Identification and Estimation of Categorical Random Coefficient Models

IAAE 2023, Oslo

Zhan Gao* and M. Hashem Pesaran*,†

* University of Southern California
† Trinity College, University of Cambridge

RANDOM COEFFICIENT MODELS

$$y_i = a + \beta_i x_i + u_i$$

- Nonparametric identification and estimation of the dist. of β_i and u_i w/ i.i.d. β_i and u_i (Beran and Hall, 1992)
- Beran (1993); Beran and Millar (1994); Beran, Feuerverger, and Hall (1996); Hoderlein, Klemelä, and Mammen (2010); Hoderlein, Holzmann, and Meister (2017); Breunig and Hoderlein (2018)

RANDOM COEFFICIENT MODELS

$$y_i = a + \beta_i x_i + u_i$$

- Scenario 1: estimate the consumer surplus based on a linear demand system (Hausman and Newey, 1995; Foster and Hahn, 2000)

RANDOM COEFFICIENT MODELS

$$y_i = a + \beta_i x_i + u_i$$

- Scenario 1: estimate the consumer surplus based on a linear demand system (Hausman and Newey, 1995; Foster and Hahn, 2000)
- Scenario 2: heterogeneous treatment effect / return to education (Lemieux, 2006; Bick et al., 2022)

RANDOM COEFFICIENT MODELS

$$y_i = a + \beta_i x_i + u_i$$

- Scenario 1: estimate the consumer surplus based on a linear demand system (Hausman and Newey, 1995; Foster and Hahn, 2000)
- Scenario 2: heterogeneous treatment effect / return to education (Lemieux, 2006; Bick et al., 2022)

$\uparrow$

parametric assumptions on β_i

CATEGORICAL RANDOM COEFFICIENT MODEL

- Suppose the single cross-section observations $\{y_i, x_i, \boldsymbol{z}_i\}_{i=1}^{n}$ follow the linear model

$$y_i = x_i \beta_i + {\boldsymbol{z}_i}' \boldsymbol{\gamma} + u_i$$

CATEGORICAL RANDOM COEFFICIENT MODEL

$$y_i = x_i \beta_i + \boldsymbol{z}_i{}' \boldsymbol{\gamma} + u_i$$

- $\beta_i \in \{b_1, b_2, \cdots, b_K\} \sim_{\text{i.i.d.}}$ a categorical distribution,

$$\beta_i = b_k, \ \text{w.p. } \pi_k,$$

with $\pi_k \in (0,1)$, $\sum_{k=1}^{K} \pi_k = 1$, $b_1 < b_2 < \cdots < b_K$

CATEGORICAL RANDOM COEFFICIENT MODEL

$$y_i = x_i \beta_i + \boldsymbol{z}_i{}' \boldsymbol{\gamma} + u_i$$

- $\beta_i \in \{b_1, b_2, \cdots, b_K\} \sim_{\text{i.i.d.}}$ a categorical distribution,

$$\beta_i = b_k, \text{ w.p. } \pi_k,$$

  with $\pi_k \in (0, 1)$, $\sum_{k=1}^{K} \pi_k = 1$, $b_1 < b_2 < \cdots < b_K$
    - K is assumed to be known.

CATEGORICAL RANDOM COEFFICIENT MODEL

$$y_i = x_i\beta_i + \boldsymbol{z}_i{}'\boldsymbol{\gamma} + u_i$$

- $\beta_i \in \{b_1, b_2, \cdots, b_K\} \sim_{\text{i.i.d.}}$ a categorical distribution,

$$\beta_i = b_k, \text{ w.p. } \pi_k,$$

  with $\pi_k \in (0,1)$, $\sum_{k=1}^{K} \pi_k = 1$, $b_1 < b_2 < \cdots < b_K$
- $\boldsymbol{\gamma} \in \mathbb{R}^p$ is homogeneous, which may include the intercept.

