# Derivation of Asymptotic Covariance Matrix

Xu Cheng and Eric Renault and Paul Sangrey

August 12, 2018

replaced by $\xi$ later

called $\omega_s$ later

Let $\omega := (\rho, c, \delta, \gamma, \beta, \psi, \phi, \pi, \theta)'$ be the vector of parameters. We split $\omega$ into two parts. The first, $\omega_r := (\rho, c, \delta, \gamma, \beta, \psi, \phi)'$, is the vector of reduced-form parameters. The second, $\xi := (\phi, \pi, \theta)'$, is the vector of structural parameters. We further split $\xi$ into three parts. $\xi_1 := (\rho, c, \delta)'$ consists of the reduced-form parameters that we estimate using $GMM$ and will remain in the limit. The second $\xi_2 := (\gamma, \beta, \psi)'$ consists of three parameters that are eliminated in our final results. We view them as functions of the other parameters, and estimate them in the first stage, but "invert" them in the second stage, to estimate the final parameters. The third $\xi_3 := \phi^2$ will be estimated by itself, but it will not remain in the final analysis. In the second stage, we use $\xi_3$ to estimate $\phi$.

## 1 Stage 1

We view estimating the first stage as a particular form of GMM. From standard GMM theory, we know that the following holds.

$\Omega_\xi$

$$\sqrt{T}(\hat{\xi} - \xi) \xrightarrow{d} N\left(0, \Omega_\xi\right) \tag{1}$$

We will construct $\Omega$ in two steps. First, we will derive the asymptotic covariance matrices for each of $\xi_1, \xi_2, \xi_3$. Then we will show how to combine them into one joint covariance matrix.

### 1.1 Step 1: $\xi_1$

The conditional moments (**) & (**) implies $E[h(\sigma_t^2, \sigma_{t+1}^2, \xi_1)] = 0$

add

$$h(\sigma_t^2, \sigma_{t+1}^2, \xi_1) := \begin{bmatrix} -c\delta - \rho\sigma_t^2 + \sigma_{t+1}^2 \\ \sigma_t^2\left(-c\delta - \rho\sigma_t^2 + \sigma_{t+1}^2\right) \\ -c^2\delta - 2c\rho\sigma_t^2 + \sigma_{t+1}^4 - \left(c\delta + \rho\sigma_t^2\right)^2 \\ \sigma_t^2\left(-c^2\delta - 2c\rho\sigma_t^2 + \sigma_{t+1}^4 - \left(c\delta + \rho\sigma_t^2\right)^2\right) \\ \sigma_t^4\left(-c^2\delta - 2c\rho\sigma_t^2 + \sigma_{t+1}^4 - \left(c\delta + \rho\sigma_t^2\right)^2\right) \end{bmatrix} \tag{2}$$

They read better if change the sign of $c\delta$ and $c^2\delta$ to "+"

By standard GMM theory, the following holds if we construct the weight matrix $W_{\xi_1,T}$ such that $W_{\xi_1,T} \xrightarrow{p} \mathbb{V}\mathrm{ar}(h(\sigma_t^2, \sigma_{t+1}^2, \xi_1))^{-1}$ We have

weighting

$$\sqrt{T}(\hat{\xi}_1 - \xi_1) \xrightarrow{d} N\left(0, \Omega_{\xi_1}\right). \tag{3}$$

~~In addition,~~ where

plug in

$$\Omega_{\xi_1} := \left[\mathbb{E}\left[h_{\xi_1}(\sigma_t^2, \sigma_{t+1}^2, \xi_1)\right]' W_{\xi_1} \mathbb{E}\left[h_{\xi_1}(\sigma_t^2, \sigma_{t+1}^2, \xi_1)\right]\right]^{-1} \tag{4}$$

typically, we write it as $E[hh']$, not var.

We estimate this by replacing the population expectations and covariances by their sample counterparts.

