Market-wide Events and Time Fixed Effects*

Elvira Sojli†
University of New South Wales

Wing Wah Tham‡
University of New South Wales

Wendun Wang§
Erasmus University and Tinbergen Institute

Abstract

Market-wide events such as financial crises and regulatory changes have been empirically shown to have heterogenous impact on firm outcomes. Time-fixed effect, a widely used approach to account for market-wide shocks in existing empirical studies, assumes a homogenous response to shocks. This paper investigates the effect of not accounting for heterogenous responses to common shocks for existing panel studies. We demonstrate that ignoring time-varying unobserved heterogeneity in current empirical practices leads to biased estimates, due to omitted variable bias. We theoretically derive the degree of bias under the data generating process (DGP) with time-varying heterogeneity, for different widely-used methods, such as the two-way fixed effect and interacted fixed effect models. We propose the use of the "group fixed effect" model to overcome the omitted variable bias problem and demonstrate its economic importance through a simulation and two empirical applications. Finally, we provide researchers with guidance and user-written functions in statistical packages to overcome the limitations of existing approaches.

Keywords: Time-varying unobserved heterogeneity; clustering; common shocks; group fixed effects; fixed effects; omitted variable bias
JEL Classification: G10; G20; G14.

---

*We are very grateful to Renee Adams, Chang Mo Kang, and Rik Sen for insightful and helpful discussions. We thank seminar participants at UNSW for helpful comments.

†Address: School of Banking and Finance, UNSW Business School, UNSW Sydney, NSW 2052 Australia. Email: e.sojli@unsw.edu.au.

‡Address: School of Banking and Finance, UNSW Business School, UNSW Sydney, NSW 2052 Australia. Email: w.tham@unsw.edu.au.

§Address: Erasmus School of Economics, Erasmus University, PO Box 1738, Rotterdam, 3000DR, the Netherlands. Email: Wang@ese.eur.nl

1 Introduction

Financial regulations, e.g., Sarbanes-Oxley Act (SOX), and market-wide events, e.g. financial crises, affect firms' outcomes and have been the focus of many papers. One of the salient findings of these papers is that market-wide shocks have heterogenous effects on firm outcomes.[1] These results imply the existence of time-varying unobserved heterogeneity that might be correlated with the treatment variable of interest. From standard econometric textbooks, we know that the failure to account for time-varying unobserved heterogeneity results in biased estimates because of omitted variable bias. Furthermore, the degree of bias depends on the correlation and the relative size of the heterogeneity with respect to the variations in treatment and explanatory variables.

In this paper, we study the econometric implications of the heterogenous effects from market-wide shocks for the current econometric specifications, practices, and findings in the finance and the economics literature. First, we theoretically derive the degree of bias under the data generating process (DGP) with time-varying heterogeneity for different widely-used methods, such as the two-way fixed effect and interacted fixed effect models. Second, we propose the use of the "group fixed effect" (GFE) model of Bonhomme and Manresa (2015a) to overcome the omitted variable bias problem and to obtain consistent estimates of interest, and provide guidance on its use.

Through simulations across different DGPs, we evaluate the efficacy of the two-way fixed effect and interacted fixed effect models with respect to the GFE model. Simulation results suggest that GFE provides not only consistent estimates, under the assumption of time-varying unobserved heterogeneity, but it also produces consistent estimates, when firms/individuals respond homogenously to market-wide shocks. As a result, we propose the use of GFE in addition to existing approaches for future empirical work in finance. We bring these competing approaches to a real data example, and we show that the economic magnitude of the bias is substantial, when scholars fail to account for heterogenous firm responses to market-wide shocks in empirical studies. Finally, we provide researchers with guidance on GFE and user-written functions in statistical packages to overcome the limitations of existing approaches.

Controlling for unobserved heterogeneity that correlates with the variable of interest is important

[1] For example see Mitton (2002), Joh (2003), Lemmon and Lins (2003), Banerjee, Jenner, and Nanda (2015), Thakor (2015), Linck, Netter, and Yang (2009), Duchin, Matsusaka, and Ozbas (2010a), Duchin, Ozbas, and Sensoy (2010b) among many others.

in empirical research, and scholars often model heterogeneity as an individual/firm-specific, time-invariant fixed effect. In addition, a time-related binary variable (time fixed effect) is included to control for unexpected variations or special events, such as SOX and the 2008 global financial crisis, which might affect the outcome variable of interest. Hence, a group (or individual) and time two-way fixed effect model allows for time-varying unobserved heterogeneity through the time-fixed effect that changes homogenously across all individual firms. In other words, using time fixed effect assumes that a particular special event or a common shock has identical impact on each firm. The two-way fixed effect model is the major workhorse for any accounting, economics and finance scholar working with panel data.

However, the homogenous impact assumption for time-fixed effects is unlikely to be true, due to heterogeneity across firms. The heterogeneous impact arises from heterogeneity in managers (skills and preferences), executive boards, and other intermediaries and empirical papers studying market-wide shocks regularly document different impacts on firm-specific outcome variables. Two recent examples of the heterogeneous impact of common shocks are the implementation of Sarbanes-Oxley Act (SOX) and the 2008 financial crisis.[2] Several recent papers show the heterogenous effect of SOX on corporate governance, due to the heterogeneity in observable and unobservable managers and firm characteristics.[3] Work on the various recent financial crises finds that observable firm characteristics (e.g., industry, profitability, and credit rating), as well as unobservable characteristics (e.g., managerial qualities and characteristics) affects firms' outcome variables in the cross-section (see Mitton, 2002; Joh, 2003; Lemmon and Lins, 2003; Campello et al., 2010, 2011; Thakor, 2015; Ho et al., 2016).

While scholars in finance regularly observe heterogenous responses to common (regulatory) shocks in financial markets, they routinely use time-fixed effects with homogenous impact to control for these shocks. Such practices will result in omitted variable bias, if the heterogenous responses are correlated with the treatment variable of interest. We survey recently published papers in the

[2] The Securities and Exchange Commission SOX Act in 2002 affected a myriad of market participants, who are regularly expected to make firm-related decisions, like firm managers and executive boards, and who affect firms and managers decisions like auditors and institutional investors. SOX was implemented to strengthen the independence of auditing firms, to improve the quality and transparency of financial statements and corporate disclosure, to enhance corporate governance, to improve the objectivity of research, and to strengthen the enforcement of the federal securities laws.

[3] For example, see Engel et al. (2007), Chhaochharia and Grinstein (2007), Heron and Lie (2007), Koh et al. (2008), Cohen et al. (2008), Linck et al. (2009), Bargeron et al. (2010), Carter et al. (2009), Duchin et al. (2010a), Duchin et al. (2010b), Brochet (2010), and Banerjee et al. (2015).

top three finance journals for the prevalence of such practices. From 2008 to 2016, there are 100 published empirical papers in the Journal of Finance (25), the Review of Financial Studies (38), and the Journal of Financial Economics (37) using panel data. There are 92 papers that use panel regression and 8 without panel regression. Among those using panel regression, 80 papers use fixed effect models with 40 using one-way fixed effect and 40 using two-way (individual and time) fixed effect models. Seven papers use pooled regression and five use random and fixed effect models. There is only one paper out of these 100 papers that uses an interacted fixed effect (firm and industry-year) model relaxing the assumption of homogeneous impact of market-wide shocks.

Heterogenous responses to market-wide shocks also imply that standard errors should be clustered by groups with similar responses to the shock, given that the market-wide responses of market participants, managers and executive board members are likely to be correlated to responses of the treatment variable of interest. Petersen (2009) and Thompson (2011) highlight the importance of appropriately adjusting the standard errors to account for correlation across firms and across time in finance.[4] However, correcting biased standard errors through clustering by fixed firm characteristics, such as firm and industry, remains problematic under the data structure with time-varying unobserved heterogeneity. All the 100 surveyed papers above clustered standard errors by observable characteristics, such as firm and industry, but not by unobserved characteristics, such as managerial characteristics.

Our central research question is how studies in finance and accounting should account for time-varying unobserved heterogeneity. The current literature mainly controls for time-varying heterogeneity using only time dummies and few use interacted fixed effect (see Gormley and Matsa, 2014). While using time dummies provides an easy and quick solution to the time-varying unobserved heterogeneity problem, the efficacy of the method depends on the true data generating process. A heterogenous effect of a market-wide shock will lead to biased coefficient estimates.

We first investigate the existence of group structure in firms' responses to market-wide events in the current empirical literature. We find that the heterogenous responses of firms to market-wide events are related to industry, managerial characteristics, and firm characteristics such as governance and leverage. This suggests the need to consider firms as part of a group based on observable and

[4]In particular, one should appropriately account for time-series dependence due to the residual of a given firm being correlated across time. In addition, residuals in a given year might be correlated across different firms. Petersen (2009) and Thompson (2011) provide finance scholars with guidance of how to do this appropriately.

unobservable characteristics, rather than just industry or size, as is the current practice. This is consistent with Gormley and Matsa (2014) that unobserved heterogeneity is common within groups like size, local economic environment, management quality, among many others.

One way to group firms by unobserved characteristics is to exploit the advances in the statistical and econometric literature in clustering. In particular, we consider grouping firms whose time profiles of outcomes, net of the effect of covariates, are most similar. Accounting for the net effect of covariates simplifies the estimation problem to the standard minimum sum-of-squares partitioning problem, which can be estimated by the "kmeans" algorithm (Forgy, 1965; Steinley, 2006). The grouping of firms allows for a simple iterative procedure for the GFE estimation.[5]

To study the bias in the estimated parameter of interest for two-way and interacted fixed effect models, we theoretically derive the bias in the estimate and the standard errors under the data structure with time-varying group heterogeneity. We find that the degree of bias depends on the correlation between the grouped time fixed effect and the explanatory variables, as well as on the relative size of the heterogeneity with respect to the variation of the treatment variable X. Particularly, the bias of coefficient estimates is high when heterogenous responses to market-wide shocks are large and highly correlated with X. If the heterogeneous time fixed effect is independent from X, the slope coefficients can still be consistently estimated. However, existing empirical findings that firms have heterogenous responses to market-wide events suggest that such bias is likely to be present.

We then conduct Monte Carlo simulations to study the efficacy of existing approaches, time-fixed effect and interacted fixed effect models, across multiple data generating processes (DGPs). The simulation study allows us to examine the efficacy of GFE and fixed effect models under different conditions of the market-wide shock frequency and size, as well as the heterogeneity of firms' responses to shocks. We find that using time fixed effect and interacted (industry-time) fixed effect models exhibits a substantial bias in the estimates of treatment variables of interest, when there are heterogenous responses to market-wide shocks. On the other hand, GFE produces consistent estimates under the condition of both homogenous and heterogenous shocks. These results are not sensitive to the frequency and size of market-wide shocks. In addition, we also find bias in the standard errors due to cross-sectional dependence in the time-varying unobserved

[5] See Bonhomme and Manresa (2015a) for asymptotic and finite-sample properties of the GFE estimator.

heterogeneity, when scholars fail to cluster the standard errors appropriately.

We estimate the GFE and the fixed effect models using a real data set and compare their relative performance. Specifically, we investigate the relation between CEO total compensations and characteristics and the relation between CEO incentives and risk taking behavior. In both exercises demonstrates that the methods used in some published papers may produce large biases in the estimate of the treatment effect of interest and in the standard errors. This suggests that appropriately modeling time-varying heterogeneity and estimating standard errors is very important. Examining actual data demonstrates how differences in approaches in modeling time-varying heterogenous unobserved effects and standard error estimates can provide information about the deficiency in a model and directions for improving it.

Overall, our results challenge the standard assumptions of empirical estimation in the presence of unobserved group heterogeneity in empirical finance research, by showing the impact of heterogenous responses to market-wide events on empirical studies. Our paper shows that ignoring time-varying unobserved heterogeneity produces inconsistent estimates and distorts statistical inference. This paper extends the work of Gormley and Matsa (2014), which provides practical guidance on how to handle unobserved group heterogeneity in empirical research. Our paper is inline with the burgeoning work on econometric challenges faced by researchers in empirical methods. Bertrand et al. (2004); Petersen (2009); Thompson (2011) discuss methods on appropriately computing standard errors in the presence of cross-sectional and time series dependence across residuals. Erickson and Whited (2012) and Koh, Reeb, Sojli and Tham (2016) discuss issues on measurement error in investment regressions and missing data problem in R&D and patents.

2 Unobserved Heterogeneity

2.1 Two way fixed effect model

We begin by discussing the most widely used method in accounting for unobserved heterogeneity, the two-way fixed effects model with individual and time fixed effects. In particular, we consider the following model:

$$y_{it} = \alpha_i + \lambda_t + X_{it}'\beta + \epsilon_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T \tag{1}$$

where y_{it} is the dependent variable for individual i at time t, α_i captures the individual specific fixed effects, λ_t models the time fixed effects, X_{it} is a $K \times 1$ vector of regressors, and ϵ_{it} is the error term. With strict exogeneity, it is typically assumed that

$$\begin{aligned}
&\epsilon_{it} \sim \text{i.i.d } (0, \sigma^2_\epsilon),\\
&\text{var}(X_{it}) = \sigma^2_X,\\
&\text{cov}(\epsilon_{it}, X_{it}) = 0,\\
&\text{cov}(\alpha_i, X_{it}) \neq 0,\ \text{cov}(\lambda_t, X_{it}) \neq 0,\\
&\text{cov}(\epsilon_{it}, \alpha_i) = 0,\ \text{cov}(\epsilon_{it}, \lambda_t) = 0.
\end{aligned}$$

These assumptions imply that the regressors, individual fixed effects and time fixed effects are uncorrelated with errors. However, regressors can be arbitrarily correlated with the individual and time fixed effects as in Gormley and Matsa (2014). As pointed out by Gormley and Matsa (2014), ignoring the nonzero covariance of the unobserved characteristics and the explanatory variable of interest results in inconsistent estimates. For notation convenience, we assume that the means of the fixed effects and of independent variables are zero (i.e., $\mu_X = \mu_\alpha = \mu_\lambda = 0$); these assumptions simplify the analysis but do not influence the estimates of β.

To estimate this model, we define $y_i = (y_{i1}, \dots, y_{iT})'$ and $X_i = (X_{i1}, \dots, X_{iT})'$, for a unit i over T time periods (time-series dimension). Furthermore, we define $y = (y_1, \dots, y_N)'$ and $X = (X_1, \dots, X_N)'$ across N cross-sectional units (cross-sectional dimension).[6] Two-way fixed effect estimators can be obtained by applying least square on the transformed data, data that is demeaned both in the cross-section and time series dimension. In particular, we define the transformation matrix T as:

$$T = I_N \otimes I_T - I_N \otimes \bar{J}_T - \bar{J}_N \otimes I_T + \bar{J}_N \otimes \bar{J}_T, \tag{2}$$

where I_N is an identity matrix of dimension N, I_T is an identity matrix of dimension T, J_T is a matrix of ones of dimension T, J_N is a matrix of ones of dimension N, $\bar{J}_N = J_N/N$, $\bar{J}_T = J_T/T$ and $\otimes$ is the Kronecker product. The transformed data $\widetilde{y} = Ty$ and $\widetilde{X} = TX$ have typical elements as $\widetilde{y}_{it} = (y_{it} - \bar{y}_{i\cdot} - \bar{y}_{\cdot t} + \bar{y}_{\cdot\cdot})$ and $\widetilde{X}_{it} = (X_{it} - \bar{X}_{i\cdot} - \bar{X}_{\cdot t} + \bar{X}_{\cdot\cdot})$, respectively, where $\bar{y}_{i\cdot} = \sum_t y_{it}/T$,

[6] A unit can be thought of as a firm or an individual.

