Network Local Projections

Aureo de Paula[1] Edoardo Rainone[2]

[1]UCL
[2]Bank of Italy

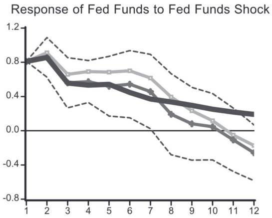
I thank Christian Brownlees, Vasco Carvalho, Oscar Jorda and Xiaodong Liu for comments. The views expressed are solely those of the author and do not necessarily represent those of the Bank of Italy. These slides are a sketch of a research idea, please do not circulate without permission. The illustration of the methodology closely follows the slides of a EABCN course held by prof. Oscar Jorda in Barcelona at UPF during September 2023. All the errors are my own.

Table of Contents

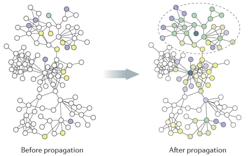
Table of Contents

Motivation

- **local projections** (LP) widely used in **time series** to estimate impulse responses (IRs)
  - see Jorda (2006), Plagborg-Moller and Wolf (2021), Jorda (2023) among others
- **IRs on cross-sectional dimension** are gaining interest, can NLP be a new tool?
  - measure spillovers, peer effects
  - on networks, space
  - eventually spatio-temporal settings



From AER Jorda (2006)



From Nature RG Cowen et al. (2017)

Potentially Interesting Applications

- **Production networks**
  - How shocks to suppliers affect customers and viceversa
  - Transmission of prices changes through **PN**

- **Financial networks**
  - How shocks to certain asset classes affect others
  - Contagion across intermediaries

- **Social networks**
  - Effects of cash transfers
  - Diffusion of adoption of new products

Preview of the preliminary results

- **what is different** between LP and NLP?
  - Forward and Backward
  - Endogenous effects ($y_{i-d} \rightarrow y_i$) may not be identified
  - Because of recursivity ($y_{i-d} \leftrightarrow y_i$)
  - Network Embedded IV can be used to identify them
  - NLP-SAR (as for LP-AR) relationship
  - NLP more robust to misspecification
  - but more demanding for identification
- were they **implicitly already used** in applied work?
  - Specifications close to NLP were used
  - Using production networks data
  - 2 close examples:
    - Carvalho et al. QJE (2021)
    - Huremovic et al. WP (2024)
    - Barrot and Sauvagnat QJE (2016)

Table of Contents

Impulse responses as a comparison of two 'averages'

Over time

$$R_{sy}(h, \delta) = E[y_{t+h} | u_t = u_0 + \delta, x_t] - E[y_{t+h} | u_t = u_0, x_t] \qquad (1)$$

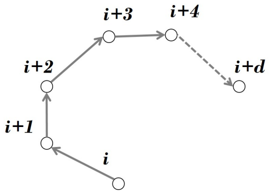
Over a network (or space)

$$R_{sy}(d, \delta) = E[y_{i+d} | u_i = u_0 + \delta, x_i] - E[y_{i+d} | u_i = u_0, x_i] \qquad (2)$$

y: outcome
t: time
h: time interval
i: individual unit
d: distance in a network, $d(i,j) = 1$ if $g^*_{ij} = 1$, $d(i,j) = 2$ if $g^*_{ij} = 0$, $\sum_k g^*_{ik} g^*_{kj} > 0$. $g_{ij} = g^*_{ij} / \sum_i g^*_{ij}$ (row-normalized)
$g_{ij} = 1$ if i and j have a link.
s: intervention
u_0: baseline, e.g.,s0=0
δ: treatment
x: vector of exogenous and predetermined variables. trivial example

Estimation by Network Local Projections (NLP)

Suppose the **network is direct "circular"**, i.e.
$g_{ij} = g_{jk} = g_{kt} = \cdots = 1$, intervention is $u_i = \delta$ and $u_{-i} = 0$.



Linear case:

$$y_{i+d} = \alpha_d + \beta_d u_i + v_{i+d}; \tag{3}$$

As long as u_i exogenous w.r.t. v_{i+d}

$$R_{sy}(d, \delta) = \beta_d \delta \tag{4}$$

This is very **similar to time series**.

