Variance attribution with an uncertain policy reaction: a conditional-moment identification hierarchy

6 November 2026
Preprint

Abstract

A policy action can reflect private news, uncertainty about the policy reaction, and an independent disturbance. When markets learn about news from the action, the associated outlook response depends on a nonlinear posterior. We study identification of five variance contributions in a one-period model with independent Gaussian primitives and a bilinear policy rule. Entire conditional covariance functions can coincide while every contribution curve differs. Even adding the entire conditional third cumulant of the action need not identify the contributions: we construct an admissible pair that also matches the actual posterior-dependent outlook variance throughout an interval. The construction uses a regular crossing of two posterior variances, established by uniform differentiated small-noise expansions. As a complementary result, the conditional third and fourth action cumulants together recover all model parameters and the private-news variance function. These are population identification results under a specified outlook equation and invariant scalar parameters. They establish neither empirical policy effects nor a stable finite-sample estimation procedure.

1 Introduction

A policy surprise is not necessarily a policy disturbance. A change in an observed action may combine news about economic conditions with uncertainty about how strongly the policymaker responds to those conditions. An accompanying forecast or valuation response may seem to resolve the decomposition. But when that response reflects inference from the same action, its information content depends on the posterior induced by the unknown model.

This note asks a specific population question. Suppose the econometrician knows the conditional covariance matrix of a policy action and an outlook response at every public state in an interval. Can these functions separate the variances attributable to private news, reaction uncertainty, their interaction, an independent disturbance, and primitive payoff news? We show that the answer can remain negative after adding the action's entire third-cumulant function. The key construction matches the nonlinear posterior variance itself. Matching only the action moments would not suffice.

The contribution is deliberately narrow. Higher moments are an established identification device, and the positive inversion using third and fourth cumulants in this model is elementary. The substantive negative result is a whole-interval, posterior-compatible observational equivalence under third-cumulant information. It demonstrates why a linear projection cannot silently replace Bayesian inference in an attribution argument. The result is existential: it does not say that ambiguity is generic or empirically prevalent.

Relation to existing work. Learning about changing policy reactions is already central to Bauer and Swanson's model [2, Section II]. Their noisy action reveals information about an unobserved reaction coefficient. Sastry's model [5, Sections 2.1 and 2.5] likewise distinguishes

observing a policy action from correctly inferring the underlying private signal, and uses additional forecast-error moments for identification. Our result concerns the particular bilinear action and posterior outlook equations below; neither learning from actions nor state-dependent policy uncertainty is claimed as new.

Higher-moment identification also has an extensive precedent. Andrade, Ferroni, and Melosi [1, Section 2 and Proposition 1] use higher-moment inequality restrictions to restrict orthogonal rotations in a linear structural-shock model. Guay [4, Section 3, Proposition 1] gives a sufficient Jacobian-rank condition for local statistical identification of linear structural vector autoregressions from unconditional second, third, and fourth moments. Those results concern fixed linear impact matrices and structural-shock higher-moment restrictions. Here the primitives are Gaussian, but their product creates non-Gaussian action distributions, and the second observable contains a conditional expectation. The relevant observables are functions conditional on a public state, and our negative result is an exact all-state equivalence. A linear-shock theorem cannot be transferred to this setting without an embedding that verifies its cross-moment restrictions and accounts for the posterior equation. The comparison with Guay uses the October 2020 working-paper version; we do not assume that every statement in the later journal version is identical.

For the product component, Gaunt [3, Theorem 2.1 and Corollary 2.10] supplies noncentral Gaussian-product representations and cumulants. Such results explain why the marginal moment algebra is tractable; they do not impose equality of the posterior outlook variance. Our claimed addition relative to these specified comparisons is the regular posterior crossing that makes coefficient-dual models compatible with that extra restriction on an interval. This bounded comparison is not a literature-wide priority claim, and the recovery formulas are presented as a complementary model-specific calculation.

2 Model, information, and targets

Let I be a compact interval with zero in its interior. Conditional on the public state $x \in I$, let η, u, ϵ be mutually independent centered Gaussian variables with variances $G(x), U, E$. The scalars satisfy

$$a, U, E > 0, \qquad b > 1, \qquad 0 < \rho \leq 1, \qquad G \in C(I), \quad 0 < \inf_I G \leq \sup_I G < \infty. \tag{1}$$

Only G may vary with x. The action surprise and the outlook response are

$$r = a\eta + u(x + \eta) + \epsilon, \qquad q = \rho m_x(r) - br, \qquad m_x(r) = \mathbb{E}[\eta \mid r, x]. \tag{2}$$

The market knows its own model parameters and observes the action and public state. The econometrician is given moment functions, not the posterior regression, the full conditional law, or realized primitive shocks. Both observables have conditional mean zero.

One interpretation of $x + \eta$ is the state to which policy responds and of $a + u$ is its uncertain reaction coefficient. Subtracting the known component ax gives the action equation in (2). The outlook equation is a maintained measurement relation: news raises the outlook through $\rho m_x(r)$ and the action lowers it through br. It is not asserted to be a general equity-pricing equation. We do not derive the sign and scale restrictions from an optimizing monetary model.

The target vector is

$$\mathcal{C}(x) = \big(a^2 G(x),\ Ux^2,\ UG(x),\ E,\ \rho^2 G(x)\big). \tag{3}$$

The first four entries sum to $\mathrm{Var}(r \mid x)$ because the four action components are pairwise uncorrelated, although they are not mutually independent. The fifth entry is the variance of

primitive payoff news $\rho\eta$. It differs from the variance $\rho^2 M(x)$ of the posterior news entering q, where

$$M(x) = \mathbb{E}[m_x(r)^2 \mid x], \qquad P(x) = M(x) - \frac{a^2 G(x)^2}{V(x)} \geq 0. \tag{4}$$

Here P measures the nonlinear part of the posterior after projecting on r.