$\bar{y}_{\cdot t} = \sum_i y_{it}/N$, $\bar{y}_{\cdot\cdot} = \sum_i \sum_t y_{it}/NT$, and similar notations apply to X. The slope coefficient can then be estimated by applying least square on the transformed data:

$$\widehat{\beta}_{\mathrm{TFE}} = (X'TX)^{-1}X'Ty.$$

In the two-way fixed effect model, the effect of regressors of interest β can be consistently estimated by within individual-time transformation, if the data generating process coincides with model (1). A key assumption of model (1) is that all individual units are exposed to the same time-varying shock, λ_t. This assumption is vulnerable in practice because of the heterogeneity of individual units. In particular, when the dependent variable y_{it} is influenced by some unobserved shocks, units typically respond to such shocks differently. However, it is also often observed that there are similar behavior and responses to shocks within groups of units, even though there are differences across groups. In this case, imposing the homogeneous time fixed effect as in model (1) leads to biased estimates of the interested variables, because the within transformation cannot completely eliminate the heterogeneous time fixed effect.

2.2 Bias of estimates in homogenous time fixed effect models

To illustrate the mechanism behind the bias and to determine the size of the bias, we use the following example with two groups of units. Within the group, individual units have the same response to shocks, while the response differs across the two groups. Then the data generating process can be written as:

$$y_{it} = \alpha_i + \theta_{g_i,t} + X_{it}'\beta + \epsilon_{it}, \quad g_i = 1, 2. \tag{3}$$

In model (3), the outcome variable is affected by heterogeneous time-varying shocks $\theta_{1,t}$ for group 1 and by $\theta_{2,t}$ for group 2. Without loss of generality, we can decompose $\theta_{g_i,t}$ as:

$$\theta_{1,t} = \bar{\theta}_t + (\theta_{1,t} - \bar{\theta}_t) \quad \text{for} \quad g_i = 1$$

and

$$\theta_{2,t} = \bar{\theta}_t + (\theta_{2,t} - \bar{\theta}_t) \quad \text{for} \quad g_i = 2,$$

where $\bar{\theta}_t$ can be interpreted as the common time fixed effect. A simple way to extract the common component of the time fixed effect is to use the arithmetic average, i.e. $\bar{\theta} = (\theta_{1,t} + \theta_{2,t})/2$. In this case, we can rewrite the group-wise time fixed effect as:

$$\theta_{1,t} = \bar{\theta}_t + (\theta_{1,t} - \theta_{2,t})/2 \quad \text{and} \quad \theta_{2,t} = \bar{\theta}_t + (\theta_{2,t} - \theta_{1,t})/2.$$

The within individual-time transformation can only eliminate $\bar{\theta}_t$, while $\theta_{1,t} - \theta_{2,t}$ and $\theta_{2,t} - \theta_{1,t}$ remain for each group, which influence the coefficient estimates as omitted variables. To quantify this bias, we define $\widetilde{\theta}_{g_i,t} = I(g_i = 1)(\theta_{1,t} - \theta_{2,t})/2 + I(g_i = 2)(\theta_{2,t} - \theta_{1,t})/2$, where $I(\cdot)$ is an indicator function. Then the bias of the two-way fixed effect estimator can be obtained by:

$$b_{\text{TFE}} = E\left[\widehat{\beta}_{\text{TFE}} - \beta\right] = \left[\sum_{i=1}^{N}\sum_{t=1}^{T} \widetilde{X}_{it}'\widetilde{X}_{it}\right]^{-1} \left[\sum_{i=1}^{N}\sum_{t=1}^{T} \widetilde{X}_{it}'\widetilde{\theta}_{g_i,t}\right].$$

We note that the degree of bias clearly depends on the correlation between the grouped time fixed effect and the regressors, as well as on the relative size of the heterogeneity $(\theta_{2,t} - \theta_{1,t})/2$ with respect to the variation of X_{it}. Particularly, the bias of coefficient estimates is high when $(\theta_{2,t} - \theta_{1,t})/2$ is large and highly correlated with X_{it}. If the heterogeneous time fixed effect is independent from X, the slope coefficients can still be consistently estimated. However, existing empirical findings that firms have heterogeneous responses to market-wide events suggest that such bias is likely to be present. For example, Campello et al. (2010, 2011)'s findings that the 2008 financial crisis has different effects on firms due to unobservable characteristics, such as managerial liquidity and credit management qualities, implies that the heterogeneous time fixed effect is unlikely to be independent from X, especially when the treatment variable X is likely to be a function of managerial characteristics.

Table 1 provides a numerical example of how omitted variable bias is caused. In this example, there are four individual units in total, and they are realizations of two population groups. Individual 1 and 2 belong to the same group in the sense that their response to shocks is homogeneous $\theta_{1,t}$, while 3 and 4 are in the other group, sharing a common $\theta_{2,t}$ that is different from $\theta_{1,t}$. The original data is presented in columns (4)-(7). If we implement the two-way fixed effects transformation, i.e. demean cross-sectionally and in the time series, the resulting data are given in columns (8)-(11) of Table 1. It is clear that the transformed time fixed effect $\widetilde{\theta}_{g,t}$ is not a zero vector, suggesting

Table 1
Demonstrating example for transformation

The table presents a simple numerical example of how omitted variable bias is caused by ignoring heterogeneity in responses to shocks, when using a two-way fixed effects model.

	Unit		Original data				Transformed data			
Indiv. ID	Time	Group ID	α_i	$\theta_{g,t}$	X	y	$\widetilde{\alpha}_i$	$\widetilde{\theta}_{g,t}$	$\widetilde{X}$	$\widetilde{y}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	1	1	2	3	...	...	0	1	...	...
1	2	1	2	4			0	0		
1	3	1	2	5			0	-1		
2	1	1	4	3			0	1		
2	2	1	4	4			0	0		
2	3	1	4	5			0	-1		
3	1	2	6	3			0	-1		
3	2	2	6	6			0	0		
3	3	2	6	9			0	1		
4	1	2	8	3			0	-1		
4	2	2	8	6			0	0		
4	3	2	8	9			0	1		

that the time fixed effect is still not completely wiped out by the transformation. Therefore, if we estimate the slope parameters simply using transformed $\widetilde{X}$ and $\widetilde{y}$, we fail to incorporate the effect of $\widetilde{\theta}_{g,t}$, which is non-zero and possibly correlated with $\widetilde{X}$, and thus suffer from omitted variable bias.

Note that if the heterogeneous time fixed effect only differs in a constant shift, then the within individual-time transformation can successfully wipe out both the individual and time fixed effect (see example in Table 2). This means that if the cross-sectional difference in their response to shocks can be explained by the individual fixed effect, or in other words, the cross-sectional difference of response is time-invariant, then the coefficient estimates are still consistent. In practice, time-varying heterogeneity seems a more plausible situation since individuals experience different types of shocks over time and the degree of heterogeneity in responses often depends on the feature of shocks.

2.3 Time-varying heterogeneity: Group fixed effect model

One way of modeling time-varying heterogeneity is to consider a linear panel data model, where the additive fixed effects have a group structure, namely grouped fixed effects (GFE), as in Bonhomme and Manresa (2015a). They propose a *K-means* algorithm to estimate the model and to study its

Table 2
Demonstrating example: Time-invariant heterogeneity

The table presents a simple numerical example of the two-way fixed effect absorbing the time invariant unobserved heterogeneity.

			Original data		Transformed data	
Indiv. ID	Time	Group ID	α_i	$\theta_{g,t}$	$\tilde{\alpha}_i$	$\tilde{\theta}_{g,t}$
1	1	1	2	3	0	0
1	2	1	2	4	0	0
1	3	1	2	5	0	0
2	1	1	4	3	0	0
2	2	1	4	4	0	0
2	3	1	4	5	0	0
3	1	2	6	4	0	0
3	2	2	6	5	0	0
3	3	2	6	6	0	0
4	1	2	8	4	0	0
4	2	2	8	5	0	0
4	3	2	8	6	0	0

asymptotic properties.

The idea of "grouped fixed effects" is consistent with the financial and accounting empirical literature that the heterogeneous responses to market-wide shocks are due to unobserved time-varying heterogeneity, such as unobservable managerial characteristics and corporate culture.

The GFE model precisely coincides with the DGP in model (3) and can be formulated more generally for any fixed number of groups G as:

$$y_{it} = \alpha_i + \theta_{g_i,t} + X_{it}'\beta + \epsilon_{it}, \quad g_i \in \{1, \dots, G\}. \tag{4}$$

This model differs from the standard fixed effect model, because it includes a time-varying group specific variable $\theta_{g_i,t}$, in addition to individual fixed effect α_i. There are two types of parameters to estimate, the group membership parameter g_i and the standard regression parameters α_i, $\theta_{g_i,t}$, and β. The individual fixed effect captures the individual firm's characteristics, while the grouped-fixed effect parameter $\theta_{g_i,t}$ models how firms respond to common shocks that vary over time. This response is common within a group but differs across groups.

If the group pattern is known, we can obtain consistent estimates of regression parameters θ and

β by applying least squares on the following model with interactions of population group dummies and time dummies

$$\dot{y}_{it} = \dot{X}_{it}'\beta + (GD_i \times TD_t)\theta_{g_i,t} + u_{it}, \tag{5}$$

where the transformed data $\dot{y}_{it} = y_{it} - 1/T\sum_{t=1}^{T} y_{it}$ and $\dot{X}_{it} = X_{it} - 1/T\sum_{t=1}^{T} X_{it}$ are used to integrate out individual fixed effects α_i. $GD_i = \{GD_{i1}, \ldots, GD_{iG}\}$ is a (time-invariant) group dummy with $GD_{ig} = 1$ if unit i belongs to group g, and $TD_t = \{TD_{t1}, \ldots, TD_{tT}\}$ contains (individual-invariant) time dummies for each period. However, in practice, the group membership is typically unknown, so that precise group dummies are not readily available. Therefore, the estimator obtained from the "infeasible" model (5) is called infeasible estimator. In real applications, if we misspecify the group structure, this approach leads to biased and inconsistent estimates.

To consistently estimate parameters in (4) without a priori knowledge of the group membership, one can employ the least square technique and jointly estimate the group and regression parameters by solving the following minimization problem:

$$Q_{NT} = \min_{\theta,\beta,g} \sum_{i=1}^{N}\sum_{t=1}^{T} (\dot{y}_{it} - \dot{X}_{it}'\beta - \theta_{g_i,t})^2. \tag{6}$$

The least square objective function in (6) can be solved by an iterative procedure suggested by Bonhomme and Manresa (2015a) as follow:

1. Let $g^{(0)}$ be an initial value of grouping. Set $s = 0$.

2. For the given $g^{(s)}$, compute:

$$(\theta^{(s+1)}, \beta^{(s+1)}) = \arg\min_{\beta,\theta} \sum_{i=1}^{N}\sum_{t=1}^{T} (\dot{y}_{it} - \dot{X}_{it}'\beta - \theta_{g_i^{(s)},t})^2.$$

3. Compute for all $i \in \{1, \ldots, N\}$:

$$g_i^{(s+1)} = \arg\min_{g\in\{1,\ldots,G\}} \sum_{t=1}^{T} (\dot{y}_{it} - \dot{X}_{it}'\beta^{(s+1)} - \theta_{g_i,t}^{(s+1)})^2.$$

4. Set $s = s + 1$ and go to Step 2 (until numerical convergence).

This algorithm iterates between estimating the regression parameters and group membership parameters and can be viewed as an EM-type algorithm (see, for example, Dempster et al., 1977). Specifically, Step 2 estimates the regression parameters for a given group structure as in the usual least squares problem. Then Step 3 finds the optimal group in terms of minimum sum of squared residuals over time for each individual, based on the estimated regression parameters from the previous step. This algorithm typically converges quickly, but it clearly depends on the chosen initial values. To avoid local optima, a large number of initial values should be tried and the one yielding the lowest sum of squared residuals is selected.

Bonhomme and Manresa (2015) show that this "grouped fixed effects" approach has the asymptotic "oracle" property that the estimated group membership of each individual converges to the true population membership as both dimensions of the panel increase. Given the consistently estimated group membership, the estimates of regression parameter $\theta_{g,t}$ and β are consistent and asymptotically equivalent to the infeasible estimate from model (5), as if we knew the true group membership. Bonhomme and Manresa (2015) suggest that this approach should yield good finite sample performance, even when some units have as few as 10 observations.

Several important issues arise when applying the GFE estimator, namely parameter identification, standard error estimation, and specification of the number of groups. We discuss these issues in turn. First, to identify the grouped specific parameter, we need the group separation condition that

$$\text{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(\theta^0_{g,t}-\bar{\theta}^0_{g,t}-\theta^0_{\widetilde{g},t}+\bar{\theta}^0_{\widetilde{g},t})^2>0, \quad \text{for} \quad g\neq\widetilde{g},$$

where $\bar{\theta}^0_{g,t} = 1/T\sum_{t=1}^{T}\theta^0_{g,t}$ and $\theta^0_{g,t}$ denotes the true value of $\theta_{g,t}$. This implies that group heterogeneity exists after controlling for individual specific effects. In other words, we require that the group specific responses to shocks are not time-invariant (see Table 2). In fact, time-varying heterogeneity seems a more plausible situation since individuals experience different types of shocks over time and the degree of heterogeneity in responses often depends on the feature of shocks.

