Learning from Survey Density Forecasts

Martín Almuzara Richard K. Crump Dennis Kongoli Gizem Koşar

FRBNY

NBER Summer Institute

Forecasting and Empirical Methods
July 2026

Eliciting probabilistic forecasts in surveys

- Density forecasts are now standard in surveys of forecasters, households, and firms:
  - Early example: Survey of Professional Forecasters (SPF).
  - Survey expectations are increasingly used to discipline models and test their predictions.
- More information and (possibly) cleaner interpretation than point forecasts:
  - Target of point forecasts is ambiguous: mean, median, mode?
  - Probabilistic forecasts go beyond central tendency: variance, skewness, multimodality.
- Typical question format:
  - Partition the support into a small number of intervals.
  - Respondents assign probabilities to each interval.
- Common: fit parametric density (e.g., generalized beta) to get moments/probabilities.

This paper

- We propose an approach to construct nonparametric lower and upper bounds for functionals of each respondent's subjective forecast density:
    - Measures of central tendency (mean, median).
    - Measures of uncertainty (variance, IQ range).
    - CDFs, quantile functions, skewness, etc.
- More flexible than parametric approach:
    - Avoids committing to a specific family of distributions.
    - Avoids restrictions on features such as skewness or number of modes.
    - Matches reported probabilities exactly.
- Mild smoothness restrictions deliver informative bounds.
- We also show how to use our bounds in subsequent regression analyses:
    - Valid estimation and inference when outcome, regressor, or both are interval-valued.

Selected literature

- **Recovering subjective distributions from probability questions:** Dominitz, Manski (1997), Engelberg, Manski, Williams (2009), Bellemare, Bissonnette, Kröger (2012), Bissonnette, De Bresser, van Soest, Vonkova (2017), Becker, Duersch, Eife, Glas (2021), Bassetti, Casarin, Del Negro (2024) ...
- **Partial identification and interval-valued data regression:** Manski, Tamer (2002), Beresteanu, Molchanov, Molinari (2011), Molinari (2020) ...

Outline

Setup and motivating example

SCE example: one-year ahead inflation

Question Q9. What is the percent chance that, over the next 12 months, prices will …

· decrease by 12% or more	π_1
· decrease by 8% to 12%	π_2
· decrease by 4% to 8%	π_3
· decrease by 2% to 4%	π_4
· decrease by 0% to 2%	π_5
· increase by 0% to 2%	π_6
· increase by 2% to 4%	π_7
· increase by 4% to 8%	π_8
· increase by 8% to 12%	π_9
· increase by 12% or more	π_{10}
Total	$\sum_{k=1}^{10} \pi_k = 1$

- A response takes the form of a vector of bin probabilities $\pi = (\pi_1, \ldots, \pi_K)'$.

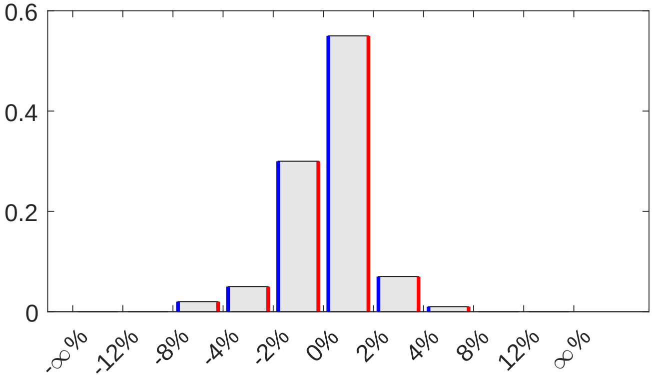
What we know and what we don't

- Denote bins as $A_k = (a_{k-1}, a_k]$ with probabilities $\pi_k = \mathrm{P}(X \in A_k)$.
- Reported probabilities identify the CDF at bin boundaries:

$$F(a_k) = \sum_{\ell \leq k} \pi_\ell.$$

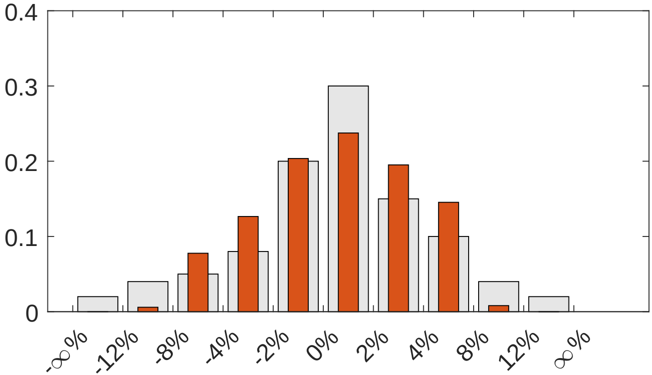
- But we do not know how probability is distributed within each bin:
  - Density, mean, median, variance, other functionals are generally not point identified.
- Following Engelberg, Manski, Williams (2009, EMW), two common routes:
  - Hands-off bounds ➡ impose no within-bin structure.
  - Parametric approach ➡ impose a specific family of densities (e.g., generalized beta).

Hands-off bounds



- Even in this favorable example, mean interval spans about two percentage points.
- With positive mass in tail bins, hands-off mean bounds are $(-\infty, \infty)$.

Parametric approach



- Even with perfect fit, many other densities are also consistent with the same π_k.
- Point identification therefore comes entirely from the assumed parametric family.

Learning from survey densities

Bounds for functionals

Formulation of the problem

- Take one respondent with unknown subjective density $f^\star$ for X.
    - Collect the reported probabilistic data into $\mathcal{D} = \{A_k, \pi_k\}_{k=1}^K$.

- The survey response is a finite set of probabilistic restrictions:

$$\pi_k = \mathrm{P}_{f^\star}(X \in A_k) = \int_{A_k} f^\star(x)\,dx, \qquad k = 1, \ldots, K.$$

- We want to learn about the functional

$$\theta^\star = \mathrm{E}_{f^\star}[G(X)] = \int G(x) f^\star(x)\,dx.$$

- Goal: match $\mathcal{D}$ exactly while using credible minimal structure on $f^\star$.

Bounds for functionals of survey densities

- Space $\mathcal{F}_{\mathcal{D}}$ of candidate densities consistent with the data and maintained restrictions.
- Core insight: $\theta^\star$ is bounded by

$$\underline{\theta} = \inf_{f \in \mathcal{F}_{\mathcal{D}}} \int G(x) f(x)\, dx, \qquad \overline{\theta} = \sup_{f \in \mathcal{F}_{\mathcal{D}}} \int G(x) f(x)\, dx.$$

  - Every $f \in \mathcal{F}_{\mathcal{D}}$ matches the respondent's reported probabilities in $\mathcal{D}$.
  - Bounds are sharp relative to the candidate density class $\mathcal{F}_{\mathcal{D}}$.
- If $\mathcal{F}_{\mathcal{D}}$ is defined by linear restrictions, these are functional linear programs.
- Bounds are respondent-specific, no cross-sectional pooling (yet).

Main elements: target functions

- Target function G is chosen by the researcher.
- Many objects of applied interest are linear density functionals of the form

$$\theta^\star = \int G(x) f^\star(x)\, dx.$$

  - Mean $\mathrm{E}_{f^\star}[\,X\,]$ obtained by $G(x) = x$.
  - Uncentered moment $\mathrm{E}_{f^\star}\left[\,X^j\,\right]$ obtained by $G(x) = x^j$.
  - CDF at a point $F^\star(a)$ obtained by $G(x) = \mathbf{1}\{\,x \leq a\,\}$.

- Other objects (variance, quantiles, etc.) are discussed below.

Main elements: candidate densities

- Baseline candidate space:

$$\mathcal{F}_{\mathcal{D}}(\phi) = \left\{ f : f \geq 0,\ \int f = 1,\ \int_{A_k} f = \pi_k \text{ for all } k,\ \left\| D^p f \right\|_\infty \leq \phi \right\}.$$

  - First three restrictions make f a density that matches the survey response exactly.
  - Last restriction imposes smoothness:

$$-\phi \leq D^p f(x) \leq \phi \qquad \text{for all } x,$$

  where D^p is p-th derivative and $\phi > 0$ controls this derivative.

- Tail restrictions are needed for unbounded supports/targets.

- ▸ Other restrictions (modes, monotonicity, rounding, etc.)

