1 Proof of Theorem 1: General Case Size

Proof. The setup of the problem and the structure of the proof for the general class size case mimics the roommate case illustrated in Theorem 1. We continue to assume a homogeneous peer effect and consider the limiting case where

1. We observe students for at most two time periods.
2. Within each class there is at most one student that is observed for two periods. All other students are observed for only one time period.

Remark 1: Clearly if the estimator is consistent for $T = 2$, it is also consistent for $T > 2$. The second simplification is equivalent to allowing all of the individual effects in a class but one to vary over time. For example, suppose class size was fixed at $M + 1$ and there were $(M + 1)\mathcal{N}$ students observed for two periods, implying that $(M + 1)\mathcal{N}$ individual effects would be estimated. We could, however, allow the individual effects to vary over time for all students but one in each group, making sure to choose these students in such a way that they are matched with someone in both periods whose individual effect does not vary over time.[1] $(2M + 1)\mathcal{N}$ individual effects would then be estimated. Having M individuals whose effect varies over time is equivalent to estimating $2M$ individual effects—it is the same as having two sets of M individuals who are each observed once. If the estimator is consistent in this case, then it is also consistent under the restricted case when all of the individual effects are time invariant (fixed effects).

[1] To see how these assignments work, consider a two period model where the groups in period 1 are $\{A, B, C\}$ and $\{D, E, F\}$ and the groups in period 2 are $\{A, B, F\}$ and $\{D, E, C\}$. We could let the individual effects for $\{B, C, E, F\}$ vary over time. Each group in each time period will have one student observed twice and one student observed once. The number of individual effects would then increase from six to ten. More generally, with a common class size of $M + 1$, the most severe overlap that still allows variation in the peer group is to have M individuals in each class remain together in both periods. In this case, we could allow all individual effects to vary over time except for one of the individual effects of the M individuals in each class that stay together in both periods. Things become more complicated when class size is not constant, but allowing all individual effects to vary over time except for a set of individuals who never share a class will grow linearly in $\mathcal{N}$. Hence, while the asymptotic variance would be affected, identification, consistency, and asymptotic normality are unaffected.

Consider the set of students that are observed for two time periods. Each of these students has M_{1n} peers in period one and M_{2n} peers in period two. Denote a student block as one student observed for two periods plus his $M_{1n} + M_{2n}$ peers. There are then $\mathcal{N}$ blocks of students, one block for each student observed twice. Denote the first student in each block as the student who is observed twice, where α_{1n} is the individual effect. For ease of exposition we will also write α_{1n} as α_{11n} or α_{12n}. The time subscripts are irrelevant here since time does not indicate a different individual. The individual effect for the ith classmate in block n at time period t is α_{itn}, where $i \geq 2$. For these individuals the time subscript is relevant for identifying each individual.

The optimization problem is then:

$$\min_{\alpha,\gamma} \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} \left[\left(y_{11n} - \alpha_{1n} - \frac{\gamma}{M_{1n}} \sum_{j=2}^{M_{1n}+1} \alpha_{j1n} \right)^2 + \left(y_{12n} - \alpha_{1n} - \frac{\gamma}{M_{2n}} \sum_{j=2}^{M_{2n}+1} \alpha_{j2n} \right)^2 \right.$$
$$\left. + \sum_{i=2}^{M_{1n}+1} \left(y_{i1n} - \alpha_{i1n} - \frac{\gamma}{M_{1n}} \sum_{j \neq i}^{M_{1n}+1} \alpha_{j1n} \right)^2 + \sum_{i=2}^{M_{2n}+1} \left(y_{i2n} - \alpha_{i2n} - \frac{\gamma}{M_{2n}} \sum_{j \neq i}^{M_{2n}+1} \alpha_{j2n} \right)^2 \right] \quad (1)$$

Within each block there are four terms, two residuals for the student observed twice, and peer residuals in time period one and two.

Again, conditional on γ, the estimates of individual effects in one block will not affect the estimates of the individual effects in another block. Hence, we are able to focus on individual blocks in isolation from one another when concentrating out the α's as a function of γ.

Our proof in the general class size case then consists of the following five lemmas, each of which is proven later in this appendix.

Lemma 1.G

The vector of unobserved student abilities, α, can be concentrated out of the least squares problem and written strictly as a function of γ and y.

Due to the complexity of these expressions we only provide them in the following proof.

We then show the form of the minimization problem when the α's are concentrated out.

Lemma 2.G

Concentrating the α's out of the original least squares problem results in an optimization problem over γ that takes the following form:

$$\min_{\gamma} \sum_{n=1}^{\mathcal{N}} \frac{\left(\sum_{t=1}^{2} \sum_{j=1}^{M_{tn}+1} W_{jtn} y_{jtn}\right)^2}{\sum_{t=1}^{2} \sum_{j=1}^{M_{tn}+1} W_{jtn}^2}$$

where

$$\begin{aligned} W_{11n} &= (\gamma - M_{2n})(M_{1n} + \gamma(M_{1n} - 1)) \\ W_{12n} &= -(\gamma - M_{1n})(M_{2n} + \gamma(M_{2n} - 1)) \\ W_{j1n} &= -\gamma(\gamma - M_{2n}) \ \forall \ j > 1 \\ W_{j2n} &= \gamma(\gamma - M_{1n}) \ \forall \ j > 1 \end{aligned}$$

Our nonlinear least squares problem has only one parameter, γ. We are now in a position to investigate the properties of our estimator of γ_o. For ease of notation, define $q(w, \gamma)$ as:

$$q(w, \gamma) = \frac{\left(\sum_{t=1}^{2} \sum_{j=1}^{M_t+1} W_{jt} y_{jt}\right)^2}{\sum_{t=1}^{2} \sum_{j=1}^{M_t+1} W_{jt}^2}$$

where $w \equiv (y, M)$. We let $\mathcal{W}$ denote the subset of $\mathbb{R}^{2+2\overline{M}} \times \mathcal{J}^2$ representing the possible values of w, where $\mathcal{J}$ is the number of possible class sizes, $\overline{M} - \underline{M} + 1$.

Our key result is then Lemma 3.G, which establishes identification.

Lemma 3.G

$$E[q(w,\gamma_o)] < E[q(w,\gamma)], \quad \forall\ \gamma \in \Gamma, \ \ \gamma \neq \gamma_o$$

Theorem 12.2 of Wooldridge (2002) establishes that sufficient conditions for consistency are identification and uniform convergence. Having already established identification, Lemma 4 shows uniform convergence.

Lemma 4.G

$$\max_{\gamma\in\Gamma} \left| \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} q(w_n,\gamma) - E[q(w,\gamma)] \right| \xrightarrow{p} 0$$

Consistency then follows from Theorem 12.2 of Wooldridge: $\gamma \xrightarrow{p} \gamma_o$.

Finally, we establish asymptotic normality of γ. Denote $s(w,\gamma_o)$ and $H(w,\gamma_o)$ as the first and second derivative of $q(w,\gamma)$ evaluated at γ_o. Then, Lemma 5 completes the proof.

Lemma 5.G

$$\sqrt{\mathcal{N}}(\gamma - \gamma_o) \xrightarrow{d} N(0, A_o^{-1} B_o A_o^{-1})$$

where

$$A_o \equiv E[H(w,\gamma_o)]$$

and

$$B_o \equiv E[s(w,\gamma_o)^2] = Var[s(w,\gamma_o)]$$

QED.

Proof of Lemma 1.G

Our objective is to show that the system of equations obtained by differentiating Equation (**??**) with respect to α can be expressed as a series of equations in terms of γ, y, and M. Again, conditional on γ, the estimates of individual effects in one block will not affect the estimates of the individual effects in another block. Thus, we can work with the system of first-order conditions within one block and then generalize the results to the full system of equations.

The first-order condition for α_{1n} is given by:

$$\begin{aligned} 0 \;=\; & -2\Big(y_{11n} - \alpha_{1n} - \frac{\gamma}{M_{1n}} \sum_{j=2}^{M_{1n}+1} \alpha_{j1n}\Big) - 2\Big(y_{12n} - \alpha_{1n} - \frac{\gamma}{M_{2n}} \sum_{j=2}^{M_{2n}+1} \alpha_{j2n}\Big) \\ & - \frac{2\gamma}{M_{1n}} \sum_{i=2}^{M_{1n}+1} \Big(y_{i1n} - \alpha_{i1n} - \frac{\gamma}{M_{1n}} \sum_{j\neq i}^{M_{1n}+1} \alpha_{j1n}\Big) - \frac{2\gamma}{M_{2n}} \sum_{i=2}^{M_{2n}+1} \Big(y_{i2n} - \alpha_{i2n} - \frac{\gamma}{M_{2n}} \sum_{j\neq i}^{M_{2n}+1} \alpha_{j2n}\Big) \end{aligned}$$

while the first-order condition for α_{i1n} (applicable to all block n students observed once in time period 1) is given by:

$$\begin{aligned} 0 \;=\; & -\frac{2\gamma}{M_{1n}}\Big(y_{1tn} - \alpha_{1n} - \frac{\gamma}{M_{1n}} \sum_{j=2}^{M_{1n}+1} \alpha_{j1n}\Big) - \frac{2\gamma}{M_{1n}} \sum_{j=2\,,\,j\neq i}^{M_{1n}+1} \Big(y_{j1n} - \alpha_{j1n} - \frac{\gamma}{M_{1n}} \sum_{k\neq j}^{M_{1n}+1} \alpha_{k1n}\Big) \\ & -2\Big(y_{i1n} - \alpha_{i1n} - \frac{\gamma}{M_{1n}} \sum_{j\neq i}^{M_{1n}+1} \alpha_{j1n}\Big) \end{aligned}$$

The first order condition for α_{i2n} is identical to above formulation except that all the time subscripts are changed from 1 to 2. Within each block n, we are left with a system of $(1 + M_{1n} + M_{2n})$ equations and $(1 + M_{1n} + M_{2n})$ unknown abilities.

We can re-arrange the above first-order conditions such that all the parameters to be estimated (α's and γ) are on the left and all the observed grades (y) are on the right. Doing this for the first-order conditions derived for α_{1n} and α_{i1n} yields the following two equations

$$\Big(2+\frac{\gamma^2(M_{1n}+M_{2n})}{M_{1n}M_{2n}}\Big)\alpha_{1n}+\sum_{t=1}^{2}\left(\Big(\frac{2\gamma}{M_{tn}}+\frac{(M_{tn}-1)\gamma^2}{M_{tn}^2}\Big)\sum_{j=2}^{M_{tn}+1}\alpha_{jtn}\right) = y_{11n}+y_{12n}+\sum_{t=1}^{2}\left(\frac{\gamma}{M_{tn}}\sum_{j=2}^{M_{tn}+1}y_{jtn}\right)$$

and

$$\Big(1+\frac{\gamma^2}{M_{1n}}\Big)\alpha_{i1n}+\Big(\frac{2\gamma}{M_{1n}}+\frac{(M_{1n}-1)\gamma^2}{M_{1n}^2}\Big)\Big(\alpha_{1n}+\sum_{j=2,\,j\neq i}^{M_{1n}+1}\alpha_{j1n}\Big) = y_{i1n}+\frac{\gamma}{M_{1n}}\Big(y_{11n}+\sum_{j=2,\,j\neq i}^{M_{1n}+1}y_{j1n}\Big)$$

Again, the first-order condition for α_{i2n} can be written in a form identical to the above equation where all the time subscripts are changed from 1 to 2.