Second, we discuss the standard error of the slope coefficient estimates. Asymptotically, the GFE estimator is equivalent to an infeasible consistent estimator from model (5), as if the true group is used. Therefore, if a large number of time series observations is available, we can estimate standard errors using the sample analogue of the asymptotic robust variance, namely the robust variance of the least squares estimator for each group. To compute this variance, let $\bar{x}_{g^0_i,t}$ be the mean of $\dot{x}_{it}$ in group $g_i^0 = g$, and define:

$$\Sigma_\beta = \text{plim}_{(N,T)\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\dot{X}_{it} - \bar{\dot{X}}_{g_i^0,t})(\dot{X}_{it} - \bar{\dot{X}}_{g_i^0,t})',$$

$$\Omega_\beta = \lim_{(N,T)\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} E\left[\epsilon_{it}\epsilon_{js}(\dot{X}_{it} - \bar{\dot{X}}_{g_i^0,t})(\dot{X}_{js} - \bar{\dot{X}}_{g_j^0,s})'\right].$$

Then, we can obtain the asymptotic variance of $\widehat{\beta}_{\text{gfe}}$ as:

$$var(\widehat{\beta}_{\text{gfe}}) = \frac{1}{NT} \Sigma_\beta^{-1} \Omega_\beta \Sigma_\beta^{-1}.$$

Note that this variance is robust to heteroscedasticity and correlation across groups and time, but it does not take into account the uncertainty caused by estimating group membership. Given that the group membership estimator converges very quickly as the number of time periods increases, asymptotic variance can often be a good approximate of variance. However, if the panel is short with small T, one should account for error caused by the group estimation, because misclassification may increase the small-sample dispersion of the estimator. A possible way to account for the uncertainty caused by the group estimation is to use a bootstrap standard error estimate based on re-sampling unit blocks of (y_i, X_i) from the complete sample. This is easy to implement but computationally more intense.[7]

Finally, we discuss the specification of the number of groups G. We first consider the consequences of misspecifying the number of groups, and then discuss how to determine this number. When the postulated number of groups is less than the true number, the GFE estimator of slope coefficients $\widehat{\beta}_{\text{gfe}}$ is inconsistent, if the heterogeneous responses are correlated with regressors. Nevertheless, if one over-specifies the number of groups, he or she can still achieve a consistent estimator of β, but probably with less efficiency. Because of these properties, we may determine the number of groups by examining the values of $\widehat{\beta}_{\text{gfe}}$ for different choices of G. In particular, since $\widehat{\beta}_{\text{gfe}}$ remains consistent for an over-specified G but not for an under-specified G, we should expect that $\widehat{\beta}_{\text{gfe}}$ converges to a certain value and remains relatively stable as G increases. Thus, we can choose

[7] Alternatively, one can also compute an analytical version of finite-sample standard errors using the formulas provided in Bonhomme and Manresa (2015b).

G, when $\widehat{\beta}_{\text{gfe}}$ is stable. To determine G, one can either use information criteria as suggested by Bonhomme and Manresa (2015a) and Su et al. (2016), or via hypothesis testing as proposed by Lu and Su (2017). In this paper, we focus on the Bayesian information criterion defined as:

$$\text{BIC}(G) = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}(\dot{y}_{it} - \dot{X}_{it}'\widehat{\beta}_{\text{gfe}}^{(G)} - \widehat{\theta}_{g_i,t}^{(G)})^2 + \widehat{\sigma}^2\frac{GT+K+N}{NT}\log(NT), \tag{7}$$

where $\widehat{\beta}_{\text{gfe}}^{(G)}$ and $\widehat{\theta}_{g_i,t}^{(G)}$ are the GFE estimators with G groups and $\widehat{\sigma}^2$ is a consistent estimate of the variance of ϵ_{it}, e.g. the sum of squared residuals obtained from $G_{\max}$ groups.

In Section 4, we evaluate the efficacy of two-way fixed effect, firm with interacted industry-time fixed effect, and group fixed effect models in a simulation study, before comparing their relative performance using two real data sets.

3 Alternative methods to model time-varying heterogeneity

In this section, we discuss alternative ways of modeling time-varying heterogeneity, some of which are widely studied in econometrics but not yet popular in finance, while some have been extensively used in finance applications.

3.1 Interacted fixed effect model

A popular approach to control for time-varying heterogeneity is to include interactions of group and time fixed effects (see e.g. Gormley and Matsa, 2014) as in model (5). We follow Gormley and Matsa (2014) and refer to it as an interacted fixed effects model. Typically, the grouping is based on observed characteristics, such as industry and/or the size of firms etc.. If the postulated grouping is correct, this model produces consistent estimates of slope coefficients, when the data generating process coincides with that in model (4), i.e. the group is correct and the model is also GFE. One issue with this approach is that the group structure may not be exactly driven by a small set of observables, but often may be driven by unobserved characteristics or a mixture of observables and unobservables. If the grouping is misspecified, this interacted fixed effects model leads to inconsistent slope coefficient estimates, if group-specific heterogeneity is correlated with regressors.

For time fixed effects, a simple way is to include one dummy for each time period, assuming that there is a time effect in each period. Over-specifying time effects does not cause inconsistency, but only affects efficiency. The efficiency loss is generally negligible for financial data, since there is sufficient cross-sectional variation with large N. In Section 4, we investigate numerically how the misspecification of groups and shocks affects the estimate of the interacted fixed effects model.

3.2 Event studies

Another popular method to control for heterogeneous responses in finance studies is event studies. There are two potential problems with event studies. First, event studies require that there are relatively few shocks over the entire time span, and the timing of the shocks needs to be precisely specified. However, in some cases, there may be many shocks over the whole sample period, or simply different time effects in each period. Then event studies are less reliable, since the time periods within each subsample is short. More importantly, if the timing of shocks is incorrectly identified or several shocks are ignored, event studies generally result in inconsistent estimates.

Second, the conventional standard errors provided by event studies are often inappropriate, because they do not take into account cross-sectional correlation. The group specific responses suggest that there exist an unobserved group pattern of features, e.g. managerial and institutional characteristics, that drives heterogeneity in responses, and it is likely that these characteristics will also affect firm's responses to any kind of corporate related events. Therefore, we can expect unobserved group characteristics in both the independent variables and the residuals that are possibly changing over time. These time-varying group components cannot be canceled out by the fixed effects transformation.

Derivations in the appendix show that the fixed effects standard error estimates underestimate the true standard errors. To obtain appropriate standard error estimates in event studies, one should explicitly take into account the latent group pattern, and calculate the group-clustered standard error. Our GFE estimates of the group membership thus provide a natural and reliable estimate of the latent group pattern, which can be used to construct clustered standard errors.

4 Simulation

4.1 Data generating process

We generate the data from the linear panel data model with time-varying and shocks (homogeneous and heterogeneous):

$$y_{it} = \alpha_i + \theta_{g_i,t} + \sum_{l=1}^{2} X_{l,it}\beta_l + \epsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T; \quad \epsilon_{it}\ N(0, \sigma_\epsilon^2 \tag{8}$$

where $\alpha_i = 1/T \sum_{t=1}^{T} X_{1,it}$, $X_{it} = (X_{1,it}, X_{2,it})'$ is a vector of regressors with loadings $\beta = (\beta_1, \beta_2)' = (1, 2)'$. The idiosyncratic errors ϵ_{it} are independently and normally distributed with mean zero and variance σ_ϵ^2. The major difference between equation (8) and standard panel data models lies in $\theta_{g_i,t}$ for $g = 1, \ldots, G$ with $G = 5$ that captures the group-wise heterogenous response to time-varying shocks. To generate the correlation between $\theta_{g_i,t}$ and X_{it}, we let them both be driven by some group components, i.e.

$$X_{l,it} = c_\tau \tau_{l,g_i,t} + \nu_{it}, \quad \nu_{it}\ N(1, \sigma_\nu^2 = 5)$$

and

$$\theta_{g_i,t} = g_i^2(\tau_{1,g_i,t} + \tau_{2,g_i,t})/c_\theta, \quad \text{for } l = 1, 2,$$

where $\tau_{l,g_i,t}$ is the group component drawn from $G(= 5)$ different normal distributions with mean g and variance $2g$ for $g = 1, \ldots, G$, and ν_{it} is the idiosyncratic component drawn from the normal distribution with mean 1 and variance $\sigma_\nu^2 = 5$. The constant c_τ reflects the importance of the group component in X and further determines the degree of correlation between X and θ. The constant c_θ controls the size of shocks. To make the variation of error terms comparable to that of the regressors, we set $\sigma_\epsilon^2 = 5$.[8]

Since the nature and characteristics of shocks play an important role for the consistency of each methods, we consider three different types of shocks:

DGP1 (Frequent shocks): Shocks occur every time period t.

DGP2 (Sparse shocks): Shocks occur only 25% of the time.

[8]Setting these variances to 5 is not essential, but they have to be of comparable sizes to control the size of the shocks and the signal to noise ratio.

DGP3 (Homogeneous shocks): Common shocks with homogeneous impact on all units.

DGP1 and DGP2 allow for heterogeneous responses to shocks. DGP2 contains less shocks than DGP1, and from theory we should expect the bias of TFE in DGP2 to be smaller than that in DGP1. DGP3 assumes that the effect of shocks is homogeneous across units, and therefore it is the same data generating process as the two-way fixed effect models with only individual and time fixed effects. Hence, GFE and TFE are expected to produce similar results under DGP3. Within each DGP, we also consider different choices of key parameters, i.e. c_τ, c_θ, and the number of groups G. These extensions help us to better understand how estimation bias is affected by increasing the correlation between regressors and heterogenous responses, increasing the size of shocks, and increasing the number of groups. In addition, we also consider the case of misspecifying G for GFE. To mimic typical empirical data in corporate finance studies (including the two empirical data sets analyzed in Section 5 with ($N = 1859$, $T = 16$) and ($N = 1627$, $T = 23$), respectively), we consider panels with large N and moderate T. In particular, our simulation here is conducted under three different sample sizes, ($N = 1000$, $T = 10$), ($N = 2000$, $T = 10$), and ($N = 1000$, $T = 20$). These sample sizes allow us to examine how the dimension of N and T affects the relative performance of various methods.

4.2 Method comparison and evaluation

We compare three methods: grouped fixed effects (GFE), standard two-way fixed effects (TFE), and group-time interacted fixed effects (IFE). The major difference between GFE and IFE lies in that the grouping in IFE is specified exogenously by researchers, typically based on industry classification, while the group pattern in GFE is determined by the data and is case specific. The group pattern produced by GFE depends on the nature of the shock and data characteristics, i.e. the same sample of firms will be grouped differently for different shocks. The details of these methods are discussed in Section 2. The validity of IFE obviously depends on the specification of groups and shocks, and we consider four scenarios for IFE. Scenario 1 exactly specifies the group memberships of all units and the time of all shocks (denoted as IFE$_{gt}$), which is hardly achieved in practice. Therefore, we also refer to this as an infeasible estimator. Scenario 2 exactly specifies the groups but not the shocks (denoted as IFE$_g$). Scenario 3 correctly specifies the shocks but not the groups (denoted as

IFE_t). Scenario 4 misspecifies both groups and shocks (denoted as IFE_{null}). When the shocks are misspecified, we generate a time dummy at every 1/3 of the time (IFE_{g2}). When the groups are misspecified, we assign 30% of units to the closest neighboring groups.

We evaluate all methods based on three criteria, the bias, the estimated standard deviation, and the root mean square error (RMSE), which are computed by:

$$\text{Bias}(\hat{\beta}) = \frac{1}{R}\sum_{r=1}^{R}\hat{\beta}_r - \beta, \quad \text{Std}(\hat{\beta}) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}(\hat{\beta}_r - \bar{\hat{\beta}})^2}$$

and

$$\text{RMSE}(\hat{\beta}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}||\hat{\beta}_r - \beta||^2},$$

where $\bar{\hat{\beta}} = 1/R\sum_{r=1}^{R}\hat{\beta}_r$ is the sample average of $\hat{\beta}$ across replications, and $R = 1000$ is the number of replications. We report the *empirical* standard deviation across replications here because it can capture the uncertainty caused by estimating the unknown group structure.[9] Note that RMSE($\hat{\beta}$) is not precisely the square root of the summation of squared Bias($\hat{\beta}$) and Std($\hat{\beta}$), since Std($\hat{\beta}$) is the *estimated* standard deviation of $\hat{\beta}$ and not the true one.

4.3 Results

In this section, we present the simulation results, which are in line with our theoretical discussion in Section 2. We first present the detailed results for DGP1 – DGP3 under a benchmark parametrization. Then, we consider alternative parametrizations in order to examine how the performance of different methods is affected by increasing the size of the shock, increasing the correlation between responses and regressors, and increasing the number of groups. We also examine how GFE performs when the number of groups is misspecified.

[9]This standard deviation differs from the quantity $\frac{1}{R}\sum_{r=1}^{R}\hat{\sigma}_{\hat{\beta}}$, where $\hat{\sigma}_{\hat{\beta}}$ is the estimated asymptotic standard error of $\hat{\beta}$, which does not take into account the grouping uncertainty. With a moderate time-series dimension, these two standard deviation estimates are equivalent, because the group estimation is super-consistent with the convergence rate of an arbitrarily large exponential order of T (see Theorem 2 of Bonhomme and Manresa (2015a). However, they differ slightly from each other when T is small. Our unreported simulation results (available upon request) suggest that when T is larger than 10, the difference between these two estimates is small.

4.3.1 Frequent shocks

Table 3 presents the bias and the mean square error of estimated coefficients under DGP1. We first consider the parameterization $c_\tau = 0.5$, $c_\theta = 15$, and $G = 5$ as a benchmark. This translates to moderate shocks, occurring 35% of the dependent variable time series, and the correlation between the time aggregated shocks ($\sum_{t=1}^{T} \theta_t$) and X is around 0.3. The results in Panel A show that when units respond differently to shocks, TFE ignores this heterogeneity and results in around 60% bias for β_1 (Bias=0.595 where β_1=1) and 30% bias for β_2 (Bias=0.599 for β_2=2). GFE successfully captures the heterogeneous response and is not biased (Bias=0.013). IFE$_{gt}$ is the best among all estimators as it correctly *specifies* (not estimates) all the groups and the shocks. IFE$_{gt}$ is more accurate than GFE because there is no estimation of groups involved. However, it is very difficult, if not impossible, to always exactly specify all groups and shocks in practice. If there is any misspecification, the performance of IFE is no longer guaranteed. To assess how misspecification influences the IFE estimator performance, we compare IFE$_g$, IFE$_t$, and IFE$_{null}$. Results in Table 3 show that incorrectly specifying the shocks (IFE$_g$) leads to a bias of 0.5, and misspecification in grouping (IFE$_t$) causes a bias of 0.25. Naturally, misspecifying both shocks and groups (IFE$_{null}$) leads to an even larger bias.[10]

Comparing the standard deviation of different estimators, we find that the standard deviation produced by GFE is only slightly larger than that of IFE$_{gt}$, suggesting that estimating the groups does not introduce much extra uncertainty, at least sample sizes of 1,000 observations or more. The standard deviation of TFE is almost twice as large as GFE, and that of IFE under misspecification is also much larger than GFE.

The RMSE of GFE is only slightly higher than that of the infeasible estimator IFE$_{gt}$. Nevertheless, the conventional TFE is highly inaccurate, with the RMSE almost 50 times larger than that of GFE. Incorrectly specifying the groups and shocks also dramatically decreases the performance of IFE (with RMSE more than 45 times larger than that of GFE), and the performance depends on the degree of misspecification.

Panels B and C of Table 3 examine how the sample size affects the performance of different

[10] The impact of misspecification obviously depends on the degree of misspecification. Unreported results (available upon demand from the authors) confirm that if we increase the proportion of individuals that are misclassified or decrease the number of specified shocks, the bias of IFE increases.

methods. Increasing the sample size improves the accuracy of GFE and IFE_{gt}, but not TFE and other versions of IFE. For GFE, increasing the time dimension decreases the bias but increasing the individual dimension does not, because only the time series variation contributes to the accurate classification. In fact, the convergence rate of group estimation is $1/T^{\delta}$, where δ is any arbitrary positive number, and N is not involved in the convergence rate. The standard deviation of GFE decreases equally by increasing N and T, confirming the convergence rate of GFE slope-coefficient estimators to be $1/\sqrt{NT}$. The accuracy of IFE_{gt} also equally improves by increasing any dimension of the sample, since it is asymptotically equivalent to GFE. On the contrary, the RMSEs of TFE and other versions of IFE do not decrease by increasing the sample size. The improvement in GFE and IFE_{gt} performance confirms the consistency of these two estimators.