Our observation sets are the exact functions

$$\begin{aligned} V(x) &= \mathrm{Var}(r \mid x), \qquad & S(x) &= \mathrm{Cov}(r, q \mid x), \qquad & T(x) &= \mathrm{Var}(q \mid x), \\ K_3(x) &= \mathbb{E}[r^3 \mid x], \qquad & K_4(x) &= \mathbb{E}[r^4 \mid x] - 3V(x)^2. \end{aligned} \tag{5}$$

Identification means uniqueness among all tuples obeying (1) that produce the stated functions on the same interval. It is not identification of event-level realizations. In particular, access to functions arbitrarily close to zero is much stronger than access to a finite collection of estimated conditional moments.

3 The identification hierarchy

Theorem 1. *Under (1)–(2):*

1. *For every such interval I, two admissible models can have identical (V, S, T) and different curves for all five entries of $\mathcal{C}$. The curves Ux^2 necessarily agree at $x = 0$. The witnesses can have $V \notin \mathrm{span}\{1, x^2\}$.*
2. *On an admissible interval, two models can have identical (V, S, T, K_3) while their a^2G, UG, and $\rho^2 G$ differ at every state.*
3. *The functions (V, S, T, K_3, K_4) identify all five scalars a, U, E, ρ, b and the entire function G, including when G is constant or affine in x^2.*

The third assertion is sufficiency for this particular augmented information set. The theorem makes no claim that K_4 alone suffices, that every identification procedure needs fourth moments, or that higher-moment estimation is practically reliable. The first assertion is proved in Appendix B; the more demanding second assertion is developed in Section 5 and Appendix A. We first give the transparent forward identities and positive inverse.

4 Forward moments and complementary recovery result

Write $g = G(x)$, $B = Ug$, and $A = a^2 g + Ux^2$. Gaussian integration gives, for $|t| < B^{-1/2}$,

$$\log \mathbb{E}[e^{tr} \mid x] = \frac{Et^2}{2} - \frac{1}{2}\log(1 - Bt^2) + \frac{At^2 + 2axBt^3}{2(1 - Bt^2)}. \tag{6}$$

Consequently,

$$\begin{aligned} V &= (a^2 + U)g + Ux^2 + E, \\ K_3 &= 6axUg, \qquad K_4 = 6(Ug)^2 + 12(a^2 g + Ux^2)Ug. \end{aligned} \tag{7}$$

The tower property gives $\mathbb{E}[r m_x(r) \mid x] = aG(x)$, so

$$\begin{aligned} S &= \rho a G - bV, \qquad & T &= \rho^2 M - 2\rho b a G + b^2 V, \\ D &:= T - S^2/V = \rho^2 P. \end{aligned} \tag{8}$$

For $x \neq 0$, define $H(x) = K_3(x)/(6x)$ and use its continuous extension at zero. In every compatible model, $H = aUG > 0$. The following quantities are observable:

$$d = \frac{K_4(0)}{6H(0)^2}, \qquad k = \frac{K_4(x)/(6H(x)) - dH(x)}{2x^2} \quad (x \neq 0). \tag{9}$$

Substitution shows $d = a^{-2} + 2/U$ and $k = U/a$, independent of the chosen nonzero state. Thus, setting $z = 1/a$,

$$z = \frac{d}{\sqrt{k^{-2} + d} + k^{-1}}, \quad a = z^{-1}, \quad U = ka, \quad G(x) = \frac{H(x)}{aU}, \quad E = V(x) - (a^2 + U)G(x) - Ux^2. \tag{10}$$

The equation $d = z^2 + 2z/k$ has exactly one positive solution. This proves global uniqueness of the rate parameters and G, without differentiability of G or a rank condition.

To recover the outlook scalars, define φ_v as the centered normal density with variance v and write

$$\begin{aligned} f_x(y) &= \int_{\mathbb{R}} \varphi_{G(x)}(z)\varphi_{E+U(x+z)^2}(y - az)\, dz, \\ h_x(y) &= \int_{\mathbb{R}} z\varphi_{G(x)}(z)\varphi_{E+U(x+z)^2}(y - az)\, dz, \qquad M(x) = \int_{\mathbb{R}} \frac{h_x(y)^2}{f_x(y)}\, dy. \end{aligned} \tag{11}$$

The positive noise variance makes $f_x > 0$. Weighted Cauchy–Schwarz gives

$$h_x(y)^2/f_x(y) \leq \int z^2\varphi_G(z)\varphi_{E+U(x+z)^2}(y - az)\, dz,$$

hence $M \leq G$. These are uniquely specified posterior integrals, not additional observations.

Lemma 2. *For every admissible tuple, $P(0) > 0$.*

Proof. If $P(0) = 0$, then $m_0(r) = \alpha r$ almost surely, with $\alpha = ag/V$. Let $F(t) = \mathbb{E}[e^{tr} \mid 0]$ and $B = Ug$. Completing the Gaussian square gives

$$\frac{\mathbb{E}[\eta e^{tr} \mid 0]}{F(t)} = \frac{agt}{1 - Bt^2}, \qquad \frac{F'(t)}{F(t)} = Et + \frac{Bt}{1 - Bt^2} + \frac{a^2gt}{(1 - Bt^2)^2}.$$

All expectations and derivatives exist on a neighborhood of zero by (6) and Cauchy–Schwarz. Multiplying the linear-posterior identity by e^{tr} and taking expectations would imply

$$V(1 - Bt^2) = E(1 - Bt^2)^2 + B(1 - Bt^2) + a^2g.$$

The right-hand coefficient of t^4 is $EB^2 > 0$ and the left-hand coefficient is zero, a contradiction. □

It now follows from (8) that

$$\rho = \sqrt{\frac{D(0)}{P(0)}}, \qquad b = \frac{\rho aG(0) - S(0)}{V(0)}. \tag{12}$$

The positive root is fixed by $\rho > 0$. Equations (10) and (12) prove part 3 of Theorem 1. Compatibility requires positivity, constancy of the recovered scalars, and all-state agreement with the posterior formulas. The formulas are not a claim that arbitrary supplied functions are compatible.