Computational strategy: discretized problem

- Assume (for now) a bounded support $\mathcal{X} = [\underline{x}, \overline{x}]$. Discretize $\mathcal{X}$ as

$$\mathcal{X}_M = \{x_0, \ldots, x_M\}, \qquad x_m = \underline{x} + mh, \qquad h = \frac{\overline{x} - \underline{x}}{M}.$$

- The discrete counterpart of density f are probability masses:

$$\widetilde{f} = (\widetilde{f}_0, \ldots, \widetilde{f}_M)'.$$

- Discretized candidate space:

$$\mathcal{F}_{\mathcal{D},M}(\phi) = \left\{ \widetilde{f} : \widetilde{f} \geq 0, \sum_{m=0}^{M} \widetilde{f}_m = 1, \sum_{x_m \in A_k} \widetilde{f}_m = \pi_k \text{ for all } k, \left\| \Delta_h^p \widetilde{f} \right\|_\infty \leq h\phi \right\},$$

where Δ_h^p is p-th discrete derivative.

Computational strategy: approximation result

- Discretized bounds solve two finite-dimensional linear programs:

$$\underline{\theta}_M(\phi) = \min_{\tilde{f} \in \mathcal{F}_{\mathcal{D},M}(\phi)} \sum_{m=0}^{M} G(x_m)\tilde{f}_m, \qquad \overline{\theta}_M(\phi) = \max_{\tilde{f} \in \mathcal{F}_{\mathcal{D},M}(\phi)} \sum_{m=0}^{M} G(x_m)\tilde{f}_m.$$

Proposition

Suppose $\mathcal{X} = [\underline{x}, \overline{x}]$ is compact, $\{A_k\}_{k=1}^{K}$ is a partition of $\mathcal{X}$, G is of bounded variation, and $\mathcal{F}_{\mathcal{D}}(\phi)$ and $\mathcal{F}_{\mathcal{D},M}(\phi)$ (for large enough M) are nonempty. Then, as $M \to \infty$,

$$\underline{\theta}_M(\phi) \longrightarrow \underline{\theta}(\phi), \qquad \overline{\theta}_M(\phi) \longrightarrow \overline{\theta}(\phi).$$

- This justifies approximating $(\underline{\theta}(\phi), \overline{\theta}(\phi))$ by discretized bounds $(\underline{\theta}_M(\phi), \overline{\theta}_M(\phi))$.
  - Good since finite-dimensional linear programs are easy to solve (even for large M).

A few remarks

- Proof sketch

- Bounded variation assumption covers $G(x) = x^j$ and $G(x) = \mathbb{1}\{x \leq a\}$.

- $\mathcal{F}_{\mathcal{D}}(\phi)$ and $\mathcal{F}_{\mathcal{D},M}(\phi)$ are continuous compact-valued correspondences, so $(\underline{\theta}(\phi), \overline{\theta}(\phi))$ and $(\underline{\theta}_M(\phi), \overline{\theta}_M(\phi))$ are continuous (and monotonic) in ϕ.
  - There is a minimum $\phi = \phi^*$ independent of G which gives tightest possible bounds.

- The argmax/argmin $\widetilde{f}$ are generally not interpretable. Focus on the bounds.

- Approximation result extends to unbounded $\mathcal{X}$ under tail constraints + truncation:
  - Suppose $A_1 = (-\infty, \underline{a}]$ and $A_K = [\overline{a}, +\infty)$.
  - Tail constraint: $f(x) \leq \gamma(x)$ all $x \in A_1, A_K$, for a known γ with $\int_{A_1 \cup A_K} |G\gamma| < \infty$.
  - Bounds obtained by discretizing $\mathcal{X}_R = [\underline{a} - R, \overline{a} + R]$, letting $M, R \to \infty$.

Learning from survey densities

Other features of the distribution

Centered moments: sharp bounds

- For $f_1, f_2 \in \mathcal{F}_{\mathcal{D}}$, define

$$m_j(f_1, f_2) = \mathrm{E}_{f_2}\left[\left(X - \mathrm{E}_{f_1}[X]\right)^j\right] = \int \left(x - \int x f_1(x)\,dx\right)^j f_2(x)\,dx.$$

  This function is linear in f_2 but nonlinear in f_1 whenever $j > 1$.

- The j-th centered moment is $\theta^\star = m_j(f^\star, f^\star)$.
  - For example, $j = 2$ recovers the variance of X.