CATEGORICAL RANDOM COEFFICIENT MODEL

$$y_i = x_i\beta_i + \boldsymbol{z}_i{}'\boldsymbol{\gamma} + u_i$$

- $\beta_i \in \{b_1, b_2, \cdots, b_K\} \sim_{\text{i.i.d.}}$ a categorical distribution,

$$\beta_i = b_k, \text{ w.p. } \pi_k,$$

  with $\pi_k \in (0,1)$, $\sum_{k=1}^K \pi_k = 1$, $b_1 < b_2 < \cdots < b_K$
- $\boldsymbol{\gamma} \in \mathbb{R}^p$ is homogeneous, which may include the intercept.
- The idiosyncratic shock $u_i \sim \left(0, \sigma_i^2\right) \perp \boldsymbol{w}_i = (x_i, \boldsymbol{z}_i')'$.

CATEGORICAL RANDOM COEFFICIENT MODEL

$$y_i = x_i \beta_i + \boldsymbol{z}_i{}' \boldsymbol{\gamma} + u_i$$

- $\beta_i \in \{b_1, b_2, \cdots, b_K\} \sim_{\text{i.i.d.}}$ a categorical distribution,

$$\beta_i = b_k, \;\text{ w.p. }\; \pi_k,$$

  with $\pi_k \in (0,1)$, $\sum_{k=1}^{K} \pi_k = 1,\; b_1 < b_2 < \cdots < b_K$
- $\boldsymbol{\gamma} \in \mathbb{R}^p$ is homogeneous, which may include the intercept.
- The idiosyncratic shock $u_i \sim \left(0, \sigma_i^2\right) \perp \boldsymbol{w}_i = (x_i, \boldsymbol{z}_i')'$.
- Allow for distributions of $\boldsymbol{w}_i$ and u_i are not identical across i while independence across i is maintained. Details

CATEGORICAL RANDOM COEFFICIENT MODEL

Goal: Identify and estimate $\boldsymbol{\gamma}$ and the distribution of β_i,

$$\boldsymbol{\theta} = \left(\boldsymbol{\pi}' = (\pi_1, \pi_2, ..., \pi_K), \boldsymbol{b}' = (b_1, b_2, ..., b_K)\right).$$

IDENTIFITION OF THE MOMENTS OF β_i

- Let $\tilde{y}_i = y_i - \mathbf{z}_i'\boldsymbol{\gamma}$, consider

$$\tilde{y}_i^r = (x_i\beta_i + u_i)^r \,,$$
$$\tilde{y}_i^r x_i^r = (x_i\beta_i + u_i.)^r \, x_i^r$$

IDENTIFITION OF THE MOMENTS OF β_i

- Let $\tilde{y}_i = y_i - \mathbf{z}_i'\boldsymbol{\gamma}$, consider

$$\tilde{y}_i^r = (x_i\beta_i + u_i)^r \,,$$
$$\tilde{y}_i^r x_i^r = (x_i\beta_i + u_i.)^r \, x_i^r$$

- Take expectations, sum over i and take limits,

$$\rho_{0,r}\mathrm{E}\left(\beta_i^r\right) + \sigma_r = \rho_{r,0} - \sum_{q=2}^{r-1}\binom{r}{q}\rho_{0,r-q}\sigma_q\mathrm{E}\left(\beta_i^{r-q}\right),$$
$$\rho_{0,2r}\mathrm{E}\left(\beta_i^r\right) + \rho_{0,r}\sigma_r = \rho_{r,r} - \sum_{q=2}^{r-1}\binom{r}{q}\rho_{0,2r-q}\sigma_q\mathrm{E}\left(\beta_i^{r-q}\right).$$

Def. of $\rho_{r,s}$ and σ_r

IDENTIFITION OF THE MOMENTS OF β_i

- Let $\tilde{y}_i = y_i - \boldsymbol{z}_i'\boldsymbol{\gamma}$, consider

$$\tilde{y}_i^r = (x_i\beta_i + u_i)^r \,,$$
$$\tilde{y}_i^r x_i^r = (x_i\beta_i + u_i.)^r \, x_i^r$$

- Take expectations, sum over i and take limits,

$$\rho_{0,r}\mathrm{E}\left(\beta_i^r\right) + \sigma_r = \rho_{r,0} - \sum_{q=2}^{r-1}\binom{r}{q}\rho_{0,r-q}\sigma_q\mathrm{E}\left(\beta_i^{r-q}\right),$$
$$\rho_{0,2r}\mathrm{E}\left(\beta_i^r\right) + \rho_{0,r}\sigma_r = \rho_{r,r} - \sum_{q=2}^{r-1}\binom{r}{q}\rho_{0,2r-q}\sigma_q\mathrm{E}\left(\beta_i^{r-q}\right).$$

