Table 1. The estimated values of $E(\gamma_i)$ under OLS-FE are now positively biased with values 0.493, 0.496, 0.494, and 0.497 for $T = 5$, 10, 20, and 30, respectively. The bias is insensitive by increases in the times series dimension. At the same time, the OLS estimate is still biased and have the largest RMSEs. The MG estimator performs uniformly better than all of the other estimators, both in terms of bias magnitude and RMSE. Indeed, the MG method produces virtually unbiased estimates.
When $c = 0.5$, we see a negative bias in $E(\gamma_i)$ for OLS-FE. In this case, the bias in OLS is smaller than when $c = -0.5$, which is an artifact of conflicting bias directions from the intercept effects, $\alpha_i$, and the slope effects, $\gamma_i$. A more interesting observation is that the sign of the OLS-FE bias changes when $c = 0.5$. This change highlights the inferential instability of the OLS-FE estimator in the presence of individual slope heterogeneity.

Finally, as we increase the magnitude of the correlation between $\gamma_i$ and $\sigma^2_{z_i}$ via $c = -1$, the bias for OLS-FE is approximately $-1$, while that of the MG remains virtually equal to zero. That is, under this form of individual slope heterogeneity the OLS-FE suffers from a severe attenuation-like bias that assigns no relevance to estimates associated with the affected variable, even though the variable has a strong predictive power in the true economic model. For $c = 1$, the sign of the OLS-FE bias changes, but the magnitudes are similar to those found when $c = -1$. That is, estimates associated with the affected variable are grossly overestimated and appear to be twice as important as they are in the true model. Notably, the magnitudes of these biases are insensitive to $T$.

Results For Experimental Design 2

We now change the data generating process by also letting $\theta_i$ vary across individuals. This new experiment allows one to have correlation between both slope parameters of the model and the variance of the data. The results of these experiments appear in Table 2.

| Table 2 About Here |

When $c = 0$, by design, the variable $w_{it}$ is uncorrelated with the individual-specific intercept, $\alpha_i$, so that $E(\theta_i)$ should be unbiased. The results in Table 2 confirm this prediction and show approximately unbiased estimates for $E(\theta_i)$ for the OLS, OLS-FE, and MG estimators. There is, however, a significant bias for $E(\gamma_i)$ under OLS estimation, and this bias is insensitive to the times series dimension. Finally, both OLS-FE and MG are approximately
unbiased for \( \text{E}(\gamma_i) \), with MG having the smallest RMSE.

As \( c \) increases in magnitude, the bias and RMSE are expected to increase for OLS, and OLS-FE estimators. Table 2 confirms these predictions. The bias of the OLS-FE estimator is significant for both slope coefficients, whereas MG is largely immune to biases, even for short panels. In addition, the RMSE of the MG estimator declines as the time dimension of the panels increases.

**Results For Experimental Design 3**

The results from the experiment with the factors is presented in Table 3.\(^8\) There are several important points to note from this experiment. First, the parameter \( \theta_i \) is not heterogeneous in this experiment. However, it is the coefficient for the variable \( w_{it} \) which is a function of the factor \( f_t \). Hence, this variable is endogenous. Therefore, we see that OLS, FE, and MG experience bias for each of the cases considered. The estimator that corrects for this effect, MGC, is not affected. As the correlation between the variance of the regressors \( z_{it} \) and the parameter \( \gamma_i \) increases through \( c \), both OLS and OLS-FE perform worse. Again, MGC is unaffected by the correlation that causes heterogeneity bias.

| Table 3 About Here |

### 4.2.2 Depicting The Bias

As a complementary simulation exercise and to illustrate the size of the bias, we reproduce the experiment for \( c \) between 0 and \(-1\), graphing the bias of \( \text{E}(\gamma_i) \) as a function of the parameter \( c \) for the OLS, OLS-FE, and MG estimators. The results are plotted in Figure 1. One can see that the MG estimator is unaffected by the correlation of the heterogeneous slope parameter, \( \gamma_i \), with the variance of the data, \( \sigma_{z_{it}}^2 \). The other estimators are, in contrast, very

\(^8\)The tests are not computed for \( T = 5 \) as there are not enough observations for each cross-sectional regression when the additional cross-sectional means are added.
Figure 1 shows the bias of $E(\gamma_i)$ for OLS, OLS-FE, and MG estimators as a function of the parameter $c$ in the simulations, where $c$ is between 0 and $-1$. It is sensitive to heterogeneity in $\gamma_i$ as it becomes more correlated with its regressor variance.

### 4.2.3 Results For The Heterogeneity Bias Test

We now explore the use of the $HB$ and $HBC$ tests in measuring the extent to which slope heterogeneity causes OLS-FE estimates to be biased. We report the empirical size and power of the tests. We use a 5% nominal size for all computations.

The Experimental designs 1 and 2 assess the finite sample performance of the $HB$ test. They produce equivalent results. Thus, we report results for the first only in Figure 2. The results illustrate the costs of parameter heterogeneity. The $HB$ test is applied to experimental design 1 for $N = 200, 500,$ and $1,000$, as well as values of $T = 10, 20,$ and $30$. When $c = 0$ we obtain the size of the test. As the parameter $c$ increases (in absolute values), there is more bias from parameter heterogeneity. For each of the respective sample sizes, we plot the percentage of rejections of the null hypothesis that OLS-FE estimates are unbiased for 5,000 replications.
Figure 2 shows the power of the Heterogeneity Bias test as a function of the parameter $c$ in the simulations, where $c$ is between 0 and $-1$. For each sample size $N$, power is plotted for $T = 10, 20,$ and 30 respectively.

