# Does Calibration Pick the Right Model? Likelihood Evidence from Income Risk[1]

Jiuyi Cheng[a] Alexis Akira Toda[a,b]

[a]Emory University

[b]Kobe University

UC Irvine
November 4, 2026

[1]Link to paper: https://arxiv.org/abs/2611.02029

# Calibration in quantitative macroeconomics

- Quantitative macroeconomics typically chooses parameters to match selected moments (Kydland and Prescott, 1982, 1996)
- Calibrated model then delivers answers on behavior, welfare, and policy
- Known concern (Hansen and Heckman, 1996; Browning et al., 1999):
  - target fit is evidence only about targets
  - other models or parameters may fit equally well yet answer differently
- How much this matters in practice is empirical question: answering it requires criterion that uses all data

# Likelihood as empirical criterion

- Likelihood evaluates joint distribution of complete observed histories
- Models can match selected statistics yet assign low probability to observed histories
- Latent states make likelihood evaluation costly:
  - we integrate over unobserved histories
  - nonlinear or non-Gaussian models generally lack Kalman recursion
- Particle filters approximate these integrals by simulation (Fernández-Villaverde and Rubio-Ramírez, 2007), but simulated likelihood is noisy and not smooth in parameters
- Hence likelihood comparison remains rare in quantitative work

# Questions

1. Can we evaluate likelihood of rich latent-state models on short, irregular panels without simulation?

2. Does likelihood pick among models that calibration cannot distinguish?

3. Does choice of criterion matter for economic answers?

# Main results

1. Likelihood estimation is feasible
   - approximate latent process by hidden Markov model (HMM); no simulation
   - allow non-Gaussian risk, permanent heterogeneity, and irregular short panels
2. Likelihood separates models that calibration cannot
   - four earnings processes can share same covariance function, yet likelihood ranks them
   - calibrations to 70 moments fit earnings histories worse than even likelihood-estimated Gaussian AR(1) persistent dynamics
3. Calibrations with similar fit give wildly different answers
   - optimal saving with earnings dynamics calibrated to 70 moments
   - calibrations with similar target fit: mean wealth from 2.2 to 8.2 times mean earnings

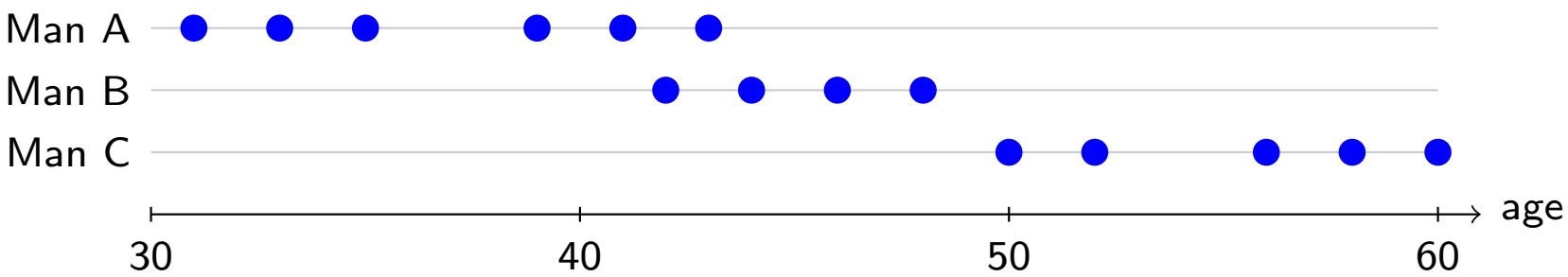
# Related literature

- Empirical foundations of calibration: Hansen and Heckman (1996), Browning et al. (1999), Watson (1993), and Hnatkovska et al. (2012)
- HMM approximations and random effects: Kitagawa (1987), Altman (2007), Langrock et al. (2012), Farmer (2021), and Farmer and Toda (2017)
- Misspecification and model comparison: White (1982) and Vuong (1989)
- Earnings risk and choice of targets: Guvenen et al. (2021, 2026) and Crawley et al. (2026)

# Illustration: PSID earnings data

- Panel Study of Income Dynamics (PSID), SRC core sample
- 28,422 observations of 4,683 men aged 30–60
- Biennial interviews, 1997–2023; annual earnings, 1996–2022
- Remove deterministic effects of age, education, and calendar year
- Require earnings above minimum threshold (Guvenen et al., 2021) and at least two observations
- Unequal entry ages and missing observations remain in likelihood

# Short, irregular earnings histories




Illustrative observation schedules: biennial interviews, unequal entry ages, gaps

- On average 6.1 observations per man; 22,619 two-year pairs
- Decompose earnings residual:

$$u_{ij} = \underbrace{\alpha_i}_{\text{permanent}} + \underbrace{X_i(t_{ij})}_{\text{persistent}} + \underbrace{\epsilon_{ij}}_{\text{transitory}}$$

- Only $u_{ij}$ observed: likelihood integrates over $\alpha_i$ and path of $X_i$