Table of Contents

Relationship with spatial autoregressive (SAR) models

Suppose:

$$y_i = \phi \sum_j g_{ij} y_j + u_i; \qquad (5)$$

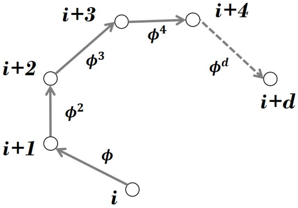
In matrix form:

$$y = \phi G y + u; \qquad (6)$$

by recursive substitution:

$$y = (I - \phi G)^{-1} u = [I + \phi G + (\phi G)^2 + (\phi G)^3 + \cdots + (\phi G)^{\text{inf}}] u; \quad (7)$$

$$R_{sy}(d, \delta) = E[y_{i+d} | u_i = \delta] - E[y_{i+d} | u_i = 0] = \phi^d \delta; \qquad (8)$$



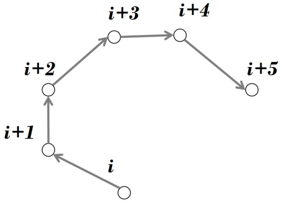
Propagation in spatial autoregressive (SAR) models

$$R_{sy}(d, \delta) = E[y_{i+d} | u_i = \delta] - E[y_{i+d} | u_i = 0] = \phi^d \delta;$$

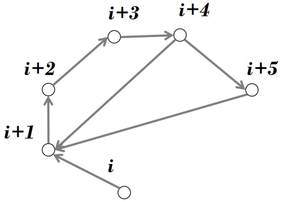
More complex if the network is not circular!

Suppose intervention is $u = \Sigma$ then:

$$R_{sy}(d, \Sigma) = (\phi G)^d \Sigma. \tag{9}$$



(circle)



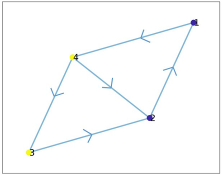
(recursive)

Main **difference 1**: LP - Forward and Backward

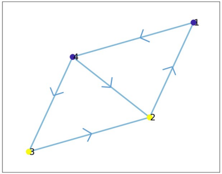
Main **difference 1**: NLP - Forward and Backward

Simple Example

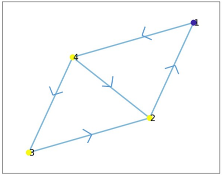
$$G = \begin{pmatrix} g_{1,1} & g_{1,2} & g_{1,3} & g_{1,4} \\ g_{2,1} & g_{2,2} & g_{2,3} & g_{2,4} \\ g_{3,1} & g_{3,2} & g_{3,3} & g_{3,4} \\ g_{4,1} & g_{4,2} & g_{4,3} & g_{4,4} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, i = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$



(a) Impulse ($i > 0$)

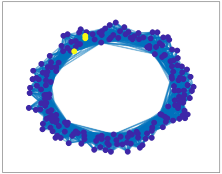


(b) 1 step F ($Gi > 0$)

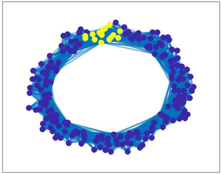


(c) 3 steps F and 2 B ($G'^2 G^3 i > 0$)

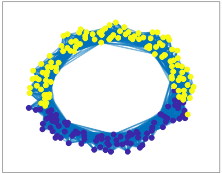
More Complex Example



(d) Impulse ($i > 0$)



(e) 1 step F ($Gi > 0$)



(f) 3 steps F and 2 B ($G'^2 G^3 i > 0$)

Main **difference 1**: NLP Forward and Backward

Forward

$$G'^{d+1}y = \alpha_{d+1} + \phi_{d+1}y + v$$

needs $E[(Gy)'v] = 0$ for identification.

Backward

$$y = \alpha_1 + \phi_{d+1}G^{d+1}y + u$$

needs $E[(G^{d+1})'u] = 0$ for identification.

- Differently from time series, F and B are conceptually similar, but different.
- Same root cause: recursivity

formal difference under SAR

Table of Contents

General Networks

Focus on **backward NLP**. If DGP is SAR and u are iid.

$$y = \alpha_{d+1} + \phi_{d+1} G^{d+1} y + v \tag{10}$$

$$v = \sum_{k=0}^{d} (\phi G)^k u \tag{11}$$

It follows that ϕ_{d+1} can be identified if

$$E[y_{i-1} v_{i+d}] \quad = \quad E[(G^{d+1} y)' v] = 0$$

Main **difference 2**: **Identification**

which translates into

$$E[(G^{d+1}y)'(\sum_{k=0}^{d} G^k u)] = E[(G^{d+1}Mu)'(\sum_{k=0}^{d} G^k u)] = 0$$

where $M = (I - \phi G)^{-1} = \sum_{l=0}^{\infty}(\phi G)^l$.