$h_{\xi_1}(\cdot)$ not defined

In OLS with homosked- $\text{var}(\hat\beta) = E(X_t X_t')^{-2} \cdot \sigma^2$

### 1.2 Step 2: $\xi_2$

We estimate $\xi_2$ by weighted least squares, a special case of GMM.

$$\mathbb{E}\left[r_{t+1} \,\middle|\, \sigma_t^2, \sigma_{t+1}^2\right] = \gamma + \beta\sigma_t^2 + \psi\sigma_{t+1}^2 \tag{5}$$

this is just $\sigma^2$

The only unusual part is we know that $\mathbb{V}\text{ar}\left(r_{t+1} \,\middle|\, \sigma_t^2, \sigma_{t+1}^2\right) = (1-\phi^2)\sigma_{t+1}^2$. Consequently, the regression results are more efficient if we adjust for heteroskedasticity. However, we do not know $\phi$, and so this might seem impossible. However, since time-invariant parts of heteroskedasticity adjustments cancel, reweighting by the inverse of $\sigma_{t+1}^2$ achieves is equivalent to the optimal reweighting. Also, since $\sigma_{t+1}^2$ is contained in the conditioning set, the fact that it is viewed as a random variable in other parts of the regression is irrelevant.

Since this regression is exactly identified, any positive-definite weight matrix, including the identity is optimal. Consequently, we have the following result, where $\Omega_{\xi_1} = \mathbb{V}\text{ar}(\frac{r_{t+1}-\gamma-\beta\sigma_t^2-\psi\sigma_{t+1}^2}{\sigma_{t+1}^2})$, i.e. the standard WLS covariance matrix.

$\Omega_{\xi_2}$

$X_t = \begin{pmatrix} \frac{1}{\sigma_t^2} \\ \sigma_{t+1}^2 \end{pmatrix}$

### 1.3 Step 3: $\xi_3$

consistently

We know that $\mathbb{V}\text{ar}(r_{t+1}|\sigma_{t+1}^2, \sigma_t^2) = (1-\phi^2)\sigma_{t+1}^2$. This implies $\mathbb{V}\text{ar}(\frac{r_{t+1}}{\sigma_{t+1}}|\sigma_{t+1}^2, \sigma_t^2) = 1-\phi^2$. Since we estimated the conditional mean of $r_{t+1}$ in section 1.2, ~~this implies that~~ the residuals $\widehat{u}_t = \frac{r_{t+1}-\widehat{\gamma}-\widehat{\beta}\sigma_t^2-\widehat{\psi}\sigma_{t+1}^2}{\sigma_{t+1}}$ satisfy ies $\frac{1}{T}\sum_{t=1}^T \hat{u}_t^2 \xrightarrow{p} (1-\phi^2)$.

Also, $\phi^2 \xrightarrow{d} N(0, \Omega_{\xi_3})$, since this is a GMM estimator. The question is what is $\Omega_{\xi_3}$. Since we are just identified, standard GMM theory tells us that it will be the covariance of the moment condition scaled by the appropriate derivative. However, since we are estimating a mean shifted by a constant, this derivative is just one. Consequently, $\Omega_{\xi_3} = \mathbb{V}\text{ar}(\frac{u_t^2}{\sigma_{t+1}^2})$, which can be estimated by $\frac{1}{T}\sum_{t=1}^T(\frac{\widehat{u}_t^2}{\sigma_{t+1}^2} - \frac{1}{T}\sum_{t=1}^T \frac{\widehat{u}_t^2}{\sigma_{t+1}^2})^2$, i.e. the sample covariance of the squared residuals from the regression in the previous stage.

Ut is not defined

## 2 Combining $\Omega_{\xi_1}, \Omega_{\xi_2}$, and $\Omega_{\xi_3}$

Each of $\Omega_{\xi_i}$ are of the form $\mathbb{E}[h_{\xi_i}]' \,\mathbb{V}\text{ar}(h(\sigma_{t+1}^2, \sigma_t^2, \xi_i))^{-1}\,\mathbb{E}[h_{\xi_i}]$. Consequently, off-diagonal blocks of the joint covariance matrix $\Omega$ can come from two places. They can come from the derivatives or the covariance of the moments. Since the moments in the first stage do not depend on the parameters in the second stage, and vice-versa, no co-movement can be coming from the derivatives between them. The other cases are trickier, and so we will consider them each in turn.

