Supplement to "Estimating spillovers using panel data, with an application to the classroom"

(*Quantitative Economics*, 3, No. 3, November 2012, 421–470)

Peter Arcidiacono
Duke university

Gigi Foster
University of New South Wales

Natalie Goodpaster
Analysis Group, Inc.

Josh Kinsler
University of Rochester

Appendix 1: Proof of Theorem 1: General Case Size

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 following limiting case:

1. We observe students for at most two time periods.

2. Within each class, there is at most one student who is observed for two periods. All other students are observed for only one time period.

Remark S.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 but one of the individual effects in a class 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 but one student 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

Peter Arcidiacono: psarcidi@econ.duk.edu
Gigi Foster: gigi.foster@unsw.edu.au
Natalie Goodpaster: natalie.goodpaster@gmail.com
Josh Kinsler: joshua.kinsler@rochester.edu

DOI: 10.3982/QE145

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).

Consider the set of students who are observed for two time periods. Each of these students has M_{1n} peers in period 1 and M_{2n} peers in period 2. 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 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

$$\begin{aligned} \min_{\alpha,\gamma} \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} \Bigg[& \left(y_{11n} - \alpha_{1n} - \frac{\gamma}{M_{1n}} \sum_{j=2}^{M_{1n}+1} \alpha_{j1n} \right)^2 + \left(y_{12n} - \alpha_{1n} - \frac{\gamma}{M_{2n}} \sum_{j=2}^{M_{2n}+1} \alpha_{j2n} \right)^2 \\ & + \sum_{i=2}^{M_{1n}+1} \left(y_{i1n} - \alpha_{i1n} - \frac{\gamma}{M_{1n}} \sum_{j\neq i}^{M_{1n}+1} \alpha_{j1n} \right)^2 \\ & + \sum_{i=2}^{M_{2n}+1} \left(y_{i2n} - \alpha_{i2n} - \frac{\gamma}{M_{2n}} \sum_{j\neq i}^{M_{2n}+1} \alpha_{j2n} \right)^2 \Bigg]. \end{aligned} \tag{S1}$$

Within each block, there are four terms: two residuals for the student observed twice, and peer residuals in time periods 1 and 2.

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 supplement.

[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 6 to 10. 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.

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 subsequent 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 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

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

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

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

where $w \equiv (y, M)$. We let $\mathcal{W}$ denote the subset of $\mathbb{R}^{2+2\overline{M}} \times \mathcal{J}^2$ that represents 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. *We have*

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

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. *We have*

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

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

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

LEMMA 5.G. *We have*

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

where

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

and

$$B_o \equiv E\big[s(w, \gamma_o)^2\big] = \mathrm{Var}\big[s(w, \gamma_o)\big].$$

PROOF OF LEMMA 1.G. Our objective is to show that the system of equations obtained by differentiating equation (S1) 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

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

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

$$0 = -\frac{2\gamma}{M_{1n}}\left(y_{1tn} - \alpha_{1n} - \frac{\gamma}{M_{1n}}\sum_{j=2}^{M_{1n}+1}\alpha_{j1n}\right)$$
$$-\frac{2\gamma}{M_{1n}}\sum_{j=2,j\neq i}^{M_{1n}+1}\left(y_{j1n} - \alpha_{j1n} - \frac{\gamma}{M_{1n}}\sum_{k\neq j}^{M_{1n}+1}\alpha_{k1n}\right)$$
$$-2\left(y_{i1n} - \alpha_{i1n} - \frac{\gamma}{M_{1n}}\sum_{j\neq i}^{M_{1n}+1}\alpha_{j1n}\right).$$

The first-order condition for α_{i2n} is identical to the 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 rearrange 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 two equations

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

and

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

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 $\mathbf{X}_n\boldsymbol{\alpha}_n = \mathbf{Y}_n$, where $\boldsymbol{\alpha}_n$ is simply a $((1 + M_{1n} + M_{2n}) \times 1)$ vector of the individual student abilities in block n. Recall that because the student blocks are independent conditional on γ, we can solve for $\boldsymbol{\alpha}_n$ separately from $\boldsymbol{\alpha}_s$ for $s \neq n$. The form of $\mathbf{X}_n$, $\boldsymbol{\alpha}_n$, and $\mathbf{Y}_n$ are given by

$$\mathbf{X}_{n((1+M_{1n}+M_{2n})\times(1+M_{1n}+M_{2n}))} = \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix},$$
$$\boldsymbol{\alpha}_{n((1+M_{1n}+M_{2n})\times 1)} = [\alpha_{1n}, \alpha_{21n}, \ldots, \alpha_{(M_{1n}+1)1n}, \alpha_{22n}, \ldots, \alpha_{(M_{2n}+1)2n}]',$$

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

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

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

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

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

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

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

The solution to the system of equations for $\boldsymbol{\alpha}_n$ is now given by the simple expression

$$\boldsymbol{\alpha}_n = \mathbf{X}_n^{-1}\mathbf{Y}_n.$$

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

$$\mathbf{X}_n^{-1} = \begin{bmatrix} (A_n - B_nD_n^{-1}C_n)^{-1} & -(A_n - B_nD_n^{-1}C_n)^{-1}B_nD_n^{-1} \\ -D_n^{-1}C_n(A_n - B_nD_n^{-1}C_n)^{-1} & D_n^{-1} + D_n^{-1}C_n(A_n - B_nD_n^{-1}C_n)^{-1}B_nD_n^{-1} \end{bmatrix}. \tag{S2}$$

Since $(A_n - B_nD_n^{-1}C_n)^{-1}$ 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 $M_{tn} \times M_{tn}$ matrix composed of only two components. Thus, to get the form of D_n^{-1}, we just need to invert one of these $M_{tn} \times M_{tn}$ matrices.