▸ Def. of $\rho_{r,s}$ and σ_r

- Recursively solve the linear systems for $\mathrm{E}\left(\beta_i^r\right)$ and σ_r, $r = 2, 3, \cdots, 2K-1$,

IDENTIFICATION OF THE DISTRIBUTION OF β_i

- Show the system

$$\mathrm{E}\left(\beta_i^r\right) = \sum_{k=1}^{K} \pi_k b_k^r,\ r = 0, 1, 2, \cdots, 2K-1,$$

has a unique solution $\left(\boldsymbol{\pi}', \boldsymbol{b}'\right)$ based on a linear recurrence structure, under the conditions

$$b_1 < b_2 < \cdots < b_K,\ \text{and}\ \pi_k \in (0, 1)\,.$$

$\sqrt{n}$-CONSISTENT ESTIMATOR FOR γ

- Let $\boldsymbol{w}_i = (x_i, \boldsymbol{z}_i')'$, $\boldsymbol{\phi} = (\mathrm{E}(\beta_i), \boldsymbol{\gamma}')'$, the OLS estimator

$$\hat{\boldsymbol{\phi}} = \left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{w}_i \boldsymbol{w}_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{w}_i y_i\right)$$

- Under regularity conditions,
  - [Consistency] $\hat{\boldsymbol{\phi}}$ is a consistent estimator for $\boldsymbol{\phi}$
  - [Asymptotic normality]

$$\sqrt{n}\left(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right) \to_d N\left(\mathbf{0}, \boldsymbol{Q}_{ww}^{-1} \boldsymbol{V}_{w\xi} \boldsymbol{Q}_{ww}^{-1}\right)$$

GMM ESTIMATOR FOR $\boldsymbol{\theta} = \left(\boldsymbol{\pi}', \boldsymbol{b}'\right)'$

- The moment conditions,

$$\mathrm{E}\left(\tilde{y}_i^r x_i^{s_r}\right) - \sum_{q=0}^{r} \binom{r}{q} \mathrm{E}\left(x_i^{r-q+s_r}\right) \mathrm{E}\left(u_i^q\right) \underbrace{\mathrm{E}\left(\beta_i^{r-q}\right)}_{=\sum_{k=1}^{K} \pi_k b_k^{r-q}} = 0$$

  $s_r = 0, 1, \cdots, S - r$, where $S > 2K - 1$.

- Plug in $\hat{\boldsymbol{\gamma}}$ for $\boldsymbol{\gamma}$ in the sample analogue, $\hat{\boldsymbol{g}}_n\left(\boldsymbol{\theta}, \boldsymbol{\sigma}, \hat{\gamma}\right)$.
- GMM estimator

$$\left(\hat{\boldsymbol{\theta}}', \hat{\boldsymbol{\sigma}}'\right)' = \arg \min_{\boldsymbol{\theta} \in \Theta, \boldsymbol{\sigma} \in \mathcal{S}} \hat{\boldsymbol{g}}_n\left(\boldsymbol{\theta}, \boldsymbol{\sigma}, \hat{\gamma}\right)' \boldsymbol{A}_n \hat{\boldsymbol{g}}_n\left(\boldsymbol{\theta}, \boldsymbol{\sigma}, \hat{\gamma}\right)$$

GMM ESTIMATOR FOR $\boldsymbol{\theta} = \left(\boldsymbol{\pi}', \boldsymbol{b}'\right)'$

Under regularity conditions,

- [Consistency] Let $\boldsymbol{\eta} = \left(\boldsymbol{\theta}', \boldsymbol{\sigma}'\right)'$ and $\boldsymbol{\eta}_0 = \left(\boldsymbol{\theta}_0', \boldsymbol{\sigma}_0'\right)'$,

$$\hat{\boldsymbol{\eta}} \to_p \boldsymbol{\eta}_0,$$

  as $n \to \infty$.

