

Robust Bayesian Inference under Partially Specified Priors

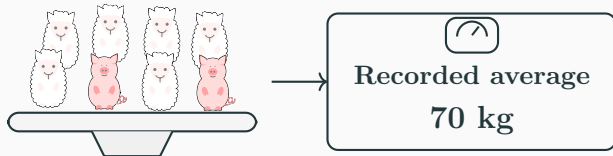
Martin Fankhauser

October 2026

Bocconi University (PhD candidate)

Motivating example

Suppose we are interested in the average weight of **sheep**, $\mathbf{W}$, but some **pigs** accidentally jumped on the scale.



Suppose we know that the **average pig weight** is 100 kg, but the **share of pigs** on the scale, q , is unknown. Then

$$70 = (1 - q) \times W + q \times 100.$$

Since sheep weights are nonnegative, the data only tells us

$$\mathbf{W} \in [0, 70].$$

Thus, $\mathbf{W}$ is **set-identified**.

Can additional information help?

The data does not reveal $\mathbf{q}$, but the farmer might have some additional information.

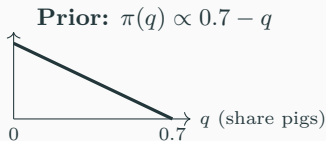


Farmer

“I only saw a few pigs jump on the scale. There may have been more, but a small number seems more likely.”

Interpretation: *Smaller pig shares are weakly more plausible.*

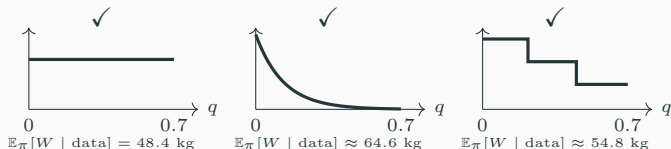
Formally we might want to model this using a **prior distribution** for $\mathbf{q}$, $\pi(q)$, and combine it with the **data** $\mathbf{W} \in [0, 70]$:



$$\mathbb{E}_{\pi}[W \mid \text{data}] \approx 58.5 \text{ kg}$$

But actually we only had incomplete prior information

All we know from the farmer is: *Smaller pig shares are at least as plausible as larger ones.*



All three priors encode what the farmer told us. Why should we arbitrarily choose one of them?

I propose instead to **bound posterior expectations** under **all weakly decreasing priors** which results in

$$\mathbb{E}_{\pi}[W \mid \text{data}] \in [48.4, 70] \text{ kg}$$

Data alone: $W \in [0, 70] \Rightarrow$ **with partial prior information:** $[48.4, 70]$

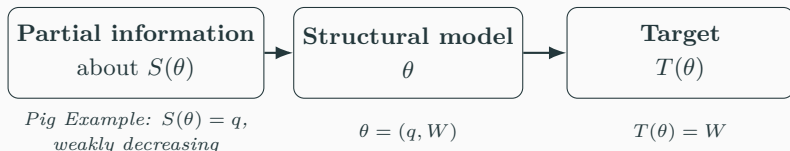
Why?

Contribution

In applications, θ may be high-dimensional and our prior knowledge is often incomplete in **two ways**:

- it concerns only a **low-dimensional feature** $S(\theta)$,
Examples: elasticities, treatment effects, impulse responses, ...
- even for $S(\theta)$, it **might not determine a unique prior**.

Examples: shape restrictions, event probabilities, ...



I develop Bayesian inference under **partially specified priors** in **set-identified models**.

The resulting posterior bounds can be computed using linear programs.

Set-Identification + Partial Prior Knowledge

Combinations of **set-identification (SI)** (*data only told us $W \in [0, 70]$*) and **partial prior knowledge (PP)** (*pig share weakly decreasing*) are common:

- **SVAR with sign or narrative restrictions**

SI: Many impulse responses consistent with Uhlig (2005) signs + data

PP: Output effects of mon. pol. shock small, Variance share high, ...

- **Missing data problems**

SI: Mean of unobserved data unknown, thus overall mean ambiguous.

PP: Most missing individuals are missing-at-random but not all, ...

- **Games with multiple equilibria**

SI: Multiplicity + unknown equilibrium selection only bound payoff parameters.

PP: One equilibrium may be weakly more likely than another, ...

Set-Identification + Partial Prior Knowledge

Combinations of **set-identification (SI)** (*data only told us $W \in [0, 70]$*) and **partial prior knowledge (PP)** (*pig share weakly decreasing*) are common:

- **SVAR with sign or narrative restrictions**

SI: Many impulse responses consistent with Uhlig (2005) signs + data

PP: *Output effects of mon. pol. shock small, Variance share high, ...*

- **Missing data problems**

SI: Mean of unobserved data unknown, thus overall mean ambiguous.

PP: *Most missing individuals are missing-at-random but not all, ...*

- **Games with multiple equilibria**

SI: Multiplicity + unknown equilibrium selection only bound payoff parameters.

PP: *One equilibrium may be weakly more likely than another, ...*

Set-Identification + Partial Prior Knowledge

Combinations of **set-identification (SI)** (*data only told us $W \in [0, 70]$*) and **partial prior knowledge (PP)** (*pig share weakly decreasing*) are common:

- **SVAR with sign or narrative restrictions**

SI: Many impulse responses consistent with Uhlig (2005) signs + data

PP: Output effects of mon. pol. shock small, Variance share high, ...

- **Missing data problems**

SI: Mean of unobserved data unknown, thus overall mean ambiguous.

PP: Most missing individuals are missing-at-random but not all, ...

- **Games with multiple equilibria**

SI: Multiplicity + unknown equilibrium selection only bound payoff parameters.

PP: One equilibrium may be weakly more likely than another, ...

Set-Identification + Partial Prior Knowledge

Combinations of **set-identification (SI)** (*data only told us $W \in [0, 70]$*) and **partial prior knowledge (PP)** (*pig share weakly decreasing*) are common:

- **SVAR with sign or narrative restrictions**

SI: Many impulse responses consistent with Uhlig (2005) signs + data

PP: Output effects of mon. pol. shock small, Variance share high, ...