# Gaussian AR(1) benchmark

$$u_{ij} = \alpha_i + X_{i,t_{ij}} + \epsilon_{ij}, \qquad X_{i,t+1} = \rho X_{it} + \eta_{i,t+1},$$
$$\eta_{it} \sim \mathcal{N}(0, \sigma_\eta^2), \quad \epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2), \quad X_{i0} \sim \mathcal{N}(0, \sigma_0^2) \text{ at age 30}$$

| Permanent effect $\alpha_i$ | (1) None | (2) Unrestricted | (3) Gaussian | (4) Gaussian | (5) Normal–Laplace |
|---|---|---|---|---|---|
| Initial dispersion | Free | Free | Stationary | Free | Free |
| $\rho$ | 0.962 | 0.947 | 0.900 | 0.946 | 0.923 |
| $\sigma_\eta$ | 0.219 | 0.240 | 0.246 | 0.228 | 0.237 |
| $\sigma_0$ | 0.531 | 0.000[b] | 0.566 | 0.417 | 0.213 |
| $\sigma_G$ | | | 0.399 | 0.323 | 0.000[b] |
| $\sigma_\epsilon$ | 0.327 | 0.310 | 0.304 | 0.320 | 0.314 |
| $a,\ b$ | | | | | 3.81, 2.79 |
| LL | −22,694 | | −22,793 | −22,690 | −22,595 |

- Same as Ornstein-Uhlenbeck process at annual dates (later: OU); Kalman filter gives exact likelihood with actual gaps

# Beyond linear-Gaussian models

- Gaussian AR(1): Kalman filter evaluates likelihood exactly and quickly
- But data show features this benchmark cannot produce:
  - heavy-tailed earnings changes ▸ Changes
  - volatility that depends on current earnings ▸ Conditional moments
  - losses that reverse more often than gains ▸ Reversals
- Models with these features are nonlinear or non-Gaussian: Kalman filter no longer exact

## Need

General likelihood method for nonlinear, non-Gaussian latent processes on short, irregular panels

# Short panels and latent states

- Individual $i$ has observations $Y_{ij}$ at dates $t_{ij}$, $j = 0, \ldots, J_i$ (PSID: $Y_{ij} = u_{ij}$ at age $t_{ij}$)
- Let $\Delta_{ij} = t_{ij} - t_{i,j-1}$ denote observation gap
- Latent Markov state $X_i(t)$ has parameters $\theta$ and transition kernel $K_{\theta,\Delta}$
- Conditional observation density:

$$Y_{ij} \mid X_i(t_{ij}) = x \quad \sim \quad g_{\theta,ij}(\cdot \mid x)$$

- Many individuals, short histories; initialization remains important
- We condition on covariates and observation dates

▸ Observation schedules

# Likelihood of observed history

- We integrate over unobserved state at every date:

$$L_i(\theta) = \int \nu_{\theta,t_{i0}}(\mathrm{d}x_0)\, g_{\theta,i0}(y_{i0} \mid x_0) \times \prod_{j=1}^{J_i} K_{\theta,\Delta_{ij}}(x_{j-1}, \mathrm{d}x_j) g_{\theta,ij}(y_{ij} \mid x_j)$$

- $\nu_{\theta,t_{i0}}$: state distribution at first observation
- Nonlinear dynamics and non-Gaussian shocks generally prevent closed-form recursion
- Finite-state approximation reuses filtered probabilities

# Finite-state approximation

- Approximate latent state on grid $\mathcal{X}_N = \{x_1, \ldots, x_N\}$
- Keep specified continuous observation density
- Let $P_{\theta,N}(\Delta)$ be transition matrix over gap $\Delta$
  - discrete time: conditional moment matching (Farmer and Toda, 2017)
  - diffusion: approximate its generator $Q_\theta$, then use $\exp(\Delta Q_\theta)$
- Define diagonal observation matrix

$$G_{ij}(\theta) = \operatorname{diag}\{g_{\theta,ij}(y_{ij} \mid x_n)\}_{n=1}^{N}$$

- Actual observation gaps enter through $P_{\theta,N}(\Delta_{ij})$

# Forward recursion

Let $\xi_{i,j-1}$ denote filtered state probabilities after observation $j-1$.

Predict: $$q_{ij} = P_{\theta,N}(\Delta_{ij})^\top \xi_{i,j-1},$$

Update: $$c_{ij} = \mathbf{1}^\top G_{ij}(\theta) q_{ij}, \qquad \xi_{ij} = \frac{G_{ij}(\theta) q_{ij}}{c_{ij}},$$

Log likelihood: $$\ell_{i,N}(\theta) = \sum_{j=0}^{J_i} \log c_{ij}.$$

- Initialize $q_{i0}$ at individual's entry date
- $c_{ij}$: conditional observation density
- Normalization avoids multiplying many small probabilities
- Dense-matrix cost is $O(J_i N^2)$ per history

# Permanent effects and observation gaps

- Let $\alpha_i$ denote permanent individual effect with common density $f_{\alpha,\theta}$
- Run filter conditional on $\alpha_i = \alpha$, then integrate with $K_\alpha$ quadrature nodes:

$$L_i(\theta) = \int L_i(\theta \mid \alpha) f_{\alpha,\theta}(\alpha)\, \mathrm{d}\alpha$$
$$\approx \sum_{k=1}^{K_\alpha} w_{ik}(\theta) L_{i,N}(\theta \mid \alpha_{ik})$$