At this point, it is useful to introduce some further notation 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 rewrite the components of $\mathbf{X}_n$ in the manner

$$\begin{aligned}
A_n &= a, \\
B_n &= [\underbrace{b_{1n}, \dots, b_{1n}}_{M_{1n}\text{ terms}}, \underbrace{b_{2n}, \dots, b_{2n}}_{M_{2n}\text{ terms}}], \\
C_n &= [\underbrace{b_{1n}, \dots, b_{1n}}_{M_{1n}\text{ terms}}, \underbrace{b_{2n}, \dots, b_{2n}}_{M_{2n}\text{ terms}}]',
\end{aligned}$$

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

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 $(M_{tn} \times M_{tn})$ 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 $M_{tn} \times M_{tn}$ matrix until we finally get to the point where we only have to invert a 2×2 matrix. Following this procedure, we can show that D_n^{-1} takes the 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 (S2). According to equation (S2), $\mathbf{X}_n^{-1}(1,1)$ 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,

$$\begin{aligned} B_n D_n^{-1} = \bigg[&\frac{b_{1n}(c_{1n} - b_{1n})}{d_{1n}}, \ldots, \frac{b_{1n}(c_{1n} - b_{1n})}{d_{1n}}, \\ &\frac{b_{2n}(c_{2n} - b_{2n})}{d_{2n}}, \ldots, \frac{b_{2n}(c_{2n} - b_{2n})}{d_{2n}} \bigg]. \end{aligned}$$

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,

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

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 (S2), $\mathbf{X}_n^{-1}(1,2)$ 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

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

The expression for $\mathbf{X}_n^{-1}(2,1)$, $-\tilde{A}_n D_n^{-1} C_n$, will be the transpose of the above equality, 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

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

The final component of $\mathbf{X}_n^{-1}$ is also the most complicated. The expression for $\mathbf{X}_n^{-1}(2,2)$ in equation (S2) is $D_n^{-1} + \tilde{A}_n D_n^{-1} C_n B_n D_n^{-1}$. Again we proceed in steps. Premultiplying $B_n D_n^{-1}$ by C_n yields an $((M_{1n} + M_{2n}) \times (M_{1n} + M_{2n}))$ 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 premultiply $C_n B_n D_n^{-1}$ by D_n^{-1}. This yields a symmetric $((M_{1n} + M_{2n}) \times (M_{1n} + M_{2n}))$

that takes the following form

$$\left[\begin{array}{ccc} \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \cdots \\ \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \frac{b_{1n}^2(c_{1n}-b_{1n})^2}{d_{1n}^2} & \cdots \\ \vdots & \vdots & \ddots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \cdots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \cdots \\ \vdots & \vdots & \ddots \end{array}\right.$$

$$\left.\begin{array}{ccc} \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \cdots \\ \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \frac{b_{1n}b_{2n}(c_{1n}-b_{1n})(c_{2n}-b_{2n})}{d_{1n}d_{2n}} & \cdots \\ \vdots & \vdots & \ddots \\ \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \cdots \\ \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \frac{b_{2n}^2(c_{2n}-b_{2n})^2}{d_{2n}^2} & \cdots \\ \vdots & \vdots & \ddots \end{array}\right].$$

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 $((M_{1n} + M_{2n}) \times (M_{1n} + M_{2n}))$ matrix that takes the form

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

where

$$\tilde{C}_{1n} = \frac{d_{1n}(c_{1n} + (M_{1n} - 1)b_{1n}) + \tilde{A}_n b_{1n}^2 (c_{1n} - b_{1n})^2}{d_{1n}^2},$$

$$\tilde{C}_{2n} = \frac{d_{2n}(c_{2n} + (M_{2n} - 1)b_{2n}) + \tilde{A}_n b_{2n}^2 (c_{2n} - b_{2n})^2}{d_{2n}^2},$$

$$\tilde{D}_{1n} = \frac{\tilde{A}_n b_{1n}^2 (c_{1n} - b_{1n})^2 - b_{1n} d_{1n}}{d_{1n}^2},$$

$$\tilde{D}_{2n} = \frac{\tilde{A}_n b_{2n}^2 (c_{2n} - b_{2n})^2 - b_{2n} d_{2n}}{d_{2n}^2},$$

$$\tilde{E}_n = \frac{\tilde{A}_n b_{1n} b_{2n} (c_{1n} - b_{1n})(c_{2n} - b_{2n})}{d_{1n} d_{2n}}.$$

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

Substituting into equation (S2) with the terms just calculated, we get the general form

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

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 $\mathbf{X}_n^{-1}$:

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

We can rearrange 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 $\tilde{B}_{2n}$

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

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 $\mathbf{X}_n^{-1}$.