- **Missing data problems**

SI: Mean of unobserved data unknown, thus overall mean ambiguous.

PP: Most missing individuals are missing-at-random but not all, ...

- **Games with multiple equilibria**

SI: Multiplicity + unknown equilibrium selection only bound payoff parameters.

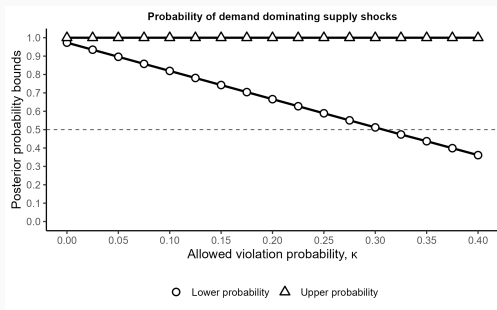
PP: One equilibrium may be weakly more likely than another, ...

Application: Is oil price supply or demand driven?

Kilian and Murphy (2014) restrict the short-run oil supply elasticity:

$0 \leq \alpha_s \leq 0.0258$. debated in Baumeister and Hamilton (2019), Kilian (2022)

Partial distrust: All priors π s.t. $P_\pi(0 \leq \alpha_s \leq 0.0258) \geq 1 - \kappa$



How much distrust can we allow before we can no longer conclude that demand shocks dominate supply shocks?

Dominance is measured using the 12-month FEVD: the share of oil-price forecast variance explained by each shock. Relation: Uncertain Identification

Robustness / Robust Bayes / Marginal Priors:

Giacomini, Kitagawa, and Uhlig (2019), Giacomini and Kitagawa (2021), Giacomini, Kitagawa, and Volpicella (2022), Moon and Schorfheide (2012), Ho (2023), Lavine, Wasserman, and Wolpert (1991), Berger (1984), Hansen and Sargent (2001), Lam, Liu, and Singham (2024), Baumeister and Hamilton (2015)

- **Prior classes built on economic knowledge**
- **Linear program population formulation**
- **Simulation approximation**

Details

SVAR Identification:

Baumeister and Hamilton (2019), Kilian and Murphy (2014), Caldara, Cavallo, and Iacoviello (2019), Kilian (2022), Antolin-Diaz and Rubio-Ramirez (2018), Arias, Caldara, and Rubio-Ramirez (2019), Ludvigson, Ma, and Ng (2017), Rubio-Ramirez, Waggoner, and Zha (2010), Uhlig (2005) etc.

- **New probabilistic identification schemes**
- **Alternative to full ambiguity and Haar prior**
- **Oil demand and supply debate**

Outline

Why is this more than prior choice?

- Some further intuition

Theory

- Constructing prior class
- Population characterization: Linear program
- Simulated base measure and reweighting

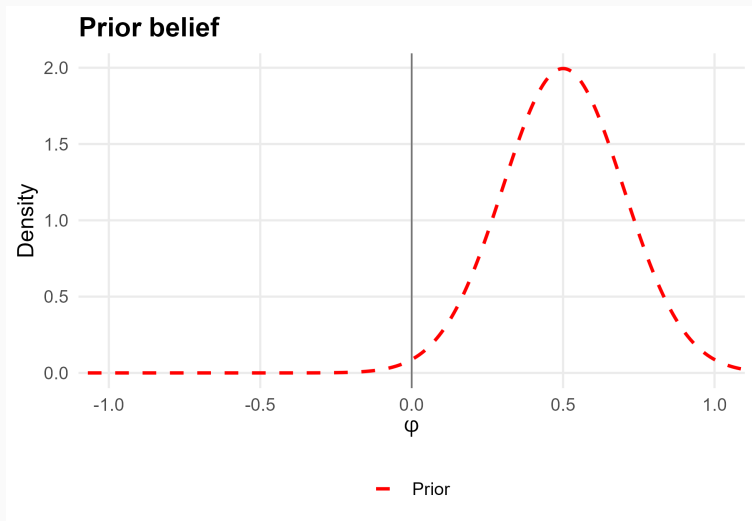
Revisiting Oil Supply and Demand Debate

- A brief summary of the debate
- Demand dominance is robust, supply irrelevance not

Why is this more than just
prior choice?

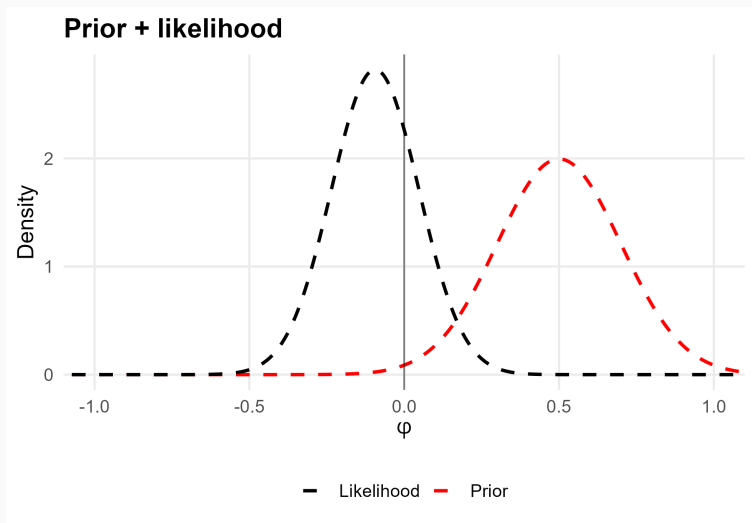
Bayesian updating

Suppose we are interested in $\phi = \mathbb{E}[Y]$. Our prior belief is:



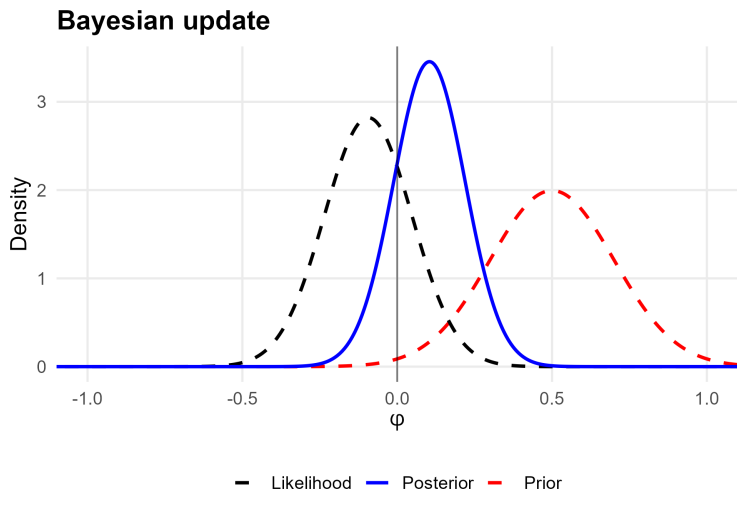
Bayesian updating

Now we obtain an iid sample $Y_1, \dots, Y_n$:



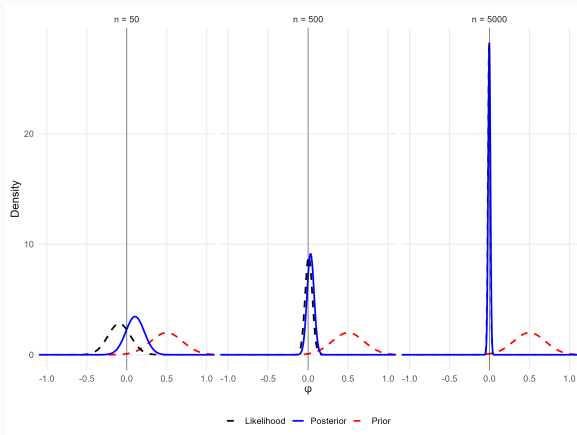
Bayesian updating

The posterior will be a combination of prior and likelihood.



As the sample grows, the likelihood dominates

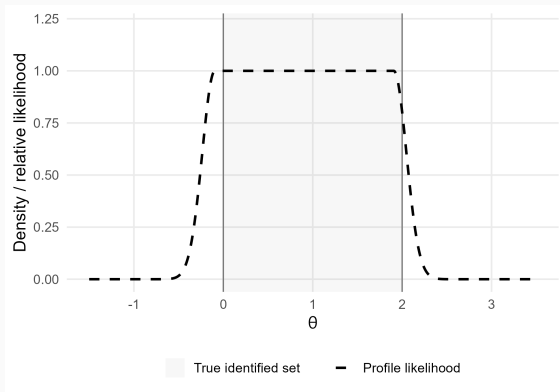
In a **point-identified** model, the likelihood eventually dominates:



Bayesians with different priors agree asymptotically.

What happens under set identification?

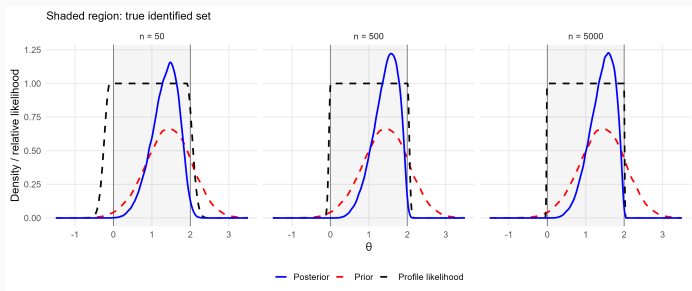
Suppose now our parameter of interest is not ϕ but θ and the only information economic theory is giving us is $\phi \leq \theta \leq \phi + 2$



In the pig example, data + assumptions told us $W \in [0, 70]$, which is the identified set.

Prior influence never vanishes

The data can not distinguish parameter values $\phi \leq \theta \leq \phi + 2$.



A flat likelihood is not telling us the parameter values are equally likely, rather we cannot distinguish them.

Knowing all values are equally likely is different from knowing nothing about them.

Any prior needs to be justified well, thus I propose to only impose what we really know.

There is no uninformative prior

Full ambiguity is extreme

Theory

Setting

Let $\phi \in \Phi$ denote reduced-form parameters and $\theta \in \Theta$ structural parameters, linked by a many-to-one map

$$\phi = g(\theta).$$

All assumptions imposed with certainty, $\mathcal{A}_0$, define the identified set

$$\text{IS}_\theta(\phi) = \{\theta : g(\theta) = \phi, \mathcal{A}_0 \text{ hold}\} \quad (\text{Pig example: } W \in [0, 70]).$$

Write the prior as

$$\pi(\phi, \theta) = \pi_\phi(\phi) \pi_{\theta|\phi}(\theta).$$

We keep π_ϕ fixed and study sensitivity to the **conditional structural prior**: $\pi_{\theta|\phi}$.

How do posterior conclusions change when we change the weights assigned to obs. equivalent structural models?

Defining admissible priors

Fix ϕ and write $s = S(\theta)$ and decompose $\pi_{\theta|\phi}(\theta) = \rho_{\phi}(\theta | s) \mu_{\phi}(s)$.

Marginal
 μ_{ϕ} for $S(\theta)$
Pig example: beliefs about q

+

Completion
 $\rho_{\phi}(\cdot | s)$
Pig example: beliefs about W given q

→ μ_{ϕ} : beliefs about the economically meaningful feature $S(\theta)$.

elasticities, impulse responses, assumptions, narratives, ...

