The econometrics of wage inequality: Issues of identification and inference

Maximilian Kasy

Two papers

1. "Who wins, who loses? Tools for distributional policy evaluation" (complete)
2. "Labor demand and wage inequality in Europe – an empirical Bayes approach" (work in progress)

- Talk: focus on conceptual issues
- Papers: technical details, empirical applications

Agenda

Develop methods to better understand:

- income inequality, in particular wage inequality
- changes in inequality
- historical causes, policy counterfactuals, predictions
- winners and losers; political economy

Econometric challenges

What if

1. we care about unobserved welfare instead of observed earnings?
   - welfare in the utilitarian / optimal tax sense
   - need to identify causal effects
     conditional on vectors of endogenous outcomes

   $\Rightarrow$ Part I: Who wins, who loses?

2. we are interested in the impact of changing labor supply?
   - many types of workers $\Rightarrow$ many regressors, few observations
   - conventional solution: parametric structural model
   - nonrobust conclusions
   - proposed solution: shrinking towards structural model
     in data-dependent, optimal way

   $\Rightarrow$ Part II: Labor demand and wage inequality / empirical Bayes

Literature

1. **Public - optimal taxation**: Samuelson (1947), Mirrlees (1971), Saez (2001), Chetty (2009), Hendren (2013), Saez and Stantcheva (2013)
2. **Labor - determinants of wage distribution**: Borjas et al. (1996), Autor et al. (1998), Autor et al. (2008), Card (2001), Card (2009), Autor and Dorn (2013)
3. **Distributional decompositions**: Oaxaca (1973), DiNardo et al. (1996), Firpo et al. (2009), Rothe (2010), Chernozhukov et al. (2013)
4. **Sociology - class analysis**: Wright (2005)
5. **Mathematical physics - fluid dynamics, differential forms**: Rudin (1991), Evans (1998)
6. **Empirical Bayes**: Robbins (1956), James and Stein (1961), Efron and Morris (1973). Morris (1983), Laird and Louis (1987), Carlin and Gelfand (1990), Efron (2010).
7. **Econometrics - various**: Koenker (2005), Hoderlein and Mammen (2007), Abbring and Heckman (2007), Matzkin (2003), Altonji and Matzkin (2005)

Part I: Who wins, who loses?

Conjecture:

- Few policy changes result in Pareto improvements.
- Most generate **winners** and **losers**.
- Certainly true for controversial policy changes.

$\Rightarrow$ If we evaluate social welfare based on individuals' welfare:

1. To **evaluate** a **policy** effect, we need to
   1.1 define how we measure individual gains and losses,
   1.2 estimate them, and
   1.3 take a stance on how to aggregate them.
2. To **understand political economy**, we need to characterize the sets of winners and losers of a policy change.

My objective:

1. tools for distributional evaluation
2. utility-based framework, arbitrary heterogeneity, endogenous prices

Proposed procedure

1. Impute money-metric welfare effect to each individual
2. Then:
   2.1 Report average effects given income / other covariates
   2.2 Construct sets of winners and losers (in expectation)
   2.3 Aggregate using welfare weights

Contrast with program evaluation approach:

1. Effect on average
2. of observed outcome

Notation

- **policy** $\alpha \in \mathbb{R}$
  individuals i
- potential outcome w^α
  realized outcome w
- partial derivatives $\partial_w := \partial / \partial w$
  with respect to policy $\dot{w} := \partial_\alpha w^\alpha$
- density f
  cdf F
  quantile Q
- wage w
  labor supply l
  consumption vector c
  covariates W

Setup

Assumption (Individual utility maximization)

individuals choose c and l to solve

$$\max_{c,l} u(c,l) \quad s.t. \quad c \cdot p \leq l \cdot w + y_0.$$

$$v := \max u$$

- *u, c, l, w vary arbitrarily across i*
- p, w, y_0 *depend on* α
  $\Rightarrow$ *so do c, l, and v*
- *u differentiable, increasing in c, decreasing in l, quasiconcave, does not depend on* α

