# Supplement to “A Simple and Computationally Trivial Estimator for Grouped Fixed Effects Models”

Martin Mugnier*

**Abstract**

This supplement contains additional material to accompany the main text. First, I show that some results can be extended to allow for an increasing number of groups. Next, I discuss estimation in models with individual-specific and/or time effects. I report results from Monte Carlo simulations of a simple version of the model. Finally, I provide a short illustration to Acemoglu, Johnson, Robinson, and Yared (2008)'s panel data set of countries.

*CREST, ENSAE, Institut Polytechnique de Paris, martin.mugnier@ensae.fr

# S1 Increasing Number of Groups

Consider a simpler version of the model without covariates ($\beta^0 = 0$ is known). Proposition S1.1 shows that uniformly consistent classification of cross-sectional units remains possible, even when there are "almost" as many groups as individuals asymptotically.

**Proposition S1.1** *Suppose* $\beta^0 = 1$, *Assumption 1 holds,* $G^0 = G^0_{NT} \leq N/2$, *and* $\lambda_2 = o(1)$ *such that* $\lambda_2 \gtrsim T^{-\kappa}$ *for some* $0 < \kappa < 1/2$ *as* $T \to \infty$. *Let* $\widehat{\beta}^1(\lambda_1) = 0$. *Then, as* $N$ *and* $T$ *tend to infinity,*

$$\left\| \widehat{W} - W^0 \right\|_{\max} = o_p(1) + o_p(N^2 T^{-\delta}) \text{ for all } \delta > 0.$$

**Proof of Proposition S1.1 :** Similar arguments as in the proof of Proposition 3.1 show that the probability of false-positive for the test $H_{0,ij} : g^0_i = g^0_j$ based on the statistics $1 - \widehat{W}_{ij} = \mathbf{1}\left\{\widehat{d}^2_\infty(i,j) > \lambda_2\right\}$ is bounded asymptotically by $o_p(NT^{-\delta})$ uniformly across pairs of units $(i,j) \in \{1, ..., N\}^2$. Similarly, outside of an event whose probability tends to zero (when there are at least two units in each group), the probability of false-negative can be shown to be bounded asymptotically as, for all $\delta > 0$,

$$G^0_{NT}(G^0_{NT} - 1) o_p(T^{-\delta}) = o_p(N^2 T^{-\delta}).$$

□

# S2 Extension of the Baseline Model

## S2.1 Individual and Time-Specific Unobserved Heterogeneity

The original model (1.1) can be augmented with unit-specific and time-specific efects $\nu^0_i$ and $\xi^0_t$:

$$y_{it} = x'_{it}\theta^0 + \nu^0_i + \xi^0_t + \alpha^0_{g^0_i t} + v_{it}. \tag{S2.1}$$

Without covariates, the two-step estimator could still be applied to within/across differenced versions of equation (S2.1). With covariates, one needs strong factors to apply Moon and Weidner (2019). Then, one can differentiate to eliminate $\nu_i^0$ and $\xi_t^0$ to estimate the groups. The last OLS linear regression step must be augmented with time and individual dummies. From a theoretical point of view, one needs to redefine separation assumptions as in Bonhomme and Manresa (2015)'s Supplemental Material but the overall methodology still applies. Allowing for group-specific slope is theoretically possible if one has access to a preliminary consistent estimator, as well as allowing for grouping in the longitudinal $t$ dimension.

### S2.2 Computational Gains in the Time-Invariant Model

If the clustered unobserved heterogeneity is time-invariant, one can relax Assumption 2(b) to Bonhomme and Manresa (2015)'s Assumption 2(e). Also, it is recommended to use a simple within, first-difference, or instrumental differencing estimator as a first step in this case (see, e.g., Arellano and Bonhomme, 2011; Wooldridge, 2010). The method becomes $O(N^2)$ without changing the asymptotic properties by considering the new distance

$$\widehat{d}^2_{\infty,\text{time-inv}}(i,j) = \left(\frac{1}{T}\sum_{t=1}^{T}(\widehat{v}_{it} - \widehat{v}_{jt})\right)^2. \tag{S2.2}$$

I call the estimator based upon $\widehat{d}^2_{\infty,\text{time-inv}}(i,j)$ the pairwise-distance (PWD) estimator.

## S3 Monte Carlo Simulations

In this section, I investigate (i) the finite sample performance of the PWD estimator in the correctly specified model:

$$y_{it} = \alpha^0_{g_i^0} + v_{it}, \tag{S3.1}$$

and (ii) finite sample sensitivity to the choice of the thresholding parameter $c_T \equiv \lambda_2$. First, I assess consistency of the PWD estimator for $c_T = 2\log(T)/\sqrt{T}$.[1] For each $(G^0, N) \in \{2, 5, 10, 50\} \times \{50, 100, 200, 500\}$ and $T$ in a linearly equally spaced grid of 4 elements from $\lceil\sqrt{N}\rceil$ to $N$, I draw $1,000$ Monte Carlo samples from model (S3.1), in

[1] I study sensitivity to this choice later.

which $(\alpha_1^0, ..., \alpha_{G^0}^0)' = \text{linspace}(-G^0/2, G^0/2, G^0)$ and $(g_1^0, ..., g_N^0) = (1, ..., 1, ..., G^0, ..., G^0)$ are deterministic and each group has equal size $N/G^0$. I consider three DGPs for the noise random variable $v_{it}$, summarized in Table 1. Tables 2-4 report Monte Carlo