5 Why the third cumulant can still fail

The transformation

$$a' = U/a, \qquad G' = a^2 G/U, \qquad (U', E', \rho', b') = (U, E, \rho, b) \tag{13}$$

preserves V, S, K_3 and exchanges the rate contributions a^2G and UG. It does not automatically preserve T: that requires equality of the actual posterior variances. Establishing that equality on a whole interval is the central step.

Let Z, W, N be independent standard normals and define

$$\begin{aligned} A(t,s) &= \mathbb{E}\Big[\mathbb{E}[Z \mid Z(1+tW) + \sqrt{s}N]^2\Big], \\ B(t,s) &= \mathbb{E}\Big[\mathbb{E}[Z \mid Z(W+t) + \sqrt{s}N]^2\Big]. \end{aligned} \tag{14}$$

Write $M(a, U, E, g, x)$ for the posterior variance in (11). If $k = 1/t$, $g = t^2/s$, and $U = E = 1$, normalization of the observations gives

$$F(0,g) := M(k,1,1,g,0) - M(1/k,1,1,k^2 g,0) = g\{A(t,s) - t^{-2}B(t,s)\}. \tag{15}$$

The factor t^{-2} reflects the different latent variances and is essential.

Proposition 3 (Regular posterior crossing)**.** *There exist $k \geq 3$ and $g_0 > 0$ with $F(0, g_0) = 0$ and $F_g(0, g_0) > 0$.*

Proof. Appendix A proves, with positive constants C and $c \in [1, 5/3]$, that for $L_t = \log(1/t)$ and $s(t,z) = zt^{8/3}L_t^{2/3}$,

$$\frac{A(t, s(t,z)) - t^{-2}B(t, s(t,z))}{t^4} \longrightarrow 7 - c - \frac{3C}{4}z^{3/2} \quad \text{in } C^1_{\text{loc}}((0,\infty)). \tag{16}$$

Its limit has a unique simple root $z_* = (4(7-c)/(3C))^{2/3}$. Choose $t = 1/n$ with an integer $n \geq 3$ sufficiently large that the left side has opposite signs at $z_*/2$ and $2z_*$ and strictly negative derivative between them. Uniform C^1 convergence ensures such an integer exists. Let z_0 be its unique root, $k = n$, $s_0 = s(t, z_0)$, and $g_0 = t^2/s_0$. At this root, differentiating (15) with $s = t^2/g$ gives

$$F_g(0, g_0) = -s_0 \partial_s \{A(t, s_0) - t^{-2}B(t, s_0)\} > 0.$$

□

For completeness, M is continuously differentiable in positive parameters and x. Differentiate the joint Gaussian density underlying (11). Each derivative is the density times a polynomial in standardized Gaussian latent variables with bounded coefficients on positive compact parameter neighborhoods. Differentiation of h^2/f produces products of conditional moments. Conditional Hölder inequalities bound their integrals by complete-data Gaussian moments; high moments also bound the tails uniformly. On compact y intervals the positive density has a uniform positive minimum. Compact convergence and uniform tail bounds justify differentiation. Appendix B gives the same argument at the auxiliary $U = 0$ boundary, where positive E remains available.

Apply the implicit-function theorem to

$$F(x,g) = M(k,1,1,g,x) - M(1/k,1,1,k^2 g,x).$$

Proposition 3 supplies a C^1 positive root function $g = G(x)$ on some compact interval $I = [-\delta, \delta]$ with $\delta > 0$, and $G(0) = g_0$. Shrink the interval so G is bounded above and away from zero. The two admissible tuples are

$$(a, U, E, \rho, b, G) = (k, 1, 1, 1/2, 2, G), \qquad (1/k, 1, 1, 1/2, 2, k^2 G). \tag{17}$$

They share the actual M function, and hence share

$$V = (k^2+1)G + x^2 + 1, \quad S = kG/2 - 2V, \quad K_3 = 6kxG, \quad T = M/4 - 2kG + 4V.$$

Their target vectors are respectively

$$(k^2G, x^2, G, 1, G/4), \qquad (G, x^2, k^2G, 1, k^2G/4).$$

Three entries differ everywhere because $k \geq 3$ and $G > 0$. This proves part 2. The construction is exact but existential: it specifies neither a certified numerical n nor an empirically calibrated interval. No numerical root fitting is used as evidence for all-state equality.

6 Interpretation and limitations

The contrast concerns information sets within one model. Covariance variation across public states is not universally sufficient for the attribution in (3); the entire third-cumulant function can also fail. This remains true when the outlook response uses each model's own Bayesian posterior, so the counterexample is not created by assigning an arbitrary second measurement equation after matching the rate moments.

The fourth-cumulant addition succeeds because it separates the coefficients of the private-news and interaction terms after normalization by K_3/x. The outlook covariance residual then identifies the transmission parameter through the strictly nonlinear posterior at zero. This reasoning relies on positive interaction and disturbance variances, sign restrictions, and common scalar parameters across states. It makes no claim for zero variances, state-varying a, U, E, ρ, b, non-Gaussian primitives, misspecified learning, or a different outlook equation.

Several practical conclusions do not follow. The inverse divides by x and x^2, and can be ill-conditioned when moments are estimated near zero or parameters approach excluded boundaries. No sampling scheme, uniform conditioning bound, estimator, or finite-sample inference is supplied. The identified primitive variances are not causal impulse responses, and the existential counterexamples do not measure ambiguity in actual policy data. The results also do not describe sharp covariance-only identified sets or prove universal finiteness of their structural preimages.

7 Conclusion

In the maintained bilinear Gaussian model, third-cumulant information can leave exact variance attribution ambiguous even after the nonlinear posterior-dependent outlook covariance is imposed throughout an interval. Adding both third and fourth action cumulants yields global population recovery. The distinction identifies a precise limit of covariance and skewness information, while keeping the positive reconstruction conditional on the model and information assumptions.