4.3.2 Sparse shocks

Next, we study the case where shocks occur infrequently, as presented in Table 4. In this DGP, we additionally compare another IFE estimator that correctly specifies groups but over-specifies the shocks by interacting groups with time dummies of each time period, denoted as IFE_{got}. Since the shocks remain the same size but appear less frequently, the heterogeneous responses of the dependent variable to the shocks are of a smaller magnitude. Hence, although the estimators that omit these heterogeneous responses, i.e. TFE, IFE_{g}, IFE_{t}, and IFE_{null}, continue to be biased, the biases are all smaller than in the case of frequent shocks.

GFE remains highly accurate, although its standard deviation is slightly larger than in DGP1. This is because GFE unnecessarily models the shock at each time period while there are actually fewer shocks. This also explains the difference between GFE and IFG_{gt} that exactly identifies both grouping and shocks. Similarly, we find that imposing more shocks (IFE_{got}) does not affect the consistency but only weakens the efficiency. Given sufficient samples in our case, the efficiency loss is ignorable (compared to IFE_{gt}). This result suggests that one could simply include time dummies at all periods, and not necessary specify the time of shocks precisely.

4.3.3 Homogeneous shocks

Previous results have shown that when individuals respond differently to shocks, ignoring such heterogeneity can lead to severely misleading results for conventional two-way fixed effect and difference-in-difference estimation with misspecified groups. A natural question is, if the data generation process is of homogeneous shocks, i.e. the impacts of shocks are homogeneous across individuals, then how much inefficiency will be caused by assuming a relatively complicated model of grouped fixed effect. To answer this question, we generate the data by (8) but set $\mathsf{G} = 1$. Table 5 compares various estimators in this homogeneous-shock case. We see that TFE and IFE_{gt} now perform equivalently well, because both models coincide with the DGP. As for GFE, we consider two cases with estimated number of groups $\hat{\mathsf{G}} = 1$ and $\hat{\mathsf{G}} = 4$. In the first case of $\hat{\mathsf{G}} = 1$, GFE boils down to TFE and IFE_{gt}, and it performs identically to the other two methods. In the second case of $\hat{\mathsf{G}} = 4$, we over-specify the number of groups, which only marginally increases the RMSE. This confirms the theoretical result of Bonhomme and Manresa (2015a) that over-specifying the number of groups does not affect the consistency of the slope coefficient estimators.[11] In such sample sizes, the potential efficiency loss is ignorable. We also observe that when the sample size increases, the increase in RMSE by over-specifying $\widehat{\mathsf{G}}$ becomes smaller as expected.

4.3.4 Larger and more correlated shocks

Next we change the parametrization to investigate how the size of the shocks and their correlation with regressors affect the performance of different methods. We first consider larger shocks ($c_\theta = 5$ and other parameters remaining the same) that take up roughly 60% of the dependent variable on average, and the results are reported in the upper panel of Table 6. As expected, the bias of TFE and misspecified IFE (IFE_g, IFE_t, and IFE_{null}) all increase to a large extent. On the contrary, GFE is not affected by the size of shocks, and it remains as accurate as the infeasible estimator IFE_{gt}.

The bottom panel of Table 6 examines the case when the correlation between shocks and regressors is high. To mimic this situation, we set $c_\tau = 1$ while other parameters keep the same, and this generates correlation around 0.5 and the size of shocks remain 35% of the dependent variable on average. We find that high correlation increases the bias of all estimators including GFE. Even though, GFE remains the most accurate among all feasible estimator, with RMSE twice as much as the infeasible estimator IFE_{gt}. The larger bias of GFE is possibly because the regressors are also characterized by a strong group structure, making it more difficult to identifying the latent group pattern of fixed effects. The RMSEs produced by TFE and IFE_g are both about 100 times larger

[11]However, over-specifying G may lead to inconsistent estimates of group fixed effects.

than that of IFE_{gt}.

4.3.5 Incorrect $\widehat{G}$

From the experiment of DGP3, we have seen that overestimating the number of groups does not affect the consistency of GFE slope coefficient estimates. However, underestimation of G does affect consistency. This is because if the number of groups is under-specified, then some groups will contain heterogeneous individuals. Since individuals within a group are treated homogeneously, this leads to ignoring heterogeneity and a biased GFE estimator. Table 7 presents the case where the true number of groups $G = 4$ is incorrectly specified as $\widehat{G} = 2$ in the GFE estimation. It shows that GFE is biased when G is under-specified, and this bias does not decrease/disappear as the sample size increases. However, we also find that the degree of bias and the RMSE of GFE are still much smaller than those of TFE. The direction of the bias of GFE estimates depends on the correlation between regressors and heterogenous responses.

Based on these results, a recommendation of the best practices in finance would be to choose a larger G, if the number of groups is uncertain and the slope coefficients are of central interest. In general, the cost of over-specification seems much smaller than under-specification, especially for large sample sizes.

4.3.6 More groups

Here we consider the situation where the data are generated by a large number of groups, i.e. $G = 10$ and $G = 20$, respectively. Enlarging the number of groups results in an even large degree of heterogeneity. Therefore, we should expect that the bias of TFE and misspecified IFE should be greater. GFE is expected to remain consistent.[12] Table 8 shows the performance under $G = 10$ and $G = 20$. We see that the RMSEs of TFE and misspecified IFE increase at least 7 times compared to those in DGP1, while the accuracy of GFE is unaffected.

[12] Bonhomme and Manresa (2015b) propose an alternative algorithm (variable neighborhood search) that potentially works more efficiently than the iterative algorithm for high-dimensional data.

5 Empirical application

5.1 Heterogeneous responses to market-wide shocks

Many papers find that personal characteristics, past experiences, and management styles of executives and board members are different and significantly affect corporate policies and culture.[13] Given this heterogeneity, it is not surprising that responses to market-wide shocks, such as regulation changes and financial crises, are different across firms. Given the broad scope of the empirical literature documenting heterogeneous responses to market shocks, below we discuss the work on the Sarbanes-Oxley Act and financial crises, for brevity.

Linck et al. (2009) finds that SOX resulted in heterogeneous demand for directors, changes in the workload, the structure of corporate boards, the liability risk faced by directors, and the composition of the director pool. These changes affect the corporate governance of firms differently. Engel et al. (2007) reports a heterogeneous effect of SOX on ownership and on the decision to go private. Bargeron et al. (2010) finds that SOX has a heterogeneous impact on managerial risk-taking behaviour across firms with different board structure, firm size, and managerial characteristics and incentives. Several recent papers find that SOX has a different impact across firms with different firm size, and find different effects of managerial opportunistic behavior on firm disclosure, financial constraints, pre-cautionary saving, information cost, financial reporting, and managerial compensation practices (see e.g. Chhaochharia and Grinstein, 2007; Heron and Lie, 2007; Koh et al., 2008; Cohen et al., 2008; Carter et al., 2009; Duchin et al., 2010a,b; Brochet, 2010; Banerjee et al., 2015).

Similar results related to heterogeneous responses to market-wide shocks are found in the financial crises literature. Mitton (2002) and Joh (2003) show that firm heterogeneity in for example, corporate governance, affects firm performance and firm profitability, when subjected to shocks like the East Asian financial crisis. Lemmon and Lins (2003) shows that heterogeneity in firm ownership structure affects firm value, investment opportunities, and stock returns during the East Asian crisis. During the recent 2008-2009 financial crisis, Lins et al. (2013) finds that heterogeneity in family ownership affects corporate investments. Campello et al. (2010) and Campello et al. (2011) document heterogeneous firm characteristics, such as size, credit rating, and profitability, affect

[13] See examples in Opler and Titman (1994), Bertrand and Schoar (2003), Graham et al. (2013), Frank and Goyal (2009), Malmendier et al. (2011), Cronqvist et al. (2012), and Malmendier and Tate (2015) among many others.

the use of credit lines, which are critical to corporate policies, during a period of financial crisis. Thakor (2015) and Ho et al. (2016) indicate that managerial characteristic differentials, such as ability, career concerns, and over-confidence, across firms can explain cross-sectional differences in lending, risk-taking behavior, operating performances, CEO turnover, and firm value.

There is a vast literature documenting differential responses of many firm outcomes and of treatment variables to market-wide regulatory and crisis shocks. These results point to heterogeneity in decision making. Next, we examine the efficacy of various approaches in handling time-varying heterogeneity in practice. We estimate the GFE and the fixed effect models using real data sets and compare their relative performance. Specifically, we investigate the relation between CEO total compensation and CEO characteristics and the relation between CEO incentives and their risk taking behavior.

5.2 CEO compensation

In this section, we benchmark our exercise to the empirical example of investigating the relation between CEO compensation and firm and manager characteristics. We compare estimates from the following estimators: OLS, FE, TFE, IFE, and GFE. Our GFE specification for CEO compensation is as follows:

$$\ln(\text{Total compensation})_{it} = \alpha_i + \theta_{g_{i,t}} + X_{it}'\beta + \epsilon_{it}, \;\; g_i \in \{1, ..., G\} \tag{9}$$

where ln(Total compensation)$_{it}$ is the natural log of the total CEO compensation at firm i in year t. X_{it} is a 9×1 vector of variables thought to affect CEO compensation. We include the following commonly used explanatory variables in our regressions: log total assets (TA), lagged market-to-book ratio (MB_{t-1}), contemporaneous (r_t) and lagged stock returns (r_{t-1}), contemporaneous and lagged return on assets (ROA_t, ROA_{t-1}), volatility of daily log stock returns (σ_t), indicator variables for female CEOs ($Female$) and the dual CEO and Chairman role ($CEO\&Chairman$). α_i is the firm fixed effect and $\theta_{g_{i,t}}$ is a time-varying group effect. We exclude both the firm fixed effect and time-varying group effect for the OLS estimator and retain the firm fixed effect for the FE estimator. For TFE, $\theta_{g_{i,t}} = \theta_t$ and for IFE $\theta_{g_{i,t}}$ is industry interacted with year fixed effects.

To implement the GFE estimation, we first need to determine the number of groups G. We consider two evaluation criteria, the Bayesian Information Criterion (BIC) given by (7) and the convergence of slope coefficient estimates. The BIC serves as an effective criterion to tradeoff between the model fitness and the number of parameters (degree of complication of the model), and we select the number of groups that minimizes this criterion. Figure 1(a) depicts how the BIC varies for different numbers of groups in the CEO compensation regression 9. The value of BIC decreases dramatically when we increase the number of groups from 1 to 10, and reaches its minimum when $\mathsf{G} = 20$. From $\mathsf{G} = 20$ onwards, increasing the number of groups does not significantly improve the model fitness, while introducing a much larger set of parameters. Next, we examine the behaviour of slope coefficient estimates as we increase the number of groups. Figure 2 depicts the GFE coefficient estimates (the dashed lines represent the 95% confidence intervals) for G ranging from 1 to 60. We find that for highly significant variables, such as log total assets, stock returns, and volatility of daily log stock returns, the estimated coefficients converge to a constant and become stable, when G is larger than 20. The estimates of insignificant variables, such as lagged market-to-book ratio and CEO-chairman indicator are less stable, even when G is as large as 45, due to their large degree of uncertainty. In general, both criteria suggest that 20 groups is sufficient to obtain consistent estimates for the effect of the determinants of managerial compensation. Thus, we focus on comparing the results of GFE with $\mathsf{G} = 20$ to those of the alternative methods discussed above.

Table 9 presents the coefficient estimates of the determinants of CEO compensation obtained from OLS, fixed effect (FE), two-way fixed effect (TFE), interacted fixed effect (IFE) based on industry grouping, and grouped fixed effect (GFE) based on $\mathsf{G} = 20$. In general, we observe non-trivial differences between GFE and other estimators. Particularly, the coefficient estimates for total assets are significant for all methods, but OLS, FE, and IFE produce estimates 11%, 25%, and 6% larger than those of GFE, respectively. TFE gives the closest estimate to GFE, but it still is about 2% larger. The effect of contemporaneous stock returns reported by GFE is generally larger than estimates obtained from other methods. We find that GFE produces a much smaller estimate than any other method of the effect of (both contemporaneous and lagged) return on assets, and the lagged effect even switches from significant to insignificant by using GFE. Also GFE produces an estimate half of that of other methods for the estimated effect of volatility.

To examine the significance of the discrepancy among various methods, we test the difference in coefficient estimates. Since the GFE estimates of slope coefficients behave asymptotically equivalent to least square estimates under true group structure (due to super-consistency of group membership estimation, see Corollary 1 of Bonhomme and Manresa (2015a)), we can test the difference between two normally distributed estimates using a standard t test. Assuming the two estimates are independent, the t statistic for the null hypothesis $H_0 : \hat{\beta}_k^* = \hat{\beta}_k^{GFE}$ can be written as

$$t = \frac{\hat{\beta}_k^* - \hat{\beta}_k^{GFE}}{\sqrt{V_k^* + V_k^{GFE}}}, \quad \text{for } k = 1, \dots, 9$$

where $\hat{\beta}_k^*$ and V_k^* represent the estimated coefficient and its variance of the k-th regressor using OLS, FE, TFE, or IFE, and $\hat{\beta}_k^{GFE}$ and V_k^{GFE} stand for the GFE estimated counterparts. The discrepancy in the effect of total assets and lagged stock returns is strongly significant for GFE versus OLS and FE (p-value of the t statistics lower than 0.02), but insignificant for GFE versus TFE and IFE. Due to relatively large standard deviation, the difference in the effect of contemporaneous and lagged return on assets is only weakly significant or insignificant. The GFE estimate of the volatility effect significantly differs from that of all other methods (p-value ranging from 0.003 to 0.085). As for the female dummy, GFE produces a significantly different estimate from that of OLS and FE (p-value lower than 0.1) and weakly significant from IFE (p-value around 0.2). If the two estimated coefficient $\hat{\beta}_k^*$ and $\hat{\beta}_k^{GFE}$ are positively correlated (which is highly likely because they are obtained from the same data), then the t statistic would be even larger, indicating more significant difference between the estimates.

Given the significant differences in the point estimates, it is important to understand the mechanism behind why FE, TFE and IFE produce estimates that are different from GFE. For this purpose, we first examine how firms across groups respond differently to shocks. We plot the time-varying group fixed effects $\theta_{g,t}$ in Figure 3. From these results, we see that responses to shocks vary considerably across firms. Failing to take into account this heterogeneity, such as TFE, leads to bias in the estimated effects of the variables of interest.

Next, we compare the groups estimated by GFE with the industry grouping. We see that the grouping produced by GFE differs substantially from the industry classification. Hence, IFE based on industry grouping cannot completely capture the heterogeneous responses, thus producing biased estimates as well.