$$\begin{aligned} \alpha_{21n} = {} & \tilde{B}_{1n} \left(y_{11n} + y_{12n} + \sum_{t=1}^{2} \left(\frac{\gamma}{M_{tn}} \sum_{j=2}^{M_{tn}+1} y_{jtn} \right) \right) \\ & + \tilde{C}_{1n} \left(y_{21n} + \frac{\gamma}{M_{1n}} y_{11n} + \frac{\gamma}{M_{1n}} \sum_{j=3}^{M_{1n}+1} y_{j1n} \right) \end{aligned}$$

$$+\tilde{D}_{1n}\left(\sum_{i=3}^{M_{1n}+1}\left(y_{i1n}+\frac{\gamma}{M_{1n}}y_{11n}+\frac{\gamma}{M_{1n}}\sum_{j=2,j\neq i}^{M_{1n}+1}y_{j1n}\right)\right)$$

$$+\tilde{E}_{n}\left(\sum_{i=2}^{M_{2n}+1}\left(y_{i2n}+\frac{\gamma}{M_{2n}}y_{12n}+\frac{\gamma}{M_{2n}}\sum_{j=2,j\neq i}^{M_{2n}+1}y_{j2n}\right)\right).$$

Again, we can rearrange the above equality, grouping on the y's:

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

The formula for α_{i1n} for $i>2$ takes the same form as above, except that (i) y_{21n} becomes y_{i1n} and (ii) the first summation on the second line is 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. □

Proof of Lemma 2.G. Lemma 1.G 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

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

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

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

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

The form of e_{12n} is identical to the above equality except all of the time subscripts on the y's, M's, and inverse components are 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

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

The residual e_{i1n} for $i > 2$ looks identical to the above equality except the leading y term is y_{i1n} rather than y_{21n}, and the summation term in the fourth line is over all $j \neq i$. The

M_{2n} residuals for the individuals observed once in the second period look identical to the above equality 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.

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. 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]:

$$\begin{aligned}
e_{11n} = {} & \big(\big(\gamma(M_{1n}-1)+M_{1n}\big)(\gamma - M_{2n})\big/\big((\gamma-M_2)^2\big(\big(M_1+\gamma(M_1-1)\big)^2+\gamma^2 M_1\big) \\
& + (\gamma - M_1)^2\big(\big(M_2+\gamma(M_2-1)\big)^2+\gamma^2 M_2\big)\big)\big) \\
& \times \Bigg(\big(\gamma(M_{1n}-1)+M_{1n}\big)(\gamma-M_{2n})y_{11n} \\
& - (\gamma - M_{1n})\big(\gamma(M_{2n}-1)+M_{2n}\big)y_{12n} \\
& - \gamma(\gamma-M_{2n})\sum_{j=2}^{M_{1n}+1} y_{j1n} + \gamma(\gamma-M_{1n})\sum_{j=2}^{M_{2n}+1} y_{j2n}\Bigg), \\
e_{12n} = {} & \big(\big(\gamma(M_{2n}-1)+M_{2n}\big)(\gamma - M_{1n})\big/\big((\gamma-M_2)^2\big(\big(M_1+\gamma(M_1-1)\big)^2+\gamma^2 M_1\big) \\
& + (\gamma - M_1)^2\big(\big(M_2+\gamma(M_2-1)\big)^2+\gamma^2 M_2\big)\big)\big) \\
& \times \Bigg(-\big(\gamma(M_{1n}-1)+M_{1n}\big)(\gamma-M_{2n})y_{11n} \\
& + (\gamma - M_{1n})\big(\gamma(M_{2n}-1)+M_{2n}\big)y_{12n} \\
& + \gamma(\gamma-M_{2n})\sum_{j=2}^{M_{1n}+1} y_{j1n} - \gamma(\gamma-M_{1n})\sum_{j=2}^{M_{2n}+1} y_{j2n}\Bigg), \\
e_{21n} = {} & \big(\gamma(\gamma - M_{2n})\big/\big((\gamma-M_2)^2\big(\big(M_1+\gamma(M_1-1)\big)^2+\gamma^2 M_1\big) \\
& + (\gamma - M_1)^2\big(\big(M_2+\gamma(M_2-1)\big)^2+\gamma^2 M_2\big)\big)\big) \\
& \times \Bigg(-\big(\gamma(M_{1n}-1)+M_{1n}\big)(\gamma-M_{2n})y_{11n} \\
& + (\gamma - M_{1n})\big(\gamma(M_{2n}-1)+M_{2n}\big)y_{12n} \\
& + \gamma(\gamma-M_{2n})\sum_{j=2}^{M_{1n}+1} y_{j1n} - \gamma(\gamma-M_{1n})\sum_{j=2}^{M_{2n}+1} y_{j2n}\Bigg),
\end{aligned}$$

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

$$
\begin{aligned}
e_{22n} = \big(&\gamma(\gamma - M_{1n}) \big/ \big((\gamma - M_2)^2\big(\big(M_1 + \gamma(M_1 - 1)\big)^2 + \gamma^2 M_1\big) \\
&+ (\gamma - M_1)^2\big(\big(M_2 + \gamma(M_2 - 1)\big)^2 + \gamma^2 M_2\big)\big)\big) \\
&\times \Bigg(\big(\gamma(M_{1n} - 1) + M_{1n}\big)(\gamma - M_{2n}) y_{11n} \\
&- (\gamma - M_{1n})\big(\gamma(M_{2n} - 1) + M_{2n}\big) y_{12n} \\
&- \gamma(\gamma - M_{2n}) \sum_{j=2}^{M_{1n}+1} y_{j1n} + \gamma(\gamma - M_{1n}) \sum_{j=2}^{M_{2n}+1} y_{j2n}\Bigg).
\end{aligned}
$$