- Estimate common heterogeneity distribution jointly with dynamics
- Propagate initial law to each individual's entry age
- Predict across missing years using actual gaps
- Likelihood evaluation requires no simulation

Earnings specification

# Likelihood under misspecification

- Maximize average individual log likelihood:

$$\widehat{\mathcal{Q}}_{I,N}(\theta) = \frac{1}{I}\sum_{i=1}^{I} \ell_{i,N}(\theta), \qquad \widehat{\theta}_{I,N} \in \arg\max_{\theta} \widehat{\mathcal{Q}}_{I,N}(\theta)$$

- Population target maximizes expected log likelihood
- Equivalently, it minimizes Kullback–Leibler (KL) divergence: measure of distributional discrepancy
- This interpretation does not require correctly specified candidate model
- Comparison uses complete histories, including levels and dependence across dates

# Inference on short panels

- Sampling unit: individual's complete history
- Fixed grid: standard $M$-estimation gives

$$\sqrt{I}(\widehat{\theta}_{I,N} - \theta_N) \xrightarrow{\mathrm{d}} \mathcal{N}(0, A_N^{-1} B_N A_N^{-1})$$

  under identification and regularity conditions (White, 1982)
- $\theta_N$ maximizes discretized population objective; $A_N, B_N$ measure curvature and score variation
- Inference for continuous-state target also requires sufficiently accurate numerical approximation
- In particular, expected score error must be $o(I^{-1/2})$

▸ Conditions and sandwich covariance

# Comparing misspecified models

- For fitted models $A$ and $B$, compute each individual's log likelihood difference:

$$d_i = \ell_i^A(\widehat{\theta}^A) - \ell_i^B(\widehat{\theta}^B)$$

- Compare their expected fit using

$$z = \frac{\sqrt{I}\,\bar{d}}{\widehat{\sigma}_d}$$

- Under relevant regularity and distinguishability conditions, equal expected fit gives standard normal limit (Vuong, 1989)
- Positive values favor $A$; negative values favor $B$
- Numerical ranking need not imply clear statistical separation

# Numerical accuracy

Gaussian Ornstein–Uhlenbeck (OU) benchmark: exact integrated likelihood.

| Approximation | LL error | Implied bias / SE |
|---|---|---|
| Rouwenhorst | 89.18 | 5.10 |
| Farmer–Toda | $< 10^{-9}$ | $< 0.001$ |
| Generator | 6.70 | 0.32 |

- $N = 129$ on tested grid; evaluate at exact estimate
- Bias/SE: largest first-order parameter bias, measured in standard errors
- We check grid spacing, range, quadrature, and score accuracy

▸ Refinement ▸ Empirical estimates

# Candidate earnings processes

$$u_{ij} = \alpha_i + X_i(t_{ij}) + \epsilon_{ij}$$

| Process | Persistent risk | Motivating fact |
|---|---|---|
| OU | Gaussian AR(1) | benchmark (5) |
| GKOS | AR(1) with Gaussian-mixture innovations | heavy tails, skewness |
| GMR | Guvenen–McKay–Ryan: losses and scarring | asymmetric reversals |
| PLV | Piecewise linear variance diffusion | state-dependent variance |

- Common permanent heterogeneity
- Gaussian or Gaussian-mixture independent transitory shocks give $4 \times 2 = 8$ specifications

▸ Mixtures ▸ GKOS ▸ GMR

# Piecewise linear variance

$$dX_i(t) = -\kappa X_i(t)\,\mathrm{d}t + \sqrt{v(X_i(t))}\,\mathrm{d}B_i(t),$$
$$v(x) = \sigma^2 + c_+ x^+ + c_- x^-, \qquad x^\pm = \max\{\pm x, 0\}.$$

- $\sigma > 0$, $c_+, c_- \geq 0$; $B_i$ is Brownian motion
- $c_+ = c_- = 0$ gives OU
- Variance rises away from zero, potentially asymmetrically
- Linear conditional mean reversion:

$$\mathrm{E}[X_i(t+h) \mid X_i(t) = x] = \mathrm{e}^{-\kappa h} x$$

▸ Stationary law ▸ Change symmetry

# Covariance equivalence

## Proposition

*Each stationary GKOS, GMR, or PLV specification has stationary OU counterpart with same persistence, permanent-effect variance, and observed earnings autocovariances at every integer lag:*

$$\text{Cov}(u_t, u_{t+h}) = V_\alpha + V\rho^h + T\mathbf{1}\{h = 0\}$$

- $V_\alpha$ is permanent-effect variance; $V$, $T$ are effective persistent and transitory variances
- Common covariances leave conditional earnings risk undetermined
- Likelihood estimates that follow are unrestricted

▸ Matching processes

# Likelihood comparison

| | Independent transitory shocks | |
|---|---|---|
| Persistent process | Gaussian | Gaussian mixture |
| OU | −22,595 | −13,694 |
| GKOS | −14,808 | −12,850 |
| GMR | −14,610 | −13,038 |
| PLV | −14,290 | −12,765 |

- Mixture transitory shocks improve every model by over 1,500 log likelihood points
- OU with mixture shocks fits better than any process with Gaussian transitory shocks
- Economic comparisons use mixture transitory shocks