Table 1: Data Generating Processes

| DGP | Noise |
|---|---|
| 1 | $v_{it} \sim \mathcal{N}(0,1)$ i.i.d. across $i$ and $t$. |
| 2 | $v_{it} = 0.5v_{it-1} + \xi_{it}$ with $\xi_{it} \sim \mathcal{N}(0,1)$ i.i.d. across $i$ and $t$, independent of $v_{i0} \sim \mathcal{N}(0,1)$. |
| 3 | $\sigma_i \sim \mathcal{U}[0.5, 1.5]$ and $v_{it} \mid \sigma_i \sim \mathcal{N}(0, \sigma_i)$ independent across $i$ and i.i.d. across $t$ for a given $i$. |

averages of the estimated number of groups $\widehat{G}^{\text{PWD}}$, the Hausdorff distance between estimated effects $(\widehat{\alpha}_1^{\text{PWD}}, ..., \widehat{\alpha}_{\widehat{G}^{\text{PWD}}}^{\text{PWD}})'$ and true effects $(\alpha_1^0, ..., \alpha_{G^0}^0)'$, Rand Index (RI), and CPU time. RI is the proportion of correctly predicted pair (true or false) returned by the PWD estimator:

$$\text{RI} = \frac{TP + TN}{TP + TN + FP + FN},$$

where

$$\begin{aligned}
TP &\equiv \text{True Positives} := \sum_{i<j} \mathbf{1}\left\{\widehat{g}_i^{\text{PWD}} = \widehat{g}_j^{\text{PWD}}\right\} \mathbf{1}\left\{g_i^0 = g_j^0\right\}, \\
TN &\equiv \text{True Negatives} := \sum_{i<j} \mathbf{1}\left\{\widehat{g}_i^{\text{PWD}} \neq \widehat{g}_j^{\text{PWD}}\right\} \mathbf{1}\left\{g_i^0 \neq g_j^0\right\}, \\
FP &\equiv \text{False Positives} := \sum_{i<j} \mathbf{1}\left\{\widehat{g}_i^{\text{PWD}} = \widehat{g}_j^{\text{PWD}}\right\} \mathbf{1}\left\{g_i^0 \neq g_j^0\right\}, \\
FN &\equiv \text{False Negatives} := \sum_{i<j} \mathbf{1}\left\{\widehat{g}_i^{\text{PWD}} \neq \widehat{g}_j^{\text{PWD}}\right\} \mathbf{1}\left\{g_i^0 = g_j^0\right\}.
\end{aligned}$$

Results suggest good finite sample performance, although deteriorating with the degree of time-dependence of the idiosyncratic shocks. In the most favorable case of independent normal errors (Tables 2 and 4), it is remarkable how perfect or almost perfect classification is achieved for moderate sample sizes and even for a large number of groups (e.g., for $(N, T, G^0) = (50, 36, 2)$ or $(N, T, G^0) = (500, 500, G^0 = 50)$).

Second, I investigate the finite sample sensitivity of the results to the choice of the thresholding parameter $c_T$. I consider DGP 1 only, fix $N = 120$, and let $(G^0, T) \in \{2, 3, 4\} \times \{11, 66, 120\}$. Figures 2-6 plot Monte Carlo averages of $\widehat{G}^{\text{PWD}}$, HD, RI, Precision (P) and Recall (R) rates as functions of $c \in \text{linspace}(0.1, 20, 40)$ with $c_T = c\log(T)/\sqrt{T}$, where each colored line corresponds to $\sigma \in \{0.25, 0.5, 1\}$, where $\sigma$ is the standard-deviation of the random noise $v_{it}$. The Recall rate (R) measures the ability of the PWD estimator to identify pairs that truly belong to the same group. The Precision rate (P) measures how precise the pairing prediction is: among all predicted pairs of units, what is the proportion of correct ones? Both formally write

$$\text{R} = \frac{TP}{TP + FN}, \quad \text{P} = \frac{TP}{TP + FP}.$$

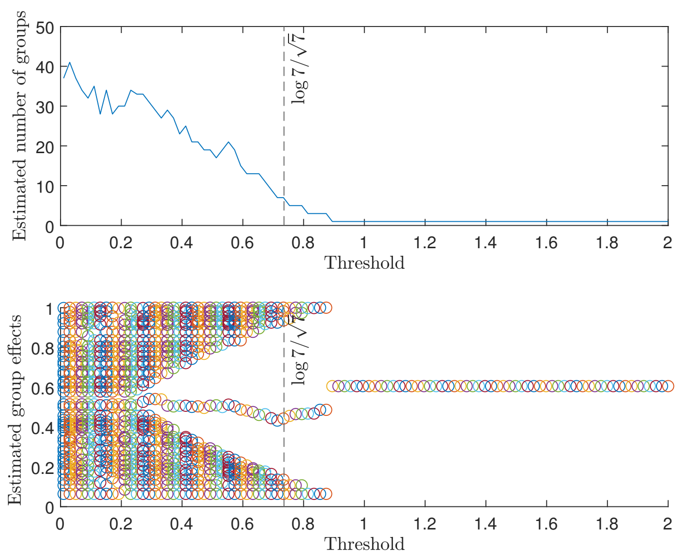
Figure 2 suggests that the larger the $T$, the larger the range of values for $c$ for which $\widehat{G}^{\text{PWD}} = G^0$. Figures 3-5, which inform about the grouping composition, confirm that the larger $T$ the better in term of RI, Recall and Precision and that classification is fast (e.g., compare with $T = 66$). This is also exactly what the small Hausdorff distances suggest in Figure 6.

# S4 Empirical Illustration

In this section, I give a short illustration of the PWD estimator which allows to visualize the grouping "regularization path". I use the balanced subsample of Acemoglu, Johnson, Robinson, and Yared (2008) which contains the Freedom House Index of democracy for $N = 74$ countries (after dropping missing values) observed during $T = 7$ periods over the time span 1970-2000. I estimate model (S3.1) using the PWD estimator for different values of $c_T \in (0, 2)$. Figure 1 reports the estimated number of groups and group-specific effects as a function of $c_T$. The top-panel shows that the regularization path for $\widehat{G}^{\text{PWD}}(c_T)$ is smooth and exhibits a general decreasing pattern from $\widehat{G}^{\text{PWD}}(0.01) \approx 40$ to $\widehat{G}^{\text{PWD}}(2) = 1$. The exact same pattern is observed when estimating model (1.1) under the constraint $\beta^0 = 0$. The bottom-panel suggests a convergence toward $\widehat{G}^{\text{PWD}} = 3$ groups before a sudden phase-transition to $\widehat{G}^{\text{PWD}} = 1$.

