Cockrell School of Engineering

Modeling Uncertainty: Energy Policy Applications and Quantile-Based Distribution Systems

Connor Colombe
Operations Research and Industrial Engineering
The University of Texas at Austin
Committee Members:
Ben Leibowicz
Eric Bickel
John Hasenbein
Andrew Waxman

Acknowledgments

Benjamin D. Leibowicz

Co-Chair

J. Eric Bickel

Co-Chair

John Hasenbein

Member

Andrew Waxman

Member

- **Funding:** Alfred P. Sloan Foundation, Cockrell School of Engineering, Center for Engineering and Decision Analytics, University of Texas Graduate School
- Thank you to my family, friends, and everyone in the audience today!

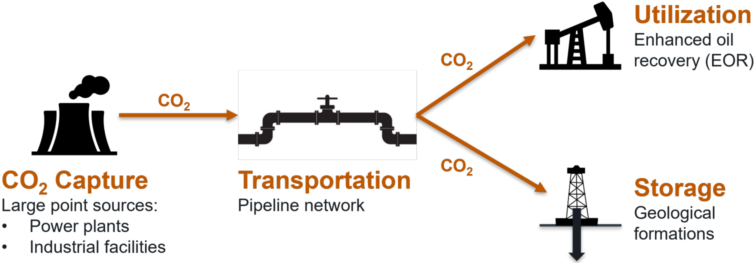
Outline

Table of Contents

Motivation: Uncertainty Is Central to Decision-Making

- Decision-makers routinely face uncertainty: future policy, costs, outcomes etc.
- This dissertation develops models and tools for quantifying and managing uncertainty:
  - **Applied:** How does uncertainty shape the adoption of carbon capture technology and the design of subsidies to accelerate it? (Chapters 2–3)
  - **Methodological:** How do we construct flexible probability distributions from limited expert-assessed data? (Chapters 4–5)
- Building the CCUS models in Chapters 2–3 made clear that the *modeling of uncertainty itself* plays a critical role in the decisions those models recommend

Background: Carbon Capture, Utilization, and Storage (CCUS)



Background: Why CCUS Matters

- CCUS is a promising technology to reduce CO_2 emissions in electricity generation and heavy industries (chemicals, steel, cement).

- Results from EMF 37 (Binsted et al., 2024) show that:
  - Most models find net-zero is impossible without some form of CCUS.
  - Achieving net-zero without CCUS would be 2–10X more costly.

- Most IPCC scenarios limiting warming to 1.5–2°C include significant CCUS deployment (IPCC, 2018, 2023).

Background: U.S.Incentives for Carbon Capture (45Q)

- A subsidy payment for each ton of CO_2 captured at qualifying facility
  - either stored geologically or utilized for an industrial process (e.g., enhanced oil recovery (EOR))

- First introduced in 2008

- Generally has enjoyed bipartisan support and increased over time
- Simple version: Capture a ton of $CO_2 \rightarrow$ U.S. government pays you $85

Table of Contents

Motivation and Research Contributions

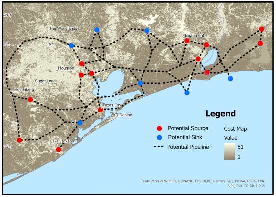
- Historically, CCUS development has been slow, why?
  - 45Q values have changed numerous times and are subject to change in the future
  - CCUS infrastructure is expensive and takes a decade or more to recoup costs

- *Might some combination of policy uncertainty (45Q values) and investor risk aversion be holding the development of CCUS infrastructure back?*
- We built a novel optimization model (a multi-period, two-stage stochastic program) to investigate the extent to which policy uncertainty and investor risk aversion may explain the limited buildout of CCUS infrastructure

Methodology: Optimization Model

Mixed-integer linear program (MILP) for CCUS infrastructure network optimization

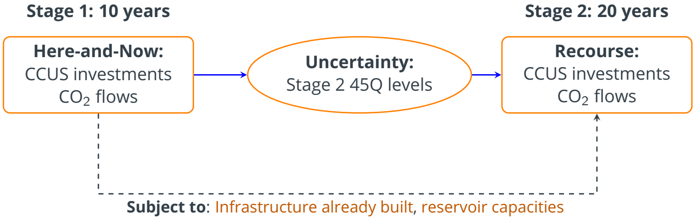
- **Objective:** Maximize expected profit
  - Revenues from 45Q + CO_2 sales for EOR
- **Policy:** 45Q levels for storage and utilization
- **By choosing:**
  - Which facilities, storage sites, pipelines to open (and when)
  - Quantities of CO_2 to capture, transport, store
- **Outputs of interest:**
  - Infrastructure investments (and timing)
  - Expected total CO_2 captured



