BANK OF ENGLAND

Appendix to Staff Working Paper No. 640

A discrete choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance

Lena Boneva and Oliver Linton

January 2017

BANK OF ENGLAND

Appendix to Staff Working Paper No. 640

A discrete choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance

Lena Boneva[1] and Oliver Linton[2]

(1) Bank of England. Email: lena.boneva@bankofengland.co.uk
(2) University of Cambridge. Email: obl20@cam.ac.uk

Information on the Bank's working paper series can be found at
www.bankofengland.co.uk/research/Pages/workingpapers/default.aspx

Publications and Design Team, Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 7601 4030 email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

Supplementary material for: A Discrete Choice Model For Large Heterogeneous Panels with Interactive Fixed Effects with an Application to the Determinants of Corporate Bond Issuance*

Lena Boneva† Bank of England and CEPR

Oliver Linton‡ University of Cambridge

January 3, 2017

1 Proofs

Proof of Lemma 4.1

We have

$$\max_{1\leq t\leq T}\|\widehat{h}_t-\overline{h}_t\|=\max_{1\leq t\leq T}\|\overline{u}_t\|=\max_{1\leq t\leq T}\left|\frac{1}{N}\sum_{i=1}^{N}u_{it}\right|.$$

Define the event

$$B=\left\{|u_{it}|\leq\tau_{N,T}\text{ for all }i\leq N,t\leq T\right\}, \tag{1}$$

where $\tau_{N,T}<\infty$ is to be determined below. We have for any $x>0$

$$\Pr\left(\sqrt{N}\max_{1\leq t\leq T}\|\widehat{h}_t-\overline{h}_t\|>x\right)\leq\Pr\left(\left\{\sqrt{N}\max_{1\leq t\leq T}\|\widehat{h}_t-\overline{h}_t\|>x\right\}\cap B\right)+\Pr(B^c)$$

*We would like to thank Peter Robinson, our editor Thierry Magnac and 3 anonymous referees for their helpful comments and Hashem Pesaran, Martin Weidner, Lu Liu, Menno Middeldorp and Magda Rutkowska for many useful discussions. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of England, the Monetary Policy Committee, the Financial Policy Committee or the Prudential Regulation Authority.

†Bank of England, Threadneedle St, London, EC2R 8AH. Email: Lena.Boneva@bankofengland.co.uk. Financial support from Cusanuswerk and the ESRC (grant no.: K67783X) is gratefully acknowledged.

‡Faculty of Economics, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD. Email: obl20@cam.ac.uk. Financial support from the Keynes Fund is gratefully acknowledged.

Then by Bonferroni and Bernstein's inequality (Van der Vaart, 1998, p285)

$$\Pr\left(\left\{\sqrt{N}\max_{1\le t\le T}\|\widehat{h}_t-\overline{h}_t\|>x\right\}\cap B\right)\le 2T\exp\left(-\frac{1}{2}\frac{x^2}{\sigma_u^2+x\tau_{N,T}/\sqrt{N}}\right),$$

where we use that u_{it} are i.i.d with mean zero and finite variance σ_u^2. Then, taking $x=\log T$ we have

$$\Pr\left(\left\{\sqrt{N}\max_{1\le t\le T}\|\widehat{h}_t-\overline{h}_t\|>\log T\right\}\cap B\right)\le 2T\exp\left(-\frac{1}{2}\frac{\log^2 T}{\sigma_u^2+\tau_{N,T}\log T/\sqrt{N}}\right)=o(1)$$

provided $\tau_{N,T}\log T/\sqrt{N}\to 0$. Furthermore, we note that with $\tau_{N,T}=(NT)^{\pi}$ for some $\pi>0$, we have

$$\Pr(B)=F_{|u|}\left(\tau_{N,T}\right)^{NT}\ge\left(1-\frac{c}{(NT)^{\pi\alpha}}\right)^{NT},$$

where $F_{|u|}$ denotes the c.d.f. of the random variable $|u_{it}|$. The moment conditions imply that $F_{|u|}(x)\ge 1-cx^{-\alpha}$ for x large for $\alpha\ge 4$. If $\pi\alpha>1$, then $\Pr(B)\to 1$. Therefore, provided $\pi>1/\alpha$ and $N^{\pi-1/2}T^{\pi}\log T\to 0$ the result is established.

Proof of Theorem 4.1

Because the infeasible estimator $\widetilde{\theta}_i$ is consistent, it suffices to show that estimating the unobserved factors does not affect the criterion function. We have

$$\begin{aligned}&\Pr\left(\sup_{\theta\in\Theta}\left|\widehat{Q}_T^i(\theta)-Q_T^i(\theta)\right|\ge\epsilon\right)\\&\le\Pr\left(\sup_{\|h-\overline{h}\|_{\mathcal{H}}\le\delta_T}\sup_{\theta\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\left|q_t^i(\theta,h_t)-q_t^i(\theta,\overline{h}_t)\right|\ge\epsilon\right)+\Pr\left(\|\widehat{h}-\overline{h}\|_{\mathcal{H}}>\delta_T\right)\qquad(2)\\&\to 0,\qquad(3)\end{aligned}$$

Proof of Theorem 4.2

To show that $\widehat{\theta}_i$ is asymptotically normal, it suffices to show that estimating the unobserved factors does not affect the limiting distribution, that is,

$$\sqrt{T}\left(\widehat{\theta}_i-\widetilde{\theta}_i\right)=o_P(1).\qquad(4)$$