Reproducibility

All existence and uniformity arguments are given below. The accompanying supplement contains a Python 3 standard-library checker for the Gaussian moment identities, coefficient-duality algebra, inverse equations, and finite coefficient balances. It uses exact rational arithmetic and generates no simulated data. These checks do not mechanically certify analytical remainder bounds or provide a numerical posterior-crossing witness. Compile this source using `latexmk -pdf article.tex`. The source archive includes every local TEX input.

A Uniform posterior expansions and the regular crossing

This appendix proves (16). All random variables Z, W, N below are independent standard normals. All derivatives with respect to s are at strictly positive s. The symbol O is uniform in the stated parameter sets.

A.1 Conditional-moment differentiation

We first state the integrability argument used repeatedly. Let X be a Gaussian latent vector and $Y_s = p(X) + \sqrt{s}N$, with a fixed polynomial p. For a latent function H, let $h_{H,s}(y) = f_s(y)\mathbb{E}[H \mid Y_s = y]$ be its weighted observation density. At fixed y,

$$s\partial_s h_{H,s}(y) = h_{HQ,s}(y), \qquad Q = (N^2 - 1)/2, \tag{18}$$

when H does not itself depend on s. This follows by differentiating the Gaussian noise kernel; $N = (y - p(X))/\sqrt{s}$ in that integral.

Lemma 4 (Uniform integral and tail bounds)**.** *Let $H_1, \dots, H_m$ be independent of s in the fixed-observation kernel: they are functions of X and possibly the fixed observed value y, not of the standardized noise $N = (y - p(X))/\sqrt{s}$. Suppose they have all moments uniformly over the parameter ranges under consideration, and the same holds after multiplication by Q. Any fixed finite sum of terms*

$$f_s(y) \prod_{i=1}^{m} \mathbb{E}[H_i \mid Y_s = y]$$

has uniformly bounded integral of its absolute value and uniformly bounded $L^p(f_s(y)dy)$ norm after removing the leading f_s, for any fixed finite p. Its logarithmic noise derivative has the same property. If conditional denominators $\mathbb{E}[\cosh(tW) \mid Y_s]^{-j}$ occur, with $j \geq 0$ and $|t| \leq t_0$, these conclusions remain true for any fixed number of t derivatives when the numerators are Gaussian polynomial–exponential moments. The bounds are uniform for $0 < s \leq s_0$.

Proof. Conditional Hölder and Jensen bound products of conditional expectations in $L^p(f_s dy)$ by unconditional higher moments of the H_i. Formula (18) and the quotient rule replace a conditional moment μ_H by

$$s\partial_s \mu_H = \mu_{HQ} - \mu_H \mu_Q.$$

The derivative of the leading density contributes μ_Q. These are again finite products with finite higher moments. The score-generated factors $H_i Q$ are used for these bounds, not treated as s-independent targets under a second noise differentiation. The conditional cosh denominator is at least one. Differentiating its inverse produces only inverse powers and products of its derivatives, which are conditional moments of $W^j \cosh(tW)$ or $W^j \sinh(tW)$. The latter have uniformly finite moments for $|t| \leq t_0$. The same score rule handles their noise derivatives. In particular an integral restricted to an event of observation probability ε is bounded by $C\varepsilon^{1-1/p}$. This proves uniform integrability and the tail assertion implicit in the statement. Differentiation under the integral follows first on compact observation intervals, where the positive density is bounded below, and then by these bounds. □

The lemma also applies with indicators of latent tail events inserted in H_i. Hölder's inequality then bounds the associated terms by a positive power of that event's probability and a Gaussian higher moment. It does not claim pointwise boundedness of conditional moments or unweighted score ratios.

A.2 Information in a smoothed Gaussian product

Let $P = ZW$, and let f_s be the density of $P + \sqrt{s}N$. Define the weighted density $h_s(y)dy = \mathbb{E}[W^2 \mathbf{1}_{\{P+\sqrt{s}N \in dy\}}]$. Direct Gaussian integration gives

$$f_0(y) = \frac{1}{2\pi} \int_0^\infty e^{-(v+y^2/v)/2} \frac{dv}{v}, \qquad h_0(y) = \frac{1}{2\pi} \int_0^\infty e^{-(v+y^2/v)/2}\, dv. \tag{19}$$

Thus $h_0(0) = 1/\pi$, $h_0' = -yf_0$, and for a constant d_0,

$$f_0(y) = \pi^{-1} \log(1/|y|) + d_0 + O(y^2(1 + |\log|y||)) \quad (y \to 0). \tag{20}$$

For example, splitting the first integral at y^2 and at 1 gives the logarithmic coefficient. Differentiation away from zero gives $f_0'' + f_0'/y - f_0 = 0$, which supplies the corresponding differentiated remainders. Both weighted densities and their derivatives are smooth off zero. Splitting the integral at fixed multiples of $|y|$, or rescaling $v = |y|w$, bounds their tails by a polynomial in $|y|$ times $e^{-c|y|}$, with $c > 0$.

Let $L_s = \log(1/\sqrt{s})$, $H(z) = \mathbb{E}[-\log|z - N|]$, and

$$C = \pi^{-1} \int_{\mathbb{R}} H'(z)^2\, dz. \tag{21}$$

Gaussian convolution makes H smooth. Splitting its convolution into $|N| \le |z|/2$ and its complement gives $H'(z) = -1/z + O(z^{-3})$ and $H''(z) = 1/z^2 + O(z^{-4})$ as $|z| \to \infty$. Thus $0 < C < \infty$.

Lemma 5 (Differentiated information asymptotics)**.** *For $I(s) = \int f_s'(y)^2/f_s(y)\, dy$,*

$$I(s) \sim \frac{C}{\sqrt{s}L_s}, \qquad I'(s) \sim -\frac{C}{2s^{3/2}L_s} \quad (s \downarrow 0). \tag{22}$$

Proof. Convolving (20) and differentiating the Gaussian kernel yields, locally uniformly in z,

$$\begin{aligned} f_s(\sqrt{s}z) &= \pi^{-1}(L_s + H(z)) + d_0 + o(1), \\ \sqrt{s}f_s'(\sqrt{s}z) &\longrightarrow \pi^{-1}H'(z), \qquad sf_s''(\sqrt{s}z) \longrightarrow \pi^{-1}H''(z). \end{aligned} \tag{23}$$

Here differentiating the locally integrable logarithm is understood by convolution, not by integrating a nonintegrable pointwise derivative.

