Let
$$C(\hat{\tilde{\boldsymbol{\tau}}}) = \frac{1}{\hat{c}^{-1}\left(\hat{G}_j(\hat{\tau}_i^{j\star}) - \hat{G}_j(\hat{\tilde{\boldsymbol{\tau}}})\right) - 1}$$
and
$$Y_{ji}(\hat{\tilde{\boldsymbol{\tau}}}; \boldsymbol{b}) = \hat{G}_j(\hat{\tilde{\boldsymbol{\tau}}}) - \hat{G}_j(\hat{\tau}_i^{j\star})$$
Note first that
$$\begin{aligned}
\chi_{ji}(\epsilon_{ji}^{\star}) &= \frac{e^{-\boldsymbol{\alpha}^T \boldsymbol{W}_{ji} + \epsilon_{ji}^{\star}} m_{ji}}{e^{-\boldsymbol{\alpha}^T \boldsymbol{W}_{ji} + \epsilon_{ji}^{\star}} m_{ji} + m_{ii}} \\
&= \frac{e^{-\ln\left(\frac{m_{ji}}{m_{ii}}\right) + \ln\left(C(\hat{\tilde{\boldsymbol{\tau}}})\right)} m_{ji}}{e^{-\ln\left(\frac{m_{ji}}{m_{ii}}\right) + \ln\left(C(\hat{\tilde{\boldsymbol{\tau}}})\right)} m_{ji} + m_{ii}} \\
&= \frac{\frac{m_{ii}}{m_{ji}} C(\hat{\tilde{\boldsymbol{\tau}}}) m_{ji}}{\frac{m_{ii}}{m_{ji}} C(\hat{\tilde{\boldsymbol{\tau}}}) m_{ji} + m_{ii}} \\
&= \frac{m_{ii} C(\hat{\tilde{\boldsymbol{\tau}}})}{m_{ii} C(\hat{\tilde{\boldsymbol{\tau}}}) + m_{ii}} \\
&= \frac{C(\hat{\tilde{\boldsymbol{\tau}}})}{C(\hat{\tilde{\boldsymbol{\tau}}}) + 1}
\end{aligned}$$
and that
$$\chi_{ji}(\bar{\epsilon}_{ji}) = 1$$
Now the expected utility for government i in stage 1 is
$$\begin{aligned}
\mathrm{E}[L(\hat{\tilde{\boldsymbol{\tau}}}, \boldsymbol{m})] &= \mathrm{E}\left[G_i(\hat{\tilde{\boldsymbol{\tau}}}) - \sum_{j \neq i} \eta_{ij} \hat{G}_j(\hat{\tilde{\boldsymbol{\tau}}}; b_j) - \sum_{j \neq i} \lambda_{ji}^{\chi}(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) \left(Y_{ji}(\hat{\tilde{\boldsymbol{\tau}}}; \boldsymbol{b}) + \hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-1}\right)\right] \\
&= G_i(\hat{\tilde{\boldsymbol{\tau}}}) - \sum_{j \neq i} \int_{\eta_{ij}} \eta_{ij} \hat{G}_j(\hat{\tilde{\boldsymbol{\tau}}}; b_j) f_{\eta}(\eta_{ij}) d\eta_{ij} - \sum_{j \neq i} \int_{\epsilon_{ji}} \lambda_{ji}^{\chi}(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) \left(Y_{ji}(\hat{\tilde{\boldsymbol{\tau}}}; \boldsymbol{b}) + \hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-1}\right) f_{\epsilon}(\epsilon_{ji}) \\
&= G_i(\hat{\tilde{\boldsymbol{\tau}}}) - \sum_{j \neq i} \int_{\epsilon_{ji}} \lambda_{ji}^{\chi}(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) \left(Y_{ji}(\hat{\tilde{\boldsymbol{\tau}}}; \boldsymbol{b}) + \hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-1}\right) f_{\epsilon}(\epsilon_{ji}) d\epsilon_{ji} \\
&= G_i(\hat{\tilde{\boldsymbol{\tau}}}) - \sum_{j \neq i} \int_{\epsilon_{ji}^{\star}}^{\bar{\epsilon}_{ji}} \lambda_{ji}^{\chi}(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) \left(Y_{ji}(\hat{\tilde{\boldsymbol{\tau}}}; \boldsymbol{b}) + \hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-1}\right) f_{\epsilon}(\epsilon_{ji}) d\epsilon_{ji}
\end{aligned}$$
because $\lambda_{ji}^{\chi}(\epsilon_{ji}) = 0$ whenever $\epsilon_{ji} \leq \epsilon_{ji}^{\star}$.

Note that I'm assuming multipliers don't depend on $\boldsymbol{m}$ *for epsilon of interest... can I show this?*

A few facts:
$$Y_{ji}(\hat{\tilde{\boldsymbol{\tau}}}^{\star}; \boldsymbol{b}) + \hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-1} = 0$$

THIS GIVES US THE VALUE OF $C(\hat{\boldsymbol{\tau}}^\star)$. And pins down lower bound on induced chi distribution as a function of $\boldsymbol{m}$ and parameters. So if we do change of variables and integrate over distribution of χ maybe this gives us a closed form for the integral?

Also,

$$\begin{aligned}
\frac{\partial \hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-1}}{\partial m_{ii}} &= -\hat{c}\chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji})^{-2} \frac{\partial \chi_{ji}(\boldsymbol{m}; \boldsymbol{\alpha}, \epsilon_{ji}}{\partial m_{ii}} \\
&= \hat{c} \left(\frac{\rho_{ji} m_{ji} + m_{ii}}{\rho_{ji} m_{ji}} \right)^2 \frac{\rho_{ji} m_{ji}}{(\rho_{ji} m_{ji} + m_{ii})^2} \\
&= \hat{c} \left(\rho_{ji} m_{ji} \right)^{-1}
\end{aligned}$$

Then,

$$\frac{\partial \mathrm{E}[L(\hat{\boldsymbol{\tau}}^\star, \boldsymbol{m})]}{\partial m_{ii}} = \frac{\hat{c}}{m_{ji}} \sum_{j \neq i} \int_{\epsilon^\star_{ji}(\hat{\boldsymbol{\tau}}^\star, \boldsymbol{m})}^{\bar{\epsilon}} \lambda^\chi_{ji}(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) \rho_{ji}(\boldsymbol{\alpha}, \epsilon_{ji})^{-1} f(\epsilon_{ji}) d\epsilon_{ji}$$

and the integral can be calculated by simulating over the government's choice problem. *Where to fix $\boldsymbol{m}$ values here? Does this choice matter?*

Change of variables to integrate over values of χ?

If we know the distribution of χ then we don't even have to calculate $\epsilon^\star$... just simulate distribution of χ from epsilons. Then draw from this induced distribution and solve the problem many times.

Still not sure where to fix $\boldsymbol{m}$ values... bounds of integration depend indirectly on these through effect on $\boldsymbol{\tau}^\star$

If we drop allocation stage altogether does this help? Just assume all wars are bilateral and all or nothing (war stage is Schlieffen's dream). Then ms are data (strengths) and we only have to iterate on stages 2 and 3. I think try this for first cut. We could easily add gamma back. This setup would also potentially let us estimate strengths.

Alternative model that preserves some of what I have now assumes that all troops return home to fight defensive wars... interpretation is a little more strained though I think.