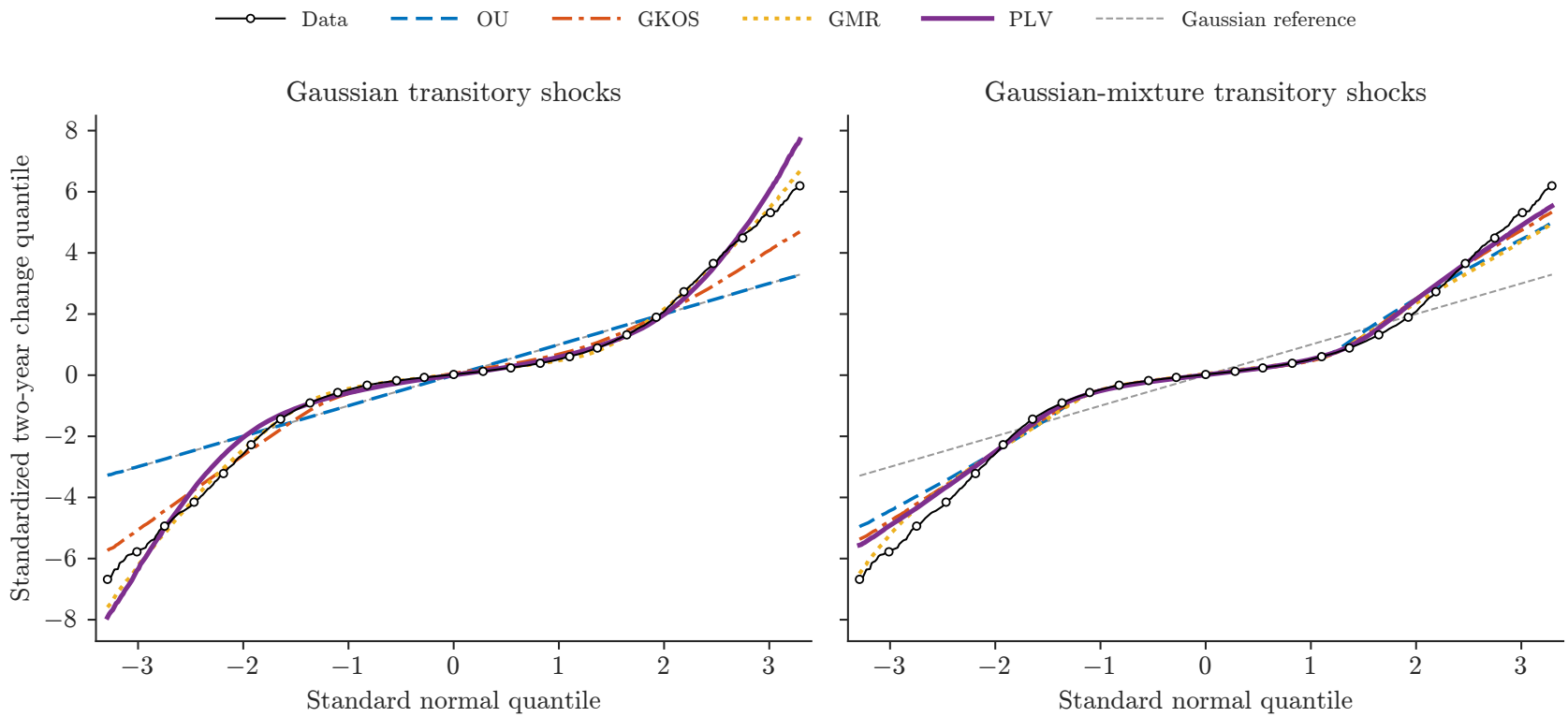
▸ Selected parameters

# Statistical separation

| Comparison | LL difference | Vuong $z$ |
|---|---|---|
| PLV vs GKOS | 85.4 | 1.56 |
| PLV vs GMR | 273.3 | 4.43 |
| GKOS vs GMR | 187.8 | 4.28 |