- Sharp bounds for $\theta^\star$ are

$$\underline{\theta}^s = \inf_{f \in \mathcal{F}_{\mathcal{D}}} m_j(f, f), \qquad \overline{\theta}^s = \sup_{f \in \mathcal{F}_{\mathcal{D}}} m_j(f, f).$$

- Functional polynomial programs, computationally hard even after discretization.

Centered moments: sequential bounds

- Instead, we propose to solve multiple programs:

$$\underline{\theta} = \inf_{f_1 \in \mathcal{F}_{\mathcal{D}}} \left\{ \inf_{f_2 \in \mathcal{F}_{\mathcal{D}}} m_j(f_1, f_2) \right\}, \qquad \overline{\theta} = \sup_{f_1 \in \mathcal{F}_{\mathcal{D}}} \left\{ \sup_{f_2 \in \mathcal{F}_{\mathcal{D}}} m_j(f_1, f_2) \right\}.$$

  - Given $\mathrm{E}_{f_1}[X]$, the inner problem is linear.
  - The outer problem searches over the interval for the mean $\mathrm{E}_{f_1}[X]$, also linear.

- Discretized version enjoys the same guarantees in the proposition above.

- Resulting bounds are valid but conservative:

$$\underline{\theta} \leq \underline{\theta}^s \leq \theta^\star \leq \overline{\theta}^s \leq \overline{\theta}.$$

- Similar approach works for ratios $m_j(f_1, f_3)/\{m_2(f_1, f_2)\}^{j/2}$ (skewness/kurtosis coef's).

Learning other features: quantiles

- Quantiles are also nonlinear functionals, but can be bounded through the CDF.
- For each x, compute bounds on $F(x) = \mathrm{P}(X \leq x)$ using target $G_x(\tilde{x}) = \mathbb{1}\{\tilde{x} \leq x\}$,

$$\underline{F}(x) \leq F(x) \leq \overline{F}(x).$$

- Valid lower/upper bounds for the τ-quantile use the upper/lower CDF envelopes:

$$\underline{Q}(\tau) = \inf\{x : \overline{F}(x) \geq \tau\}, \qquad \overline{Q}(\tau) = \inf\{x : \underline{F}(x) \geq \tau\}.$$

  These bounds are valid (but slightly conservative).

- Same idea works for interquartile range and quantile-based skewness measures.

Implementation and SCE illustration

Implementation: smoothness, tails and grid

- We impose smoothness through fourth discrete derivatives, $p = 4$.

- Baseline choice: fit pilot densities and set ϕ to the maximum absolute fourth derivative of the best-fitting pilot.
    - Candidate pilots: generalized beta and skew normal. Denote choice by ϕ_{pilot}.

- We also report results using the tightest bounds based on minimum ϕ:

$$\phi_{\text{min}} = \min\{\phi : \mathcal{F}_{\mathcal{D},M}(\phi) \neq \emptyset\}.$$

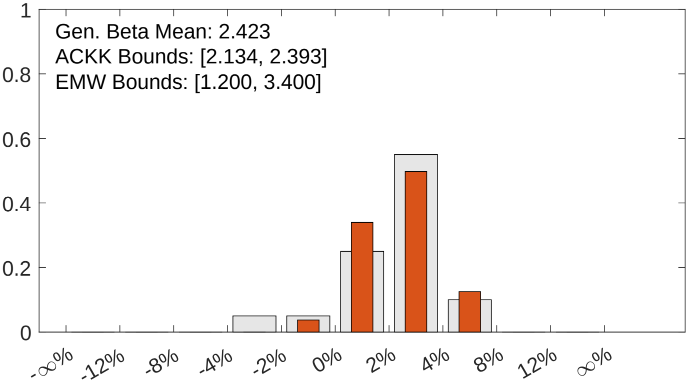
- To deal with open-ended bins, we use exponential tail envelopes:
    - Calibrate two rate parameters $\lambda_{\text{LT}}, \lambda_{\text{UT}}$ to control probability of extreme outliers. ▸ Tails

- $M + 1 \leq 1{,}000$ gridpoints to balance approximation quality/numerical stability.