Finally, we examine how the direction of the bias is explained by the theoretical results in

Section 2.2. For example, for the TFE estimator, the bias comes from the correlation of the heterogeneous responses to market-wide shocks and variables of interest as

$$b_{\mathrm{TFE}} = E\left[\widehat{\beta}_{\mathrm{TFE}} - \beta\right] = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{X}_{it}'\widetilde{X}_{it}\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{X}_{it}'\widetilde{\theta}_{g_{i,t}}\right].$$

Since $\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{X}_{it}'\widetilde{X}_{it}$ is a quadratic term, which is asymptotically positive definite, the sign of the bias depends on the correlation between the explanatory variables and the group fixed effects. For example, if female CEOs react more strongly to market-wide shocks than male CEOs, one would expect a upward bias of TFE as suggested by the above equation. To illustrate the economic mechanism of the bias, we present the correlation between $\theta_{g,t}$ and all explanatory variables in Equation 9 in Panel A of Table 10. We find the signs of correlations are precisely in line with the direction of bias. In particular, we find a significantly positive correlation between the heterogeneous response and the log of total assets, lagged stock returns, contemporary and lagged return on assets, and the female dummy. All these variables are upward biased given by TFE and IFE. On the contrary, a significant negative correlation is found between $\theta_{g_{i,t}}$ and contemporary stock returns, which is reflected in a downward bias reported by TFE and IFE.

The results from this empirical example show that failure to account for heterogeneity to market-wide shocks in the OLS, FE, TFE and IFE estimations might lead to the various estimates being inconsistent and result in incorrect statistical inference.

5.3 Managerial incentives and risk-taking

In next empirical example, we study the economic importance of accounting for heterogeneous responses to market-wide shocks through the investigation of the relation managerial incentives on risk-taking behavior using OLS, FE, TFE, IFE, and GFE and compare the resulting estimates. We start by estimating the following model for managerial risk-taking following Coles et al. (2006) for a sample period from 1992-2014:

$$\text{Risk taking behavior}_{it} = \alpha_{\mathrm{Ind}_i} + \theta_{g_i t} + X_{it}'\beta + \epsilon_{it},\ \ g_i \in \{1, ..., G\}, \tag{10}$$

where Risk taking behavior$_{it}$ is research and development expenditures scaled by total assets (R&D) at firm i in year t. X_{it} is a 9×1 vector of explanatory variables, among which we are particularly interested in the two variables that measure managerial incentives, lagged vega (the dollar change in the value of the CEO's wealth for a 0.01 change in standard deviation of returns) and lagged delta (the dollar change in the value of the CEO's wealth for a 1% change in stock price). Using lagged vega and delta as proxies of managerial incentives is consistent with Coles et al. (2006).[14] In addition to these two covariates, we include the following commonly used controls thought to affect managerial incentives: managerial cash compensation, log sales, market to book ratio, firm's surplus cash, sales growth, stock returns and book leverage. Unlike the previous example, we follow Coles et al. (2006) to focus on exploiting the within industry variation through industry fixed effects α_{Ind_i} (rather than firm fixed effect).[15] $\theta_{g_{i,t}}$ is a time-varying group effect with an unknown group pattern as usual. Note that our model mainly differs from Coles et al. (2006)'s in the specification of the time effect. As in the previous example, we exclude both the industry fixed and time-varying group effect for the OLS estimator and retain the industry fixed effect for the FE estimator. For TFE, $\theta_{gi,t} = \theta_t$ and IFE specifies $\theta_{gi,t}$ as industry (2-digit SIC controls) is interacted with year fixed effects.

Before implementing the GFE estimation, we again employ the BIC and plot the coefficient estimates against G to determine the number of groups. Figure 1 (b) reports the BIC for different numbers of groups, and the criterion reaches its minimum when $G = 15$. This choice of the number of groups is also supported by the plot of estimated slope coefficients as a function of G in Figure 4. For significant variables, such as vega, log sales, market to book ratio, excess cash, sales growth, and total returns, their estimated coefficients all stabilize when G is larger than 10. To ensure the consistency of slope coefficient estimates, we choose $G = 15$ for the following analysis.[16]

Table 11 presents the coefficients estimates of the determinants of risk taking behaviour obtained from OLS, fixed effect (FE), two-way fixed effect (TFE), interacted fixed effect (IFE) based on industry grouping, and grouped fixed effect (GFE) based on $G = 15$. Consistent with Coles et al.

[14]We are thankful to Lalitha Naveen for sharing measures of managerial incentives. The Vega and Delta variables are collected from https://sites.temple.edu/lnaveen/data/.

[15]Coles et al. (2006) find that the firm fixed effect specification for R&D yields insignificant estimated coefficient on Vega due to limited within firm variation.

[16]Recall that over-specification of G does not harm consistency but efficiency, while under-specifying G leads to inconsistency slope coefficient estimates.

(2006), we find the estimated coefficients on vega are significant at 1% with industry effect.[17] For dependent variable R&D, the coefficient on vega is positive, and for CAPEX, the estimate is negative suggesting that higher vega is associated with higher R&D and lower capital expenditures. In addition, we find that the estimated coefficients of all control variables using GFE is consistent with Coles et al. (2006), which uses industry and year fixed effect specification.

From the econometric theory in Section 2.2, we expect the GFE estimates of the treatment effect to be the same as industry fixed effect estimates, if there is homogeneous responses to market wide shocks. However, we observe non-trivial differences between the treatment effect of managerial incentives using GFE and other estimators (industry fixed effect and industry interacted with year specification), suggesting the importance of heterogeneous responses to market wide shocks. The estimates in Table 11 illustrate that TFE and IFE can be severely biased and lead to incorrect inferences. For example, for the coefficient on vega, TFE and IFE report an estimate that is 17% and 24% larger than GFE with the p-value of the t statistic less than 12% and 6%, respectively. There is also substantial differences in the estimated effect of excess cash for various methods. TFE and IFE report 31% and 12% smaller estimates than that of GFE. The differences in the point estimates of managerial risk-taking behavior on vega and delta, due to heterogeneous responses to market-wide shocks, has important corporate policy implications on the executive compensation structure.

We report the responses of all groups in Figure 5. The group structure produced by GFE again differs from industry grouping to a large extent. This implies that IFE with industry-year interaction may not fully capture the heterogeneity. The correlations between group fixed effects and explanatory variables presented in Table ?? again provide sound support for these results. We find that $\theta_{g_{i,t}}$ is strongly and positively correlated with Vega_{t-1} and MB, which is reflected by the upward bias of the FE, TFE, and IFE estimators.

6 Conclusions

Firms are often exposed to market-wide shocks such as regulatory changes and financial crises. Failing to incorporate these market-wide shocks can lead to inconsistent estimates of treatment effect

[17] The point estimates is be very close to those of Coles et al. (2006) (Table 2) when we use the same sample period in their paper. Results can be found in the appendix.

and can affect statistical inference. The most widely used approach to model market-wide shocks is the time fixed effects models, which assume that all firms or individuals respond homogeneously to these shocks. However, the homogeneity assumption is unlikely to be correct, because it is well-documented that regulatory changes like the Sarbanes-Oxley Act and financial crises have heterogenous impact on firms outcomes.

In this paper, we demonstrate how ignoring this heterogeneity results in biased slope coefficient estimates, when heterogeneous shocks are correlated with explanatory variables. More specifically, we formally quantify the bias of the estimates of existing methods in using OLS, two-way fixed effects and interacted industry-year fixed effects models.

We propose the use of the "grouped fixed effects" estimator to account for heterogenous responses to market-wide shocks. The group fixed effect estimators assume that there is an underlying latent group pattern for the population of firms. Within the group, firms respond similarly to each shock, while they respond differently across groups. One important advantage of this approach is that the group structure can be consistently estimated from the data, jointly with the regression parameters. Therefore, no a priori knowledge of grouping is required. This avoids possible estimate inconsistency caused by the misspecification of the group structure.

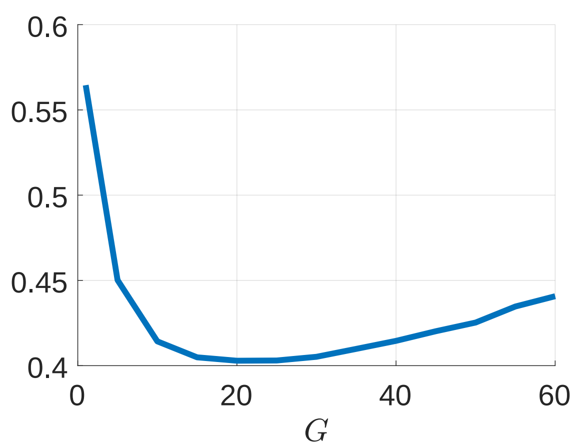
We theoretically compare the grouped fixed effect (GFE) estimators with popular alternatives, such as two-way fixed effects (TFE), interacted fixed effects (IFE) with artificially imposed group structure (e.g. industry-based grouping), and event studies. We show that two-way fixed effects estimates are inconsistent if firms respond differently to the shocks. On the contrary, grouped fixed effects always produce consistent slope coefficient estimates, when the number of groups is not under-specified. GFE is asymptotically equivalent to the interacted fixed effects model, if the imposed group structure in IFE happens to coincide with the data generating process. However, if the group structure of IFE is misspecified, IFE estimates are no longer consistent. Although event studies can address heterogeneous responses by sub-sample analysis, the standard deviation estimates are often misleading in the sense that it may fail to take into account within-group correlations. We provide guidelines on how to appropriately correct for standard errors based on the group structure.

We conduct a large number of Monte Carlo simulation experiments, controlling for the size and the frequency of shocks, the degree of heterogeneity among units, the correlation between responses and regressors, the number of groups, etc. Simulation results confirm our theoretical argument that conventional methods are biased to different extents, depending on the feature of shocks, the correlation, and specification of groups, etc. On the contrary, GFE remains a consistent approach even when all firms respond homogeneously.

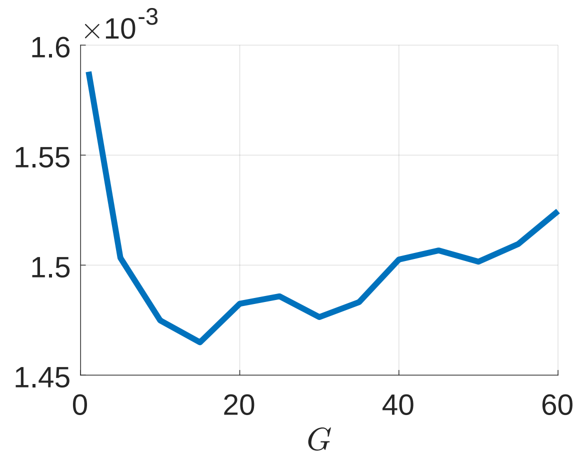
We demonstrate the economic importance of accounting for heterogeneity through an empirical model of CEO total compensation and executive risk taking behavior. We find that accounting for time-varying heterogeneity plays an important role in practice. Our findings and the proposed GFE estimator are likely to be of increasing importance and of practical use to empirical researchers.

Figure 1
Bayesian information criterion to determine the number of groups

This figure provides the BIC for different numbers of groups in the two empirical applications. The figure in the left panel, Figure (a), considers the case of CEO compensation, and the one in the right panel, Figure (b), concerns the managerial incentives and risk taking. The x-axis is the number of groups, and the y-axis is the BIC.



(a) CEO compensation



(b) Managerial incentives

Figure 2
CEO compensation: Slope coefficient estimates as a function of the number of groups

This figure plots the estimated slope coefficients as the number of groups increases from 1 to 60 in the study of CEO compensation. The x-axis is the number of groups, and the y-axis is the estimated coefficient.

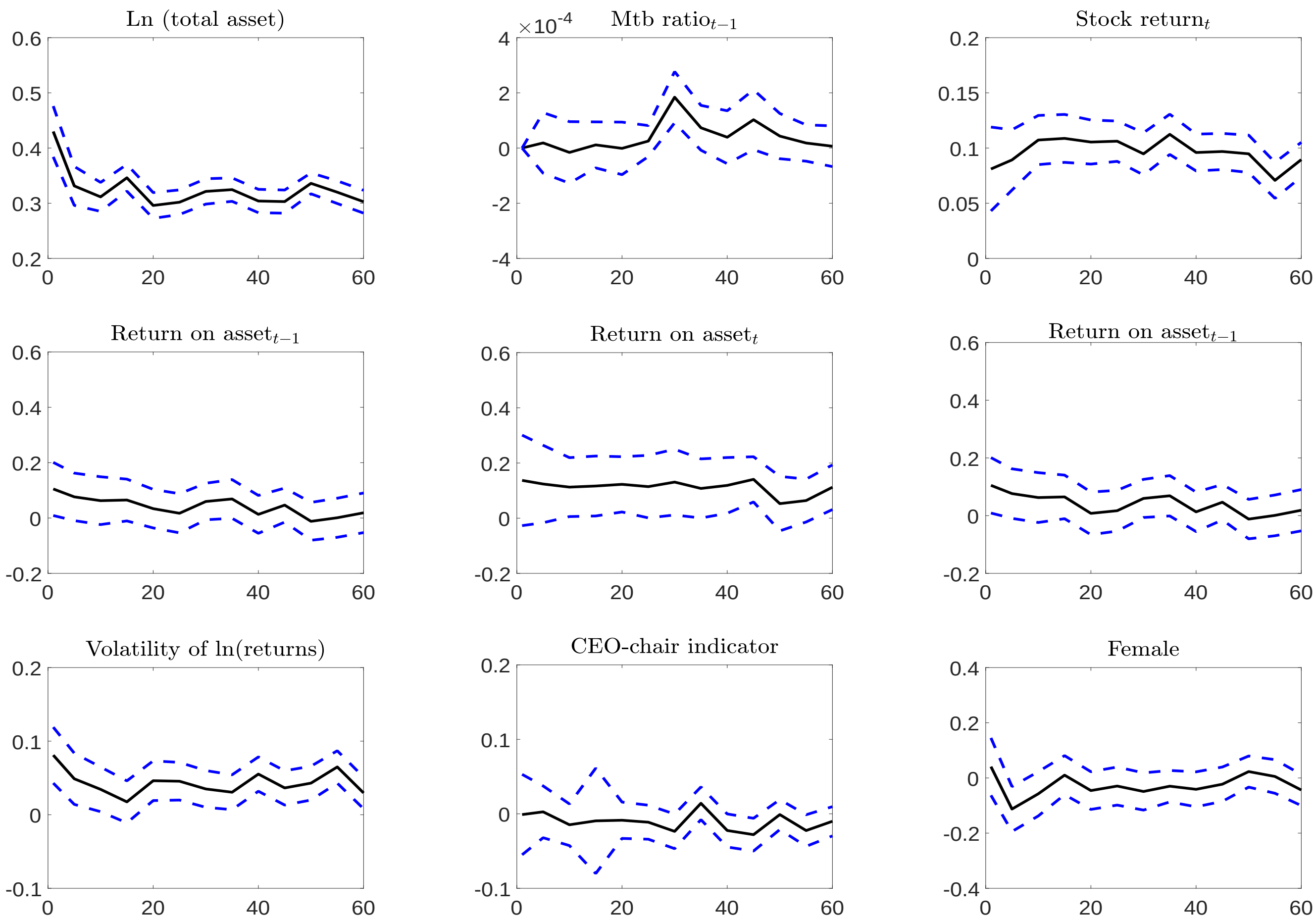


Figure 3
CEO compensation: Heterogeneous responses of groups

This figure plots the heterogeneous responses to the shocks (group fixed effects) for 20 groups. The x-axis is the year, and the y-axis is the estimated group fixed effects $\theta_{g,t}$.