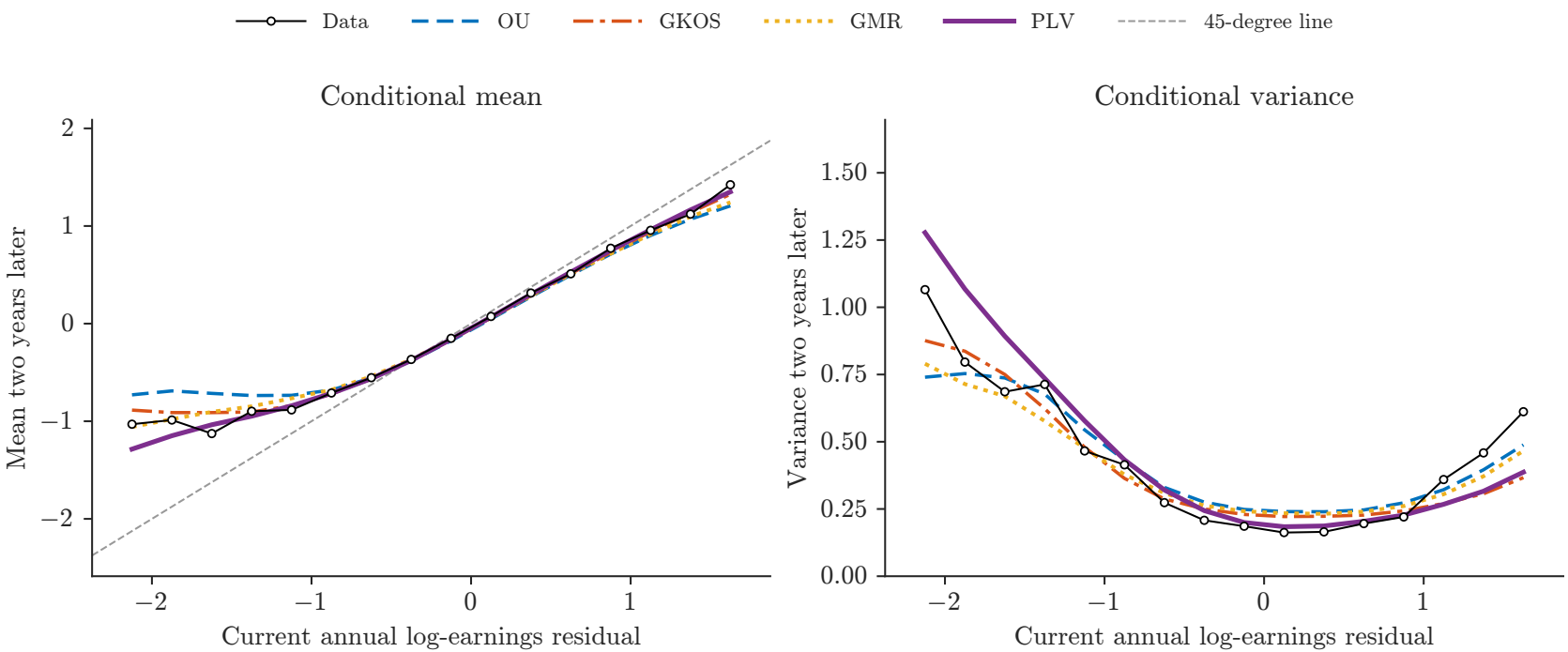
- Gaussian-mixture independent transitory shocks throughout
- PLV and GKOS fit significantly better than GMR
- PLV has highest likelihood, but its advantage over GKOS is not decisive: one-sided $p = 0.060$
- OU is nested in other models and is omitted from these tests

# Two-year earnings changes




- Mixture transitory shocks improve tail fit, but every model still misses most extreme standardized changes

# Conditional moments with mixture shocks




- U-shaped observed variance profile also arises under OU with mixture transitory shocks

▸ Earnings levels ▸ Reversals

# Saving problem

$$\begin{aligned} &\text{maximize} && \mathsf{E} \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \\ &\text{subject to} && a_{t+1} = R a_t + y_t + b - c_t \geq 0 \end{aligned}$$

- $\beta = 0.96$, $\gamma = 2$, $R = 1.02$
- Common mean raw earnings $\mathsf{E}\, y_t = 1$ and transfer $b = 0.1$
- Households observe persistent state and current income
- Use stationary earnings laws and set permanent effects to zero

▸ Welfare measure

# Saving under fitted and matched processes

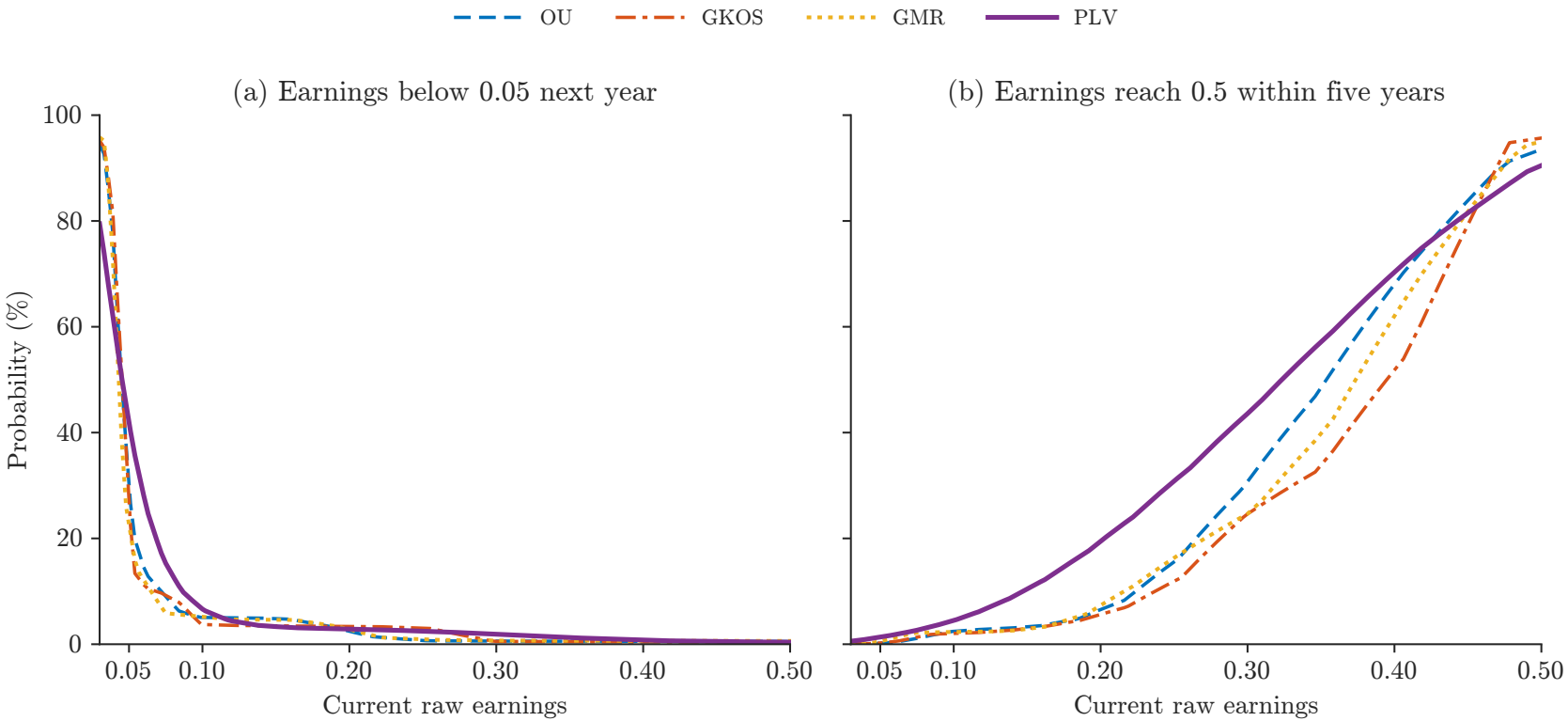
Stationary mean assets / mean raw earnings

| | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| Likelihood estimates | 4.805 | 7.374 | 5.592 | 10.508 |
| Common covariances | 4.805 | 4.636 | 4.313 | 4.193 |

- Match other processes to fitted OU persistence and persistent/transitory variances
- Keep preferences, prices, mean earnings, and transfers fixed
- PLV moves from highest to lowest mean assets; range narrows from 4.8–10.5 to 4.2–4.8

▸ Transformations ▸ Saving policies

# Risk under common covariances




- Common stationary covariances leave different probabilities of large losses and recovery

# Calibration targets

| Targets | Number |
|---|---|
| Two- and six-year change moments by age / recent earnings | 36 |
| Mean subsequent changes after extreme changes | 8 |
| Change percentiles and tail probabilities | 8 |
| Level variances and autocovariances by age | 12 |
| Level percentiles by age | 6 |
| Total | 70 |

- Adapt targets from GKOS and GMR to panel
- Calibrate all parameters, including permanent effects and initial dispersion
- Fix targets and weights before search; minimize squared relative discrepancies with equal weight across eight sets

▸ Eight sets ▸ Search ▸ Restrictions

# Calibration and history fit

| | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| Moment objective, MLE | 0.187 | 0.114 | 0.148 | 0.149 |
| Moment objective, calibrated | 0.0406 | 0.0206 | 0.0197 | 0.0276 |
| LL, MLE | −13,694 | −12,850 | −13,038 | −12,765 |
| LL, calibrated | −15,104 | −14,209 | −13,877 | −13,956 |

- Calibration lowers moment objective by 78–87%
- Calibrations closely match percentiles of changes and levels, but none matches all targets within sampling error
- Every calibration fits histories worse than likelihood-estimated OU process with mixture transitory shocks

# Saving and welfare

| | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| *Mean assets / mean raw earnings* | | | | |
| Likelihood estimates | 4.80 | 7.37 | 5.59 | 10.51 |
| Calibrated | 1.63 | 4.69 | 4.98 | 3.76 |
| *Consumption equivalent (CE)* | | | | |
| Likelihood estimates | 0.793 | 0.746 | 0.802 | 0.739 |
| Calibrated | 1.029 | 0.893 | 0.854 | 0.677 |

- Every calibration implies less wealth than its likelihood estimate
- Welfare rises for OU, GKOS, and GMR and falls for PLV

# Variation within one model

| Process | Mean assets | CE |
|---|---|---|
| OU | 1.4–3.6 | 0.92–1.04 |
| GKOS | 2.2–8.2 | 0.70–1.00 |
| GMR | 2.9–7.2 | 0.66–0.97 |
| PLV | 3.1–9.3 | 0.67–0.77 |

- Retain evaluated candidates with objectives less than 0.0092 above reported value
- Within each model, mean assets vary by factor of 2.5–3.7
- Ranges are descriptive: neither confidence intervals nor exhaustive bounds

# Choice of targets

Stationary mean assets / mean raw earnings

| | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| Likelihood estimates | 4.80 | 7.37 | 5.59 | 10.51 |
| 70 changes/levels targets | 1.63 | 4.69 | 4.98 | 3.76 |
| 18 changes-only targets | 4.31 | 18.05 | 13.25 | 5.30 |

- Changing targets can reverse comparison with likelihood estimation
- Changes-only targets eliminate permanent effects; that exercise holds permanent-effect parameters and initial dispersion at their likelihood estimates

