Are Unobservables Separable?

Andrii Babii[1] Jean-Pierre Florens[2]

[1]UNC Chapel Hill

[2]Toulouse School of Economics

Yale Econometrics Seminar, May 2021

Preview

Motivation

Heterogeneity of economic effects in unobservables.

Preview

Motivation

Heterogeneity of economic effects in unobservables.

This paper

1. Testing separability of unobservables for nonparametric models with endogeneity.
2. Independence as a testable implication of separability.
3. Does not estimate complex nonseparable models.
4. Residual-based tests.

Preview

Motivation

Heterogeneity of economic effects in unobservables.

This paper

1. Testing separability of unobservables for nonparametric models with endogeneity.
2. Independence as a testable implication of separability.
3. Does not estimate complex nonseparable models.
4. Residual-based tests.

Results

1. Distribution of nonparametric IV residuals.
2. Residual-based Kolmogorov-Smirnov and Cramér-von Mises tests.
3. Separability of Engel curves.

Table of Contents

Nonseparable IV model

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W,$$

where

- $Z \in \mathbf{R}^p$ are observable (potentially) endogenous variables;
- $\varepsilon \in \mathbf{R}^r$ are unobservables;
- $W \in \mathbf{R}^q$ is a vector of instrumental variables.

Nonseparable IV model

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

1. $\varepsilon \perp\!\!\!\perp W$ is stronger than mean-independence in separable IV models:
   - allows for heteroskedasticity of $\mathrm{Var}(Y|Z)$, $\mathrm{Var}(Y|W)$, and $\mathrm{Var}(\varepsilon|Z)$;
   - implies that $\mathrm{Var}(\varepsilon|W)$ is constant.

Nonseparable IV model

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

1. $\varepsilon \perp\!\!\!\perp W$ is stronger than mean-independence in separable IV models:
   - allows for heteroskedasticity of $\mathrm{Var}(Y|Z)$, $\mathrm{Var}(Y|W)$, and $\mathrm{Var}(\varepsilon|Z)$;
   - implies that $\mathrm{Var}(\varepsilon|W)$ is constant.
2. However, since we do not estimate the nonseparable model, the following conditions are not needed:
   - existence of triangular structure;
   - scalar unobservable ε;
   - monotonocity of $e \mapsto \Phi(z, e)$;
   - nonlinear completeness, etc.

Example 1: Heterogeneous Production Function

ECONOMETRICA

VOLUME 12 JULY-OCTOBER, 1944 NUMBERS 3 & 4

RANDOM SIMULTANEOUS EQUATIONS AND THE THEORY OF PRODUCTION*,[1]

By JACOB MARSCHAK and WILLIAM H. ANDREWS, JR.[2]

INTRODUCTION

§1

To describe and measure causation, the economist cannot perform experiments. That is, he cannot choose one variable as "dependent," and, while keeping the other, "independent," ones under control (i.e., while making them assume deliberately chosen sets of values), watch the values taken by the dependent, i.e., uncontrolled variable. The economist has no independent variables at his disposal because he has to take the values of all variables as they come, produced by a mechanism outside his control. This mechanism is expressed by a system of simultaneous equations, as many of them as there are variables. The experimenter can isolate one such equation, substituting his own action for all the other equations. The economist cannot.

For example, in agricultural experimentation preassigned quantities

Figure: Jacob Marschak, director of Cowles Comission, 1943-1948

Example 1: Heterogeneous Production Function

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

where

- $Y \in \mathbf{R}$ is output;
- $Z \in \mathbf{R}^p$ are observed inputs (labor, capital, etc.);
- $\varepsilon \in \mathbf{R}^r$ are unobserved technical efficiency, entrepreneurial factor (ability, urge, luck, etc.);
- $W \in \mathbf{R}^q$ are instrumental variables, e.g., input prices.

Example 2: Heterogeneous Demand Function

Random utility maximization leads to nonseparable demand function

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

Brown and Walker (Ecma, 1989), Blundell, Horowitz, and Parey (ReStat, 2017):

- $Y \in \mathbf{R}$ is demand;
- $Z \in \mathbf{R}^p$ are prices, income, and other observables;
- $\varepsilon \in \mathbf{R}^r$ are unobserved taste factors;
- $W \in \mathbf{R}$ is supply shifter.

