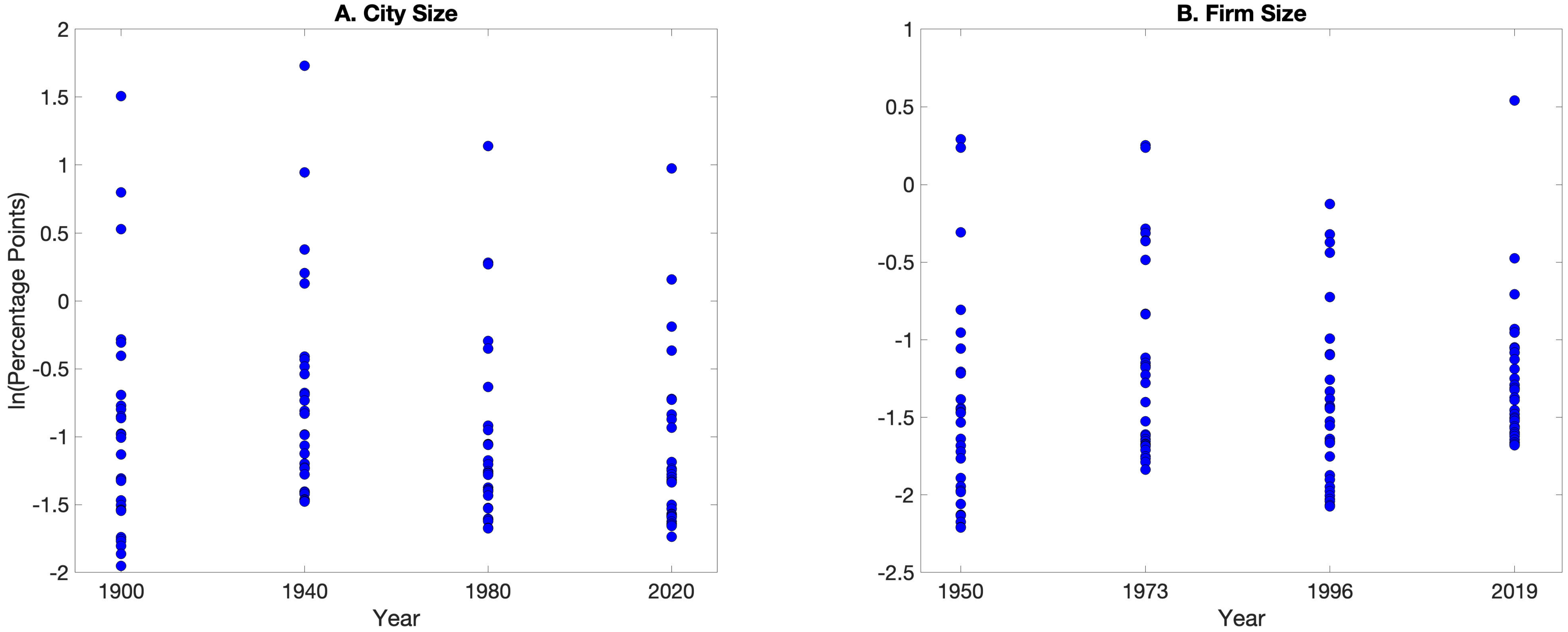
Time Varying Extremes

Ulrich K. Müller and Mark W. Watson
Princeton University

EVA 2025

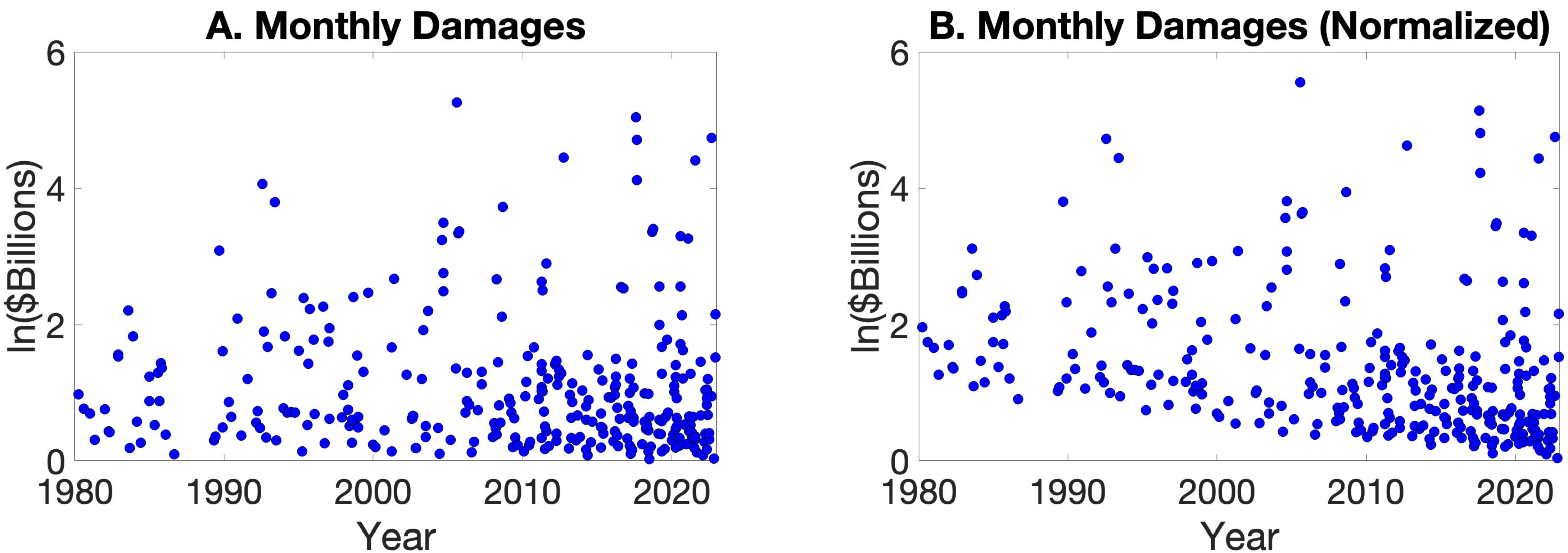
30 Largest Cities and Firms in U.S. Over Time



Size relative to total population and total private sector employment, respectively.

⇒ Is this data compatible with constant underlying tail properties?

Monthly Weather Related Damage Events Exceeding $1B



Normalization by total U.S. capital stock.

⇒ Is this data compatible with constant underlying tail properties?

This Paper

- Frequentist tests of tail properties in panel of k extreme observations for each of $t = 1, \ldots, T$ periods
  - For each t, consider $n_t \to \infty$ asymptotics for k-fixed to argue for convergence to joint GEV distribution

    Parameters of GEV distribution a function of sample size n_t and tail parameters of population distribution
  - Yields panel of T independent approximately GEV-valued random vectors
  - Derive hypothesis tests about tail parameters that are asymptotically valid for *fixed* (k, T)
  - Extension to observing k_τ exceedances over (extreme) threshold τ, for $k_\tau = O_p(1)$