The results in Figure 2 show that when $c = 0$ the empirical size of the $HB$ test approximates the nominal 5%. The results also show strong evidence that the test is very powerful against deviations for the null hypothesis. The power of $HB$ test improves as the number of individuals, $N$, increases. There is also additional improvement as the number of time periods, $T$, increases. The main point is that as the parameter $c$ increases, the OLS-FE bias increases, and so does the probability of the HB test rejecting the null of no bias in the OLS-FE estimator. The simulations show the usefulness of our test in detecting cases where individual heterogeneity harms inference based on OLS-FE.

Now we use the Experimental design 3 to assess the finite sample performance of the $HBC$ test. In this case, the DGP allows for both cross-sectional dependence in the errors and endogeneity. The result appear in Figure 3. The results indicate the same pattern as in the previous case. The empirical size approximates the nominal size, and the $HBC$ test has large power against deviations from the null hypothesis of slope homogeneity. The results in Figure 3 also show evidence that the $HBC$ test is able to accommodate both cross-sectional
Figure 3 shows the power of the Heterogeneity Bias test as a function of the parameter $c$ in the simulations, where $c$ is between 0 and –1. For each sample size $N$, power is plotted for $T = 10, 20,$ and 30 respectively.

dependence in the errors and endogeneity without size distortion or power loss. This is due to the fact that the test makes use of the proper point estimates and their corresponding variance-covariance matrix.

To summarize our Monte Carlo experiments, we show that heterogeneity in individual responses to a given economic driver may introduce severe biases in methods commonly used to estimate policy parameters. Focusing on estimators designed to recover individuals’ average response (slope) coefficients, we find that the bias of fixed effects estimation can be made arbitrarily large by increasing the magnitude of the covariance between the regression slope and the data variance. The bias can be positive or negative depending on the sign of the covariance. Critically, the mean group estimator we employ is unaffected by the slope heterogeneity bias. In addition, the mean group estimator has the smallest RMSE (hence, the best performance under this inference metric) over the range of our experiments. The
bias and RMSE for the MG estimator is uniformly smaller than those of the other estimators. Moreover those statistics are largely insensitive to the time dimension, $T$. Finally, we show that our new $HB$ and $HBC$ tests have correct empirical size and power to detect precisely the cases where OLS-FE is biased. The $HBC$ test is expected to perform well in empirical setting since it is able to allow for both cross-sectional dependence and endogeneity.

5 Empirical Application: Investment Models

We illustrate our proposed techniques for estimating models with slope heterogeneity using a traditional corporate finance application. In particular, we compare OLS-FE, MG, and MGC estimators in the context of the Fazzari et al.’s (1988) investment model. In the model, a firm’s investment spending is regressed on a proxy for investment demand (Tobin’s $Q$) and operating cash flows. A review of the corporate investment literature shows that virtually all empirical work in the area considers panel data models with firm-fixed effects (see, among others, Kaplan and Zingales (1997) and Rauh (2006)). At the same time that there is a consensus about the inclusion of firm-specific intercepts, existing studies assume homogeneous slope coefficients for $Q$ and cash flow across individual firms.\footnote{A number of papers estimate investment–cash flow sensitivities for sample partitions based on proxies for financial constraints (e.g., firm size or existence of bond ratings). These estimations are also subject to the firm heterogeneity biases that we highlight in our paper.}

The Fazzari et al. model is commonly represented as

$$\text{Investment}_{it} = \alpha_i + \theta Q_{it} + \gamma \text{Cash Flow}_{it} + u_{it},$$

where $\text{Investment}$ is the ratio of current investment spending scaled by the firm’s lagged capital stock, $Q$ is the ratio of the firm’s market value over the book value of assets, and $\text{Cash Flow}$ is the firm’s operating income divided by its lagged capital stock. The parameter $\alpha_i$ is the firm-specific fixed effect and $u_{it}$ is the innovation term.
Suppose that investment is governed by

\[ \text{Investment}_{it} = \alpha_i + \theta_i Q_{it} + \gamma_i \text{Cash Flow}_{it} + u_{it}, \]  

(17)

with \( \theta_i \) and \( \gamma_i \) possibly different across firms. When data is generated according equation (17), assuming an homogeneous \( \theta \) and \( \gamma \) across a panel of firms and estimating model (16) may result in severely biased parameters and incorrect inferences. In what follows, we concretely characterize the problems that arise from estimating (16) in lieu of (17).

5.1 Data Description

Our data are taken from COMPUSTAT from 1970 through 2010. The sample consists of manufacturing firms with fixed capital of more than $5 million (with 1976 as the base year for the CPI). We only study firms whose annual assets and sales growth are less than 100% (e.g., Almeida and Campello (2007)). Summary statistics for investment, \( Q \), and cash flow are presented in Table 4. Since these statistics are similar to those found in other studies, we omit a discussion of their properties.