Example 3: Heterogeneous Engel Curves

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

Imbens and Newey (Ecma, 2009)

- $Y \in \mathbf{R}$ is demand for some commodity;
- $Z \in \mathbf{R}^p$ total budget;
- $\varepsilon \in \mathbf{R}^r$ are taste factors;
- $W \in \mathbf{R}$ is gross income.

Nonseparable IV model

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

1. Allows for heterogeneity in unobservables – the structural function $z \mapsto \Phi(z, e)$ can be different for each $e \in \mathbf{R}^r$.
2. Challenging to identify and to estimate nonparametrically.
3. The nonparametric identification and estimation of separable nonparametric IV models is easier

$$Y = \varphi(Z) + \psi(\varepsilon).$$

Nonseparable IV model

$$Y = \Phi(Z, \varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

1. Allows for heterogeneity in unobservables – the structural function $z \mapsto \Phi(z, e)$ can be different for each $e \in \mathbf{R}^r$.
2. Challenging to identify and to estimate nonparametrically.
3. The nonparametric identification and estimation of separable nonparametric IV models is easier

$$Y = \varphi(Z) + \psi(\varepsilon).$$

Are nonseparable models empirically relevant?

Related literature

Separability of unobservables

Lu and White (JoE, 2014) and Su, Tu, and Ullah (ER, 2015).

Separable nonparametric IV models (NPIV)

Newey and Powell (Ecma, 2003), Hall and Horowitz (AS, 2005), Blundell, Chen, and Kristensen (Ecma, 2007), and Darolles et al (Ecma, 2011).

Nonseparable nonparametric IV models

Chernozhukov, Imbens, and Newey (2007, JoE), Florens, Heckman, Meghir, and Vytlacil (Ecma, 2008) Imbens and Newey (Ecma, 2009), Torgovitsky (Ecma, 2015).

Distribution of regression residuals

Durbin (AS, 1973), Loynes (AS, 1980), Akritas and Van Keilegom (SJS, 2001), Einmahl and Van Keilegom (JoE, 2008).

Contributions

New separability test

Sidestep estimation of nonseparable nonparametric IV model.

Contributions

New separability test

Sidestep estimation of nonseparable nonparametric IV model.

Nonparametric IV residuals

Donsker CLT for the empirical distribution of NPIV residuals and residual-based independence tests.

Contributions

New separability test

Sidestep estimation of nonseparable nonparametric IV model.

Nonparametric IV residuals

Donsker CLT for the empirical distribution of NPIV residuals and residual-based independence tests.

Empirical illustration

Testing separability of Engel curves.

Definitions

1. There exists a separable representation if for some measurable $\varphi : \mathbf{R}^p \to \mathbf{R}$ and $\psi : \mathbf{R}^r \to \mathbf{R}$

$$Y = \varphi(Z) + \psi(\varepsilon), \qquad \varepsilon \perp\!\!\!\perp W.$$

2. The distribution of $Z|W$ is L_2-complete if

$$\mathbb{E}[\phi(Z)|W] = 0 \implies \phi = 0, \qquad \forall \phi \in L_2.$$

   Newey and Powell (2003), Andrews (JoE, 2017).

Comment: the nonseparable model can be nonidentified.

Separability: a testable implication

Theorem

Suppose that (i) there exists a separable representation; (ii) the distribution of $Z|W$ is L_2-complete. Let φ be a unique solution to

$$\mathbb{E}[\varphi(Z)|W] = \mathbb{E}[Y|W].$$

Then for $U \triangleq Y - \varphi(Z)$, we have

$$U \perp\!\!\!\perp W.$$

A testable implication

$$H_0: \; U \perp\!\!\!\perp W \qquad vs. \qquad H_1: \; U \not\perp\!\!\!\perp W,$$

where $U = Y - \varphi(Z)$ and φ solves $\mathbb{E}[\varphi(Z)|W] = \mathbb{E}[Y|W]$.