Objects of interest

Definition

1. *Money metric utility impact of policy:*

$$\dot{e} := \dot{v} \big/ \partial_{y_0} v$$

2. *Average conditional policy effect on welfare:*

$$\gamma(y, W) := E[\dot{e} | y, W, \alpha]$$

3. *Sets of winners and losers:*

$$\mathscr{W} := \{(y, W) : \gamma(y, W) \geq 0\}$$
$$\mathscr{L} := \{(y, W) : \gamma(y, W) \leq 0\}$$

4. *Policy effect on social welfare: $SWF: \; v(.) \to \mathbb{R}$*

$$\dot{SWF} = E[\omega \cdot \gamma]$$

Marginal policy effect on individuals

Lemma

$$\begin{aligned} \dot{y} &= (\dot{l} \cdot w + l \cdot \dot{w}) + \dot{y}_0, \\ \dot{e} &= \qquad\quad\;\; l \cdot \dot{w} \; + \dot{y}_0 - c \cdot \dot{p}. \end{aligned}$$

Proof: Envelope theorem.

1. wage effect $l \cdot \dot{w}$,
2. effect on unearned income $\dot{y}_0$,
3. behavioral effect $b := \dot{l} \cdot w$,
4. price effect $-c \cdot \dot{p}$.

Income vs utility:

$$\dot{y} - \dot{e} = \dot{l} \cdot w + c \cdot \dot{p}.$$

Identification of disaggregated welfare effects

- Goal: **identify** $\gamma(\mathbf{y}, \mathbf{W}) = \mathbf{E}[\dot{\mathbf{e}}|\mathbf{y}, \mathbf{W}, \alpha]$
- Simplified case:
  no change in prices, unearned income
  no covariates
- Then

$$\gamma(y) = E[l \cdot \dot{w} | l \cdot w, \alpha]$$

- Denote $x = (l, w)$.
  Need to identify

$$g(x, \alpha) = E[\dot{x} | x, \alpha]$$

  from

$$f(x|\alpha).$$

- Made necessary by combination of
  1. utility-based social welfare
  2. heterogeneous wage response.

Assume :

1. $x = x(\alpha, \varepsilon)$, $x \in \mathbb{R}^k$
2. $\alpha \perp \varepsilon$
3. $x(., \varepsilon)$ differentiable

Physics analogy:

- $x(\alpha, \varepsilon)$: position of particle ε at time α
- $f(x|\alpha)$: density of gas / fluid at time α, position x
- $\dot{f}$ change of density
- $h(x, \alpha) = E[\dot{x}|x, \alpha] \cdot f(x|\alpha)$: "flow density"

Stirring your coffee

- If we know densities $f(x|\alpha)$,
- what do we know about flow
  $g(x,\alpha) = E[\dot{x}|x,\alpha]$?

Problem: Stirring your coffee

- does not change its density,
- yet moves it around.
- $\Rightarrow$ different flows $g(x,\alpha)$
  consistent with a constant density $f(x|\alpha)$

Will show:

- Knowledge of $f(x|\alpha)$
  - identifies $\nabla \cdot h = \sum_{j=1}^{k} \partial_{x^j} h^j$
  - where $h = E[\dot{x}|x,\alpha] \cdot f(x|\alpha)$,
  - identifies nothing else.
- Add to h
  - $\tilde{h}$ such that $\nabla \cdot \tilde{h} \equiv 0$
  - $\Rightarrow f(x|\alpha)$ does not change
  - "stirring your coffee"
- Additional conditions
  - e.g.: "wage response unrelated to initial labor supply"
  - $\Rightarrow$ just-identification of $g(x,\alpha) = E[\dot{x}|x,\alpha]$
  - $g^j(x,\alpha) = \partial_\alpha Q(v^j|v^1,\ldots,v^{j-1},.\alpha)$

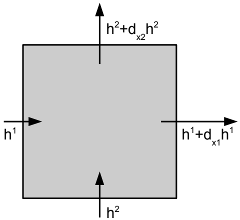
Density and flow

Recall

$$h(x,\alpha) := E[\dot{x}|x,\alpha]\cdot f(x|\alpha)$$

$$\nabla\cdot h := \sum_{j=1}^{k} \partial_{x^j} h^j$$

$$\dot{f} := \partial_\alpha f(x|\alpha)$$



Theorem

$$\dot{\mathbf{f}} = -\nabla\cdot\mathbf{h}$$

Proof:

1. For some $a(x)$, let

$$A(\alpha) := E[a(x(\alpha,\varepsilon))|\alpha] = \int a(x(\alpha,\varepsilon))dP(\varepsilon)$$
$$= \int a(x)f(x|\alpha)dx.$$

2. By partial integration:

$$\dot{A}(\alpha) = E[\partial_x a \cdot \dot{x}|\alpha] = \sum_{j=1}^{k} \int \partial_{x^j} a \cdot h^j dx$$
$$= -\int a \cdot \sum_{j=1}^{k} \partial_{x^j} h^j dx = -\int a \cdot (\nabla \cdot h) dx.$$

3. Alternatively:

$$\dot{A}(\alpha) = \int a(x)\dot{f}(x|\alpha)dx.$$

4. 2 and 3 hold for any $a \Rightarrow \dot{f} = -\nabla \cdot h$. $\square$

The identified set

Theorem

The identified set for h is given by

$$h^0 + \mathscr{H}$$

where

$$\begin{aligned}
\mathscr{H} &= \{\tilde{h} : \nabla \cdot \tilde{h} \equiv 0\} \\
h^{0j}(x,\alpha) &= f(x|\alpha) \cdot \partial_\alpha Q(v^j | v^1, \ldots, v^{j-1}, \alpha) \\
v^j &= F(x^j | x^1, \ldots, x^{j-1}, \alpha)
\end{aligned}$$

Point identification

Theorem

Assume

$$\frac{\partial}{\partial x^j} E[\dot{x}^i | x, \alpha] = 0 \; for \; j > i.$$

Then h is point identified, and equal to h^0 as defined before.

In particular

$$\begin{aligned} g^j(x, \alpha) &= E[\dot{x}^j | x, \alpha] \\ &= \partial_\alpha Q(v^j | v^1, \ldots, v^{j-1}, \alpha). \end{aligned}$$

Recap Part I

Application: impact of EITC

- Most policy changes generate winners and losers
- Motivates interest in
  1. disaggregated welfare effects
  2. sets of winners and losers (political economy!)
  3. weighted average welfare effects (optimal policy!)
- Consider framework which allows for
  1. endogenous prices / wages (vs. public finance)
  2. utility-based social welfare (vs. labor, distributional decompositions)
  3. arbitrary heterogeneity (vs. labor)

Part II: Labor demand and wage inequality – Empirical Bayes

- Previous part: one-dimensional treatment variable
  e.g. EITC-expansion
- Much of labor literature on wage inequality:
  impact of changing labor supply of various types
  $\Rightarrow$ high-dimensional treatment
- Examples:
  - Impact of migration on native wage inequality
  - Skill biased technical change versus expansion of higher education

Example: Impact of migration

- Literature: Estimate CES-production function model, consider historical counterfactual of no immigration.
- "Migration increased inequality"
  - Borjas et al. (1996)
  - CES-model with 4 types, by education
  - national economy, time series variation
- "Migration did not increase inequality"
  - Card (2001), Card (2009)
  - nested CES, 2 education types, natives vs migrants (justified by pre-tests)
  - cross-city, Bartik-type instrument

$\Rightarrow$ Conclusions depend on functional form choices!