We can write the above system of equations in matrix form such that $\mathbf{X}_n \alpha_n = \mathbf{Y}_n$, where α_n is simply a $((1 + M_{1n} + M_{2n})$ x $1)$ vector of the individual student abilities in block n. Recall that because the student blocks are independent conditional on γ, we can solve for α_n separately from α_s for $s \neq n$. The form of $\mathbf{X}_n$, α_n, and $\mathbf{Y}_n$ are given by the following

$$\mathbf{X}_{n_{((1+M_{1n}+M_{2n})\,\mathrm{x}\,(1+M_{1n}+M_{2n}))}} = \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix}$$

$$\alpha_{n_{((1+M_{1n}+M_{2n})\,\mathrm{x}\,1)}} = \left[\alpha_{1n},\ \alpha_{21n},\ \ldots,\ \alpha_{(M_{1n}+1)1n},\ \alpha_{22n},\ \ldots,\ \alpha_{(M_{2n}+1)2n}\right]'$$

$$\mathbf{Y}_{n_{((1+M_{1n}+M_{2n})\,\mathrm{x}\,1)}} = \begin{bmatrix} y_{11n} + y_{12n} + \sum_{t=1}^{2}\left(\frac{\gamma}{M_{tn}}\sum_{j=2}^{M_{tn}+1} y_{jtn}\right) \\ y_{21n} + \frac{\gamma}{M_{1n}} y_{11n} + \frac{\gamma}{M_{1n}}\sum_{j=3}^{M_{1n}+1} y_{j1n} \\ \vdots \\ y_{(M_{1n}+1)1n} + \frac{\gamma}{M_{1n}} y_{11n} + \frac{\gamma}{M_{1n}}\sum_{j=2}^{M_{1n}} y_{j1n} \\ y_{22n} + \frac{\gamma}{M_{2n}} y_{12n} + \frac{\gamma}{M_{2n}}\sum_{j=3}^{M_{2n}+1} y_{j2n} \\ \vdots \\ y_{(M_{2n}+1)2n} + \frac{\gamma}{M_{2n}} y_{12n} + \frac{\gamma}{M_{2n}}\sum_{j=2}^{M_{2n}} y_{j2n} \end{bmatrix}$$

where the components of $\mathbf{X}_n$ are defined below:

$$A_n = 2 + \frac{\gamma^2(M_{1n} + M_{2n})}{M_{1n}M_{2n}}$$

$$B_n = \left[\underbrace{\frac{2\gamma}{M_{1n}} + \frac{(M_{1n}-1)\gamma^2}{M_{1n}^2},\ \ldots\ ,\frac{2\gamma}{M_{1n}} + \frac{(M_{1n}-1)\gamma^2}{M_{1n}^2}}_{M_{1n}\text{ terms}},\ \underbrace{\frac{2\gamma}{M_{2n}} + \frac{(M_{2n}-1)\gamma^2}{M_{2n}^2},\ \ldots\ \frac{2\gamma}{M_{2n}} + \frac{(M_{2n}-1)\gamma^2}{M_{2n}^2}}_{M_{2n}\text{ terms}}\right]$$

$$C_n = \left[\underbrace{\frac{2\gamma}{M_{1n}} + \frac{(M_{1n}-1)\gamma^2}{M_{1n}^2}, \ \dots \ , \frac{2\gamma}{M_{1n}} + \frac{(M_{1n}-1)\gamma^2}{M_{1n}^2}}_{M_{1n} \text{ terms}}, \ \underbrace{\frac{2\gamma}{M_{2n}} + \frac{(M_{2n}-1)\gamma^2}{M_{2n}^2}, \ \dots \ \frac{2\gamma}{M_{2n}} + \frac{(M_{2n}-1)\gamma^2}{M_{2n}^2}}_{M_{2n} \text{ terms}} \right]'$$

$$D_n = \begin{bmatrix} 1+\frac{\gamma^2}{M_{1n}} & \frac{2\gamma}{M_{1n}} + \frac{(M_{1n}-1)\gamma^2}{M_{1n}^2} & \dots & 0 & 0 & \dots \\ \frac{2\gamma}{M_{1n}} + \frac{(M_{1n}-1)\gamma^2}{M_{1n}^2} & 1+\frac{\gamma^2}{M_{1n}} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & \dots & 1+\frac{\gamma^2}{M_{2n}} & \frac{2\gamma}{M_{2n}} + \frac{(M_{2n}-1)\gamma^2}{M_{2n}^2} & \dots \\ 0 & 0 & \dots & \frac{2\gamma}{M_{2n}} + \frac{(M_{2n}-1)\gamma^2}{M_{2n}^2} & 1+\frac{\gamma^2}{M_{2n}} & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

where D_n is an $((M_{1n}+M_{2n})$ x $(M_{1n}+M_{2n}))$ symmetric matrix. The form of A_n, B_n, C_n, and D_n is driven by the coefficients on the α's in the re-arranged system of first-order conditions.

The solution to the system of equations for α_n is now given by the following simple expression

$$\alpha_n = \mathbf{X}_n^{-1}\mathbf{Y}_n$$

The difficulty in calculating the solution arises in finding the inverse of $\mathbf{X}_n$. Using the formula derived by Banachiewicz(1937), the inverse of $\mathbf{X}_n$ can be calculated blockwise according to

$$\mathbf{X_n^{-1}} = \begin{bmatrix} (A_n - B_n D_n^{-1} C_n)^{-1} & -(A_n - B_n D_n^{-1} C_n)^{-1} B_n D_n^{-1} \\ -D_n^{-1} C_n (A_n - B_n D_n^{-1} C_n)^{-1} & D_n^{-1} + D_n^{-1} C_n (A_n - B_n D_n^{-1} C_n)^{-1} B_n D_n^{-1} \end{bmatrix} \quad (2)$$

Since $(A_n - B_n D_n^{-1} C_n)^{-1}$ is just a scalar, the only difficult component of this formula is D_n^{-1}. However, notice that D_n is block diagonal where each block is a symmetric M_{tn} x M_{tn} matrix composed of only two components. Thus to get the form of D_n^{-1} we just need to invert one of these M_{tn} x M_{tn} matrices.

At this point it is useful to introduce some further notation in order to keep the matrix algebra for calculating $\mathbf{X}_n^{-1}$ palatable. Define:

$$
\begin{aligned}
a_n &= \frac{2M_{1n}M_{2n} + \gamma^2(M_{1n}+M_{2n})}{M_{1n}M_{2n}} \\
b_{1n} &= \frac{2\gamma M_{1n} + \gamma^2(M_{1n}-1)}{M_{1n}^2} \\
b_{2n} &= \frac{2\gamma M_{2n} + \gamma^2(M_{2n}-1)}{M_{2n}^2} \\
c_{1n} &= \frac{M_{1n}+\gamma^2}{M_{1n}} \\
c_{2n} &= \frac{M_{2n}+\gamma^2}{M_{2n}} \\
d_{1n} &= \frac{(\gamma - M_{1n})^2(\gamma^2 - \gamma(2+\gamma)M_{1n} + (1+\gamma)^2M_{1n}^2)}{M_{1n}^4} \\
d_{2n} &= \frac{(\gamma - M_{2n})^2(\gamma^2 - \gamma(2+\gamma)M_{2n} + (1+\gamma)^2M_{2n}^2)}{M_{2n}^4}
\end{aligned}
$$

Using these terms we can re-write the components of $\mathbf{X}_n$ in the following way

$$A_n = a$$

$$B_n = \left[\underbrace{b_{1n},\ \ldots\ , b_{1n},}_{M_{1n}\text{ terms}}\ \underbrace{b_{2n},\ \ldots\ b_{2n}}_{M_{2n}\text{ terms}}\right]$$

$$C_n = \left[\underbrace{b_{1n},\ \ldots\ , b_{1n},}_{M_{1n}\text{ terms}}\ \underbrace{b_{2n},\ \ldots\ b_{2n}}_{M_{2n}\text{ terms}}\right]'$$

$$D_n = \begin{bmatrix} c_{1n} & b_{1n} & \ldots & 0 & 0 & \ldots \\ b_{1n} & c_{1n} & \ldots & 0 & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & \ldots & c_{2n} & b_{2n} & \ldots \\ 0 & 0 & \ldots & b_{2n} & c_{2n} & \ldots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

Again, the key challenge in finding $\mathbf{X}_n^{-1}$ is finding D_n^{-1}. Since D_n^{-1} is block diagonal, this boils down to finding the inverse of a symmetric (M_{tn} x M_{tn}) matrix that consists of two components, b_{tn} and c_{tn}. Depending on the size of M_{tn} this may in itself be difficult. However, we can recursively apply the same blockwise formula to this M_{tn} x M_{tn} matrix until we finally get to the point where we only have to invert a two-by-two matrix. Following this procedure one can show that D_n^{-1} takes the following simple form

$$D_n^{-1} = \begin{bmatrix} \frac{c_{1n}+b_{1n}(M_{1n}-2)}{d_{1n}} & \frac{-b_{1n}}{d_{1n}} & \cdots & 0 & 0 & \cdots \\ \frac{-b_{1n}}{d_{1n}} & \frac{c_{1n}+b_{1n}(M_{1n}-2)}{d_{1n}} & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & \cdots & \frac{c_{2n}+b_{2n}(M_{2n}-2)}{d_{2n}} & \frac{-b_{2n}}{d_{2n}} & \cdots \\ 0 & 0 & \cdots & \frac{-b_{2n}}{d_{2n}} & \frac{c_{2n}+b_{2n}(M_{2n}-2)}{d_{2n}} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

We now have all the components required to calculate $\mathbf{X}_n^{-1}$ using Equation (**??**). According to Equation (**??**), $\mathbf{X}_n^{-1}(1,1)$ is given by $(A_n - B_n D_n^{-1} C_n)^{-1}$. To calculate this expression, we proceed step by step, starting with the first term in $B_n D_n^{-1}$.

$$\begin{aligned} B_n D_n^{-1}(1,1) &= \frac{b_{1n}(c_{1n} + b_{1n}(M_{1n}-2)) - b_{1n}^2(M_{1n}-1)}{d_{1n}} \\ &= \frac{b_{1n}(c_{1n} - b_{1n})}{d_{1n}} \end{aligned}$$

Given the simple structure of B_n and the symmetric nature of D_n^{-1}, it is obvious that the first M_{1n} terms of $B_n D_n^{-1}$ will be identical to the expression derived above. In addition, the final M_{2n} terms will take the same form as the above expression, however, all the time subscripts will change from 1 to 2. As a result,

$$B_n D_n^{-1} = \left[\frac{b_{1n}(c_{1n} - b_{1n})}{d_{1n}}, \ \ldots, \ \frac{b_{1n}(c_{1n} - b_{1n})}{d_{1n}}, \ \frac{b_{2n}(c_{2n} - b_{2n})}{d_{2n}}, \ \ldots, \ \frac{b_{2n}(c_{2n} - b_{2n})}{d_{2n}} \right]$$

Calculating $B_n D_n^{-1} C_n$ is rather simple, since it is just a scalar.

$$B_n D_n^{-1} C_n = \frac{M_{1n} b_{1n}^2 (c_{1n} - b_{1n})}{d_{1n}} + \frac{M_{2n} b_{2n}^2 (c_{2n} - b_{2n})}{d_{2n}}$$

Finally,

$$\mathbf{X}_n^{-1}(1,1) = a - \Big(\frac{M_{1n}b_{1n}^2(c_{1n}-b_{1n})}{d_{1n}} + \frac{M_{2n}b_{2n}^2(c_{2n}-b_{2n})}{d_{2n}}\Big)$$

Because this terms appears in all of the other components of $\mathbf{X}_n^{-1}$, for expositional ease we define $\tilde{A}_n = \mathbf{X}_n^{-1}(1,1)$.

According to Equation (**??**), $\mathbf{X}_n^{-1}(1,2)$ is given by $-\tilde{A}_n B_n D_n^{-1}$. We calculated the expression for $B_n D_n^{-1}$ in the previous step, thus

$$\mathbf{X}_n^{-1}(1,2) = -\tilde{A}_n \left[\frac{b_{1n}(c_{1n}-b_{1n})}{d_{1n}},\ \ldots,\ \frac{b_{1n}(c_{1n}-b_{1n})}{d_{1n}},\ \frac{b_{2n}(c_{2n}-b_{2n})}{d_{2n}},\ \ldots,\ \frac{b_{2n}(c_{2n}-b_{2n})}{d_{2n}}\right]$$

The expression for $\mathbf{X}_n^{-1}(2,1)$, $-\tilde{A}_n D_n^{-1} C_n$, will be the transpose of the above, since D_n^{-1} is symmetric and $B_n^T = C_n$. Again for expositional ease, define $\tilde{B}_{1n} = -\frac{\tilde{A}_n b_{1n}(c_{1n}-b_{1n})}{d_{1n}}$ and $\tilde{B}_{2n} = -\frac{\tilde{A}_n b_{2n}(c_{2n}-b_{2n})}{d_{2n}}$. Using this definition we can write

$$\mathbf{X}_n^{-1}(1,2) = \left[\tilde{B}_{1n},\ \ldots,\ \tilde{B}_{1n},\ \tilde{B}_{2n},\ \ldots,\ \tilde{B}_{2n}\right]$$

The final component of $\mathbf{X}_n^{-1}$ is also the most complicated. The expression for $\mathbf{X}_n^{-1}(2,2)$ in Equation (**??**) is $D_n^{-1} + \tilde{A}_n D_n^{-1} C_n B_n D_n^{-1}$. Again we proceed in steps. Pre-multiplying $B_n D_n^{-1}$ by C_n will yield an $((M_{1n}+M_{2n})$ x $(M_{1n}+M_{2n}))$ matrix that takes the form

$$C_n B_n D_n^{-1} = \begin{bmatrix} \frac{b_{1n}^2(c_{1n}-b_{1n})}{d_{1n}} & \frac{b_{1n}^2(c_{1n}-b_{1n})}{d_{1n}} & \cdots & \frac{b_{1n}b_{2n}(c_{2n}-b_{2n})}{d_{2n}} & \frac{b_{1n}b_{2n}(c_{2n}-b_{2n})}{d_{2n}} & \cdots \\ \frac{b_{1n}^2(c_{1n}-b_{1n})}{d_{1n}} & \frac{b_{1n}^2(c_{1n}-b_{1n})}{d_{1n}} & \cdots & \frac{b_{1n}b_{2n}(c_{2n}-b_{2n})}{d_{2n}} & \frac{b_{1n}b_{2n}(c_{2n}-b_{2n})}{d_{2n}} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})}{d_{1n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})}{d_{1n}} & \cdots & \frac{b_{2n}^2(c_{2n}-b_{2n})}{d_{2n}} & \frac{b_{2n}^2(c_{2n}-b_{2n})}{d_{2n}} & \cdots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})}{d_{1n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})}{d_{1n}} & \cdots & \frac{b_{2n}^2(c_{2n}-b_{2n})}{d_{2n}} & \frac{b_{2n}^2(c_{2n}-b_{2n})}{d_{2n}} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

Notice that within any quadrant of the matrix all the terms are identical. Finally we need to pre-multiply $C_n B_n D_n^{-1}$ by D_n^{-1}. This yields a symmetric $((M_{1n}+M_{2n})$ x $(M_{1n}+M_{2n}))$ that takes the following form

$$\begin{bmatrix} \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \dots & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \dots \\ \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \dots & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \dots & \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \dots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \dots & \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