Comments:

1. Classical independence tests between $(U_i)_{i=1}^n$ and $(W_i)_{i=1}^n$.
2. Can estimate the nuisance parameter φ using the nonparametric IV regression (NPIV).
3. Critical values: asymptotic distribution can be affected by the estimation of nuisance parameter φ.
4. Bootstrap?

Nonparametric IV regression

$$Y = \varphi(Z) + U, \qquad \mathbb{E}[U|W] = 0.$$

1. Conditional moment restriction

$$\mathbb{E}[Y|W] = \mathbb{E}[\varphi(Z)|W].$$

2. Equivalently, $\varphi : \mathbf{R}^p \to \mathbf{R}$ is a solution to integral equation $r = T\varphi$:

$$r(w) \triangleq \mathbb{E}[Y|W = w]f_W(w) = \int \varphi(z)f_{ZW}(z, w)\mathrm{d}z \triangleq (T\varphi)(w).$$

3. Ill-posed inverse problem: generalized inverse of T^*T is not continuous.

Regularization in Sobolev spaces

1. Sobolev spaces with smoothness $s \in \mathbf{R}$

$$H^s(\mathbf{R}^p) = \left\{ f \in L_2(\mathbf{R}^p) : \; \|f\|_s \triangleq \|L^s f\| < \infty \right\},$$

   where $L^s f = F^{-1}(\omega_s F f)$, F is Fourier transform, and ω_s is a weight function.

2. Tikhonov regularization of $T\varphi = r$ in Sobolev spaces

$$\hat{\varphi} = \arg\min_{\phi} \left\| \hat{T}\phi - \hat{r} \right\|^2 + \alpha_n \|\phi\|_s^2,$$

   where $(\hat{T}, \hat{r})$ are estimates of (T, r), e.g., kernel smoothing.

Hilbert scale assumptions

1. Operator smoothing:

$$\|T\phi\|_v \sim \|\phi\|_{v-a}, \qquad \forall \phi \in L_2(\mathbf{R}^p), \quad \forall v \in \mathbf{R}.$$

2. Parameter smoothness:

$$\varphi \in H^b(\mathbf{R}^p).$$

Comments:

- the Hilbert scale framework ensures sufficiently fast convergence rates for the Tikhonov-regularized estimator

$$\|\hat{\varphi} - \varphi\| = o_P(n^{-1/4}).$$

- Provides a useful connection between the empirical processes theory and regularization theory.

Nonparametric IV residuals

1. Nonparametric IV residuals: $\hat{U}_i = Y_i - \hat{\varphi}(Z_i)$.
2. Empirical distribution functions for $F_U(u) = \Pr(U \leq u)$

$$\hat{F}_U(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{U_i \leq u\}} \qquad \text{and} \qquad \hat{F}_{\hat{U}}(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{\hat{U}_i \leq u\}}.$$

3. Can we have the following weak convergence

$$\sqrt{n}(\hat{F}_{\hat{U}} - F_U) \rightsquigarrow \mathbb{G}?$$

4. Complication: $\hat{F}_{\hat{U}}$ is a non-smooth function of regularized estimator $\hat{\varphi}$.

Distribution of nonparametric IV residuals

Theorem

Under Hilbert scale assumptions (+ some technical conditions)

$$\sqrt{n}(\hat{F}_{\hat{U}} - F_U) \rightsquigarrow \mathbb{G}$$

where $\mathbb{G}$ *is a centered Gaussian process with covariance structure*

$$\begin{aligned}(u, u') \mapsto\ & F_U(u \wedge u') - F_U(u)F_U(u') \\ & + \mathbb{E}\left[U^2 \left[T(T^*T)^{-1} f_{UZ}(u, .)\right](W) \left[T(T^*T)^{-1} f_{UZ}(u', .)\right](W)\right] \\ & + \mathbb{E}\left[\mathbb{1}_{\{U \leq u\}} U \left[T(T^*T)^{-1} f_{UZ}(u', .)\right](W)\right] \\ & + E\left[\mathbb{1}_{\{U \leq u'\}} U \left[T(T^*T)^{-1} f_{UZ}(u, .)\right](W)\right].\end{aligned}$$