By the Mean Value Theorem, we have

$$0=\frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widehat{\theta}_i,\widehat{h}_t)}{\partial\theta}$$

$$
\begin{aligned}
&= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widetilde{\theta}_i,\widehat{h}_t)}{\partial\theta} + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^*)}{\partial\theta\partial\theta^\intercal}\left(\widehat{\theta}_i-\widetilde{\theta}_i\right)\\
&= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widetilde{\theta}_i,\overline{h}_t)}{\partial\theta} + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^{**})}{\partial\theta\partial h^\intercal}\left(\widehat{h}_t-\overline{h}_t\right) + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^*)}{\partial\theta\partial\theta^\intercal}\left(\widehat{\theta}_i-\widetilde{\theta}_i\right)\\
&= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^{**})}{\partial\theta\partial h^\intercal}\left(\widehat{h}_t-\overline{h}_t\right) + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^*)}{\partial\theta\partial\theta^\intercal}\left(\widehat{\theta}_i-\widetilde{\theta}_i\right),
\end{aligned}
$$

where h_t^* and h_t^{**} are intermediate values. Then, provided

$$
\liminf_{N,T\to\infty}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^*)}{\partial\theta\partial\theta^\intercal}\right\| > 0
$$

$$
\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^{**})}{\partial\theta\partial h^\intercal}\left(\widehat{h}_t-\overline{h}_t\right)\right\| = o_P(T^{-1/2}).
$$

the result (4) follows. These properties follow from the uniform convergence of $\widehat{h}_t-\overline{h}_t$, the Cauchy-Schwarz inequality and conditions D. For example, with probability tending to one

$$
\begin{aligned}
\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^{**})}{\partial\theta\partial h^\intercal}\left(\widehat{h}_t-\overline{h}_t\right)\right\|^2 &\leq \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{\partial^2 q_t^i(\widetilde{\theta}_i,h_t^{**})}{\partial\theta\partial h^\intercal}\right\|^2 \times \left\|\widehat{h}-\overline{h}\right\|_{\mathcal{H}}^2\\
&\leq \sup_{\left\|h-\overline{h}\right\|_{\mathcal{H}}<\delta_T}\ \sup_{\|\theta-\theta_{0i}\|\leq\delta_T}\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{\partial^2 q_t^i(\theta,h_t)}{\partial\theta\partial\theta^\intercal}\right\|^2 \times \left\|\widehat{h}-\overline{h}\right\|_{\mathcal{H}}^2\\
&= O_P\left(\frac{\log^2 T}{N}\right) = o_P(T^{-1})
\end{aligned}
$$

Proof of Theorem 4.3

It suffices to show that the feasible objective functions are uniformly close to the infeasible ones, see CJL (2016). Thus

$$
\begin{aligned}
&\Pr\left(\max_{1\leq i\leq N}\sup_{\theta\in\Theta}\left|\widehat{Q}_T^i(\theta)-Q_T^i(\theta)\right|\geq\epsilon\right)\\
&\leq \Pr\left(\max_{1\leq i\leq N}\ \sup_{||h-\overline{h}\|_{\mathcal{H}}\leq\delta_T}\ \sup_{\theta\in\Theta}\left|\frac{1}{T}\sum_{t=1}^{T}q_t^i(\theta,h_t)-q_t^i(\theta,\overline{h}_t)\right|\geq\epsilon\right) + \Pr\left(||\widehat{h}-\overline{h}\|_{\mathcal{H}}>\delta_T\right)\\
&\leq \Pr\left(\max_{1\leq i\leq N}\ \sup_{||h-\overline{h}\|_{\mathcal{H}}\leq\delta_T}\ \sup_{\theta\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\left|q_t^i(\theta,h_t)-q_t^i(\theta,\overline{h}_t)\right|\geq\epsilon\right) + \Pr\left(||\widehat{h}-\overline{h}\|_{\mathcal{H}}>\delta_T\right)\\
&\to 0. \qquad (5)
\end{aligned}
$$

Proof of Theorem 4.4

We show that

$$\widehat{\theta} - \widetilde{\theta} = o_P(N^{-1/2}). \tag{6}$$

The estimators $\widetilde{\theta}_i$ and $\widehat{\theta}_i$, $i = 1, \ldots, N$ satisfy the first order conditions

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widetilde{\theta}_i, \overline{h}_t)}{\partial \theta} = 0 = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widehat{\theta}_i, \widehat{h}_t)}{\partial \theta}.$$

We first work with a linear approximation to $\widehat{\theta}_i$. Define

$$L_{Ti}(\theta) = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widetilde{\theta}_i, \widehat{h}_t)}{\partial \theta} + M_i\left(\theta - \widetilde{\theta}_i\right) \tag{7}$$

from which we obtain for $\widehat{\theta}_i^*$ such that $L_{Ti}(\widehat{\theta}_i^*) = 0$,

$$\widehat{\theta}_i^* - \widetilde{\theta}_i = -M_i^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widetilde{\theta}_i, \widehat{h}_t)}{\partial \theta}.$$

We first establish the result for this linear approximation. By the Mean-Value Theorem, we have for $r = 1, \ldots, p$

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\widetilde{\theta}_i, \widehat{h}_t)}{\partial \theta_r} &= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\theta_{0i}, \overline{h}_t)}{\partial \theta_r} + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\theta_{0i}, \overline{h}_t)}{\partial \theta_r \partial h^\intercal}\left(\widehat{h}_t - \overline{h}_t\right) + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\theta_{0i}, \overline{h}_t)}{\partial \theta_r \partial \theta^\intercal}\left(\widetilde{\theta}_i - \theta_{0i}\right) \\
&+ \frac{1}{2T}\sum_{t=1}^{T}(\widehat{h}_t - \overline{h}_t)^\intercal \frac{\partial^3 q_t^i}{\partial \theta_r \partial h \partial h}(\theta_i^*, h_t^*)(\widehat{h}_t - \overline{h}_t) + \left(\widetilde{\theta}_i - \theta_{0i}\right)^\intercal \frac{1}{2T}\sum_{t=1}^{T}\frac{\partial^3 q_t^i(\theta_i^*, \overline{h}_t^*)}{\partial \theta_r \partial \theta \partial h^\intercal}\left(\widehat{h}_t - \overline{h}_t\right) \\
&+ \left(\widetilde{\theta}_i - \theta_{0i}\right)^\intercal \frac{1}{2T}\sum_{t=1}^{T}\frac{\partial^3 q_t^i(\theta_i^*, \overline{h}_t^*)}{\partial \theta_r \partial \theta \partial \theta^\intercal}\left(\widetilde{\theta}_i - \theta_{0i}\right) \\
&= \sum_{k=1}^{6} J_{rk;i},
\end{aligned}$$

where h_t^* and θ_i^* are intermediate values. It follows that

$$\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\theta}_i^* - \widetilde{\theta}_i\right) = \sum_{k=1}^{6}\frac{1}{N}\sum_{i=1}^{N} M_i^{-1} \times J_{\cdot k;i} \equiv \sum_{k=1}^{6} R_{rk}, \tag{8}$$

where $J_{\cdot k;i}$ denotes the vector with r^{th} element $J_{rk;i}$. We consider in sequence the vector random variables $R_1 - R_6$.