We record bounds that permit integration of these limits. Fix $0 < \delta < 1/4$. For $|y| \le \delta$ and small s, set $r = |y| + \sqrt{s}$. There are positive constants independent of s such that

$$\begin{aligned} f_s(y) &\ge c_0(1 + \log(1/r)), & |f_s^{(j)}(y)| &\le C_j r^{-j} \quad (j = 1, 2, 3), \\ |h_s(y)| &\le C_0, & |h_s''(y)| &\le C_0(1 + |\log r|). \end{aligned} \tag{24}$$

To obtain these estimates, when $|y| \le 2\sqrt{s}$ rescale the convolution by $\sqrt{s}$; constants cancel against each differentiated Gaussian kernel and the remaining logarithm is integrable. When $|y| > 2\sqrt{s}$, split the convolution at $|u - y| = |y|/2$, use differentiated (20) on the near part, and Gaussian tail bounds on the complement. The lower bound follows by retaining a fixed-probability noise interval on which $|y - \sqrt{s}N| \le Cr$. The bound for h_s'' also follows from $h_s' = -yf_s - sf_s'$.

The outer estimates required here are *density-weighted*: for $j \le 3$, the integrals on $|y| \ge \delta$ of f_s times each fixed product of the score ratios $f_s^{(j)}/f_s$ used below are bounded uniformly as $s \downarrow 0$. The analogous weighted expressions involving h_s have the same property. To see this, on a fixed annulus use smooth convergence and positivity. At infinity, (19) gives upper and lower bounds of the form $e^{-|y|}$ times powers of $|y|$ for f_0, h_0 and polynomially bounded derivative ratios. Convolution for $s \le s_0$ preserves polynomial-exponential envelopes with a common integrable

exponential factor after multiplication by the density. The exponentially small contribution from the logarithmic singularity is bounded using Gaussian kernel derivatives. The ratios themselves need not decay.

For $I(s)$, after $y = \sqrt{s}z$, the normalized integrand is bounded on the inner region by

$$C_1(1+|z|)^{-2}\frac{L_s}{1+\log(1/(\sqrt{s}(1+|z|)))}.$$

For $J(s) = \int f_s((\log f_s)'')^2$, the corresponding bound has $(1+|z|)^{-4}$. On $|z| \le s^{-1/4}$ the logarithmic ratio is bounded; beyond that point its bound $O(L_s)$ times either integrable power has vanishing integral. The outer weighted bounds make its normalized contribution negligible. Dominated convergence and (23) therefore give

$$\sqrt{s}L_s I(s) \to C, \qquad s^{3/2}L_s J(s) \to \pi^{-1}\int H''(z)^2\,dz.$$

The heat equation $\partial_s f_s = f_s''/2$ and integration by parts yield $I'(s) = -J(s)$; the stated weighted estimates justify the boundary terms at every $s > 0$. Scaling of the heat convolution of $-\log|y|$ gives $H'' + zH' = -1$. Differentiate, multiply by H', and integrate to obtain $\int H''^2 = \frac{1}{2}\int H'^2$. This proves (22), including its derivative constant. □

A.3 The small-mean posterior

Under the base observation $Y = P + \sqrt{s}N$, write

$$n_j(y)dy = \mathbb{E}[ZW^j\mathbf{1}_{\{Y\in dy\}}], \qquad f_j(y)dy = \mathbb{E}[W^j\mathbf{1}_{\{Y\in dy\}}].$$

In this subsection $f_0 = f_s$ and $f_2 = h_s$. Shifting the mean of W by t tilts its density by $e^{tW-t^2/2}$. Conditional sign symmetry gives the exact expression

$$B(t,s) = e^{-t^2/2}\int f_s(y)\frac{\mathbb{E}[Z\sinh(tW)\mid Y=y]^2}{\mathbb{E}[\cosh(tW)\mid Y=y]}\,dy. \tag{25}$$

Lemma 6 (Uniform small-mean expansion)**.** *For small $|t|$, uniformly over $0 < s \le s_0$,*

$$t^{-2}B(t,s) = A_0(s) + t^2 b_2(s) + t^4 b_4(s) + R_B(t,s), \quad |R_B| + |s\partial_s R_B| \le C_2 t^6, \tag{26}$$

where

$$\begin{aligned} &A_0(s) = 1 - s + s^2 I(s), \\ &b_2(s) = -1 + O(s), \qquad b_2'(s) = O(1), \\ &b_4(s) \to c, \qquad sb_4'(s) \to 0, \qquad c = \frac{1}{9}\mathbb{E}[P^2\mathbb{E}[W^2\mid P]^2] \in [1, 5/3]. \end{aligned} \tag{27}$$

Proof. Apply Taylor's theorem through order seven to (25). Lemma 4 bounds its eighth t derivative and its logarithmic noise derivative uniformly. Parity removes odd powers and the expression starts at order t^2. Applying the same argument to $s\partial_s B$ proves the differentiated remainder in (26); no differentiation of an unproved asymptotic is used.