Candidate investments for our Texas-Louisiana Gulf Coast case study

New Features: Policy Uncertainty and Risk Aversion

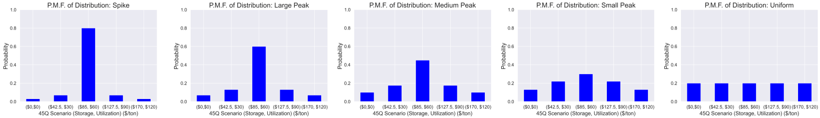
- Uncertainty in future 45Q tax credit levels captured through two-stage stochastic program



- **Risk aversion:** replace expected profit with expected utility using $u(w) = \frac{w^{1-\eta}}{1-\eta}$, where larger $\eta \rightarrow$ more risk averse

Case Study: How Do We Represent Uncertainty?

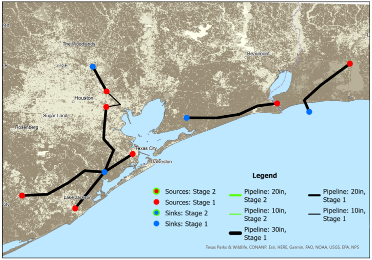
- We have a set of distributions to represent the level of "uncertainty" about the future 45Q values
- Each distribution assigns probabilities to a set of 5 hypothetical future 45Q values (scenarios).



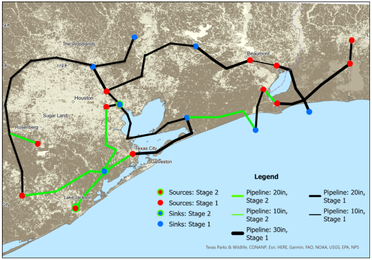
Increasing Uncertainty →

Results: Deterministic Policy Scenarios

- Here we see the infrastructure developments when the second stage incentive is known



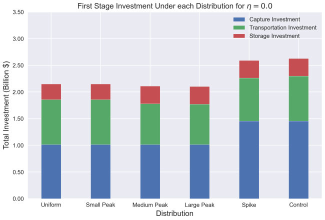
Stage 2 Incentives = ($0, $0)



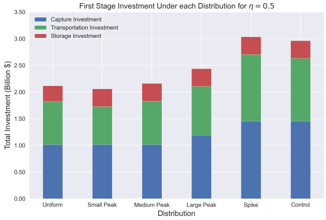
Stage 2 Incentives = ($127.50, $90)

- Even if 45Q were to disappear in a decade, current* values are sufficient to spur near-term investment

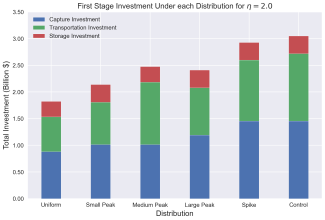
Results: Policy Uncertainty and Risk Aversion



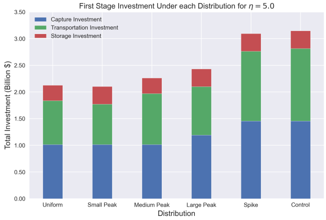
$\eta = 0.0$ (risk neutral)



$\eta = 0.5$



$\eta = 2.0$



$\eta = 5.0$ (highly risk averse)

Main Takeaways

- The current 45Q levels are sufficiently generous to drive meaningful CCUS investment for the next decade, even if they are certain to be phased out after 10 years
- Utilization opportunities tend to be leveraged first, then geological storage sites
- Policy uncertainty and risk aversion reduce expected CO_2 captured by at most 16%, with incentive uncertainty appearing to be a more important factor than risk aversion

Table of Contents

Research Questions

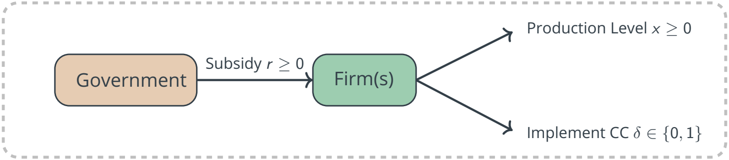
- Are the current levels of the 45Q tax credits about right, too low, or too generous?
- What CC subsidy level maximizes social welfare?
- Under what conditions would a CC subsidy lead to undesirable outcomes, such as higher CO_2 emissions?
- How does government uncertainty about firms' investment costs of CC affect the optimal subsidy level and associated outcomes?