We have $E(R_1) = 0$ and $\partial q_t^i(\theta_{0i}, \overline{h}_t)/\partial\theta$ is i.i.d. across i and t conditional on X, d, f, so that

$$R_1 = \frac{1}{N}\sum_{i=1}^{N} M_i^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\theta_{0i}, \overline{h}_t)}{\partial\theta} = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} M_i^{-1}\frac{\partial q_t^i(\theta_{0i}, \overline{h}_t)}{\partial\theta} = O_P(N^{-1/2}T^{-1/2}).$$

Consider

$$R_2 = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} M_i^{-1}\frac{\partial^2 q_t^i}{\partial\theta\partial h}(\theta_{0i}, \overline{h}_t)\right)(\widehat{h}_t - \overline{h}_t). \tag{9}$$

We have $\partial\Phi_{it0}/\partial\overline{h}_t = \phi_{it0}\overline{\kappa}_i$ and $\partial\phi_{it0}/\partial\overline{h}_t = -\theta_{i0}^{\intercal} z_{it}\phi_{it0}\overline{\kappa}_i$, whence

$$\begin{aligned}\frac{\partial^2 q_t^i}{\partial\theta\partial h}(\theta_{0i}, \overline{h}_t) = {} & \frac{1}{T}\frac{-\phi_{it0}^2 z_{it}\overline{\kappa}_i}{\Phi_{it0}(1-\Phi_{it0})} \\ & - \frac{1}{T}\frac{Y_{it} - \Phi_{it0}}{(\Phi_{it0}(1-\Phi_{it0}))^2}\left(1 - 2\Phi_{it0}\right))\,\phi_{it0}^2 z_{it}\overline{\kappa}_i \\ & - \frac{1}{T}\frac{Y_{it} - \Phi_{it0}}{\Phi_{it0}(1-\Phi_{it0})}\left(\theta_{i0}^{\intercal} z_{it}\right)\phi_{it0}^2\overline{\kappa}_i z_{it}.\end{aligned}$$

We decompose (9) into three terms: the second and third terms are just linear combinations of the random variables $Y_{it} - \Phi_{it0}$, which are i.i.d. mean zero conditional on the factors; the first term is different and we treat this more carefully. This term can be written as

$$W_{NT} = \frac{1}{N^2 T}\sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{t=1}^{T} r_i(d_t, f_t, \overline{u}_t, u_{it})u_{lt}$$

for some function $r_i(\cdot)$. Write for each $l = 1, \ldots, N$

$$r_i(d_t, f_t, \overline{u}_{it}, u_{it}) = r_i(d_t, f_t, \overline{u}_t^{-l}, u_{it}) + r_{i;3}(d_t, f_t, \overline{u}_t^{-l}, u_{it})\frac{u_{lt}}{N} + \frac{1}{2} + r_{i;33}(d_t, f_t, \overline{u}_t^{-l*}, u_{it})\frac{u_{lt}^2}{N^2}$$

for some intermediate value $\overline{u}_t^{-l*}$, where $\overline{u}_t^{-l} = \sum_{j\neq l} u_{jt}/N$ so that $\overline{u}_t - \overline{u}_t^{-l} = u_{lt}/N$. We have for $l \neq i$, $E\left(r_{i;3}(d_t, f_t, \overline{u}_t^{-l}, u_{it})u_{lt}\right) = 0$ and $E\left(\left|r_{i;33}(d_t, f_t, \overline{u}_t^{-l*}, u_{it})\right| u_{lt}^2\right) < \infty$. From this we obtain that

$$E(W_{NT}) = O(N^{-1}).$$

By similar arguments we obtain

$$E\left(W_{NT}^2\right) = \frac{1}{N^4T^2}\sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{i'=1}^{N}\sum_{l'=1}^{N}\sum_{t=1}^{T}\sum_{t'=1}^{T} E\left[r_i(d_t, f_t, \overline{u}_{it}, u_{it})r_{i'}(d_{t'}, f_{t'}, \overline{u}_{i't'}, u_{i't'})^{\intercal} u_{lt}u_{l't'}\right] = O(N^{-1}T^{-1}),$$

because whenever either $t \neq t'$ or all four indices in $\{i, i', l, l'\}$ are distinct, then the expectation is zero for the leave out case, or small otherwise. Therefore,

$$W_{NT} = O_P(N^{-1}) + O_P(N^{-1/2}T^{-1/2}). \tag{10}$$

In conclusion, $R_2 = O_P(N^{-1}) + O_P(N^{-1/2}T^{-1/2})$. We have

$$\begin{aligned}
R_3 &= \frac{1}{N}\sum_{i=1}^{N} M_i^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\theta_{0i},\overline{h}_t)}{\partial\theta\partial\theta^{\intercal}}\left(\widetilde{\theta}_i-\theta_{0i}\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(M_i^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\theta_{0i},\overline{h}_t)}{\partial\theta\partial\theta^{\intercal}}\right)\left(\widetilde{\theta}_i-\theta_{0i}\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\widetilde{\theta}_i-\theta_{0i}\right)+\frac{1}{N}\sum_{i=1}^{N}\left(M_i^{-1}\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\theta_{0i},\overline{h}_t)}{\partial\theta\partial\theta^{\intercal}}-I_p\right)\left(\widetilde{\theta}_i-\theta_{0i}\right)\\
&= o_P(N^{-1/2})+O_P(T^{-1}) = o_P(N^{-1/2}),
\end{aligned}$$