The final step is to subtract $\tilde{A}_n D_n^{-1} C_n B_n D_n^{-1}$ from D_n^{-1}. The result is a symmetric $((M_{1n} + M_{2n})$ x $(M_{1n} + M_{2n}))$ matrix that takes the following form

$$D_n^{-1} - \tilde{A}_n D_n^{-1} C_n B_n D_n^{-1} = \begin{bmatrix} \tilde{C}_{1n} & \tilde{D}_{1n} & \dots & \tilde{E}_n & \tilde{E}_n & \dots \\ \tilde{D}_{1n} & \tilde{C}_{1n} & \dots & \tilde{E}_n & \tilde{E}_n & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \tilde{E}_n & \tilde{E}_n & \dots & \tilde{C}_{2n} & \tilde{D}_{2n} & \dots \\ \tilde{E}_n & \tilde{E}_n & \dots & \tilde{D}_{2n} & \tilde{C}_{2n} & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

where

$$\begin{aligned} \tilde{C}_{1n} &= \frac{d_{1n}(c_{1n} + (M_{1n} - 1)b_{1n}) + \tilde{A}_n b_{1n}^2 (c_{1n} - b_{1n})^2}{d_{1n}^2} \\ \tilde{C}_{2n} &= \frac{d_{2n}(c_{2n} + (M_{2n} - 1)b_{2n}) + \tilde{A}_n b_{2n}^2 (c_{2n} - b_{2n})^2}{d_{2n}^2} \\ \tilde{D}_{1n} &= \frac{\tilde{A}_n b_{1n}^2 (c_{1n} - b_{1n})^2 - b_{1n} d_{1n}}{d_{1n}^2} \\ \tilde{D}_{2n} &= \frac{\tilde{A}_n b_{2n}^2 (c_{2n} - b_{2n})^2 - b_{2n} d_{2n}}{d_{2n}^2} \\ \tilde{E}_n &= \frac{\tilde{A}_n b_{1n} b_{2n} (c_{1n} - b_{1n})(c_{2n} - b_{2n})}{d_{1n} d_{2n}} \end{aligned}$$

Recall that the a_n, b_n's, c_n's, and d_n's were defined earlier and are functions solely of γ, M_{1n}, and M_{2n}.

Substituting into Equation (**??**) with the terms just calculated, we get the general form of $\mathbf{X}_n^{-1}$,

$$\mathbf{X}_n^{-1} = \begin{bmatrix} \tilde{A}_n & \tilde{B}_{1n} & \tilde{B}_{1n} & \ldots & \tilde{B}_{2n} & \tilde{B}_{2n} & \ldots \\ \tilde{B}_{1n} & \tilde{C}_{1n} & \tilde{D}_{1n} & \ldots & \tilde{E}_n & \tilde{E}_n & \ldots \\ \tilde{B}_{1n} & \tilde{D}_{1n} & \tilde{C}_{1n} & \ldots & \tilde{E}_n & \tilde{E}_n & \ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \tilde{B}_{2n} & \tilde{E}_n & \tilde{E}_n & \ldots & \tilde{C}_{2n} & \tilde{D}_{2n} & \ldots \\ \tilde{B}_{2n} & \tilde{E}_n & \tilde{E}_n & \ldots & \tilde{D}_{2n} & \tilde{C}_{2n} & \ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

Using $\mathbf{X}_n^{-1}$ and the formula for $\mathbf{Y}_n$ we can solve for the α_n's as functions of γ, y, and M. As an example, the solution for α_{1n} can be obtained by multiplying $\mathbf{Y}_n$ by the first row of $\mathbf{X}_n^{-1}$.

$$\begin{aligned} \alpha_{1n} &= \tilde{A}_n \left(y_{11n} + y_{12n} + \sum_{t=1}^{2} \Big(\frac{\gamma}{M_{tn}} \sum_{j=2}^{M_{tn}+1} y_{jtn} \Big) \right) \\ &\quad + \sum_{t=1}^{2} \left(\tilde{B}_{tn} \sum_{i=2}^{M_{tn}+1} \left(y_{itn} + \frac{\gamma}{M_{tn}} y_{1tn} + \frac{\gamma}{M_{tn}} \sum_{j=2, j\neq i}^{M_{tn}+1} y_{jtn} \right) \right) \end{aligned}$$

We can re-arrange this formula such that we group all the common y terms together. Doing so yields the solution for α_{1n} in terms of $\tilde{A}_n$, $\tilde{B}_{1n}$, and $\tilde{B}_{2n}$

$$\alpha_{1n} = \Big(\tilde{A}_n + \gamma \tilde{B}_{1n}\Big) y_{11n} + \Big(\tilde{A}_n + \gamma \tilde{B}_{2n}\Big) y_{12n} + \sum_{t=1}^{2} \left(\Big(\tilde{A}_n \frac{\gamma}{M_{tn}} + \tilde{B}_{tn} \frac{\gamma(M_{tn}-1) + M_{tn}}{M_{tn}} \Big) \sum_{j=2}^{M_{tn}+1} y_{jtn} \right)$$

Finding the solution for any α in block n other than α_{1n} follows the same basic procedure. Simply multiply $\mathbf{Y}_n$ by the appropriate row from $\mathbf{X}_n^{-1}$. As an example, below is the formula for α_{21n}. To arrive at this formula simply multiply $\mathbf{Y}_n$ by the second row of $\mathbf{X}_n^{-1}$.

$$\begin{aligned} \alpha_{21n} &= \tilde{B}_{1n} \left(y_{11n} + y_{12n} + \sum_{t=1}^{2} \Big(\frac{\gamma}{M_{tn}} \sum_{j=2}^{M_{tn}+1} y_{jtn} \Big) \right) + \tilde{C}_{1n} \left(y_{21n} + \frac{\gamma}{M_{1n}} y_{11n} + \frac{\gamma}{M_{1n}} \sum_{j=3}^{M_{1n}+1} y_{j1n} \right) \\ &+ \tilde{D}_{1n} \left(\sum_{i=3}^{M_{1n}+1} \big(y_{i1n} + \frac{\gamma}{M_{1n}} y_{11n} + \frac{\gamma}{M_{1n}} \sum_{j=2, j\neq i}^{M_{1n}+1} y_{j1n} \big) \right) \\ &+ \tilde{E}_n \left(\sum_{i=2}^{M_{2n}+1} \big(y_{i2n} + \frac{\gamma}{M_{2n}} y_{12n} + \frac{\gamma}{M_{2n}} \sum_{j=2, j\neq i}^{M_{2n}+1} y_{j2n} \big) \right) \end{aligned}$$

Again, we can re-arrange the above, grouping on the y's

$$
\begin{aligned}
\alpha_{21n} &= \Big(\tilde{B}_{1n} + \tilde{C}_{1n}\frac{\gamma}{M_{1n}} + \tilde{D}_{1n}\frac{(M_{1n}-1)\gamma}{M_{1n}}\Big)y_{11n} + \Big(\tilde{B}_{1n} + \tilde{E}_n\gamma\Big)y_{12n} \\
&+ \Big(\tilde{B}_{1n}\frac{\gamma}{M_{1n}} + \tilde{C}_{1n} + \tilde{D}_{1n}\frac{(M_{1n}-1)\gamma}{M_{1n}}\Big)y_{21n} \\
&+ \Big(\tilde{B}_{1n}\frac{\gamma}{M_{1n}} + \tilde{C}_{1n}\frac{\gamma}{M_{1n}} + \tilde{D}_{1n}\frac{M_{1n}+(M_{1n}-2)\gamma}{M_{1n}}\Big)\sum_{j=3}^{M_{1n}+1} y_{j1n} \\
&+ \Big(\tilde{B}_{1n}\frac{\gamma}{M_{2n}} + \tilde{E}_n\frac{M_{2n}+(M_{2n}-1)\gamma}{M_{2n}}\Big)\sum_{j=2}^{M_{2n}+1} y_{j2n}
\end{aligned}
$$

The formula for α_{i1n} for $i > 2$ takes the same form as above, except that (1) y_{21n} becomes y_{i1n} and (2) the first summation on the second line will be over all $j \neq i$. The formula for α_{i2n} for $i > 1$ also takes the same general form, except that all of the subscripts denoting period 1 need to be changed to denote period 2, and vice versa. All the terms in the formulas for α_{1n} and α_{itn} consist solely of γ, y, and M.

QED

Proof of Lemma 2.G

Lemma 1 provides a solution for α strictly as a function of y, γ, and M. We can substitute this solution back into the original optimization problem to derive the result in Lemma 2.G.

Consider minimizing the sum of squared residuals within a particular block n. There are $2 + M_{1n} + M_{2n}$ residuals within each block, two for the student observed twice, and one each for the peers in both time periods. We begin by simplifying the residual for the first observation of the student observed twice, which is given by the expression below

$$e_{11n} = y_{11n} - \alpha_{1n} - \frac{\gamma}{M_{1n}} \sum_{j=2}^{M_{1n}+1} \alpha_{j1n}$$

Substituting for α_{1n} and α_{j1n} in e_{11n} with the results from Lemma 1.G and collecting terms on the y's results in

$$\begin{aligned}
e_{11n} \quad = \quad & y_{11n}\left(1 - \tilde{A}_n - 2\gamma\tilde{B}_{1n} - \frac{\gamma^2}{M_{1n}}\tilde{C}_{1n} - \frac{(M_{1n}-1)\gamma^2}{M_{1n}}\tilde{D}_{1n}\right) \\
& - y_{12n}\left(\tilde{A}_n + \gamma\tilde{B}_{1n} + \gamma\tilde{B}_{2n} + \gamma^2\tilde{E}_n\right) \\
& - \left(\sum_{j=2}^{M_{1n}+1} y_{j1n}\right)\left(\frac{\gamma}{M_{1n}}\tilde{A}_n + \frac{\gamma^2 + M_{1n} + \gamma(M_{1n}-1)}{M_{1n}}\tilde{B}_{1n} + \frac{\gamma(M_{1n} + \gamma(M_{1n}-1))}{M_{1n}^2}\tilde{C}_{1n} + \frac{\gamma(\gamma + M_{1n}(M_{1n} + \gamma M_{1n} - 2\gamma - 1))}{M_{1n}^2}\tilde{D}_{1n}\right) \\
& - \left(\sum_{j=2}^{M_{2n}+1} y_{j2n}\right)\left(\frac{\gamma}{M_{2n}}\tilde{A}_n + \frac{\gamma^2}{M_{2n}}\tilde{B}_{1n} + \frac{M_{2n} + \gamma(M_{2n}-1)}{M_{2n}}\tilde{B}_{2n} + \frac{\gamma(M_{2n} + \gamma(M_{2n}-1))}{M_{2n}}\tilde{E}_n\right)
\end{aligned}$$

The form of e_{12n} will be identical to the above except all of the time scripts on the y's, M's, and inverse components will be swapped. In other words, 1's become 2's, and 2's become 1's. Similarly, substituting for α in e_{21n} and collecting terms yields

$$
\begin{aligned}
e_{21n} \quad = \quad & y_{21n}\left(1 - \tilde{A}_n \frac{\gamma^2}{M_{1n}^2} - 2\tilde{B}_{1n}\frac{M_{1n}\gamma + \gamma^2(M_{1n}-1)}{M_{1n}^2} - \tilde{C}_{1n}\frac{M_{1n}^2 + \gamma^2(M_{1n}-1)}{M_{1n}^2} - \tilde{D}_{1n}\frac{2\gamma M_{1n}(M_{1n}-1) + \gamma^2(M_{1n}-1)(M_{1n}-2)}{M_{1n}^2}\right) \\
& -y_{11n}\left(\tilde{A}_n\frac{\gamma}{M_{1n}} + \tilde{B}_{1n}\frac{M_{1n} + (M_{1n}-1)\gamma + \gamma^2}{M_{1n}} + \tilde{C}_{1n}\frac{\gamma M_{1n} + (M_{1n}-1)\gamma^2}{M_{1n}^2} + \tilde{D}_{1n}\frac{(M_{1n}-1)\gamma(M_{1n} + (M_{1n}-1)\gamma)}{M_{1n}^2}\right) \\
& -y_{12n}\left(\tilde{A}_n\frac{\gamma}{M_{1n}} + \tilde{B}_{1n}\frac{M_{1n} + (M_{1n}-1)\gamma}{M_{1n}} + \tilde{B}_{2n}\frac{\gamma^2}{M_{1n}} + \tilde{E}\frac{\gamma M_{1n} + (M_{1n}-1)\gamma^2}{M_{1n}}\right) \\
& -\left(\sum_{j=3}^{M_{1n}+1} y_{j1n}\right)\left(\tilde{A}_n\frac{\gamma^2}{M_{1n}^2} + 2\tilde{B}_{1n}\frac{\gamma M_{1n} + \gamma^2(M_{1n}-1)}{M_{1n}^2} + \tilde{C}_{1n}\frac{2\gamma M_{1n} + \gamma^2(M_{1n}-2)}{M_{1n}^2} + \tilde{D}_{1n}\frac{(M_{1n} + \gamma(M_{1n}-2))^2 + (M_{1n}-1)\gamma^2}{M_{1n}^2}\right) \\
& -\left(\sum_{j=2}^{M_{2n}+1} y_{j2n}\right)\left(\tilde{A}_n\frac{\gamma^2}{M_{1n}M_{2n}} + \tilde{B}_{1n}\frac{\gamma M_{1n} + \gamma^2(M_{1n}-1)}{M_{1n}M_{2n}} + \tilde{B}_{2n}\frac{\gamma M_{2n} + \gamma^2(M_{2n}-1)}{M_{1n}M_{2n}} + \tilde{E}_n\frac{M_{1n} + \gamma(M_{1n}-1)}{M_{1n}}\frac{M_{2n} + \gamma(M_{2n}-1)}{M_{2n}}\right)
\end{aligned}
$$

The residual e_{i1n} for $i > 2$ will look identical to the above except the leading y term will be y_{i1n} rather than y_{21n}, and the summation term in the fourth line will be over all $j \neq i$. The M_{2n} residuals for the individuals observed once in the second period will look identical to the above except that all of the time subscripts are swapped - 1's become 2's and 2's become 1's - for all the y's, M's, and inverse components.