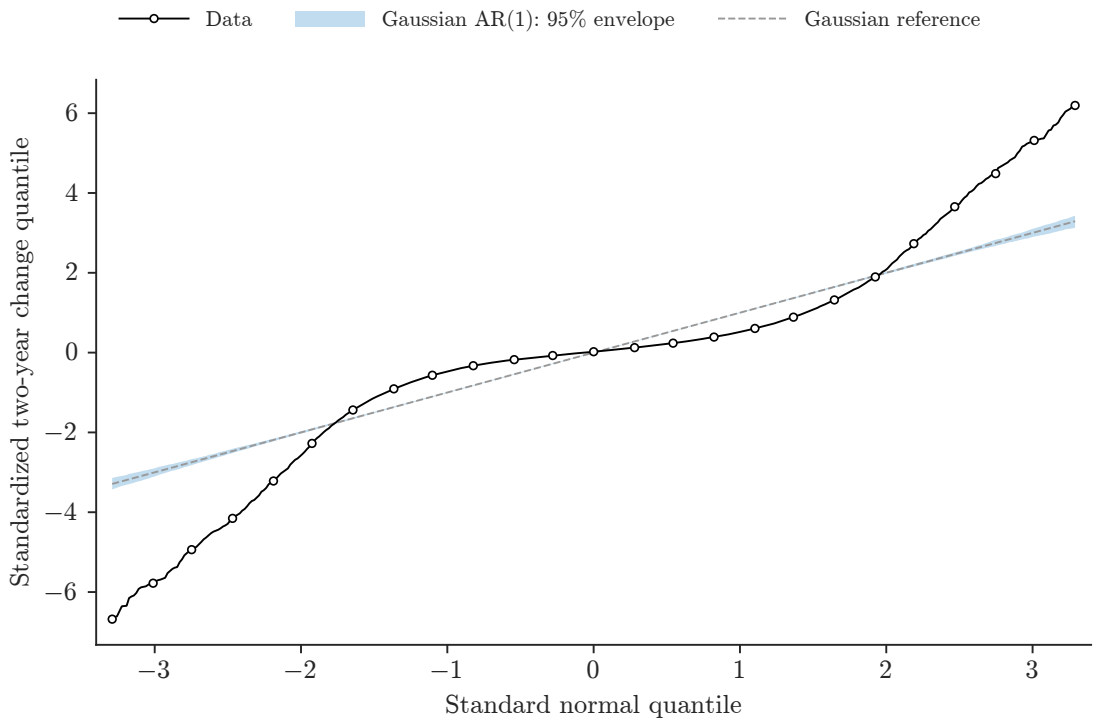
▸ Changes-only fit

# Conclusion

- HMM approximations make likelihood estimation feasible for rich latent processes on short, irregular panels
- Likelihood comparison provides common empirical discipline before counterfactual analysis
- In earnings illustration, calibrations improve target fit but assign lower likelihood to observed histories
- Model selection and parameter choice materially change saving and welfare
- Every candidate remains misspecified: targeted moments help diagnose what likelihood-selected model misses

