Response to Reviewer 1: `Cover letter for authors.pdf`

Jilei Lin and Daniel J. Eck

Thank you very much for taking the time to provide a thoughtful review of our manuscript. We think that these comments will go a long way to improve the exposition of our manuscript.

Comment 1: How could the method be applied to forecast two or more period after a shock?

Response: Thanks for pointing it out. The method can be extended to forecast two or more periods after a shock. There are two situations.

First, suppose that there is no additional shock after $T_1^* + 1$. From equation (1) in the manuscript, the forecast at time point $T_1^* + 2$ is

$$\hat{y}_{i,T_i^*+2} = \hat{\eta}_1 + \hat{\phi}_i \hat{y}_{1,T_1^*+1}^2 + \hat{\theta}_i' \mathbf{x}_{1,T_1^*+2}, \tag{*}$$

where $\hat{y}_{1,T_1^*+1}^2 = \hat{y}_{1,T_1^*+1}^1 + \hat{\alpha}_{T_1^*+1}$, $\hat{y}_{1,T_1^*+1}^1$ is the forecast without adjustment, $\mathbf{x}_{1,T_1^*+2}$ is either known or forecasted, and $\hat{\alpha}_{T_1^*+1}$ is the shock-effect estimator for the shock at $T_1^* + 1$. The forecast for $d > 2$ more periods is similarly stated as

$$\hat{y}_{i,T_i^*+d} = \hat{\eta}_1 + \hat{\phi}_i \hat{y}_{1,T_1^*+d-1} + \hat{\theta}_i' \mathbf{x}_{1,T_1^*+d},$$

where $\mathbf{x}_{1,T_1^*+d}$ is either known or forecasted.

Second, suppose that there is one more shock effect after $T_1^* + 1$, say, $\alpha_{T_1^*+d^*}$ for some $d^* > 1$. Without loss of generality, assume $d^* = 2$. Under our framework, users may want to construct another set of donor pool with time series that experienced shock effects similar to $\alpha_{T_1^*+2}$. Next, apply our proposed method to compute $\hat{\alpha}_{T_1^*+2}$, the estimator for the shock effect at $T_1^* + 2$. The forecast at the point $T_1^* + 2$ will be,

$$\hat{y}_{i,T_i^*+2} = \hat{\eta}_1 + \hat{\phi}_i \hat{y}_{1,T_1^*+1}^2 + \hat{\theta}_1' \mathbf{x}_{1,T_1^*+2} + \hat{\alpha}_{T_1^*+2}.$$

These two cases can be further combined to generalize forecasting for two or more periods. We added aspects of this response to the Discussion.

Comment 2: Page 3, line 30: Blundell and Bond [1998] model is a fixed effect dynamic panel and estimate the parameters using the panel analysis. It would be nice if you clarify in what sense your work is similar to them.

1

Response: The model of Blundell and Bond [1998] is

$$y_{it} = \alpha y_{i,t-1} + \beta_1' x_{it} + \beta_2' x_{it-1} + \eta_i + v_{it}, \quad i = 1,\ldots,N, \quad t = 2,\ldots,T.$$

where $|\alpha| < 1$. In contrast, our model is

$$y_{i,t} = \eta_i + \alpha_i D_{i,t} + \phi_i y_{i,t-1} + \theta_i' \mathbf{x}_{i,t} + \varepsilon_{i,t},$$

where $|\phi_i| < 1$. Our model is similar to the one of Blundell and Bond [1998] in the sense that we have the fixed-effect decomposition of the error term $\eta_i + \varepsilon_{i,t}$ and the autoregressive structure. However, our model allows for η_i as random effects, different time lengths for each time series, different autoregressive parameters, and a shock component. To avoid confusion, we remove this reference in the revised manuscript.

Comment 3: Page 4, line 50: Your paper considers forecasting a time-series model. It is confusing for readers to read your sentence In this article, we consider a dynamic panel data model...".

Response: Sorry for the confusion. It is corrected in the revised version.

Comment 4: Figures on page 5: Figure labels, (a) and (b), are missing. What does the black line in the top graph represent? Is it the realized observations?