- Bayes inference about time varying parameters
  - Use Hamiltonian Monte Carlo to overcome non-Gaussian smoothing problem

Related Literature

- Very large and mature literature on inference based on $k = k_n \to \infty$ extreme observations

- Müller and Wang (2017, JASA)

  Inference about tail properties based on k extreme observations with k fixed for single sample

- Bücher and Segers (2017, Extremes)

  Asymptotic properties of maximum likelihood of GEV as $T \to \infty$

Fixed-k Asymptotics

- $X_1 \geq X_2 \geq \cdots \geq X_k$ order statistics of i.i.d. sample W_i, $i = 1, \ldots, n$ of population with (approximate) Generalized Pareto tail with parameters (ν, ω, ξ)

- Induces
$$\mathbf{X} = \{X_j\}_{j=1}^{k} \overset{a}{\sim} \mathsf{GEV}_k\left(\omega\frac{n^\xi - 1}{\xi} + \nu, \omega n^\xi, \xi\right) \sim \mathsf{GEV}_k(\mu, \sigma, \xi)$$
(with exact GP W_i, only approximation is $n^{-1}\sum_{i=1}^{n+1} E_i \approx 1$ for E_i i.i.d. exponential RVs)

- For $\xi > 0$, standard further approximation sets
$$\mu = \omega\frac{n^\xi - 1}{\xi} + \nu \approx \omega\frac{n^\xi}{\xi}.$$
But not uniformly valid: For every n, can set ν large enough so that it is non-negligible in μ...