<table>
<thead>
<tr>
<th>Table 4 About Here</th>
</tr>
</thead>
</table>

To assess the performance of different estimators over the times series dimension, we classify the sample into cases where firms provide, alternatively, a minimum of 10, 20, or 30 years of data.

5.2 Estimation Results

We estimate the Fazzari et al. model using simple least squares (OLS) and least squares with firm intercepts (OLS-FE). Since papers in the literature capture intertemporal variation by
adding time dummies to the OLS method, we also estimate OLS with both firm- and year-
fixed effects (OLS-FE2). Moreover, we compare these methods with the random effects (RE),
the mean group estimator (MG) examined in Section 2.2 as well as the mean group estimator
that allows for cross-sectional correlation and a form of endogeneity (MGC) described in
Section 3.3.

Before performing our estimations, we test for the presence of slope heterogeneity using
the Pesaran-Yamagata-Swamy (PYS) test. We also test for biases in the OLS-FE estimator
using the heterogeneity bias ($HB$ and $HBC$) tests. The results are displayed in Table
5. The $p$-values for both tests are less than 0.000 in all specifications we consider. These
tests systematically reject the null hypothesis of slope homogeneity and show that slope
heterogeneity causes OLS-FE estimates to be biased.

Table 5 About Here

The results for OLS, RE, OLS-FE, and OLS-FE2 in Table 5 resemble those reported in
the literature (e.g., Baker, Stein, and Wurgler (2003), Cummins et al. (2006), and Polk and
Sapienza (2009)). The estimates returned across these methods are similar, suggesting that
little variation is coming from either firm- or time-specific effects. In a related paper Schaller
(1990) examines the empirical performance of investment models, and allows for estimates
for the adjustment cost parameters to vary over firms. The results document empirical ev-
dence of heterogeneity across firms. However, Schaller (1990) model does not control for
cash flow sensitivity and also does not provide formal tests for slope heterogeneity.

The most salient feature of Table 5 is the difference between the cash flow coefficient
that is returned by the MG and MGC estimators and those returned by the other methods.
For the case where we allow firms with 10 or more observations to enter the sample, the
cash flow coefficient under the MG method is 0.291 ($p$-value < 0.001). Under the OLS-FE
estimation, that same coefficient is only 0.043. In other words, the MG estimation suggests
that the impact of cash flow on investment is about seven times larger than what is implied from the standard OLS-FE. Notably, for longer panels, the MG estimates of the cash flow coefficient are about 15 times larger than their OLS-FE counterparts. And the same holds true when we incorporate cross-sectional effects under the MGC estimator. In all, estimation methods that allow for heterogeneity in firm policies suggest that investment responses to cash flow innovations are about one order of magnitude larger than what is estimated under the OLS-FE framework. The mean group estimation also yields larger coefficients for $Q$, but differences across methods are less notable, implying that cross-firm heterogeneity in investment responses to $Q$ is somewhat less pronounced.

5.3 Graphical Evidence

For each sample composition of Table 5, the PYS tests reject the null of slope homogeneity at better than the 0.01% level. Given the statistical evidence of firm slope heterogeneity, we provide a graphical representation of the distribution of estimated firm coefficients. A histogram of $\gamma_i$, the individual firm sensitivity of investment to cash flow, is shown in Figure 4.

Given the evidence that firms are heterogeneous in their responses and the fact that the OLS-FE and MG estimates are very different, one might want to understand and address the nature of the OLS-FE bias. We explore the calculation of the firm heterogeneity bias in the next section, noting that similar calculations can be performed in any other applications considered by the researcher.

5.4 Assessing The OLS-FE Heterogeneity Bias

The OLS-FE produces negatively biased estimates of cash flow coefficients. Such negative bias would result from a negative correlation between the sensitivity of investment to cash flow, $\gamma_i$, and the variance of cash flow, $\sigma^2_{C_{F,i}}$. To assess the degree of that correlation, we
Figure 4 shows the histogram of \( \hat{\gamma}_i \), which represent the individual firm sensitivity to cash flow in the investment equation example. The vertical lines are the estimates for OLS-FE, MG and, Median of \( \hat{\gamma}_i \).

contrast the distribution of those two parameters in Figure 5. We plot two measures of association. The red line shows a nonparametric smoother capturing local features of the relationship.\(^\text{10}\)

Figure 5 shows that there is a strong negative relation between the sensitivity of investment to cash flow and the variance of cash flow for each firm. It is this parameter–data correlation structure that generates the policy heterogeneity bias that arises under the fixed effects framework.

5.5 Investment–Cash Flow Sensitivity Over Time

Our results have a number of implications for the investment literature. Recent work by Chen and Chen (2012) suggests that the sensitivity of investment to cash flow has declined over time. Critically, the authors argue that point estimates of investment–cash flow sen-\(^\text{10}\)The smoother we employ is known as the Friedman Super Smoother (see Friedman (1984)).
Figure 5 confirms the relationship between $\hat{\gamma}_i$ and $\hat{\sigma}^2_{CF,i}$. The red line represents a smooth non-parametric regression of $\hat{\gamma}_i$ on $\hat{\sigma}^2_{CF,i}$.

