Lecture 3

Matthew Osborne

8 September 2006

Partitioned Fit

- How does OLS break $\hat{\boldsymbol{\mu}}$ into $\mathbf{X}_k\widehat{\beta}_k$?

- Assume $\mathbf{X}$ is full column rank (3.1).

- Consider $\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2$.

- Interpret $\hat{\boldsymbol{\beta}}_1$:

  1. Take out part of $\mathbf{y}$ and $\mathbf{X}_1$ collinear with $\mathbf{X}_2$.

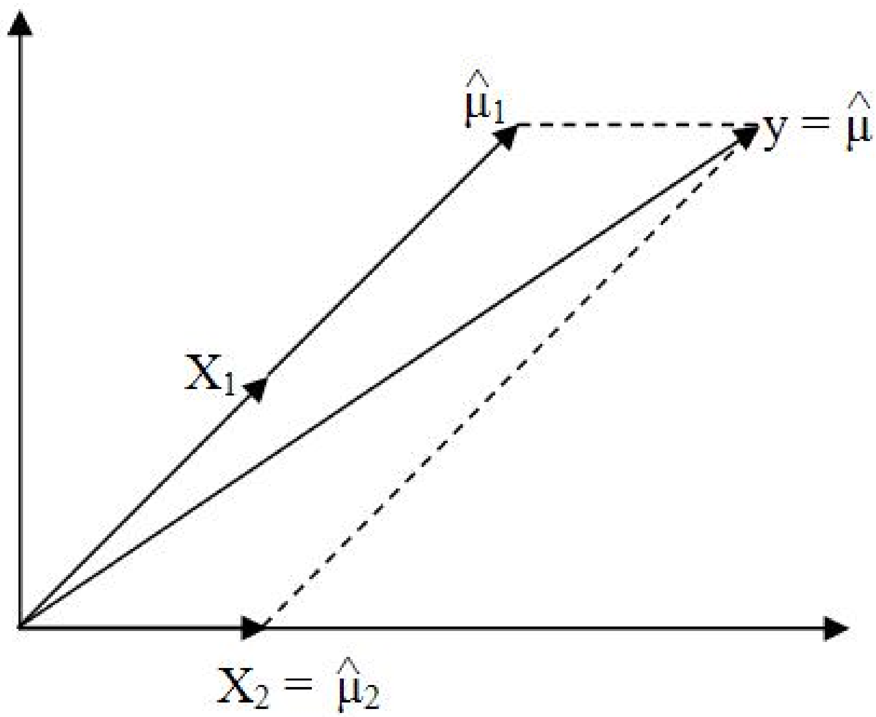
  2. Fit transformed $\mathbf{y}$ on transformed $\mathbf{X}_1$.

Partition $\boldsymbol{\mu}$ as follows:

$$\boldsymbol{\mu} \equiv \mathrm{X}\boldsymbol{\beta} = \begin{bmatrix} \mathrm{X}_1 & \mathrm{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} = \mathrm{X}_1\boldsymbol{\beta}_1 + \mathrm{X}_2\boldsymbol{\beta}_2 = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2$$

An Example

Let $\mathbf{y} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Let $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.



- We see graphically how to get $\hat{\boldsymbol{\mu}}_2$.

- Want: projection $\mathbf{P}\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_2$.

- Mathematically: annihilate cols of $\mathbf{X}_1$.

$$\mathbf{P}_{\mathbf{X}_1} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' =$$

$$\mathbf{I} - \mathbf{P}_{\mathbf{X}_1} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' = \mathbf{I} - \qquad\qquad =$$

$$(\mathbf{I} - \mathbf{P}_{\mathbf{X}_1})\mathbf{X}_1 =$$

- We need to preserve the columns of $\mathbf{X}_2$.