In order to write the least squares problem strictly as a function of γ, we can simply substitute the above expressions directly into the least squares problem. However, before doing so it is helpful to simplify the expressions for the residuals by substituting in for the inverse components, $\tilde{A}_n$, $\tilde{B}_{1n}$, $\tilde{B}_{2n}$, $\tilde{C}_{1n}$, $\tilde{C}_{2n}$, $\tilde{D}_{1n}$, $\tilde{D}_{2n}$, and $\tilde{E}_n$. At this point, the algebra required to show how these equations simplify is extremely cumbersome. Web Appendix 2 shows the full derivation for the case where $M_{1n} = M_{2n}$. Here we jump directly to the simplified versions of the individual residuals:[2]

$$e_{11n} = \left(\frac{(\gamma(M_{1n}-1)+M_{1n})(\gamma-M_{2n})}{(\gamma-M_2)^2((M_1+\gamma(M_1-1))^2+\gamma^2M_1)+(\gamma-M_1)^2((M_2+\gamma(M_2-1))^2+\gamma^2M_2)} \right)$$
$$\times \left((\gamma(M_{1n}-1)+M_{1n})(\gamma-M_{2n})y_{11n} - (\gamma-M_{1n})(\gamma(M_{2n}-1)+M_{2n})y_{12n} - \gamma(\gamma-M_{2n}) \sum_{j=2}^{M_{1n}+1} y_{j1n} + \gamma(\gamma-M_{1n}) \sum_{j=2}^{M_{2n}+1} y_{j2n} \right)$$

$$e_{12n} = \left(\frac{(\gamma(M_{2n}-1)+M_{2n})(\gamma-M_{1n})}{(\gamma-M_2)^2((M_1+\gamma(M_1-1))^2+\gamma^2M_1)+(\gamma-M_1)^2((M_2+\gamma(M_2-1))^2+\gamma^2M_2)} \right)$$
$$\times \left(-(\gamma(M_{1n}-1)+M_{1n})(\gamma-M_{2n})y_{11n} + (\gamma-M_{1n})(\gamma(M_{2n}-1)+M_{2n})y_{12n} + \gamma(\gamma-M_{2n}) \sum_{j=2}^{M_{1n}+1} y_{j1n} - \gamma(\gamma-M_{1n}) \sum_{j=2}^{M_{2n}+1} y_{j2n} \right)$$

$$e_{21n} = \left(\frac{\gamma(\gamma-M_{2n})}{(\gamma-M_2)^2((M_1+\gamma(M_1-1))^2+\gamma^2M_1)+(\gamma-M_1)^2((M_2+\gamma(M_2-1))^2+\gamma^2M_2)} \right)$$
$$\times \left(-(\gamma(M_{1n}-1)+M_{1n})(\gamma-M_{2n})y_{11n} + (\gamma-M_{1n})(\gamma(M_{2n}-1)+M_{2n})y_{12n} + \gamma(\gamma-M_{2n}) \sum_{j=2}^{M_{1n}+1} y_{j1n} - \gamma(\gamma-M_{1n}) \sum_{j=2}^{M_{2n}+1} y_{j2n} \right)$$

$$e_{22n} = \left(\frac{\gamma(\gamma-M_{1n})}{(\gamma-M_2)^2((M_1+\gamma(M_1-1))^2+\gamma^2M_1)+(\gamma-M_1)^2((M_2+\gamma(M_2-1))^2+\gamma^2M_2)} \right)$$
$$\times \left((\gamma(M_{1n}-1)+M_{1n})(\gamma-M_{2n})y_{11n} - (\gamma-M_{1n})(\gamma(M_{2n}-1)+M_{2n})y_{12n} - \gamma(\gamma-M_{2n}) \sum_{j=2}^{M_{1n}+1} y_{j1n} + \gamma(\gamma-M_{1n}) \sum_{j=2}^{M_{2n}+1} y_{j2n} \right)$$

The simplified versions of e_{i1n} and e_{i2n} for $i > 2$ exactly match the above expressions for e_{21n} and e_{22n} respectively.

[2]The algebra required to simplify these expressions is available upon request.

Close inspection of the residual equations indicates that they are all closely related. In fact, the residuals can be derived from one another according to:

$$\begin{aligned}
e_{i1n} &= -e_{11n}\frac{\gamma}{M_{1n}+\gamma(M_{1n}-1)} \\
e_{i2n} &= -e_{i1n}\frac{\gamma - M_{1n}}{\gamma - M_{2n}} \\
e_{12n} &= -e_{11n}\left(\frac{\gamma - M_{1n}}{\gamma - M_{2n}}\right)\left(\frac{M_{2n}+\gamma(M_{2n}-1)}{M_{1n}+\gamma(M_{1n}-1)}\right) \qquad (3)
\end{aligned}$$

Using these relationships, the sum of the squared residuals in block n, $e_{11n}^2 + e_{12n}^2 + \sum_{j=2}^{M_{1n}+1} e_{j1n}^2 + \sum_{j=2}^{M_{2n}+1} e_{j2n}^2$, can be written:

$$\begin{aligned}
&= e_{11n}^2 + \frac{(\gamma - M_{1n})^2}{(\gamma - M_{2n})^2}\frac{(M_{2n}+\gamma(M_{2n}-1))^2}{(M_{1n}+\gamma(M_{1n}-1))^2}e_{11n}^2 + \frac{\gamma^2 M_{1n}}{(M_{1n}+\gamma(M_{1n}-1))^2}e_{11n}^2 + \frac{\gamma^2 M_{2n}(\gamma - M_{1n})^2}{(\gamma - M_{2n})^2(M_{1n}+\gamma(M_{1n}-1))^2}e_{11n}^2 \\
&= e_{11n}^2\left[\frac{(\gamma - M_{2n})^2((M_{1n}+\gamma(M_{1n}-1))^2+\gamma^2 M_{1n}) + (\gamma - M_{1n})^2((M_{2n}+\gamma(M_{2n}-1))^2+\gamma^2 M_{2n})}{(\gamma - M_{2n})^2(M_{1n}+\gamma(M_{1n}-1))^2}\right]
\end{aligned}$$

Finally, substituting for e_{11n}, we arrive at the least squares problem.

QED

Proof of Lemma 3.G

Recall that $q(w,\gamma)$ is given by

$$q(w,\gamma) = \frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}y_{jt}\right)^2}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2}$$

where the W's are defined in the outline of Lemma 2.G. Substituting in for y_{jt} with the data generating process yields:

$$q(w,\gamma) = \frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}\left[\alpha_{jto} + \frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}\alpha_{kto} + \epsilon_{jt}\right]\right)^2}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2}$$

Collecting the α_{jto} terms yields:

$$q(w,\gamma) = \frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}\left(W_{jt} + \frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1} W_{kt}\right)\alpha_{jto} + W_{jt}\epsilon_{jt}\right)^2}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2} \tag{4}$$

Note that the coefficient on α_{1o} is given by the weight at $t=1$ plus the weight at $t=2$:

$$\sum_{t=1}^{2}\left(W_{1t} + \frac{\gamma_o}{M_t}\sum_{k\neq 1}^{M_t+1} W_{kt}\right) = (\gamma - M_2)(M_1 + \gamma(M_1-1) - \gamma_o\gamma) - (\gamma - M_1)(M_2 + \gamma(M_2-1) - \gamma_o\gamma)$$

Because of the symmetry, after multiplying out, any terms involving M_1M_2 will drop out as will any terms where neither M_1 and M_2 enter. The expression then reduces to:

$$\begin{aligned}\sum_{t=1}^{2}\left(W_{1t} + \frac{\gamma_o}{M_t}\sum_{k\neq 1}^{M_t+1} W_{kt}\right) &= (M_1-M_2)\gamma + (M_1-M_2)\gamma^2 - (M_1-M_2)\gamma_o\gamma + (M_2-M_1)\gamma \\ &= (M_1-M_2)(\gamma^2 - \gamma_o\gamma) \\ &= (M_1-M_2)(\gamma-\gamma_o)\gamma\end{aligned}$$

Now consider the coefficient on α_{j1o} for $j > 1$ which can be split into three components: 1) the own weight, the weight from observation 1, and the weight from classmates besides 1:

$$W_{j1} + \frac{\gamma_o}{M_1}\sum_{k\neq j}^{M_1+1} W_{k1} = (\gamma - M_2)\left(-\gamma + \left[1 + \gamma - \frac{\gamma}{M_1}\right]\gamma_o - \left[\frac{\gamma(M_1-1)}{M_1}\right]\gamma_o\right)$$

which reduces to:

$$W_{j1} + \frac{\gamma_o}{M_1}\sum_{k\neq j}^{M_1+1} W_{k1} = (\gamma - M_2)(\gamma_o - \gamma)$$

We then know that the coefficient on W_{j2} for $j > 1$ is given by:

$$W_{j2} + \frac{\gamma_o}{M_2} \sum_{k \neq j}^{M_2+1} W_{k2} = (\gamma - M_1)(\gamma - \gamma_o)$$

Substituting for these expressions in (**??**) yields:

$$\begin{aligned} q(w,\gamma) &= \Bigg[\Bigg((M_1 - M_2)(\gamma - \gamma_o)\gamma\alpha_{1o} + (\gamma - M_2)(\gamma_o - \gamma)\left(\sum_{j=2}^{M_1+1} \alpha_{j1o}\right) \\ &\quad + (\gamma - M_1)(\gamma - \gamma_o)\left(\sum_{j=2}^{M_2+1} \alpha_{j2o}\right) + \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}\epsilon_{jt}\Bigg)^2\Bigg] \\ &\quad \Bigg/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right] \end{aligned}$$

We next take expectations conditional on M_1, M_2:

$$\begin{aligned} E[q(w,\gamma)|M_1,M_2] &= E\Bigg\{\Bigg(\Bigg[\Bigg((M_1 - M_2)(\gamma - \gamma_o)\gamma\alpha_{1o} + (\gamma - M_2)(\gamma_o - \gamma)\left(\sum_{j=2}^{M_1+1} \alpha_{j1o}\right) \\ &\quad + (\gamma - M_1)(\gamma - \gamma_o)\left(\sum_{j=2}^{M_2+1} \alpha_{j2o}\right) + \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}\epsilon_{jt}\Bigg)^2\Bigg] \\ &\quad \Bigg/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]\Bigg) |M_1, M_2\Bigg\} \end{aligned}$$

Expanding the square and noting that 1) $E(\alpha_{jto}\epsilon_{kt'}) = 0$ for all j, k, t, t' by assumption 3 and 2) $E(\epsilon_{jt}\epsilon_{kt'}) = 0$ for all $j \neq k$ or $t \neq t'$ by assumption 2 yields:

$$\begin{aligned} E[q(w,\gamma)|M_1,M_2] &= E\Bigg\{\Bigg(\Bigg[\Bigg((M_1 - M_2)(\gamma - \gamma_o)\gamma\alpha_{1o} + (\gamma - M_2)(\gamma_o - \gamma)\left(\sum_{j=2}^{M_1+1} \alpha_{j1o}\right) \\ &\quad + (\gamma - M_1)(\gamma - \gamma_o)\left(\sum_{j=2}^{M_2+1} \alpha_{j2o}\right)\Bigg)^2 + \left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\epsilon_{jt}^2\right)\Bigg] \\ &\quad \Bigg/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]\Bigg) |M_1, M_2\Bigg\} \end{aligned}$$

which can be rewritten as:

$$
\begin{aligned}
E[q(w,\gamma)|M_1,M_2] &= (\gamma-\gamma_o)^2 E\left[\left((M_1-M_2)\gamma\alpha_{1o}-(\gamma-M_2)\left(\sum_{j=2}^{M_1+1}\alpha_{j1o}\right)\right.\right.\\
&\quad \left.\left.+\ (\gamma-M_1)\left(\sum_{j=2}^{M_2+1}\alpha_{j2o}\right)\right)^2 |M_1,M_2\right] \Big/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]\\
&\quad +\left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2 E\left(\epsilon_{jt}^2|M_1,M_2\right)\right] \Big/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]
\end{aligned}
$$

Note that assumption 4 implies that the conditioning in the expectation over the squared errors is not needed. Further, $E(\epsilon_{jt}^2) = E(\epsilon_{kt}^2)$ for all j,k by assumption 7. We can then express the expectation over the squared errors solely as a function of the first observation's squared error:

$$
\begin{aligned}
E[q(w,\gamma)|M_1,M_2] &= (\gamma-\gamma_o)^2 E\left[\left((M_1-M_2)\gamma\alpha_{1o}-(\gamma-M_2)\left(\sum_{j=2}^{M_1+1}\alpha_{j1o}\right)\right.\right.\\
&\quad \left.\left.+\ (\gamma-M_1)\left(\sum_{j=2}^{M_2+1}\alpha_{j2o}\right)\right)^2 |M_1,M_2\right] \Big/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]\\
&\quad +\left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2 E\left(\epsilon_{1t}^2\right)\right] \Big/ \left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]
\end{aligned}
$$

Note that the weights in the numerator of the second expectation are the same weights as in the denominator. Assumption 4 implies that these weights are orthogonal to the squared first and second period errors. Further, $E(\epsilon_{11}^2) = E(\epsilon_{12}^2)$. Taking the unconditional expectation then yields:

$$
\begin{aligned}
E[q(w,\gamma)] &= (\gamma-\gamma_o)^2 E\left[\left((M_1-M_2)\gamma\alpha_{1o}-(\gamma-M_2)\left(\sum_{j=2}^{M_1+1}\alpha_{j1o}\right)\right.\right.\\
&\quad \left.\left.+\ (\gamma-M_1)\left(\sum_{j=2}^{M_2+1}\alpha_{j2o}\right)\right)^2 \Big/ \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right]\\
&\quad +E\left(\epsilon^2\right)
\end{aligned}
$$

The first term in the above expression is strictly greater than 0 for all $\gamma \neq \gamma_o$ and the second term does not depend upon γ. As a result, $E[q(w,\gamma_o)] < E[q(w,\gamma)]$ for all $\gamma \in \Gamma$ when $\gamma \neq \gamma_o$.