Cash flow sensitivities are approximately zero since the mid 1990's. To show how these inferences are a direct result of the estimation method employed, we present cash flow sensitivity estimates from both OLS-FE and MG estimators. We estimate equation (17) for each year from 1970 to 2010 using a rolling window of 10 years. Rolling estimates of cash flow sensitivities with 95% confidence bands are presented in Figure 6.

Our results contrast with those of Chen and Chen (2012). Even though we observe a downward trend in investment–cash flow sensitivities, the magnitude of the estimated coefficients are far from zero, and imply economically significant effects through present times. For example, the MG-estimated coefficient for 2006 (which is gauged from the 2001–2010 time window) is 0.183, which contrasts with the OLS-FE estimate of only 0.024.

Given the evidence of biases in OLS-FE based estimations in environments with heterogeneous firms, the use of the MG estimator appears to be a robust way to estimate the magnitude of the population parameters in the corporate investment model. Our results
Figure 6 shows the estimated coefficients of $\hat{\gamma}$ using OLS-FE and MG estimation using a rolling window of $T = 10$ time-series observations for each year. The standard deviation of cash flow divided by 10 is shown for the same rolling window periods. The estimated coefficients show that the sensitivity of investment to cash flow remains — in a statistical sense — a potentially important factor in understanding investment behavior.

6 Conclusion

We show that heterogeneity in individual responses to economic stimuli can distort estimates that are commonly reported. The theorems and Monte Carlo evidence advanced in this paper show that standard fixed effects models may often render biased estimates of individual responses in light of slope heterogeneity. Critically, the bias of the fixed effects estimator can be very pronounced and lead to mistaken economic conclusions.

As part of the exploration of the potential bias in fixed effects estimation, we propose and analyze new statistical tests that are designed to indicate whether fixed effects is the appropriate estimator. One interpretation of the proposed tests is that they are testing the
moment conditions required for fixed effects estimator to be consistent for corresponding population parameters of interest. The tests are simple to calculate, governed by a simple distribution, and shown to be effective in an extensive Monte Carlo experiment. The tests are intuitive, in that it directly estimates the bias associated with the fixed effects estimator, and standardizes the estimated bias. The structure of our heterogeneity bias ($HB$ and $HBC$) tests are similar to a Wald test, and have chi-square limiting distributions under the null hypothesis of no heterogeneity bias in the fixed effects estimator. The $HBC$ test is a generalization of the $HB$ test by allowing for cross-section dependence in the errors and a form of endogeneity.

Finally, it is worth noting that even if one fails to reject the null hypothesis of no heterogeneity bias in fixed effects estimation, the fixed effects estimator may not be efficient. The results from our Monte Carlo experiment illustrate this fact (when we know that there is no bias by construction, yet mean group estimators (MG and MGC) have smaller MSE). Moreover, in our empirical example, we reject the null of no heterogeneity bias for the fixed effects estimator, yet the standard errors are sometimes smaller or larger for the fixed effects estimated coefficients. Our results suggest that the Mean Group Correlated estimator of Pesaran (2006) provides a useful robustness check. The test we propose is a formal way to check if the mean group estimators (MG and MGC) are the appropriate tool for a given empirical study.
Appendix

Recall the following definitions: \( y_i = (y_{i1}, \ldots, y_{iT})^\top \), \( X_i = (x_{i1}, \ldots, x_{iT})^\top \), \( u_i = (u_{i1}, \ldots, u_{iT})^\top \), \( M_i = I_T - \upsilon_T (\upsilon_T^\top \upsilon_T)^{-1} \upsilon_T \), \( I_T \) is an identity matrix of order \( T \), \( \upsilon_T \) is a \( T \times 1 \) vector of ones.

Proof of Theorem 1: First,

\[
\hat{\beta}_i - \hat{\beta}_{MG} = \beta_i + (X_i^\top M_i X_i)^{-1} X_i^\top M_i u_i - \frac{1}{N} \sum_{j=1}^{N} (X_j^\top M_i X_j)^{-1} X_j^\top M_i y_j
\]

\[
= \beta_i - \beta + (X_i^\top M_i X_i)^{-1} X_i^\top M_i u_i - \frac{1}{N} \sum_{i=1}^{N} (\beta_i - \beta) - \frac{1}{N} \sum_{j=1}^{N} (X_j^\top M_i X_j)^{-1} X_j^\top M_i u_j.
\]

Then,

\[
\sqrt{N} \hat{\delta} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i (\hat{\beta}_i - \hat{\beta}_{MG})
\]

\[
= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i (\beta_i - \beta) + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T-1} X_i^\top M_i u_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i \left[ \frac{1}{N} \sum_{i=1}^{N} (\beta_i - \beta) \right]
\]

\[
- \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i \left[ \frac{1}{N} \sum_{j=1}^{N} (X_j^\top M_i X_j)^{-1} X_j^\top M_i u_j \right]
\]

\[
= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{1}{T} X_i^\top M_i X_i - \frac{1}{N} \sum_{j=1}^{N} \frac{1}{T} X_j^\top M_i X_j \right) (\beta_i - \beta)
\]

\[
+ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i u_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i \left[ \frac{1}{N} \sum_{j=1}^{N} (X_j^\top M_i X_j)^{-1} X_j^\top M_i u_j \right]
\]

\[
= A_1 + A_2.
\]