by Cauchy-Schwarz and the assumption that $T^2/N \to 0$. The remaining terms, $R_4 - R_6$ are also treated by crude bounding. For example, with probability tending to one

$$\begin{aligned}
&\max_{1\le i\le N}\left\|\frac{1}{N}\sum_{i=1}^{N}M_i^{-1}\frac{1}{2T}\sum_{t=1}^{T}(\widehat{h}_t-\overline{h}_t)^{\intercal}\frac{\partial^3 q_t^i}{\partial\theta_r\partial h\partial h}(\theta_{0i},h_t^*)(\widehat{h}_t-\overline{h}_t)\right\|\\
&\le \|\widehat{h}-\overline{h}\|_{\mathcal{H}}^2\times\frac{1}{2NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\left\|M_i^{-1}\right\|\sup_{\left\|h-\overline{h}\right\|_{\mathcal{H}}<\delta_T}\left\|\frac{\partial^3 q_t^i}{\partial\theta_r\partial h\partial h}(\theta_{0i},h_t)\right\|\\
&= O_P\left(\frac{\log^2 T}{N}\right) = o_P(N^{-1/2}).
\end{aligned}$$

$$\begin{aligned}
&\left\|\frac{1}{N}\sum_{i=1}^{N}M_i^{-1}\left(\widetilde{\theta}_i-\theta_{0i}\right)^{\intercal}\frac{1}{2T}\sum_{t=1}^{T}\frac{\partial^3 q_t^i(\theta_i^*,\overline{h}_t^*)}{\partial\theta_r\partial\theta\partial\theta^{\intercal}}\left(\widetilde{\theta}_i-\theta_{0i}\right)\right\|^2\\
&\le \max_{1\le i\le N}\left\|\widetilde{\theta}_i-\theta_{0i}\right\|^2\times\frac{1}{2NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|M_i^{-1}\right\|\sup_{\left\|h-\overline{h}\right\|_{\mathcal{H}}<\delta_T}\sup_{\|\theta-\theta_{0i}\|\le\delta_T}\left\|\frac{\partial^3 q_t^i(\theta,h_t)}{\partial\theta_r\partial\theta\partial\theta^{\intercal}}\right\|^2\\
&= O_P\left(\frac{\log^2 N}{T}\right) = o_P(N^{-1/2}).
\end{aligned}$$

Finally, we show that the linear approximation is very close to the actual score function of the feasible estimator

$$\begin{aligned}
&\sqrt{N}\max_{1\le i\le N}\sup_{\|\theta-\theta_{0i}\|\le\delta_N}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial q_t^i(\theta,\widehat{h}_t)}{\partial\theta}-L_{Ti}(\theta)\right\|\\
&= \sqrt{N}\max_{1\le i\le N}\sup_{\|\theta-\theta_{0i}\|\le\delta_T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\theta-\widetilde{\theta}_i\right)^{\intercal}\frac{\partial^2 q_t^i(\theta,\widehat{h}_t)}{\partial\theta\partial\theta^{\intercal}}\left(\theta-\widetilde{\theta}_i\right)\right\|\\
&\le \sqrt{N}\times\max_{1\le i\le N}\left\|\widetilde{\theta}_i-\theta_{0i}\right\|^2\times\max_{1\le i\le N}\sup_{\|\theta-\theta_{0i}\|\le\delta_T}\sup_{\left\|h-\overline{h}\right\|_{\mathcal{H}}<\delta_T}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t^i(\theta,h_t)}{\partial\theta\partial\theta^{\intercal}}\right\|
\end{aligned}$$

$$= o_P(1).$$

The argument follows as in CJL (2016, p68).

2 Additional tables for the simulation study

Tables 1-8 report the results for experiments 2-5 that are discussed in Section 5 of the main paper.

3 Tables for the robustness checks

Tables 9 to 12 report the results of our robustness checks that are discussed in Section 6.3 of the main paper.

Table 1: Small sample properties of the mean group estimator $\widehat{\beta}$: Experiment 2

T/N	Bias (× 1000)				RMSE (× 1000)				Power				Size				Coverage probability			
	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300
									Infeasible estimator											
50	2.588	3.869	2.846	3.258	36.37	25.64	17.93	14.53	0.2875	0.556	0.851	0.9605	0.0545	0.049	0.042	0.0525	0.9455	0.951	0.958	0.9475
100	0.107	1.415	0.8318	0.9228	23.31	16.49	11.33	9.243	0.551	0.882	0.995	0.9995	0.0465	0.0455	0.0375	0.039	0.9535	0.9545	0.9625	0.961
200	0.1966	0.2106	0.03057	0.03981	15.79	11.19	8.039	6.47	0.874	0.9935	1	1	0.044	0.0465	0.0475	0.0475	0.956	0.9535	0.9525	0.9525
300	-0.4161	-0.04835	0.3275	0.1518	12.74	9.074	6.46	5.131	0.965	1	1	1	0.0405	0.036	0.0445	0.039	0.9595	0.964	0.9555	0.961
									CCEMG estimator											
50	-1.487	0.6606	-0.0882	0.6682	36.15	25.12	17.64	14.18	0.2525	0.508	0.805	0.945	0.0565	0.0435	0.0485	0.0435	0.9435	0.9565	0.9515	0.9565
100	-4.385	-2.043	-2.021	-1.73	23.74	16.48	11.53	9.341	0.4755	0.837	0.9905	0.999	0.047	0.0505	0.0455	0.0475	0.953	0.9495	0.9545	0.9525
200	-4.581	-3.372	-2.91	-2.7	16.53	11.72	8.584	7.009	0.7905	0.984	1	1	0.0535	0.061	0.07	0.0655	0.9465	0.939	0.93	0.9345
300	-5.233	-3.586	-2.63	-2.578	13.71	9.784	6.978	5.745	0.925	0.999	1	1	0.063	0.064	0.0675	0.0665	0.937	0.936	0.9325	0.9335
									Naive estimator											
50	158.7	159.4	158.7	158.1	162.5	161.9	160.4	159.6	1	1	1	1	0.9985	1	1	1	0.0015	0	0	0
100	157.2	158.6	158.3	158.2	159	159.8	159.1	158.9	1	1	1	1	1	1	1	1	0	0	0	0
200	158.1	158.1	157.9	157.9	159	158.7	158.3	158.2	1	1	1	1	1	1	1	1	0	0	0	0
300	157.8	158	158.3	158.4	158.3	158.4	158.6	158.6	1	1	1	1	1	1	1	1	0	0	0	0