QED.

Proof of Lemma 4.G

Uniform convergence requires that

$$\max_{\gamma\in\Gamma}\left|\frac{1}{\mathcal{N}}\sum_{n=1}^{\mathcal{N}} q(w_n,\gamma) - E[q(w,\gamma)]\right| \xrightarrow{p} 0$$

Theorem 12.1 in Wooldridge states four conditions that the data and q must satisfy in order for the above condition to hold.

1. Γ is compact

   This condition is satisfied by assumption 8.

2. For each $\gamma \in \Gamma$, $q(\cdot,\gamma)$ is Borel measurable on $\mathcal{W}$

   $q(\cdot,\gamma)$ is measurable with respect to product σ-algebra of $\mathcal{B}(\mathbb{R}^{2+2\overline{M}}) \times 2^{\mathcal{J}}$, where $2^{\mathcal{J}}$ is the power set over the possible class sizes.

3. For each $w \in \mathcal{W}$, $q(w,\cdot)$ is continuous on Γ

   Our concentrated objective function is continuous in γ.

4. $|q(w,\gamma)| \leq b(w)$ for all $\gamma \in \Gamma$, where b is a nonnegative function on $\mathcal{W}$ such that $E[b(w)] < \infty$

Recall that $q(w,\gamma)$ is given by:

$$q(w,\gamma) = \frac{\left(\sum_{j=1}^{M_1+1} W_{j1}y_{j1} + \sum_{j=1}^{M_2+1} W_{j2}y_{j2}\right)^2}{\sum_{j=1}^{M_1+1} W_{j1}^2 + \sum_{j=1}^{M_2+1} W_{j2}^2}$$

Expanding the square and noting that $W_{jt}^2 y_{jt}^2 + W_{kt'}^2 y_{kt'}^2 \geq 2W_{jt}W_{kt'}y_{jt}y_{kt'}$ for all j,k,t,t' (the triangle inequality), we have:

$$q(w,\gamma) \leq \frac{(2+M_1+M_2)\left(\sum_{j=1}^{M_1+1} W_{j1}^2 y_{j1}^2 + \sum_{j=1}^{M_2+1} W_{j2}^2 y_{j2}^2\right)}{\sum_{j=1}^{M_1+1} W_{j1}^2 + \sum_{j=1}^{M_2+1} W_{j2}^2}$$

where the the leading term arises from replacing all the cross products using the triangle inequality.

Note that each of the terms in the denominator is positive, implying that:

$$q(w, \gamma) < (2 + M_1 + M_2) \left(\sum_{j=1}^{M_1+1} y_{j1}^2 + \sum_{j=1}^{M_2+1} y_{j2}^2 \right) = b(w)$$

where we have shown that $b(w) > q(w, \gamma)$ for all w.

We now show that $E[b(w)] < \infty$. Note that $E[b(w)]$ is given by:

$$E[b(w)] = E \left[(2 + M_1 + M_2) \sum_{t=1}^{2} \sum_{j=1}^{M_t+1} y_{jt}^2 \right]$$

Note also that, by the law of iterated expectations $E[b(w)] = E(E[b(w)|M_1, M_2])$. We first show that the inner expectation is bounded for all M_1, M_2 and then show that this guarantees the outer expectation is finite. Substituting in for y with the data generating process into the inner expectation yields:

$$E[b(w)|M_1, M_2] = (2 + M_1 + M_2) E \left[\sum_{t=1}^{2} \sum_{j=1}^{M_t+1} \left(\alpha_{jto} + \frac{\gamma_o}{M_t} \sum_{k \neq j}^{M_t+1} \alpha_{kto} + \epsilon_{jt} \right)^2 |M_1, M_2 \right]$$

Repeatedly using the triangle inequality after expanding the square implies:

$$E[b(w)|M_1, M_2] \leq (2+M_1+M_2) E \left[\sum_{t=1}^{2} \sum_{j=1}^{M_t+1} (M_t + 2) \left(\alpha_{jto}^2 + \frac{\gamma_o^2}{M_t^2} \sum_{k \neq j}^{M_t+1} \alpha_{kto}^2 + \epsilon_{jt}^2 \right) |M_1, M_2 \right]$$

Collecting α_{jto} terms and recognizing that $\gamma_o^2/M_t \leq \gamma_o^2$ implies that:

$$E[b(w)|M_1, M_2] \leq (2 + M_1 + M_2) E \left[(1 + \gamma_o^2) \sum_{t=1}^{2} \sum_{j=1}^{M_t+1} (M_t + 2) \left(\alpha_{jto}^2 + \epsilon_{jt}^2 \right) |M_1, M_2 \right]$$

We can take the expectation operator through yielding:

$$E[b(w)|M_1, M_2] \leq (2 + M_1 + M_2)(1 + \gamma_o^2) \sum_{t=1}^{2} \sum_{j=1}^{M_t+1} (M_t + 2) \left[E \left(\alpha_{jto}^2 \middle| M_1, M_2 \right) + E \left(\epsilon_{jt}^2 \right) \right]$$

where the conditioning is not necessary for the second expectation by assumption 4. Assumptions 5, 6, and 8 ensure that $E(\alpha_{jto}^2|M_1, M_2)$, $E(\epsilon_{jt}^2)$, and γ_o are all finite. Thus, $E[b(w)|M_1, M_2] < \infty$ for all M_1, M_2. Now, note that $E[b(w)] = E(E[b(w)|M_1, M_2]) \leq \max_{M_1,M_2} E[b(w)|M_1, M_2] < \infty$, where the last inequality arises from assumption 1.

QED

Proof of Lemma 5.G

To establish asymptotic normality, we now show that the six conditions of Theorem 12.3 in Wooldridge (2002) are satisfied.

1. γ_o must be in the interior of Γ

   This condition is satisfied by assumption 8.

2. Each element of $H(w,\gamma)$ is bounded in absolute value by a function $b(w)$ where $E[b(w)] < \infty$

   Recall that $q(w,\gamma)$ be written as:

$$\begin{aligned} q(w,\gamma) &= \frac{\left(\sum_{j=1}^{M_1+1} W_{j1}y_{j1} + \sum_{j=1}^{M_2+1} W_{j2}y_{j2}\right)^2}{\sum_{j=1}^{M_1+1} W_{j1}^2 + \sum_{j=1}^{M_2+1} W_{j2}^2} \\ &= \frac{\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1} W_{jt}W_{kt'}y_{jt}y_{kt'}}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2} \end{aligned}$$

   Denoting W'_{jt} as the first partial derivative with respect to γ,

$$\begin{aligned} W'_{11} &= [(1+2\gamma-M_2)(M_1-1)+1] \\ W'_{12} &= -[(1+2\gamma-M_1)(M_2-1)+1] \\ W'_{j1} &= -2\gamma+M_2 \qquad \text{for all } j>1 \\ W'_{j2} &= 2\gamma-M_1 \qquad \text{for all } j>1 \end{aligned}$$

   and W''_{jt} as the second partial derivative of W_{jt} with respect to γ,

$$\begin{aligned} W''_{11} &= 2(M_1-1) \\ W''_{12} &= -2(M_2-1) \\ W''_{j1} &= -2 \qquad \text{for all } j>1 \\ W''_{j2} &= 2 \qquad \text{for all } j>1 \end{aligned}$$

   We can then write the score as:

$$\begin{aligned} s(w,\gamma) &= \frac{2\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1} W'_{jt}W_{kt'}y_{jt}y_{kt'}}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2} \\ &- \frac{\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1} W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W'_{jt}W_{jt}\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right)^2} \end{aligned}$$

and the hessian as:

$$
\begin{aligned}
H(w,\gamma) &= \frac{2\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}(W''_{jt}W_{kt'}+W'_{jt}W'_{kt'})y_{jt}y_{kt'}}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2} \\
&- \frac{4\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W'_{jt}W_{kt'}y_{jt}y_{kt'}\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2} \\
&- \frac{\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(W'_{jt}W'_{jt}+W''_{jt}W_{jt})\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2} \\
&+ \frac{\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^3}
\end{aligned}
$$

We need to derive a bounding function such that $b(w) \geq |H(w,\gamma)|$ for all $\gamma \in \Gamma$. Note that:

$$
\begin{aligned}
|H(w,\gamma)| &\leq \frac{\left|2\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}(W''_{jt}W_{kt'}+W'_{jt}W'_{kt'})y_{jt}y_{kt'}\right|}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2} \\
&+ \frac{\left|4\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W'_{jt}W_{kt'}y_{jt}y_{kt'}\right)\right|\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2} \\
&+ \frac{\left|\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\right|\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(W'_{jt}W'_{jt}+W''_{jt}W_{jt})\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2} \\
&+ \frac{\left|\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\right|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^3}
\end{aligned}
$$

Repeatedly applying the triangle inequality and collecting terms yields:

$$
\begin{aligned}
|H(w,\gamma)| \quad \leq \quad & \frac{2\,(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W''_{jt})^2+(W'_{jt})^2+W_{jt}^2)y_{jt}^2\right)}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2} \\
+ \quad & \frac{4\,(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W'_{jt})^2+W_{jt}^2)y_{jt}^2\right)\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2} \\
+ \quad & \frac{(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(W'_{jt}W'_{jt}+W''_{jt}W_{jt})\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2} \\
+ \quad & \frac{(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^3}
\end{aligned}
$$

Denote W^*_{jt} as the weight given to y_{jt}^2 in the above expression:

$$
\begin{aligned}
W^*_{jt} \quad = \quad & \frac{2\,(2+M_1+M_2)\left[(W''_{jt})^2+(W'_{jt})^2+W_{jt}^2\right]}{\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2} \\
+ \quad & \frac{4\,(2+M_1+M_2)\,(W'^2_{jt}+W_{jt}^2)\left|\left(2\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W'_{kt}W_{kt}\right)\right|}{\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2\right)^2} \\
+ \quad & \frac{(2+M_1+M_2)\,W_{jt}^2\left|\left(2\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}(W'_{kt}W'_{kt}+W''_{kt}W_{kt})\right)\right|}{\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2\right)^2} \\
+ \quad & \frac{(2+M_1+M_2)\left(W_{jt}W_{jt'}y_{jt}\right)\left(2\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W'_{kt}W_{kt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2\right)^3}
\end{aligned}
$$

implying that:

$$|H(w,\gamma)| \leq \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W^*_{jt}y_{jt}^2$$

Note that W^*_{jt} is function only of the class sizes and γ and for any class sizes and γ it is finite. Since the expression on the left hand side of the above equation is increasing in W^*_{jt}, define B^*_{jt} as:

$$B^*_{jt} = \max_{\gamma} W^*_{jt}$$

which exists and is finite due to all elements of Γ being finite. Our bounding function is then:

$$b(w) = \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt}y_{jt}^2$$

We then need to establish that $E[b(w)] < \infty$. We first show that $E[b(w)|M_1, M_2] < \infty$:

$$\begin{aligned} E[b(w)|M_1,M_2] &= \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt} E(y^2_{jt}|M_1,M_2) \\ &= \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt} E\left[\left(\alpha_{jto} + \frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}\alpha_{kto} + \epsilon_{jt}\right)^2 |M_1,M_2\right] \end{aligned}$$

Repeatedly using the triangle inequality after expanding the square implies:

$$E[b(w)|M_1,M_2] \leq \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt}(M_t+2) E\left[\left(\alpha^2_{jto} + \frac{\gamma^2_o}{M^2_t}\sum_{k\neq j}^{M_t+1}\alpha^2_{kto} + \epsilon^2_{jt}\right) |M_1,M_2\right]$$

Collecting α_{jto} terms and recognizing that $\gamma^2_o/M_t \leq \gamma^2_o$ implies that:

$$E[b(w)|M_1,M_2] \leq (1+\gamma^2_o)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt}(M_t+2) E\left[\left(\alpha^2_{jto} + \epsilon^2_{jt}\right)|M_1,M_2\right]$$

Assumptions 1, 5, 6, and 8 ensure that B^*, $E(\alpha^2_{jto}|M_1, M_2)$, $E(\epsilon^2_{jt})$, and γ_o are all finite, implying that $E[b(w)|M_1, M_2] < \infty$. Now, note that $E[b(w)] = E(E[b(w)|M_1, M_2]) \leq \max_{M_1,M_2} E[b(w)|M_1, M_2] < \infty$, where the last inequality arises from assumption 1.