# S5 Tables and Figures

Figure 1: PWD Regularization Path




*Notes: Acemoglu, Johnson, Robinson, and Yared (2008)'s democracy data. The top-panel plots* $\widehat{G}^{\mathrm{PWD}}$ *as a function* $c_T$*. The bottom-panel plots* $\left\{\widehat{\alpha}_g^{\mathrm{PWD}} : g \in \left\{1, ..., \widehat{G}^{\mathrm{PWD}}\right\}\right\}$ *as a function of* $c_T$*.*

Table 2: Consistency of the PWD estimator under i.i.d. errors

| | | $G^0 = 2$ | | | | $G^0 = 5$ | | | | $G^0 = 10$ | | | | $G^0 = 50$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $T$ | $\widehat{G}^{\mathrm{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\mathrm{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\mathrm{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\mathrm{PWD}}$ | HD | RI | CPU time |
| 50 | 8 | 13.448 | 0.853 | 0.689 | 0.0231 | 31.569 | 0.6837 | 0.8347 | 0.0245 | 33.595 | 0.6025 | 0.9285 | 0.0268 | 41.097 | 1.0521 | 0.9918 | 0.0271 |
| | 22 | 3.12 | 0.2724 | 0.9679 | 0.0197 | 31.078 | 0.5124 | 0.8398 | 0.028 | 33.978 | 0.5033 | 0.9316 | 0.033 | 44.195 | 0.7252 | 0.9952 | 0.038 |
| | 36 | 2.057 | 0.0523 | 0.9987 | 0.0202 | 29.031 | 0.4071 | 0.8472 | 0.0351 | 34.036 | 0.4676 | 0.9318 | 0.0385 | 43.049 | 0.6741 | 0.9943 | 0.0384 |
| | 50 | 2.0 | 0.0318 | 1.0 | 0.0179 | 25.567 | 0.3422 | 0.8598 | 0.0392 | 34.006 | 0.4309 | 0.9332 | 0.0388 | 41.194 | 0.6594 | 0.9928 | 0.0363 |
| 100 | 10 | 20.265 | 0.8568 | 0.6847 | 0.0356 | 61.661 | 0.7014 | 0.8192 | 0.0514 | 66.048 | 0.6207 | 0.9167 | 0.0473 | 76.09 | 0.6061 | 0.9909 | 0.0468 |
| | 40 | 2.067 | 0.0431 | 0.9991 | 0.0365 | 54.776 | 0.4204 | 0.8283 | 0.1461 | 65.684 | 0.468 | 0.9184 | 0.159 | 77.127 | 0.521 | 0.9916 | 0.186 |
| | 70 | 2.0 | 0.0195 | 1.0 | 0.0366 | 36.093 | 0.3167 | 0.8603 | 0.1492 | 62.926 | 0.3492 | 0.9211 | 0.2321 | 75.669 | 0.5423 | 0.9907 | 0.2714 |
| | 100 | 2.0 | 0.0164 | 1.0 | 0.0464 | 19.101 | 0.2653 | 0.9213 | 0.1268 | 53.595 | 0.2724 | 0.9283 | 0.2673 | 75.805 | 0.5456 | 0.9911 | 0.3696 |
| 200 | 15 | 20.541 | 0.7577 | 0.7738 | 0.0388 | 121.746 | 0.6481 | 0.8105 | 0.2366 | 129.174 | 0.5911 | 0.9095 | 0.2539 | 142.493 | 0.518 | 0.9877 | 0.2808 |
| | 77 | 2.0 | 0.0129 | 1.0 | 0.0331 | 61.0 | 0.3285 | 0.8491 | 0.4463 | 119.47 | 0.3402 | 0.9122 | 0.8661 | 142.23 | 0.4934 | 0.9874 | 1.058 |
| | 139 | 2.0 | 0.0098 | 1.0 | 0.0387 | 13.602 | 0.2447 | 0.9704 | 0.2244 | 66.249 | 0.2485 | 0.934 | 0.8318 | 133.374 | 0.375 | 0.9892 | 1.719 |
| | 200 | 2.0 | 0.0082 | 1.0 | 0.0377 | 5.492 | 0.0657 | 0.9988 | 0.1662 | 26.174 | 0.2079 | 0.9774 | 0.4734 | 96.023 | 0.2161 | 0.9935 | 1.7135 |
| 500 | 23 | 17.832 | 0.6652 | 0.8941 | 0.2407 | 294.97 | 0.6124 | 0.8047 | 1.7975 | 319.05 | 0.5614 | 0.9041 | 1.9425 | 338.79 | 0.5145 | 0.9837 | 2.1004 |
| | 182 | 2.0 | 0.0054 | 1.0 | 0.2244 | 9.838 | 0.2209 | 0.993 | 0.5474 | 80.128 | 0.2383 | 0.9483 | 3.1459 | 236.88 | 0.2404 | 0.9868 | 9.3569 |
| | 341 | 2.0 | 0.0038 | 1.0 | 0.309 | 5.0 | 0.0085 | 1.0 | 0.5465 | 10.566 | 0.0589 | 0.9997 | 0.9849 | 60.138 | 0.1724 | 0.9992 | 4.3177 |
| | 500 | 2.0 | 0.0031 | 1.0 | 0.4193 | 5.0 | 0.0071 | 1.0 | 0.7088 | 10.0 | 0.012 | 1.0 | 1.2732 | 50.056 | 0.039 | 1.0 | 5.2066 |

*Notes*: Results are averaged over 1,000 Monte Carlo replications. $G^0$ ≡True number of groups; $\widehat{G}^{\mathrm{PWD}}$ ≡Estimated number of groups; HD≡Hausdorff Distance between estimated and true group effects; RI≡Rand Index; CPU time≡MATLAB's `cputime`.