Research Contributions

- First analytical exploration of optimal subsidies for carbon capture using a Stackelberg game model
- Analytically derive the optimal CC subsidy and establish conditions under which subsidizing CC can perversely increase CO_2 emissions
- Show how government uncertainty about firm investment costs shapes the optimal subsidy level
- Perform a numerical case study on a coal power plant:
  - optimal subsidy is $83/ton with full information and $88/ton with cost uncertainty
  - very close to the current 45Q level of $85/ton

Methodology

- **Stackelberg (sequential) game model**: Government sets the level of the 45Q subsidy, then the firm chooses (i) whether to implement CC or not and (ii) its production level



- The government can anticipate the best response of the firm to any subsidy offered, and sets the subsidy level accordingly to maximize social welfare

Methodology: Uncertainty and Case Study

- Uncertainty extension: firm's CC investment cost modeled as a random variable in the government's problem
- Numerical case study: firm considers retrofitting its coal-fired power plant with CC

Coal-fired power plant: Petra Nova

Analytical Results: Full Information

Proposition 1 (Subsidy Threshold)

Under quadratic costs, the unique threshold subsidy is

$$\hat{r} := \eta + \frac{p(\sqrt{2aI/p^2 + 1} - 1)}{\varepsilon}$$

Proposition 3 (Threshold or Nothing)

The optimal subsidy r^* is either $r^* = \hat{r}$ (just enough to induce CC investment) or any $r^* < \hat{r}$ (no investment).

Analytical Results: Full Information (cont.)

Proposition 4 (Optimal Subsidy Level)

$\hat{r}$ maximizes social welfare if and only if

$$p(p-\beta)+2al \leq \sqrt{2al+p^2}\,(p-(1-\varepsilon)\beta-\varepsilon\eta)$$

Proposition 5 (CO_2 Emissions Impact)

The threshold subsidy leads to a net decrease in emissions iff

$$\varepsilon > \hat{\varepsilon} := 1 - \frac{1}{\sqrt{\frac{2la}{p^2}+1}}$$

Analytical Results: Cost Uncertainty

Proposition 8 (Effect of Uncertainty on Optimal Subsidy)

Under cost uncertainty $\tilde{I} \sim U([I - \rho, I + \rho])$, the optimal subsidy is greater than $\hat{r}$ when $\rho < \hat{\rho}$, and less than $\hat{r}$ when $\rho > \hat{\rho}$.

- In our case study, $\hat{\rho}$ (uncertainty threshold) is large enough that the optimal subsidy increases under all practically reasonable parameterizations
- Firms may have a strategic incentive to obscure their true costs

Case Study: Coal-Fired Power Plant

Metric	No Subsidy (control)	Full Information	Cost Uncertainty
Subsidy Level ($/t CO_2)	0	83.4	87.8
Firm Production Level	131.4	147.1	155.2
Firm Profit (Billion $)	3.9	3.9	4.5
CO_2 Emissions (Million t)	131.4	22.1	23.9
Gov. Expenditure (Billion $)	0	10.4	11.6
Social Welfare (Billion $)	−11.8	−8.1	−8.9

- Optimal subsidy is $83/ton with full information and $88/ton with cost uncertainty
- Emissions threshold requires $\hat{\varepsilon} > 10\%$, easily satisfied in practice. Substantial emissions reduction ($\sim$ 80%) in either case
- Firm profit increases by $600M under uncertainty, confirming the incentive to obscure costs

Key Findings and Takeaways

- In a world with perfect information, the government maximizes social welfare by offering a CC subsidy that is just high enough to induce investment
  - Numerical case study suggests current 45Q level of $85/ton may be very close to the optimal threshold
- Theoretically possible for a CC subsidy to increase CO_2 emissions (Prop. 5), but unlikely
- The firm may have an incentive to obscure its true investment cost

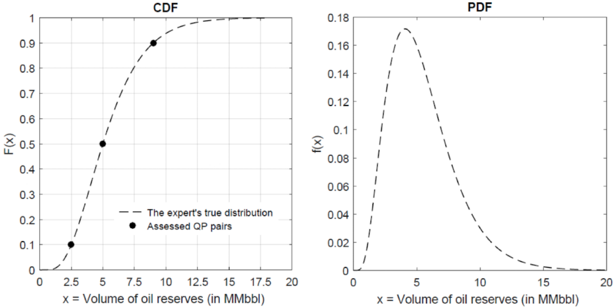
Table of Contents

Motivation: From Uncertainty to Distributions

- As we have seen, uncertainty plays a central role in CCUS decision-making:
  - Shifting infrastructure investment levels, raising optimal subsidy levels, and affecting policy outcomes
- But how do we represent that uncertainty in the first place?
- In practice, experts rarely provide a full probability distribution. Instead, an analyst elicits a finite number of quantile-probability pairs:
  - "The 10th percentile of project cost is $530M, the median is $620M, the 90th percentile is $810M."
- **The problem:** How do we construct a valid probability distribution that is consistent with these assessments?

Motivation

- Given a set of points on an expert's CDF curve (unknown to you), what is the best guess for the expert's distribution?
- Quantile parameterized distributions (QPDs) aim to specify distributions based on a set of input quantiles



Background: Quantile Functions

- Instead of working with CDF's, it is easier to work with quantile functions.