Notes: The data generating process is defined in (17)-(21) (in the main paper) except that $\beta_{1i} = 0.5$, for all i. The nominal size is 5% and power is computed under the alternative $\beta_1 = 0.45$. The number of replications is set to 2000.

Table 2: Small sample properties of the mean group estimator $\widehat{\beta}$: Experiment 3

T/N	Bias (× 1000)				RMSE (× 1000)				Power				Size				Coverage probability			
	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300
Infeasible estimator																				
50	1.604	2.332	2.829	2.293	34.93	25.5	17.7	14.86	0.2645	0.5355	0.862	0.954	0.045	0.05	0.0455	0.045	0.955	0.95	0.9545	0.955
100	1.35	0.4999	0.4462	0.686	23.16	16.77	11.83	9.817	0.567	0.8585	0.9925	1	0.046	0.0525	0.0475	0.052	0.954	0.9475	0.9525	0.948
200	0.8779	0.1244	-0.1658	0.4899	15.91	10.92	7.969	6.4	0.8725	0.994	1	1	0.039	0.036	0.0425	0.0365	0.961	0.964	0.9575	0.9635
300	0.3728	0.01739	0.1115	0.09328	12.83	9.209	6.444	5.294	0.9675	1	1	1	0.0365	0.0415	0.039	0.0415	0.9635	0.9585	0.961	0.9585
CCEMG estimator																				
50	-47.5	-46.79	-45.93	-45.71	57.33	52.61	49.07	48.35	0.0345	0.049	0.0485	0.084	0.276	0.477	0.7375	0.8565	0.724	0.523	0.2625	0.1435
100	-49.81	-49.47	-49.06	-48.69	54.28	52.06	50.51	49.76	0.0245	0.033	0.049	0.06	0.512	0.815	0.975	0.997	0.488	0.185	0.025	0.003
200	-50.81	-50.54	-50.53	-49.83	53	51.62	51.16	50.29	0.01	0.0155	0.016	0.0215	0.784	0.989	1	1	0.216	0.011	0	0
300	-51.54	-50.86	-50.3	-50.2	53.01	51.63	50.74	50.52	0.0075	0.0085	0.012	0.0205	0.9055	0.998	1	1	0.0945	0.002	0	0
Naive estimator																				
50	56.46	56.73	56.68	57.78	66.45	63.91	61.78	62.62	0.9115	0.9855	0.9995	0.9995	0.428	0.679	0.847	0.907	0.572	0.321	0.153	0.093
100	56.55	55.25	56.1	55.67	61.54	58.81	58.53	58.02	0.997	1	1	1	0.6635	0.8775	0.974	0.987	0.3365	0.1225	0.026	0.013
200	56.32	55.77	55.1	55.82	58.73	57.34	56.29	56.87	1	1	1	1	0.891	0.9915	1	1	0.109	0.0085	0	0
300	55.23	55.31	55.67	55.55	56.93	56.39	56.49	56.26	1	1	1	1	0.958	0.999	1	1	0.042	0.001	0	0

Notes: The data generating process is defined in (17)-(21) (in the main paper) except that $k_{ji2} \sim NID(0, 0.1)$. The nominal size is 5% and power is computed under the alternative $\beta_1 = 0.45$. The number of replications is set to 2000.

Table 3: Small sample properties of the mean group estimator $\widehat{\beta}$: Experiment 4

T/N	Bias (× 1000)				RMSE (× 1000)				Power				Size				Coverage probability			
	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300
	Infeasible estimator																			
50	2.221	1.969	2.262	2.177	38.46	26.95	19.07	15.65	0.2515	0.466	0.7935	0.929	0.0495	0.045	0.043	0.0475	0.9505	0.955	0.957	0.9525
100	-0.06905	0.765	0.8125	0.9479	24.85	17.22	12.25	10.08	0.505	0.8375	0.985	0.999	0.0535	0.045	0.0425	0.0475	0.9465	0.955	0.9575	0.9525
200	0.251	0.58	0.43	0.3177	16.47	11.69	8.278	6.73	0.838	0.9905	1	1	0.0365	0.039	0.037	0.0445	0.9635	0.961	0.963	0.9555
300	0.1099	-0.1997	-0.03769	-0.17	13.23	9.397	6.761	5.432	0.9575	0.9995	1	1	0.039	0.039	0.041	0.0335	0.961	0.961	0.959	0.9665
	CCEMG estimator																			
50	-52.04	-51.04	-50.59	-50.92	62.59	56.95	54.08	53.54	0.0595	0.058	0.0815	0.1005	0.344	0.548	0.7995	0.915	0.656	0.452	0.2005	0.085
100	-55.59	-54.52	-53.57	-53.44	60	56.86	54.99	54.55	0.049	0.0565	0.0875	0.1135	0.6535	0.9035	0.99	0.998	0.3465	0.0965	0.01	0.002
200	-56.03	-55.2	-54.65	-54.54	58.13	56.31	55.33	55.02	0.061	0.07	0.1085	0.141	0.926	0.9985	1	1	0.074	0.0015	0	0
300	-56.61	-55.81	-55.47	-55.12	57.94	56.56	55.88	55.43	0.065	0.0955	0.1475	0.181	0.9895	1	1	1	0.0105	0	0	0
	Naive estimator																			
50	102.3	103.1	102.5	101.3	109.3	108	106.7	105.1	0.9955	1	1	1	0.8765	0.9715	0.995	0.9995	0.1235	0.0285	0.005	0
100	100.4	100.7	101.3	101.5	103.6	102.9	103.1	103.1	0.9995	1	1	1	0.9915	0.9995	1	1	0.0085	0	0	0
200	101.1	100.6	101.1	101.1	102.7	101.6	102	101.9	1	1	1	1	1	1	1	1	0	0	0	0
300	100.7	100.4	100.5	100.7	101.8	101.1	101.1	101.3	1	1	1	1	1	1	1	1	0	0	0	0

Notes: The data generating process is defined in (17)-(21) and (22) (in the main paper). The nominal size is 5% and power is computed under the alternative $\beta_1 = 0.45$. The number of replications is set to 2000.