$$\mathbf{P} = \mathbf{X}_2[\mathbf{X}_2'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_1})\mathbf{X}_2]^{-1}\mathbf{X}_2'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_1})$$

$$[\mathbf{X}_2'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_1})\mathbf{X}_2] =$$

$$\mathbf{X}_2'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_1}) =$$

- So $\mathbf{P} =$

- Notice $\mathbf{P}\mathbf{X}_2 =$

- Notice $\mathbf{P}\hat{\boldsymbol{\mu}} =$

Partitioned Fit

Prop. 2: If Assumption (3.1) holds, and

$$\mathbf{X}_{1\perp 2} \equiv (\mathbf{I} - \mathbf{P}_{\mathrm{X}2})\mathbf{X}_1$$
$$\mathbf{y}_{1\perp 2} \equiv (\mathbf{I} - \mathbf{P}_{\mathrm{X}2})\mathbf{y}$$
$$\mathbf{P}_{\mathrm{X}2} \equiv \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'$$

then

1. The OLS fitted vector $\hat{\boldsymbol{\mu}}_1 \equiv \mathbf{X}_1\hat{\boldsymbol{\beta}}_1$ is the unique projection of $\mathbf{y}$ (or $\hat{\boldsymbol{\mu}}$) onto Col($\mathbf{X}_1$) s.t. Col($\mathbf{X}_2$) and Col$^{\perp}$($\mathbf{X}$) are annihilated. Furthermore, $\hat{\boldsymbol{\mu}}_1 = \mathbf{P}_{12}\mathbf{y} = \mathbf{P}_{12}\hat{\boldsymbol{\mu}}$, where the unique projector is

$$\mathbf{P}_{12} = \mathbf{X}_1(\mathbf{X}_{1\perp 2}'\mathbf{X}_1)^{-1}\mathbf{X}_{1\perp 2}'$$

2. The OLS fitted coefficients $\hat{\boldsymbol{\beta}}_1$ are

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_{1\perp 2}'\mathbf{X}_{1\perp 2})^{-1}\mathbf{X}_{1\perp 2}'\mathbf{y}_{\perp 2}$$

- Fitted Values of RHS regressions:

$$\mathbf{P_{X_2}X_1} = \left[\ \mathbf{P_{X_2}X_{11}} \quad \mathbf{P_{X_2}X_{12}} \quad \cdots \quad \mathbf{P_{X_2}X_{1K}}\ \right]$$

- Residuals: $\mathbf{X}_{1\perp 2} \equiv (\mathbf{I}$ - $\mathbf{P}_{\mathrm{X}_2})\mathbf{X}_1$.

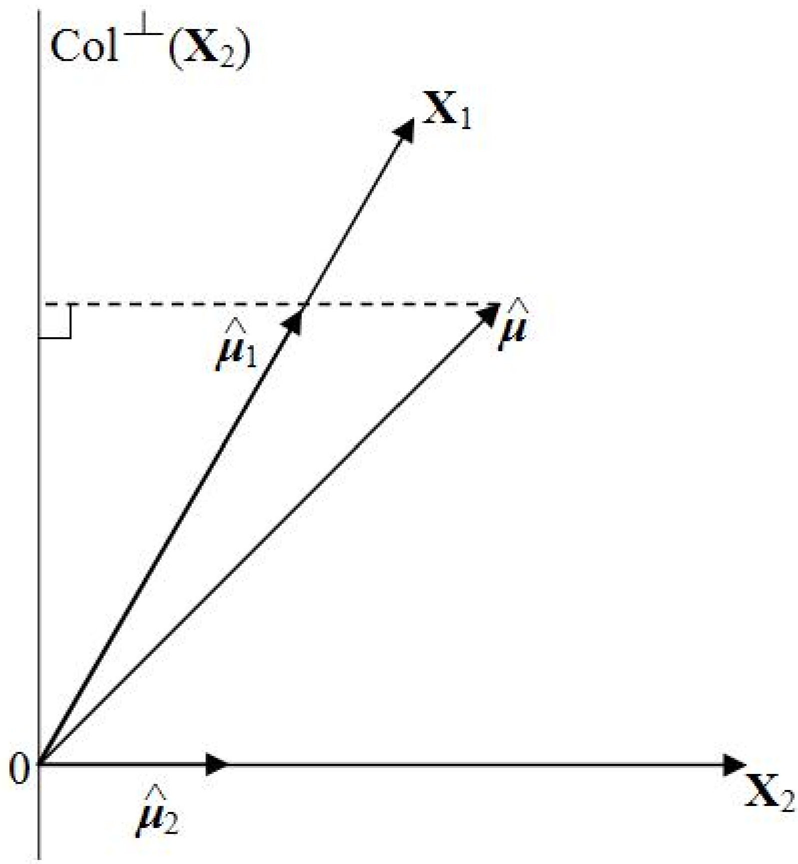
- Fitted vector $\hat{\boldsymbol{\mu}}_1 = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 = \mathbf{P}_{12}\hat{\boldsymbol{\mu}}$:

$$\begin{aligned}
&\mathbf{P}_{\mathbf{X}_1} \equiv \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' \\
&\mathbf{P}_{12} = \mathbf{X}_1(\mathbf{X}_{1\perp 2}'\mathbf{X}_1)^{-1}\mathbf{X}_{1\perp 2}' \\
&\mathbf{X}_{1\perp 2} \equiv (\mathbf{I} - \mathbf{P}_{\mathbf{X}_2})\mathbf{X}_1
\end{aligned}$$

- $\mathbf{P}_{12}\ \mathbf{X}_1 = \mathbf{X}_1$; $\mathbf{P}_{12}\ \mathbf{X}_2 = \mathbf{0}$.

- So $\mathbf{P}_{12}\hat{\boldsymbol{\mu}} = \mathbf{P}_{12}(\hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\mu}}_2) = \hat{\boldsymbol{\mu}}_1$.

$$\mathbf{P}_{12} = \mathbf{X}_1[\mathbf{X}_1'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_2})\mathbf{X}_1]^{-1}\mathbf{X}_1'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_2})$$



Proof of Prop 2, Part 2

Orthogonal decomposition of **y** - $\boldsymbol{\mu}$:

$$\begin{aligned}
\mathbf{y} - \boldsymbol{\mu} &= (\mathbf{I} - \mathbf{P}_{\mathrm{X2}})(\mathbf{y} - \boldsymbol{\mu}) + \mathbf{P}_{\mathrm{X2}}(\mathbf{y} - \boldsymbol{\mu}) \\
||\mathbf{y} - \boldsymbol{\mu}||^2 &= ||(\mathbf{I} - \mathbf{P}_{\mathrm{X2}})(\mathbf{y} - \boldsymbol{\mu})||^2 + ||\mathbf{P}_{\mathrm{X2}}(\mathbf{y} - \boldsymbol{\mu})||^2
\end{aligned}$$

$$\begin{aligned}
(\mathbf{I} - \mathbf{P}_{\mathrm{X2}})(\mathbf{y} - \boldsymbol{\mu}) &= (\mathbf{I} - \mathbf{P}_{\mathrm{X2}})\mathbf{y} - (\mathbf{I} - \mathbf{P}_{\mathrm{X2}})\mathbf{X}_1\boldsymbol{\beta}_1 - (\mathbf{I} - \mathbf{P}_{\mathrm{X2}})\mathbf{X}_2\boldsymbol{\beta}_2 \\
&= \mathbf{y}_{\perp 2} - \mathbf{X}_{1\perp 2}\boldsymbol{\beta}_1
\end{aligned}$$

$$||\mathbf{y} - \boldsymbol{\mu}||^2 = ||\mathbf{y}_{\perp 2} - \mathbf{X}_{1\perp 2}\boldsymbol{\beta}_1||^2 + ||\mathbf{P}_{\mathrm{X2}}(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1) - \mathbf{X}_2\boldsymbol{\beta}_2||^2$$