Definition

Let X be a random variable with CDF $F(x)$. The *quantile function* (QF) is

$$Q(p) := \inf\{x : F(x) \geq p\}, \quad p \in (0,1).$$

If F is continuous and strictly increasing on its support then

$$Q(p) = F^{-1}(p), \quad p \in (0,1).$$

- The quantile function is simply the inverse of the CDF (axes are swapped).
- The only requirement of a quantile function is that it is increasing on $(0,1)$.

How Should We Proceed?

- We would like to specify a probability distribution based on a finite set of assessments. There are several approaches:
- **Discretize** the distribution with three points and weight the P10, P50, P90 by 0.25, 0.50, 0.25, respectively, or some other weighting (Hammond and Bickel, 2013).
  - → **Introduces discretization error.**
- **Fit a known parametric family**, such as a Beta distribution, to these assessments.
  - → **May not match the assessed points.**
- **Specify a quantile-parameterized distribution (QPD)** that matches these assessments exactly.
  - → **QPD may not be feasible (monotonic).**

Research Objectives

- Design a new QPD system with simple, interpretable conditions for monotonicity that can be easily verified or enforced ex ante, avoiding complex post-fit numerical repair.
- Establish key theoretical properties: full-rank design matrix and exact interpolation of any finite set of quantile assessments
- Derive transparent conditions that control monotonicity and modality directly through coefficient constraints.
- Benchmark against the a leading QPD system (Metalog) across distributions and demonstrate competitive or superior accuracy.

Quantile Function Transformation Rules

- Gilchrist (2000) presents transformation rules that ensure any transformed quantile function remains valid.

Transformation	Form	Conditions Ensuring Monotonicity
Reflection	$-Q(1-p)$	Always valid
Additive combination	$Q_1(p) + Q_2(p)$	Both valid QFs
Positive linear combination	$a\,Q_1(p) + b\,Q_2(p)$	$a, b > 0$
Product combination	$Q_1(p) \cdot Q_2(p)$	Both nonnegative on $(0, 1)$
Monotone transformation	$T(Q(p))$	T continuous and non-decreasing

- **Corollary (Odd-Power Rule):** If $Q(p)$ is a quantile function and k is any odd positive integer, then $Q(p)^k$ is also a quantile function (Bickel, 2025).
- Gilchrist defines a linear quantile distribution as the conical combination:

$$Q(p) = a_1 + \sum_{k=2}^{K} a_k\, Q_k(p), \quad a_k \geq 0 \text{ for } k \geq 2$$

where each Q_k is referred to as a basis quantile function

Fitting Assessed Points: The Design Matrix

- Given K assessed pairs (p_i, x_i), we require $Q(p_i) = x_i$ for all i, giving the linear system $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{a}$:

$$\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{pmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{pmatrix} 1 & Q_2(p_1) & \cdots & Q_K(p_1) \\ 1 & Q_2(p_2) & \cdots & Q_K(p_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Q_2(p_K) & \cdots & Q_K(p_K) \end{pmatrix}}_{\boldsymbol{B}} \underbrace{\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{pmatrix}}_{\boldsymbol{a}}$$

- If $\boldsymbol{B}$ is full rank: unique solution $\boldsymbol{a} = \boldsymbol{B}^{-1}\boldsymbol{x}$, exact interpolation.

State of the Art: Two Leading QPD Systems

- **J-QPD** (Hadlock and Bickel, 2019):
  - Based on the Johnson distribution system
  - Guarantees a valid quantile function for any input
  - Limited to 3 symmetric quantile–probability pairs that include the median
- **Metalog** (Keelin, 2016):
  - Can match any n quantile–probability pairs via a linear system $\boldsymbol{x} = \boldsymbol{Ba}$
  - Highly flexible and widely adopted in practice
  - No guarantee that the result is a valid quantile function
- There is a fundamental tradeoff between these systems: J-QPD guarantees validity but is restrictive; Metalog is flexible but may produce invalid distributions

Metalog: How It Works and Why It Breaks

- The Metalog begins with a base quantile function $S(p)$ in location-scale form

$$Q(p) = \mu + \sigma\, S(p)$$

- To generate flexibility, μ and σ are expanded as polynomials in $(p - 0.5)$
- A K-term Metalog quantile function is

$$Q_{\mathrm{ML}}^{[K]}(p) = \sum_{k=1}^{K} a_k\, (p - 0.5)^{\lfloor (k-1)/2 \rfloor} \left(\mathbb{1}_{k \in \mu} + S(p)\, \mathbb{1}_{k \in \sigma} \right)$$

- Key limitation: the basis functions are not quantile functions — products like $(p - 0.5) \cdot S(p)$ and $(p - 0.5)^2$ violate Gilchrist's transformation rules
- Monotonicity must be checked and repaired ex post through numerical optimization

Can We Have Both?