Table 3: Consistency of the PWD estimator under weakly dependent errors

| | | $G^0=2$ | | | | $G^0=5$ | | | | $G^0=10$ | | | | $G^0=50$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $T$ | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time |
| 50 | 8 | 28.595 | 1.4024 | 0.5414 | 0.0216 | 32.107 | 1.1536 | 0.8214 | 0.0275 | 33.375 | 0.9449 | 0.9176 | 0.028 | 35.682 | 1.3475 | 0.9855 | 0.0273 |
| | 22 | 17.803 | 0.9754 | 0.6238 | 0.0267 | 31.736 | 0.7796 | 0.832 | 0.03 | 33.706 | 0.6663 | 0.9254 | 0.0315 | 38.623 | 1.1209 | 0.9894 | 0.0325 |
| | 36 | 10.578 | 0.7982 | 0.7543 | 0.0246 | 31.378 | 0.6548 | 0.8374 | 0.039 | 33.919 | 0.5848 | 0.9293 | 0.0382 | 39.864 | 0.9718 | 0.9912 | 0.0386 |
| | 50 | 6.528 | 0.6328 | 0.8646 | 0.0245 | 30.584 | 0.5979 | 0.8417 | 0.0412 | 34.134 | 0.545 | 0.9309 | 0.0418 | 39.854 | 0.867 | 0.9915 | 0.0368 |
| 100 | 10 | 54.019 | 1.4465 | 0.5244 | 0.0466 | 63.822 | 1.2461 | 0.8121 | 0.0525 | 66.017 | 1.0834 | 0.9102 | 0.0495 | 69.514 | 0.9166 | 0.9864 | 0.0441 |
| | 30 | 17.864 | 0.8408 | 0.7239 | 0.051 | 61.334 | 0.6931 | 0.8205 | 0.1498 | 65.385 | 0.613 | 0.917 | 0.1548 | 74.838 | 0.6368 | 0.9904 | 0.178 |
| | 70 | 6.447 | 0.5802 | 0.9223 | 0.0457 | 55.833 | 0.5859 | 0.827 | 0.2099 | 65.453 | 0.5384 | 0.9187 | 0.2397 | 75.422 | 0.5907 | 0.9906 | 0.2736 |
| | 100 | 3.413 | 0.3072 | 0.9795 | 0.0419 | 48.448 | 0.5243 | 0.8368 | 0.2423 | 63.861 | 0.504 | 0.9204 | 0.3224 | 75.413 | 0.5825 | 0.9907 | 0.3614 |
| 200 | 15 | 91.877 | 1.3407 | 0.5195 | 0.1777 | 125.656 | 1.1787 | 0.8074 | 0.2408 | 130.224 | 1.0442 | 0.9063 | 0.2491 | 136.702 | 0.7222 | 0.9849 | 0.2714 |
| | 77 | 10.08 | 0.6585 | 0.9134 | 0.1068 | 106.55 | 0.6098 | 0.816 | 0.7937 | 127.67 | 0.5544 | 0.9102 | 0.9589 | 142.40 | 0.5066 | 0.9873 | 1.0767 |
| | 139 | 2.965 | 0.2254 | 0.9924 | 0.0463 | 71.937 | 0.4896 | 0.8375 | 0.9182 | 115.835 | 0.4826 | 0.9131 | 1.4926 | 142.416 | 0.4901 | 0.9878 | 1.8468 |
| | 200 | 2.166 | 0.0551 | 0.999 | 0.0502 | 41.683 | 0.4085 | 0.8811 | 0.778 | 92.895 | 0.4117 | 0.92 | 1.676 | 136.125 | 0.4528 | 0.9889 | 2.469 |
| 500 | 23 | 170.186 | 1.2131 | 0.5171 | 1.057 | 310.33 | 1.1007 | 0.8035 | 1.869 | 322.65 | 1.0035 | 0.9031 | 1.9837 | 337.404 | 0.7414 | 0.9827 | 2.0861 |
| | 182 | 3.232 | 0.2498 | 0.9958 | 0.2965 | 122.31 | 0.4737 | 0.8396 | 4.8164 | 247.13 | 0.4717 | 0.9081 | 9.8664 | 326.424 | 0.469 | 0.9841 | 13.09 |
| | 341 | 2.016 | 0.0115 | 1.0 | 0.3321 | 22.837 | 0.3443 | 0.9669 | 1.8112 | 96.738 | 0.3481 | 0.9392 | 6.9564 | 215.331 | 0.3505 | 0.9878 | 15.6834 |
| | 500 | 2.0 | 0.0063 | 1.0 | 0.4076 | 6.274 | 0.1528 | 0.9985 | 0.8521 | 30.89 | 0.2868 | 0.985 | 3.3249 | 108.39 | 0.2895 | 0.9949 | 11.2575 |

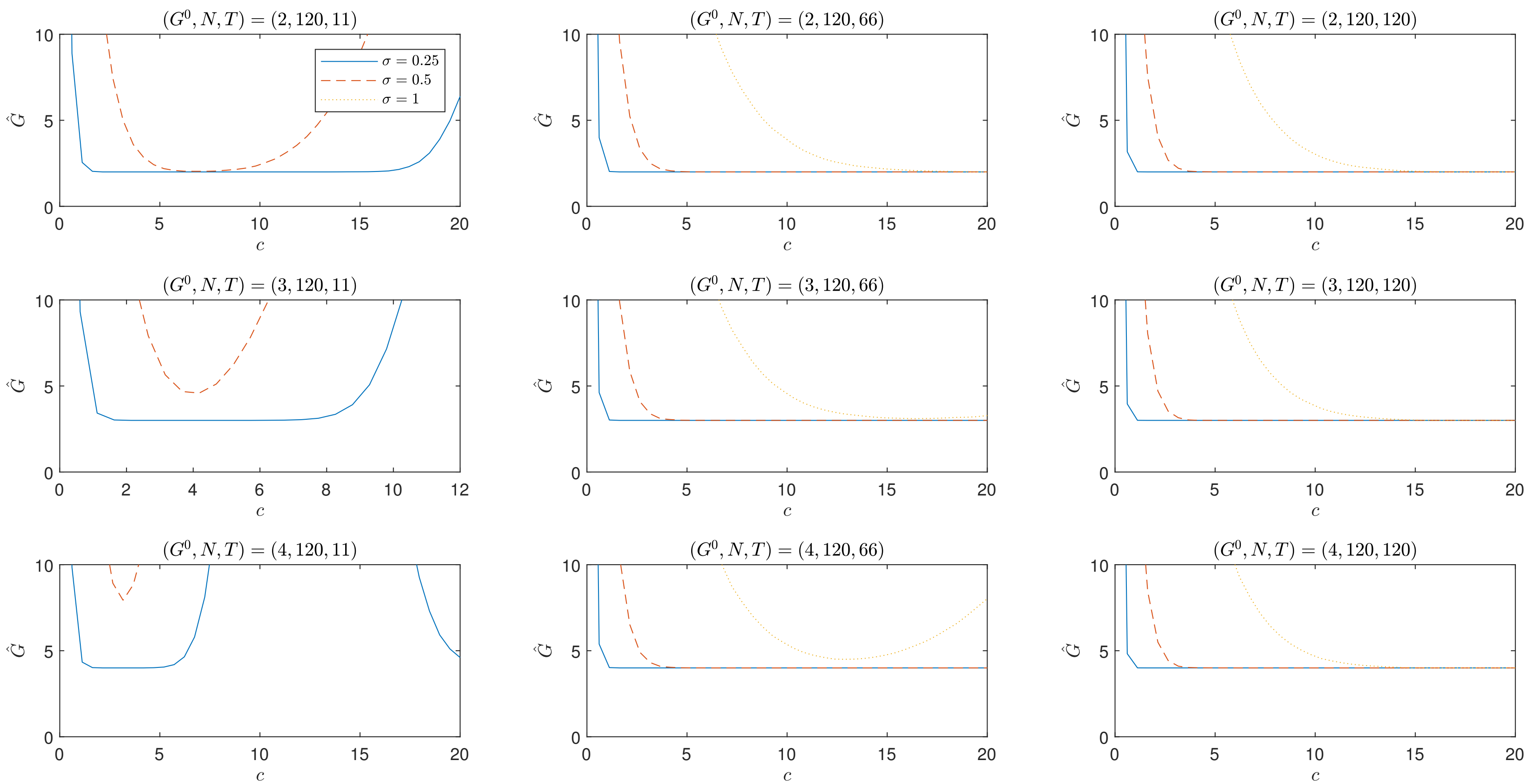
*Notes*: Results are averaged over $1,000$ Monte Carlo replications. $G^0$ ≡True number of groups; $\widehat{G}^{\text{PWD}}$ ≡Estimated number of groups; HD≡Hausdorff Distance between estimated and true group effects; RI≡Rand Index; CPU time≡MATLAB's `cputime`.