- [Asymptotic normality] Let $\sqrt{n}\,(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \to_d \zeta_\gamma \sim N(0, V_\gamma)$,

$$\sqrt{n}\,(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \to_d \left(\boldsymbol{G}_0' \boldsymbol{A} \boldsymbol{G}_0\right)^{-1} \boldsymbol{G}_0' \boldsymbol{A} \left(\boldsymbol{\zeta} + \boldsymbol{G}_{0,\gamma} \boldsymbol{\zeta}_\gamma\right),$$

  as $n \to \infty$.

MULTIPLE REGRESSORS WITH RANDOM COEFFICIENTS

- The model

$$y_i = \boldsymbol{x}_i' \boldsymbol{\beta}_i + \boldsymbol{z}_i' \boldsymbol{\gamma} + u_i,$$

MULTIPLE REGRESSORS WITH RANDOM COEFFICIENTS

- The model

$$y_i = \boldsymbol{x}_i' \boldsymbol{\beta}_i + \boldsymbol{z}_i' \boldsymbol{\gamma} + u_i,$$

- $\boldsymbol{\beta}_i \in \mathbb{R}^p$ follows the multivariate distribution

$$\Pr\left(\beta_{i1} = b_{1k_1}, \beta_{i2} = b_{2k_2}, \cdots, \beta_{ip} = b_{pk_p}\right) = \pi_{k_1,k_2,\cdots,k_p},$$

with $k_j \in \{1, 2, \cdots, K\}$, $b_{j1} < b_{j2} < \cdots < b_{jK}$, and $\sum_{k_1,k_2,\cdots,k_p \in \{1,2,\cdots,K\}} \pi_{k_1,k_2,\cdots,k_p} = 1$.

MULTIPLE REGRESSORS WITH RANDOM COEFFICIENTS

- The model

$$y_i = \boldsymbol{x}_i' \boldsymbol{\beta}_i + \boldsymbol{z}_i' \boldsymbol{\gamma} + u_i,$$

- $\boldsymbol{\beta}_i \in \mathbb{R}^p$ follows the multivariate distribution

$$\Pr\left(\beta_{i1} = b_{1k_1}, \beta_{i2} = b_{2k_2}, \cdots, \beta_{ip} = b_{pk_p}\right) = \pi_{k_1,k_2,\cdots,k_p},$$

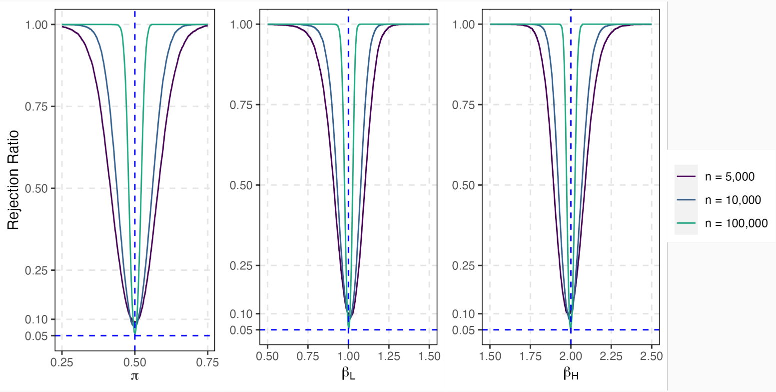
  with $k_j \in \{1, 2, \cdots, K\}$, $b_{j1} < b_{j2} < \cdots < b_{jK}$, and $\sum_{k_1,k_2,\cdots,k_p \in \{1,2,\cdots,K\}} \pi_{k_1,k_2,\cdots,k_p} = 1$.
- Identification and estimation of the marginal distributions of $\boldsymbol{\beta}_i$ follow as corollaries.
- The joint distribution can be identified in special cases, $p = 2$ and $K = 2$ for example.