Can we design a QPD system that is as flexible as Metalog but with interpretable monotonicity conditions like J-QPD?

- Match any n quantile–probability pairs (like Metalog)
- Provide simple, verifiable conditions for validity (like J-QPD)
- Control monotonicity and modality ex ante through coefficient constraints

Our Answer

The QFlex Distribution System

Introducing QFlex: Three Basis Families

- QFlex is constructed entirely from valid quantile functions
- Three basis families, each controlling a different part of the distribution:

$$R_j(p) := \left[-\ln(1-p)\right]^j \quad \text{(right-tail exponential)}$$

$$L_j(p) := -\left[-\ln(p)\right]^j \quad \text{(left-tail reflected exponential)}$$

$$C_j(p) := (p-\gamma)^{2j-1} \quad \text{(centered uniform, } \gamma \in (0,1))$$

- These are interleaved to form the basis B_k:

$$B_2 = R_1,\ B_3 = L_1,\ B_4 = C_1,\ B_5 = R_2,\ B_6 = L_2,\ B_7 = C_2,\ \ldots$$

- A K-order QFlex (unbounded) quantile function is:

$$Q^{[K]}_{\mathrm{Flex}}(p) := a_1 + \sum_{k=2}^{K} a_k\, B_k(p)$$

Building Intuition: Low-Order QFlex

- $K = 2$: $Q^{[2]}_{\text{Flex}}(p) = a_1 + a_2\big[-\ln(1-p)\big]$
  $\rightarrow$ Shifted and scaled exponential distribution
- $K = 3$: $Q^{[3]}_{\text{Flex}}(p) = a_1 + a_2\big[-\ln(1-p)\big] + a_3\ln(p)$
  $\rightarrow$ Skew-logistic distribution (logistic core + skewness via $\ln p$)
- $K = 4$: $Q^{[4]}_{\text{Flex}}(p) = a_1 + a_2\big[-\ln(1-p)\big] + a_3\ln(p) + a_4(p-\gamma)$
  $\rightarrow$ Flattened skew-logistic (Gilchrist, 2000); center term controls PDF height
- Each additional term adds flexibility while preserving quantile function structure

Fitting QFlex and The Design Matrix $\boldsymbol{B}$

- To fit QFlex, we solve the linear system $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{a}$
- If $\boldsymbol{B}$ is not full rank, the system has no unique solution
  - we may not be able to fit the assessed quantiles
- What guarantees $\boldsymbol{B}$ is full rank? $\rightarrow$ The analytic structure of the basis functions
- A set of functions forms an *Extended Complete Chebyshev* (ECT) system if their Wronskian is nonzero everywhere on the domain
- **Why this matters:** if the basis is ECT, then $\boldsymbol{B}$ is full rank for any choice of distinct assessment probabilities, so a unique solution is always guaranteed regardless of input data

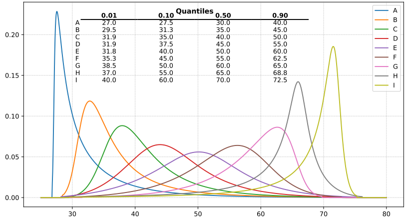
QFlex Design Matrix: Results

- For $K \leq 6$: QFlex is an ECT system
  - $\boldsymbol{B}$ is full rank for any distinct probabilities and any γ
  - The system $\boldsymbol{x} = \boldsymbol{Ba}$ always has a unique solution

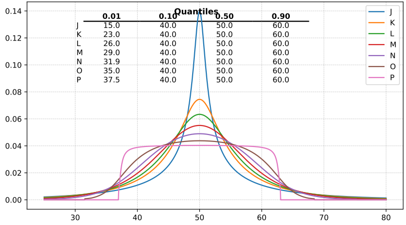
- For $K \geq 7$: QFlex is no longer ECT
  - $\boldsymbol{B}$ is full rank for almost all* assessment grids
  - Rank deficiency occurs only when the centering parameter γ equals the symmetry center ζ of the assessed probabilities — easily avoided

- By contrast, Metalog is not ECT at any order and cannot guarantee full rank even for small K

QFlex Flexibility: Skewed and Symmetric Shapes



Skewed distributions



Symmetric distributions

- All generated using 4-term QFlex fits to quantile assessments from Keelin (2011)
- Visually indistinguishable from Metalog fits at the same order