Table 4: Consistency of the PWD estimator under heteroskedastic errors

| | | $G^0=2$ | | | | $G^0=5$ | | | | $G^0=10$ | | | | $G^0=50$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | $T$ | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time | $\widehat{G}^{\text{PWD}}$ | HD | RI | CPU time |
| 50 | 8 | 15.869 | 0.9211 | 0.6471 | 0.0225 | 31.651 | 0.7309 | 0.8332 | 0.0255 | 33.744 | 0.6399 | 0.9268 | 0.026 | 40.022 | 1.0997 | 0.9906 | 0.0275 |
| | 22 | 4.841 | 0.4884 | 0.9134 | 0.0224 | 31.374 | 0.5593 | 0.8386 | 0.0311 | 34.255 | 0.5243 | 0.9313 | 0.0332 | 43.216 | 0.8085 | 0.9943 | 0.0365 |
| | 36 | 2.257 | 0.1016 | 0.9937 | 0.022 | 30.082 | 0.454 | 0.8442 | 0.033 | 34.272 | 0.4806 | 0.9316 | 0.0363 | 42.551 | 0.7016 | 0.9939 | 0.0443 |
| | 50 | 2.006 | 0.0381 | 0.9999 | 0.0185 | 26.943 | 0.3792 | 0.8545 | 0.0364 | 34.135 | 0.4532 | 0.9328 | 0.0355 | 41.12 | 0.679 | 0.9928 | 0.0415 |
| 100 | 10 | 29.069 | 0.9689 | 0.605 | 0.0456 | 62.248 | 0.7981 | 0.8179 | 0.0568 | 65.794 | 0.7018 | 0.9153 | 0.0474 | 74.017 | 0.6838 | 0.9898 | 0.0549 |
| | 40 | 2.344 | 0.1132 | 0.9955 | 0.0457 | 56.172 | 0.4677 | 0.8265 | 0.1434 | 65.958 | 0.4917 | 0.9182 | 0.159 | 76.597 | 0.5338 | 0.9915 | 0.1772 |
| | 70 | 2.0 | 0.0213 | 1.0 | 0.0416 | 39.427 | 0.3467 | 0.8523 | 0.1642 | 63.817 | 0.3783 | 0.9205 | 0.2393 | 75.513 | 0.5473 | 0.9907 | 0.2803 |
| | 100 | 2.0 | 0.0177 | 1.0 | 0.0402 | 22.032 | 0.2861 | 0.9072 | 0.1375 | 55.448 | 0.2937 | 0.9267 | 0.2832 | 75.832 | 0.5458 | 0.9911 | 0.3649 |
| 200 | 15 | 35.896 | 0.8967 | 0.6458 | 0.0561 | 122.235 | 0.7455 | 0.8101 | 0.226 | 129.324 | 0.6703 | 0.9089 | 0.2543 | 142.211 | 0.5459 | 0.987 | 0.2748 |
| | 77 | 2.0 | 0.0143 | 1.0 | 0.0307 | 68.329 | 0.3566 | 0.8408 | 0.4835 | 121.611 | 0.3689 | 0.9117 | 0.9058 | 142.751 | 0.4924 | 0.9874 | 1.0951 |
| | 139 | 2.0 | 0.0101 | 1.0 | 0.0396 | 16.56 | 0.2582 | 0.9591 | 0.2576 | 72.042 | 0.2621 | 0.9302 | 0.9223 | 134.743 | 0.3999 | 0.9891 | 1.8094 |
| | 200 | 2.0 | 0.0084 | 1.0 | 0.0454 | 5.76 | 0.087 | 0.9981 | 0.1657 | 28.772 | 0.2154 | 0.9736 | 0.5166 | 99.223 | 0.2239 | 0.9931 | 1.8396 |
| 500 | 33 | 37.936 | 0.7906 | 0.7672 | 0.3116 | 298.003 | 0.6836 | 0.8046 | 1.8629 | 319.607 | 0.6228 | 0.904 | 1.9931 | 339.888 | 0.5361 | 0.9835 | 2.1764 |
| | 182 | 2.0 | 0.0056 | 1.0 | 0.2115 | 11.672 | 0.2355 | 0.9897 | 0.599 | 88.894 | 0.2462 | 0.9435 | 3.6233 | 245.004 | 0.2504 | 0.9865 | 10.57 |
| | 341 | 2.0 | 0.0041 | 1.0 | 0.3299 | 5.0 | 0.0088 | 1.0 | 0.5171 | 10.966 | 0.0819 | 0.9995 | 0.9751 | 63.247 | 0.18 | 0.9989 | 4.7129 |
| | 500 | 2.0 | 0.0033 | 1.0 | 0.4094 | 5.0 | 0.0073 | 1.0 | 0.7156 | 10.0 | 0.0124 | 1.0 | 1.271 | 50.098 | 0.0433 | 1.0 | 5.4568 |