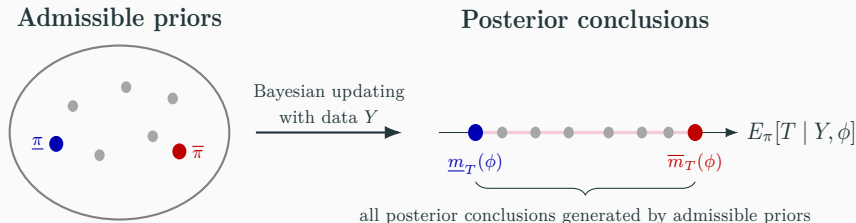
→ ρ_{ϕ} : how probability is distributed among models with the same $S(\theta) = s$. translates from marginal to joint. **Example**

The set of **admissible marginals**, $\mathcal{M}_S(\phi)$, and **completions**, $\mathcal{R}(\phi)$, define the induced class of structural priors:

$$\Pi(\phi) \equiv \Pi_{\mathcal{M}, \mathcal{R}}(\phi) \equiv \{\pi_{\mu, \rho} : \mu \in \mathcal{M}_S(\phi), \rho \in \mathcal{R}(\phi)\}.$$

From a class of priors to posterior bounds

Each admissible prior implies a potentially different posterior conclusion for some target $T : \Theta \rightarrow \mathbb{R}$.



To characterize the entire range, we need to find the two worst-case priors.

$$\underline{m}_T(\phi) = \inf_{\pi \in \Pi(\phi)} \mathbb{E}_\pi[T | Y, \phi] \quad \text{and} \quad \overline{m}_T(\phi) = \sup_{\pi \in \Pi(\phi)} \mathbb{E}_\pi[T | Y, \phi]$$

Pig example: $\underline{m}_T(\phi) = 48.4$ and $\overline{m}_T(\phi) = 70$, where T is the sheep weight.

Set identification makes prior sensitivity linear

In general, changing the prior changes both the **posterior weighting** and its **normalization**:

$$\mathbb{E}_\pi[T \mid Y, \phi] = \frac{\int T(\theta) p(Y \mid \theta, \phi) \pi_{\theta|\phi}(d\theta)}{\int p(Y \mid \theta, \phi) \pi_{\theta|\phi}(d\theta)}.$$

Set identification implies $p(Y \mid \theta, \phi) = p(Y \mid \phi)$ for all $\theta \in \text{IS}_\theta(\phi)$, hence it cancels:

$$\mathbb{E}_\pi[T \mid Y, \phi] = \frac{p(Y \mid \phi) \int T(\theta) \pi_{\theta|\phi}(d\theta)}{p(Y \mid \phi) \int \pi_{\theta|\phi}(d\theta)} = \int T(\theta) \pi_{\theta|\phi}(d\theta).$$

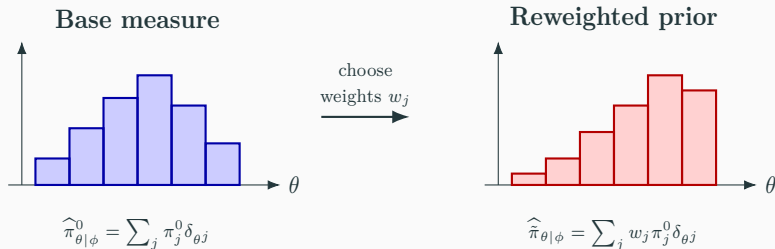
Conditional posterior sensitivity becomes linear in the prior.

Proposition 1

$\underline{m}_T(\phi)$ and $\bar{m}_T(\phi)$, are solutions to **infinite dimensional linear programs** if $\Pi_{\mathcal{M}, \mathcal{R}}(\phi)$ is defined by affine restrictions.

Simulation turns priors into weights

We approximate the class $\Pi_{\mathcal{M},\mathcal{R}}(\phi)$ by simulation:



Choose a convenient **reference/base measure** $\pi_{\theta|\phi}^0$ with full support on the identified set: $\text{supp}(\pi_{\theta|\phi}^0) = \text{IS}_{\theta}(\phi)$.

see Schennach (2014), Lam, Liu, and Singham (2024), etc.

An alternative prior is obtained by reweighting:

$$\hat{\tilde{\pi}}_{\theta|\phi} = \sum_{j=1}^J w_j \pi_j^0 \delta_{\theta j}, \quad w_j \geq 0, \quad \sum_{j=1}^J w_j \pi_j^0 = 1.$$

Does simulation recover the population bounds?

The simulated bounds are obtained from finite-dimensional LPs:

$$\widehat{m}_{T,J}(\phi) = \min_{w \in W_J(\phi)} \sum_{j=1}^J w_j \pi_j^0 T(\theta^j) \quad \widehat{m}_{T,J}(\phi) = \max_{w \in W_J(\phi)} \sum_{j=1}^J w_j \pi_j^0 T(\theta^j)$$

where $W_J(\phi) \equiv \left\{ w \in \mathbb{R}_+^J : \sum_{j=1}^J w_j \pi_j^0 = 1, \text{ s.t. } \mu \in \mathcal{M}(\phi) \text{ and } \rho \in \mathcal{R}(\phi) \right\}$.

Theorem 1

Let

$$\theta^1, \dots, \theta^J \stackrel{iid}{\sim} \pi_{\theta|\phi}^0, \quad \text{supp}(\pi_{\theta|\phi}^0) = IS_{\theta}(\phi).$$

Under regularity and support conditions, almost surely,

$$\widehat{m}_{T,J}(\phi) \rightarrow \underline{m}_T(\phi), \quad \widehat{m}_{T,J}(\phi) \rightarrow \overline{m}_T(\phi),$$

as $J \rightarrow \infty$.