Let \( r_i = (X_i^\top M_i X_i)^{-1} X_i^\top M_i u_i \), so that

\[
A_2 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i (r_i - \bar{r}).
\]

By Assumption 1 and under the null hypothesis, the central limit theorem applies to both
terms. Write the terms

\[ A_1 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} A_{1i}, \]
\[ A_2 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} A_{2i}. \]

The estimated variance matrix is

\[
\hat{\Omega} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T} X_i^\top M_i X_i \left( \beta_i - \hat{\beta}_{MG} \right) \left( \hat{\beta}_i - \hat{\beta}_{MG} \right) \top \frac{1}{T} X_i^\top M_i X_i \\
= \frac{1}{N} \sum_{i=1}^{N} \left( A_{1i} A_{1i}^\top + A_{1i} A_{2i}^\top + A_{2i} A_{1i}^\top + A_{2i} A_{2i}^\top \right).
\]

Since \( A_{1i} \) and \( A_{2i} \) are uncorrelated,

\[
\frac{1}{N} \sum_{i=1}^{N} A_{1i} A_{2i}^\top \xrightarrow{p} 0.
\]

Hence, the estimated variance converges to

\[
\frac{1}{N} \sum_{i=1}^{N} \left[ E(A_{1i} A_{1i}^\top) + E(A_{2i} A_{2i}^\top) \right].
\]

The dominant term of the first block contains \( T^{-2} E[(X_i^\top M_i X_i)(\beta_i - \beta)(\beta_i - \beta)^\top (X_i^\top M_i X_i)] \).

The rank of this term vary from 0 to \( k \). The second block has leading term

\[
\frac{1}{NT^2} \sum_{i=1}^{N} E \left[ (X_i^\top M_i X_i)^{-1} X_i^\top M_i \Omega_i M_i X_i (X_i^\top M_i X_i)^{-1} \right],
\]

which is full rank, with

\[
\Omega_i = E(u_i u_i^\top | X_i).
\]
B Appendix

Proof of Theorem 2: We find the distribution of the normalized estimate of the population moment condition in question. First,

\[
\hat{\delta}_C = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_Q X_i}{T} \right) (\hat{\beta}_{Ci} - \hat{\beta}_{MGC})
\]

\[
= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_Q X_i}{T} \right) \left[ (\hat{\beta}_{Ci} - \beta_i) + (\beta_i - \beta) + (\beta - \hat{\beta}_{MGC}) \right]
\]

\[
= B_1 + B_2 + B_3.
\]

The first term, \(B_1\), can be simplified as

\[
\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_Q X_i}{T} \right) (\hat{\beta}_{Ci} - \beta_i) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_Q X_i}{T} \right) (X_i^\top M_Q X_i)^{-1} X_i^\top M_Q u_i
\]

\[
= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_Q u_i}{T} \right).
\]

Given Assumptions 4, 5, 7, and 10, Lemma 2 of Pesaran (2006) still holds, and we have \(M_Q = M_G + O_p(N^{-1}) + O_p(N^{-1/2}T^{-1/2})\). Hence, we eliminate the factors in \(u_i\) by the matrix \(M_G\) so that

\[
\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_G \epsilon_i}{T} \right) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_Q \epsilon_i}{T} \right) + O_p(N^{-1/2}) + O_p(T^{-1/2}).
\]

Given the independence across \(i\), we have

\[
E \left[ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^\top M_G \epsilon_i}{T} \right) \right]^2 = \frac{1}{N} \sum_{i=1}^{N} E \left( \frac{X_i^\top M_G \epsilon_i}{T} \right)^2
\]

\[
= O(T^{-1}),
\]

so that \(B_1 = O_p(N^{-1/2}) + O_p(T^{-1/2})\).
Considering $B_3$, we have

$$\hat{\beta}_{MGC} = \frac{1}{N} \sum_{i=1}^{N} \hat{\beta}_{Ci}$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left[ \beta_i + (X_i^T M_Q X_i)^{-1} X_i^T M_Q u_i \right]$$

$$= \beta + \frac{1}{N} \sum_{i=1}^{N} (\beta_i - \beta) + \frac{1}{N} \sum_{i=1}^{N} (X_i^T M_Q X_i)^{-1} X_i^T M_Q u_i$$

$$= \beta + \frac{1}{N} \sum_{i=1}^{N} (\beta_i - \beta) + \frac{1}{N} \sum_{i=1}^{N} (X_i^T M_G X_i)^{-1} X_i^T M_G \epsilon_i + O_p(N^{-1}) + O_p(N^{-1/2}T^{-1/2}).$$

Noting that

$$\frac{1}{N} \sum_{i=1}^{N} (X_i^T M_G X_i)^{-1} X_i^T M_G \epsilon_i = O_p(N^{-1/2}T^{-1/2}),$$

we have

$$B_3 = - \left[ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^T M_G X_i}{T} \right) \frac{1}{N} \sum_{i=1}^{N} (\beta_i - \beta) \right] + O_p(N^{-1/2}) + O_p(T^{-1/2}).$$

Now combining $B_2$ and $B_3$ we have that

$$B_2 + B_3 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \frac{X_i^T M_G X_i}{T} \right) \left[ (\beta_i - \beta) - \frac{1}{N} \sum_{j=1}^{N} (\beta_j - \beta) \right] + O_p(N^{-1/2}) + O_p(T^{-1/2}).$$