Table 4: Small sample properties of the mean group estimator $\widehat{\beta}$: Experiment 5

T/N	Bias (× 1000)				RMSE (× 1000)				Power				Size				Coverage probability			
	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300	50	100	200	300
									Infeasible estimator											
50	2.741	2.228	2.859	2.14	33.89	23.81	17.09	13.41	0.3005	0.5715	0.8825	0.9705	0.041	0.0465	0.0475	0.038	0.959	0.9535	0.9525	0.962
100	1.931	0.5019	0.5716	0.4187	22.8	15.45	11.26	9.172	0.5995	0.8905	0.996	1	0.053	0.0365	0.043	0.0505	0.947	0.9635	0.957	0.9495
200	0.1784	0.1349	0.1921	0.1349	15.37	10.74	7.521	6.425	0.8875	0.9935	1	1	0.0495	0.032	0.04	0.0405	0.9505	0.968	0.96	0.9595
300	-0.1834	-0.2191	0.0583	0.1146	12.55	8.898	6.36	5.097	0.9695	1	1	1	0.041	0.0435	0.0435	0.046	0.959	0.9565	0.9565	0.954
									CCEMG estimator											
50	0.9652	1.81	2.908	2.42	34.33	24.39	17.38	13.64	0.271	0.5495	0.87	0.973	0.041	0.0465	0.044	0.037	0.959	0.9535	0.956	0.963
100	-0.2707	-0.532	0.08839	0.1874	22.99	15.63	11.32	9.253	0.549	0.866	0.9945	0.9995	0.0495	0.04	0.046	0.047	0.9505	0.96	0.954	0.953
200	-2.298	-1.033	-0.3968	-0.2565	15.65	10.82	7.547	6.443	0.8455	0.99	1	1	0.047	0.035	0.041	0.0415	0.953	0.965	0.959	0.9585
300	-2.6	-1.443	-0.5476	-0.3022	12.94	9.048	6.409	5.125	0.949	1	1	1	0.051	0.0475	0.0445	0.046	0.949	0.9525	0.9555	0.954
									Naive estimator											
50	113.4	113.8	113.5	112.5	118.4	116.8	115.5	114.1	1	1	1	1	0.941	0.998	1	1	0.059	0.002	0	0
100	113.5	112.3	112.4	112.1	115.8	113.6	113.3	112.9	1	1	1	1	0.9995	1	1	1	5e-04	0	0	0
200	112.6	112.6	112.6	112.3	113.6	113.3	113	112.6	1	1	1	1	1	1	1	1	0	0	0	0
300	112.6	112.3	112.8	112.5	113.3	112.7	113.1	112.7	1	1	1	1	1	1	1	1	0	0	0	0

Notes: The data generating process is defined in (17)-(21) (in the main paper) except that $k_{ji2} = 0 \ \forall i, j$ and $\kappa_{2i} = 0 \ \forall i$. The nominal size is 5% and power is computed under the alternative $\beta_1 = 0.45$. The number of replications is set to 2000.

Table 5: Small sample properties of the marginal effect $\widehat{ME}$: Experiment 4

T/N	Bias (× 1000)				RMSE (× 1000)			
	50	100	200	300	50	100	200	300
				Infeasible estimator				
50	-4.939	-4.988	-4.945	-4.932	10.07	7.935	6.572	6.071
100	-2.63	-2.426	-2.451	-2.412	6.627	4.859	3.864	3.451
200	-1.231	-1.146	-1.193	-1.216	4.334	3.157	2.4	2.087
300	-0.8167	-0.9001	-0.8656	-0.9018	3.458	2.563	1.922	1.651
				CCEMG estimator				
50	-5.156	-5.015	-4.912	-4.96	10.06	7.908	6.482	6.087
100	-2.663	-2.493	-2.457	-2.43	6.666	4.881	3.875	3.449
200	-1.264	-1.176	-1.204	-1.222	4.411	3.19	2.416	2.108
300	-0.9125	-0.9155	-0.903	-0.9022	3.542	2.617	1.949	1.672
				Naive estimator				
50	54.03	54.25	54.07	53.73	55.17	55.17	54.87	54.46
100	57.75	57.95	58.01	58.21	58.32	58.37	58.38	58.55
200	59.99	59.9	59.97	59.97	60.27	60.1	60.15	60.14
300	60.52	60.38	60.53	60.49	60.7	60.52	60.64	60.61
				Linear probability estimator				
50	4.402	4.511	4.652	4.57	10.16	7.901	6.498	5.96
100	2.369	2.574	2.565	2.575	6.82	5.066	4.039	3.636
200	1.296	1.401	1.365	1.368	4.527	3.356	2.553	2.232
300	0.812	0.8319	0.8303	0.8355	3.631	2.657	1.95	1.664

Notes: The mean group estimator of the average marginal effect is is reported. The data generating process is defined in (17)-(21) and (22) (in the main paper). The number of replications is set to 2000.