Fitting $\neq$ Validity

- QFlex can match any finite set of coherent assessments, but the resulting $Q_{\mathrm{Flex}}^{[K]}$ is not necessarily monotonic
- QFlex monotonicity depends on the coefficient vector $\boldsymbol{a}$
- For low orders, we have **exact** (necessary and sufficient) conditions on $\boldsymbol{a}$:
  - $K = 2$: valid iff $a_2 > 0$
  - $K = 3$: valid iff $a_2 > 0$ and $a_3 > 0$
  - $K = 4$: valid iff $a_2 > 0$, $a_3 > 0$, and $(\sqrt{a_2} + \sqrt{a_3})^2 + a_4 > 0$
- We are now solving a least squares fit problem when we constrain coefficients
- For $K \geq 5$: conditions become transcendental — no closed form possible

QFlex Monotonicity Conditions

- $Q^{[K]}_{\mathrm{Flex}}$ is valid if and only if its *quantile density function* $q^{[K]}_{\mathrm{Flex}}(p)$ is positive.

$$q^{[K]}_{\mathrm{Flex}}(p) := Q^{[K]}_{\mathrm{Flex}}(p)' = \sum_{k=2}^{K} a_k \, B'_k(p) > 0 \quad \forall\, p \in (0,1)$$

- Every basis derivative $B'_k(p) \geq 0$ on $(0,1)$ because QFlex is built from valid quantile functions

Theorem 1 (Monotonicity Sufficiency Guarantee)

If $a_k \geq 0$ for all $k \geq 2$ and at least one $a_k > 0$, then $Q^{[K]}_{\mathrm{Flex}}$ is strictly increasing and defines a valid quantile function.

- Intuition: a nonnegative sum of nonnegative functions is nonnegative

Enforcing Monotonicity for Higher Orders

- When exact conditions are unavailable, we have two ex-ante constraint sets that can be imposed during fitting:

1. **All coefficients nonnegative** (Thm. 1)
   - $a_k \geq 0$ for all $k \geq 2 \Rightarrow$ valid
   - Most restrictive, but simplest to enforce
2. **All tails nonneg., tail–center coefficient bound** (Prop. 5)
   - All tail coefficients ≥ 0, center bounded: $a_2 + a_3 > \overline{M}_{\text{center}}$
   - Less restrictive; allows negative center coefficients

- Both are linear constraints on $\boldsymbol{a}$, no specialized software needed

QFlex Modality: Where Do Modes Come From?

- Modes of the distribution correspond to roots of $q'(p)$
- Each basis curvature $B_k''(p)$ is either zero or strictly increasing

Theorem 5 (Only Negative Coefficients Create Modes)

If $a_k \geq 0$: the component $a_k B_k''(p)$ is increasing — cannot create a mode. If $a_k < 0$: may introduce at most one additional mode.

QFlex Modality: Design-Controlled

- Theorem 5 gives us direct control over how many modes QFlex can exhibit:

Coefficient Constraints	Valid QF?	Max Modes
All $a_k \geq 0$ ($k \geq 2$) (Thm. 1)	Yes	1
Prop. 5 enforced (+tails, +T–C bound; centers free)	Yes	$\max\{1, \lfloor (K-1)/6 \rfloor\}$
Prop. 4 holds (+leading tails, +T–C magnitude)	Yes	$\max\{1, \lfloor (K-3)/2 \rfloor\}$

- An analyst can target any desired modality by selecting the appropriate coefficient constraints
- By contrast, Metalog modality bounds require a nontrivial global analysis and do not condition on validity

Bounded and Semi-Bounded QFlex

- Many applications require distributions on a half-line $[\ell, \infty)$ or finite interval $[\ell, u]$.
- **Semi-bounded:** apply the exponential transformation

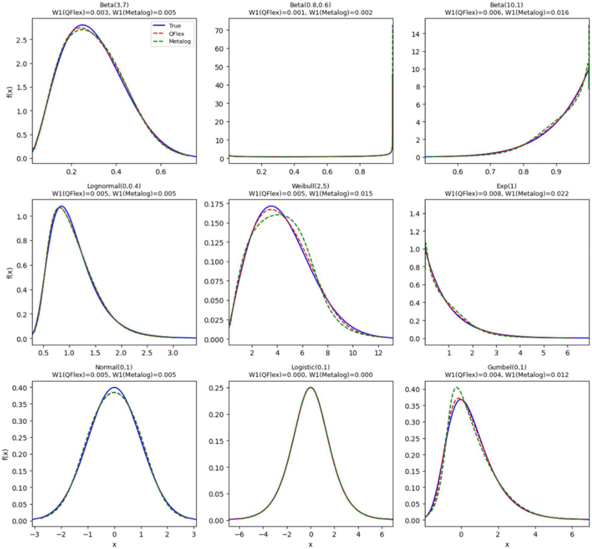
$$Q^{[K]}_{\text{Flex-S}}(p) = \ell + \exp\big(Q^{[K]}_{\text{Flex}}(p)\big)$$