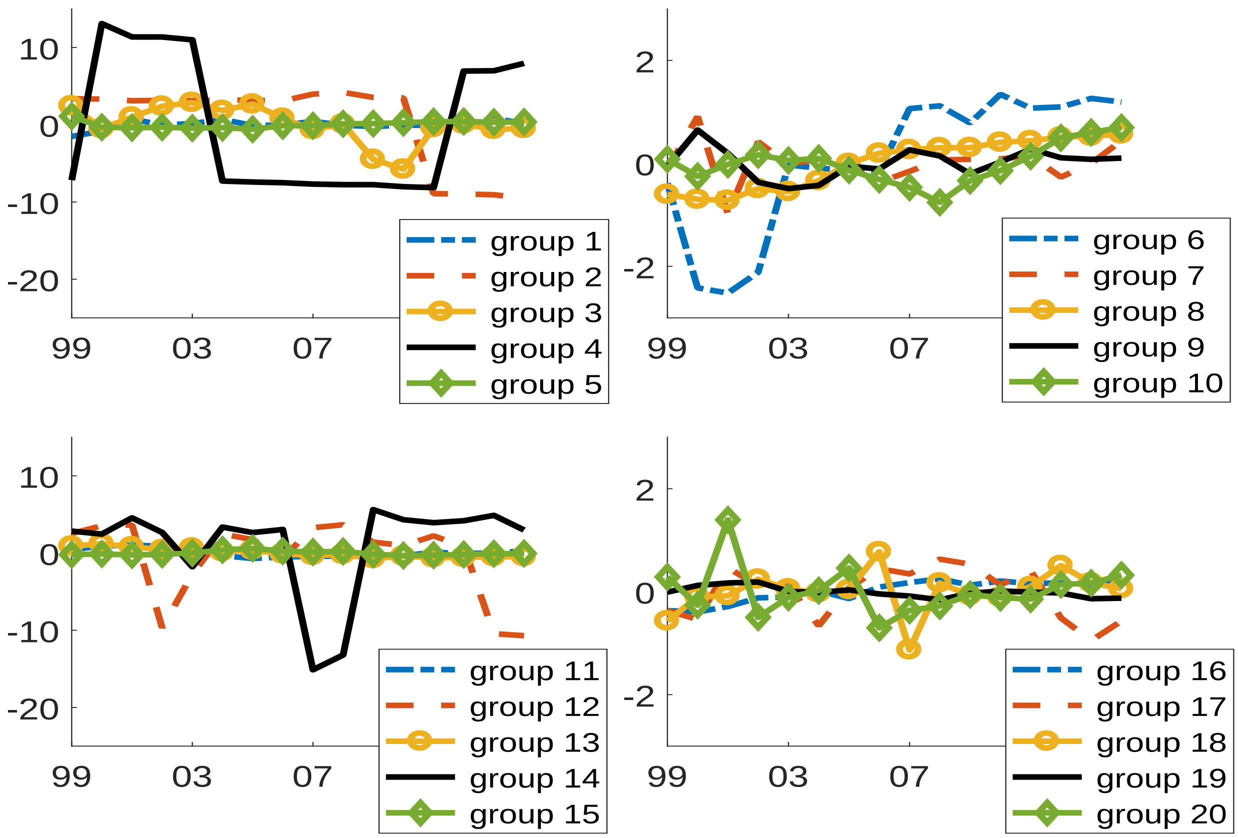


Figure 4
Managerial incentives and risk taking behaviour: Slope coefficients as a function of the number of groups

This figure plots the estimated slope coefficient estimates as the number of groups increases from 1 to 60 in the study of managerial incentives and risk taking behaviour. The x-axis is the number of groups, and the y-axis is the estimated coefficient.

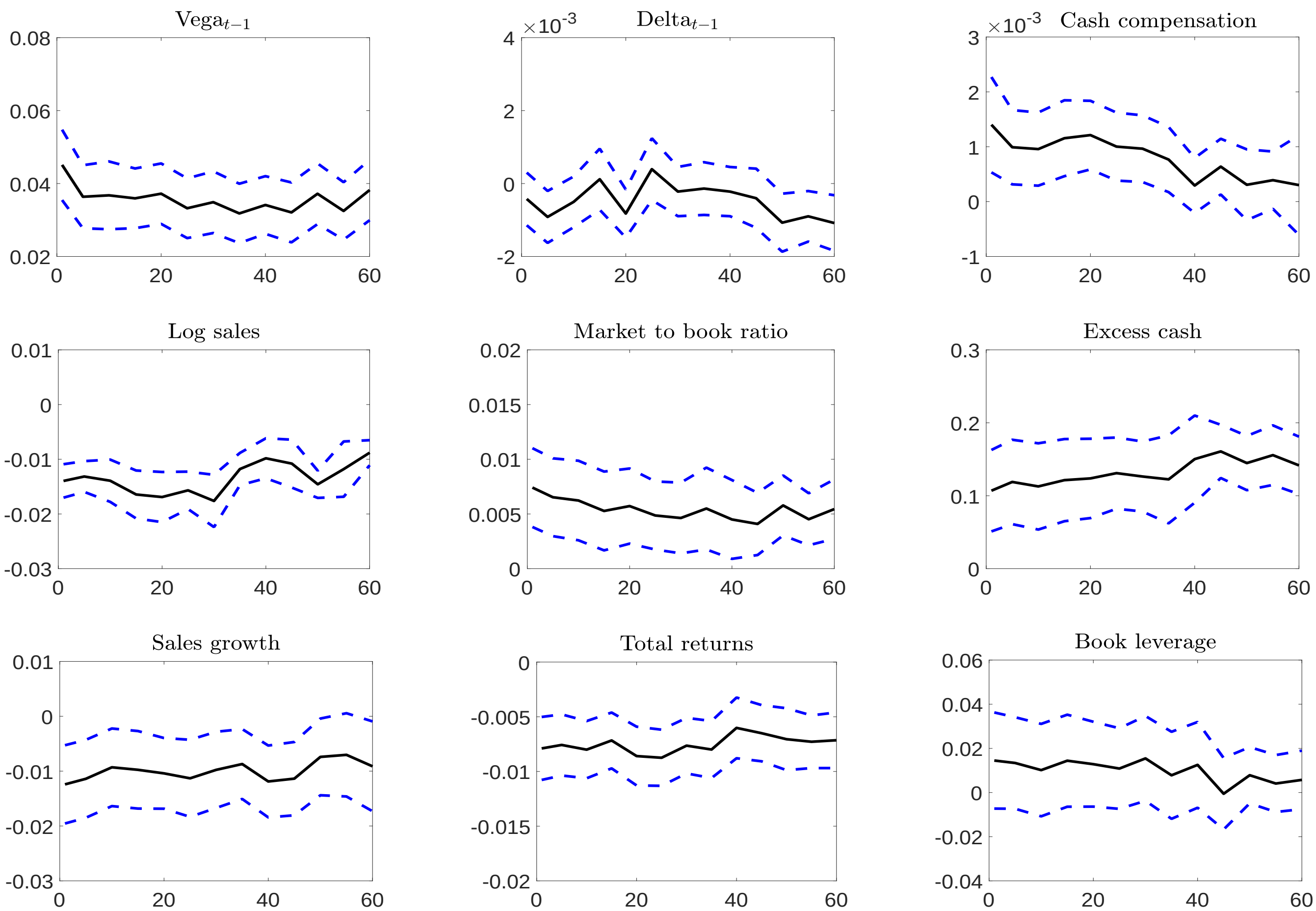


Figure 5

Managerial incentive and risk taking : Heterogeneous responses of groups

This figure plots the heterogeneous responses to the shocks (group fixed effects) for 15 groups. The x-axis is the year, and the y-axis is the estimated group fixed effects $\theta_{g,t}$.

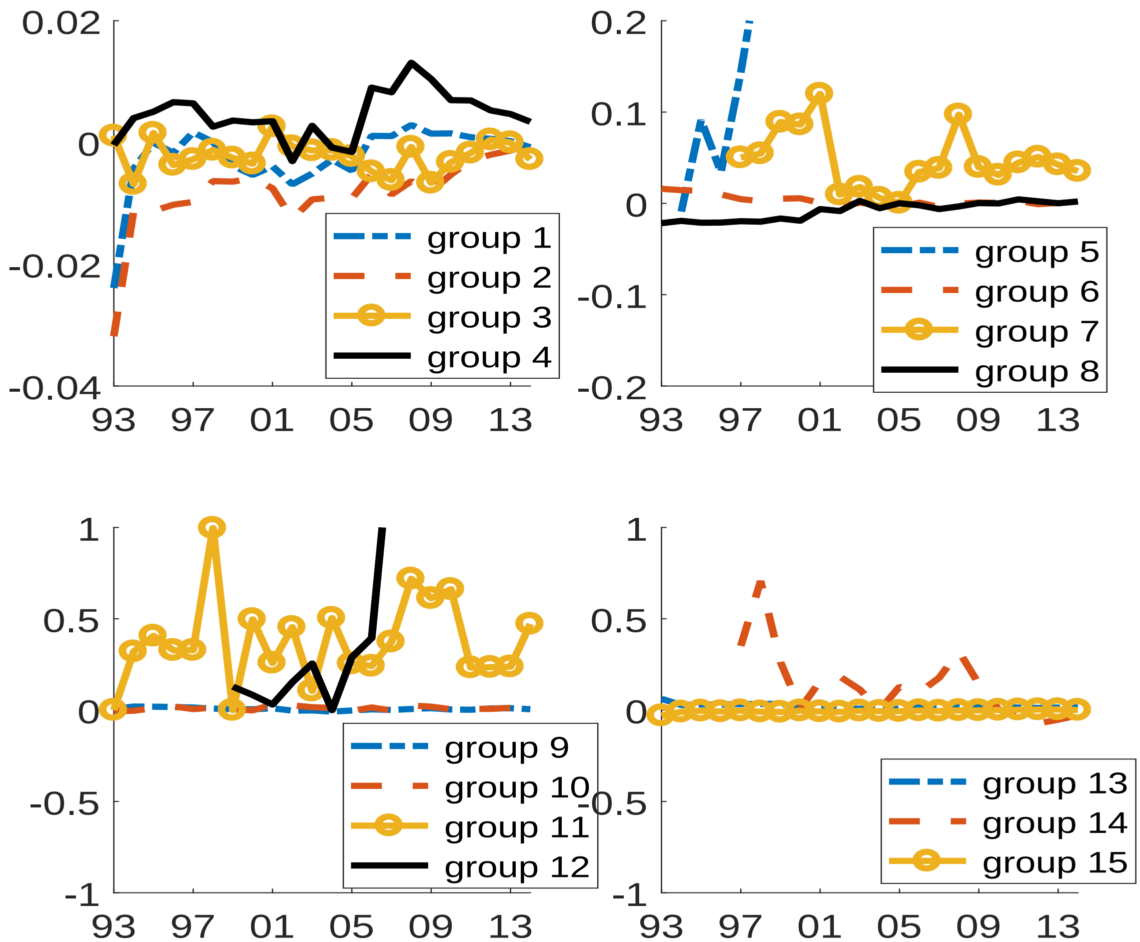


Table 3
Frequent moderate shocks

The table presents the bias, estimated standard deviation and root mean square error of estimated coefficients for simulations under DGP1 for different econometric specifications. DGP1 is generated using the procedure described in Section 4, where shocks occur at every time t and heterogenous responses to shocks are allowed. The parameters in the DGP are set as $c_\tau = 0.5$, $c_\theta = 15$, and $G = 5$. c_τ is a constant which determines the degree of correlation between X and θ, c_θ is a constant which controls the size of shocks, G is the number of groups. The table reports the results using the following methods: grouped fixed effects (GFE), two-way fixed effects (TFE), and group-time interacted fixed effects (IFE). The IFE method includes 4 different specifications: IFE_{gt} exactly specifies the group memberships of all individuals and the time points of all shocks, IFE_{g} exactly specifies the groups but correctly specifies the shocks every 1/3 of the time, IFE_{t} correctly specifies the shocks but not the groups, and IFE_{null} misspecifies both groups and shocks. When the groups are misspecified, we assign 30% of individuals to the closest neighboring groups. The details of these methods are presented in Section 4.2. N is the number of units, and T is the number of time series observations.

		GFE	TFE	IFE_{gt}	IFE_{g}	IFE_{t}	IFE_{null}
				Panel A. $N = 1000$, $T = 10$			
β_1	Bias	0.0139	0.5954	0.0002	0.5125	0.2548	0.5693
	Std.	0.0119	0.2057	0.0108	0.2203	0.1061	0.2148
β_2	Bias	0.0132	0.5991	−0.0006	0.5145	0.2590	0.5720
	Std.	0.0124	0.2102	0.0109	0.2206	0.1089	0.2133
β	RMSE	0.0183	0.8858	0.0144	0.9898	1.0035	0.8500
				Panel B. $N = 1000$, $T = 20$			
β_1	Bias	0.0074	0.6107	0.0004	0.6336	0.2625	0.6577
	Std.	0.0074	0.1383	0.0072	0.1465	0.0733	0.1449
β_2	Bias	0.0071	0.6110	0.0002	0.6347	0.2591	0.6588
	Std.	0.0077	0.1381	0.0072	0.1453	0.0719	0.1433
β	RMSE	0.0150	0.8869	0.0102	0.9235	0.3831	0.9558
				Panel C. $N = 2000$, $T = 10$			
β_1	Bias	0.0126	0.6045	0.0005	0.5208	0.2571	0.5778
	Std.	0.0093	0.2006	0.0077	0.1978	0.1055	0.1950
β_2	Bias	0.0122	0.6072	−0.0002	0.5217	0.2666	0.5806
	Std.	0.0089	0.2064	0.0072	0.2102	0.1088	0.2052
β	RMSE	0.0215	0.9008	0.0105	0.7929	0.3969	0.8664

Table 4
Sparse moderate shocks

The table presents the bias, estimated standard deviation and root mean square error of estimated coefficients for simulations under DGP2 for different econometric specifications. DGP2 is generated using the procedure described in Section 4, where shocks occur 25% of the time and heterogenous responses to shocks are allowed. The parameters in the DGP are set as $c_\tau = 0.5$, $c_\theta = 15$, and $G = 5$. c_τ is a constant which determines the degree of correlation between X and θ, c_θ is a constant which controls the size of shocks, G is the number of groups. The table reports the results using the following methods: grouped fixed effects (GFE), two-way fixed effects (TFE), and group-time interacted fixed effects (IFE). The IFE method includes 5 different specifications: IFE_{gt} exactly specifies the group memberships of all individuals and the time points of all shocks, IFE_{g} exactly specifies the groups but correctly specifies the shocks every 1/3 of the time, IFE_{got} exactly specifies the groups but over-specifies the shocks at each time period, IFE_{t} correctly specifies the shocks but not the groups, and IFE_{null} misspecifies both groups and shocks. When the groups are misspecified, we assign 30% of individuals to the closest neighboring groups. The details of these methods are presented in Section 4.2. N is the number of units, and T is the number of time series observations.

