Discussion of paper
"Scalable Nonparametric Price Elasticity Estimation"
by *Wang and Huang*

Discussion by Vineet Kumar (Yale School of Management)

March 2023
UTD FORMS

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
    - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
    - Minimal assumptions – nonparametric
    - No demand model required even
    - Closed form for function, conditional on function values at nearest neighbors
    - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
    - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
    - Minimal assumptions – nonparametric
    - No demand model required even
    - Closed form for function, conditional on function values at nearest neighbors
    - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\Longrightarrow$ Computationally light and scalable

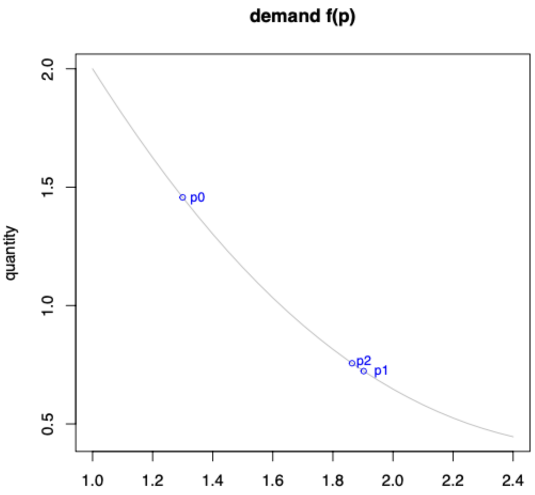
- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\implies$ Computationally light and scalable

- Motivation: Obtain **price elasticities** from aggregate data.
- Our typical approach to demand models is parametric (e.g. logit)
- Want to do this using a nonparametric approach
  - Why?
- Use ML method, specifically bagged (bootstrap averaged) nearest neighbors
- Desiderata:
  - Minimal assumptions – nonparametric
  - No demand model required even
  - Closed form for function, conditional on function values at nearest neighbors
  - $\Longrightarrow$ Computationally light and scalable

How it works – with exogenous prices

- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
    - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
    - Use a ML method to obtain functional approximation
    - Universal function approximators for different classes of functions

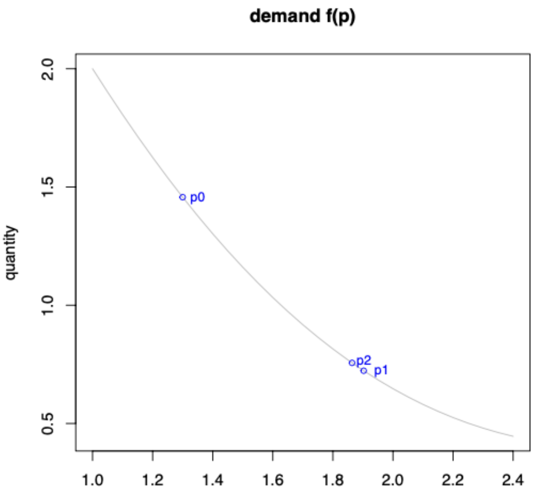
Endogenous prices are a challenge



How it works – with exogenous prices

- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
  - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
  - Use a ML method to obtain functional approximation
  - Universal function approximators for different classes of functions

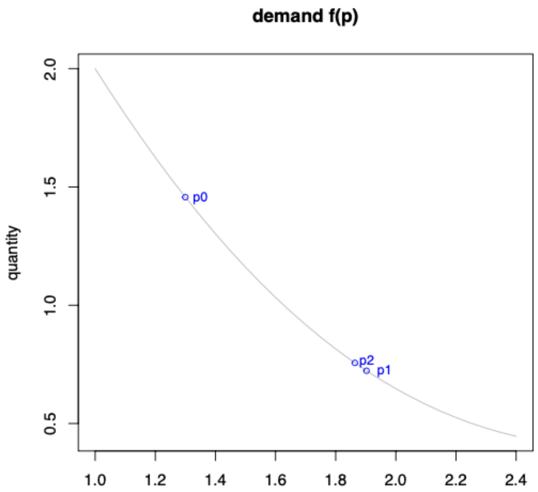
Endogenous prices are a challenge



How it works – with exogenous prices

- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
  - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
  - Use a ML method to obtain functional approximation
  - Universal function approximators for different classes of functions