The first coefficients obtained by division of the series are

$$A_0 = \int \frac{n_1^2}{f_s}, \qquad b_2 = \int\left(\frac{n_1 n_3}{3f_s} - \frac{n_1^2 h_s}{2f_s^2} - \frac{n_1^2}{2f_s}\right).$$

Integration by parts gives $n_1 = yf_s + sf_s'$, $n_3 = yh_s + sh_s'$, and $h_s' = -n_1$. Since $\mathbb{E}[Y^2] = 1 + s$, the first identity in (27) follows. Set $\ell_s = f_s'/f_s$,

$$R_1 = \int yh_s\ell_s, \qquad Q_1 = \int h_s\ell_s^2.$$

Using $\mathbb{E}[Y^2W^2] = 3 + s$, direct expansion gives

$$b_2 = -1 - \frac{2s}{3}R_1 - \frac{s^2}{2}Q_1 + \frac{s^2}{3} - \left(\frac{s^2}{2} + \frac{s^3}{3}\right) I. \tag{28}$$

The inner bounds (24), the heat equation, and the weighted outer bounds give

$$R_1 = O(1),\ R_1' = O(s^{-1/2}),\ Q_1 = O(s^{-1/2}),\ Q_1' = O(s^{-3/2}),\quad I = O(s^{-1/2}),\ I' = O(s^{-3/2}).$$

For clarity, the possibly singular integrands in R_1' are bounded by a constant times $|y|/(|y| + \sqrt{s})^3$, and those in Q_1' by a constant times $(|y| + \sqrt{s})^{-4}$. Terms from h_s'' contribute at most $O(s^{-1/2}\log(1/s))$ to Q_1', absorbed by the displayed bound. Substitution in (28), and then differentiation of that exact identity, gives $b_2 = -1 + O(s)$ and $b_2' = O(1)$.

To make the sixth-order coefficient explicit, let $v_j = n_j/f_s$ and $\mu_j = f_j/f_s$. Then

$$b_4(s) = \int f_s \Big\{ \frac{v_3^2}{36} + \frac{v_1 v_5}{60} - \frac{v_1 v_3 \mu_2}{6} + \frac{v_1^2 \mu_2^2}{4} - \frac{v_1^2 \mu_4}{24} - \frac{v_1 v_3}{6} + \frac{v_1^2 \mu_2}{4} + \frac{v_1^2}{8} \Big\}\, dy. \tag{29}$$

At $s = 0$, let $m_j(P) = \mathbb{E}[W^j \mid P]$. The coefficient of t^6 in (25) integrates to

$$\mathbb{E}\left[P^2\left(\frac{m_2^2}{9} - \frac{m_4}{40}\right)\right] + \frac{\mathbb{E}[P^2 m_2]}{12} + \frac{\mathbb{E}[P^2]}{8} = c,$$

since $\mathbb{E}[P^2] = 1$, $\mathbb{E}[P^2W^2] = 3$, and $\mathbb{E}[P^2W^4] = 15$. Weighted Cauchy–Schwarz gives $\mathbb{E}[P^2 m_2^2] \geq 9$; conditional Jensen gives $\mathbb{E}[P^2 m_2^2] \leq 15$.

It remains to justify both limits involving b_4. Each coefficient is an integral of a finite product of conditional Gaussian-polynomial moments. For fixed $y \neq 0$, the weighted product densities and their derivatives converge smoothly as $s \downarrow 0$. Their logarithmic noise derivatives tend to zero there. Lemma 4 gives uniform $L^p(f_s dy)$ bounds, with some $p > 1$, for the coefficient integrands and their score-differentiated integrands. The probabilities of $|Y| < \delta$ and $|Y| > R$ can be made uniformly small as $s \downarrow 0$: $Y \to P$ in probability, P has no atom at zero, and moments control the far tail. On $\delta \leq |y| \leq R$ smooth convergence applies. Removing the two tails by uniform integrability proves $b_4 \to c$ and $sb_4' \to 0$. □

A.4 The large-mean posterior

Lemma 7 (Joint large-mean expansion)**.** *For $s = zt^{8/3}\log(1/t)^{2/3}$, uniformly for z in every fixed compact subinterval of $(0, \infty)$,*

$$A(t,s) = 1 - t^2 - s + 7t^4 + R_A(t,s), \qquad |R_A| + |s\partial_s R_A| = o(t^4). \tag{30}$$

Proof. Set $V_t = 1 + tW$ and $Y = V_t Z + \sqrt{s}N$. Its density is $p_{t,s}(y) = \mathbb{E}[\varphi_{V_t^2+s}(y)]$. Since $\mathbb{E}[ZY] = 1$ and $\mathbb{E}[Y^2] = 1 + t^2 + s$, projection gives the exact identity

$$A(t,s) = 1 - t^2 - s + \int \frac{w(y)^2}{p_{t,s}(y)}\, dy, \tag{31}$$
$$w(y) = p_{t,s}(y)\mathbb{E}[Z - Y \mid Y = y] = y\mathbb{E}\left[\left(\frac{V_t}{V_t^2 + s} - 1\right)\varphi_{V_t^2+s}(y)\right].$$

Retain the symmetric latent set $|W| \leq t^{-1/8}$ and observation set $|y| \leq t^{-1/32}$. There V_t stays uniformly away from zero, and $(|tW| + s)y^2 \to 0$. Taylor expansion of the smooth normal density and its prefactor, using odd-W cancellation, gives

$$w(y) = \{t^2(2y - y^3) - sy\}\varphi_1(y) + O_{\mathrm{wt}}(t^4 + t^2 s + s^2).$$

More explicitly, if $w_{\text{in}}, p_{\text{in}}$ denote the symmetric-W retained integrals and $e_w = w_{\text{in}} - \{t^2(2y - y^3) - sy\}\varphi_1(y)$, then for fixed m and C,

$$\begin{aligned} |e_w| + |s\partial_s e_w| &\le C(t^4 + t^2 s + s^2)\varphi_1(y)(1 + |y|^m), \\ |p_{\text{in}} - \varphi_1| + |s\partial_s(p_{\text{in}} - \varphi_1)| &\le C(t^2 + s)\varphi_1(y)(1 + |y|^m). \end{aligned} \tag{32}$$

The omitted mass of the truncated normal contributes an exponentially small term, absorbed by these bounds. These are the bounds denoted by O_{wt} above. This follows directly from Taylor's integral remainder: on the retained set all variance denominators are bounded below, derivatives are Gaussian densities times polynomials, and their exponential correction relative to φ_1 is uniformly bounded because $(|tW| + s)y^2 \to 0$. Integrating the required powers of W gives fixed Gaussian moments. The retained density divided by φ_1 tends uniformly to one and is at least $1/2$ for small t.