		GFE	TFE	IFE_{gt}	IFE_{g}	IFE_{got}	IFE_{t}	IFE_{null}
				Panel A. $N = 1000$, $T = 10$				
β_1	Bias	0.0032	0.3584	−0.0002	0.2712	−0.0001	0.1340	0.3178
	Std.	0.0130	0.3537	0.0093	0.3835	0.0105	0.1328	0.3609
β_2	Bias	0.0035	0.3386	−0.0008	0.2410	−0.0005	0.1268	0.2897
	Std.	0.0161	0.3191	0.0086	0.3345	0.0100	0.1257	0.3159
β	RMSE	0.0212	0.6852	0.0127	0.6246	0.0144	0.2596	0.6438
				Panel B. $N = 1000$, $T = 20$				
β_1	Bias	0.0018	0.4515	0.0003	0.4489	0.0001	0.1657	0.4744
	Std.	0.0077	0.2559	0.0062	0.3171	0.0068	0.0978	0.3068
β_2	Bias	0.0016	0.4450	0.0001	0.4425	0.0003	0.1618	0.4672
	Std.	0.0081	0.2639	0.0062	0.3214	0.0071	0.0973	0.3069
β	RMSE	0.0114	0.7326	0.0087	0.7751	0.0098	0.2695	0.7945
				Panel C. $N = 2000$, $T = 10$				
β_1	Bias	0.0037	0.3636	0.0004	0.2754	0.0004	0.1280	0.3186
	Std.	0.0177	0.3438	0.0063	0.3622	0.0072	0.1240	0.3456
β_2	Bias	0.0016	0.3682	−0.0005	0.2866	−0.0005	0.1262	0.3316
	Std.	0.0125	0.3216	0.0063	0.3655	0.0071	0.1178	0.3381
β	RMSE	0.0220	0.6993	0.0089	0.6498	0.0101	0.2480	0.6669

Table 5
Homogeneous shocks

The table presents the bias, estimated standard deviation and root mean square error of estimated coefficients for simulations under DGP3 for different econometric specifications. DGP3 is generated using the procedure described in Section 4, where shocks occur at every time t and there is a homogenous response to the shock. The parameters in the DGP are set as $c_\tau = 0.5$, $c_\theta = 15$, and $G = 1$. c_τ is a constant which determines the degree of correlation between X and θ, c_θ is a constant which controls the size of shocks, G is the number of groups. The table reports the results using the following methods: grouped fixed effects (GFE), two-way fixed effects (TFE), and group-time interacted fixed effects (IFE). GFE is estimated under two specifications of groups, $\widehat{G} = 1$ and $\widehat{G} = 4$. The IFE method includes 4 different specifications:IFE$_{gt}$ exactly specifies the group memberships of all individuals and the time points of all shocks, IFE$_g$ exactly specifies the groups but correctly specifies the shocks every 1/3 of the time, IFE$_t$ correctly specifies the shocks but not the groups, and IFE$_{null}$ misspecifies both groups and shocks. The details of these methods are presented in Section 4.2. N is the number of units, and T is the number of time series observations.

		GFE ($\widehat{G} = 1$)	GFE ($\widehat{G} = 4$)	TFE	IFE$_{gt}$	IFE$_g$	IFE$_t$	IFE$_{null}$
				Panel A. $N = 1000$, $T = 10$				
β_1	Bias	0.0003	0.0007	0.0003	0.0003	0.0557	0.0003	0.0531
	Std.	0.0106	0.0095	0.0106	0.0106	0.0632	0.0106	0.0384
β_2	Bias	−0.0002	0.0000	−0.0002	−0.0002	0.0542	−0.0001	0.0494
	Std.	0.0110	0.0095	0.0110	0.0110	0.0633	0.0110	0.0381
β	RMSE	0.0152	0.0167	0.0152	0.0152	0.1184	0.0153	0.0905
				Panel B. $N = 1000$, $T = 20$				
β_1	Bias	0.0002	0.0003	0.0002	0.0002	0.0666	0.0002	0.0658
	Std.	0.0073	0.0069	0.0073	0.0073	0.0385	0.0073	0.0289
β_2	Bias	0.0001	0.0000	0.0001	0.0001	0.0667	0.0001	0.0669
	Std.	0.0071	0.0069	0.0071	0.0071	0.0402	0.0071	0.0275
β	RMSE	0.0102	0.0106	0.0102	0.0102	0.1089	0.0102	0.1006
				Panel C. $N = 2000$, $T = 10$				
β_1	Bias	0.0000	−0.0001	0.0000	0.0000	0.0566	0.0000	0.0527
	Std.	0.0071	0.0068	0.0071	0.0071	0.0627	0.0071	0.0362
β_2	Bias	0.0000	−0.0001	0.0000	0.0000	0.0535	0.0000	0.0490
	Std.	0.0074	0.0068	0.0074	0.0074	0.0623	0.0074	0.0344
β	RMSE	0.0102	0.0113	0.0102	0.0102	0.1177	0.0103	0.0875

Table 6
Large and highly correlated shocks

The table presents the bias, estimated standard deviation and root mean square error of estimated coefficients for simulations under DGP1 for different shock and correlation sizes. DGP1 is generated using the procedure described in Section 4, where shocks occur at every time t and heterogenous responses to shocks are allowed. The table reports the results using the following methods: grouped fixed effects (GFE), two-way fixed effects (TFE), and group-time interacted fixed effects (IFE). The IFE method includes 4 different specifications: IFE_{gt} exactly specifies the group memberships of all individuals and the time points of all shocks, IFE_{g} exactly specifies the groups but correctly specifies the shocks every 1/3 of the time, IFE_{t} correctly specifies the shocks but not the groups, and IFE_{null} misspecifies both groups and shocks. When the groups are misspecified, we assign 40% of individuals to the closest neighboring groups. The details of these methods are presented in Section 4.2. Panel A presents the results for $c_\tau = 0.5$, a larger time series shock $c_\theta = 5$, and $G = 5$, and Panel B presents the results for a larger correlation c_τ=1. c_τ is a constant which determines the degree of correlation between X and θ, c_θ is a constant which controls the size of shocks, G is the number of groups. We report the results for 1000 cross-sectional units (N), 20 time periods (T), and 5 groups (G).

$(N = 1000, T = 20)$		GFE	TFE	IFE_{gt}	IFE_{g}	IFE_{t}	IFE_{null}
Panel A. Large shocks: $c_\tau = 0.5$, $c_\theta = 5$, and $G = 5$							
β_1	Bias	0.0012	1.8553	−0.0002	1.9234	0.8076	1.9969
	Std.	0.0073	0.4343	0.0068	0.4613	0.2225	0.4565
β_2	Bias	0.0013	1.8431	0.0000	1.9104	0.7990	1.9844
	Std.	0.0076	0.4295	0.0073	0.4588	0.2189	0.4542
β	RMSE	0.0107	2.6855	0.0100	2.7878	1.1781	2.8878
Panel B. Highly correlated shocks: $c_\tau = 1$, $c_\theta = 15$, and $G = 5$							
β_1	Bias	0.0148	0.6809	−0.0004	0.6930	0.3723	0.7051
	Std.	0.0091	0.1079	0.0071	0.1174	0.0828	0.1119
β_2	Bias	0.0144	0.6833	−0.0004	0.6961	0.3745	0.7085
	Std.	0.0087	0.1101	0.0069	0.1161	0.0826	0.1125
β	RMSE	0.0242	0.9769	0.0099	0.9960	0.5409	1.0120

Table 7
Under-specify G

The table presents the bias, estimated standard deviation and root mean square error of estimated coefficients for simulations under DGP1 for an underspecified number of groups G. DGP1 is generated using the procedure described in Section 4, where shocks occur at every time t and heterogenous responses to shocks are allowed. The parameters in the DGP are set as $c_\tau = 0.5$, $c_\theta = 15$, and $G = 4$. c_τ is a constant which determines the degree of correlation between X and θ, c_θ is a constant which controls the size of shocks, G is the number of groups. The table reports the results using grouped fixed effects (GFE) and two-way fixed effects (TFE). The estimates are generated using two groups, $\widehat{G} = 2$, where the true DGP of G=4. The details of these methods are presented in Section 4.2. N is the number of units, and T is the number of time series observations.

		$N = 1000, T = 10$		$N = 1000, T = 20$		$N = 2000, T = 10$	
		GFE	TFE	GFE	TFE	GFE	TFE
β_1	Bias	0.0799	0.2772	0.0808	0.2848	0.0781	0.2833
	Std.	0.0397	0.1082	0.0277	0.0793	0.0365	0.1027
β_2	Bias	0.0778	0.2762	0.0819	0.2826	0.0777	0.2782
	Std.	0.0416	0.1061	0.0266	0.0764	0.0391	0.1117
β	RMSE	0.1255	0.4196	0.1213	0.4160	0.1224	0.4250

Table 8
More groups G

The table presents the bias, estimated standard deviation and root mean square error of estimated coefficients for simulations under DGP1 with two different number of groups G. DGP1 is generated using the procedure described in Section 4, where shocks occur at every time t and heterogenous responses to shocks are allowed. The parameters in the DGP are set as $c_\tau = 0.5$, $c_\theta = 15$. c_τ is a constant which determines the degree of correlation between X and θ, c_θ is a constant which controls the size of shocks. We consider the number of groups $G = 10$ and $G = 20$ in DGP. The table reports the results using the following methods: grouped fixed effects (GFE), two-way fixed effects (TFE), and group-time interacted fixed effects (IFE). The IFE method includes 4 different specifications: IFE_{gt} exactly specifies the group memberships of all individuals and the time points of all shocks, IFE_{g} exactly specifies the groups but correctly specifies the shocks every 1/3 of the time, IFE_{t} correctly specifies the shocks but not the groups, and IFE_{null} misspecifies both groups and shocks. When the groups are misspecified, we assign 30% of individuals to the closest neighboring groups. The details of these methods are presented in Section 4.2. We report the results for 1000 cross-sectional units (N) and 50 time periods (T).

		GFE	TFE	IFE_{gt}	IFE_{g}	IFE_{t}	IFE_{null}
$G = 10$ ($N = 1000$, $T = 20$)							
β_1	Bias	0.0040	5.0593	−0.0002	4.9302	2.8958	5.0519
	Std.	0.0073	0.6131	0.0071	0.6419	0.4988	0.6227
β_2	Bias	0.0048	5.0731	0.0006	4.9358	2.8899	5.0528
	Std.	0.0079	0.5423	0.0077	0.5978	0.4745	0.5704
β	RMSE	0.0124	7.2112	0.0105	7.0312	4.1485	7.1947
$G = 20$ ($N = 1000$, $T = 20$)							
β_1	Bias	0.0051	5.0402	0.0011	4.9088	2.8528	5.0238
	Std.	0.0078	0.6268	0.0075	0.6510	0.4893	0.6313
β_2	Bias	0.0039	5.0259	−0.0002	4.8785	2.8488	4.9996
	Std.	0.0076	0.5917	0.0075	0.6173	0.4883	0.6025
β	RMSE	0.0126	7.1697	0.0107	6.9785	4.0904	7.1411

Table 9
CEO compensation

The table presents a regression of CEO compensation on explanatory variables using different estimation models. The data contains 1859 firms spanned over 16 years from 1999–2014. OLS is ordinary least squares without any fixed effects estimator, FE includes firm fixed effects, TFE is firm and year fixed effect estimator, IFE is a firm fixed effect with industry and year interaction estimator. GFE estimators are obtained based on $G = 20$, selected from a range of G from 1 to 60 based on various information criterion, and include firm fixed effects. For IFE, preliminary within-transformation is taken to integrate out the individual specific effect as in GFE. The industry IFE is based on 1-digit SIC code. TA is log of total assets, MB is the market to book ratio, r_t is stock returns at time t, r_{t-1} is stock returns at time $t-1$, ROA is return on assets, σ_r is the volatility of returns, CEO & chairman is a dummy variable equal to 1 when the CEO is also the chairman of the firm, and zero otherwise, and Female is a dummy equal to 1 if the CEO is a female and zero otherwise. One-way clustered standard errors are reported in the parentheses except for GFE. For GFE, the asymptotic standard errors are presented. ***, **, * denote significance at 1%, 5%, and 10% level, respectively.

	OLS	FE	TFE	IFE	GFE
TA	0.362***	0.430***	0.327***	0.342***	0.322***
	(0.013)	(0.023)	(0.027)	(0.026)	(0.012)
MB_{t-1}	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
r_t	0.099***	0.081***	0.067***	0.073***	0.107***
	(0.021)	(0.019)	(0.023)	(0.020)	(0.010)
r_{t-1}	0.203***	0.187***	0.157***	0.156***	0.153***
	(0.020)	(0.015)	(0.018)	(0.016)	(0.009)
ROA_t	0.226	0.137*	0.180**	0.198**	0.099**
	(0.102)	(0.082)	(0.089)	(0.087)	(0.048)
ROA_{t-1}	0.177***	0.105**	0.122**	0.116**	0.034
	(0.062)	(0.048)	(0.050)	(0.053)	(0.035)
σ_r	0.128***	0.081***	0.090***	0.067***	0.034**
	(0.031)	(0.019)	(0.021)	(0.020)	(0.014)
CEO&chairman	0.112***	−0.001	0.026	0.018	−0.000
	(0.032)	(0.027)	(0.028)	(0.027)	(0.012)
Female	0.064	0.041	−0.012	0.011	−0.046
	(0.079)	(0.052)	(0.049)	(0.054)	(0.034)
Obs	20,060	20,060	20,060	20,060	20,060
Overall R^2	0.298	0.294	0.299	0.115	0.447

Table 10
CEO compensation correlations

Panel A presents the correlation between the group fixed effects and the explanatory variables for the regression of CEO total compensation. The data contains 1859 firms spanned over 16 years from 1999–2014. TA is log of total assets, MB is the market to book ratio, r_t is stock returns at time t, r_{t-1} is stock returns at time $t-1$, ROA is return on assets, σ_r is the volatility of returns, CEO & chairman is a dummy variable equal to 1 when the CEO is also the chairman of the firm, and zero otherwise, and Female is a dummy equal to 1 if the CEO is a female and zero otherwise. Bold correlation coefficients are statistically significant at the 5% level. Panel B presents CEO characteristics and compensation structure for high and low ROA pre and post crisis of 2008. Executive age is the age of the CEO. Tot. Comp. and Cash Comp. is the total and cash component of the CEO compensation respectively. Delta is the dollar change to the CEO wealth for a 1% change in stock price. Vega is the dollar change in the CEO's wealth for a 0.01 change in the standard deviation of returns.