- For repeated cross sections obtain $\mathbf{X}_t \overset{a}{\sim} iid\mathsf{GEV}_k(\mu_t, \sigma_t, \xi_t)$, $t = 1, \ldots, T$

Inference for Parameters in Stable Model

- Model is $\mathbf{X}_t \overset{a}{\sim} iid\text{GEV}_k(\mu, \sigma, \xi)$, $t = 1, \ldots, T$

- Test $H_0 : (\mu, \sigma, \xi) \in \Theta_0$ using generalized LR statistic (as in MW (2017))

$$\text{LR} = \sup_{\mu,\sigma,\xi} \ln \prod_{t=1}^{T} f_{\text{GEV}}(\mathbf{X}_t|\mu, \sigma, \xi) - \sup_{(\mu,\sigma,\xi)\in\Theta_0} \ln \prod_{t=1}^{T} f_{\text{GEV}}(\mathbf{X}_t|\mu, \sigma, \xi).$$

- In general, distribution of LR depends on nuisance parameters.
  - But for many hypotheses of interest, distribution is known or only depends on ξ: $H_0 : \xi = \xi_0$, $H_0 : \mu = \sigma/\xi$, $\xi = 1$ (Zipf's Law), tests that specify quantiles of $\text{GEV}_1(\mu, \sigma, \xi)$, etc.
  - Rather than simple **sup** bound on critical value (as in as in MW (2017)), set $\text{cv}(\hat{\xi}) = \exp(a_0 + a_1\hat{\xi} + a_2\hat{\xi}^2)$ with $\hat{\xi}$ the MLE and pick a_i to numerically solve

$$\min_{\xi} \mathbb{P}_{\xi}(\text{LR} > \text{cv}(\hat{\xi})) \quad \text{subject to} \quad \max_{\xi} \mathbb{P}_{\xi}(\text{LR} > \text{cv}(\hat{\xi})) \leq \alpha$$

Choice of k

- Without further assumptions, cannot use data to determine appropriate k: For any procedure that has $\hat{k}_n \xrightarrow{p} \infty$ if W_i is exactly Pareto, there exists sequence of underlying populations for which that choice of $\hat{k}_n$ yields arbitrarily misleading inference, yet GEV theory holds for any fixed k (MW 2017)

- Substantive assumption that enables informative inference: Tail of mass k/n of underlying population is well approximated by GP

- Consider different k and see how conclusions change

Exceedances Data

- Let $\mathbf{Y} = \{W_i : W_i \geq \tau\}$ be the exceedances over τ, of random dimension $k_\tau \geq 0$

- For $\tau = \tau_n$ such that $k_\tau = O_p(1)$, derive asymptotic distribution of $\mathbf{Y}$ assuming GP(ν, ω, ξ) tail of W_i

  $\Rightarrow$ Distribution depends on τ_n and GEV parameters $(\mu, \sigma, \xi) = \left(\omega \frac{n^\xi - 1}{\xi} + \nu, \omega n^\xi, \xi\right)$

- Straightforwardly generalizes to panel of independent $\mathbf{Y}_t$, $t = 1, \ldots, T$

- Apply LR statistic to data $\mathbf{Y}_t$, $t = 1, \ldots, T$, similar to above

Empirical Results Assuming Parameter Stability

	MLE				95% CI		$LR/cv(\hat{\xi})$ for
	ξ	σ	μ	$q_{0.9}$	ξ	$q_{0.9}$	Zipf's Law
Weather damages	0.84	0.86	0.45	6.22	(0.66,1.06)	(5.11,7.82)	1.25
Weather normalized	1.00	0.87	0.90	8.25	(0.80,1.22)	(6.68,10.61)	0.02
City Size	0.65	1.34	2.02	8.89	(0.43,0.93)	(3.92,31.43)	1.57
Firm Size	0.42	0.48	1.00	2.81	(0.23,0.69)	(1.56,7.92)	3.49

5% critical value of $LR/cv(\hat{\xi})$ is equal to 1.0 by construction.