2nd term can be made zero:

$$\begin{aligned}
\min_{\beta} ||\mathbf{y} - \mathbf{X}\boldsymbol{\beta}||^2 &= \min_{\boldsymbol{\beta}_1}\{||\mathbf{y}_{\perp 2} - \mathbf{X}_{1\perp 2}\boldsymbol{\beta}_1||^2 \\
&\quad + \min_{\boldsymbol{\beta}_2} ||\mathbf{P}_{\mathrm{X2}}(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1) - \mathbf{X}_2\boldsymbol{\beta}_2||^2\} \\
&= \min_{\boldsymbol{\beta}_1} ||\mathbf{y}_{\perp 2} - \mathbf{X}_{1\perp 2}\boldsymbol{\beta}_1||^2
\end{aligned}$$

Proof of Prop 2, Part 2

Prop. 1 tells us we can solve $\min_{\boldsymbol{\beta}_1} ||\mathbf{y}_{\perp 2} - \mathbf{X}_{1\perp 2}\boldsymbol{\beta}_1||^2$, if $\mathbf{X}_{1\perp 2}$ is full rank.

$$\mathbf{X} = (\mathbf{I} - \mathbf{P}_{\mathrm{X}2} + \mathbf{P}_{\mathrm{X}2})\mathbf{X} = \left[\ \mathbf{X}_{1\perp 2} + \mathbf{P}_{\mathrm{X}2}\mathbf{X}_1 \quad \mathbf{X}_2\ \right]$$

$\mathbf{X}$ is full rank. Since $\mathbf{P}_{\mathrm{X}2}$ $\mathbf{X}_1$ is l.d. with $\mathbf{X}_2$, so $\mathbf{X}_{1\perp 2}$ must be full rank.

Thus,

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}'_{1\perp 2}\mathbf{X}_{\mathbf{1\perp 2}})^{-\mathbf{1}}\mathbf{X}'_{\mathbf{1\perp 2}}\mathbf{y}_{\perp \mathbf{2}}$$

Back to our Example

- Regression of $\mathbf{X}_2$ on $\mathbf{X}_1$: $\mathbf{P}_{\mathbf{X}_1}\mathbf{X}_2 = \begin{bmatrix} \quad \end{bmatrix}$.

- Residual $= \mathbf{X}_{2\perp 1} = \begin{bmatrix} \quad \end{bmatrix}$

- $\mathbf{y}_{\perp 1} = \begin{bmatrix} \quad \end{bmatrix} - \begin{bmatrix} \quad \end{bmatrix} = \begin{bmatrix} \quad \end{bmatrix}$.

- $\hat{\boldsymbol{\beta}}_2 =$

- Alternatively, since $\mathbf{X}_2\widehat{\beta}_2 = \hat{\boldsymbol{\mu}}_2$,

$$\widehat{\beta}_2 = (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\hat{\boldsymbol{\mu}}_2.$$

Generalized Distance

- Notice that we can write:

$$
\begin{aligned}
&||\mathbf{y}_{\perp 2} - \mathbf{X}_{1\perp 2}\boldsymbol{\beta}_1||^2 = ||(\mathbf{I} - \mathbf{P}_{\mathbf{X}2})(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1)||^2 \\
&= (\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1)'(\mathbf{I} - \mathbf{P}_{\mathbf{X}2})'(\mathbf{I} - \mathbf{P}_{\mathbf{X}2})(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1) \\
&= (\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1)'(\mathbf{I} - \mathbf{P}_{\mathbf{X}2})(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1)
\end{aligned}
$$

- Generalization of OLS:

$$
\mathbf{X}_1(\mathbf{X}_1'\mathbf{A}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{A}\mathbf{y} = \underset{\mathbf{z}\in \mathrm{Col}(\mathbf{X}_1)}{\operatorname{argmin}} (\mathbf{y} - \mathbf{z})'\mathbf{A}(\mathbf{y} - \mathbf{z})
$$

- Let $\mathbf{Z} \equiv \mathbf{A}'\mathbf{X}$.

$$
\mathbf{X}(\mathbf{X}'\mathbf{A}\mathbf{X})^{-1}\mathbf{X}'\mathbf{A} = \mathbf{X}(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'
$$

- i) Preserves Col($\mathbf{X}$); ii) Annihilates Col$^{\perp}$($\mathbf{Z}$).