Given Assumption 9, we apply the central limit theorem. The variance of the leading term is

$$\mathbb{E} \left[ \left( \frac{X_i^T M_G X_i}{T} \right) (\beta_i - \beta)(\beta_i - \beta)^T \left( \frac{X_i^T M_G X_i}{T} \right) \right].$$
As in Theorem 1, we can estimate this via

$$
\hat{\Omega}_C = \frac{1}{N-1} \sum_{i=1}^{N} \hat{\delta}_C \hat{\delta}_C^\top 
$$

$$
= \frac{1}{N-1} \sum_{i=1}^{N} \left( \frac{X_i^\top M Q X_i}{T} \right) \left( \hat{\beta}_{C_i} - \hat{\beta}_{MGC} \right) \left( \hat{\beta}_{C_i} - \hat{\beta}_{MGC} \right)^\top \left( \frac{X_i^\top M Q X_i}{T} \right).
$$

This estimator is the analogue of (67) of Pesaran (2006), and consistent for the population variance.
References


Table 1. OLS, OLS-FE, and MG estimators: heterogeneity in $\gamma_i$

<table>
<thead>
<tr>
<th>$c$</th>
<th>OLS Bias</th>
<th>RMSE</th>
<th>OLS-FE Bias</th>
<th>RMSE</th>
<th>MG Bias</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E(\theta_i)$</td>
<td>$E(\gamma_i)$</td>
<td>$E(\theta_i)$</td>
<td>$E(\gamma_i)$</td>
<td>$E(\theta_i)$</td>
<td>$E(\gamma_i)$</td>
</tr>
<tr>
<td>$T = 5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = 0$</td>
<td>0.000</td>
<td>0.334</td>
<td>0.000</td>
<td>0.333</td>
<td>0.000</td>
<td>0.333</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.039</td>
<td>0.345</td>
<td>0.028</td>
<td>0.344</td>
<td>0.020</td>
<td>0.343</td>
</tr>
<tr>
<td>$c = -0.5$</td>
<td>-0.001</td>
<td>0.664</td>
<td>0.001</td>
<td>0.663</td>
<td>0.000</td>
<td>0.662</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.042</td>
<td>0.673</td>
<td>0.030</td>
<td>0.671</td>
<td>0.021</td>
<td>0.670</td>
</tr>
<tr>
<td>$c = 0.5$</td>
<td>0.001</td>
<td>0.004</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.045</td>
<td>0.128</td>
<td>0.032</td>
<td>0.116</td>
<td>0.022</td>
<td>0.112</td>
</tr>
<tr>
<td>$c = -1$</td>
<td>0.000</td>
<td>0.994</td>
<td>0.000</td>
<td>0.995</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.051</td>
<td>1.009</td>
<td>0.036</td>
<td>1.008</td>
<td>0.025</td>
<td>1.013</td>
</tr>
<tr>
<td>$c = 1$</td>
<td>0.001</td>
<td>-0.335</td>
<td>0.000</td>
<td>-0.331</td>
<td>0.000</td>
<td>-0.335</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.057</td>
<td>0.386</td>
<td>0.040</td>
<td>0.375</td>
<td>0.028</td>
<td>0.376</td>
</tr>
<tr>
<td>OLS-FE Bias</td>
<td>-0.000</td>
<td>-0.999</td>
<td>-0.001</td>
<td>-0.994</td>
<td>0.000</td>
<td>-0.998</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.053</td>
<td>1.026</td>
<td>0.035</td>
<td>1.016</td>
<td>0.024</td>
<td>1.019</td>
</tr>
<tr>
<td>MG Bias</td>
<td>-0.000</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.036</td>
<td>0.072</td>
<td>0.015</td>
<td>0.066</td>
<td>0.009</td>
<td>0.064</td>
</tr>
</tbody>
</table>

This table shows the bias and the RMSE associated with the estimation of the model in equation (15) using the OLS, OLS-FE, and MG estimators in simulated panel data in the presence of heterogeneity in the parameter $\gamma_i$. The parameter $c$ controls the amount and direction of the heterogeneity bias in $\gamma_i$. Sample size is $N = 500$ and $T = \{5, 10, 20, 30\}$. The number of replications is 5,000.
Table 2.
OLS, OLS-FE, and MG estimators: heterogeneity in $\theta_i$ and $\gamma_i$