3. $s(w, \cdot)$ is continuously differentiable on the interior of Γ for all $w \in \mathcal{W}$

   Since $H(w, \gamma)$ is continuous in γ, $s(w, \cdot)$ is continuously differentiable.

4. $A_o \equiv E[H(w, \gamma_o)]$ is positive definite

   With only one parameter, this implies that the Hessian is strictly greater than zero when evaluated at the true γ. To test this condition, we evaluate the expected Hessian at γ_o. We first note that we can interchange the expectations and the partial derivatives: $E[H(w, \gamma)] = \partial^2 E[q(w, \gamma)]/\partial\gamma^2$. From Lemma 3.G, we know that $E[q(w, \gamma)]$ can be written as:

$$\begin{aligned} E[q(w,\gamma)] = (\gamma-\gamma_o)^2 E\Bigg[\Bigg((M_1-M_2)\gamma\alpha_{1o} - (\gamma - M_2)\left(\sum_{j=2}^{M_1+1}\alpha_{j1o}\right) \\ + (\gamma - M_1)\left(\sum_{j=2}^{M_2+1}\alpha_{j2o}\right)\Bigg)^2 \Bigg/ \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W^2_{jt}\Bigg] \\ + E\left(\epsilon^2\right) \end{aligned}$$

Note that γ affects three terms: $(\gamma - \gamma_o)^2$, the term inside the expectation, and the denominator. However, because we are going to evaluate the expected Hessian at γ_o, we only need the second derivative of the first term, $(\gamma-\gamma_o)^2$. All of the other partial derivatives will either be multiplied by $(\gamma-\gamma_o)^2$ or $(\gamma-\gamma_o)$, both of which are zero when $\gamma=\gamma_o$. The second derivative of $(\gamma-\gamma_o)^2$ with respect to γ is positive. This second derivative is then multiplied by the expectation of a squared object in the numerator and divided by the sum of squared objects in the denominator. Thus, the expectation of the Hessian evaluated at γ_o is strictly positive.

5. $E[s(w,\gamma_o)]=0$

   Note that $E[s(w,\gamma)]=\partial E[q(w,\gamma)]/\partial\gamma$. Differentiating with respect to γ leaves terms that are multiplied by $(\gamma-\gamma_o)$ or by $(\gamma-\gamma_o)^2$, implying that if we evaluate the derivative at $\gamma=\gamma_0$ then the expected score is zero.

6. Each element of $s(w,\gamma_o)$ has finite second moment.

   We first take the expected squared score conditional on M_1, M_2 which is given by:

$$E[s(w,\gamma)^2|M_1,M_2]=E\left(\left[\frac{2\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W'_{jt}W_{kt'}y_{jt}y_{kt'}}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2}\right.\right.$$
$$\left.\left.-\frac{\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1}W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right]^2\middle|M_1,M_2\right)$$

   Applying the triangle inequality and collecting terms yields:

$$\begin{aligned}E[s(w,\gamma)^2|M_1,M_2] \;\leq\; & E\left(\left[\frac{2\,(2+M_1+M_2)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W'_{jt})^2+W_{jt}^2)y_{jt}^2}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2}\right.\right.\\ & \left.\left.-\frac{\left((2+M_1+M_2)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right]^2\middle|M_1,M_2\right)\end{aligned}$$

Repeatedly applying the triangle inequality we can write:

$$
\begin{aligned}
E[s(w,\gamma)^2|M_1,M_2] \leq & \quad E\left(2\left[\frac{2(2+M_1+M_2)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W'_{jt})^2+W_{jt}^2)y_{jt}^2}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2}\right]^2 \Big| M_1,M_2\right) \\
& + E\left(2\left[\frac{\left((2+M_1+M_2)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right]^2 \Big| M_1,M_2\right) \\
\leq & \quad E\left(8(2+M_1+M_2)^2\left[\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W'_{jt})^2+W_{jt}^2)y_{jt}^2\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right] \Big| M_1,M_2\right) \\
& + E\left(8(2+M_1+M_2)^2\left[\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)^2\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^4}\right] \Big| M_1,M_2\right) \\
\leq & \quad E\left(8(2+M_1+M_2)^3\left[\frac{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W'_{jt})^2+W_{jt}^2)^2y_{jt}^4}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right] \Big| M_1,M_2\right) \\
& + E\left(8(2+M_1+M_2)^3\left[\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^4y_{jt}^4\right)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^4}\right] \Big| M_1,M_2\right)
\end{aligned}
$$

Note that the expectation is taken with respect to the y's conditional on the M's and γ. Denote W_{jt}^* as the aggregate weight given to y_{jt}^4 in the above expression:

$$
W_{jt}^* = \frac{8(2+M_1+M_2)^3((W'_{jt})^2+W_{jt}^2)^2}{\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2\right)^2} + \frac{8(2+M_1+M_2)^3(W_{jt}^4\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W'_{kt}W_{kt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2\right)^4}
$$

where we know that W_{jt}^* is finite as the denominator is greater than zero, M_1 and M_2 are finite, and γ is finite. Substituting in with W_{jt}^* in the inequality yields:

$$
E[s(w,\gamma)^2|M_1,M_2] \leq \sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^*E(y_{jt}^4|M_1,M_2)
$$

Substituting in for y_{jt} and repeatedly applying the triangle inequality yields:

$$
\begin{aligned}
E[s(w,\gamma)^2|M_1,M_2] \leq & \sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^*E\left[\left(\alpha_{jto}+\frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}\alpha_{kto}+\epsilon_{jt}\right)^4|M_1,M_2\right] \\
\leq & \sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^*E\left[(M_t+2)^2\left(\alpha_{jto}^2+\left(\frac{\gamma_o}{M_t}\right)^2\sum_{k\neq j}^{M_t+1}\alpha_{kto}^2+\epsilon_{jt}^2\right)^2|M_1,M_2\right] \\
\leq & \sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^*E\left[(M_t+2)^3\left(\alpha_{jto}^4+\left(\frac{\gamma_o}{M_t}\right)^4\sum_{k\neq j}^{M_t+1}\alpha_{kto}^4+\epsilon_{jt}^4\right)|M_1,M_2\right]
\end{aligned}
$$

Collecting terms we have:

$$
E[s(w,\gamma)^2|M_1,M_2] \leq \sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^*\left[(M_t+2)^3\left(1+\frac{\gamma_o^4}{M_t^3}\right)E\left(\alpha_{jto}^4|M_1,M_2\right)+E(\epsilon_{jt}^4)\right]
$$

W^*_{jt}, γ_o, and M_t are all finite and since the fourth moments of α and ϵ's are finite by assumptions 5 and 6, the expression is finite for all $\gamma \in \Gamma$ and for all M_t. Further, $E[s(w,\gamma_o)^2] \leq \max_{M_1,M_2} E[s(w,\gamma_o)^2|M_1,M_2] < \infty$.

QED □

2 Proof of Lemma 2.G: $M_{1n} = M_{2n} = M_n$

Proof. The algebra required to derive the simplified residual expressions for the general class size case is terribly cumbersome. For a sense of how the algebra works, we instead show how to derive the residual equations for a slightly simpler problem, the case where $M_{1n} = M_{2n} = M_n$.

We take as a staring point here the results of Lemma 1.G when $M_{1n} = M_{2n} = M_n$. While we do not derive the result here, following the steps in Lemma 1.G would yield the following solutions for α_{1n} and α_{i1n}.

$$\alpha_{1n} = \left(\tilde{A}_n + \hat{\gamma}\tilde{B}_n\right)(y_{11n} + y_{12n}) + \left(\tilde{A}_n\frac{\hat{\gamma}}{M_n} + \tilde{B}_n\frac{\hat{\gamma}(M_n-1)+M_n}{M_n}\right)\sum_{t=1}^{2}\sum_{j=2}^{M_n+1} y_{jtn}$$

$$\begin{aligned}\alpha_{i1n} &= \left(\tilde{B}_n + \tilde{C}_n\frac{\hat{\gamma}}{M_n} + \tilde{D}_n\frac{(M_n-1)\hat{\gamma}}{M_n}\right)y_{11n} + \left(\tilde{B}_n + \tilde{E}_n\hat{\gamma}\right)y_{12n} + \left(\tilde{B}_n\frac{\hat{\gamma}}{M_n} + \tilde{C}_n + \tilde{D}_n\frac{(M_n-1)\hat{\gamma}}{M_n}\right)y_{i1n} \\ &+ \left(\tilde{B}_n\frac{\hat{\gamma}}{M_n} + \tilde{C}_n\frac{\hat{\gamma}}{M_n} + \tilde{D}_n\frac{M_n+(M_n-2)\hat{\gamma}}{M_n}\right)\sum_{j=2,\, j\neq i}^{M_n+1} y_{j1n} + \left(\tilde{B}_n\frac{\hat{\gamma}}{M_n} + \tilde{E}_n\frac{M_n+(M_n-1)\hat{\gamma}}{M_n}\right)\sum_{j=2}^{M_n+1} y_{j2n}\end{aligned}$$

where

$$\tilde{A}_n = \frac{\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2M_n^2}{2(1+\gamma)^2(\gamma-M_n)^2}$$

$$\tilde{B}_n = \frac{\gamma^2 - \gamma(2+\gamma)M_n}{2(1+\gamma)^2(\gamma-M_n)^2}$$

$$\tilde{C}_n = \frac{\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2(8+\gamma(12+5\gamma))M_n^2 - 4\gamma(1+\gamma)^2(2+\gamma)M_n^3 + 2(1+\gamma)^4M_n^4}{2(1+\gamma)^2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}$$

$$\tilde{D}_n = \frac{\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2(6+\gamma(8+3\gamma))M_n^2 - 2\gamma(1+\gamma)^2(2+\gamma)M_n^3}{2(1+\gamma)^2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}$$

$$\tilde{E}_n = \frac{\gamma^2(\gamma(M_n-1)+2M_n)^2}{2(1+\gamma)^2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}$$

The form of α_{i2n} is identical to the above formulation for α_{i1n} except that the time indices are swapped on all the terms. Notice that here we have written the inverse components directly as functions of γ and M_n. The extra notation utilized in the general M case is not necessary here since we are not going to show how to derive $\mathbf{X}_n^{-1}$ directly. However, it is immediately clear that finding a simplified version for the residual equations will be easier in this case since there are simply fewer terms to deal with.

With the equations for the abilities in hand, we can begin substituting into the residual equations. Consider the residual for individual 1 in block n at time period 1,

$$e_{11n} = y_{11n} - \alpha_{1n} - \frac{\gamma}{M_n}\sum_{j=2}^{M_n+1}\alpha_{j1n}$$

Substituting in the solutions for α_{1n} and α_{j1n} and combining like terms yields the following:

$$\begin{aligned} e_{11n} \;=\; & y_{11n}\left(1 - \tilde{A}_n - 2\gamma\tilde{B}_n - \tilde{C}_n\frac{\gamma^2}{M_n} - \tilde{D}_n\frac{\gamma^2(M_n-1)}{M_n}\right) - y_{12n}\left(\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n\right) \\ & - \left(\sum_{j=2}^{M_n+1} y_{j1n}\right)\left(\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{C}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n^2} + \tilde{D}_n\frac{\gamma(\gamma+M_n(M_n+M_n\gamma-2\gamma-1))}{M_n^2}\right) \\ & - \left(\sum_{j=2}^{M_n+1} y_{j2n}\right)\left(\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{E}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n}\right) \end{aligned} \tag{5}$$

Using the formulas for $\tilde{A}_n$, $\tilde{B}_n$, $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ we show that the coefficients on the y's simplify quite nicely. First we illustrate how $\tilde{A}_n$, $\tilde{B}_n$, $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ are functionally related.