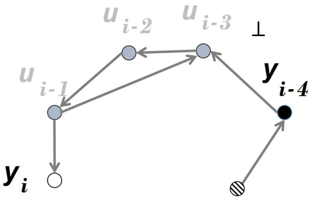
$$E[(\sum_{l=0}^{\infty} \phi^l G^{(d+1)+l} u)'(\sum_{k=0}^{d} G^k u)] \quad = \quad E[u'\Delta u] = 0$$

where $\Delta = \sum_{l=0}^{\infty} \sum_{k=0}^{d} \phi^l G'^{(d+1)+l} G^k$

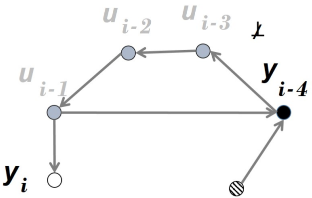
Result 1: as long as $\phi \neq 0$ and $trace(\Delta) \neq 0$, ϕ_{d+1} is not identified.

Endogeneity

- if no links from $i-1, i-2, ..., i-d$ to $i-(d+1)$,
  - $y_{i-(d+1)}$ is econometrically **exogenous**
  - estimation follows standard LP.
- if there are links,
  - $y_{i-(d+1)}$ is econometrically **endogenous**
  - no standard LP.
- endogeneity is endemic in networks because of recursivity.



(exogeneity)



(endogeneity)

A more General Specification

$$y = \alpha_{d+1} + \phi_{d+1} G^{d+1} y + G^{d+1} x \gamma_{d+1} + \sum_{k=0}^{d} G^k x \gamma_k + v. \quad (12)$$

$$v = \sum_{k=0}^{d} G^k u \quad (13)$$

- Same issues with ϕ_{d+1}
- Identification possible through **instrumental variables**.

Network Embedded Instrumental Variables

$G^{(d+1)+l}x, l > 1$ can be used as instrument if (it is relevant):

$$\begin{aligned} E[(G^{d+1}y)'(G^{(d+1)+l}x)] &= E[(G^{d+1}Mx)'(G^{(d+1)+l}x)] \\ &= E[(\sum_{l=0}^{\infty} \phi^l G^{(d+1)+l}x)' G^{(d+1)+l}x] \neq 0 \end{aligned}$$

which is true if $\phi \neq 0$, and exogenous:

$$E[v'(G^{(d+1)+l}x)] = E[(\sum_{k=0}^{d} G^k u)'(G^{(d+1)+l}x)] = 0$$

which is true because x is orthogonal to u also on different net-lags.

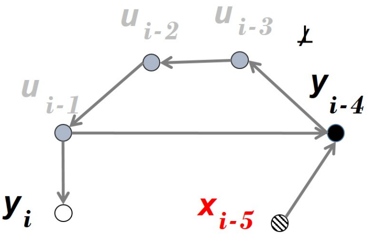
Network Embedded Instrumental Variables/2

Second step:

$$y = \alpha_{d+1} + \phi_{d+1} G^{d+1} y + G^{d+1} x \gamma_{d+1} + \sum_{k=0}^{d} G^k x \gamma_k + v. \quad (14)$$

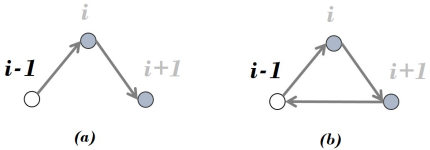
1-lag approximation first step:

$$G^{d+1} y = \alpha^*_{d+1} + \sum_{k=0}^{d+1} G^k x \gamma^*_k + G^{d+1+1} x \beta^* + v^*$$



Identification

- NLP identified if $I, G, G^2, ..., G^{d+1}, G^{d+1+1}, ..., G^{d+1+p}$ linearly independent and $\beta^* \neq 0, \gamma_k^* \neq 0, \forall k$.
- in SAR ϕ is not dentified if I, G, G^2 linearly dependent (Bramoulle et al. 2009).
- I.e. intransitive triads in the network.