Table 6: Small sample properties of the marginal effect $\widehat{ME}$: Experiment 5

T/N	Bias (× 1000)				RMSE (× 1000)			
	50	100	200	300	50	100	200	300
				Infeasible estimator				
50	-4.699	-4.763	-4.657	-4.821	10.04	7.882	6.44	5.971
100	-2.127	-2.531	-2.511	-2.549	6.709	5.037	4.044	3.619
200	-1.262	-1.289	-1.274	-1.286	4.599	3.35	2.509	2.242
300	-0.9358	-0.9462	-0.8643	-0.8519	3.759	2.741	2.037	1.703
				CCEMG estimator				
50	-5.962	-5.954	-5.863	-6.014	10.76	8.716	7.395	6.977
100	-2.843	-3.216	-3.225	-3.244	7.028	5.451	4.532	4.148
200	-1.672	-1.66	-1.652	-1.662	4.76	3.517	2.723	2.477
300	-1.195	-1.2	-1.119	-1.11	3.872	2.847	2.162	1.849
				Naive estimator				
50	40.28	40.56	40.45	40.03	41.5	41.38	41.04	40.56
100	44.19	43.91	43.92	43.82	44.74	44.25	44.19	44.05
200	45.81	45.84	45.82	45.75	46.06	46.01	45.94	45.85
300	46.49	46.36	46.53	46.41	46.65	46.47	46.61	46.48
				Linear probability estimator				
50	4.937	4.867	4.945	4.805	10.63	8.276	6.815	6.093
100	2.921	2.508	2.51	2.492	7.253	5.199	4.134	3.663
200	1.263	1.279	1.294	1.28	4.745	3.441	2.586	2.275
300	0.7918	0.791	0.857	0.87	3.859	2.752	2.082	1.734

Notes: The mean group estimator of the average marginal effect is is reported. The data generating process is defined in (17)-(21) (in the main paper) except that $k_{ji2} = 0 \; \forall i, j$ and $\kappa_{2i} = 0 \; \forall i$. The number of replications is set to 2000.

Table 7: Small sample properties of the marginal effect $\widehat{ME}$: Experiment 2

T/N	Bias (× 1000)				RMSE (× 1000)			
	50	100	200	300	50	100	200	300
				Infeasible estimator				
50	-4.741	-4.475	-4.7	-4.618	10.2	7.754	6.463	5.827
100	-2.622	-2.282	-2.425	-2.409	6.707	4.909	3.868	3.424
200	-1.268	-1.253	-1.313	-1.304	4.476	3.286	2.546	2.194
300	-0.9855	-0.8788	-0.7807	-0.8276	3.626	2.627	1.937	1.633
				CCEMG estimator				
50	-4.727	-4.542	-4.78	-4.672	10.25	7.779	6.516	5.86
100	-2.629	-2.333	-2.454	-2.432	6.764	4.935	3.903	3.445
200	-1.336	-1.308	-1.35	-1.338	4.575	3.336	2.589	2.219
300	-1.064	-0.9193	-0.8168	-0.851	3.673	2.665	1.964	1.648
				Naive estimator				
50	58.65	58.76	58.74	58.45	59.42	59.3	59.14	58.82
100	62.27	62.64	62.62	62.55	62.65	62.91	62.81	62.73
200	64.44	64.45	64.38	64.34	64.62	64.57	64.47	64.43
300	64.96	65.03	65.12	65.12	65.07	65.11	65.18	65.17
				Linear probability estimator				
50	5.321	5.5	5.257	5.382	11.03	8.676	7.041	6.571
100	2.726	2.936	2.806	2.847	7.057	5.369	4.222	3.827
200	1.328	1.407	1.34	1.355	4.684	3.468	2.626	2.264
300	0.7718	0.9124	0.992	0.9621	3.68	2.719	2.082	1.742

Notes: The mean group estimator of the average marginal effect is reported. The data generating process is defined in (17)-(21) (in the main paper) except that $\beta_{1i} = 0.5$, for all i. The number of replications is set to 2000.

Table 8: Small sample properties of the marginal effect $\widehat{ME}$: Experiment 3

T/N	Bias (× 1000)				RMSE (× 1000)			
	50	100	200	300	50	100	200	300
	Infeasible estimator							
50	-4.954	-4.816	-4.75	-4.867	10.01	7.908	6.449	6.085
100	-2.297	-2.499	-2.527	-2.469	6.54	5.096	4.029	3.582
200	-1.088	-1.267	-1.356	-1.176	4.451	3.223	2.554	2.094
300	-0.7677	-0.8723	-0.8377	-0.8485	3.592	2.663	1.949	1.685
	CCEMG estimator							
50	-6.268	-6.303	-6.2	-6.24	10.68	8.915	7.572	7.232
100	-3.381	-3.574	-3.594	-3.515	7.047	5.732	4.77	4.367
200	-1.975	-2.096	-2.229	-2.034	4.813	3.652	3.115	2.682
300	-1.55	-1.643	-1.594	-1.618	3.955	3.071	2.393	2.204
	Naive estimator							
50	36.45	36.53	36.5	36.91	38.17	37.86	37.56	37.9
100	40.32	39.82	40.22	39.99	41.16	40.46	40.69	40.48
200	42.07	41.96	41.72	42	42.47	42.25	41.96	42.22
300	42.32	42.34	42.47	42.46	42.59	42.54	42.63	42.61
	Linear probability estimator							
50	4.269	4.249	4.381	4.296	10.02	7.887	6.366	5.808
100	2.259	2.009	1.987	2.088	6.83	5.049	3.796	3.399
200	0.9125	0.7905	0.649	0.8506	4.57	3.164	2.33	1.986
300	0.4143	0.2925	0.3499	0.3362	3.726	2.657	1.858	1.56

Notes: The mean group estimator of the average marginal effect is reported. The data generating process is defined in (17)-(21) (in the main paper) except that $k_{ji2} \sim NID(0, 0.1)$. The number of replications is set to 2000.