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

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

$$
\begin{aligned}
e_{i1n} &= -e_{11n} \frac{\gamma}{M_{1n} + \gamma(M_{1n} - 1)}, \\
e_{i2n} &= -e_{i1n} \frac{\gamma - M_{1n}}{\gamma - M_{2n}}, \\
e_{12n} &= -e_{11n} \left(\frac{\gamma - M_{1n}}{\gamma - M_{2n}}\right)\left(\frac{M_{2n} + \gamma(M_{2n} - 1)}{M_{1n} + \gamma(M_{1n} - 1)}\right).
\end{aligned}
\tag{S3}
$$

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

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

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

Proof of Lemma 3.G. Recall that $q(w, \gamma)$ is given by

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

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

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

Collecting the α_{jto} terms yields

$$q(w,\gamma)=\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}\left(W_{jt}+\frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}W_{kt}\right)\alpha_{jto}+W_{jt}\varepsilon_{jt}\right)^2}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2}. \tag{S4}$$

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

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

Because of the symmetry, after multiplying out, any terms involving M_1M_2 will drop out as will any terms where neither M_1 nor M_2 enter. The expression then reduces to

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

Now consider the coefficient on α_{j1o} for $j>1$, which can be split into three components: 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 (S4) yields

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

We next take expectations conditional on M_1 and M_2:

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

Expanding the square and noting that $E(\alpha_{jto}\varepsilon_{kt'}) = 0$ for all j, k, t, t' by Theorem 1(ii) and $E(\varepsilon_{jt}\varepsilon_{kt'}) = 0$ for all $j \neq k$ or $t \neq t'$ by Theorem 1(i) yields

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

which can be rewritten as

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

Note that Theorem 1(v) implies that the conditioning in the expectation over the squared errors is not needed. Furthermore, $E(\varepsilon_{jt}^2)=E(\varepsilon_{kt}^2)$ for all j,k by Theorem 1(v). We can then express the expectation over the squared errors solely as a function of the first observation's squared error:

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

Note that the weights in the numerator of the second expectation are the same weights as in the denominator. Theorem 1(v) implies that these weights are orthogonal to the squared first and second period errors. Furthermore, $E(\varepsilon_{11}^2)=E(\varepsilon_{12}^2)$. Taking the unconditional expectation then yields

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

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 $\gamma\neq\gamma_o$. □

Proof of Lemma 4.G. Uniform convergence requires that

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

Theorem 12.1 in Wooldridge states four conditions that the data and q must satisfy so that the above condition holds.

1. *The parameter space Γ is compact.* This condition is satisfied by Theorem 1(vi).

2. *For each $\gamma\in\Gamma$, $q(\cdot,\gamma)$ is Borel measurable on $\mathcal{W}$.* The function $q(\cdot,\gamma)$ 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. *For all $\gamma\in\Gamma$, $|q(w,\gamma)|\le b(w)$, 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 $W_{jt}^2y_{jt}^2+W_{kt'}^2y_{kt'}^2\ge 2W_{jt}W_{kt'}y_{jt}y_{kt'}$ for all j,k,t,t' (the triangle inequality), we have

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

where 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

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

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 and M_2, and then show that this guarantees that the outer expectation is finite. With the data generating process substituting for y into the inner expectation yields

$$\begin{aligned}
&E\big[b(w)|M_1, M_2\big] \\
&\quad = (2 + M_1 + M_2)E\left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}\left(\alpha_{jto} + \frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}\alpha_{kto} + \varepsilon_{jt}\right)^2 \Bigg| M_1, M_2\right].
\end{aligned}$$

Repeatedly using the triangle inequality after expanding the square implies

$$\begin{aligned}
&E\big[b(w)|M_1, M_2\big] \\
&\quad \leq (2 + M_1 + M_2)E\left[\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(M_t + 2)\left(\alpha_{jto}^2 + \frac{\gamma_o^2}{M_t^2}\sum_{k\neq j}^{M_t+1}\alpha_{kto}^2 + \varepsilon_{jt}^2\right)\Bigg| M_1, M_2\right].
\end{aligned}$$

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

$$\begin{aligned}
&E\big[b(w)|M_1, M_2\big] \\
&\quad \leq (2 + M_1 + M_2)E\left[\big(1 + \gamma_o^2\big)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(M_t + 2)\big(\alpha_{jto}^2 + \varepsilon_{jt}^2\big)\Big| M_1, M_2\right].
\end{aligned}$$

We can take the expectation operator through, which yields

$$\begin{aligned}
&E\big[b(w)|M_1, M_2\big] \\
&\quad \leq (2 + M_1 + M_2)\big(1 + \gamma_o^2\big)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(M_t + 2)\big[E\big(\alpha_{jto}^2|M_1, M_2\big) + E\big(\varepsilon_{jt}^2\big)\big],
\end{aligned}$$

where the conditioning is not necessary for the second expectation by Theorem 1(v). Theorem 1(iii), (iv), and (vi) ensures that $E(\alpha_{jto}^2|M_1, M_2)$, $E(\varepsilon_{jt}^2)$, 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 Theorem 1(i). □

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. *The parameter* γ_o *must be in the interior of* Γ. This condition is satisfied by Theorem 1(vi).