MONTE CARLO

	$\pi = 0.5$			$\beta_L = 1$			$\beta_H = 2$		
n	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
500	-0.0234	0.2384	0.3678	-0.0882	0.6297	0.3217	-0.0216	0.5816	0.2186
1K	-0.0185	0.1769	0.2981	-0.0362	0.4285	0.2767	-0.0198	0.3216	0.2083
2K	-0.0069	0.1233	0.2376	-0.0151	0.2274	0.2370	-0.0123	0.1574	0.1828
5K	-0.0029	0.0677	0.1586	-0.0020	0.0988	0.1504	-0.0060	0.0735	0.1434
10K	-0.0010	0.0414	0.1112	0.0008	0.0535	0.1060	-0.0032	0.0463	0.1050
100K	0.0001	0.0114	0.0610	0.0006	0.0135	0.0666	-0.0003	0.0135	0.0620

Data generating process: $y_i = \alpha + x_i\beta_i + z_{i1}\gamma_1 + z_{i2}\gamma_2 + u_i$,for $i = 1, 2, ..., n$. generate $x_i = (\tilde{x}_{1i} - 2)/2$ where $\tilde{x}_{1i} \sim \text{IID}\chi^2(2)$ for $i = 1, 2, \cdots, \lfloor n/2 \rfloor$, and $x_i = (\tilde{x}_{2i} - 2)/4$ where $\tilde{x}_{2i} \sim \text{IID}\chi^2(4)$, for $i = \lfloor n/2 \rfloor + 1, \cdots, n$. The additional regressors, z_{ij}, for $j = 1, 2$ with homogeneous slopes are generated as $z_{i1} = x_i + v_{i1}$ and $z_{i2} = z_{i1} + v_{i2}$, with $v_{ij} \sim \text{IID } N(0, 1)$, for $j = 1, 2$. The error term u_i is generated as $u_i = \sigma_i\varepsilon_i$, where σ_i^2 are generated as $0.5(1 + \text{IID}\chi^2(1))$, and $\varepsilon_i \sim \text{IID}N(0, 1)$.

MONTE CARLO



Empirical Power Functions

HETEROGENEOUS RETURN TO EDUCATION

- Estimate the Mincerian equation with repeated cross-sectional data by education groups,

$$\log \text{wage}_{it} = \alpha_t + \beta_{it}\text{edu}_{it} + \rho_{1t}\text{exper}_{it} + \rho_{2t}\text{exper}_{it}^2 + \tilde{\mathbf{z}}_{it}'\tilde{\boldsymbol{\gamma}}_t + u_{it},$$

where $\beta_{it} = \begin{cases} b_{tL} & \text{w.p. } \pi_t, \\ b_{tH} & \text{w.p. } 1-\pi_t. \end{cases}$

HETEROGENEOUS RETURN TO EDUCATION

- Estimate the Mincerian equation with repeated cross-sectional data by education groups,

$$\log \text{wage}_{it} = \alpha_t + \beta_{it}\text{edu}_{it} + \rho_{1t}\text{exper}_{it} + \rho_{2t}\text{exper}_{it}^2 + \tilde{\mathbf{z}}_{it}'\tilde{\boldsymbol{\gamma}}_t + u_{it},$$

$$\text{where } \beta_{it} = \begin{cases} b_{tL} & \text{w.p. } \pi_t, \\ b_{tH} & \text{w.p. } 1 - \pi_t. \end{cases}$$

- Data: The May and Outgoing Rotation Group (ORG) supplements of the Current Population Survey (CPS) data, 1973 - 2003.

HETEROGENEOUS RETURN TO EDUCATION

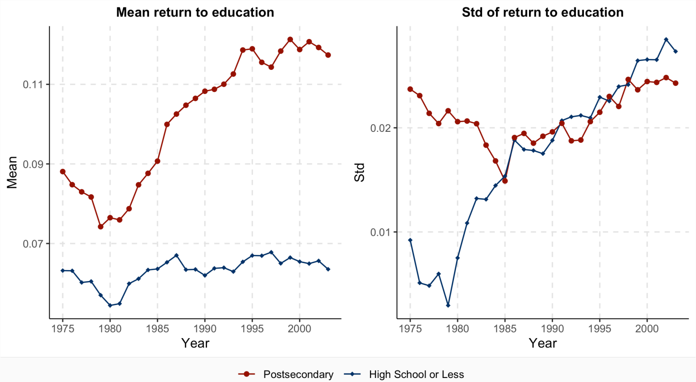
- Estimate the Mincerian equation with repeated cross-sectional data by education groups,

$$\log \text{wage}_{it} = \alpha_t + \beta_{it}\text{edu}_{it} + \rho_{1t}\text{exper}_{it} + \rho_{2t}\text{exper}_{it}^2 + \tilde{\boldsymbol{z}}_{it}'\tilde{\boldsymbol{\gamma}}_t + u_{it},$$

$$\text{where } \beta_{it} = \begin{cases} b_{tL} & \text{w.p. } \pi_t, \\ b_{tH} & \text{w.p. } 1 - \pi_t. \end{cases}$$