Review of CES-production functions

Notation:

- types of workers $j = 1, \ldots, J$,
  cross-section of labor markets $i = 1, \ldots, n$
- wages w, labor supply N
- $Y_{ij} = \log(w_{ij})$, $X_{ij} = \log(N_{ij})$

Assumptions:

1. marginal productivity theory of wages:

$$w_{ij} = \frac{\partial f_i(N_{i1}, \ldots, N_{iJ})}{\partial N_{ij}}$$

2. CES production function:

$$f_i(N_{i1}, \ldots, N_{iJ}) = \left(\sum_{j=1}^{J} \gamma_j N_{ij}^{\rho} \right)^{1/\rho}$$

Wage equation

- This yields

$$w_{ij} = \frac{\partial f_i(N_{i1}, \ldots, N_{iJ})}{\partial N_{ij}} = \left(\sum_{j'=1}^{J} \gamma_j N_{ij'}^{\rho} \right)^{1/\rho - 1} \cdot \gamma_j \cdot N_j^{\rho - 1}.$$

- relative wage between groups j and j' is equal to

$$\frac{w_{ij}}{w_{ij'}} = \frac{\gamma_j}{\gamma_{j'}} \cdot \left(\frac{N_{ij}}{N_{ij'}} \right)^{\rho - 1}.$$

- Taking logs yields

$$Y_{j,i} - Y_{j',i} = \log(\gamma_j) - \log(\gamma_{j'}) + \beta_0 \cdot (X_{j,i} - X_{j',i}),$$

where $\beta_0 = \rho - 1$.

Unrestricted estimation

- Equivalent to:

$$Y_{j,i} = \alpha_i + \gamma_j + \sum_{j'} \beta_{j,j'} X_{j',i} + \varepsilon_{j,i},$$

$$\beta_{j,j'} = \beta_0 \cdot \begin{cases} \left(1 - \frac{1}{J}\right) & j = j' \\ -\frac{1}{J} & j \neq j' \end{cases}$$

- Could drop restrictions across $\beta_{j,j'}$,
  estimate the differenced, unrestricted model:

$$\begin{aligned} \Delta \cdot Y_i &= \Delta \cdot \gamma + \delta \cdot X_i + \Delta \cdot \varepsilon_i. \\ \delta &= \Delta \cdot \beta, \\ \Delta &= (-e, I_{J-1}) \end{aligned}$$

Drawbacks of either approach

- Structural:
  - biased, inconsistent
  - non-robust to choices of types, functional form
  - cf. literature on migration
- Unrestricted:
  - large variance
  - under-identified if many types, few observations

Intermission: Theory in empirical economics

- Structural models seem to cause non-robustness, possibly inconsistency.
  $\Rightarrow$ Should we use theory in empirical research at all?

The positivist ideal:

- Follow the example of physics.
- Develop theories which
  1. are assumed to be universally true, and
  2. have testable implications.
- Maintain these theories while they have not been rejected by statistical tests.
- When they have been rejected, replace them with new theories that are consistent with all available evidence.

The reality of economic theory

- We have no theories that are even approximately universally true.
- People don't universally – or even consistently in well defined contexts –
  - maximize utility,
  - discount exponentially,
  - maximize expected utility under risk,
  - play Nash equilibrium,
  - act as price takers on markets, ...
- Even less
  1. do people maximize utility with additive EV1 errors,
  2. does aggregate production follow a CES production function with 3 inputs, ...
- All of these theories can be, and have been, rejected.

What to do?

Several options:

1. Ignore this, keep following the positivist ideal, argue that theories don't actually have to be true.
   (Wasn't there something about Billiard players?)
2. Forget about economic theory, just try to do good statistics / mostly harmless econometrics.
3. Try to find a middle ground that makes reasonable use of theory.

An attempt at a middle ground

- Shrink “towards theory”
- Advantages:
  - Improves estimator performance if theory is (approximately) true.
  - Is not dogmatic – yields consistent estimates either way.
- Bayesian interpretation: (improper) priors that put low weight on parameter values deviating a lot from theory.
- Can do empirical Bayes version – avoids critique of subjectivism / arbitrary choice of tuning parameters.
- Might yield James-Stein type shrinkage informed by theory.
- Coming next: an implementation of this program in the context of labor demand.