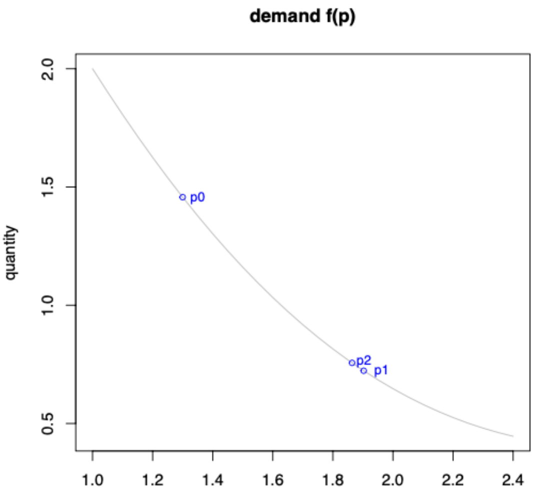
Endogenous prices are a challenge



How it works – with exogenous prices

- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
  - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
  - Use a ML method to obtain functional approximation
  - Universal function approximators for different classes of functions

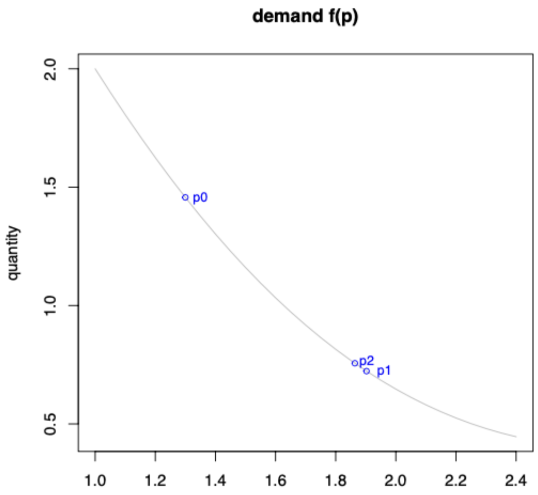
Endogenous prices are a challenge



How it works – with exogenous prices

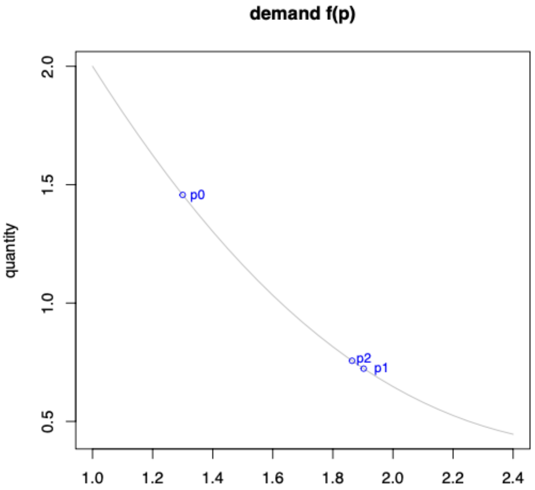
- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
  - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
  - Use a ML method to obtain functional approximation
  - Universal function approximators for different classes of functions

$\boxed{\text{Endogenous prices are a challenge}}$



How it works – with exogenous prices

- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
  - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
  - Use a ML method to obtain functional approximation
  - Universal function approximators for different classes of functions