Additional results in the paper

Many more properties are discussed in the paper

- Implication for joint prior [Details](#)
- Conditional vs unconditional restriction [Details](#)
- Relation to other robust Bayes approaches [Details](#) [Toy Example](#)
Giacomini and Kitagawa (2021), Giacomini, Kitagawa, and Uhlig (2019) and Ho (2023)
- Relation to “Uncertain identification” [Details](#)
Giacomini, Kitagawa, and Volpicella (2022)
- Relation to marginally specified priors [Details](#)
Kessler, Hoff, and Dunson (2015)

Revisiting the Oil Supply and Demand Debate

Set identification in SVARs

Consider a reduced-form VAR(1):

$$y_t = Ay_{t-1} + u_t, \quad \mathbb{E}[u_t u_t'] = \Sigma_u.$$

Structural shocks satisfy

$$u_t = B\varepsilon_t, \quad \mathbb{E}[\varepsilon_t \varepsilon_t'] = I, \quad \Sigma_u = BB'.$$

The reduced form identifies $\phi = (A, \Sigma_u)$, but not a unique structural impact matrix B . Let P be any factorization such that $PP' = \Sigma_u$:

$$B = PQ, \quad QQ' = I, \quad BB' = PQQ'P' = PP' = \Sigma_u.$$

Different rotations Q imply the same reduced-form.

b/c likelihood determined by first two moments.

Sign, zero, or narrative restrictions restrict $Q \in \mathcal{Q}(\phi)$ but typically do not select a unique $Q \Rightarrow$ **set identification**.

Prior information is usually incomplete

Different admissible rotations, Q , imply different

impulse responses, elasticities, FEVDs, policy rules, ...

A precise Bayesian analysis requires a **joint prior over the entire structural model**, i.e. about all of these features.

E.g. Haar-based priors (Arias and Waggoner 2024) or economically motivated priors (Baumeister and Hamilton 2015).

We may have **no prior view** on some impulse response at horizon 20, yet a full joint prior **necessarily induces one**.

⇒ With my framework we do not have to specify a full prior.

A brief incomplete history of the literature

Based on exclusion restrictions Kilian (2009) finds that oil demand shocks not supply shocks are dominating the movement in the real price of oil.

Evidence based on **sign restricted VARs** in Kilian and Murphy (2012), Kilian and Murphy (2014) (KM2014) support this.

Baumeister and Hamilton (2019) challenged this finding by replacing **hard restrictions** $0 \leq \alpha_s \leq 0.0258$ with **precise priors**.

These precise priors have been criticized e.g. by Kilian (2022) for their implications on other parts of the VAR.

Starting from KM2014 how much distrust in their assumption is needed to reach different conclusions?

Oil supply elasticity assumption

As discussed above Kilian and Murphy (2014) impose bounds for the supply elasticity, α_s :

$$0 \leq \alpha_s \leq 0.0258$$

For both flow demand and speculative supply elasticities.

(KM2014 also restrict demand elasticities, etc.)

This might not be what we literally believe or we want to understand what happens under **soft relaxation** of those bounds:

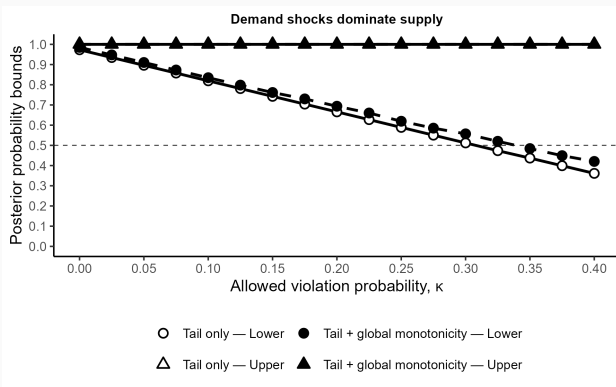
$$P(0 \leq \alpha_s \leq 0.0258) \geq 1 - \kappa$$

Here, $S(\theta)$ are the elasticities above.

Breakdown-frontier for demand dominance

$$\underline{p}(\kappa) = \inf_{\pi \in \Pi_{\kappa}} P_{\pi}(\text{FEVD}_{12}^{\text{demand}} - \text{FEVD}_{12}^{\text{supply}} > 0)$$

$$\bar{p}(\kappa) = \sup_{\pi \in \Pi_{\kappa}} P_{\pi}(\text{FEVD}_{12}^{\text{demand}} - \text{FEVD}_{12}^{\text{supply}} > 0)$$

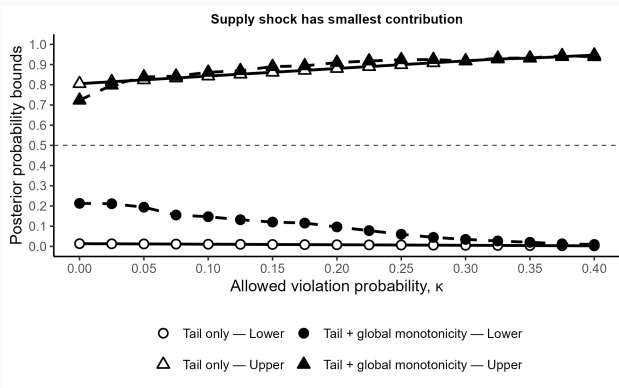


Breakdown-frontier for supply irrelevance

Similarly for supply irrelevance

$$\underline{p}(\kappa) = \inf_{\pi \in \Pi_{\kappa}} P_{\pi}(\text{FEVD}_{12}^{\text{supply}} \text{ is smallest shock})$$

$$\bar{p}(\kappa) = \sup_{\pi \in \Pi_{\kappa}} P_{\pi}(\text{FEVD}_{12}^{\text{supply}} \text{ is smallest shock})$$



Conclusion

Set-identification often leaves us with uninformative bounds (*Pig example: $W \in [0, 70]$*). Prior information might help but:

Our economic assumptions often only partially identify prior beliefs just as data can partially identify parameters.

⇒ I show how to bound posterior quantities based on partial priors.

Imposing some prior features without committing to a precise prior is often possible and this paper shows how to do it.

For SVARs, easier to justify than Haar prior or **full-ambiguity** robust framework (Giacomini and Kitagawa (2021)).