Property 1

The components of $\mathbf{X}_n^{-1}$ are interrelated according to the following:

$$\tilde{A}_n = \tilde{B}_n + \frac{M_n^2}{2(\gamma-M_n)^2}, \quad \tilde{C}_n = \tilde{D}_n + \frac{M_n^2}{(\gamma-M_n)^2}, \quad \tilde{D}_n = \tilde{B}_n + \frac{V_n}{2}, \quad \tilde{E}_n = \tilde{B}_n - \frac{V_n}{2}$$

where

$$V_n = \frac{\gamma M_n^2(-\gamma + 2M_n + \gamma M_n)}{(\gamma-M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Proof of Property 1

Solving for $\tilde{A}_n$ as a function of $\tilde{B}_n$ is rather straightforward as they have the same denominator.

$$\begin{aligned}\tilde{A}_n - \tilde{B}_n &= \frac{\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2 - (\gamma^2 - \gamma(2+\gamma)M_n)}{2(1+\gamma)^2(\gamma - M_n)^2} \\ &= \frac{M_n^2}{2(\gamma - M_n)^2}\end{aligned}$$

In order to relate $\tilde{C}_n$ to $\tilde{B}_n$, we first show how $\tilde{C}_n$ is related to $\tilde{D}_n$ and then how $\tilde{D}_n$ is related to $\tilde{B}_n$. Below are the formulas for $\tilde{C}_n$ and $\tilde{D}_n$.

$$\tilde{C}_n = \frac{\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2(8+\gamma(12+5\gamma))M_n^2 - 4\gamma(1+\gamma)^2(2+\gamma)M_n^3 + 2(1+\gamma)^4 M_n^4}{2(1+\gamma)^2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

$$\tilde{D}_n = \frac{\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2(6+\gamma(8+3\gamma))M_n^2 - 2\gamma(1+\gamma)^2(2+\gamma)M_n^3}{2(1+\gamma)^2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Both terms share the same denominator, and in fact share the same first two terms in the numerator. Subtracting $\tilde{D}_n$ from $\tilde{C}_n$ and simplifying yields

$$\tilde{C}_n - \tilde{D}_n = \frac{M_n^2}{(\gamma - M_n)^2}$$

Next we want to find the difference between $\tilde{D}_n$ and $\tilde{B}_n$. This difference is more complicated than the first two since they do not share the same denominator. However we can easily get a common denominator since the denominator for $\tilde{B}_n$ is simply missing one term present in the denominator of $\tilde{D}_n$. Thus we can write the difference as

$$\begin{aligned}\tilde{D}_n - \tilde{B}_n &= \frac{\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2(6+\gamma(8+3\gamma))M_n^2 - 2\gamma(1+\gamma)^2(2+\gamma)M_n^3}{2(1+\gamma)^2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} \\ &- \frac{(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)(\gamma^2 - \gamma(2+\gamma)M_n)}{2(1+\gamma)^2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}\end{aligned}$$

Combining terms and simplifying yields

$$\begin{aligned}\tilde{D}_n - \tilde{B}_n &= \frac{\gamma M_n^2(\gamma - 2M_n - \gamma M_n)}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} \\ &= \frac{V_n}{2}\end{aligned}$$

The last piece is to relate $\tilde{E}_n$ to $\tilde{B}_n$. Just as with $\tilde{D}_n$ we need to find a common denominator.

$$\begin{aligned}
\tilde{E}_n - \tilde{B}_n &= \frac{\gamma^2(\gamma(M_n-1)+2M_n)^2 - (\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)(\gamma^2-\gamma(2+\gamma)M_n)}{2(1+\gamma)^2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)} \\
&= \frac{\gamma M_n^2(-\gamma+2M_n+\gamma M_n)}{2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)} \\
&= -\frac{V_n}{2}
\end{aligned}$$

QED

Using Property 1, we now show that the coefficients on the observed grades in Equation (**??**) have other appealing properties. Then we use these properties to simplify Equation (**??**), in an effort to arrive at a simplified version of the least squares problem as a function of γ.

Property 2

Equation (**??**) describes the prediction error for the first outcome of the individual observed twice in block n. In Equation (**??**), the coefficient on y_{11n} is equal to the coefficient on y_{12n}, and the coefficient on $\sum_{j=2}^{M_n+1} y_{j1n}$ is equal in magnitude but of the opposite sign as the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$.

Proof of Property 2

The coefficient on y_{11n} is given by

$$1 - \tilde{A}_n - 2\gamma\tilde{B}_n - \tilde{C}_n\frac{\gamma^2}{M_n} - \tilde{D}_n\frac{\gamma^2(M_n-1)}{M_n}$$

Substituting in for $\tilde{A}_n$, $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ as a function of $\tilde{B}_n$ using Property 1 and simplifying yields the following

$$1-\tilde{A}_n-2\gamma\tilde{B}_n-\tilde{C}_n\frac{\gamma^2}{M_n}-\tilde{D}_n\frac{\gamma^2(M_n-1)}{M_n} = \frac{2\gamma^2-2\gamma(\gamma-2)M_n+M_n^2-V_n\gamma^2(\gamma-M_n)^2}{2(\gamma-M_n)^2}-\tilde{B}_n(1+\gamma)^2$$

The coefficient on y_{12n} is given by

$$\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n$$

Again making the appropriate substitutions allowed by Property 1, we can re-write this expression as

$$\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n = \frac{M_n^2 - V_n\gamma^2(\gamma-M_n)^2}{2(\gamma-M_n)^2} + \tilde{B}_n(1+\gamma)^2$$

Finally, taking the difference between the coefficients on y_{11n} and y_{12n}, we find

$$\begin{aligned}
&\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n - (1 - \tilde{A}_n - 2\gamma\tilde{B}_n - \tilde{C}_n\frac{\gamma^2}{M_n} - \tilde{D}_n\frac{\gamma^2(M_n-1)}{M_n}) \\
&= 2\tilde{B}_n(1+\gamma)^2 + \frac{M_n^2 - V_n\gamma^2(\gamma-M_n)^2 - 2\gamma^2 + 2\gamma(\gamma-2)M_n - M_n^2 + V_n\gamma^2(\gamma-M_n)^2}{2(\gamma-M_n)^2} \\
&= 2\tilde{B}_n(1+\gamma)^2 + \frac{-2\gamma^2 + 2\gamma(\gamma-2)M_n}{2(\gamma-M_n)^2} \\
&= \frac{\gamma^2 - \gamma(2+\gamma)M_n}{(\gamma-M_n)^2} + \frac{-2\gamma^2 + 2\gamma(\gamma-2)M_n}{2(\gamma-M_n)^2} \\
&= 0
\end{aligned}$$

where the second to last line results from substituting in our formula for $\tilde{B}_n$ given in Equation (**??**).

Now we show that the coefficient on $\sum_{j=2}^{M_n+1} y_{j1n}$ is equal in magnitude but of the opposite sign as the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$. The coefficient on $\sum_{j=2}^{M_n+1} y_{j1n}$ is given by

$$\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{C}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n^2} + \tilde{D}_n\frac{\gamma(\gamma+M_n(M_n+M_n\gamma-2\gamma-1))}{M_n^2}$$

and the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$ is given by

$$\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{E}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n}$$

If we add these two coefficients together we arrive at the following expression

$$\begin{aligned}
& 2\tilde{A}_n\tfrac{\gamma}{M_n} + 2\tilde{B}_n\tfrac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{C}_n\tfrac{\gamma M_n(1+\gamma)-\gamma^2}{M_n^2} \\
+ \quad & \tilde{D}_n\tfrac{\gamma(\gamma+M_n(M_n+M_n\gamma-2\gamma-1))}{M_n^2} + \tilde{E}_n\tfrac{\gamma M_n(1+\gamma)-\gamma^2}{M_n}
\end{aligned}$$

Now, we substitute for $\tilde{A}_n$, $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ as functions of $\tilde{B}_n$ from Property 1. After some manipulation, we can write the above expression in the following form

$$2\tilde{B}_n(1+\gamma)^2 + \frac{4\gamma M_n + 2\gamma^2 M_n - 2\gamma^2}{2(\gamma-M_n)^2}$$

Notice that this expression contains no V_n terms as they cancel out when substituting in for $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$. The last step is to substitute in for $\tilde{B}_n$ from Equation (**??**).

$$\frac{\gamma^2 - \gamma(2+\gamma)M_n}{(\gamma-M_n)^2} + \frac{4\gamma M_n + 2\gamma^2 M_n - 2\gamma^2}{2(\gamma-M_n)^2}$$

All of the terms in the above expression cancel out, proving that the sum of the coefficients on $\sum_{j=2}^{M_n+1} y_{j1n}$ and $\sum_{j=2}^{M_n+1} y_{j2n}$ are equal in magnitude and of the opposite sign.

QED

Now we return to Equation (**??**), which describe the prediction error for the first observation of the student observed twice in block n. Using Properties 1 and 2 we will show how to simplify this expression, and in turn describe how the prediction errors for all of the other outcomes in block n can be similarly simplified. This will yield a simplified version of the original least squares problem strictly as a function of γ, M_n, and y.

Property 2 indicates that

$$1 - \tilde{A}_n - 2\gamma\tilde{B}_n - \tilde{C}_n\frac{\gamma^2}{M_n} - \tilde{D}_n\frac{\gamma^2(M_n-1)}{M_n} = \tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n$$

and

$$\begin{aligned}&\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{C}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n^2} + \tilde{D}_n\frac{\gamma(\gamma+M_n(M_n+M_n\gamma-2\gamma-1))}{M_n^2}\\ &= -(\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{E}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n})\end{aligned}$$

We now proceed to solve for each of these coefficients strictly as a function of γ. First, we solve for the coefficient on y_{12n}.

By substituting for $\tilde{A}_n$ and $\tilde{E}_n$ from Property 1, we can write the coefficient on y_{12n} in the following way

$$\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n = \frac{M_n^2 - V_n\gamma^2(\gamma-M_n)^2}{2(\gamma-M_n)^2} + \tilde{B}_n(1+\gamma)^2$$

To solve for this as a function of γ we need to substitute in for $\tilde{B}_n$ and V_n. Substituting in for $\tilde{B}_n$ from Equation (**??**) and V_n from Property 1 yields

$$= \frac{(\gamma(\gamma-(2+\gamma)M_n)+M_n^2)(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2) - \gamma^2M_n^2(\gamma^2-2\gamma M_n-\gamma^2M_n)}{2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}$$

We can re-arrange this expression in the following manner:

$$\begin{aligned}= \quad & \frac{(\gamma^2-2\gamma M_n+M_n^2)(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}{2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}\\ & -\frac{\gamma^2M_n(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2+\gamma^2M_n-2\gamma M_n^2-\gamma^2M_n^2)}{2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}\end{aligned}$$

where we split the expression simply for ease of presentation. The numerator in the second line simplifies greatly, such that the entire expression simplifies to

$$= \frac{(\gamma^2 - 2\gamma M_n + M_n^2)(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2) - \gamma^2 M_n(\gamma^2 - 2\gamma M_n + M_n^2)}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

The numerator then factors to produce

$$= \frac{(\gamma - M_n)^2(M_n + \gamma(M_n - 1))^2}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Finally, we cancel out the common terms in the numerator and denominator to yield

$$\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n = \frac{(M_n + \gamma(M_n - 1))^2}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

This gives us the coefficient on y_{11n} and y_{12n} in the expression for e_{11n} as a function of γ. Now we proceed to solve for the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$ as a function of γ.

Using Property 1, we can write the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$ in the following fashion:

$$\tilde{A}_n \frac{\gamma}{M_n} + \tilde{B}_n \frac{\gamma(\gamma + M_n - 1) + M_n}{M_n} + \tilde{E}_n \frac{\gamma M_n(1+\gamma) - \gamma^2}{M_n} = \tilde{B}_n(1+\gamma)^2 + \frac{\gamma M_n}{2(\gamma - M_n)^2} - \frac{V_n(\gamma M_n(1+\gamma) - \gamma^2)}{2M_n}$$

Substituting for $\tilde{B}_n$ from Equation (??)and re-arranging yields

$$\frac{\gamma^2 - \gamma M_n - \gamma^2 M_n}{2(\gamma - M_n)^2} - \frac{V_n(\gamma M_n(1+\gamma) - \gamma^2)}{2M_n}$$

Substituting for V_n from Property 1 and finding a common denominator yields

$$\frac{(\gamma^2 - \gamma M_n - \gamma^2 M_n)(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2) - \gamma M_n(\gamma - 2M_n - \gamma M_n)(\gamma M_n(1+\gamma) - \gamma^2)}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

After some manipulation the numerator of the above expression simplifies to yield

$$\frac{-(\gamma - M_n)^2(\gamma M_n(1+\gamma) - \gamma^2)}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Canceling out the common terms in the numerator and denominator yields

$$\frac{-(\gamma M_n(1+\gamma) - \gamma^2)}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Finally we can substitute our simplified versions of the coefficients on y_{11n}, y_{12n}, $\sum_{j=2}^{M_n+1} y_{j2n}$, and $\sum_{j=2}^{M_n+1} y_{j2n}$ back into the equation for e_{11n}, described in Equation (??).

$$e_{11n} = \frac{(M_n + \gamma(M_n - 1))^2}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}(y_{11n} - y_{12n}) + \frac{\gamma(M_n + \gamma(M_n - 1))}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n})$$

This simplifies further to produce

$$e_{11n} = \frac{(M_n + \gamma(M_n - 1))}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} \left((M_n + \gamma(M_n - 1))(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n}) \right)$$

We now have the component of the least squares problem that corresponds to the residual for student 1 in block n as a function of γ with the α's concentrated out. Next, we need to find similar expressions for e_{12n} and e_{itn}.

Finding a version of e_{12n} as function of γ is simple since it takes a form that is essentially identical to e_{11n}. The expression for e_{12n} is given by

$$e_{12n} = y_{12n} - \alpha_{1n} - \frac{\gamma}{M_n} \sum_{j=2}^{M_n+1} \alpha_{j2n}$$

which after substituting for α using the results from Lemma 1 yields

$$\begin{aligned} e_{12n} \;=\; & y_{12n}\left(1 - \tilde{A}_n - 2\gamma\tilde{B}_n - \tilde{C}_n\frac{\gamma^2}{M_n} - \tilde{D}_n\frac{\gamma^2(M_n-1)}{M_n}\right) - y_{11n}\left(\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n\right) \\ & - \left(\sum_{j=2}^{M_n+1} y_{j2n}\right)\left(\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{C}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n^2} + \tilde{D}_n\frac{\gamma(\gamma+M_n(M_n+M_n\gamma-2\gamma-1))}{M_n^2}\right) \\ & - \left(\sum_{j=2}^{M_n+1} y_{j1n}\right)\left(\tilde{A}_n\frac{\gamma}{M_n} + \tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n} + \tilde{E}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n}\right) \end{aligned}$$

This equation is identical to the equation for e_{11n} except that all the time subscripts are changed. However, we know from Property 2 that the coefficients on y_{11n} and y_{12n} are equal in this expression and that coefficients on $\sum_{j=2}^{M_n+1} y_{j2n}$ and $\sum_{j=2}^{M_n+1} y_{j2n}$ are equal but of the opposite sign. Thus, $e_{12n} = -e_{11n}$.