We justify the truncation before using the quotient. On the retained y interval, the omitted-W density and numerator, together with their logarithmic noise derivatives, are bounded by a power of s^{-1} and t^{-1} times $\exp(-ct^{-1/4})$. This follows from the supremum bounds on a Gaussian kernel of variance at least s and Gaussian tail moments of W. Division by the retained density costs at most $C\exp(Ct^{-1/16})$, still leaving a quantity smaller than every power of t on the stated scale.

For the remaining weighted integrals, conditional Cauchy–Schwarz bounds w^2/p by $p\mathbb{E}[(Z - Y)^2 \mid Y]$. In fact, with $R = Z - y$, $m = \mathbb{E}[R \mid Y = y]$, and $Q = (N^2 - 1)/2$, its exact score derivative is

$$s\partial_s(w^2/p) = p\{2m\mathbb{E}[RQ \mid Y = y] - m^2\mathbb{E}[Q \mid Y = y]\}.$$

Here $Z - y$ is independent of s at fixed y in the kernel representation. Lemma 4 bounds each term on a tail event by a positive power of its probability times complete-data higher moments. The omitted-W event has Gaussian small probability. On retained W, $|Y| > t^{-1/32}$ requires a Gaussian tail of Z or N, since V_t is bounded and $s \to 0$. All these errors, also after $s\partial_s$, are smaller than every power of t. The rare possibility $V_t \simeq 0$ is included in the omitted-W event, rather than discarded without control.

Squaring the retained expansion in (31) and integrating gives

$$\int \frac{w^2}{p_{t,s}} = 7t^4 + O(t^6 + t^2 s + s^2),$$

with the same remainder bound for the logarithmic noise derivative. The coefficient is $\mathbb{E}[(2Z - Z^3)^2] = 7$. Since $t^2 s + s^2 = o(t^4)$ uniformly on the stated scale, (30) follows. □

A.5 Combining the expansions

Put $D(t, s) = A(t, s) - B(t, s)/t^2$. Lemmas 6 and 7 give

$$D(t, s) = (7 - b_4(s))t^4 - s^2 I(s) + o(t^4)$$

uniformly on the joint scale, with the same $o(t^4)$ control after $s\partial_s$. In particular the other terms are $t^2 O(s)$ and $O(t^6)$, whose logarithmic derivatives have the same order. Since $L_s/\log(1/t) \to 4/3$ uniformly for compact positive z sets, Lemma 5 yields

$$\frac{s^2 I(s)}{t^4} \to \frac{3C}{4} z^{3/2}, \qquad \partial_z\left(\frac{s^2 I(s)}{t^4}\right) \to \frac{9C}{8}\sqrt{z}.$$

The identity $\partial_z = (s/z)\partial_s$ converts the logarithmic derivative bounds into uniform z derivative bounds. Together with $b_4 \to c$ and $sb_4' \to 0$, these prove (16).

B Covariance equivalence with all five curves different

We prove part 1 of Theorem 1. This construction is separate from the coefficient-dual pair, which leaves U and E unchanged.

B.1 A positive-noise boundary expansion

Fix a, g, E in compact subsets of $(0, \infty)$, x in a compact interval, and temporarily allow $0 \le U \le U_0$. The joint density is

$$p(z, y) = \varphi_g(z)\varphi_v(y - az), \qquad v = E + U(x + z)^2.$$

Every fixed-order parameter derivative is p times a finite sum of monomials in z, x, and $N = (y - az)/\sqrt{v}$, with factors $v^{-j/2}$ and bounded coefficients. Because $v \ge E_{\min} > 0$, these scores are bounded by $C(1 + |z| + |N|)^m$. A derivative of the integrated posterior variance or its projection residual is a finite sum of the observation density times products of conditional score moments, possibly multiplied by powers of y.

Conditional Hölder inequalities bound their absolute integrals by complete-data moments. The variables z and N are independent Gaussian variables with bounded moments and $y = az + \sqrt{v}N$ has uniformly bounded moments of every order. Hölder and high-moment Markov bounds therefore control the tails uniformly. On compact y intervals the positive density has a uniform lower bound. This proves joint continuity and differentiation of the integrated expressions at every fixed finite order, including one-sided U derivatives at zero. It avoids any assumed square-integrability of a density ratio relative to the $U = 0$ Gaussian.

Set $A = a^2 g$, $W_0 = A + E$, $k_0 = ag/W_0$, and $w_0 = gE/W_0$. At $U = 0$, $\eta \mid r$ is $N(k_0 r, w_0)$ and the posterior is linear. Differentiation in U uses the identity $\partial_v \varphi_v = \partial_y^2 \varphi_v / 2$. With $f_0 = \varphi_{W_0}$ and $B_0(y) = x^2 + 2xk_0 y + w_0 + k_0^2 y^2$,

$$\partial_U f\big|_0 = \tfrac{1}{2}\partial_y^2(B_0 f_0), \qquad \partial_U h\big|_0 = \tfrac{1}{2}\partial_y^2((k_0 y B_0 + 2w_0(x + k_0 y))f_0).$$

For $R_U(y) = m(y) - ag\, y/V$, direct differentiation gives

$$R_0'(y) = C_2(y^2 - W_0) + C_3(y^3 - 3W_0 y), \quad C_2 = \frac{xg(E - 2A)}{W_0^3}, \quad C_3 = \frac{ag^2(E - A)}{W_0^4}.$$

The two polynomials are orthogonal under $N(0, W_0)$, with squared norms $2W_0^2$ and $6W_0^3$. Since $R_0 = 0$, Taylor's integral remainder yields

$$P = U^2 \left\{ \frac{2x^2 g^2 (E - 2A)^2}{W_0^4} + \frac{6a^2 g^4 (E - A)^2}{W_0^5} \right\} + o(U^2). \tag{33}$$

The expansion holds with any prescribed finite number of derivatives in the other parameters, by the preceding smoothness argument.