Consider the covariance between $h(\sigma_{t+1}^2, \sigma_t^2, \xi_1)$ and $h(\sigma_{t+1}^2, \sigma_t^2, \xi_2)$. For some functions $h_1$, $h_2$ we can rewrite them as follows.

This is not true. The GMM for $\xi_3$ is not this. I can revise this.

$$\mathbb{C}\text{ov}\left(h_1\left(r_{t+1}, \sigma_{t+1}^2, \sigma_t^2\right), h_2\left(\sigma_{t+1}^2, \sigma_t^2\right)\right) \tag{6}$$

Since the moments are mean zero by construction.

$$= \mathbb{E}\left[h_1(r_{t+1}, \sigma_{t+1}^2, \sigma_t^2) h_2\left(\sigma_{t+1}^2, \sigma_t^2\right)\right] \tag{7}$$

By the law of iterated expectations.

$$= \mathbb{E}\left[\mathbb{E}[h_1(r_{t+1}, \sigma_{t+1}^2, \sigma_t^2) \,|\, \sigma_{t+1}^2, \sigma_t^2]\,\mathbb{E}\left[h_2\left(\sigma_{t+1}^2, \sigma_t^2\right) \,\middle|\, \sigma_{t+1}^2, \sigma_t^2\right]\right] \tag{8}$$



could you define $h_1(r_{t+1}, \sigma_{t+1}^2, \sigma_t^2)$ & $h_2(\sigma_{t+1}^2, \sigma_t^2)$ explicitly?

we want to show: $E(u_t^3 | \sigma_t^2, \sigma_{t+1}^2) = 0$ due to conditional Gaussian.

Then by LIE, $E\left(\frac{u_t^3}{\sigma_{t+1}^3}\right) = 0$

The first term in the expression above equals zero, and, hence, so does the entire expression. In other words, the first two set of moment conditions are independent. By an identical argument, the first and third moments are also independent.

asymptotically

The question at hand is how are the second and third moments related. Since the derivatives are with respect to different parameters (and constant) no dependence arises from there. The question is how are the moment conditions in the second and third steps related. The second stage moment condition is a conditional mean and third stage moment is a conditional covariance. Let $u_t$ denote the error term in that regression (as it did above).

"|" = "|"

$$\mathbb{E}\left[\mathbb{E}\left[\frac{r_{t+1} - \mathbb{E}\left[r_{t+1} \mid \sigma_{t+1}^2, \sigma_t^2\right]}{\sigma_{t+1}}\right] \mathbb{E}\left[\frac{(r_{t+1} - \mathbb{E}\left[r_{t+1} \mid \sigma_{t+1}^2, \sigma_t^2\right])^2}{\sigma_{t+1}^2}\right]\right] \mathbb{E}\left[\frac{u_t u_t^2}{\sigma_{t+1}^2}\right] \tag{9}$$

3

should be clearer on the conditioning.

Then since $u_t$ is conditionally Gaussian, its conditional (and hence unconditional) third moment is zero.

In other words, these values are also uncorrelated. Now, the careful reader might be worried about filling in the population expectations instead of their estimators in the regression above. However, since the expectations are linear and consistently estimable, this error vanishes in the limit. Intuitively, OLS mean and variance estimates are asymptotically independent.

In addition, since all three components are asymptotically independent; the inverse of a block-diagonal matrix is block-diagonal, and we using optimal weighting matrices in each part, we are using an optimal weighting matrix for $\xi$, not just its components.

## 3 Stage 2

In this stage, we convert the reduced-form parameters we estimated in the first stage into estimates of the structural parameters. The three parameters we need to estimate are $\pi, \theta$ and $\phi$. (We estimated $\phi^2$ in the previous stage, but we could not estimate its sign. We have the following four link functions. $\psi(\omega_s, \xi_1), \gamma(\omega_s, \xi_1), \beta(\omega_s, \xi_1), \xi_3(\omega_s)$. We denote the stacked link function $g(\omega_s, \xi)$.