Generalized Projectors

Def: Direct Sum: Let $\mathbb{S}_1$ and $\mathbb{S}_2$ be two disjoint vector subspaces of $\mathbb{R}^N$ so that $\mathbb{S}_1 \cap \mathbb{S}_2 = \{\mathbf{0}\}$. The vector space

$$\mathbb{V} = \left\{ \mathbf{z} \in \mathbb{R}^N | \mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2, \mathbf{z}_1 \in \mathbb{S}_1, \mathbf{z}_2 \in \mathbb{S}_2 \right\}$$

is called the direct sum of $\mathbb{S}_1$ and $\mathbb{S}_2$ and it is denoted by $\mathbb{Z} = \mathbb{S}_1 \oplus \mathbb{S}_2$.

Def: Projector: Let $\mathbb{R}^N$ be the direct sum of two linear subspaces $\mathbb{S}_1$ and $\mathbb{S}_2$. Let $\mathbf{z} \in \mathbb{R}^N$ so that $\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2$ for unique $\mathbf{z}_i \in \mathbb{S}_i, (i = 1, 2)$. Then $\mathbf{P}$ is a projector onto $\mathbb{S}_1$ along $\mathbb{S}_2$ if $\mathbf{Pz} = \mathbf{z}_1$ for all $\mathbf{z}$.

Lemma: The projector $\mathbf{P}$ onto $\mathbb{S}_1$ along $\mathbb{S}_2$ is unique.

Generalized Projectors

Lemma: Assume $\dim(\text{Col}(\mathbf{X})) = K$. The matrix $\mathbf{P}_{12} = \mathbf{X}_1(\mathbf{X}_{1\perp 2}{}'\mathbf{X}_1)^{-1}\mathbf{X}_{1\perp 2}{}'$ is the unique projector onto $\text{Col}(\mathbf{X}_1)$ along $\text{Col}(\mathbf{X}_2) \oplus \text{Col}^{\perp}(\mathbf{X})$.

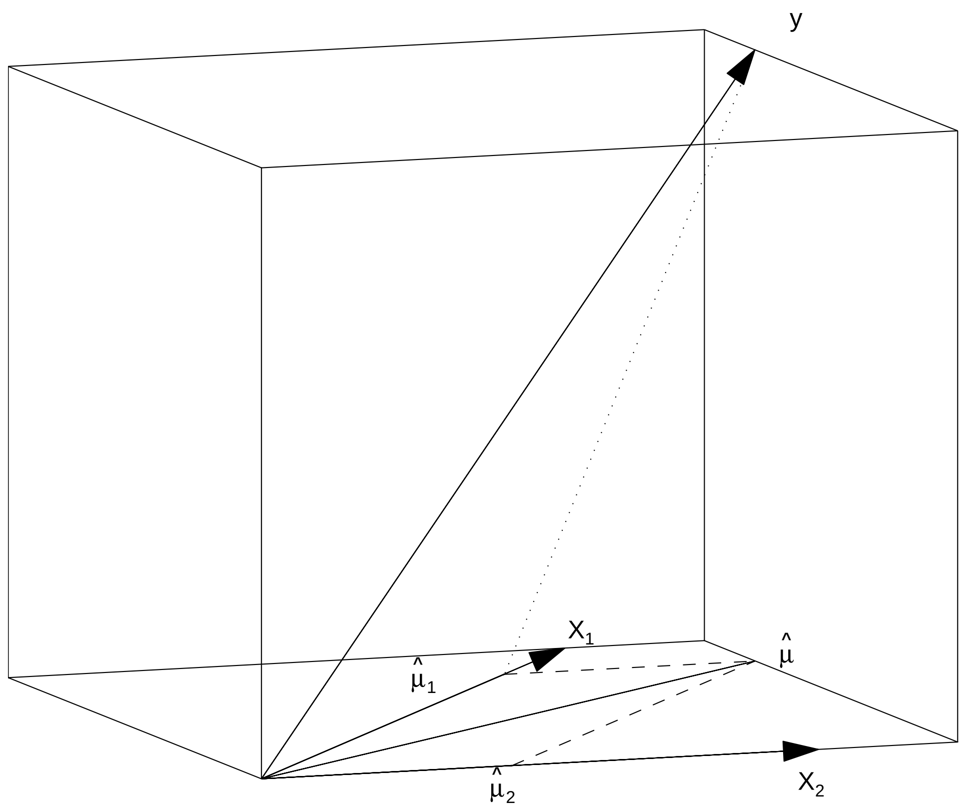
Proof: Since $\mathbf{X}$ is full rank we can write