Testing Parameter Stability

- Want to test $H_0 : (\mu_t, \sigma_t, \xi_t) = (\mu, \sigma, \xi)$, $t = 1, \ldots, T$, treating T as fixed

- Standard test statistic in time series from Nyblom (1989, JASA)

$$L_T = T^{-2} \sum_{t=1}^{T} \left(\sum_{l=1}^{t} S_l(\hat{\theta}) \right)' \hat{V}^{-1} \left(\sum_{l=1}^{t} S_l(\hat{\theta}) \right)$$

  with $S_t(\theta) = \partial \ln f(\mathbf{X}_t)/\partial\theta$, $\hat{\theta}$ the MLE, and $\hat{V} = -T^{-1} \sum_{t=1}^{T} \partial S_t(\theta)/\partial\theta'|_{\theta=\hat{\theta}}$

- As for LR statistic, distribution of L_T (only) depends on ξ under null hypothesis for fixed T

  $\Rightarrow$ Same approach for critical value construction as for LR, namely use $\text{cv}(\hat{\xi}) = \exp(a_0 + a_1 \hat{\xi} + a_2 \hat{\xi}^2)$

Empirical Results Testing Parameter Stability

	$L_T/\mathsf{cv}(\hat{\xi})$
Weather damages	11.6
Weather normalized	4.2
City Size	1.1
Firm Size	0.7

5% critical value of $L_T/\mathsf{cv}(\hat{\xi})$ is equal to 1.0 by construction.

Bayesian Estimation of Parameter Path

- Recall GEV (μ, σ, ξ) parameters are linked to GP (ν, ω, ξ) parameters via

$$(\mu, \sigma, \xi) = \left(\omega \frac{n^{\xi} - 1}{\xi} + \nu, \omega n^{\xi}, \xi\right)$$

$\Rightarrow$ (μ, σ) may time vary due to variation in n, or due to time variation in (ν, ω, ξ)

- Reparameterization GEV in terms of α and (m, s, ξ) such that time variation in n only affects α

- Independent Gaussian random walk priors for (suitable monotone transformation of) $(\alpha_t, m_t, s_t, \xi_t)$

Computational Aspects

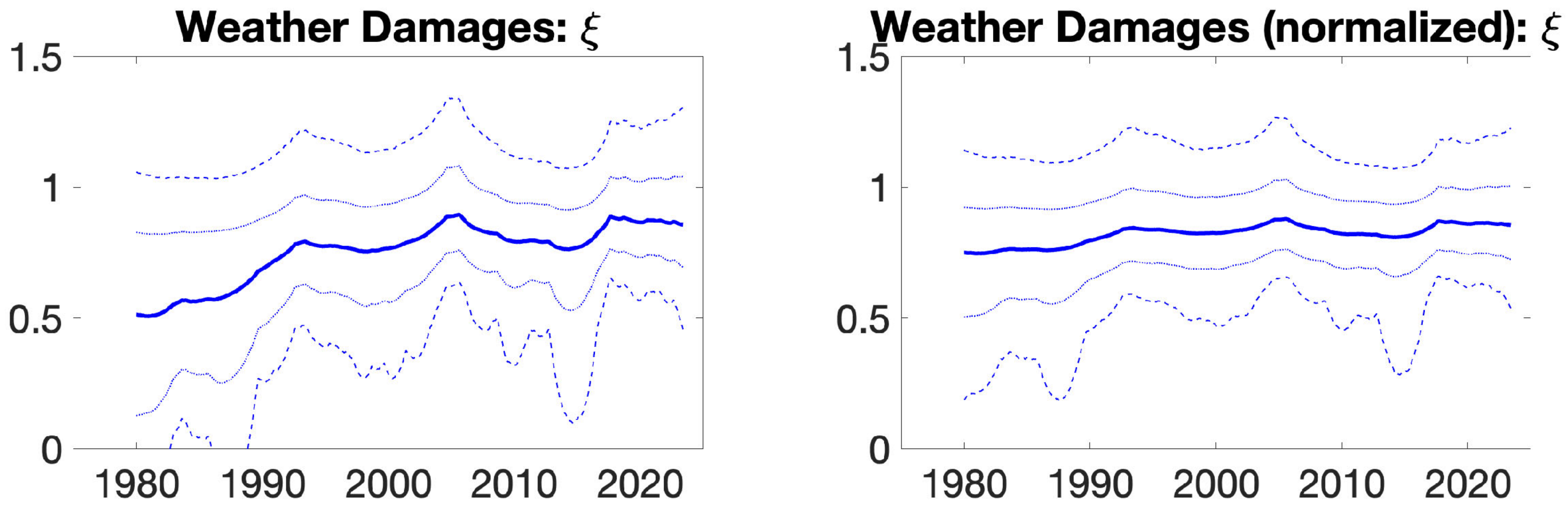
- Filtering/smoothing problem with non-Gaussian measurements $\mathbf{X}_t \sim$GEV$_k(\alpha_t, m_t, s_t, \xi_t)$