Response: Thanks for pointing it out. The labels (a) and (b) are now added in the revised manuscript. The black line stands for the realized observations whereas the violet line stands for the fitted values.

Comment 5: Page 7, line 52: Are these time series estimates or panel data estimates? If panel estimators are used, then OLS estimators in this model results in inconsistency of the estimators. In that case, how wouldn't this inconsistency affect the results in section 3 about the unbiasedness?

Response: Sorry for not making it clear. They are time series estimates or the ordinary least squares (OLS) estimates but not panel data estimates. In fact, our method can be generalized to any model whose parameter can be estimated in an unbiased fashion. In the second paragraph of Section 6, we wrote

> "Our methodology is developed for autoregressive models, but it can be generalized to any setting where model parameters can be estimated unbiasedly, there is an additive shock-effect structure, and the time series in the donor pool are independent from the one of interest."

Comment 6: Page 8, line 46: It is not clear to met how you estimate α_i's by OLS. Could you give the expressions for $\hat{\alpha}_i$? Also, how can you identify α_i from η_i? Isn't $\hat{\alpha}_{\text{adj}}$ an estimator of $E(\alpha_1)$?

Response: Sorry for not making it clear. The estimation procedures can be explained in detail as follows. Note that for $i = 2, \ldots, n+1$, the design matrix for ith time series is $(\mathbf{1}_{T_i}, \mathbf{D}_i, \mathbf{Y}_i^{(1)}, \mathbf{X}_i)$, where $\mathbf{X}_i$ is the covariates of ith time series, $\mathbf{D}_i \in \mathbb{R}^{T_i}$ with $D_{i,t} = I(t = T_i^* + 1)$, and $\mathbf{Y}_i^{(1)}$ is the lagged response $(y_{i,1}, \ldots, y_{i,T_i})'$. We define $\mathbf{M}_i = (\mathbf{1}_{T_i}, \mathbf{D}_i, \mathbf{Y}_i^{(1)}, \mathbf{X}_i)$. Then the expression for $\hat{\alpha}_i$ is

$$\hat{\alpha}_i = [(\mathbf{M}_i'\mathbf{M}_i)^{-1}\mathbf{M}_i'\mathbf{Y}_i]_2, \quad \text{for } i = 2, \ldots, n+1,$$

where $[\cdot]_2$ denotes the second element of a vector. That is, $\hat{\alpha}_i$ is the second element of the OLS estimate vector. Our model is

$$y_{i,t} = \eta_i + \alpha_i D_{i,t} + \phi_i y_{i,t-1} + \theta_i' x_{i,t} + \varepsilon_{i,t}, \quad i = 1, \ldots, n+1.$$

the intercept is denoted as η_i. Thus, we cannot identify α_i from η_i. As defined in (4) in the manuscript, through averaging, $\hat{\alpha}_{\text{adj}}$ is an estimator of $E(\alpha_1)$.

Comment 7: Page 9, line 10: It is not clear to me what $\mathbf{U}_i$'s are. An example could help understanding that better.

Response: Thanks for pointing it out. This quantity is the model (design) matrix corresponding to the ith time series. We clarify this quantity the revised version, and we change the notation to be $\mathbf{M}_i$. Our previous response describes $\mathbf{M}_i$ in more detail.

Comment 8: Page 9, line 15: Why the closed form expressions for $E(\hat{\alpha}_{\text{IVW}})$ and $Var(\hat{\alpha}_{\text{IVW}})$ are not provided?

Response: In page 8, we wrote

> "The inverse-variance weighted estimator is defined as
>
> $$\hat{\alpha}_{\text{IVW}} = \frac{\sum_{i=2}^{n+1} \hat{\alpha}_i / \hat{\sigma}_{i\alpha}^2}{\sum_{i=2}^{n+1} 1 / \hat{\sigma}_{i\alpha}^2}, \quad \text{where} \quad \hat{\sigma}_{i\alpha}^2 = \hat{\sigma}_i^2 (\mathbf{M}_i'\mathbf{M}_i)_{22}^{-1},$$
>
> where $\hat{\alpha}_i$ is the OLS estimator of α_i, $\hat{\sigma}_i$ is the residual standard error from OLS estimation, and $\mathbf{M}_i$ is the design matrix for OLS with respect to time series for $i = 2, \ldots, n+1$. Note that since σ is unknown, estimation is required and the numerator and denominator terms are dependent in general."