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

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

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

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

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

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

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

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

$$+\frac{\left|\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1} W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\right|\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(W'_{jt}W'_{jt}+W''_{jt}W_{jt})\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right)^2}$$

$$+\frac{\left|\left(\sum_{t=1}^{2}\sum_{t'=1}^{2}\sum_{j=1}^{M_t+1}\sum_{k=1}^{M_{t'}+1} W_{jt}W_{kt'}y_{jt}y_{kt'}\right)\right|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right)^3}.$$

Repeatedly applying the triangle inequality and collecting terms yields:

$$\left|H(w,\gamma)\right| \le \frac{2(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}\left((W''_{jt})^2+(W'_{jt})^2+W_{jt}^2\right)y_{jt}^2\right)}{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2}$$

$$+\frac{4(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}\left((W'_{jt})^2+W_{jt}^2\right)y_{jt}^2\right)\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right)^2}$$

$$+\frac{(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left|\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}(W'_{jt}W'_{jt}+W''_{jt}W_{jt})\right)\right|}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right)^2}$$

$$+\frac{(2+M_1+M_2)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W'_{jt}W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} W_{jt}^2\right)^3}.$$

Denote the weight given to y_{jt}^2 in the above expression as

$$W_{jt}^* = \frac{2(2+M_1+M_2)[(W''_{jt})^2+(W'_{jt})^2+W_{jt}^2]}{\sum_{t=1}^{2}\sum_{k=1}^{M_t+1} W_{kt}^2}$$

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

$$+\frac{(2+M_1+M_2)W_{jt}^2\left|\left(2\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}(W_{kt}'W_{kt}'+W_{kt}''W_{kt})\right)\right|}{\left(\sum_{t=1}^{2}\sum_{k=1}^{M_t+1}W_{kt}^2\right)^2}$$

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

implying that

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

Note that W_{jt}^* is a 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

$$B_{jt}^*=\max_{\gamma}W_{jt}^*,$$

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 $E[b(w)|M_1,M_2]<\infty$:

$$\begin{aligned}E\big[b(w)|M_1,M_2\big]&=\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}B_{jt}^*E\big(y_{jt}^2|M_1,M_2\big)\\&=\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}B_{jt}^*E\left[\left(\alpha_{jto}+\frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}\alpha_{kto}+\varepsilon_{jt}\right)^2\middle|M_1,M_2\right].\end{aligned}$$

Repeatedly using the triangle inequality after expanding the square implies:

$$E\big[b(w)|M_1,M_2\big] \leq \sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt}(M_t+2)E\left[\left(\alpha^2_{jto}+\frac{\gamma_o^2}{M_t^2}\sum_{k\neq j}^{M_t+1}\alpha^2_{kto}+\varepsilon^2_{jt}\right)\Big|M_1,M_2\right].$$

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

$$E\big[b(w)|M_1,M_2\big] \leq \big(1+\gamma_o^2\big)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1} B^*_{jt}(M_t+2)E\big[\big(\alpha^2_{jto}+\varepsilon^2_{jt}\big)|M_1,M_2\big].$$

Theorem 1(vi), (iii), (iv), and (vi) ensure that B^*, $E(\alpha^2_{jto}|M_1,M_2)$, $E(\varepsilon^2_{jt})$, 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 the fact that M_1 and M_2 are finite.

3. *The function* $s(w,\cdot)$ *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. *The equality* $A_o \equiv E[H(w,\gamma_o)]$ *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

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

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 are multiplied by either $(\gamma-\gamma_o)^2$ 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. *We have* $E[s(w,\gamma_o)]=0$. Note that $E[s(w,\gamma)] = \partial E[q(w,\gamma)]/\partial\gamma$. Differentiating with respect to γ leaves terms that are multiplied by $(\gamma_0-\gamma)$ or by $(\gamma_0-\gamma)^2$, 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 and M_2 which is given by

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

Applying the triangle inequality and collecting terms yields

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

Repeatedly applying the triangle inequality, we can write

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

$$+E\left(2\left[\frac{\left((2+M_1+M_2)\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)\left(2\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}'W_{jt}\right)}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right]^2\middle|M_1,M_2\right)$$

$$\leq E\left(8(2+M_1+M_2)^2\left[\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W_{jt}')^2+W_{jt}^2)y_{jt}^2\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right]\middle|M_1,M_2\right)$$

$$+E\left(8(2+M_1+M_2)^2\left[\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2y_{jt}^2\right)^2\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}'W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^4}\right]\middle|M_1,M_2\right)$$

$$\leq E\left(8(2+M_1+M_2)^3\left[\frac{\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}((W_{jt}')^2+W_{jt}^2)^2y_{jt}^4}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^2}\right]\middle|M_1,M_2\right)$$

$$+E\left(8(2+M_1+M_2)^3\left[\frac{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^4y_{jt}^4\right)\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}'W_{jt}\right)^2}{\left(\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W_{jt}^2\right)^4}\right]\middle|M_1,M_2\right).$$

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

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

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

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

$$E\big[s(w,\gamma)^2|M_1,M_2\big]\le\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W^*_{jt}E\big(y^4_{jt}|M_1,M_2\big).$$