- NLP is more demanding, requires more intransitivity
- **Result** 2**:** Not fully similarly to LP-AR, NLP and SAR are not identified under the same sufficient conditions.

Reduced form NLP (RF-NLP)

NLP can be used in reduced form, to estimate the effects of the treatment at distance d directly

Forward

$$G'^d y = \alpha_{d+1} + \phi_{d+1} Gx + \sum_{k=0}^{d} G'^k x \gamma_k + v$$

Backward

$$y = \alpha_1 + \phi_{d+1} G^{d+1} x + \sum_{k=0}^{d} G^k x \gamma_k + u$$

SAR mispecification

Suppose the true DGP is

$$y = \phi_1 G_1 y + \phi_2 G_2 y + \phi_3 G_3 y + x\beta + \epsilon$$

G_p can be higher order lags or due to heterogeneous transmission.
Assume $G_1 = G,\ G_2 = G^2,\ G_3 = G^3$.
If you use

$$y = \phi G y + x\beta + \epsilon$$

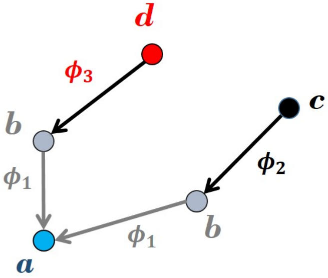
to estimate $\delta y_i / \delta x_j | d(i,j) = d \in [1,2,3,4]$,

Distance	True effect	SAR(1) estimated effect
1	$\phi_1 \beta$	$\hat{\phi}_{SAR}\beta$
2	$(\phi_2 + \phi_1^2)\beta$	$\hat{\phi}_{SAR}^2\beta$
3	$(2\phi_1\phi_2 + \phi_3 + \phi_1^3)\beta$	$\hat{\phi}_{SAR}^3\beta$
4	$(\phi_2^2 + \phi_1\phi_3 + \phi_1^4 + 3\phi_2\phi_1^2)\beta$	$\hat{\phi}_{SAR}^4\beta$

Result 3: NLP less prone to misspecification than SAR (similar to LP-AR).

Example - Heterogeneous Transmission in PN

Suppose there are 4 sectors, a, b, c and d, with such IO connections.



To sector	From sector	True effect	SAR estimated effect
a	b	$\phi_1\beta$	$\hat{\phi}_{SAR}\beta$
a	d	$\phi_1\phi_3\beta$	$\hat{\phi}^2_{SAR}\beta$
a	c	$\phi_1\phi_2\beta$	$\hat{\phi}^2_{SAR}\beta$

Table of Contents

Monte Carlo Experiments

- Study the properties of NLP in finite samples

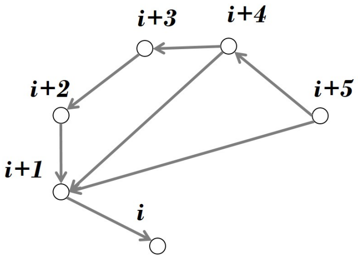
- For different specifications
  - Backward vs forward
  - Level of density

- Comparison with SAR
  - Misspecification

- SAR DGP and estimation
  - $SAR: y = +\phi Gy + x\beta + Gx\gamma + \epsilon$
- NLP estimation
  - NLP: $y = \alpha_1 + \phi_{d+1} G^{d+1} y[IV = G^{d+2}x] + \sum_{k=0}^{d+1} G^k x\gamma_k + v$
  - NLP_F: $G'^d y = \alpha_{d+1} + \phi_{d+1} Gy[IV = G^2 x] + Gx\gamma + \sum_{k=0}^{d} G'^k x\gamma_k + v$
  - NLP_{noGx}: $y = \alpha_1 + \phi_{d+1} G^{d+1} y[IV = G^{d+2}x] + G^{(d+1)} x\gamma + v$
  - $NLP_{F,noGx}$: $G'^d y = \alpha_{d+1} + \phi_{d+1} Gy[IV = G^2 x] + Gx\beta + v$
  - $NLP_{noGxnoiv}$: $y = \alpha_1 + \phi_{d+1} G^{d+1} y + G^{(d+1)} x\gamma + v$