$$\begin{aligned}\mathbb{R}^N &= \text{Col}(\mathbf{X}) \oplus \text{Col}^{\perp}(\mathbf{X}) \\ &= \text{Col}(\mathbf{X}_1) \oplus [\text{Col}(\mathbf{X}_2) \oplus \text{Col}^{\perp}(\mathbf{X})]\end{aligned}$$

We showed $\mathbf{P}_{12}$ preserves $\text{Col}(\mathbf{X}_1)$ and annihilates $\text{Col}(\mathbf{X}_2)$. Need to show annihilates $\text{Col}^{\perp}(\mathbf{X})$.

$$\mathbf{X}_{1\perp 2}'\mathbf{z} = \mathbf{X}_1'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_2})'\mathbf{z} = \mathbf{X}_1'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_2})\mathbf{z} = \mathbf{X}_1'\mathbf{z} = \mathbf{0}$$

$\mathbf{P}_{12}$ applied to $\mathbf{y}$:



Proof of Prop 2, Part 1

$$
\begin{aligned}
\hat{\boldsymbol{\mu}}_1 &\equiv \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 \\
&= \mathbf{X}_1(\mathbf{X}'_{1\perp 2}\mathbf{X}_{1\perp 2})^{-1}\mathbf{X}_{1\perp 2} \\
&= \mathbf{X}_1(\mathbf{X}'_1(\mathbf{I}-\mathbf{P}_{\mathbf{X}_2})'(\mathbf{I}-\mathbf{P}_{\mathbf{X}_2})\mathbf{X}_1)^{-1}\mathbf{X}'_1(\mathbf{I}-\mathbf{P}_{\mathbf{X}_2})'(\mathbf{I}-\mathbf{P}_{\mathbf{X}_2})\mathbf{y} \\
&= \mathbf{X}_1(\mathbf{X}'_1(\mathbf{I}-\mathbf{P}_{\mathbf{X}_2})\mathbf{X}_1)^{-1}\mathbf{X}'_1(\mathbf{I}-\mathbf{P}_{\mathbf{X}_2})\mathbf{y} \\
&= \mathbf{P}_{12}\mathbf{y}
\end{aligned}
$$

- Previous Lemma showed $\mathbf{P}_{12}\mathbf{y}$ is unique projection onto Col($\mathbf{X}_1$) annihilating Col($\mathbf{X}_2$) and Col$^{\perp}$($\mathbf{X}$).

- So $\mathbf{P}_{12}\mathbf{y} = \mathbf{P}_{12}(\mathbf{y}$ - $\hat{\mu} + \hat{\mu}) = \mathbf{P}_{12}\hat{\boldsymbol{\mu}}$.

Why we Might Care

- Panel of 10,000 individuals, 5 periods.

$$\mathbf{x}'_{nt}\boldsymbol{\beta} = \sum_{k=1}^{K_1} x_{ntk}\beta_{1k} + \sum_{k=1}^{10{,}000} d_{ntk}\beta_{2k}; \quad d_{ntk} = \begin{cases} 0 & \text{if } n \neq k \\ 1 & \text{if } n = k \end{cases}$$

- 10,000 dummy variables!

- Solution: let $[\mathbf{X}_2] = [d_{ntk}]$.