SCE illustration: data and questions

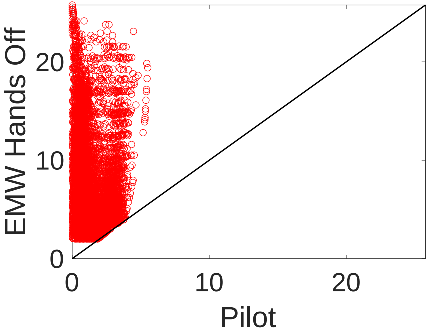
- Data: one-year-ahead inflation expectations from the Survey of Consumer Expectations.
- Restrictions:
    - Drop responses assigning 100% probability to a single bin.
    - Drop responses with zero probability between adjacent nonzero-probability bins.
    - Needed for parametric estimation to be well defined.
- Sample: June 2013 through June 2024.
    - Total after restrictions: $140{,}942$ responses.
    - $65{,}423$ responses place positive probability in at least one extreme bin.
- Questions:
    - How much do smoothness restrictions tighten EMW hands-off bounds?
    - How do parametric point estimates compare with our nonparametric bounds?
    - How sensitive are the bounds to ϕ?

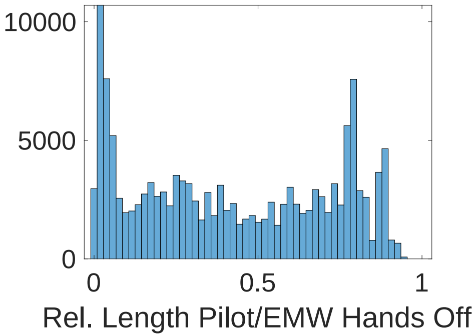
Individual responses



- Our bounds can be substantially tighter than EMW bounds while matching the data.
- Generalized-beta mean often falls outside the nonparametric interval.

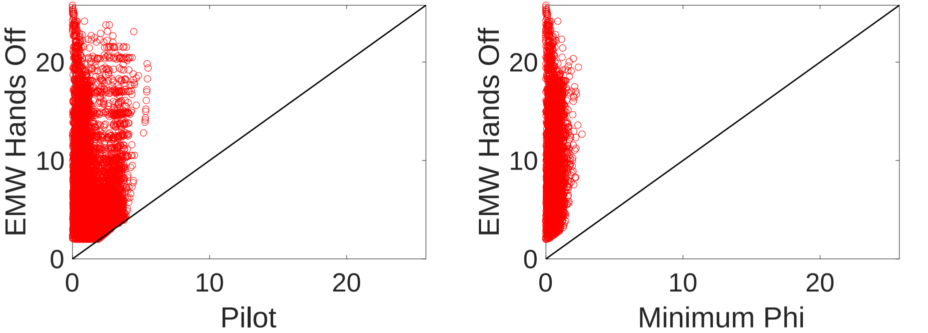
Mean bounds: gains relative to EMW





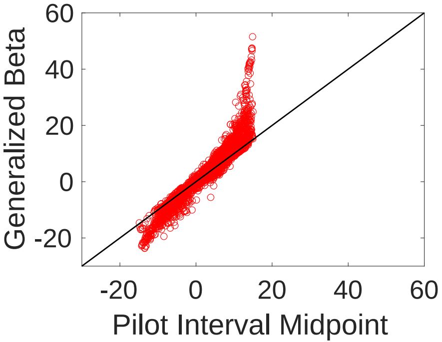
- By construction, ϕ_{pilot}-bounds nested in EMW, with substantial precision gains.
- Median interval length reduction > 1.7pp. It is > 4.1pp for 25% of respondents.

Mean bounds: role of smoothness constraint



- Using the tightest ϕ_{min}-bounds substantially narrows the identified interval.
- They can be as wide as 2.6pp, so the data do not generally point-identify the mean.

Mean bounds: parametric estimates versus bounds



- Generalized-beta mean differs materially from the midpoint of the ϕ_{pilot}-bounds.
- Frequency of generalized-beta mean in interval: 71.5% (ϕ_{pilot}), 11.9% ($\phi_{\min}$)

Regressions with bounds

Regression analysis with bounds

- Survey density functionals are often used in regression analyses downstream.
  - Question: how to account for the fact that we only observe bounds of functionals?

- Panel model for (possibly) set-valued outcome/regressors:

$$y_{it}^{\star} = \beta' x_{it}^{\star} + \alpha_i + v_{it}, \qquad \mathrm{E}\left[\, v_{it} \mid \alpha_i, x_{i1}^{\star}, \ldots, x_{iT}^{\star} \,\right] = 0,$$

  where $\underline{y}_{it} \le y_{it}^{\star} \le \overline{y}_{it}$ and $\underline{x}_{it} \le x_{it}^{\star} \le \overline{x}_{it}$.