The parametric empirical Bayes approach

- Parameters η, hyper-parameters θ
- **Model**:

$$Y|\eta \sim f(Y|\eta)$$

- **Family of priors**:

$$\eta \sim \pi(\eta|\theta)$$

- **Marginal density** of Y:

$$Y|\theta \sim g(Y|\theta) := \int f(Y|\eta)\pi(\eta|\theta)d\eta$$

- Estimation of hyperparameters: marginal MLE

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmax}}\ g(Y|\theta).$$

- Estimation of η:

$$\widehat{\eta} = E[\eta|Y, \theta = \widehat{\theta}]$$

An empirical Bayes approach for labor demand

- **Model**:
  unrestricted estimator as sufficient statistic;
  asymptotic approximation:

$$\widehat{\delta}_{\uparrow} | \eta \sim N(\delta_{\uparrow}, V)$$
$$\widehat{V} \cdot V^{-1} \to^p I.$$

- **Family of priors**:
  coefficients = structural model + noise of unknown variance

$$\beta = (\beta_{j,j'}) = \beta_0 \cdot M + \zeta$$
$$\zeta_{j,j'} \sim^{iid} N(0, \tau^2),$$

Differencing $\Rightarrow \delta = \Delta \cdot \beta = \beta_0 \cdot \Delta + \Delta \cdot \zeta$

- Parameters η, hyper-parameters θ:

$$\begin{aligned}
\eta &= (\delta, V) \\
\theta &= (\beta_0, \tau^2, V) \\
\widehat{\delta}_{\uparrow}|\eta &\sim N(\delta_{\uparrow}, V) \\
\delta_{\uparrow}|\theta &\sim N(\beta_0 \cdot \Delta_{\uparrow}, \tau^2 \cdot I_{J-1} \otimes P),
\end{aligned}$$

where $P = I_J + E$.

- **Marginal density** of Y:

$$\widehat{\delta}_{\uparrow}|\theta \sim N(\beta_0 \cdot \Delta_{\uparrow}, \Sigma(\tau^2, V)),$$

where

$$\Sigma(\tau^2, V) = \mathrm{Var}\left(\widehat{\delta}_{\uparrow}\middle|\theta\right) = \tau^2 \cdot I_{J-1} \otimes P + V.$$

Solving for the estimator

- Hyperparameters: MLE for the marginal likelihood,

$$(\widehat{\beta}_0, \widehat{\tau}^2) = \underset{b_0, t^2}{\operatorname{argmin}} \log\left(\det(\Sigma(t^2, \widehat{V}))\right) + (\widehat{\delta}_\uparrow - b_0 \cdot \Delta_\uparrow)' \cdot \Sigma(t^2, \widehat{V})^{-1} \cdot (\widehat{\delta}_\uparrow - b_0 \cdot \Delta_\uparrow)$$

- Parameter of interest δ:

$$\widehat{\delta}_\uparrow^{EB} = \widehat{\beta}_0 \cdot \Delta_\uparrow + I_{J-1} \otimes P \cdot \left(I_{J-1} \otimes P + \frac{1}{\widehat{\tau}^2} \widehat{V}\right)^{-1} \cdot (\widehat{\delta}_\uparrow - \widehat{\beta}_0 \cdot \Delta_\uparrow)$$

Geometry of the estimator

Advantages

Some Monte Carlo evidence

1. Consistent conditional on any value of δ
2. Counterfactual predictions driven by data, whenever these are informative
3. Lower variance than unrestricted OLS
4. Conjecture: uniformly dominates unrestricted approach in terms of mean squared error – confirmed by Monte Carlo simulations cf. James-Stein shrinkage!

Data-driven predictions

- Direction where unrestricted estimates are precise:

$$(I_{J-1} \otimes x') \cdot \widehat{V} \cdot (I_{J-1} \otimes x')' \approx 0,$$

- Then

$$\begin{aligned}
\widehat{\delta}^{EB} \cdot x &= (I_{J-1} \otimes x') \cdot \widehat{\delta}^{EB}_{\uparrow} \\
&= (I_{J-1} \otimes x') \cdot \left[\widehat{\delta}_{\uparrow} + \widehat{V} \cdot \left(\widehat{\tau}^2 \cdot I_{J-1} \otimes P + \widehat{V} \right)^{-1} \cdot (\widehat{\beta}_0 \cdot \Delta_{\uparrow} - \widehat{\delta}_{\uparrow}) \right] \\
&\approx (I_{J-1} \otimes x') \cdot \widehat{\delta}_{\uparrow} = \widehat{\delta} \cdot x.
\end{aligned}$$

Summary – Who wins, who loses?