- **Bounded:** apply the logistic CDF $\Lambda(x) = \frac{1}{1+e^{-x}}$

$$Q^{[K]}_{\text{Flex-B}}(p) = \ell + (u - \ell)\,\Lambda\big(Q^{[K]}_{\text{Flex}}(p)\big)$$

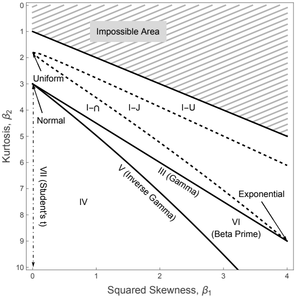
- Both transformations are monotonic, so monotonicity is inherited directly from the unbounded QFlex

QFlex Fits to Some Well Known Distributions ($K = 4$)



- True (blue), QFlex (red), Metalog (green)
- QFlex provides the closer match in most cases
- No constraints were imposed here

Comparing Performance of QFlex and Metalog



- Tested QFlex and Metalog across $\sim$3,500 Pearson-system distributions
- Spans bounded (Beta), semi-bounded (Type VI), and unbounded (Type IV) families
- Truncation orders $K = 3, \ldots, 10$
- For each fit, we evaluate:
  - Accuracy: normalized Wasserstein (W_1) error
  - Monotonicity: does the fit define a valid QF?
  - Modality: are spurious modes introduced?

Comparison: Results

- Accuracy: QFlex delivers lower W_1 error than Metalog across almost every Pearson region and truncation order
- Monotonicity: Both achieve valid fits in most cases, but QFlex provides analytic certificates while Metalog relies entirely on ex-post numerical checks
- Modality: QFlex produces no spurious modes across all tests. Metalog produces spurious modes at $K = 10$ in approximately 1% of Type IV distributions
- Takeaway: QFlex matches or exceeds Metalog accuracy while providing interpretable guarantees on validity and shape that Metalog cannot offer

QFlex: Summary of Contributions

- New QPD system QFlex: entirely from valid quantile functions, adhering to Gilchrist's transformation rules. It provides:
  1. Full-rank design matrix for $K \leq 6$; generic full rank for all K
  2. Exact interpolation of any finite set of coherent quantile assessments
  3. Two independent, easily verifiable monotonicity certificates
  4. Controllable monotonicity ex ante through simple linear constraints
  5. Controllable modality
  6. Competitive or superior accuracy to Metalog

Table of Contents

Summary of Contributions

- **Chapter 2:** Two-stage stochastic program for CCUS network design. Current 45Q levels justify some near-term investment; policy uncertainty reduces capture by $\leq 16\%$; uncertainty matters more than risk aversion
- **Chapter 3:** First Stackelberg game for optimal CC subsidies. Optimal subsidy $\approx$ \$83/ton, close to 45Q. CC subsidies can perversely increase emissions, but unlikely in practice. Cost uncertainty may raise the optimal subsidy
- **Chapter 4:** QFlex distribution system with interpretable monotonicity conditions, full-rank design matrix, modality control, and competitive accuracy vs. Metalog across $\sim$3,500 distributions

Questions?

Thank you for your time and attention!

Summary of Contributions (Not in this presentation)

- **Chapter 5:** Proved the conjectured monotonicity conditions for the truncated Johnson SL and SB power-series expansions (Bickel, 2026), establishing that odd-order truncations define valid quantile functions for all parameter values
- **Additional work (not in dissertation):** Co-authored a study on optimal resource placement for electric grid resilience (Sambasivam et al., 2024), analyzing when centralized vs. distributed strategies are preferable across different network topologies

Future Research: CCUS Policy (Chapters 2–3)

- **Larger networks:** Use decomposition methods (e.g., Benders), better formulations, and/or approximation heuristics to scale to nationwide optimization
- **Dynamic costs:** Model technological maturation and declining capture costs over time
- **Additional uncertainties:** Model uncertainty in capture costs, storage capacities, and CO_2 transport via stochastic programming
- **Multiple heterogeneous firms:** Shared infrastructure, learning spillovers, and menus of subsidies tailored to firm populations
- **Richer subsidy models:** Endogenous capture fraction ε, more realistic distributions for government uncertainty

Future Research: QFlex (Chapter 4)

- **Alternative basis functions:** Replace exponential/reflected-exponential bases with Pareto or lognormal bases for heavy-tailed variants (e.g., QFlex-Heavy)
- **Multivariate extensions:** Combine QFlex marginals with copulas while preserving monotonicity guarantees
- **Open-source software:** Release a well-documented package with fitting routines, monotonicity checks, and modality controls

Chapter 5: The Johnson Distribution System

- The Johnson system (Johnson, 1949) generates flexible distribution families via monotonic transformations of the standard normal $\Phi^{-1}(p)$