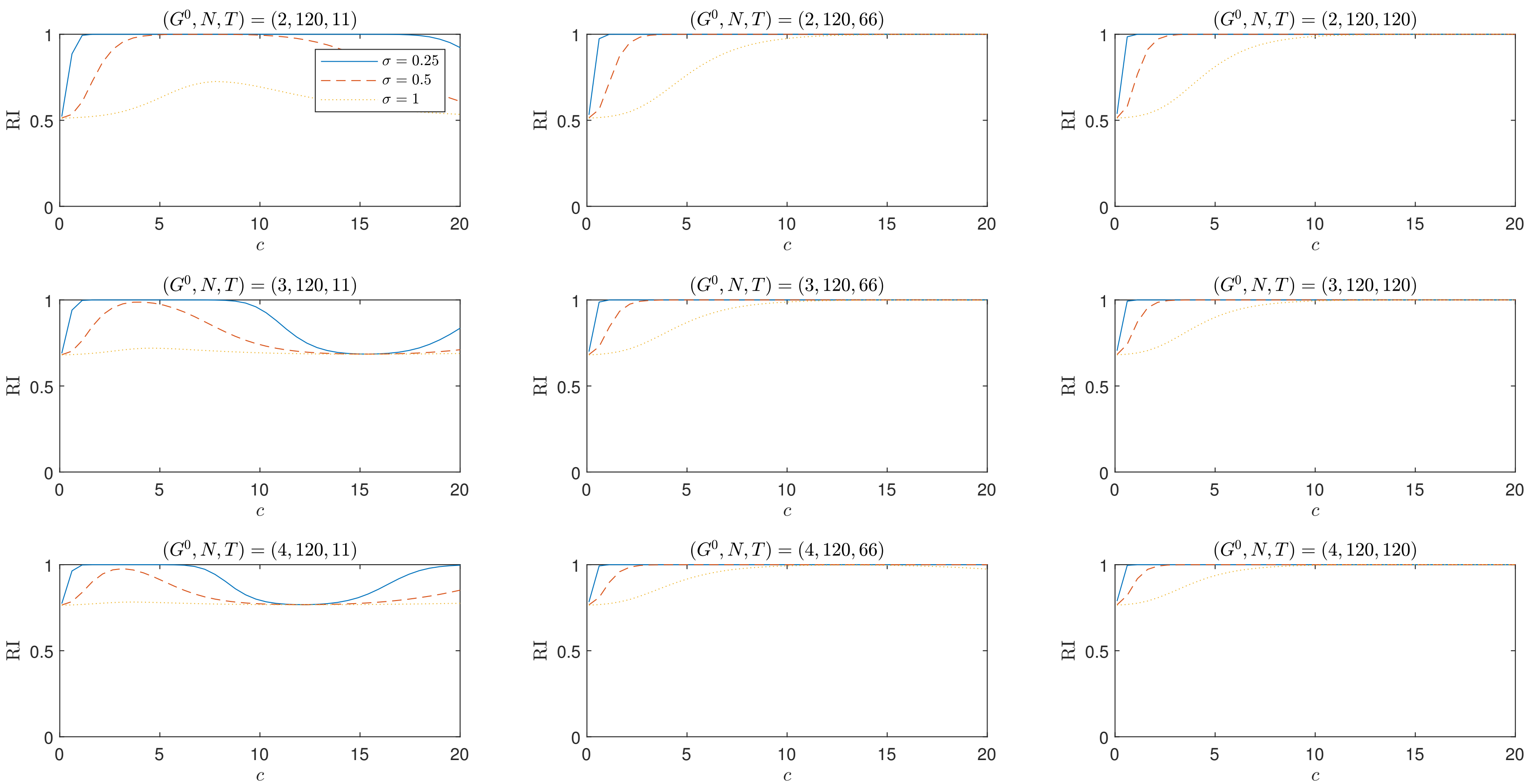
*Notes*: Results are averaged over 1,000 Monte Carlo replications. $G^0$ ≡True number of groups; $\widehat{G}^{\text{PWD}}$ ≡Estimated number of groups; HD≡Hausdorff Distance between estimated and true group effects; RI≡Rand Index; CPU time≡MATLAB's `cputime`.

Figure 2: Sensitivity of the estimated number of groups ($\widehat{G}$)