Many more applications possible:

- How much belief in sacrifice ratio positive necessary that no output puzzle?
- Share of missing-at-random observations such that treatment effect retains sign?
- ...

Thanks a lot!

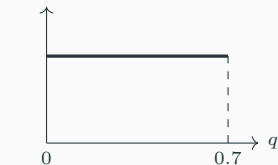
`martin.fankhauser@unibocconi.it`

`www.martin-fankhauser.at`

Inference based on only what we know

Instead, in this paper I show how to do inference considering all priors with a specific feature. For this example:

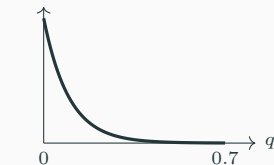
Flattest admissible prior



most weight on large q
allowed by monotonicity

$$E[W \mid \text{data}] = 48.4 \text{ kg}$$

Concentrate near zero



arbitrarily much weight
near $q = 0$

$$E[W \mid \text{data}] \rightarrow 70 \text{ kg}$$

Thus we can bound posterior expectations under **all weakly decreasing priors**:

$$E[W \mid \text{data}] \in [48.4, 70] \text{ kg}$$

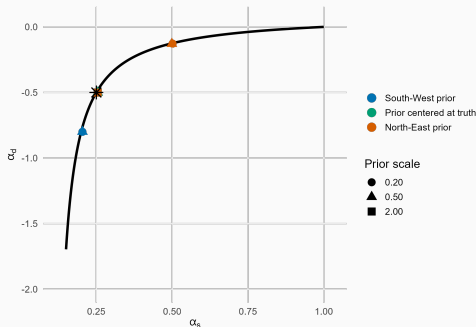
More diffuse prior does not let the likelihood decide

Suppose a supply and demand system (price p , quantity q).

$$q_t = \alpha_s p_t + u_t^s,$$

$$q_t = \alpha_d p_t + u_t^d$$

Assume $\alpha_s > 0$ and $\alpha_d < 0$.



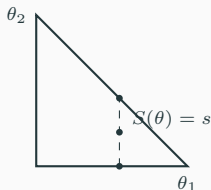
black line: identified set, colors: different prior centers, shapes: different prior variance.

The marginal does not determine the prior

Consider

$$\theta = (\theta_1, \theta_2), \quad 0 \leq \theta_1 \leq 1, \quad 0 \leq \theta_2 \leq 1 - \theta_1,$$

and suppose our prior information is: $\mathbf{S}(\theta) = \theta_1 \sim \mathbf{U}[0, 1]$.



For each s , possible completions for θ_2 include

$$\rho_1(\theta_2 | s) = \delta_0(\theta_2), \quad \rho_2(\theta_2 | s) = U[0, 1-s](\theta_2), \quad \rho_3(\theta_2 | s) = \delta_{1-s}(\theta_2).$$

All imply the same marginal $\theta_1 \sim U[0, 1]$, but for $T(\theta) = \theta_1 + \theta_2$:

$$E[T] = \frac{1}{2}, \quad E[T] = \frac{3}{4} \quad \text{and} \quad E[T] = 1. \quad \text{Back}$$

Undisputed (hard) assumptions

The identified set, i.e. the smallest set of parameter values consistent with the data and model, is defined relative to the assumptions which define the model.

Assume that there is a set of assumptions $\mathcal{A}_0$ we take for granted.

Sheeps and pigs example:

Recall we assumed $\mathbf{W} \geq \mathbf{0}$ and also that the **average weight of pig** is 100 kg. We assumed they belong to $\mathcal{A}_0$.

The identified set relative to those was $\mathbf{W} \in [\mathbf{0}, \mathbf{70}]$.

What are the elements of $\mathcal{A}_0$ depend on the application. Importantly, my framework allows to impose any further assumptions probabilistically. [Back](#)

What prior information gives a linear program?

Many economically meaningful restrictions are affine in π :

- **Moment restrictions**

$$\underline{m} \leq E_{\pi}[h(S(\theta))] \leq \overline{m}.$$

- **Probability restrictions**

$$\underline{p} \leq P_{\pi}(S(\theta) \in B) \leq \overline{p}.$$

- **Fixed marginal**

$$S_{\#}\pi = \mu^0.$$

- **Shape restrictions**

e.g. a weakly decreasing marginal density for $S(\theta)$.

Nonlinear economic objects can still generate linear restrictions on the prior.

Regularity for the simulation approximation

Assumption 1

1. **Compactness:** $(\text{IS}_\theta(\phi), d)$ is a nonempty compact metric space.
2. **Restrictions:** J is countable and L finite; $a_j \in C(\text{IS}_\theta(\phi))$, $b_j \in \mathbb{R}$; each G_ℓ is relatively open in $\text{IS}_\theta(\phi)$ and $\kappa_\ell \in [0, 1]$.
3. **Finite approximation:** each $J_r \subseteq J$ is finite, $\Pi^{(r+1)}(\phi) \subseteq \Pi^{(r)}(\phi)$ and $\bigcap_{r \geq 1} \Pi^{(r)}(\phi) = \Pi(\phi)$.
4. **Simulation support:** $\theta^{(1)}, \theta^{(2)}, \dots \stackrel{iid}{\sim} \nu$, and for every $C \in \mathcal{C}$, $x \in C$, and $\epsilon > 0$, $\nu(B(x, \epsilon) \cap C) > 0$.
5. **Strict feasibility:** for every r , there exists $\pi_r^\circ \in \Pi^{(r)}(\phi)$ such that

$$\min_{j \in J_r} \left\{ b_j - \int_{\text{IS}_\theta(\phi)} a_j(\theta) \pi_r^\circ(d\theta) \right\} > 0.$$

6. **Event targets:** for a relatively closed $A \subseteq \text{IS}_\theta(\phi)$,

$$\inf_{\pi \in \Pi(\phi)} \pi\left(\text{int}_{\text{IS}_\theta(\phi)} A\right) = \inf_{\pi \in \Pi(\phi)} \pi(A).$$