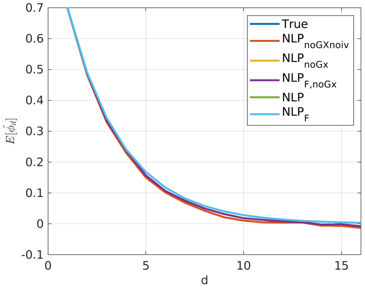
Setting

- Randomly normally generated X and ϵ, Reps $= 500$.
- Simulated recursive networks.
    - For each i, links from node $i+j$ to $i+1$ directed to i. for $j \leq z_i$.
    - $z_i = m$.
    - m governs the density
    - G is row-normalized
- pivotal baseline simulation: $N = 600$ nodes, $\phi = 0.7$, $\beta = 0.3$, $\gamma = 0.2$, $\sigma = 0.01$, $m = 4$.

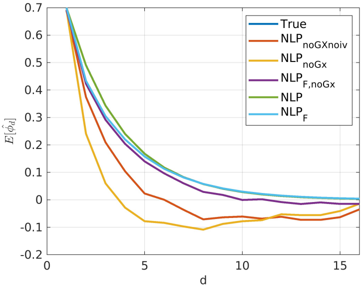


Different NLP specifications

Differently from LP, higher density biases estimates if in between lags are not included!



(g) purely circular (m=1)



(h) denser (m=4)

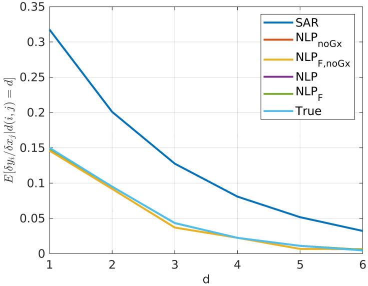
Intuition: higher density create higher recursivity, thus omitting them biases estimates. NLP precision

SAR - order misspecification

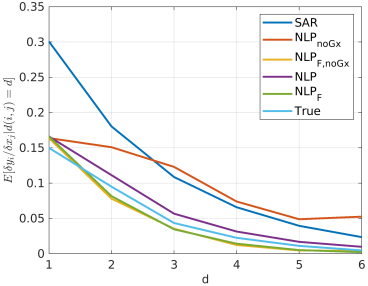
DGP

$$y = \phi_1 Gy + \phi_2 G^2 y + x\beta + \epsilon$$

$\phi_1 = 0.3$, $\phi_2 = 0.1$. estimate $\delta y_i / \delta x_j | d(i,j) = d$, with SAR or RF-NLP (RHS $G^d x$)



(i) m=1



(j) m=4

NLP precision

Table of Contents

Were NLP implicitly already used in applied work?

Some recent papers on production networks used approaches similar to RF-NLP.

- Carvalho et al. (2021) QJE "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake"
- Huremovic et al. (2024) WP "Production and Financial Networks in Interplay: Crisis Evidence from Supplier-Customer and Credit Registers"
- Others?

Less close approaches

- Barrot and Sauvagnat QJE (2016)
- Acemoglu et al. Mecroecon Annuals (2015)

Nice motivation to study NLP!

Carvalho et al. (2021) specification

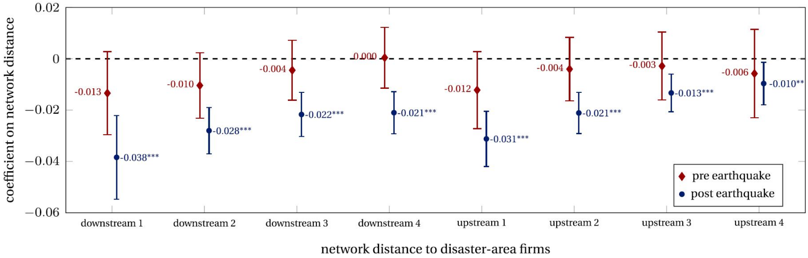
$$
\begin{aligned}
y_{ipst} &= \gamma_i + \gamma_{pst} + \sum_{k=1}^{4} \sum_{\tau \neq 2011} \beta_{k,\tau}^{down} Downstream_i^k * year_\tau \\
&+ \sum_{k=1}^{4} \sum_{\tau \neq 2011} \beta_{k,\tau}^{up} Upstream_i^k * year_\tau \\
&+ \sum_{\tau \neq 2011} \delta_\tau X_{isp} * year_\tau + \epsilon_{ispt},
\end{aligned} \tag{15}
$$

y: log sales; i: firm, p: prefecture; s: industry; t: time.
$Downstream_i^k$ and $Upstream_i^k$ dummy variables that indicate whether firm i is, respectively, a downstream or upstream distance k to disaster-area firms.
Abstract from the time dimension for now, and get rid of firm and ps FEs for simplicity.