The only confusing thing is because a few of the parameters are both structural and reduced form parameters, they show in $g(\omega_s, \xi)$ on both sides. Then the 2nd-stage sample criterion function is

This is NOT true! $\psi, \gamma, \beta, \phi$ are reduced form & $\pi, \theta, \phi$ are structural

weighting

$$Q_T(\omega) = \frac{1}{2} g\left(\omega_s, \widehat{\xi}\right)' \mathcal{W}_T g\left(\omega_s, \widehat{\xi}\right). \tag{10}$$

with second stage weight matrix $\mathcal{W}_T$. We want to estimate $\omega_s$, and so we differentiate and get the first-order condition

$$\frac{\partial Q(\omega)}{\partial \omega_s} = g_{\omega_s}\left(\omega_s, \hat{\xi}\right)' \mathcal{W}_T g\left(\omega_s, \hat{\xi}\right) = 0. \tag{11}$$

define it first

We now expand $\xi$ around $\xi_0$

$\to w_{s,0}$

$$\sqrt{T}\frac{\partial Q(\omega)}{\partial \omega_s} = g_{\omega_s}\left(\omega_s, \hat{\xi}\right)' \mathcal{W}_T \left[\sqrt{T} g\left(\omega_{s,0}, \xi_0\right) + g_\xi\left(\omega_{s,0}, \tilde{\xi}\right)' \sqrt{T}\left(\tilde{\xi} - \xi_0\right)\right], \tag{12}$$

The first term equals zero because we choose $\omega_s$ to make it hold in-sample.

$$= g_{\omega_s}\left(\omega_s, \hat{\xi}\right)' \mathcal{W}_T \left[g_\xi\left(\omega_{s,0}, \tilde{\xi}\right)' \sqrt{T}\left(\tilde{\xi} - \xi_0\right)\right]. \tag{13}$$



$g(w_{s,0}, \xi_0)$ holds in population. by the model

This is only 0 for $w_{s,0}$

We also need to compute the Hessian:

$$\frac{\partial^2 Q(\tilde{\omega}_s, \xi)}{\partial \omega_s \partial \omega_s'} = g_{\omega_s}\left(\tilde{\omega}_s, \hat{\xi}\right)' \mathcal{W}_T g_{\omega_s}\left(\tilde{\omega}_s, \hat{\xi}\right) + g_{\omega_s,\omega_s}\left(\tilde{\omega}_s, \hat{\xi}\right) \mathcal{W}_T g\left(\tilde{\omega}_s, \hat{\xi}\right)' \tag{14}$$

$$\xrightarrow{p} g_{\omega_s}(\omega_s, \xi_0)' \mathcal{W} g_{\omega_s}(\omega_s, \xi_0) + 0. \tag{15}$$

$$w_{s,0} \qquad w_{s,0} \qquad \text{for } \tilde{w}_s \to_p w_{s,0} \tag{16}$$

Let $B := g_{\omega_s}(\omega_{s,0}, \xi_0)' \mathcal{W} g_{\omega_s}(\omega_{s,0}, \xi_0)$ Then by the extremum estimator theory, we have

$$\sqrt{T}\left(\hat{\omega}_s - \omega_{s,0}\right) \xrightarrow{d} N\left(0, B^{-1} \mathbb{E}[g_{\omega_s}(\theta_0, \xi_0)]' \Omega_\xi \, \mathbb{E}[g_{\omega_s}(\theta_0, \xi_0)] B^{-1}\right). \tag{17}$$

The covariance in the middle is GMM-covariance of the reduced-form parameters. The optimal weight matrix is $\mathcal{W} = (g_\xi' \Omega_\xi g_\xi)^{-1}$. We can estimate it by plugging $\widehat{\xi}$ into to the formulas above and their derivatives. In this case, the $\Omega_{w_s}$ equals

$$B^{-1} g_{\omega_s}' \mathcal{W} g_{\omega_s} B^{-1}. \tag{18}$$

Let $g_w = g_{w_s}(w_{s,0}, \xi_0)$,

$g_\xi = g_\xi(w_{s,0}, \xi_0)$.

$g_w' W g_\xi' \Omega_\xi g_\xi \cdot W g_w$

under $W_T \to_p W$

becomes

$(g_w' \, g_\xi' \Omega_\xi g_\xi \, g_w)^{-1}$