- Object of interest is β. Examples:
  - Do older cohorts adjust inflation expectations more in response to shocks?
  - What is the coefficient of relative prudence in an Euler equation?

- We develop computationally feasible, valid inference tools for this problem. ▸ Details

Application: consumption Euler equation

- Consumption Euler equation:

$$\left(\frac{1+r_i}{1+\delta_{i,t}}\right) \mathrm{E}_i\left[\, U'(C_{i,t+1}) \,\right] = U'(C_{i,t}).$$

Literature using subjective expectations: Bertola, Guiso, Pistaferri (2005), Christelis, Georgarakos, Jappelli, van Rooij (2020), Crump, Eusepi, Tambalotti, Topa (2022).

- Second-order expansion uncovers prudence parameter ρ (Kimball, 1990):

$$\underbrace{\mathrm{E}_i\left[\, \mathrm{GC}_{i,t+1} \,\right]}_{y^\star_{it}} \approx \underbrace{\rho}_{\beta} \underbrace{\frac{\mathrm{E}_i\left[\, \mathrm{GC}^2_{i,t+1} \,\right]}{2}}_{x^\star_{it}} + \underbrace{\frac{1}{\xi}\left(\frac{r_i - \delta_{i,t}}{1+r_i}\right)}_{\alpha_i + v_{it}}.$$

- Both LHS/RHS come from the same subjective spending-growth density.

Conclusion

Conclusion

- We introduce a new approach to construct nonparametric lower and upper bounds from survey-elicited densities:
  - Avoids strong restrictions on unknown density (e.g., assuming generalized beta).
  - Satisfies the restrictions implied by the survey responses exactly.
  - Easy to implement computationally.

- We also introduce methods to incorporate these bounds into regression analyses.

- In separate work, we are exploring extensions to multivariate functionals such as covariances, correlations, and joint CDFs:
  - Additional survey questions in a special SCE module provide information about dependence complementing marginal probabilistic forecasts.

Thank you!

Backup slides

Backup: additional restrictions

- Tail constraints. If $\mathcal{X}$ and G are unbounded, bounds may be uninformative.
  - Known dominating function γ with $|G\gamma|$ integrable, and $T \subseteq \mathcal{X}$ with $\mathcal{X} \setminus T$ compact.
  - Then impose
  $$f(x) \leq \gamma(x), \qquad \text{for all } x \in T.$$
  - This is an infinite collection of linear inequality constraints on f.

- Monotonicity. Suppose the q-th derivative of f is non-increasing over a set $\mathcal{M}$.
  - Then impose
  $$D^q f(x) - D^q f(x') \geq 0, \qquad \text{for all } x, x' \in M \text{ such that } x < x'.$$
  - In particular, $q = 1$ and $\mathcal{M} = \mathcal{X}$ imposes unimodality (with possibly flat regions).

Backup: additional restrictions

- Location of the mode. Some surveys ask for the mode $\widetilde{x}$ of the forecast density.
  - This information can be imposed through

$$f(\widetilde{x}) - f(x) \geq 0, \qquad \text{for all } x \in \mathcal{X}.$$

  - This is again an infinite collection of linear inequality constraints on f.

- Rounding bias. Suppose reported probability π_k may be rounded by an amount less or equal to r, known to the researcher.
  - Replace the exact-fit restriction

$$\int_{A_k} f(x)\,dx = \pi_k$$

    with the inequalities

$$\pi_k - r \leq \int_{A_k} f(x)\,dx \leq \pi_k + r.$$

  - These restrictions remain linear in f.

Back

Backup: proof of Proposition 1

- It is enough to prove the result for the lower bound:

$$\underline{\theta}_M(\phi) \longrightarrow \underline{\theta}(\phi),$$

since the upper bound is entirely analogous.

- The proof has two steps:

$$\limsup_{M\to\infty} \underline{\theta}_M(\phi) \leq \underline{\theta}(\phi), \qquad \liminf_{M\to\infty} \underline{\theta}_M(\phi) \geq \underline{\theta}(\phi).$$

- Limit superior: show that any density $f \in \mathcal{F}_{\mathcal{D}}(\phi)$ can be discretized to a vector $\widetilde{f}$ that (approximately) satisfies the constrains in $\mathcal{F}_{\mathcal{D},M}(\phi)$.