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

$$\begin{aligned}
&E\big[s(w,\gamma)^2|M_1,M_2\big]\\
&\quad\le\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W^*_{jt}E\Bigg[\Bigg(\alpha_{jto}+\frac{\gamma_o}{M_t}\sum_{k\neq j}^{M_t+1}\alpha_{kto}+\varepsilon_{jt}\Bigg)^4\Bigg|M_1,M_2\Bigg]\\
&\quad\le\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W^*_{jt}E\Bigg[(M_t+2)^2\Bigg(\alpha^2_{jto}+\bigg(\frac{\gamma_o}{M_t}\bigg)^2\sum_{k\neq j}^{M_t+1}\alpha^2_{kto}+\varepsilon^2_{jt}\Bigg)^2\Bigg|M_1,M_2\Bigg]\\
&\quad\le\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W^*_{jt}E\Bigg[(M_t+2)^3\Bigg(\alpha^4_{jto}+\bigg(\frac{\gamma_o}{M_t}\bigg)^4\sum_{k\neq j}^{M_t+1}\alpha^4_{kto}+\varepsilon^4_{jt}\Bigg)\Bigg|M_1,M_2\Bigg].
\end{aligned}$$

Collecting terms, we have

$$E\big[s(w,\gamma)^2|M_1,M_2\big]\le\sum_{t=1}^{2}\sum_{j=1}^{M_t+1}W^*_{jt}\bigg[(M_t+2)^3\bigg(1+\frac{\gamma_o^4}{M_t^3}\bigg)E\big(\alpha^4_{jto}|M_1,M_2\big)+E\big(\varepsilon^4_{jt}\big)\bigg],$$

where W^*_{jt}, γ_o, and M_t are all finite, and since the fourth moments of α and ε's are finite by Theorem 1(iii) and (iv), the expression is finite for all $\gamma\in\Gamma$ and for all M_t. Furthermore, $E[s(w,\gamma_o)^2]\le\max_{M_1,M_2}E[s(w,\gamma_o)^2|M_1,M_2]<\infty$. □

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

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 starting 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

$$\alpha_{1n}=(\tilde A_n+\hat\gamma\tilde B_n)(y_{11n}+y_{12n})+\bigg(\tilde A_n\frac{\hat\gamma}{M_n}+\tilde B_n\frac{\hat\gamma(M_n-1)+M_n}{M_n}\bigg)\sum_{t=1}^{2}\sum_{j=2}^{M_n+1}y_{jtn},$$

$$\begin{aligned}
\alpha_{i1n}={}&\bigg(\tilde B_n+\tilde C_n\frac{\hat\gamma}{M_n}+\tilde D_n\frac{(M_n-1)\hat\gamma}{M_n}\bigg)y_{11n}+(\tilde B_n+\tilde E_n\hat\gamma)y_{12n}\\
&+\bigg(\tilde B_n\frac{\hat\gamma}{M_n}+\tilde C_n+\tilde D_n\frac{(M_n-1)\hat\gamma}{M_n}\bigg)y_{i1n}
\end{aligned}$$

$$+ \left(\tilde{B}_n \frac{\hat{\gamma}}{M_n} + \tilde{C}_n \frac{\hat{\gamma}}{M_n} + \tilde{D}_n \frac{M_n + (M_n - 2)\hat{\gamma}}{M_n} \right) \sum_{j=2, j\neq i}^{M_n+1} y_{j1n}$$

$$+ \left(\tilde{B}_n \frac{\hat{\gamma}}{M_n} + \tilde{E}_n \frac{M_n + (M_n - 1)\hat{\gamma}}{M_n} \right) \sum_{j=2}^{M_n+1} y_{j2n},$$

where

$$\tilde{A}_n = \frac{\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2}{2(1+\gamma)^2(\gamma - M_n)^2}, \tag{S5}$$

$$\tilde{B}_n = \frac{\gamma^2 - \gamma(2+\gamma)M_n}{2(1+\gamma)^2(\gamma - M_n)^2}, \tag{S6}$$

$$\begin{aligned} \tilde{C}_n = \big(&\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2\big(8 + \gamma(12+5\gamma)\big)M_n^2 \\ &- 4\gamma(1+\gamma)^2(2+\gamma)M_n^3 + 2(1+\gamma)^4 M_n^4\big) \\ &\big/\big(2(1+\gamma)^2(\gamma - M_n)^2\big(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2\big)\big), \end{aligned} \tag{S7}$$

$$\tilde{D}_n = \frac{\gamma^4 - 2\gamma^3(2+\gamma)M_n + \gamma^2(6+\gamma(8+3\gamma))M_n^2 - 2\gamma(1+\gamma)^2(2+\gamma)M_n^3}{2(1+\gamma)^2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}, \tag{S8}$$

$$\tilde{E}_n = \frac{\gamma^2(\gamma(M_n - 1) + 2M_n)^2}{2(1+\gamma)^2(\gamma - M_n)^2(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)}. \tag{S9}$$

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

$$\begin{aligned} e_{11n} = y_{11n}&\left(1 - \tilde{A}_n - 2\gamma\tilde{B}_n - \tilde{C}_n \frac{\gamma^2}{M_n} - \tilde{D}_n \frac{\gamma^2(M_n - 1)}{M_n}\right) \\ &- y_{12n}\big(\tilde{A}_n + 2\gamma\tilde{B}_n + \gamma^2\tilde{E}_n\big) \\ &- \left(\sum_{j=2}^{M_n+1} y_{j1n}\right)\left(\tilde{A}_n \frac{\gamma}{M_n} + \tilde{B}_n \frac{\gamma(\gamma + M_n - 1) + M_n}{M_n}\right. \\ &\left. + \tilde{C}_n \frac{\gamma M_n(1+\gamma) - \gamma^2}{M_n^2} + \tilde{D}_n \frac{\gamma(\gamma + M_n(M_n + M_n\gamma - 2\gamma - 1))}{M_n^2}\right) \end{aligned} \tag{S10}$$

$$
\begin{aligned}
&-\left(\sum_{j=2}^{M_n+1} y_{j2n}\right)\\
&\times\left(\tilde{A}_n\frac{\gamma}{M_n}+\tilde{B}_n\frac{\gamma(\gamma+M_n-1)+M_n}{M_n}+\tilde{E}_n\frac{\gamma M_n(1+\gamma)-\gamma^2}{M_n}\right).
\end{aligned}
$$