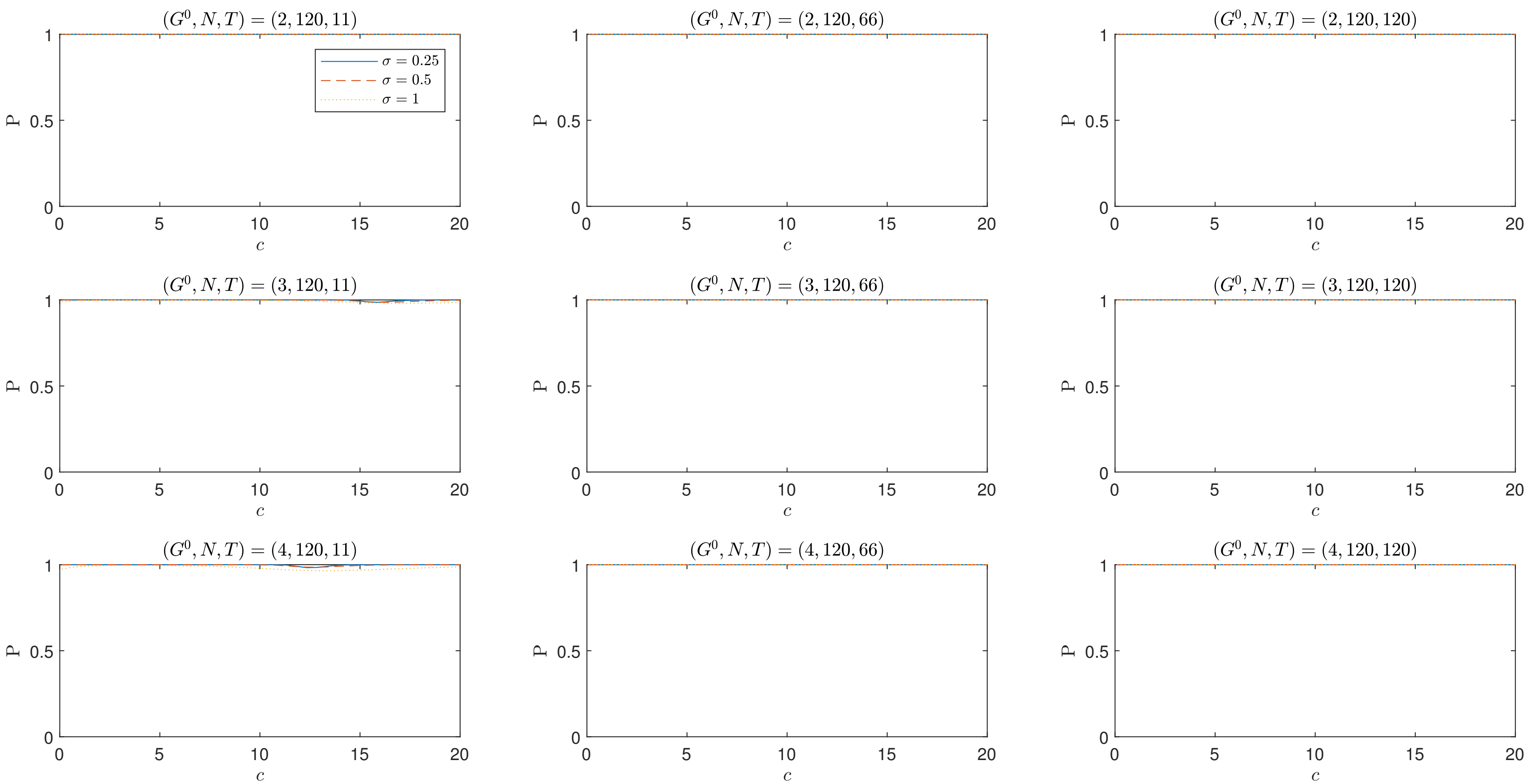

*Notes*: solid blue, dashed red, and dotted yellow lines report averages of $\widehat{G}$ with $c_T = cT^{-4/5}$ as a function of $c \in (0, 20)$ for noise standard deviation of .25, .5 and 1 respectively. In each plot, $N = 120$; from top to bottom $G^0 = 2, 3, 4$; from left to right $T = 11, 66, 120$. Results are averaged over $1,000$ Monte Carlo replications.

Figure 3: Sensitivity of the Rand Index (RI)



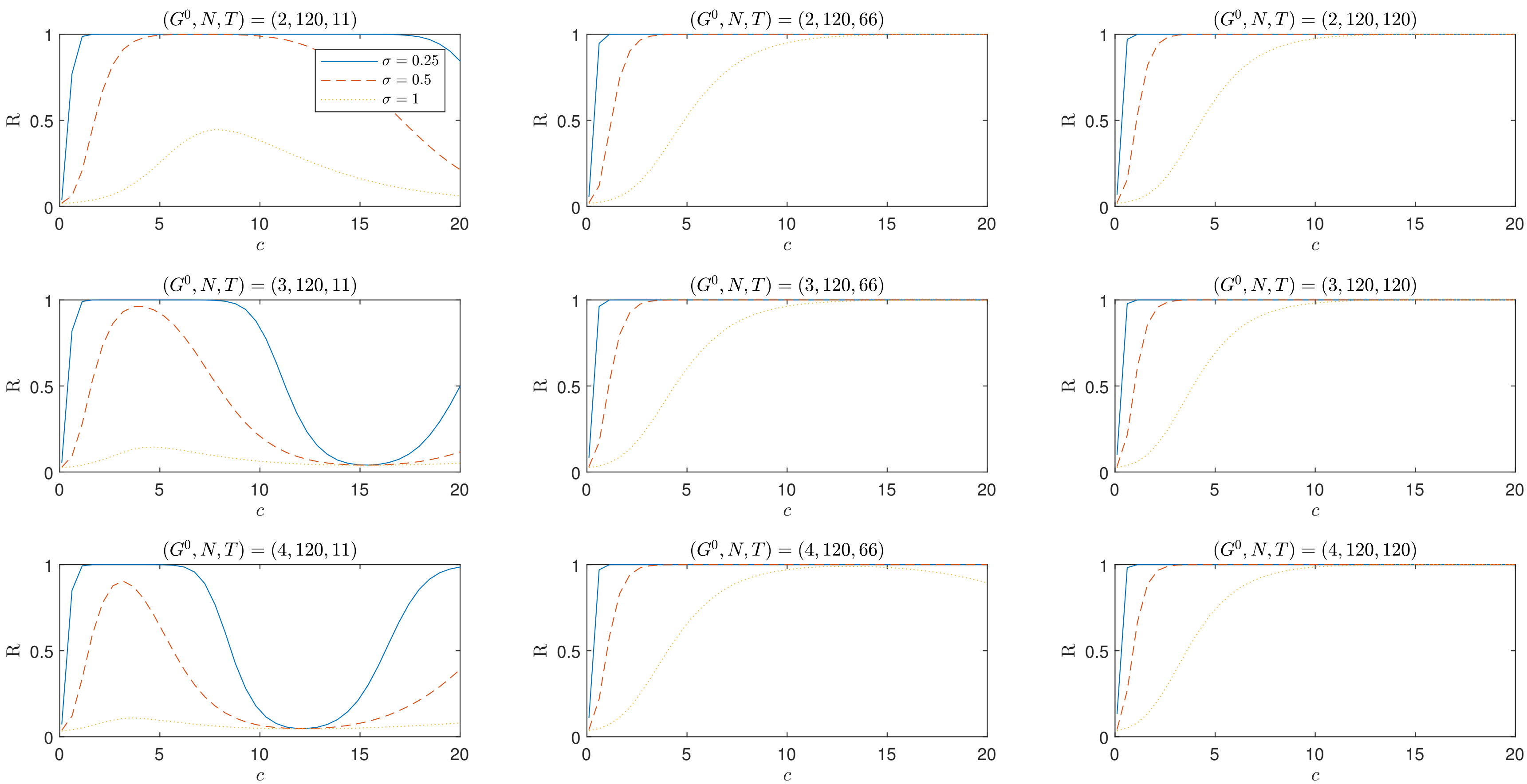

*Notes*: solid blue, dashed red, and dotted yellow lines report averages of RI with $c_T = cT^{-4/5}$ as a function of $c \in (0, 20)$ for noise standard deviation of .25, .5 and 1 respectively. In each plot, $N = 120$; from top to bottom $G^0 = 2, 3, 4$; from left to right $T = 11, 66, 120$. Results are averaged over $1,000$ Monte Carlo replications.