- Argue for interest in disaggregated welfare effects:

$$\gamma(y) = E[l \cdot \dot{w} | l \cdot w, \alpha]$$

- Need to identify causal effects conditional on endogenous outcomes:

$$g(x,\alpha) = E[\dot{x}|x,\alpha]$$

  from $f(x|\alpha)$.
- Key equation:

$$\dot{f} = -\nabla \cdot h$$

- With exclusion restrictions:

$$g^j(x,\alpha) = \partial_\alpha Q(v^j | v^1, \ldots, v^{j-1}, \alpha)$$

Summary – Empirical Bayes

- Object of interest: regressions with many regressors, one for each type of labor

$$\Delta \cdot Y_i = \Delta \cdot \gamma + \delta \cdot X_i + \Delta \cdot \varepsilon_i$$

- Restrictions of structural model: $\delta = \Delta \cdot \beta$,

$$\beta = \beta_0 \cdot M$$

- Structural model: inconsistent
  Unrestricted model: high variance
- Proposed solution: Empirical Bayes

$$\begin{aligned}\widehat{\delta_\uparrow} | \eta &\sim N(\delta_\uparrow, V) \\ \delta_\uparrow | \theta &\sim N(\beta_0 \cdot \Delta_\uparrow, \tau^2 \cdot I_{J-1} \otimes P),\end{aligned}$$

Thanks for your time!

Application: distributional impact of EITC

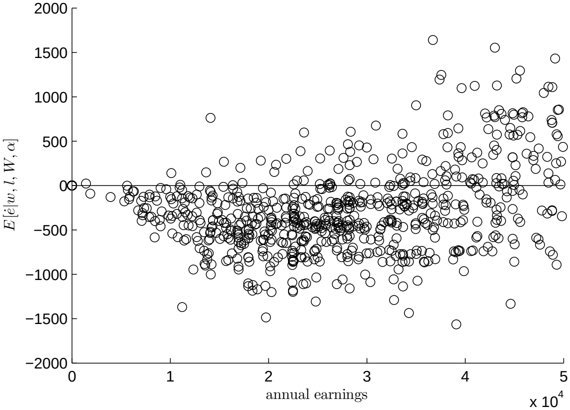
- Following Leigh (2010)
  (see also Meyer and Rosenbaum (2001), Rothstein (2010))
- CPS-MORG
- Variation in state top-ups of EITC
  across time and states
- $\alpha =$ maximum EITC benefit available
  (weighted average across family sizes)

State EITC supplements 1984-2002

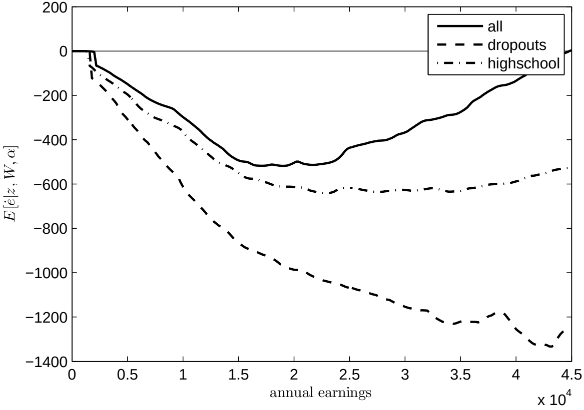
State:	CO	DC	IA	IL	KS	MA	MD	ME	MN	MN	NJ	NY	OK	OR	RI	VT	WI	WI	WI
# chld.							1+		0	1+		1+					1	2	3+
1984																	30	30	30
1985																	30	30	30
1986															22.21				
1987															23.46				
1988															22.96	23			
1989															22.96	25	5	25	75
1990			5												22.96	28	5	25	75
1991			6.5						10	10					27.5	28	5	25	75
1992			6.5						10	10					27.5	28	5	25	75
1993			6.5						15	15					27.5	28	5	25	75
1994			6.5						15	15		7.5			27.5	25	4.4	20.8	62.5
1995			6.5						15	15		10			27.5	25	4	16	50
1996			6.5						15	15		20			27.5	25	4	14	43
1997			6.5			10			15	15		20		5	27.5	25	4	14	43
1998			6.5		10	10	10		15	25		20		5	27	25	4	14	43
1999	8.5		6.5		10	10	10		25	25		20		5	26.5	25	4	14	43
2000	10	10	6.5	5	10	10	15	5	25	25	10	22.5		5	26	32	4	14	43
2001	10	25	6.5	5	10	15	16	5	33	33	15	25		5	25.5	32	4	14	43
2002	0	25	6.5	5	15	15	16	5	33	33	17.5	27.5	5	5	25	32	4	14	43