<table>
<thead>
<tr>
<th></th>
<th>$T = 5$</th>
<th>$T = 10$</th>
<th>$T = 20$</th>
<th>$T = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E(\theta_i)$</td>
<td>$E(\gamma_i)$</td>
<td>$E(\theta_i)$</td>
<td>$E(\gamma_i)$</td>
</tr>
<tr>
<td>$c = 0$</td>
<td>OLS Bias 0.002 0.331 -0.001 0.333 0.002 0.331 -0.001 0.334</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.102 0.345 0.090 0.344 0.086 0.341 0.083 0.344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.002 -0.003 -0.001 0.001 0.001 0.000 -0.001 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.103 0.102 0.088 0.088 0.084 0.082 0.082 0.080</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.000 -0.001 0.000 0.000 0.001 0.001 0.000 0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.075 0.074 0.064 0.065 0.064 0.064 0.064 0.064</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = -0.5$</td>
<td>OLS Bias 0.497 0.664 0.497 0.664 0.497 0.666 0.498 0.665</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.519 0.675 0.514 0.673 0.513 0.674 0.513 0.672</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.497 0.495 0.498 0.497 0.497 0.499 0.498 0.497</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.519 0.517 0.514 0.514 0.512 0.514 0.513 0.511</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.000 -0.001 -0.001 0.000 0.001 0.001 0.000 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.072 0.073 0.064 0.064 0.064 0.064 0.064 0.063</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = 0.5$</td>
<td>OLS Bias -0.497 0.003 -0.495 0.001 -0.497 0.002 -0.499 0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.519 0.129 0.513 0.121 0.513 0.118 0.513 0.114</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias -0.497 -0.495 -0.495 -0.497 -0.497 -0.495 -0.499 -0.498</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.519 0.516 0.512 0.514 0.512 0.514 0.514 0.512</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias -0.001 0.001 0.001 0.000 0.000 0.000 -0.001 0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.073 0.072 0.065 0.065 0.064 0.064 0.065 0.064</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = -1$</td>
<td>OLS Bias 0.990 0.995 0.996 0.994 0.997 0.998 0.994 0.998</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 1.018 1.011 1.018 1.007 1.018 1.010 1.014 1.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.990 0.991 0.995 0.988 0.997 0.994 0.994 0.995</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 1.018 1.020 1.018 1.011 1.017 1.015 1.014 1.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias -0.002 0.000 0.001 0.000 0.001 0.000 -0.001 0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.073 0.073 0.065 0.064 0.064 0.062 0.063 0.065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = 1$</td>
<td>OLS Bias -0.991 -0.328 -0.993 -0.331 -0.992 -0.332 -0.998 -0.329</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 1.018 0.381 1.015 0.376 1.013 0.375 1.019 0.369</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias -0.991 -0.988 -0.995 -0.991 -0.992 -0.997 -0.998 -0.991</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 1.019 1.017 1.016 1.014 1.012 1.017 1.019 1.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.001 0.000 0.000 0.000 0.001 -0.001 -0.001 0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 0.074 0.073 0.065 0.065 0.064 0.065 0.065 0.062</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table shows the bias and the RMSE associated with the estimation of the model in equation (15) using the OLS, OLS-FE, and MG estimators in simulated panel data in the presence of heterogeneity in the parameters $\theta_i$ and $\gamma_i$. The parameter $c$ controls the amount and direction of the heterogeneity bias in both $\theta_i$ and $\gamma_i$. Sample size is $N = 500$ and $T = \{5, 10, 20, 30\}$. The number of replications is 5,000.
Table 3.
OLS, OLS-FE, MG, and MGC estimators: heterogeneity in $\gamma_i$

<table>
<thead>
<tr>
<th></th>
<th>$T = 10$ E($\theta_i$)</th>
<th>$T = 20$ E($\gamma_i$)</th>
<th>$T = 30$ E($\theta_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c = 0$</td>
<td>OLS Bias 0.449 0.331 0.476 0.335 0.482 0.333</td>
<td>RMSE 0.464 0.341 0.483 0.345 0.487 0.342</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.540 -0.001 0.541 0.001 0.539 -0.001</td>
<td>RMSE 0.553 0.068 0.547 0.065 0.543 0.066</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.474 -0.003 0.488 0.002 0.491 -0.001</td>
<td>RMSE 0.489 0.088 0.495 0.082 0.495 0.081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MGC Bias 0.001 -0.002 0.000 0.001 0.000 -0.001</td>
<td>RMSE 0.029 0.069 0.014 0.064 0.010 0.065</td>
<td></td>
</tr>
<tr>
<td>$c = -0.5$</td>
<td>OLS Bias 0.449 0.002 0.474 0.003 0.484 0.001</td>
<td>RMSE 0.465 0.114 0.482 0.114 0.489 0.111</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.540 0.001 0.540 0.001 0.540 0.000</td>
<td>RMSE 0.553 0.067 0.546 0.065 0.545 0.064</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.474 -0.495 0.487 -0.495 0.492 -0.498</td>
<td>RMSE 0.489 0.511 0.494 0.511 0.497 0.512</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MGC Bias 0.001 0.001 0.000 0.000 0.000 -0.001</td>
<td>RMSE 0.029 0.068 0.014 0.065 0.010 0.063</td>
<td></td>
</tr>
<tr>
<td>$c = 0.5$</td>
<td>OLS Bias 0.449 0.666 0.476 0.664 0.484 0.662</td>
<td>RMSE 0.465 0.675 0.483 0.672 0.489 0.670</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.540 0.000 0.541 0.000 0.541 0.000</td>
<td>RMSE 0.553 0.070 0.547 0.065 0.545 0.064</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.474 0.498 0.488 0.497 0.492 0.495</td>
<td>RMSE 0.489 0.515 0.495 0.511 0.497 0.509</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MGC Bias 0.000 0.001 0.000 0.000 0.000 -0.001</td>
<td>RMSE 0.029 0.071 0.014 0.065 0.010 0.063</td>
<td></td>
</tr>
<tr>
<td>$c = -1$</td>
<td>OLS Bias 0.450 -0.328 0.475 -0.329 0.483 -0.332</td>
<td>RMSE 0.466 0.372 0.483 0.369 0.488 0.373</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.541 0.002 0.541 0.001 0.540 0.000</td>
<td>RMSE 0.555 0.067 0.547 0.065 0.544 0.065</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.475 -0.992 0.488 -0.993 0.492 -0.997</td>
<td>RMSE 0.490 1.014 0.495 1.013 0.497 1.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MGC Bias 0.001 0.001 0.000 0.000 0.000 -0.001</td>
<td>RMSE 0.029 0.068 0.014 0.064 0.011 0.064</td>
<td></td>
</tr>
<tr>
<td>$c = 1$</td>
<td>OLS Bias 0.450 0.994 0.475 0.999 0.485 0.998</td>
<td>RMSE 0.466 1.007 0.482 1.012 0.490 1.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG Bias 0.541 0.000 0.540 0.001 0.541 0.001</td>
<td>RMSE 0.554 0.068 0.546 0.067 0.545 0.064</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS-FE Bias 0.475 0.991 0.487 0.997 0.493 0.995</td>
<td>RMSE 0.491 1.013 0.495 1.018 0.498 1.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MGC Bias 0.000 0.000 0.000 0.002 0.000 0.001</td>
<td>RMSE 0.029 0.069 0.014 0.066 0.010 0.063</td>
<td></td>
</tr>
</tbody>
</table>