- Residuals: $\begin{array}{l} \mathbf{X}_{1\perp 2} = [x_{ntk} - \frac{1}{5}\sum_{t=1}^{5} x_{ntk}] \\ \mathbf{y}_{\perp 2} = [y_{nt} - \frac{1}{5}\sum_{t=1}^{5} y_{nt}] \end{array}$

- Do OLS with only K_1 regressors.

An Application

- Impact of reputation on price in eBay auctions.

- Reputation has long been important in theory, but not much empirical research.

- Theory: 1 seller, n bidders, independent private values, $v_i > 0$.

- $r^S \in (0, 1]$ - probability seller pays (reputation).

- $r^B \in (0, 1]$ - probability buyer pays (reputation).

- Expected utility for winning bidder:

$$r^B(r^S v_i - b)$$

- Proxy bids $(b_1, ..., b_n)$.

- i has high bid if $b_i > b_j \forall j \neq i$. High bid $= \max_{k \neq i} b_k$.

- Bidder can increase proxy bids until end of auction.

- Assume bidders always have time to increase bids.

Equilibrium

- A pair $\{(b_1^*, ..., b_n^*), i^*\}$, $b_i^* = \max_k b_k^*$.

- Bidding ends when: $\max_{k \neq i} b_k^* \geq r^s v_j \quad \forall j \neq i^*$.

- Nobody bids above his expected value: $b_j^* \leq r^S v_j \quad \forall j$.

- Prop: Order bidders according to their values.

  1. $i^* = 1$.

  2. $\max_{k \neq 1} b_k^* = b_2 = r^S v_2$; An equilibrium is $\{(r^S v_1, ..., r^S v_n), 1\}$.

Empirical Method

- $b^*_{i2} = r_i^S v_{i2}$; $\log(b^*_{i2}) = \log(r_i^S) + \log(v_{i2})$.

$$
\begin{aligned}
r_i^S &= \lambda x_{i1}^{\theta_1} x_{i2}^{\theta_2} ... x_{iK}^{\theta_K}; \quad \lambda, x_{ij} > 0 \\
v_{i2} &= \phi y_{i1}^{\pi_1} y_{i2}^{\pi_2} ... y_{iM}^{\pi_M} e^{\eta_{i2}}; \quad y_{ij} > 0
\end{aligned}
$$

- η_{i2} is a random error (we haven't really discussed this yet).

- Rewrite model as follows, where ε_{i2} is the residual.

$$
\begin{aligned}
&\log(b^*_{i2}) = c + \tilde{x}_i'\theta + \tilde{y}_i'\pi + \alpha_{t_i} + \varepsilon_{i2} \\
&\varepsilon_{i2} = \eta_{i2} - E(\eta_{i2}|t_i),\ E(\varepsilon_{i2}) = 0,\ Var(\varepsilon_{i2}) = \sigma^2_{t_i}
\end{aligned}
$$

Hedonic Regression

- $p(z) = z\beta + \varepsilon$. β's WTP, implicit prices.

- Utility: $U(x, z_1, ..., z_n)$, $I = x + p(z)$.

- Consumer FOC $\rightarrow \frac{\partial p}{\partial z} = \frac{\partial U}{\partial z} / \frac{\partial U}{\partial x}$ (Rosen (1974)).

- Firm problem implies $\frac{\partial p}{\partial z}$ = MC per unit sold.

- Used to measure WTP for clean air, public schools, house attributes.

- Research issues: dynamic issues (ie migration), omitted variables (economic activity).

Data Set

- eBay auctions for Pentium III 500 cpus. 95 auctions that closed between Sept 23, 1999 and Dec 18, 1999.

- Dependent Var: *SecondPrice* = winning bid plus shipping cost.

- Reputation: Shades, number of pos, neg, and neutral feedbacks.

- Item characteristics: Market price, Visa, Used, Retail.

- Auction characteristics: Exclude, len5, len7, len10.

- Summary statistics: know your data!

Results

- Coefficients of the fitted vector, $\hat{\boldsymbol{\mu}}$.

- Interpretation: elasticities.

- Are the results economically significant?

- Increase in reputation from 0 to 15 increases price by $12.

- Extrapolate results.