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*

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

where

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

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

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

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 $\tilde{B}_n$:

$$
\begin{aligned}
\tilde{C}_n=&\big(\gamma^4-2\gamma^3(2+\gamma)M_n+\gamma^2\big(8+\gamma(12+5\gamma)\big)M_n^2\\
&-4\gamma(1+\gamma)^2(2+\gamma)M_n^3+2(1+\gamma)^4M_n^4\big)\\
&\big/\big(2(1+\gamma)^2(\gamma-M_n)^2\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\big),\\
\tilde{D}_n=&\frac{\gamma^4-2\gamma^3(2+\gamma)M_n+\gamma^2(6+\gamma(8+3\gamma))M_n^2-2\gamma(1+\gamma)^2(2+\gamma)M_n^3}{2(1+\gamma)^2(\gamma-M_n)^2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}.
\end{aligned}
$$

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 that is present in the denominator of $\tilde{D}_n$. Thus we can write the difference as

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

Combining terms and simplifying yields

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

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.

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

□

Using Property 1, we now show that the coefficients on the observed grades in equation (S10) have other appealing properties. Then we use these properties to simplify equation (S10), in an effort to arrive at a simplified version of the least squares problem as a function of γ.

PROPERTY 2. *Equation* (S10) *describes the prediction error for the first outcome of the individual observed twice in block* n. *In Equation* (S10), *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 to the coefficient on* $\sum_{j=2}^{M_n+1} y_{j2n}$.

PROOF. 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

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

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 rewrite 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

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

where the second to last line results from substituting in our formula for $\tilde{B}_n$ given in equation (S6).

Now we show that the coefficient on $\sum_{j=2}^{M_n+1} y_{j1n}$ is equal in magnitude but to the opposite sign to 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

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

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 expression

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

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 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 (S6).

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

All of the terms in the above expression cancel out, proving that the sums 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. □

Now we return to equation (S10), which describes the prediction error for the first observation of the student observed twice in block n. Using Properties 1 and 2, we 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 yields 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

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

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 manner

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

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 (S6) and V_n from Property 1 yields

$$\begin{aligned}&\big(\big(\gamma\big(\gamma-(2+\gamma)M_n\big)+M_n^2\big)\big(\gamma^2-\gamma(2+\gamma)M_n\\&\quad+(1+\gamma)^2M_n^2\big)-\gamma^2M_n^2(\gamma^2-2\gamma M_n-\gamma^2M_n)\big)\\&\quad/\big(2(\gamma-M_n)^2\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\big).\end{aligned}$$

We can rearrange this expression in the manner

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

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

$$\begin{aligned}&\big((\gamma^2-2\gamma M_n+M_n^2)\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\\&\quad-\gamma^2M_n(\gamma^2-2\gamma M_n+M_n^2)\big)/\big(2(\gamma-M_n)^2\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\big).\end{aligned}$$

The numerator then factors to produce

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

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

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

This gives us the coefficients 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 manner

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

Substituting for $\tilde{B}_n$ from equation (S6) and rearranging yields

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

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

$$\begin{aligned}
&\big((\gamma^2-\gamma M_n-\gamma^2 M_n)\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\\
&\quad -\gamma M_n(\gamma-2M_n-\gamma M_n)\big(\gamma M_n(1+\gamma)-\gamma^2\big)\big)\\
&\quad /\big(2(\gamma-M_n)^2\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\big).
\end{aligned}$$

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

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

Canceling out the common terms in the numerator and denominator yields

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

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 (S10).

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

This simplifies further to produce

$$\begin{aligned}
e_{11n} &= \frac{(M_n+\gamma(M_n-1))}{2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)}\\
&\quad \times\left(\big(M_n+\gamma(M_n-1)\big)(y_{11n}-y_{12n})+\gamma\sum_{j=2}^{M_n+1}(y_{j2n}-y_{j1n})\right).
\end{aligned}$$

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}:

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

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

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

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

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 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}. Then the formula for e_{21n} from the basic least squares problem can be written as