Carvalho et al. (2021) specification/2

Let G^* be the input output matrix $g^*_{ij} = q_{ij} * p_{ij}$ and $t_i = 1$ if the firm is hit by the disaster.

$$\begin{aligned} y &= \sum_{k=1}^{4} \beta_k^{down} \textcolor{red}{Upstream^k} + \sum_{k=1}^{4} \beta_k^{up} \textcolor{blue}{Upstream^k} + \delta x + \epsilon, \\ &= \sum_{k=1}^{4} \beta_k^{down} I(\textcolor{red}{G'^{*k}} t > 0) + \sum_{k=1}^{4} \beta_k^{up} I(\textcolor{blue}{G^{*k}} t > 0) + \delta x + \epsilon, \end{aligned}$$



Avenues for research

- Understand better **properties**
- Study the **asymptotics**
- Extend Monte Carlo **experiments**
- **Application**
  - what if SAR and NLP estimates differ in popular settings?
  - best applications to put in the paper
    - US sectorial public data from Acemoglu et al. (2016)
    - better EU data / research question?
    - firm 2 firm data?
- **Spatio-temporal** extension sounds promising

THANKS!

edoardo.rainone@bancaditalia.it

Supporting Material

A trivial example

Suppose $u_t \in 0, 1$ is randomly assigned, then:

$$R_{sy}(d, 1) = 1/N_1 \sum_i y_{i+d} u_i - 1/N_0 \sum_i y_{i+d}(1 - u_i) \qquad (16)$$

$N_1 = \sum_i u_i; \quad N = N_0 + N_1.$

back

Backward is not Forward in Networks

Suppose the DGP is SAR:

$$\begin{aligned} y &= \phi G y + u = (I - \phi G)^{-1} u \\ &= u + \phi G u + \phi^2 G^2 u + \phi^3 G^3 u + \cdots \\ &= u + \phi G u + \phi^2 G^2 u + \phi^3 G^3 y \end{aligned} \tag{17}$$

$$G'^3 y = G'^3 u + \phi G'^3 G u + \phi^2 G'^3 G^2 u + \phi^3 \underbrace{G'^3 G^3}_{\neq I} y \tag{18}$$

Take $d = 3$

Backward

$$\begin{aligned} y &= \phi_3 G^3 y + v \\ &= \phi_3 G^3 y + (u + \phi G u + \phi^2 G^2 u) \end{aligned} \tag{19}$$

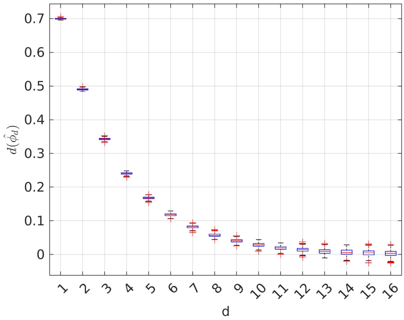
Forward

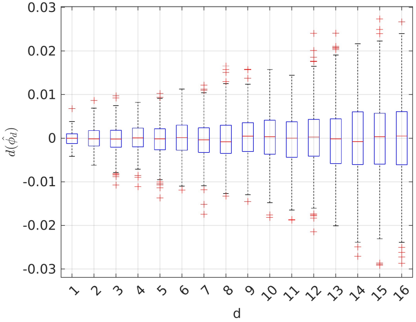
$$G'^3 y = \phi_3 y + v$$

$$y = \phi_3 G'^{-3} y + G'^{-3} v$$

back

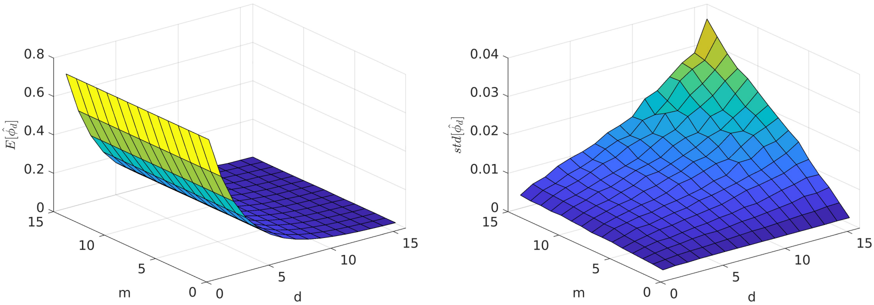
Precision decreases with distance (m=4)