Welfare effects of wage changes induced by a 10% expansion of the EITC

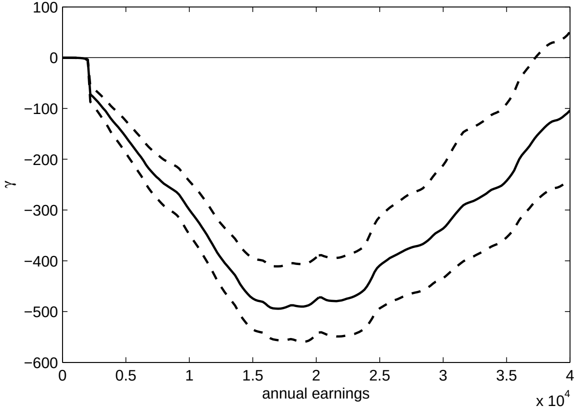
estimated welfare effect $l \cdot \dot{w}$ for a subsample of 1000 households, plotted against their earnings.



Kernel regression of welfare effects on earnings



95% confidence band for welfare effects given earnings



Back

Geometric interpretation

- Empirical Bayes maps $\widehat{\delta} \to \widehat{\delta}^{EB}$
- Geometry of this mapping?
- Simplifying assumptions (for exposition):
  1. variance of $\beta|\theta = \tau^2 \cdot I$ (no differencing)
  2. diagonal V (just a change of coordinates)
  3. $\beta_0 = 0$ (general case discussed in paper)
- Formally:

$$\widehat{\beta}|\beta \sim N(\beta, \mathrm{diag}(v))$$
$$\beta|\tau^2 \sim N(0, \tau^2 \cdot I),$$

Empirical Bayes in this simplified setting

- $\widehat{\beta}^{EB}$ given $\widehat{\tau}^2$:

$$\widehat{\beta}^{EB} = \text{diag}\left(\frac{\widehat{\tau}^2}{\widehat{\tau}^2 + v_k}\right) \cdot \widehat{\beta}.$$

- Does not directly give mapping $\widehat{\beta} \to \widehat{\beta}^{EB}$, since $\widehat{\tau}^2$ depends on $\widehat{\beta}$
- FOC for $\widehat{\tau}^2$:

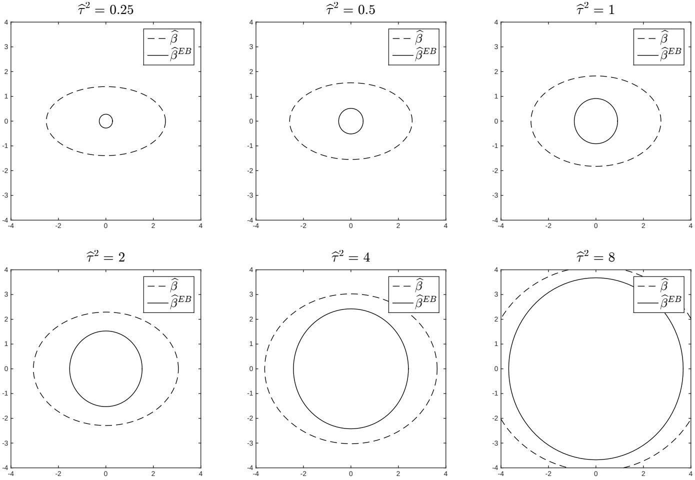
$$\sum_k \frac{1}{\widehat{\tau}^2 + v_k} = \sum_k \frac{\widehat{\beta}_k^2}{(\widehat{\tau}^2 + v_k)^2}.$$