Table 9: The effect of yields on bond issuance for US firms: controlling for liquidity

	Pre-crisis			Post-crisis		
	All	Financial	Other	All	Financial	Other
Coefficient estimates						
Yield	-0.127	-0.062	-0.184	-0.05	-0.06	-0.09
	(-1.727)	(-0.394)	(-2.116)	(-1.327)	(-1.046)	(-1.887)
Size	-0.073	-0.049	-0.09	-0.077	-0.051	-0.068
	(-0.403)	(-1.475)	(-0.35)	(-0.512)	(-1.206)	(-0.324)
Liquidity	2.474	2.398	2.604	15.561	-3.419	21.484
	(4.57)	(2.257)	(3.712)	(1.062)	(-0.724)	(1.072)
Marginal effects						
Yield	-0.016	-0.001	-0.022	-0.009	-0.004	-0.015
Size	-0.004	-0.01	-0.002	-0.011	-0.009	-0.01
Liquidity	0.39	0.456	0.389	4.79	-0.82	6.627
Observations	321	62	225	375	72	270

Notes: The dependent variable is 1 if a firm issues a bond in a particular month and zero otherwise. Yield is the firm-specific corporate bond yield, size is measured by assets/1000 and liquidity is the share of current debt among total debt. All specifications include a measure of credit supply (leverage in the broker-dealer market), the federal funds rate (pre-crisis only) and the change in Federal Reserve Holdings of Treasury Notes (post-crisis only) as common factors. Columns (1) and (4) use all firms, (2) and (5) use financial sector firms and (3) and (6) use all other firms (excluding mining and agriculture). t-statistics are shown in parenthesis.

Table 10: The effect of yields on bond issuance for US firms: corporate spreads instead of yields

	Pre-crisis			Post-crisis		
	All	Financial	Other	All	Financial	Other
Coefficient estimates						
Spread	-0.115	-0.261	-0.137	-0.044	-0.013	-0.098
	(-1.452)	(-1.334)	(-1.56)	(-1.049)	(-0.224)	(-1.845)
Size	0.085	-0.011	0.13	-0.23	-0.027	-0.29
	(0.337)	(-0.38)	(0.363)	(-3.18)	(-0.473)	(-2.965)
Marginal effects						
Spread	-0.016	-0.029	-0.018	-0.008	0.002	-0.016
Size	0.02	-0.004	0.032	-0.027	-0.006	-0.034
Observations	321	62	225	378	72	273

Notes: The dependent variable is 1 if a firm issues a bond in a particular month and zero otherwise. Spread is the firm-specific corporate bond yield minus the federal funds rate and size is measured by assets/1000. All specifications include a measure of credit supply (leverage in the broker-dealer market) and the change in Federal Reserve Holdings of Treasury Notes (post-crisis only) as common factors. Columns (1) and (4) use all firms, (2) and (5) use financial sector firms and (3) and (6) use all other firms (excluding mining and agriculture). t-statistics are shown in parenthesis.

Table 11: The effect of yields on bond issuance for US firms: alternative sample choices

a) Firms with at least 20 time series observations

	Pre-crisis			Post-crisis		
	All	Financial	Other	All	Financial	Other
Coefficient estimates						
Yield	-0.166	-0.118	-0.216	-0.057	-0.025	-0.109
	(-1.972)	(-0.691)	(-2.114)	(-1.382)	(-0.405)	(-2.099)
Size	0.043	0.063	0.031	-0.22	-0.009	-0.278
	(0.199)	(0.935)	(0.103)	(-2.921)	(-0.117)	(-2.755)
Marginal effects						
Yield	-0.02	-0.001	-0.026	-0.01	-0.002	-0.017
Size	-0.028	-0.004	-0.035	-0.028	-0.004	-0.035
Observations	337	64	237	387	74	279

b) Firms with at least 40 time series observations

	Pre-crisis			Post-crisis		
	All	Financial	Other	All	Financial	Other
Coefficient estimates						
Yield	-0.163	-0.193	-0.216	-0.034	-0.022	-0.083
	(-1.995)	(-1.149)	(-2.245)	(-0.869)	(-0.349)	(-1.695)
Size	0.105	0.022	0.12	-0.174	-0.014	-0.222
	(0.454)	(0.722)	(0.368)	(-2.369)	(-0.274)	(-2.219)
Marginal effects						
Yield	-0.02	-0.016	-0.026	-0.006	0.001	-0.013
Size	0.023	0	0.031	-0.02	-0.002	-0.026
Observations	301	58	212	366	70	263

Notes: The dependent variable is 1 if a firm issues a bond in a particular month and zero otherwise. Yield is the firm-specific corporate bond yield and size is measured by assets/1000. All specifications include a measure of credit supply (leverage in the broker-dealer market), the federal funds rate (pre-crisis only) and the change in Federal Reserve Holdings of Treasury Notes (post-crisis only) as common factors. Columns (1) and (4) use all firms, (2) and (5) use financial sector firms and (3) and (6) use all other firms (excluding mining and agriculture). t-statistics are shown in parenthesis.

Table 12: The effect of yields on bond issuance for US firms: firms with at least 2 issuances

	Pre-crisis			Post-crisis		
	All	Financial	Other	All	Financial	Other
Coefficient estimates						
Yield	-0.137	-0.252	-0.139	-0.066	0.034	-0.152
	(-1.195)	(-1.175)	(-0.981)	(-1.291)	(0.391)	(-2.424)
Size	0.108	-0.025	0.137	-0.181	-0.096	-0.216
	(0.392)	(-0.695)	(0.353)	(-2.312)	(-2.351)	(-1.999)
Marginal effects						
Yield	-0.02	-0.019	-0.021	-0.011	0.009	-0.024
Size	0.021	-0.007	0.03	-0.023	-0.015	-0.027
Observations	240	48	170	250	50	180

Notes: The dependent variable is 1 if a firm issues a bond in a particular month and zero otherwise. Yield is the firm-specific corporate bond yield and size is measured by assets/1000. All specifications include a measure of credit supply (leverage in the broker-dealer market), the federal funds rate (pre-crisis only) and the change in Federal Reserve Holdings of Treasury Notes (post-crisis only) as common factors. Columns (1) and (4) use all firms, (2) and (5) use financial sector firms and (3) and (6) use all other firms (excluding mining and agriculture). t-statistics are shown in parenthesis.