Figure 4: Sensitivity of the Precision rate (P)



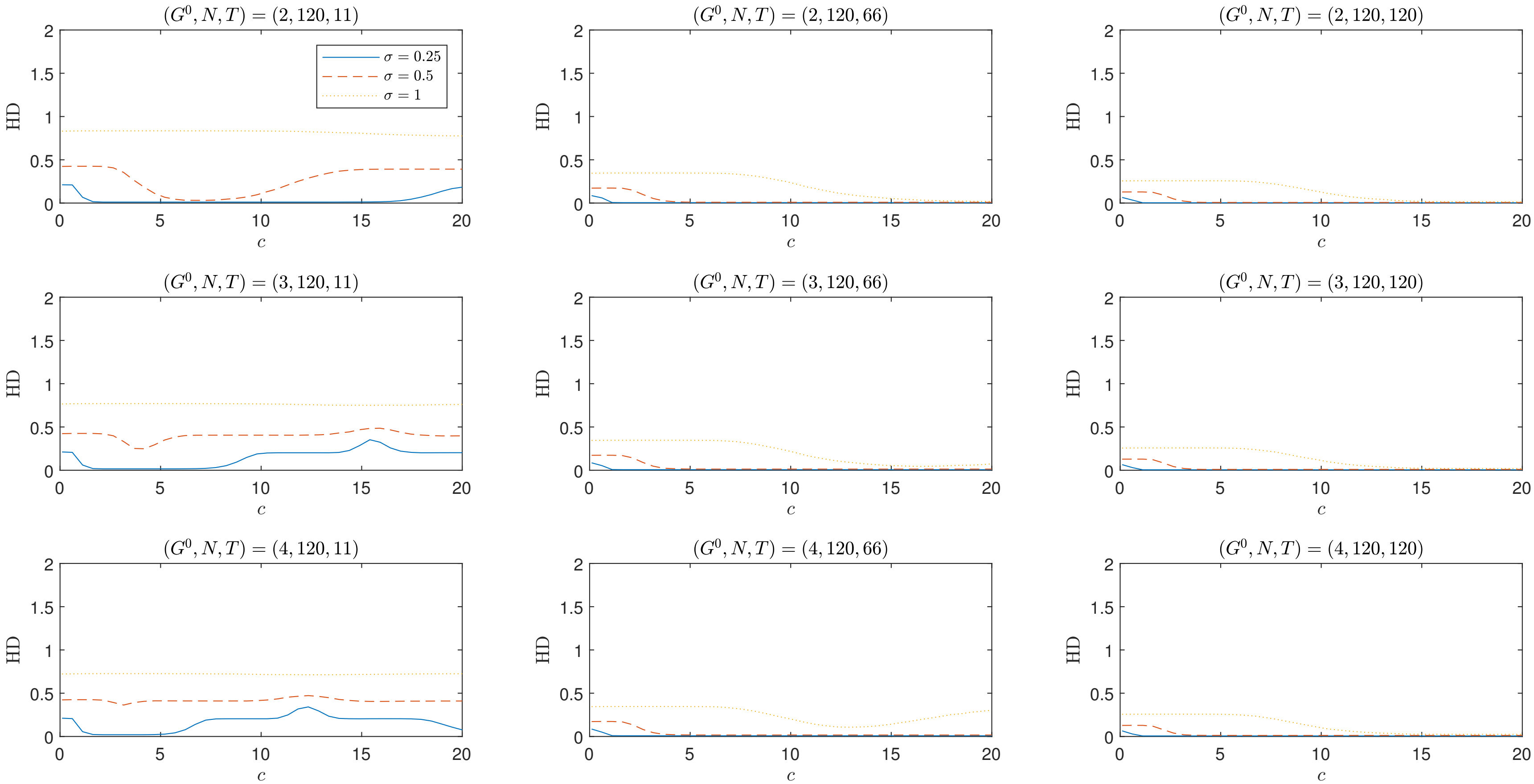

*Notes*: solid blue, dashed red, and dotted yellow lines report averages of P with $c_T = cT^{-4/5}$ as a function of $c \in (0, 20)$ for noise standard deviation of .25, .5 and 1 respectively. In each plot, $N = 120$; from top to bottom $G^0 = 2, 3, 4$; from left to right $T = 11, 66, 120$. Results are averaged over $1,000$ Monte Carlo replications.

Figure 5: Sensitivity of the Recall rate (R)




*Notes*: solid blue, dashed red, and dotted yellow lines report averages of R with $c_T = cT^{-4/5}$ as a function of $c \in (0, 20)$ for noise standard deviation of .25, .5 and 1 respectively. In each plot, $N = 120$; from top to bottom $G^0 = 2, 3, 4$; from left to right $T = 11, 66, 120$. Results are averaged over $1,000$ Monte Carlo replications.

Figure 6: Sensitivity of the Hausdorff Distance (HD)




*Notes*: solid blue, dashed red, and dotted yellow lines report averages of HD with $c_T = cT^{-4/5}$ as a function of $c \in (0, 20)$ for noise standard deviation of .25, .5 and 1 respectively. In each plot, $N = 120$; from top to bottom $G^0 = 2, 3, 4$; from left to right $T = 11, 66, 120$. Results are averaged over $1,000$ Monte Carlo replications.

# References


Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2008): "Income and Democracy," *American Economic Review*, 98(3), 808–42.

Arellano, M., and S. Bonhomme (2011): "Nonlinear Panel Data Analysis," *Annual Review of Economics*, 3(1), 395–424.

Bonhomme, S., and E. Manresa (2015): "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 83(3), 1147–1184.

Moon, H. R., and M. Weidner (2019): "Nuclear Norm Regularized Estimation of Panel Regression Models," .

Wooldridge, J. (2010): *Econometric Analysis of Cross Section and Panel Data, second edition*, Econometric Analysis of Cross Section and Panel Data. MIT Press.