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

Substituting in for α from Lemma 1.G and combining like terms yields

$$
\begin{aligned}
e_{21n} = {}& y_{21n}\left(1 - \tilde{A}_n \frac{\gamma^2}{M_n^2} - 2\tilde{B}_n \frac{M_n\gamma+\gamma^2(M_n-1)}{M_n^2} - \tilde{C}_n \frac{M_n^2+\gamma^2(M_n-1)}{M_n^2}\right. \\
&\left. - \tilde{D}_n \frac{2\gamma M_n(M_n-1)+\gamma^2(M_n-1)(M_n-2)}{M_n^2}\right) \\
&- y_{11n}\left(\tilde{A}_n \frac{\gamma}{M_n} + \tilde{B}_n \frac{M_n+(M_n-1)\gamma+\gamma^2}{M_n} + \tilde{C}_n \frac{\gamma M_n+(M_n-1)\gamma^2}{M_n^2}\right. \\
&\left. + \tilde{D}_n \frac{(M_n-1)\gamma(M_n+(M_n-1)\gamma)}{M_n^2}\right) \\
&- y_{12n}\left(\tilde{A}_n \frac{\gamma}{M_n} + \tilde{B}_n \frac{M_n+(M_n-1)\gamma+\gamma^2}{M_n}\right. \\
&\left. + \tilde{E}_n \frac{\gamma M_n+(M_n-1)\gamma^2}{M_n}\right) \\
&- \left(\sum_{j=3}^{M_n+1} y_{j1n}\right)\left(\tilde{A}_n \frac{\gamma^2}{M_n^2} + 2\tilde{B}_n \frac{\gamma M_n+\gamma^2(M_n-1)}{M_n^2}\right. \\
&\left. + \tilde{C}_n \frac{2\gamma M_n+\gamma^2(M_n-2)}{M_n^2} + \tilde{D}_n \frac{(M_n+\gamma(M_n-2))^2+(M_n-1)\gamma^2}{M_n^2}\right)
\end{aligned} \tag{S11}
$$

$$- \left(\sum_{j=2}^{M_n+1} y_{j2n} \right)$$
$$\times \left(\tilde{A}_n \frac{\gamma^2}{M_n^2} + 2\tilde{B}_n \frac{\gamma M_n + \gamma^2 (M_n - 1)}{M_n^2} + \tilde{E}_n \frac{(M_n + (M_n - 1)\gamma)^2}{M_n^2} \right).$$

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

PROPERTY 3. *The coefficients on y_{11n} and y_{12n} in the equation for e_{21n} are equal in magnitude but of 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. 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

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

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 rewrite the coefficient on $\sum_{j=3}^{M_n+1} y_{j1n}$ in the manner

$$\begin{aligned}
&\tilde{A}_n \frac{\gamma^2}{M_n^2} + 2\tilde{B}_n \frac{\gamma M_n + \gamma^2 (M_n - 1)}{M_n^2} + \tilde{C}_n \frac{2\gamma M_n + \gamma^2 (M_n - 2)}{M_n^2} \\
&\quad + \tilde{D}_n \frac{(M_n + \gamma(M_n - 2))^2 + (M_n - 1)\gamma^2}{M_n^2} \\
&\quad = \tilde{B}_n (1+\gamma)^2 + \frac{V_n(\gamma^2 - 2\gamma M_n (1+\gamma) + M_n^2 (1+\gamma)^2)}{2M_n^2} + \frac{2\gamma M_n (2+\gamma) - 3\gamma^2}{2(\gamma - M_n)^2}.
\end{aligned}$$

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

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

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 (S6) yields

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

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

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

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 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 opposite sign. □

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

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

Finding a common denominator and rearranging yields

$$\begin{aligned}
&\tilde{B}_n(1+\gamma)^2 + \big(\gamma^2\big((\gamma-M_n)^2+\gamma M_n(2M_n-\gamma+\gamma M_n)\big) \\
&\quad -\gamma(\gamma M_n+2M_n-\gamma)\big((\gamma-M_n)^2+\gamma M_n(2M_n+\gamma M_n-2\gamma)\big)\big) \\
&\quad \big/\big(2(\gamma-M_n)^2\big(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2\big)\big).
\end{aligned}$$

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

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

The above expression simplifies further to

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

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:

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

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

$$\begin{aligned}
e_{21n} = {}& \frac{\gamma}{2(\gamma^2-\gamma(2+\gamma)M_n+(1+\gamma)^2M_n^2)} \\
&\times\left(\big(M_n+\gamma(M_n-1)\big)(y_{11n}-y_{12n})+\gamma\sum_{j=2}^{M_n+1}(y_{j2n}-y_{j1n})\right).
\end{aligned}$$

Finally, we have all the components of the least squares problem strictly as functions of y, γ, and M_n. Rewriting 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 facts that $e_{12n} = -e_{11n}$, $e_{j1n} = e_{21n}$ for $i > 3$, and $e_{22n} = -e_{21n}$, we can simplify the above expression to

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

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

$$\begin{aligned}\min_{\gamma} 2 \sum_{n=1}^{N} \Bigg[& \frac{(M_n + \gamma(M_n - 1))^2}{4(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)^2} \\ & \times \left(\big(M_n + \gamma(M_n - 1)\big)(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n})\right)^2 \\ & + \frac{\gamma^2 M_n}{4(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)^2} \\ & \times \left(\big(M_n + \gamma(M_n - 1)\big)(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n})\right)^2 \Bigg].\end{aligned}$$

Notice that the terms inside the square brackets are exactly the same. We can rearrange the above expression by combining like terms:

$$\begin{aligned}\min_{\gamma} 2 \sum_{n=1}^{N} \Bigg[& \frac{(M_n + \gamma(M_n - 1))^2 + \gamma^2 M_n}{4(\gamma^2 - \gamma(2+\gamma)M_n + (1+\gamma)^2 M_n^2)^2} \\ \times & \left(\big(M_n + \gamma(M_n - 1)\big)(y_{11n} - y_{12n}) + \gamma \sum_{j=2}^{M_n+1} (y_{j2n} - y_{j1n})\right)^2 \Bigg].\end{aligned}$$

Simplifying the leading term leaves us with the least squares problem

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

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.

Submitted March, 2011. Final version accepted March, 2012.