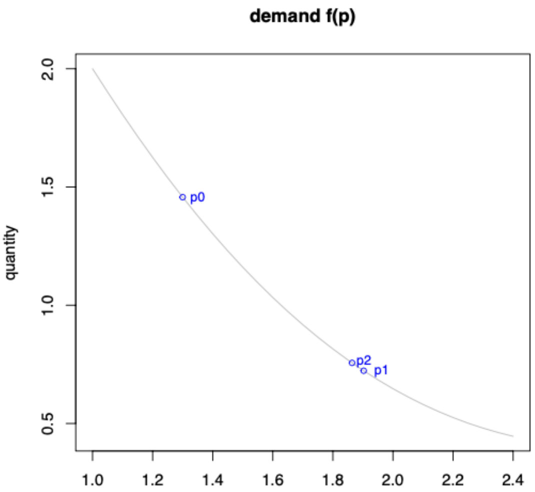


$\boxed{\text{Endogenous prices are a challenge}}$

How it works – with exogenous prices

- Consider simple case without endogeneity
- Model: $\boxed{s_{jt} = f_j(p_{jt}, x_{jt}) + \varepsilon_{jt}}$
- Elasticities: If you know function f, then we can easily obtain elasticities
  - Change price to $p + \Delta$, determine how sales change.
- How do you get f nonparametrically?
  - Use a ML method to obtain functional approximation
  - Universal function approximators for different classes of functions

$\boxed{\text{Endogenous prices are a challenge}}$



How it works – Dealing with endogeneity

- Control function approach is used to deal with endogeneity
- $\boxed{\mathbf{p = g(z) + u}}$
- where z's are instruments
- Want to obtain price derivatives: $\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \partial_p \mathbf{E}[\varepsilon_{jt}|p,x]$
- Newey, Powell and Vella (1999) provide a way to write the selection bias to get: $\partial_p \mathbf{E}[\varepsilon_{jt}|p,x] = (\partial_z g(z)' \partial_z g(z))^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$
- All quantities can be easily obtained from data

How it works – Dealing with endogeneity

- Control function approach is used to deal with endogeneity
- $\boxed{\mathbf{p} = \mathbf{g}(\mathbf{z}) + \mathbf{u}}$
- where z's are instruments
- Want to obtain price derivatives: $\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \partial_p \mathbf{E}[\varepsilon_{jt}|p,x]$
- Newey, Powell and Vella (1999) provide a way to write the selection bias to get: $\partial_p \mathbf{E}[\varepsilon_{jt}|p,x] = (\partial_z g(z)' \partial_z g(z))^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$
- All quantities can be easily obtained from data

How it works – Dealing with endogeneity

- Control function approach is used to deal with endogeneity
- $\boxed{\mathbf{p} = \mathbf{g}(\mathbf{z}) + \mathbf{u}}$
- where z's are instruments
- Want to obtain price derivatives: $\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \partial_p \mathbf{E}[\varepsilon_{jt}|p,x]$
- Newey, Powell and Vella (1999) provide a way to write the selection bias to get: $\partial_p \mathbf{E}[\varepsilon_{jt}|p,x] = (\partial_z g(z)' \partial_z g(z))^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$
- All quantities can be easily obtained from data

How it works – Dealing with endogeneity

- Control function approach is used to deal with endogeneity
- $\boxed{\mathbf{p} = \mathbf{g}(\mathbf{z}) + \mathbf{u}}$
- where z's are instruments
- Want to obtain price derivatives: $\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \partial_p \mathbf{E}[\varepsilon_{jt}|p,x]$
- Newey, Powell and Vella (1999) provide a way to write the selection bias to get: $\partial_p \mathbf{E}[\varepsilon_{jt}|p,x] = (\partial_z g(z)'\partial_z g(z))^{-1}\partial_z g(z)'\partial_z \mathbf{E}[s_j|p,x]$
- All quantities can be easily obtained from data

How it works – Dealing with endogeneity

- Control function approach is used to deal with endogeneity
- $\boxed{\mathbf{p} = \mathbf{g}(\mathbf{z}) + \mathbf{u}}$
- where z's are instruments
- Want to obtain price derivatives: $\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \partial_p \mathbf{E}[\varepsilon_{jt}|p,x]$
- Newey, Powell and Vella (1999) provide a way to write the selection bias to get: $\partial_p \mathbf{E}[\varepsilon_{jt}|p,x] = (\partial_z g(z)' \partial_z g(z))^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$
- All quantities can be easily obtained from data