Panel A: Correlation

	TA	MB_{t-1}	r_t	r_{t-1}	ROA_t	ROA_{t-1}	σ_r	CEO&chairman	Female
$\theta_{g_i,t}$	**0.022**	**0.059**	**−0.020**	**0.033**	**0.028**	**0.025**	**0.034**	**0.022**	**0.088**

Panel B: Pre. vs. Post Crisis

ROA Type	Tot. Comp.	Vega	Delta	Tenure	CEO turnover
Pre-Crisis					
Low	1179.41	0.18	0.65	13.60	0.12
High	1551.61	0.16	1.09	16.14	0.11
Post-Crisis					
Low	942.72	0.10	0.36	15.83	0.11
High	996.96	0.16	0.71	17.21	0.10

Hirfindahl-Hirschman Index across groups

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
HHI	0.03	0.5	0.5	1	0.03	0.2	0.02	0.01	0.01	0.02	0.04	0.3	0.04	0.5	0.01	0.01	0.03	0.02	0.01	0.02

Table 11
Managerial incentives and risk taking behaviour

The table presents a regression of executive risk taking on executive incentives with different fixed effects specifications. The data contains 2729 firms spanned over 23 years from 1992–2014. OLS is ordinary least squares without any fixed effects, FE includes industry fixed effects following Coles et al. (2006), TFE is an industry and year fixed effect estimator, IFE is an industry fixed effect with industry and year interaction fixed effect model. GFE estimators are obtained based on $G = 15$, selected from a range of G from 5 to 60 based on various information criterion, and includes industry fixed effects. For IFE, preliminary within-transformation is taken to integrate out the industry specific effect as in GFE. The grouping of IFE is based on 12 groups of two-digit SIC code. We include the following explanatory variables as in Coles et al. (2006): cash compensation (*Cash Comp.*), log sales (*Sales*), market to book ratio (*MB*), excess cash (*Ex. Cash*), sales growth (*Sales Growth*), total stock returns (*Tot. Return*), and book leverage (*Book Lev.*). α_i is the firm fixed effect and $\theta_{gi,t}$ is a time-varying group effect. We exclude both the firm fixed and time-varying group effect for the OLS estimator and retain the firm fixed effect for the FE estimator. For TFE, $\theta_{gi,t} = \theta_t$ and IFE specifies $\theta_{gi,t}$ as industry is interacted with year fixed effects. One-way clustered standard errors are reported in the parentheses except for GFE. For GFE, the asymptotic standard errors are presented. ***, **, * denote significance at 1%, 5%, and 10% level, respectively.

	OLS	FE	TFE	IFE	GFE
$Vega_{t-1}$	0.060***	0.044***	0.045***	0.049***	0.037***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
$Delta_{t-1}$	-0.001***	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cash comp.	0.000	0.001***	0.001***	0.001	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sales	-0.016***	-0.014***	-0.014***	-0.014***	-0.013***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
MB	0.009***	0.007***	0.007***	0.008***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ex. cash	0.143***	0.108***	0.107***	0.109***	0.122***
	(0.032)	(0.028)	(0.028)	(0.030)	(0.029)
Sales growth	-0.015***	-0.013***	-0.012***	-0.014***	-0.012***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)
Tot. return	-0.009***	-0.008***	-0.008***	-0.008***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Book lev.	0.002	0.015	0.015	0.015	0.013
	(0.011)	(0.011)	(0.011)	(0.012)	(0.010)
Obs	23,335	23,335	23,335	23,335	23,335
Overall R^2	0.251	0.367	0.368	0.382	0.590

Table 12
Managerial incentives and risk taking behaviour: Correlation analysis

This table presents the correlation between the group fixed effects and the explanatory variables for the regression of managerial incentives and risk taking behaviour. The data contains 2729 firms spanned over 23 years from 1992–2014. Vega is the dollar change in the value of the CEO's wealth for a 0.01 change in standard deviation of returns, Delta is the dollar change in the value of the CEO's wealth for a 1% change in stock price, Cash comp. is cash compensation, Sales is log sales, MB is the market to book ratio, Ex. cash is excess cash, Sales growth is the sales growth, Tot. Return is total stock returns, and Book Lev. is book leverage. Bold correlation coefficients are statistically significant at the 5% level. Pre-SOX window is from 2000-2002 and Post SOX is from 2003-2005

	$Vega_{t-1}$	$Delta_{t-1}$	Cash comp.	Sales	MB	Ex. cash	Sales growth	Tot. return	Book lev.
$\theta_{g_i,t}$	**0.055**	**0.030**	**0.020**	**-0.016**	**0.094**	-0.001	0.006	0.007	0.004

Panel B: Pre. vs. Post SOX

Vega Type	Vega	R&D	Capex
Pre-SOX			
Low	0.017	0.039	0.063
High	0.216	0.035	0.055
Low-High	-0.19***	0.004	0.008***
Post-SOX			
Low	0.018	0.041	0.043
High	0.240	0.030	0.043
Low-High	-0.22***	0.011***	0.0
Post-Pre Sox	-0.03 ***	0.007***	-0.008***

Hirfindahl-Hirschman Index across groups

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
HHI	0.008	0.010	0.011	0.012	1	0.014	0.051	0.013	0.020	0.013	1	1	0.045	1	0.007

References

Banerjee, S., M. H. Jenner, and V. Nanda (2015). "Restraining overconfident CEOs through improved governance: Evidence from the Sarbanes-Oxley Act." *Review of Financial Studies*, 28, 2812–2858.

Bargeron, L. L., K. M. Lehn, and C. J. Zutter (2010). "Sarbanes-oxley and corporate risk-taking." *Journal of Accounting and Economics*, 49(12), 34 – 52.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics*, 119(1), 249.

Bertrand, M. and A. Schoar (2003). "Managing with style: The effect of managers on firm policies*." *The Quarterly Journal of Economics*, 118(4), 1169.

Bonhomme, S. and E. Manresa (2015a). "Grouped patterns of heterogeneity in panel data." *Econometrica*, 83(3), 1147–1184.

Bonhomme, S. and E. Manresa (2015b). "Supplement to 'grouped patterns of heterogeneity in panel data'." *Econometrica Supplemental Material*, 83.

Brochet, F. (2010). "Information content of insider trades before and after the sarbanes-oxley act." *The Accounting Review*, 85(2), 419–446.

Campello, M., E. Giambona, J. R. Graham, and C. R. Harvey (2011). "Liquidity management and corporate investment during a financial crisis." *The Review of Financial Studies*, 24(6), 1944.

Campello, M., J. R. Graham, and C. R. Harvey (2010). "The real effects of financial constraints: Evidence from a financial crisis." *Journal of Financial Economics*, 97(3), 470 – 487.

Carter, M. E., L. J. Lynch, and S. L. C. Zechman (2009). "Changes in bonus contracts in the post-sarbanes–oxley era." *Review of Accounting Studies*, 14(4), 480–506.

Chhaochharia, V. and Y. Grinstein (2007). "Corporate governance and firm value: The impact of the 2002 governance rules." *The Journal of Finance*, 62(4), 1789–1825.

Cohen, D. A., A. Dey, and T. Z. Lys (2008). "Real and accrual-based earnings management in the pre- and post-sarbanes-oxley periods." *The Accounting Review*, 83(3), 757–787.

Coles, J. L., N. D. Daniel, and L. Naveen (2006). "Boards: Does one size fit all." *Journal of Financial Economics*, 79, 431–468.

Cronqvist, H., A. K. Makhija, and S. E. Yonker (2012). "Behavioral consistency in corporate finance: CEO personal and corporate leverage." *Journal of Financial Economics*, 103(1), 20 – 40.

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

Duchin, R., J. G. Matsusaka, and O. Ozbas (2010a). "When are outside directors effective?" *Journal of Financial Economics*, 96(2), 195 – 214.

Duchin, R., O. Ozbas, and B. A. Sensoy (2010b). "Costly external finance, corporate investment, and the subprime mortgage credit crisis." *Journal of Financial Economics*, 97(3), 418 – 435.

Engel, E., R. M. Hayes, and X. Wang (2007). "The sarbanes-oxley act and firms' going-private decisions." *Journal of Accounting and Economics*, 44(1&2), 116 – 145.

Forgy, E. (1965). "Cluster analysis of multivariate data: Efficiency versus interpretability of classification." *Biometrics*, 21(3), 768–769.

Frank, M. Z. and V. K. Goyal (2009). "Capital structure decisions: Which factors are reliably important?" *Financial Management*, 38(1), 1–37.

Gormley, T. A. and D. A. Matsa (2014). "Common errors: How to (and not to) control for unobserved heterogeneity." *Review of Financial Studies*, 27, 617–661.

Graham, J. R., C. R. Harvey, and M. Puri (2013). "Managerial attitudes and corporate actions." *Journal of Financial Economics*, 109(1), 103 – 121.

Heron, R. A. and E. Lie (2007). "Does backdating explain the stock price pattern around executive stock option grants?" *Journal of Financial Economics*, 83(2), 271 – 295.

Ho, P.-H., C.-W. Huang, C.-Y. Lin, and J.-F. Yen (2016). "CEO overconfidence and financial crisis: Evidence from bank lending and leverage." *Journal of Financial Economics*, 120(1), 194 – 209.

Joh, S. W. (2003). "Corporate governance and firm profitability: evidence from korea before the economic crisis." *Journal of Financial Economics*, 68(2), 287 – 322.

Koh, K., D. A. Matsumoto, and S. Rajgopal (2008). "Meeting or beating analyst expectations in the post-scandals world: Changes in stock market rewards and managerial actions." *Contemporary Accounting Research*, 25, 1067–1098.

Lemmon, M. L. and K. V. Lins (2003). "Ownership structure, corporate governance, and firm value: Evidence from the east asian financial crisis." *The Journal of Finance*, 58(4), 1445–1468.

Linck, J. S., J. M. Netter, and T. Yang (2009). "The effects and unintended consequences of the sarbanes-oxley act on the supply and demand for directors." *The Review of Financial Studies*, 22(8), 3287–3328.

Lins, K. V., P. Volpin, and H. F. Wagner (2013). "Does family control matter? International evidence from the 2008-2009 financial crisis." *The Review of Financial Studies*, 26(10), 2583–2619.

Lu, X. and L. Su (2017). "Determining the number of groups in latent panel structures with an application to income and democracy." *Quantitative Economics*, forthcoming.

Malmendier, U. and G. Tate (2015). "Behavioral CEOs: The role of managerial overconfidence." *The Journal of Economic Perspective*, 29(4), 37–60.

Malmendier, U., G. Tate, and J. Yan (2011). "Overconfidence and early-life experiences: The effect of managerial traits on corporate financial policies." *The Journal of Finance*, 66(5), 1687–1733.

Mitton, T. (2002). "A cross-firm analysis of the impact of corporate governance on the east asian financial crisis." *Journal of Financial Economics*, 64(2), 215 – 241.

Opler, T. C. and S. Titman (1994). "Financial distress and corporate performance." *The Journal of Finance*, 49(3), 1015–1040.

Petersen, M. A. (2009). "Estimating standard errors in finance panel data sets: Comparing approaches." *Review of Financial Studies*, 22, 435–480.

Steinley, D. (2006). "K-means clustering: A half-century synthesis." *British Journal of Mathematical and Statistical Psychology*, 59(1), 1–34.

Su, L., Z. Shi, and P. C. B. Phillips (2016). "Identifying latent structures in panel data." *Econometrica*, 84, 2215–2264.

Thakor, A. V. (2015). "The financial crisis of 2007-2009: Why did it happen and what did we learn?" *The Review of Corporate Finance Studies*, 4(2), 155–205.

Thompson, S. B. (2011). "Simple formulas for standard errors that cluster by both firm and time." *Journal of Financial Economics*, 99, 1–10.

Appendix

Event study standard errors

The transformed residuals in the case of standard event studies still contain two components, a group-specific time-varying component $\gamma_{g_i,t}$ and an idiosyncratic component η_{it} for each individual unit, namely

$$\dot{\epsilon}_{it} = \gamma_{g_i,t} + \eta_{it}. \tag{11}$$

Also, the transformed independent variable $\dot{X}$ is driven by a group-specific time-varying component and an idiosyncratic component, and it can specified as

$$\dot{X}_{it} = \delta_{g_i,t} + \nu_{it}. \tag{12}$$

Here all four components $\gamma_{g_i,t}$, η_{it}, $\delta_{g_i,t}$, and ν_{it} are assumed to be independent with each other and across time, and they all have a zero mean and finite variance, σ^2_γ, σ^2_η, σ^2_δ, and σ^2_ν, respectively.

To see how the correlation causes bias in standard errors, we consider an example of a single regressor and thus X_{it} is a scalar. Due to the group-specific components, the residual and regressor of individual units (firms) are correlated with each other with the following correlation structure

$$\begin{aligned} corr(\dot{X}_{it}, \dot{X}_{js}) &= 1 \quad \text{for } i = j \text{ and } t = s \\ &= \rho_x = \sigma^2_\delta/\sigma^2_X \quad \text{for } g_i = g_j \text{ and } t = s \\ &= 0 \quad \text{for all } g_i \neq g_j \text{ or } t \neq s, \end{aligned}$$

and

$$\begin{aligned} corr(\dot{\epsilon}_{it}, \dot{\epsilon}_{js}) &= 1 \quad \text{for } i = j \text{ and } t = s \\ &= \rho_\epsilon = \sigma^2_\gamma/\sigma^2_\epsilon \quad \text{for } g_i = g_j \text{ and } t = s \\ &= 0 \quad \text{for all } g_i \neq g_j \text{ or } t \neq s. \end{aligned}$$

Under such residual and regressor structure, the fixed effects coefficient estimate is still consistent, but its standard error estimate is downward biased due to the ignorance of the group correlation. In particular, we can obtain the asymptotic variance of the fixed effects coefficient estimate as

$$\begin{aligned} var(\hat{\beta}_{FE}) &= \left[\left(\sum_{i=1}^{N} \sum_{t=1}^{T} \dot{X}_{it} \dot{\epsilon}_{it} \right)^2 \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \dot{X}^2_{it} \right)^{-2} \right] \\ &= \left[\left(\sum_{i=1}^{N} \sum_{t=1}^{T} \dot{X}^2_{it} \dot{\epsilon}^2_{it} + \sum_{i,j\in g} \sum_{t=1}^{T} \dot{X}_{it} \dot{X}_{jt} \dot{\epsilon}_{it} \dot{\epsilon}_{jt} \right) \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \dot{X}^2_{it} \right)^{-2} \right] \\ &= \left[NT\sigma^2_X \sigma^2_\epsilon + \sum_{g=1}^{G} N_g (N_g - 1) T \rho_X \rho_\epsilon \sigma^2_X \sigma^2_\epsilon \right] \left[NT\sigma^2_X \right]^{-2} \\ &= \frac{\sigma^2_\epsilon}{NT\sigma^2_X} \left[1 + 1/N \sum_{g=1}^{G} N_g (N_g - 1) \rho_X \rho_\epsilon \right], \end{aligned} \tag{13}$$

where N_g is the number of individuals in the g-th group. Under a group structure, individuals within a group share the same component in both the independent variable and residuals, leading to a positive correlation. Therefore, the fixed effects standard error estimates underestimate the true standard errors.