- Limit inferior: smoothness constraints for discrete optimizers in $\mathcal{F}_{\mathcal{D},M}(\phi)$ are preserved under weak convergence to the continuous problem.

Backup: proof of Proposition 1

- Fix $f \in \mathcal{F}_{\mathcal{D}}(\phi)$ and define its scaled discretization:

$$\widetilde{f}_m^{(M)} = hf(x_m).$$

- Preserves smoothness (approx) but may not satisfy bin probabilities exactly.
- Rescale within each bin:

$$\widehat{f}_m^{(M)} = \widetilde{f}_m^{(M)} \frac{\pi_{k(m)}}{\sum_{x_j \in A_{k(m)}} \widetilde{f}_j^{(M)}},$$

  where $k(m)$ is the bin containing x_m.
- Now, $\widehat{f}^{(M)}$ satisfies the probabilistic constraints exactly.

Backup: proof of Proposition 1

- Smoothness constraint for $\widehat{f}^{(M)}$ holds with parameter $\phi_h = \phi + O(h)$, and changing ϕ_h back to ϕ changes the value of the linear program by $O(h) = o(1)$.

- Since the product Gf is of bounded variation,

$$\sum_{m=0}^{M} G(x_m)\widehat{f}_m^{(M)} = \int G(x)f(x)\,dx + O(h).$$

- Hence $\limsup_{M\to\infty} \underline{\theta}_M(\phi) \leq \int G(x)f(x)\,dx$, and taking infimum over f gives

$$\limsup_{M\to\infty} \underline{\theta}_M(\phi) \leq \underline{\theta}(\phi).$$

Backup: proof of Proposition 1

- Let $\widehat{f}^{(M)} \in \mathcal{F}_{\mathcal{D},M}(\phi)$ attain $\underline{\theta}_M(\phi)$, and define the measure

$$\mu_M = \sum_{m=0}^{M} \widehat{f}_m^{(M)} \delta_{x_m}$$

  where δ_x is Dirac mass on x.

- Compactness of probability measures on $\mathcal{X}$ gives a weakly convergent subsequence

$$\mu_M \implies \mu.$$

- The bin probability constraints pass to the limit.

Backup: proof of Proposition 1

- Define $g^{(M)}$ as the piecewise-constant interpolant of

$$h^{-1}\Delta_h^p \widehat{f}^{(M)}.$$

- The discrete smoothness constraint implies $\|g^{(M)}\|_\infty \leq \phi$, so along a further subsequence, $g^{(M)}$ converges weakly to some g with $\|g\|_\infty \leq \phi$.

- Summation by parts shows that μ has a density f with $D^p f = g$. Hence, $f \in \mathcal{F}_{\mathcal{D}}(\phi)$.

- Since G is of bounded variation,

$$\liminf_{M\to\infty} \underline{\theta}_M(\phi) \geq \int G(x)f(x)\,dx \geq \underline{\theta}(\phi).$$

▸ Back

Backup: implementation of tail restrictions

- We impose exponential tail envelopes of the form:

$$\widetilde{f}_m \leq \widetilde{f}_{\text{low}} e^{-\lambda_{\text{LT}}(x_{\text{low}} - x_m)} \quad \text{for } x_m \leq x_{\text{low}}, \qquad \widetilde{f}_m \leq \widetilde{f}_{\text{upp}} e^{-\lambda_{\text{UT}}(x_m - x_{\text{upp}})} \quad \text{for } x_m \geq x_{\text{upp}},$$

  where $x_{\text{low}}, x_{\text{upp}}$ are boundary gridpoints, $\widetilde{f}_{\text{low}}, \widetilde{f}_{\text{upp}}$ are the corresponding entries of $\widetilde{f}$.

- Following SCE practice, we calibrate rate parameters $\lambda_{\text{LT}}, \lambda_{\text{UT}}$ so that

$$f(-38) \leq 10^{-4}, \qquad f(38) \leq 10^{-4},$$

  controlling the subjective probability of extremely low/high inflation outcomes.

- Tail restrictions are omitted when respondents put zero mass on tail bins.

Back

Regressions with bounds: setup

- Latent panel regression:

$$y_{it} = x_{it}'\beta + \alpha_i + v_{it}, \qquad \mathrm{E}\left[v_{it} \mid \alpha_i, \{x_{i\tau}\}_{\tau=1}^{T} \right] = 0.$$

  Fixed-effects perspective: unrestricted distribution of α_i conditional on the regressors.