How it works – Dealing with endogeneity

- Control function approach is used to deal with endogeneity
- $\boxed{\mathbf{p} = \mathbf{g}(\mathbf{z}) + \mathbf{u}}$
- where z's are instruments
- Want to obtain price derivatives: $\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \partial_p \mathbf{E}[\varepsilon_{jt}|p,x]$
- Newey, Powell and Vella (1999) provide a way to write the selection bias to get: $\partial_p \mathbf{E}[\varepsilon_{jt}|p,x] = \left(\partial_z g(z)' \partial_z g(z)\right)^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$
- All quantities can be easily obtained from data

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)' \partial_z g(z)\right)^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)'\partial_z g(z)\right)^{-1} \partial_z g(z)'\partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)'\partial_z g(z)\right)^{-1} \partial_z g(z)'\partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has “fast” asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)'\partial_z g(z)\right)^{-1} \partial_z g(z)'\partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)' \partial_z g(z)\right)^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)' \partial_z g(z)\right)^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has “fast” asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let’s start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)' \partial_z g(z)\right)^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)' \partial_z g(z)\right)^{-1} \partial_z g(z)' \partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)'\partial_z g(z)\right)^{-1} \partial_z g(z)'\partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has “fast” asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let’s start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

What is the Role of ML here?

- One view of ML is that it is functional approximation
- If we have $\partial_p f_j(p,x)$ we can get elasticities
- Want to approximate

$$\partial_p f_j(p,x) = \partial_p \mathbf{E}[s_j|p,x] - \left(\partial_z g(z)'\partial_z g(z)\right)^{-1} \partial_z g(z)'\partial_z \mathbf{E}[s_j|p,x]$$

- Want a method that has "fast" asymptotic convergence (rate $\sqrt{m/T}$)
- Why the complexity:
  - Let's start with k-Nearest Neighbor (Bias-Variance Tradeoff)
  - Ensemble methods like Bagging can reduce variance (think of Random Forest)
  - ...But still biased.
  - use a jackknife to correct the bias (Biau, Cerou, and Guyader, 2010)
  - corrects first order bias.

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Paper is well written and motivated, which is not common in methodological papers
- Method is new and applicable to pretty much any demand setting.
- One of the biggest issues is endogeneity, which is accommodated here.
- Structural models can often take days, weeks or even months to estimate
- Anything that helps with obtaining realistic estimates of **economic quantities** like elasticity very useful
- Critical to have when **real-time** or timely decisions are required
- Scalable to lots of products, markets

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
    - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
  - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
  - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
  - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
  - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
  - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Things to Like

- Nonparametric implies the model misspecification is not likely a huge issue
- Prove closed form for bagged nearest neighbors estimator
  - In parametric models, sometimes assumptions drive the result(rather than the data).
- Bagged Nearest Neighbor has attractive computational properties
- Not easy to obtain a method that has all the desirable properties
- Authors show monte carlos to show approximation quality
- Practical value in realistic applications...fast..

Thoughts – Model and Counterfactuals

- Let's start with profit-maximizing firm: $\pi(p) = f(x, p)(p - c)$
- In equilibrium, after solving FOC, Lerner index $\dfrac{p-c}{p} = -\dfrac{1}{\varepsilon(p)}$
- So, can use this to recover marginal costs if we know $\varepsilon(p)$.

Thoughts – Model and Counterfactuals

- Let's start with profit-maximizing firm: $\pi(p) = f(x,p)(p-c)$
- In equilibrium, after solving FOC, Lerner index $\frac{p-c}{p} = -\frac{1}{\varepsilon(p)}$
- So, can use this to recover marginal costs if we know $\varepsilon(p)$.

Thoughts – Model and Counterfactuals

- Let's start with profit-maximizing firm: $\pi(p) = f(x,p)(p-c)$
- In equilibrium, after solving FOC, Lerner index $\dfrac{p-c}{p} = -\dfrac{1}{\varepsilon(p)}$
- So, can use this to recover marginal costs if we know $\varepsilon(p)$.