This table shows the bias and the RMSE associated with the estimation of the model in equation (15) using the OLS, OLS-FE, MG, and MGC estimators in simulated panel data in the presence of heterogeneity in the parameter $\gamma_i$. The parameter $c$ controls the amount and direction of the heterogeneity bias in $\gamma_i$. Sample size is $N = 500$ and $T = \{10, 20, 30\}$. The number of replications is 5,000.
### Table 4.
Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Median</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment</td>
<td>50449</td>
<td>0.219</td>
<td>0.166</td>
<td>0.181</td>
<td>0.048</td>
<td>0.526</td>
</tr>
<tr>
<td>Q</td>
<td>50449</td>
<td>1.010</td>
<td>0.713</td>
<td>0.860</td>
<td>0.537</td>
<td>1.861</td>
</tr>
<tr>
<td>Cash Flow</td>
<td>50449</td>
<td>0.343</td>
<td>0.681</td>
<td>0.299</td>
<td>-0.138</td>
<td>1.011</td>
</tr>
</tbody>
</table>

This table shows the basic descriptive statistics for Q, cash flow, and investment. The data are taken from the annual COMPUSTAT industrial files over the 1970 to 2010 period. See text for details.
Table 5.  
Empirical Results for Model of Investment

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS</th>
<th>RE</th>
<th>OLS-FE</th>
<th>OLS-FE2</th>
<th>MG</th>
<th>MGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T \geq 10$ $Q$</td>
<td>0.034*** (0.005)</td>
<td>0.033*** (0.004)</td>
<td>0.031*** (0.004)</td>
<td>0.034*** (0.005)</td>
<td>0.045*** (0.006)</td>
<td>0.041*** (0.009)</td>
</tr>
<tr>
<td>$Cash Flow$</td>
<td>0.057*** (0.021)</td>
<td>0.046** (0.019)</td>
<td>0.043** (0.018)</td>
<td>0.041** (0.017)</td>
<td>0.291*** (0.009)</td>
<td>0.288*** (0.011)</td>
</tr>
<tr>
<td>Observations</td>
<td>41254</td>
<td>PYS p-value</td>
<td>0.000</td>
<td>HB p-value</td>
<td>0.000</td>
<td>HBC p-value</td>
</tr>
</tbody>
</table>

| $T \geq 20$ $Q$ | 0.029*** (0.004) | 0.028*** (0.004) | 0.027*** (0.004) | 0.029*** (0.004) | 0.045*** (0.006) | 0.032*** (0.007) |
| $Cash Flow$    | 0.038* (0.020) | 0.031* (0.017) | 0.038* (0.016) | 0.028* (0.015) | 0.251*** (0.009) | 0.248*** (0.010) |
| Observations   | 24125 | PYS p-value | 0.000 | HB p-value | 0.000 | HBC p-value | 0.000 |

| $T \geq 30$ $Q$ | 0.023*** (0.002) | 0.022*** (0.003) | 0.022*** (0.003) | 0.024*** (0.004) | 0.030*** (0.007) | 0.026*** (0.008) |
| $Cash Flow$    | 0.020* (0.011) | 0.015* (0.009) | 0.015* (0.008) | 0.015* (0.008) | 0.228*** (0.012) | 0.231*** (0.012) |
| Observations   | 13764 | PYS p-value | 0.000 | HB p-value | 0.000 | HBC p-value | 0.000 |

Firm Effects    | No    | No        | Yes     | Yes     | Yes    | Yes     |
Time Effects     | No    | No        | No      | Yes     | Yes    | Yes     |

This table shows coefficients and standard errors returned from applying OLS, RE, OLS-FE, OLS-FE2, MG and MGC methods to estimate equation (17). Standard errors are in parentheses. p-values for Heterogeneity Bias (HB and HBC) tests are calculated for each subsample. The data are taken from the annual COMPUSTAT industrial files over the 1970 to 2010 period. See text for details. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.