B.2 Two exact covariance-compatible root functions

Choose

$$a_1 = 1/4, \quad c_1 = 4/3, \qquad a_2 = \sqrt{6/5}/5, \quad c_2 = 5/4.$$

For an auxiliary $\lambda \ge 0$, x near zero, and v near one, set

$$\begin{aligned} U_j &= \lambda/c_j, & E_j &= 1/c_j, & G_j(x, v, \lambda) &= \frac{c_j v - \lambda x^2 - 1}{c_j a_j^2 + \lambda}, \\ \rho_j &= c_j a_j + \lambda/a_j, & b_j &= c_j + 2, & D_j(x, v, \lambda) &= \rho_j^2 P(a_j, U_j, E_j, G_j, x). \end{aligned} \tag{34}$$

Substitution in the forward equations gives, identically for both models,

$$V_j = v, \qquad S_j = -2v - \lambda x^2 - 1. \tag{35}$$

For $\lambda > 0$, let $\mathcal{F} = (D_1 - D_2)/\lambda^2$. Both D_j and their first λ derivatives vanish at zero; Taylor's integral formula extends $\mathcal{F}$ smoothly to $\lambda = 0$ from the right. Write $A_j = v - 1/c_j$. Equation (33) gives

$$\mathcal{F}(x, v, 0) = \alpha(v)x^2 + \beta(v), \tag{36}$$

where α and β are the differences (index 1 minus index 2) of

$$\alpha_j(v) = \frac{2A_j^2(v - 3A_j)^2}{a_j^2 v^4}, \qquad \beta_j(v) = \frac{6A_j^4(v - 2A_j)^2}{a_j^4 v^5}.$$

Exact substitution gives $\beta_1(1) = \beta_2(1) = 3/2$, $\beta'(1) = -7$, and $\alpha(1) = -17/120$. Hence $\mathcal{F}(0, 1, 0) = 0$ and $\mathcal{F}_v(0, 1, 0) = -7$.

For clarity, a one-sided argument avoids extending to a negative variance. Choose a small closed v neighborhood about one on which $\mathcal{F}_v < 0$ for small $|x|$ and $\lambda \geq 0$. The opposite signs of $\mathcal{F}$ at its two v endpoints persist on a smaller (x, λ) rectangle. Strict monotonicity and the intermediate value theorem give a unique root $v = v_\lambda(x)$ there. Interior implicit differentiation in x and the joint smoothness just proved give its finite-order derivatives and their continuity as $\lambda \downarrow 0$. Reflection symmetry makes the root even in x.

For a fixed sufficiently small $\lambda > 0$, substitute this root into (34). Then $D_1 = D_2$ and (35) gives equal V, S, hence equal $T = S^2/V + D$ throughout a smaller compact interval. At $x = \lambda = 0$ the limiting values are

$$G_1 = 4, \quad G_2 = 25/6, \quad \rho_1 = 1/3, \quad \rho_2 = \sqrt{6/5}/4, \quad b_1 = 10/3, \quad b_2 = 13/4.$$

These are interior to all restrictions except $U = 0$. Taking small positive λ and shrinking the interval gives strictly admissible witnesses with positive U_j and uniformly positive bounded G_j.

At $x = 0$, the target limits as $\lambda \downarrow 0$ are

Target	Model 1	Model 2
a^2G	$1/4$	$1/5$
UG/λ	3	$10/3$
E	$3/4$	$4/5$
ρ^2G	$4/9$	$5/16$

The four nonzero gaps persist for small positive λ on a sufficiently small interval. Moreover $U_1 = 3\lambda/4$ and $U_2 = 4\lambda/5$, so $U_1 x^2 \neq U_2 x^2$ whenever $x \neq 0$. All five contribution curves differ. The limit varies the observed functions; it is not an extremizing sequence in a fixed identified set.

The witnesses can also have non-affine V. Direct differentiation in (36) gives

$$\beta''(1) = -241/3, \qquad \alpha'(1) = -13/30.$$

Writing $v_0(x) = 1 + px^2 + qx^4 + o(x^4)$ yields

$$p = -17/840, \qquad 2q = -32521/14817600 \neq 0.$$

Continuity of the fourth derivative preserves this nonzero coefficient for small positive λ. Thus $V = v_\lambda \notin \operatorname{span}\{1, x^2\}$.

Finally transfer the construction to any prescribed compact interval I containing zero in its interior. Choose $d \geq 1$ such that x/d lies inside the constructed interval for every $x \in I$, and transform each model by

$$a^* = a/d, \quad U^* = U/d^2, \quad E^* = E, \quad \rho^* = \rho/d, \quad b^* = b, \quad G^*(x) = d^2 G(x/d).$$

Couple $\eta^* = d\eta$ and $u^* = u/d$. The new action and outlook at x have exactly the old law at x/d, since $m_x^*(r) = d\, m_{x/d}(r)$. All restrictions persist. Each target becomes its old curve composed with x/d, preserving the required differences and non-affinity. This completes the proof.

References

[1] P. Andrade, F. Ferroni, and L. Melosi (2023). *Identification Using Higher-Order Moments Restrictions*. Federal Reserve Bank of Chicago Working Paper 2023-28, August 18. Working-paper page.

[2] M. D. Bauer and E. T. Swanson (2023). A reassessment of monetary policy surprises and high-frequency identification. *NBER Macroeconomics Annual*, 37. Author-hosted article.

[3] R. E. Gaunt (2025). *On the product of correlated normal random variables and the noncentral chi-square difference distribution*. arXiv:2408.04101v2, September 4. Versioned preprint.

[4] A. Guay (2020). *Identification of Structural Vector Autoregressions Through Higher Unconditional Moments*. Chaire en macroéconomie et prévisions Working Paper 2020-19, October. Inspected working-paper version. Later publication: *Journal of Econometrics* 225(1), 27–46 (2021), doi:10.1016/j.jeconom.2020.10.006.

[5] K. A. Sastry (2025). *Disagreement About Monetary Policy*. Author manuscript, March 31. Inspected author version.