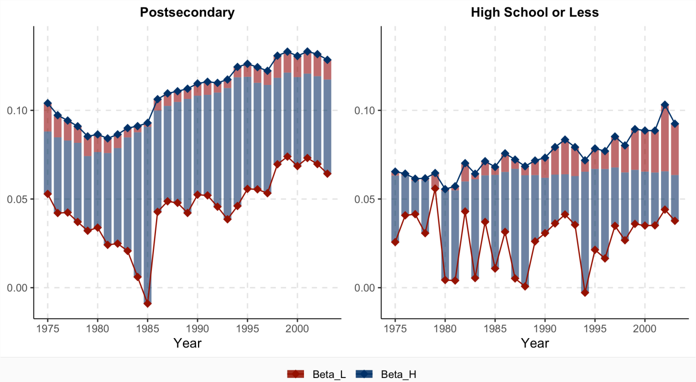
- Data: The May and Outgoing Rotation Group (ORG) supplements of the Current Population Survey (CPS) data, 1973 - 2003.
- Between group heterogeneity $\uparrow$ due to the postsecondary education group; Within group heterogeneity $\uparrow$ in high school or less education group.

HETEROGENEOUS RETURN TO EDUCATION



- Red line: the proportion of low return group.
  Blue line: the proportion of high return group.
- Left panel: the mean return to education
  Right panel: the std of the return to education.

HETEROGENEOUS RETURN TO EDUCATION



- Red bar represents the proportion of low return group. Blue bar represents the proportion of high return group.
- ◇ / ◇ represent the estimated high / low return to education.
- Left panel: Postsecondary / Right panel: High school or less group.

Thank you!

Questions and comments are welcomed

References

Beran, R. (1993). Semiparametric random coefficient regression models. Annals of the Institute of Statistical Mathematics 45(4), 639–654.

Beran, R., A. Feuerverger, and P. Hall (1996). On nonparametric estimation of intercept and slope distributions in random coefficient regression. The Annals of Statistics 24(6), 2569–2592.

Beran, R. and P. Hall (1992). Estimating coefficient distributions in random coefficient regressions. The Annals of Statistics 20(4), 1970–1984.

Beran, R. and P. W. Millar (1994). Minimum distance estimation in random coefficient regression models. The Annals of Statistics 22(4), 1976–1992.

Bick, A., A. Blandin, and R. Rogerson (2022). Hours and wages. The Quarterly Journal of Economics. Forthcoming.

Breunig, C. and S. Hoderlein (2018). Specification testing in random coefficient models. Quantitative Economics 9(3), 1371–1417.

Foster, A. and J. Hahn (2000). A consistent semiparametric estimation of the consumer surplus distribution. Economics Letters 69(3), 245–251.

Hausman, J. A. and W. K. Newey (1995). Nonparametric estimation of exact consumers surplus and deadweight loss. Econometrica 63(6), 1445–1476.

Hoderlein, S., H. Holzmann, and A. Meister (2017). The triangular model with random coefficients. Journal of Econometrics 201(1), 144–169.

Hoderlein, S., J. Klemelä, and E. Mammen (2010). Analyzing the random coefficient model nonparametrically. Econometric Theory 26(3), 804–837.

Lemieux, T. (2006). Postsecondary education and increasing wage inequality. American Economic Review 96(2), 195–199.