Proof sketch

1. Asymptotic equicontinuity-type argument gives

$$\begin{aligned}\sqrt{n}(\hat{F}_{\hat{U}}(u) - F_U(u)) &= \sqrt{n}(\hat{F}_U(u) - F_U(u)) \\ &\quad + \sqrt{n}\left(\Pr(Y \leq u + \hat{\varphi}(Z)) - F_U(u)\right) \\ &\quad + o_P(1).\end{aligned}$$

Comment: use $\|\hat{\varphi} - \varphi\|_s = o_P(1)$.

Proof sketch

1. Asymptotic equicontinuity-type argument gives

$$\begin{aligned}\sqrt{n}(\hat{F}_{\hat{U}}(u) - F_U(u)) &= \sqrt{n}(\hat{F}_U(u) - F_U(u)) \\ &\quad + \sqrt{n}\left(\Pr(Y \leq u + \hat{\varphi}(Z)) - F_U(u)\right) \\ &\quad + o_P(1).\end{aligned}$$

   Comment: use $\|\hat{\varphi} - \varphi\|_s = o_P(1)$.

2. By Taylor's theorem

$$\sqrt{n}\left(\Pr(Y \leq u + \hat{\varphi}(Z)) - F_U(u)\right) = \sqrt{n}\langle \hat{\varphi} - \varphi, f_{UZ}(u, .)\rangle + o_P(1)$$

   Comment: use $\|\hat{\varphi} - \varphi\| = o_P(n^{-1/4})$.

Proof sketch

1. Asymptotic equicontinuity-type argument gives

$$\begin{aligned}\sqrt{n}(\hat{F}_{\hat{U}}(u) - F_U(u)) &= \sqrt{n}(\hat{F}_U(u) - F_U(u)) \\ &+ \sqrt{n}\left(\Pr(Y \leq u + \hat{\varphi}(Z)) - F_U(u)\right) \\ &+ o_P(1).\end{aligned}$$

   Comment: use $\|\hat{\varphi} - \varphi\|_s = o_P(1)$.

2. By Taylor's theorem

$$\sqrt{n}\left(\Pr(Y \leq u + \hat{\varphi}(Z)) - F_U(u)\right) = \sqrt{n}\langle \hat{\varphi} - \varphi, f_{UZ}(u,.)\rangle + o_P(1)$$

   Comment: use $\|\hat{\varphi} - \varphi\| = o_P(n^{-1/4})$.

3. Approximation

$$\sqrt{n}\langle \hat{\varphi} - \varphi, f_{UZ}(u,.)\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \left[T(T^*T)^{-1} f_{UZ}(u,.)\right](W_i) + o_P(1).$$

Residual-based independence process

1. Empirical distributions

$$\hat{F}_{\hat{U}W}(u,w) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{\hat{U}_i \leq u, W_i \leq w\}},$$

$$\hat{F}_{\hat{U}}(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{\hat{U}_i \leq u\}}, \qquad \hat{F}_{W}(w) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{W_i \leq w\}},$$

where $\hat{U}_i = Y_i - \hat{\varphi}(Z_i)$ are NPIV residuals.

2. Independence empirical process

$$\mathbb{G}_n(u,w) = \sqrt{n}\left(\hat{F}_{\hat{U}W}(u,w) - \hat{F}_{\hat{U}}(u)\hat{F}_{W}(w)\right).$$