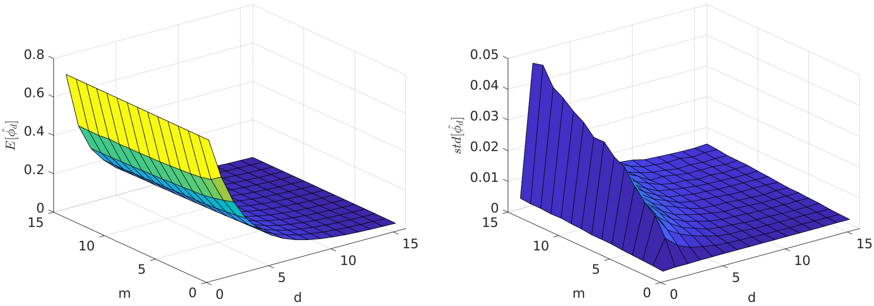
Standard errors increase with the distance back

NLP - distance and density (d,m)



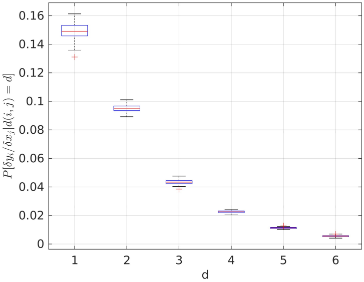
… and much more with the density of the network. back

NLP_F - distance and density (d,m)

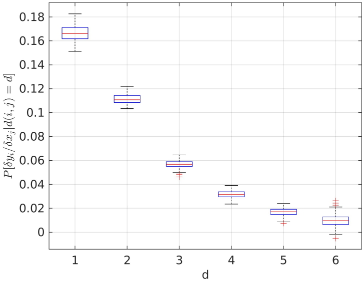


However, the forward-outcome NLP have the opposite feature, thus they can be used for longer distances. back

$RF-NLP$ precision - under SAR misspecification



(k) m=1



(l) m=4

back

Huremovic et al. (2024) specification

$$\Delta ln(s) = \underbrace{\alpha_D t}_{direct} + \underbrace{\alpha_{FD}\xi^{FD}}_{first-o,down} + \underbrace{\alpha_{HD}\xi^{HD}}_{high-o,down} + \underbrace{\alpha_{FU}\xi^{FU}}_{first-o,up} + \underbrace{\alpha_{HU}\xi^{HU}}_{high-o,up}$$
$$+ \underbrace{\alpha_{CC}\xi^{CC}}_{cust,cntrly} + \underbrace{\alpha_{SC}\xi^{SC}}_{supplier,cntrly} + Z\gamma + FE + \epsilon$$

s: sales/purchases of a firm. It can be re-written as

$$y = \alpha_D t + \alpha_{FD} G^{*\prime} t + \alpha_{HD} \sum_{k=2}^{\infty} G^{*\prime} t + \alpha_{FH} H t + \alpha_{HU} \sum_{k=2}^{\infty} H' t$$
$$+ \alpha_{CC}\xi^{CC} + \alpha_{SC}\xi^{SC} + Z\gamma + FE + \epsilon \qquad (20)$$

Where $G = AG := \{\alpha_i g_{ij}\}$, $H = GAMTV := \{[\alpha_i g_{ij}/\mu_i(1+\nu+t_i)](v_j/v_i)\}$, $\xi^{CC} = \sum_{k=0}^{\infty} G^{*\prime}$, $\xi^{SC} = \sum_{k=0}^{\infty} H'$. g^*_{ij}: row-norm intermediate IO matrix element; α_i: input elasticities; μ_i: markups to marginal cost; t_i: treatment; v_i: firm centrality γ_i: preference weight by customers. More complex connection with a theoretical model.

back