From conditional to joint posterior bounds

Keeping π_ϕ fixed induces the joint class

$$\Pi_{\mathcal{M}, \mathcal{R}}^J = \left\{ \pi_\phi(d\phi) \pi_{\theta|\phi}(d\theta | \phi) : \pi_{\theta|\phi}(\cdot | \phi) \in \Pi_{\mathcal{M}, \mathcal{R}}(\phi) \right\}.$$

Under set identification,

$$\pi(d\phi, d\theta | Y) = \pi_\phi(d\phi | Y) \pi_{\theta|\phi}(d\theta | \phi).$$

Under the measurable-selection conditions of the paper,

$$\underline{m}_T(Y) = \int \underline{m}_T(\phi) \pi_\phi(d\phi | Y),$$

$$\overline{m}_T(Y) = \int \overline{m}_T(\phi) \pi_\phi(d\phi | Y).$$

Solve pointwise in ϕ , then integrate over reduced-form uncertainty.

Do conditional restrictions survive jointly?

Let

$$\mu_\phi = S_\# \pi_{\theta|\phi}.$$

If, for π_ϕ -almost every ϕ ,

$$\mu_\phi(B) \leq \kappa,$$

then

$$P_\pi(S(\theta) \in B) \leq \kappa.$$

If $Z = |S - c|$ has a conditionally **non-increasing density** supported on $[0, \bar{z}(\phi)]$, then its unconditional density is also non-increasing.

Both conclusions also hold for the posterior

after replacing π_ϕ by $\pi_\phi(\cdot | Y)$.

Back

If we partially trust a baseline prior

Let $\pi_{\theta|\phi}^0$ be a substantive baseline prior with

$$\pi_{S|\phi}^0 = S_{\#}\pi_{\theta|\phi}^0,$$

and let ρ_{ϕ}^0 denote its completion.

Suppose we **preserve the baseline completion** $\mathcal{R}(\phi) = \{\rho_{\phi}^0\}$ but allow the marginal

$$\mu \in \mathcal{M}_S(\phi), \quad \mu \ll \pi_{S|\phi}^0.$$

Then the unique induced structural prior is

$$\pi_{\theta|\phi}^{\mu}(d\theta) = \frac{d\mu}{d\pi_{S|\phi}^0}(S(\theta)) \pi_{\theta|\phi}^0(d\theta).$$

Change beliefs about $S(\theta)$ while preserving how the baseline completes them into a joint prior.

Marginal reweighting and KL robustness

For every admissible $\mu \ll \pi_{S|\phi}^0$, $\text{KL}\left(\pi_{\theta|\phi}^\mu \parallel \pi_{\theta|\phi}^0\right) = \text{KL}\left(\mu \parallel \pi_{S|\phi}^0\right)$.

Moreover, if the divergence is finite,

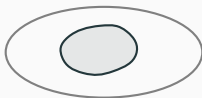
$$\pi_{\theta|\phi}^\mu = \arg \min_{\substack{q \ll \pi_{\theta|\phi}^0 \\ S_{\#} q = \mu}} \text{KL}\left(q \parallel \pi_{\theta|\phi}^0\right).$$

Fixing the marginal, preserving the baseline completion is KL-closest.

Define $\bar{\lambda}(\phi) = \sup_{\mu \in \mathcal{M}_S(\phi)} \text{KL}\left(\mu \parallel \pi_{S|\phi}^0\right)$.

Then every marginally reweighted prior lies in the structural KL ball:

$$\text{KL}\left(\pi_{\theta|\phi}^\mu \parallel \pi_{\theta|\phi}^0\right) \leq \bar{\lambda}(\phi).$$



Converse need not hold: the KL ball also allows changes in the completion.

[Back](#)

Qualitative Restrictions Can Shrink Ambiguity

Consider

$$\theta = \phi + \lambda U, \quad U \in [0, 1].$$

Under full ambiguity,

$$\overline{\{E[\theta \mid \phi]\}} = [\phi, \phi + \lambda].$$

Now require only that U has a weakly decreasing density.

The closure of the set of conditional posterior means is

$$\left[\phi, \phi + \frac{\lambda}{2} \right].$$

Qualitative prior information reduces ambiguity without choosing a precise distribution.

Corollary: Event probabilities are eventually exact

Suppose the prior class contains **only event restrictions**:

$$\Pi(\phi) = \{\pi \in \mathcal{P}(K_\phi) : \pi(G_\ell) \leq \kappa_\ell, \quad \ell = 1, \dots, L\},$$

Consider an **indicator target** $T = \mathbf{1}_A$. Partition K_ϕ into nonempty cells $\mathcal{C}$ according to membership in $G_1, \dots, G_L$ and A . Define the feasible cell masses:

$$\mathcal{D} = \left\{ p \in \Delta(\mathcal{C}) : \sum_{C \subseteq G_\ell} p_C \leq \kappa_\ell \quad \forall \ell \right\}.$$

Corollary 1. The population bounds reduce to finite LPs:

$$\underline{m}_T(\phi) = \min_{p \in \mathcal{D}} \sum_{C \subseteq A} p_C, \quad \overline{m}_T(\phi) = \max_{p \in \mathcal{D}} \sum_{C \subseteq A} p_C.$$

If $\theta^{(n)} \stackrel{iid}{\sim} \nu$ and $\nu(C) > 0$ for every $C \in \mathcal{C}$, then almost surely the simulated bounds **equal the population bounds** for all sufficiently large N .

Equality holds as soon as every cell has been sampled. With three event restrictions, there are at most $2^4 = 16$ cells.

How does this compare to other robust Bayes approaches?