(LIMITED) DEGREE OF HETEROGENEITY IN MOMENTS

- Let $\boldsymbol{\phi}_i = (\beta_i, \boldsymbol{\gamma}')'$, $\boldsymbol{\phi} = E(\boldsymbol{\phi}_i) = (E(\beta_i), \boldsymbol{\gamma}')'$, then

$$\mathrm{E}(\boldsymbol{w}_i y_i) = \mathrm{E}\left(\boldsymbol{w}_i \boldsymbol{w}_i'\right) \phi.$$

(LIMITED) DEGREE OF HETEROGENEITY IN MOMENTS

- Let $\boldsymbol{\phi}_i = (\beta_i, \boldsymbol{\gamma}')'$, $\boldsymbol{\phi} = E(\boldsymbol{\phi}_i) = (E(\beta_i), \boldsymbol{\gamma}')'$, then

$$\mathrm{E}(\mathbf{w}_i y_i) = \mathrm{E}\left(\mathbf{w}_i \mathbf{w}_i'\right) \phi.$$

- Average over i,

$$\frac{1}{n}\sum_{i=1}^{n} E(\mathbf{w}_i y_i) = \left[\frac{1}{n}\sum_{i=1}^{n} E\left(\mathbf{w}_i \mathbf{w}_i'\right)\right] \phi$$

(LIMITED) DEGREE OF HETEROGENEITY IN MOMENTS

- Let $\boldsymbol{\phi}_i = (\beta_i, \boldsymbol{\gamma}')'$, $\boldsymbol{\phi} = E(\boldsymbol{\phi}_i) = (E(\beta_i), \boldsymbol{\gamma}')'$, then

$$\mathrm{E}(\boldsymbol{w}_i y_i) = \mathrm{E}\left(\boldsymbol{w}_i \boldsymbol{w}_i'\right) \phi.$$

- Average over i,

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(\boldsymbol{w}_i y_i)}_{\to \boldsymbol{q}_{wy}} = \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(\boldsymbol{w}_i \boldsymbol{w}_i'\right)\right]}_{\to \boldsymbol{Q}_{ww}} \boldsymbol{\phi}$$

(LIMITED) DEGREE OF HETEROGENEITY IN MOMENTS

- Let $\boldsymbol{\phi}_i = (\beta_i, \boldsymbol{\gamma}')'$, $\boldsymbol{\phi} = E(\boldsymbol{\phi}_i) = (E(\beta_i), \boldsymbol{\gamma}')'$, then

$$\mathrm{E}(\mathbf{w}_i y_i) = \mathrm{E}\left(\mathbf{w}_i \mathbf{w}_i'\right) \phi.$$

- Average over i,

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(\mathbf{w}_i y_i)}_{\to \boldsymbol{q}_{wy}} = \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(\mathbf{w}_i \mathbf{w}_i'\right)\right]}_{\to \boldsymbol{Q}_{ww}} \boldsymbol{\phi}$$

- Let $n \to \infty$, $\boldsymbol{\phi}$ is identified as

$$\boldsymbol{\phi} = \boldsymbol{Q}_{ww}^{-1} \boldsymbol{q}_{wy}$$

Back

(LIMITED) DEGREE OF HETEROGENEITY IN MOMENTS

- Let $\tilde{y}_i = y_i - \mathbf{z}_i'\boldsymbol{\gamma}$. Assumptions in use:

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left(\tilde{y}_i^r x_i^s\right) - \rho_{r,s} \right| = O\left(n^{-1/2}\right),$$

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left(u_i^r\right) - \sigma_r \right| = O\left(n^{-1/2}\right).$$

(LIMITED) DEGREE OF HETEROGENEITY IN MOMENTS

- Let $\tilde{y}_i = y_i - \mathbf{z}_i'\boldsymbol{\gamma}$. Assumptions in use:

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left(\tilde{y}_i^r x_i^s\right) - \rho_{r,s} \right| = O\left(n^{-1/2}\right),$$

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left(u_i^r\right) - \sigma_r \right| = O\left(n^{-1/2}\right).$$

- Example. Let $e_{ir} = \mathrm{E}\left(u_i^r\right) - \sigma_r$,

$$\left| n^{-1} \sum_{i=1}^{n} \mathrm{E}\left(u_i^r\right) - \sigma_r \right| \leq n^{-1} \sum_{i=1}^{n} |e_{ir}| \leq O\left(n^{-1/2}\right)$$

if $\sum_{i=1}^{n} |e_{ir}| = O\left(n^{\alpha_r}\right)$ with $\alpha_r < \frac{1}{2}$, where α_r measures the degree of heterogeneity. Back