To get the final piece of the least squares problem with the α's concentrated out we need to substitute for α in e_{itn}. where $i > 1$. To find a simplified formula for e_{itn} consider first substituting in for α in e_{21n}. The formula for e_{21n} from the basic least squares problem can be written as follows:

$$e_{21n} = y_{21n} - \alpha_{21n} - \frac{\gamma}{M_n}\left(\alpha_{1n} + \sum_{j=3}^{M_n+1} \alpha_{j1n}\right)$$

Substituting in for α from Lemma 1 and combining like terms yields the following expression:

$$\begin{aligned}
e_{21n} \quad = \quad & y_{21n}\left(1-\tilde{A}_n\frac{\gamma^2}{M_n^2}-2\tilde{B}_n\frac{M_n\gamma+\gamma^2(M_n-1)}{M_n^2}-\tilde{C}_n\frac{M_n^2+\gamma^2(M_n-1)}{M_n^2}-\tilde{D}_n\frac{2\gamma M_n(M_n-1)+\gamma^2(M_n-1)(M_n-2)}{M_n^2}\right) \\
& -y_{11n}\left(\tilde{A}_n\frac{\gamma}{M_n}+\tilde{B}_n\frac{M_n+(M_n-1)\gamma+\gamma^2}{M_n}+\tilde{C}_n\frac{\gamma M_n+(M_n-1)\gamma^2}{M_n^2}+\tilde{D}_n\frac{(M_n-1)\gamma(M_n+(M_n-1)\gamma)}{M_n^2}\right) \\
& -y_{12n}\left(\tilde{A}_n\frac{\gamma}{M_n}+\tilde{B}_n\frac{M_n+(M_n-1)\gamma+\gamma^2}{M_n}+\tilde{E}_n\frac{\gamma M_n+(M_n-1)\gamma^2}{M_n}\right) \\
& -\left(\sum_{j=3}^{M_n+1}y_{j1n}\right)\left(\tilde{A}_n\frac{\gamma^2}{M_n^2}+2\tilde{B}_n\frac{\gamma M_n+\gamma^2(M_n-1)}{M_n^2}+\tilde{C}_n\frac{2\gamma M_n+\gamma^2(M_n-2)}{M_n^2}+\tilde{D}_n\frac{(M_n+\gamma(M_n-2))^2+(M_n-1)\gamma^2}{M_n^2}\right) \\
& -\left(\sum_{j=2}^{M_n+1}y_{j2n}\right)\left(\tilde{A}_n\frac{\gamma^2}{M_n^2}+2\tilde{B}_n\frac{\gamma M_n+\gamma^2(M_n-1)}{M_n^2}+\tilde{E}_n\frac{(M_n+(M_n-1)\gamma)^2}{M_n^2}\right) \qquad (6)
\end{aligned}$$

To simplify the above expression, we follow the same strategy employed in simplifying e_{11n}.

Property 3

The coefficients on y_{11n} and y_{12n} in the equation for e_{21n} are equal in magnitude but of the opposite sign. The same relationship exists between the coefficients on $\sum_{j=3}^{M_n+1} y_{j1n}$ and $\sum_{j=2}^{M_n+1} y_{j2n}$. In addition, the coefficient on y_{21n} is identical to the coefficient on $\sum_{j=3}^{M_n+1} y_{j1n}$.

Proof of Property 3

The first step is to examine the coefficients on y_{11n} and y_{12n}. Our work is simple here since the coefficients on y_{11n} and y_{12n} in the expression for e_{21n} and the coefficients on $\sum_{j=2}^{M_n+1} y_{j1n}$ and $\sum_{j=2}^{M_n+1} y_{j2n}$ in the expression for e_{11n} are exactly the same. Thus, we know they are opposite in sign, and of magnitude

$$\frac{(\gamma M_n(1+\gamma)-\gamma^2)}{2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}$$

by Property 2.

Now we turn to the coefficients on y_{21n}, $\sum_{j=3}^{M_n+1} y_{j1n}$, and $\sum_{j=2}^{M_n+1} y_{j2n}$. Using the results from Property 1 relating $\tilde{A}_n$, $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ to $\tilde{B}_n$, we can re-write the coefficient on $\sum_{j=3}^{M_n+1} y_{j1n}$ in the following fashion:

$$\tilde{A}_n\frac{\gamma^2}{M_n^2}+2\tilde{B}_n\frac{\gamma M_n+\gamma^2(M_n-1)}{M_n^2}+\tilde{C}_n\frac{2\gamma M_n+\gamma^2(M_n-2)}{M_n^2}+\tilde{D}_n\frac{(M_n+\gamma(M_n-2))^2+(M_n-1)\gamma^2}{M_n^2}$$
$$=\tilde{B}_n(1+\gamma)^2+\frac{V_n(\gamma^2-2\gamma M_n(1+\gamma)+M_n^2(1+\gamma)^2)}{2M_n^2}+\frac{2\gamma M_n(2+\gamma)-3\gamma^2}{2(\gamma-M_n)^2}$$

Next, substituting in for $\tilde{A}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ in the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$ and simplifying yields

$$\tilde{A}_n\frac{\gamma^2}{M_n^2}+2\tilde{B}_n\frac{\gamma M_n+\gamma^2(M_n-1)}{M_n^2}+\tilde{E}_n\frac{(M_n+(M_n-1)\gamma)^2}{M_n^2}$$
$$=\tilde{B}_n(1+\gamma)^2-\frac{V_n(\gamma^2-2\gamma M_n(1+\gamma)+M_n^2(1+\gamma)^2)}{2M_n^2}+\frac{\gamma^2}{2(\gamma-M_n)^2}$$

Adding together the simplified expressions for the coefficients on $\sum_{j=3}^{M_n+1} y_{j1n}$ and $\sum_{j=2}^{M_n+1} y_{j2n}$ yields

$$2\tilde{B}_n(1+\gamma)^2+\frac{\gamma^2}{2(\gamma-M_n)^2}+\frac{2\gamma M_n(2+\gamma)-3\gamma^2}{2(\gamma-M_n)^2}$$

after the terms including V_n cancel each other. Substituting in our expression for $\tilde{B}_n$ from Equation (**??**) yields

$$\frac{\gamma^2-\gamma(2+\gamma)M_n}{(\gamma-M_n)^2}+\frac{2\gamma M_n+\gamma^2M_n-\gamma^2}{(\gamma-M_n)^2}$$

All the terms in the above expression cancel out, indicating that

$$\tilde{A}_n \frac{\gamma^2}{M_n^2} + 2\tilde{B}_n \frac{\gamma M_n + \gamma^2(M_n-1)}{M_n^2} + \tilde{C}_n \frac{2\gamma M_n + \gamma^2(M_n-2)}{M_n^2} + \tilde{D}_n \frac{(M_n + \gamma(M_n-2))^2 + (M_n-1)\gamma^2}{M_n^2}$$
$$= -\left(\tilde{A}_n \frac{\gamma^2}{M_n^2} + 2\tilde{B}_n \frac{\gamma M_n + \gamma^2(M_n-1)}{M_n^2} + \tilde{E}_n \frac{(M_n + (M_n-1)\gamma)^2}{M_n^2}\right)$$

or that the coefficients on $\sum_{j=3}^{M_n+1} y_{j1n}$ and $\sum_{j=2}^{M_n+1} y_{j2n}$ are equal in magnitude but of the opposite sign.

Finally, we can substitute in for $\tilde{A}_n$, $\tilde{C}_n$, $\tilde{D}_n$, and $\tilde{E}_n$ as function of $\tilde{B}_n$ from Property 1 in the coefficient for y_{21n}. After some simplification we can show that this coefficient can be written as

$$-\tilde{B}_n(1+\gamma)^2 - \frac{V_n(\gamma^2 - 2\gamma M_n(1+\gamma) + M_n^2(1+\gamma)^2)}{2M_n^2} - \frac{2\gamma M_n(2+\gamma) - 3\gamma^2}{2(\gamma - M_n)^2}$$

Comparing this to the coefficient on $\sum_{j=3}^{M_n+1} y_{j1n}$ as shown above indicates that these two expressions are exactly the same, except that the signs are flipped on all the terms. Thus, the coefficients for y_{21n} and $\sum_{j=3}^{M_n+1} y_{j1n}$ are equal in magnitude but of the opposite sign.

QED

All that remains is to find the expression for these three coefficients as a function of γ. We can work with the easiest formula since they are all identical. Recall that the coefficient on $\sum_{j=2}^{M_n+1} y_{j2n}$ can be written

$$\tilde{B}_n(1+\gamma)^2 - \frac{V_n(\gamma^2 - 2\gamma M_n(1+\gamma) + M_n^2(1+\gamma)^2)}{2M_n^2} + \frac{\gamma^2}{2(\gamma - M_n)^2}$$

Substituting in for V_n yields

$$\tilde{B}_n(1+\gamma)^2 - \frac{\gamma(\gamma M_n + 2M_n - \gamma)(M_n^2(1+\gamma)^2 + \gamma^2 - 2\gamma M_n(1+\gamma))}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} + \frac{\gamma^2}{2(\gamma - M_n)^2}$$

Finding a common denominator and re-arranging yields

$$\tilde{B}_n(1+\gamma)^2 + \frac{\gamma^2((\gamma - M_n)^2 + \gamma M_n(2M_n - \gamma + \gamma M_n)) - \gamma(\gamma M_n + 2M_n - \gamma)((\gamma - M_n)^2 + \gamma M_n(2M_n + \gamma M_n - 2\gamma))}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Finally, substituting in for $\tilde{B}_n$, finding a common denominator, and eliminating terms yields.

$$\frac{\gamma^2((\gamma - M_n)^2 + \gamma M_n(2M_n - \gamma + \gamma M_n)) - \gamma^2(\gamma M_n + 2M_n - \gamma)}{2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

The above expression simplifies further to

$$\frac{\gamma^2}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}$$

Now we have expressions for all the terms in the equation for e_{21n}. We can substitute back in and write the residual as a simple function of γ.

$$e_{21n} = \frac{\gamma(M_n + \gamma(M_n - 1))}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}(y_{11n} - y_{12n}) + \frac{\gamma^2}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n})$$

Notice that in the residual for e_{21n}, we can combine y_{21n} and $\sum_{j=3}^{M_n+1} y_{j1n}$ since they share the exact same coefficient. This means that the form of e_{j1n} for all $j > 1$ will take the exact form as the equation for e_{21n}. In addition, if we were to write down the equation for e_{22n}, it would take the exact same form as the equation for e_{21n}, except the coefficients would be swapped across the two time periods. As a result, $e_{22n} = -e_{21n}$. These relationships will allow us to greatly simplify the least squares problem.

We can simplify the solution for e_{21n} by factoring out the common terms in the numerator and denominator of each term. Doing so yields

$$e_{21n} = \frac{\gamma}{2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)} \left((M_n + \gamma(M_n - 1))(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n}) \right)$$

Finally we have all the components of the least squares problem strictly as functions of y, γ, and M_n. Re-writing the least squares problem in terms of the residuals yields

$$\min_{\gamma} \sum_{n=1}^{N} \left(e_{11n}^2 + e_{12n}^2 + \sum_{j=2}^{M_n+1} (e_{i1n}^2 + e_{i2n}^2) \right)$$

Using the fact that $e_{12n} = -e_{11n}$, $e_{j1n} = e_{21n}$ for i greater than 3, and $e_{22n} = -e_{21n}$ we can simplify the above expression to

$$\min_{\gamma} \; 2 \sum_{n=1}^{N} (e_{11n}^2 + M_n e_{21n}^2)$$

Now substituting in for the residuals using the results previously derived yields

$$\min_{\gamma} \; 2 \sum_{n=1}^{N} \Bigg[\frac{(M_n + \gamma(M_n - 1))^2}{4(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)^2} \left((M_n + \gamma(M_n - 1))(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n}) \right)^2$$
$$+ \frac{\gamma^2 M_n}{4(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)^2} \left((M_n + \gamma(M_n - 1))(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n}) \right)^2 \Bigg]$$

Notice that the terms inside the squares are exactly the same. We can re-arrange the above expression by combining like terms.

$$\min_{\gamma} \; 2\sum_{n=1}^{N}\left[\frac{(M_n+\gamma(M_n-1))^2+\gamma^2 M_n}{4(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)^2}\left((M_n+\gamma(M_n-1))(y_{11n}-y_{12n})+\gamma\sum_{j=2}^{M_n+1}(y_{j2n}-y_{j1n})\right)^2\right]$$

Simplifying the leading term leaves us with the following least squares problem,

$$\min_{\gamma} \; \sum_{n=1}^{N}\frac{\left((M_n+\gamma(M_n-1))(y_{11n}-y_{12n})+\gamma\sum_{j=2}^{M_n+1}(y_{j2n}-y_{j1n})\right)^2}{2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}$$

Notice that if you set $M_{1n}=M_{2n}=M_n$ in the general version of the least squares problem you arrive at the above formulation.

QED □