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + +\omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + +\omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + +\omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + + \omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + + \omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
    - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + +\omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + +\omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- Typically we start with a utility model $\boxed{u_{ijt} = \beta X_{jt} + \alpha p_{jt} + \xi_{jt} + +\omega_{ijt}}$
- We recover parameters $\theta = (\alpha, \beta)$.
- We might expect the sales to j to be impacted by all the unobservables ξ_{kt}. So we should have $f(x, p, \xi)$.
- What is the impact of extant price variation?
- Local versus Global: when are we likely to have local be a good approximation to global as in Compani (2022)?
  - What classes of demand and hence profit functions?
- Here the method cannot recover preference parameters
- However, the claim is we can still do counterfactuals. How?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x, p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c + \delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\tilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c+\delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\tilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c + \delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\widetilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c + \delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\widetilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c + \delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\widetilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c+\delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\tilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c+\delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\widetilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x,p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c + \delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\widetilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – Model and Counterfactuals

- **Key Restriction**: Demand $f(x, p)$ that is recovered through ML should not change in the counterfactual
  - Demand invariance
- What class of counterfactuals would work here?
  - An excise tax would be covered but a sales tax would not. Why?
- Excise tax $\implies$ cost increases $c \to c + \delta$ but does not impact demand directly in any way.
- $\implies$ elasticities are not affected.
- However, what happens when we have a sales tax increase?
  - The elasticity recovered from the data $\varepsilon(p)$ is different from counterfactual $\widetilde{\varepsilon}(p)$
- What about something like Petrin (2002), minivan?

Thoughts – ML

1. How are the hyperparameters (tuning parameters) of the ML algorithm set? No theoretical guidelines for this.
2. What about the g function (recall $p = g(z) + u$). This needs to be parametrically specified?
3. Can we use any other ML method in place of Bagged nearest neighbors?
    - Random Forest, Deep nets etc.
4. Boils down to what you need from a computational perspective as well as a

Thoughts – ML

1. How are the hyperparameters (tuning parameters) of the ML algorithm set? No theoretical guidelines for this.
2. What about the g function (recall $p = g(z) + u$). This needs to be parametrically specified?
3. Can we use any other ML method in place of Bagged nearest neighbors?
   - Random Forest, Deep nets etc.
4. Boils down to what you need from a computational perspective as well as a

Thoughts – ML

1. How are the hyperparameters (tuning parameters) of the ML algorithm set? No theoretical guidelines for this.
2. What about the g function (recall $p = g(z) + u$). This needs to be parametrically specified?
3. Can we use any other ML method in place of Bagged nearest neighbors?
   - Random Forest, Deep nets etc.
4. Boils down to what you need from a computational perspective as well as a

Thoughts – ML

1. How are the hyperparameters (tuning parameters) of the ML algorithm set? No theoretical guidelines for this.
2. What about the g function (recall $p = g(z) + u$). This needs to be parametrically specified?
3. Can we use any other ML method in place of Bagged nearest neighbors?
   - Random Forest, Deep nets etc.
4. Boils down to what you need from a computational perspective as well as a

Thoughts – ML

1. How are the hyperparameters (tuning parameters) of the ML algorithm set? No theoretical guidelines for this.
2. What about the g function (recall $p = g(z) + u$). This needs to be parametrically specified?
3. Can we use any other ML method in place of Bagged nearest neighbors?
   - Random Forest, Deep nets etc.
4. Boils down to what you need from a computational perspective as well as a

Conclusion

1. Very interesting and practically valuable contribution to estimate elasticities in the presence of price endogeneity
2. Authors show that method can satisfy a number of desirable properties both in theory and in practice
3. Need to have invariance of demand curve in the counterfactual

Conclusion

1. Very interesting and practically valuable contribution to estimate elasticities in the presence of price endogeneity
2. Authors show that method can satisfy a number of desirable properties both in theory and in practice
3. Need to have invariance of demand curve in the counterfactual

Conclusion

1. Very interesting and practically valuable contribution to estimate elasticities in the presence of price endogeneity
2. Authors show that method can satisfy a number of desirable properties both in theory and in practice
3. Need to have invariance of demand curve in the counterfactual