To be more specific, $\hat{\sigma}_{i\alpha}$ depends on $\hat{\alpha}_i$. As a result, the numerator and denominator of $\hat{\alpha}_{\text{IVW}}$ are dependent. Without making distributional assumptions, it is impossible to evaluate $E(\hat{\alpha}_{\text{IVW}})$ and $Var(\hat{\alpha}_{\text{IVW}})$ in a closed form.

Comment 9: Page 13, line 14: μ_α is unknown, and in practice we have to estimate it. Would the results hold if one replaces μ_α with its estimate?

Response: We go for a plug-in approach. If we can estimate μ_α well, the results should hold with high probability. Section 3.2 is dedicated to this estimation problem. Using plug-in approaches, we get a decision rule to determine whether the scalar adjustment will be beneficial with correctness probability p. Section 3.3 details the leave-one-out cross validation procedures, which are tailored to our setup, to estimate p. As a result, users can be informed about the credibility of our methods.

Comment 10: Page 14, line 7: Shouldnt $E(\alpha_1)$ be μ_α? Previously μ_α was used.

Response: In the line 7 at Page 14, $E(\alpha_1)$ is taken under $\mathcal{M}_{21}$ or $\mathcal{M}_{22}$. In contrast, μ_α is used under $\mathcal{M}_1$. In Section 2.1, we define the distribution of α_i as

$$\alpha_i \sim \mathcal{F}_\alpha \quad \text{with} \quad E_{\mathcal{F}_\alpha}(\alpha_i) = \mu_\alpha, Var_{\mathcal{F}_\alpha}(\alpha_i) = \sigma_\alpha^2, \quad \text{iid under } \mathcal{M}_1$$
$$\alpha_i = \mu_\alpha + \delta_i' \mathbf{x}_{i,T_i^*+1} + \tilde{\varepsilon}_i \quad \text{independent under } \mathcal{M}_{21} \text{ or } \mathcal{M}_{22}.$$

In other words, $\mathcal{M}_{21}$ or $\mathcal{M}_{22}$ is an extended version of $\mathcal{M}_1$ by allowing it to depend on some covariates. However, in general, $E(\alpha_1) \neq \mu_\alpha$ under $\mathcal{M}_{21}$ or $\mathcal{M}_{22}$ due to the presence of $\delta_i' \mathbf{x}_{i,T_i^*+1}$.

Comment 11: Page 18, numerical example: The model setup does not contain the individual effects (intercepts) similar to the models in equations (1)-(2) on page 6. How would the result change if you include them with different variances?

Response: The result wouldn't change. In fact, in our proof listed in the Appendix, we do not make use of the assumption that η_i follows the same distribution. In other words, the results still hold when the individual effects have different variances.

Comment 12: Page 23, construction of donor pool: Your model assumes that the shocks have the same distribution. Given that these shocks are from different sources and happened because of various reasons, how could one justify them? Practically, there are many previous shocks. How can someone choose from them?

Response: Our methodology requires that a practitioner can choose a suitable donor pool of candidate shock-effects. If this cannot be done, then the methodology may fail. In real-life applications, practitioners can use qualitative and quantitative information to inform their decisions about which time series are included in the donor pool. Consider our Conoco Phillips example. On March 9th 2020, Conoco Phillips was expected to experience a negative shock in the midst of the COVID-19 pandemic and an oil supply shock. While the COVID-19 pandemic is unprecedented, several, although not all, economic measurements surrounding March 9th 2020 resembled those in the days of the 2008 financial crises. On the other hand, the oil supply shock had precedent. As a result, we selected the shocks in the past corresponding to the 2008 financial crisis and a past 2014 oil supply shock to form our donor pool. Adjustment with respect to this donor pool led to much better forecasts than doing nothing.

Comment 13: Figure 2 on Page 24: Are the black dots the real data? What does the line represent?

Response: Thanks for pointing it out. We clarify this in the revised version. The black dots represent the real data. The line in violet color represents the fitted line of our model.

References

Richard Blundell and Stephen Bond. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143, 1998.