Suppose $\widehat{\tau}^2$ is given

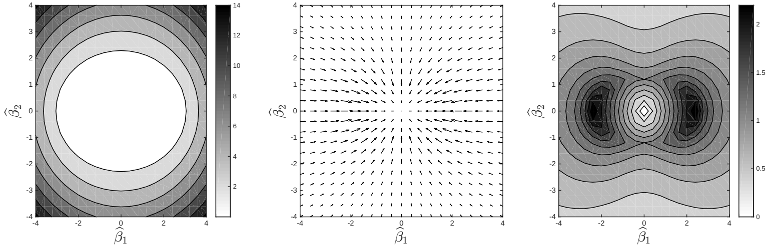
1. What's the set of $\widehat{\beta}$ yielding this $\widehat{\tau}^2$? **Ellipsoid** with semi-axes of length
$$(\widehat{\tau}^2 + v_k) \cdot \sqrt{\sum_{k'} \frac{1}{\widehat{\tau}^2 + v_{k'}}}.$$
2. What's the corresponding set of $\widehat{\beta}^{EB}$? **Circle** (!) of radius
$$\widehat{\tau}^2 \cdot \sqrt{\sum_{k'} \frac{1}{\widehat{\tau}^2 + v_{k'}}}.$$
3. $\widehat{\tau}^2$ is 0 inside ellipsoid with semi-axes of length
$$v_k \cdot \sqrt{\sum_{k'} \frac{1}{v_{k'}}}.$$

Fixing $\widehat{\tau}^2$

$\dim(\beta) = 2,\ v_1 = 2,\ v_2 = 1$



The mapping from $\widehat{\beta}$ to $\widehat{\tau}^2$ and $\widehat{\beta}^{EB}$



$\widehat{\tau}^2$, $\widehat{\beta}^{EB} - \widehat{\beta}$, and its length as a function of $\widehat{\beta}$

▸ Back

Some Monte Carlo evidence

- Comparing mean squared error of
  1. structural
  2. unrestricted
  3. empirical Bayes

  estimators.
- Two sets of simulations
  1. Conditional on θ:
     $\beta =$ structural model + random noise
  2. Conditional on η:
     β fixed

MSE relative to empirical Bayes conditional on θ

design parameters					MSE relative to empirical Bayes estimation		
n	J	σ^2	β_0	τ^2	structural	unrestricted	emp. Bayes
50	16	1.0	1.0	0.2	0.83	1.20	1.00
50	16	0.5	1.0	0.2	1.55	1.15	1.00
200	16	0.5	1.0	0.2	7.76	1.04	1.00
200	4	1.0	1.0	0.5	7.92	1.10	1.00
200	4	0.5	1.0	1.0	30.54	1.03	1.00

$$\beta = (\beta_{j,j'}) = \beta_0 \cdot M + \zeta$$
$$\zeta_{j,j'} \sim^{iid} N(0, \tau^2)$$

MSE relative to empirical Bayes conditional on η

design parameters						mean squared error		
n	J	σ^2	β_{00}	β_{01}	β_{02}	structural	unrestricted	emp. Bayes
200	4	1.0	1.0	1.0	1.0	0.18	1.47	1.00
200	4	1.0	1.0	1.0	6.0	15.04	1.03	1.00
200	4	1.0	0.0	1.0	6.0	19.37	1.01	1.00

$$\beta = \beta_{00} \cdot M_{J0} + \beta_{01} \cdot M_{J1} + \beta_{02} \cdot M_{J2},$$

- $M_{J0} = M$ in the first $J/4$ columns, zero elsewhere,
- $M_{J2} = M$ in the last $J/4$ columns, and zero elsewhere,
- $M_{J1} = M$ in the middle $J/2$ columns, and zero elsewhere.

Back