Independence tests

1. Kolmogorov-Smirnov

$$T_{\infty,n} = \sup_{u,w} |\mathbb{G}_n(u,w)|.$$

2. Cramér-von Mises

$$T_{2,n} = \iint |\mathbb{G}_n(u,w)|^2 \mathrm{d}\hat{F}_{\hat{U}W}(u,w).$$

Non-standard asymptotic distribution

Theorem

Under Hilbert scale assumptions (+ some technical conditions) under H_0

$$T_{\infty,n} \rightsquigarrow \sup_{u,w} |\mathbb{G}(u,w)| \qquad \textit{and} \qquad T_{2,n} \rightsquigarrow \iint |\mathbb{G}(u,w)|^2 \mathrm{d}F_{UW}(u,w),$$

where $\mathbb{G}$ is a centered Gaussian process with covariance structure

$$(u,w,u',w') \mapsto \mathbb{E}\Big[\Big(\mathbb{1}_{\{U\leq u, W\leq w\}} - \mathbb{1}_{\{U\leq u\}}F_W(w) - \mathbb{1}_{\{W\leq w\}}F_U(u) + F_{UW}(u,w) + \delta_{u,w}(U,W)\Big) \times \\ \times \Big(\mathbb{1}_{\{U\leq u', W\leq w'\}} - \mathbb{1}_{\{U\leq u'\}}F_W(w') - \mathbb{1}_{\{W\leq w'\}}F_U(u') + F_{UW}(u',w') + \delta_{u',w'}(U,W)\Big)\Big].$$

Comments

1. The covariance structure is contaminated by

$$\delta_{u,w}(U_i, W_i) = U_i \left(T(T^*T)^{-1}\rho(u,.,w)\right)(W_i),$$
$$\rho(u,z,w) = \int^{w} f_{UZW}(u,z,\tilde{w})\mathrm{d}\tilde{w} - f_{UZ}(u,z)F_W(w).$$

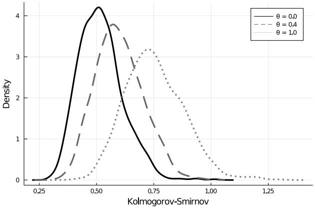
2. NB: for nonparametric conditional mean regression, we have distribution-free independence tests; see Einmahl and Van Keilegom (JoE, 2008).
3. The bootstrap may fail to estimate densities f_{UZ} and f_{UZW} consistently $\implies$ use subsampling or m-out-of-n bootstrap.
4. Consistency under fixed alternatives and power against parameteric local alternatives.

MC experiments: simulation design

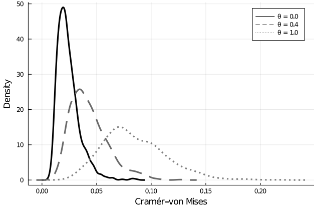
$$Y = \varphi(Z) + \theta ZU + U, \quad \begin{pmatrix} Z \\ W \\ U \end{pmatrix} \sim_{i.i.d.} N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.4 & 0.3 \\ 0.4 & 1 & 0 \\ 0.3 & 0 & 1 \end{pmatrix} \right).$$

- $\theta = 0$: separable IV model;
- $\theta \neq 0$: nonseparable IV model.

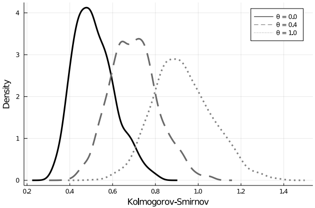
Figure: Simulated distributions of Kolmogorov-Smirnov (KS) and Cramér-von Mises (CvM) tests under $H_0 : \theta = 0$ (solid line) and H_1: $\theta = 0.4$ (dashed) and $\theta = 1$ (dotted). Structural function $\varphi(z) = \cos(z)$.



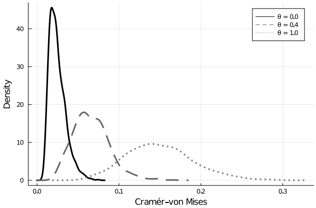
(a) KS test, $n = 500$



(b) CvM test, $n = 500$

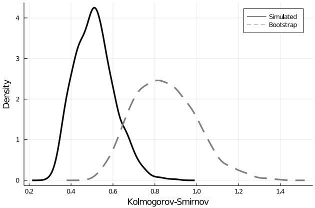


(c) KS test, $n = 1,000$

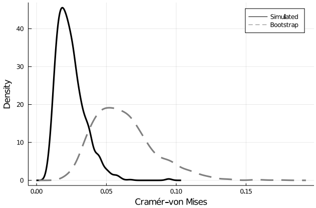


(d) CvM test, $n = 1,000$

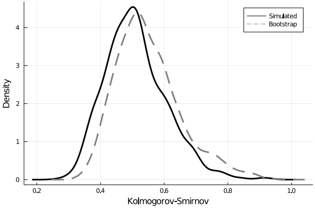
Figure: m out of n bootstrap. Simulated (solid line) and bootstrapped (dashed) distributions of KS and CvM statistics under H_0.