  $\Rightarrow$ Use Hamiltonian Monte Carlo treating parameter paths as (high-dimensional) parameters, implemented using Stan

- One complication: GEV distribution has support that depends on parameters

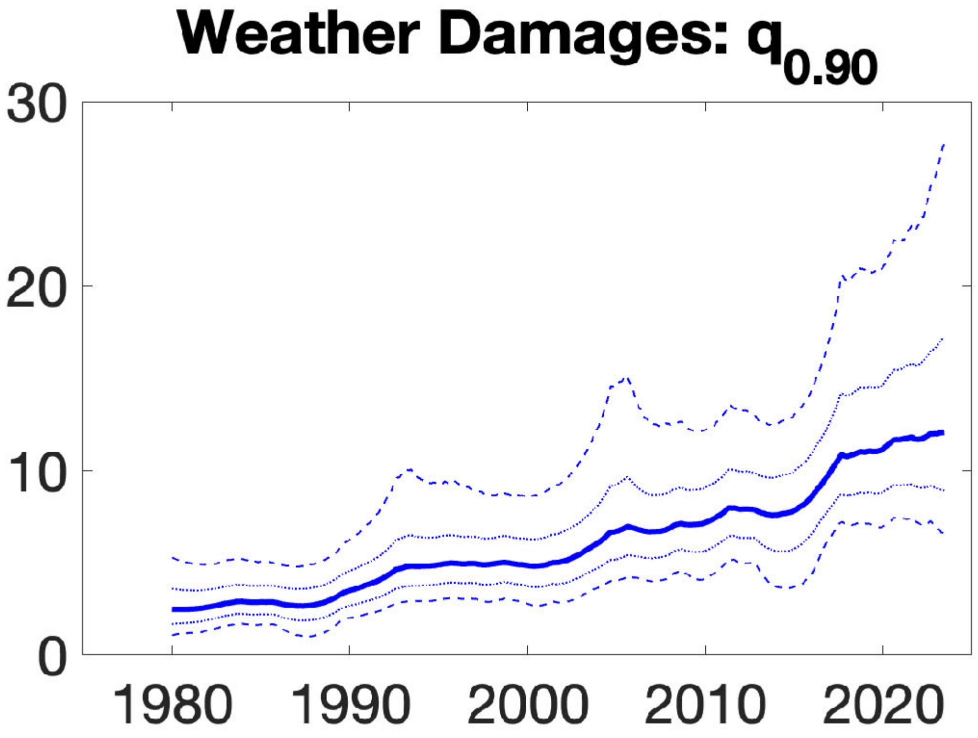
  $\Rightarrow$ Extend likelihood to full support by quadratic extrapolation, then undo by reweighing draws from extended likelihood posterior