- The latent variables are not directly observed:

$$\underline{y}_{it} \leq y_{it} \leq \overline{y}_{it}, \qquad \underline{x}_{it} \leq x_{it} \leq \overline{x}_{it}.$$

- For each unit, define compact convex sets

$$\mathcal{Y}_i = \left\{ y \in \mathbb{R}^T : \underline{y}_{it} \leq y_t \leq \overline{y}_{it} \; \forall t \right\}, \qquad \mathcal{X}_i = \left\{ X \in \mathbb{R}^{T \times K} : \underline{x}_{it} \leq x_t \leq \overline{x}_{it} \; \forall t \right\}.$$

Identifying information

- Let $D \in \mathbb{R}^{(T-1)\times T}$ eliminate fixed effects: $D1_{T\times 1} = 0_{(T-1)\times 1}$.
- Applying D gives
$$Dy_i = DX_i\beta + Dv_i, \qquad \mathrm{E}[\,Dv_i \mid X_i\,] = 0.$$
- For a matrix-valued function $G(\cdot)$,
$$\mathrm{E}\left[\,G(X_i)'D(y_i - X_i\beta)\,\right] = 0.$$
- With interval-valued data, this becomes a random set of model predictions:
$$Q_i(b;G) = \left\{G(X)'D(y - Xb) : y \in \mathcal{Y}_i,\ X \in \mathcal{X}_i\right\}.$$
- Identified set for fixed G is $B_G = \{b : 0 \in \mathrm{E}_A[\,Q_i(b;G)\,]\}$.
(Beresteanu, Molchanov, Molinari, 2011)

Support-function representation

- The support function of $Q_i(b; G)$ is

$$h_i(b, u) = \max_{y \in \mathcal{Y}_i,\, X \in \mathcal{X}_i} u' \{G(X)' D y - G(X)' D X b\}.$$

- The identified set can be represented as

$$B_G = \left\{ b : \min_{\|u\| \leq 1} \mathrm{E}[\, h_i(b, u)] \geq 0 \right\}.$$

- Computational bottleneck: evaluate $h_i(b, u)$ for each b and each direction u.

- If direct evaluation is hard, use an envelope $\widetilde{h}_i$ satisfying

$$\widetilde{h}_i(b, u) \geq h_i(b, u).$$

- This yields an outer identified set $B_G \subseteq \widetilde{B}_G$, so inference is valid but more conservative.

Inference by moment inequalities

- Define the sample mean and variance of the envelope:

$$\bar{h}_N(b,u) = \frac{1}{N}\sum_{i=1}^{N} \tilde{h}_i(b,u), \qquad \hat{\sigma}_N^2(b,u) = \frac{1}{N}\sum_{i=1}^{N}\left(\tilde{h}_i(b,u) - \bar{h}_N(b,u)\right)^2.$$

- Test statistic:

$$T_N(b) = \max_{\|u\|=1}\left[-\frac{\sqrt{N}\,\bar{h}_N(b,u)}{\hat{\sigma}_N(b,u)}\right]_+.$$

  Only $\|u\| = 1$ is needed because the support function is positively homogeneous in u.

- In the scalar case, $u \in \{-1, 1\}$:

$$T_N(b) = \max\left\{0,\ -\frac{\sqrt{N}\,\bar{h}_N(b,-1)}{\hat{\sigma}_N(b,-1)},\ -\frac{\sqrt{N}\,\bar{h}_N(b,1)}{\hat{\sigma}_N(b,1)}\right\}.$$

Confidence sets

- A $(1-\alpha)$ confidence set is obtained by test inversion:

$$\mathrm{CS}_\alpha = \{b : T_N(b) \leq c_{1-\alpha}(b)\}.$$

  where $c_{1-\alpha}(b)$ is the $(1-\alpha)$-quantile of the sampling distribution of $T_N(b)$.

- Critical values can be based on asymptotic or bootstrap approximations for conditional moment inequalities. (Andrews, Shi, 2015)

- Moment selection and recentering used to improve power. (Andrews, Jia, 2012; Romano, Shaik, Wolf, 2014)

- Practical tradeoff:
  - Exact support functions give sharper inference but can be computationally hard.
  - Envelopes are faster and preserve validity, at the cost of conservativeness.

Back