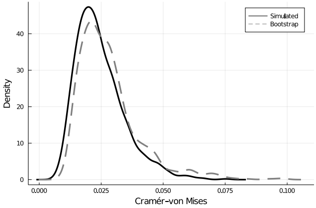
(a) KS statistics: $m_n = n$



(b) CvM statistics: $m_n = n$

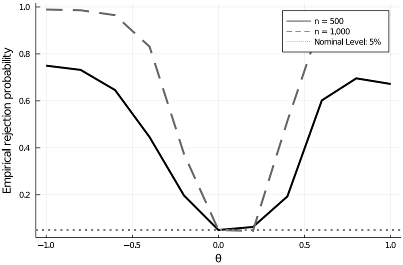


(c) KS: adaptive m_n

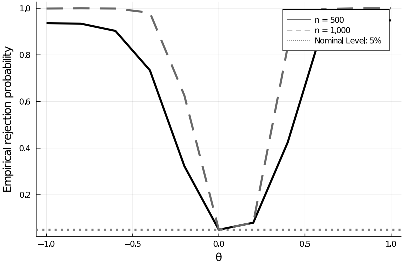


(d) CvM: adaptive m_n

Figure: Empirical rejection probabilities as a function of degree of separability $\theta \in [-1, 1]$ for samples of size $n = 500$ (solid line) and $n = 1,000$ (dashed).



(a) KS test



(b) CvM test

Separability of Engel curves

1. Engel curves describe how the demand for commodity changes while the household's budget increases:
   - Y - share spent on commodity;
   - Z - total budget allocated to subgroup of commodities.
2. Endogeneity: both Y and Z are likely to be jointly determined; see Blundell, Chen, and Kristensen (Ecma, 2007).
3. Gross earnings is a valid IV, assuming that heterogeneity in earnings is not related to households' preferences over consumption.

Data

- 2015 Consumer Expenditure Survey data in the US;
- $10,055$ married couples with positive income;
- Y share of total non-durable expenditures spent;
- Z log of total non-durable expenditures;
- W gross earnings of households' head.

Table: Testing separability of Engel curves. The table shows m out of n bootstrap p-values of Kolmogorov-Smirnov and Cramér-von Mises tests for 13 commodities.

Commodity	KS	CvM	Commodity	KS	CvM
Food home	0.00	0.00	Gas and oil	0.00	0.00
Food away	0.00	0.00	Personal care	0.00	0.00
Clothing	0.00	0.00	Health	0.00	0.00
Tobacco	0.00	0.00	Insurance	0.00	0.00
Alcohol	0.00	0.00	Reading	0.00	0.53
Trips	0.00	0.00	Transportation	0.01	0.03
Entertainment	0.08	0.00			

Comment: the test rejects separability for all commodities except for Reading (CvM test).

Comments

1. Not surprising: why would every household have the same Engel curve?
2. Rescuing the NPIV model: controlling for other observables $X \in \mathbf{R}^{r-1}$

$$Y = \varphi(Z, X) + U, \qquad \mathbb{E}[U|X, W] = 0$$

3. Semiparametric approaches:
    - mixed index models: $\varphi(Z, X) = \varphi(Z, X^\top \beta)$?
    - random coefficients with nonlinearities?

Conclusions

1. New separability test that does not require identification and estimation of nonseparable IV model.
2. Donsker-type CLTs for the empirical distribution of NPIV residuals and residual-based independence tests.
3. Heterogeneity of (nonlinear) economic effects in unobservables is empirically relevant.

Thank you!

ababii.github.io