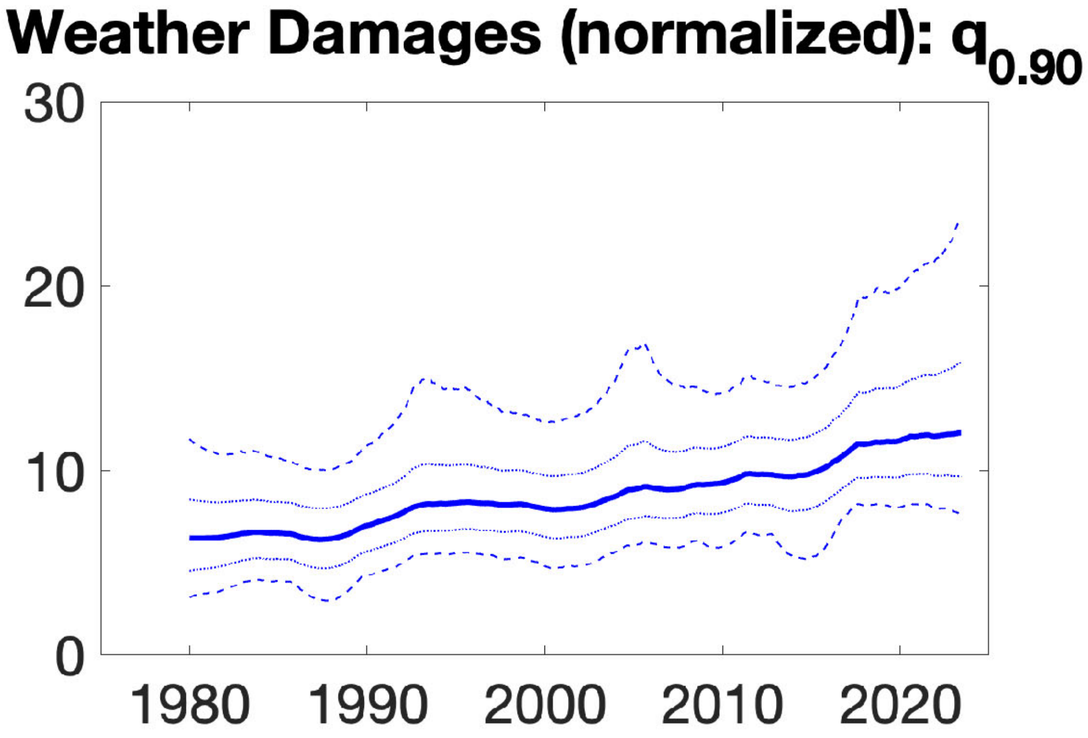
Posterior Parameter Path



Median, $\pm$1stddev and 95% pointwise credible sets

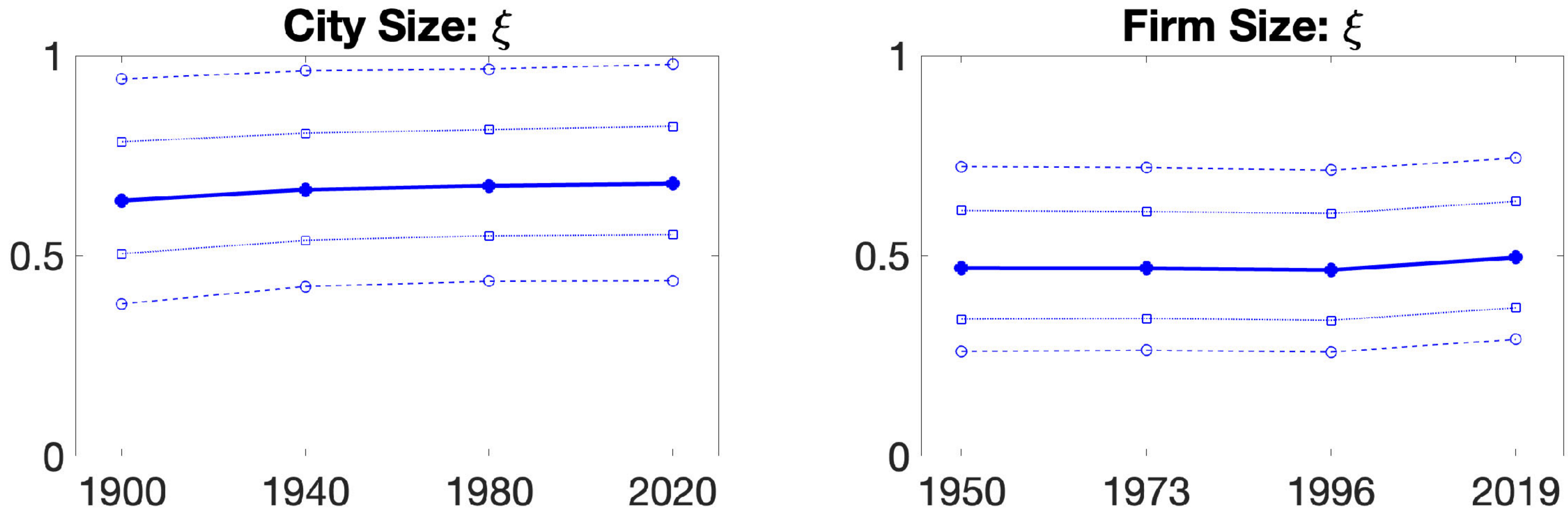
Posterior Parameter Path





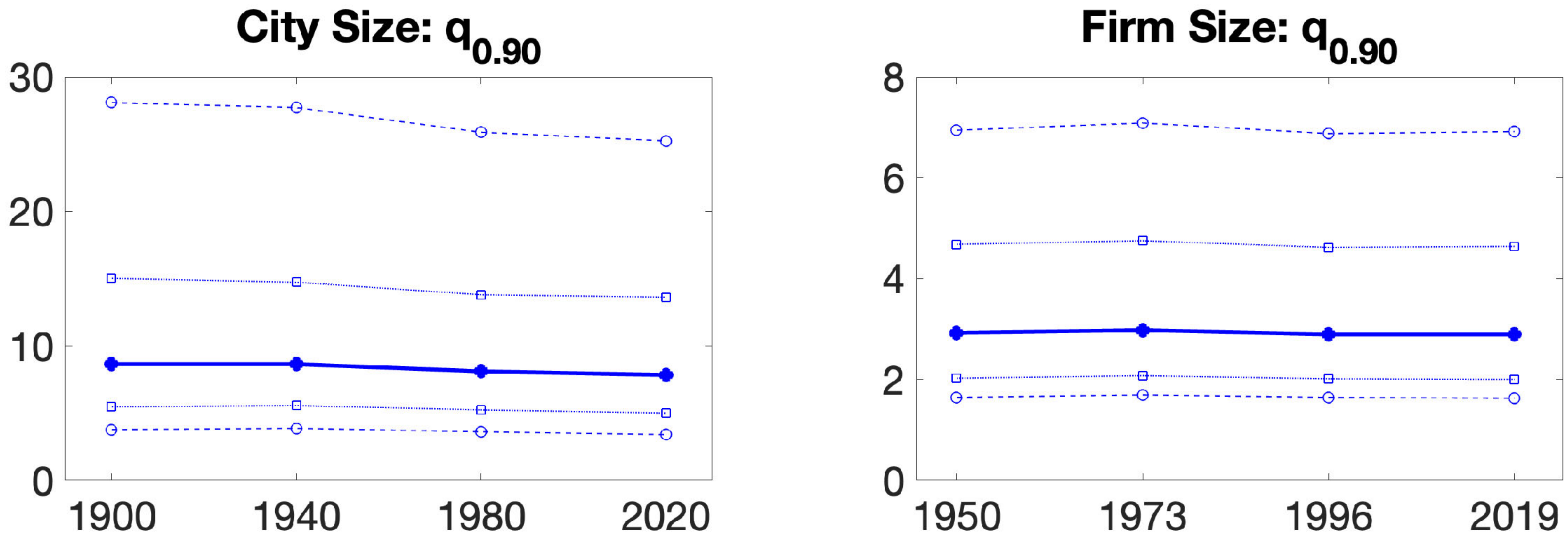
Median, ±1stddev and 95% pointwise credible sets

Posterior Parameter Path



Median, $\pm$1stddev and 95% pointwise credible sets

Posterior Parameter Path



Median, ±1stddev and 95% pointwise credible sets

Log-Bayes Factors

	Time Varying Parameters			
	all	none	α	(ξ, s, m)
Weather damages	0.0	-55.2	-5.7	-0.8
Weather normalized	0.0	-20.9	-7.6	0.1
City Size	0.0	4.1	4.0	0.1
Firm Size	0.0	2.1	2.8	0.2

Conclusions

- Valid inference for tail properties and parameter stability under (k, T) fixed asymptotics

- Bayesian approach to estimation of parameter path

  $\Rightarrow$ (surprisingly?) easy to implement using Stan

  $\Rightarrow$ could be quite useful for forecasting purposes

Thank you!