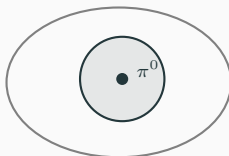
Full ambiguity



No additional
prior restrictions

Giacomini and Kitagawa (2021)

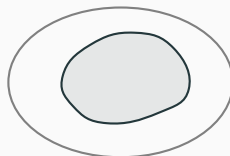
KL sensitivity



Perturbations around
a precise benchmark

Giacomini, Kitagawa, and Uhlig (2019)
Ho (2023)

This paper



Ambiguity defined by
economic knowledge

Different primitive: rather than allowing every prior or perturbing one benchmark, define the prior class directly from the features we can defend.

[Back: Literature](#)

[Back: Theory](#)

Comparison: Giacomini, Kitagawa, and Volpicella (2022)

Model averaging under uncertain identification: Assign prior probabilities to alternative models with different identifying assumptions.

What can this reproduce?

A bound on the probability of violating $S(\theta) \leq b$:

$$\underbrace{(1 - \kappa)\mathcal{P}(K_0) + \kappa\mathcal{P}(K)}_{\text{Model averaging}} = \underbrace{\{\pi \in \mathcal{P}(K) : P_\pi(S > b) \leq \kappa\}}_{\text{Our prior class}},$$

where $K_0 = \{\theta \in K : S(\theta) \leq b\}$ and K is the unrestricted identified set.

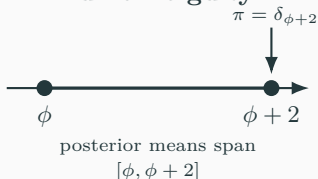
What about qualitative shape restrictions?

Requiring a **weakly decreasing density beyond** b would generally require a **continuum** of suitably specified models.

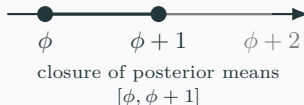
Full ambiguity versus partial prior information

Recall the toy set-identified model: $\theta \in [\phi, \phi + 2]$.

Full ambiguity



Weakly decreasing prior



Under full ambiguity, $\Pi(\phi) = \mathcal{P}([\phi, \phi + 2])$. To attain the upper posterior mean $\phi + 2$, the prior must be

$$\pi_{\theta|\phi} = \delta_{\phi+2}.$$

This prior might be considered extreme. For instance, it's ruled out by our weakly decreasing density class.

References

-  Antolin-Diaz, J. and J. Rubio-Ramirez (2018). “**Narrative Sign Restrictions for SVARs**”. In: *The American Economic Review* 108.10, pp. 2802–2829.
-  Arias, J., D. Caldara, and J. Rubio-Ramirez (2019). “**The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure**”. In: *Journal of Monetary Economics* 101, pp. 1–13.
-  Arias J. Rubio-Ramirez, J. and D. Waggoner (2024). “**Uniform Priors for Impulse Responses**”. In: *forthcoming*.
-  Baumeister, C. and J. Hamilton (2015). “**Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information**”. In: *Econometrica* 83.5, pp. 1963–1999.
-  — (2019). “**Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks**”. In: *American Economic Review* 109.5, pp. 1873–1910.
-  Berger, J. (1984). *The Robust Bayesian Viewpoint (With Discussion)*.

-  Caldara, D., M. Cavallo, and M. Iacoviello (2019). **“Oil Price Elasticities and Oil Price Fluctuations”**. In: *Journal of Monetary Economics* 103, pp. 1–20.
-  Giacomini, R. and T. Kitagawa (2021). **“Robust Bayesian Inference for Set-Identified Models”**. In: *Econometrica* 89, pp. 1519–1556.
-  Giacomini, R., T. Kitagawa, and H. Uhlig (2019). **“Estimation under Ambiguity”**. In.
-  Giacomini, R., T. Kitagawa, and A. Volpicella (2022). **“Uncertain Identification”**. In: *Quantitative Economics* 12, pp. 95–123.
-  Hansen, L. and T. Sargent (2001). **“Robust Control and Model Uncertainty”**. In: *American Economic Review* 91.2, pp. 60–66.
-  Ho, P. (2023). **“Global Robust Bayesian Analysis in Large Models”**. In: *Journal of Econometrics* 235, pp. 608–642.
-  Kessler, D., P. Hoff, and D. Dunson (2015). **“Marginally specified priors for non-parametric Bayesian estimation”**.

In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 77, pp. 35–58.



Kilian, L. (2009). **“Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market”**. In: *American Economic Review* 99.3, pp. 1053–69.



— (2022). **“Facts and Fiction in Oil Market Modeling”**. In: *Energy Economics* 110, p. 105973.



Kilian, L. and D. Murphy (2012). **“Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models”**. In: *Journal of European Economic Association* 10.5, pp. 1166–88.



— (2014). **“The Role of Inventories and Speculative Trading in the Global Market for Crude Oil”**. In: *Journal of Applied Econometrics* 29.3, pp. 454–78.



Lam, H., Z. Liu, and D. I. Singham (2024). **“Shape-Constrained Distributional Optimization via Importance-Weighted Sample Average Approximation”**. In: *arXiv preprint arXiv:2406.07825*.

-  Lavine, M., L. Wasserman, and R. Wolpert (1991). “**Bayesian inference with specified marginals**”. In: *Journal of the American Statistical Association* 86.416, pp. 964–971.
-  Ludvigson, S., S. Ma, and S. Ng (2017). “**Shock restricted structural vector-autoregressions**”. In: *National Bureau of Economic Research*.
-  Moon, H. and F. Schorfheide (2012). “**Bayesian and Frequentist Inference in Partially Identified Models**”. In: *Econometrica* 80.2, pp. 755–782.
-  Rubio-Ramirez, J., D. Waggoner, and T. Zha (2010). “**Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference**”. In: *The Review of Economic Studies* 77.2.
-  Schennach, S. (2014). “**Entropic Latent Variable Integration via Simulation**”. In: *Econometrica* 82.1, pp. 345–385.
-  Uhlig, H. (2005). “**What are the effects of monetary policy on output? Results from an agnostic identification procedure**”. In: *Journal of Monetary Economics* 52, pp. 381–419.