- Three subfamilies, each using a different transformation:
  - **SU** (unbounded): hyperbolic sine
  - **SL** (semi-bounded): exponential
  - **SB** (bounded): logistic

- Four parameters $(\gamma, \delta, \xi, \lambda)$ allow a unique distribution for any admissible combination of mean, standard deviation, skewness, and kurtosis

Johnson Power-Series Expansions (JPSE)

- Bickel (2026) introduced a novel formulation: express the Johnson quantile functions as explicit power-series expansions
- These truncated expansions provide a new analytic foundation for structured quantile-based modeling
- Bickel conjectured sufficient conditions under which truncated SL and SB expansions remain valid quantile functions (i.e., strictly increasing)
- Strong numerical evidence supported the conjectures, but no formal proofs were provided
- Our contribution: rigorous proofs of both conjectures

JPSE-SL: From Johnson to Power Series

- The Johnson SL quantile function:

$$Q_{\mathrm{SL}}(p) = \xi + \lambda \exp\left(\frac{\Phi^{-1}(p) - \gamma}{\delta}\right)$$

- Expanding $\exp(\cdot)$ as a truncated Taylor series:

$$\boxed{Q_{\mathrm{SL}}^{[K]}(p) = \xi + \lambda\left(1 + \sum_{k=1}^{K} \frac{z(p)^k}{\delta^k\, k!}\right)}, \quad z(p) = \Phi^{-1}(p) - \gamma$$

- Valid quantile function iff the derivative polynomial is strictly positive:

$$P_{\mathrm{SL}}^{[K]}(z) = \frac{1}{\delta} \sum_{k=0}^{K-1} \frac{(z/\delta)^k}{k!} > 0 \quad \forall z \in \mathbb{R}$$

JPSE-SL: Proof of Monotonicity

Theorem (JPSE-SL Monotonicity)

For all odd truncation orders K and $\delta > 0$, $P_{\mathrm{SL}}^{[K]}(z) > 0$ for all $z \in \mathbb{R}$.

- By the Lagrange remainder theorem:

$$\sum_{k=0}^{K-1} \frac{z^k}{k!} = \underbrace{e^z}_{>0} + \underbrace{\frac{e^c \cdot (-z)^K}{K!}}_{\text{sign depends on } K}$$

  for some c between 0 and z

- For $z > 0$: the truncated series is trivially positive
- For $z < 0$ and K odd: $(-z)^K > 0$, so both terms are positive
- Therefore $P_{\mathrm{SL}}^{[K]}(z) > 0$ for all $z \in \mathbb{R}$

JPSE-SB: From Johnson to Power Series

- The Johnson SB quantile function:

$$Q_{\mathrm{SB}}(p) = \xi + \lambda \,\mathrm{logistic}\left(\frac{\Phi^{-1}(p) - \gamma}{\delta}\right)$$

- Expanding $\mathrm{logistic}(\cdot)$ as a truncated power series introduces Bernoulli numbers:

$$Q_{\mathrm{SB}}^{[K]}(p) = \xi + \lambda \left(\frac{1}{2} + \sum_{k=1}^{K} \frac{B_{2k}(4^k - 1)}{\delta^{2k-1}(2k)!}\, z(p)^{2k-1}\right)$$

- Valid quantile function iff the derivative polynomial is strictly positive — considerably harder than SL since B_{2k} alternates in sign

JPSE-SB: Proof Strategy

Theorem (JPSE-SB Monotonicity)

For all odd K and $\delta > 0$, $P_{\text{SB}}^{[K]}(z) > 0$ for all $z \in \mathbb{R}$.

- **Step 1:** Substitute a closed-form identity for B_{2k} using the Riemann zeta function to obtain a double summation
- **Step 2:** Interchange the order of summation so the inner sum has a closed-form expression
- **Step 3:** After simplification, $P_K(z)$ takes the form

$$P_K(z) = \frac{-1}{z^2} \sum_{m=0}^{\infty} \sum_{k=1}^{K} (2k-1)(-u_m)^k$$

where $u_m = z^2/((2m+1)^2\pi^2) > 0$

JPSE-SB: Completing the Proof

- Need to show each inner sum $\sum_{k=1}^{K}(2k-1)(-u)^k$ is negative

- This evaluates in closed form; for odd K it reduces to showing

$$u-(2K+1)u^K-(2K-1)u^{K+1}<1 \quad \forall\, u>0$$

- Case analysis:
  - $0<u<1$: LHS $<u<1$ since the last two terms are negative
  - $u=1$: LHS $=-4K<1$
  - $u>1$: the u^K and u^{K+1} terms dominate, making LHS negative

- Each inner sum is negative, and $-1/z^2<0$, so $P_K(z)>0$ ■

The University of Texas at Austin
Cockrell School of Engineering