▸ Theory ▸ Accuracy ▸ Calibration

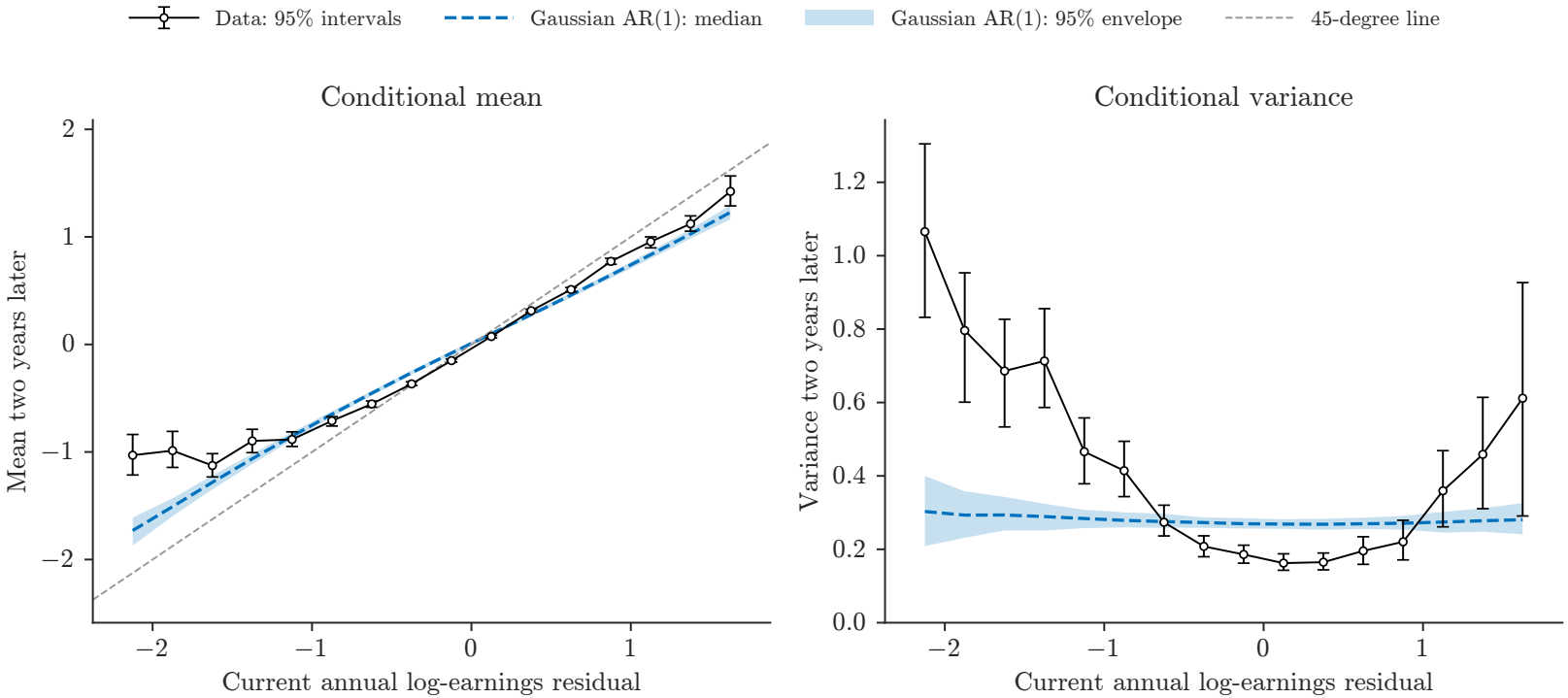
# Heavy-tailed earnings changes




- Both tails far heavier than Gaussian: excess kurtosis 10.82 in data, 0.00 under Gaussian AR(1)
- Skewness −0.50: large losses more common than large gains

◂ Beyond linear-Gaussian ◂ Models

# Conditional moments under Gaussian AR(1)




- Observed conditional variance is U-shaped, rising sharply at low earnings
- Gaussian AR(1) predicts flat profile and too little recovery from very low earnings
- AR(1) state variance does not depend on current state: motivates state-dependent variance

# Asymmetric reversals

- Large loss (gain): earnings fall to at most half (at least double) over two years
- Reversal: next two-year change has opposite sign

| Percent | Data | AR(1) |
|---|---|---|
| Opposite sign after large loss | 78.6 | 73.9 |
| Opposite sign after large gain | 59.6 | 74.1 |
| Loss–gain gap (pp) | 19.0 | 0.0 |
| At least half retraced after loss | 53.5 | 39.4 |
| At least half retraced after gain | 25.1 | 39.4 |

- Gaussian AR(1) implies symmetric reversals; in data, losses reverse more often and more fully
- Data gap has bootstrap 95% interval $[14.5, 23.7]$

◂ Beyond linear-Gaussian ◂ Models

# Observation schedules and sample

- Let $D_i$ collect covariates, observation counts, and observation dates
- We condition likelihood on $D_i$
- Conditional on included covariates, schedule is independent of latent process and measurement shocks
- Informative dropout or outcome-dependent observation requires observation model or ignorability argument
- Earnings illustration selects usable positive earnings; it does not model employment
- Empirical comparisons treat individuals as independent and first-stage earnings adjustment as fixed

◂ Short-panel model

# Centered Gaussian mixtures

$$\mathcal{M}(p, \delta, \sigma_1, \sigma_2) := p\,\mathcal{N}(-(1-p)\delta, \sigma_1^2) + (1-p)\,\mathcal{N}(p\delta, \sigma_2^2)$$

- Mixture has mean zero
- $p$ controls component weights; $\delta$ separates component means
- Standard deviations $\sigma_1, \sigma_2$ allow different dispersion
- Mixture adds three parameters relative to Gaussian noise
- We use mixtures for independent transitory shocks in every process and for persistent innovations in GKOS

◂ Models

# GKOS specification

$$X_{it} = \rho X_{i,t-1} + \eta_{it}, \qquad \eta_{it} \sim \mathcal{M}(p_\eta, \delta_\eta, \sigma_{\eta,1}, \sigma_{\eta,2})$$

- Adapt constant-mixture specification, Model 3, of Guvenen et al. (2021)
- Introduce mean reversion so process admits stationary law
- Mixture weights and component distributions are constant across age and earnings
- Allow skewness and excess kurtosis while preserving linear conditional mean
- Gaussian persistent innovations recover annual OU transition

◂ Models

# Losses and scarring in GMR

$$X_{it} = \rho X_{i,t-1} + \eta_{it} + \psi c_{it},$$
$$u_{ij} = \alpha_i + X_i(t_{ij}) + (1-\psi) c_{i,t_{ij}} + \epsilon_{ij},$$
$$\zeta_{it} = \begin{cases} 0 & \text{with probability } p_\zeta, \\ \log(1-\ell_{it}) & \text{with probability } 1-p_\zeta, \end{cases} \qquad c_{it} = \zeta_{it} - \mathsf{E}\,\zeta_{it}$$

- $\ell_{it}$: exponential loss size, conditional on $0 < \ell_{it} < 1$
- $\eta_{it}$ is Gaussian and $0 < \psi < 1$
- Loss affects earnings fully on impact and by $\psi\rho^k c_{it}$ after $k \geq 1$ years
- Adapt Model 2 of Guvenen et al. (2026), adding mean reversion and common independent transitory shocks

◂ Models ▸ Likelihood kernel

# Shared-loss likelihood kernel

- Same loss moves current earnings and persistent state
- Define $H_{nk}(u, \alpha)$ as joint transition and observation density on grid:

$$H_{nk}(u, \alpha)\,\mathrm{d}u = \Pr\{X_t = x_k,\ u_t \in \mathrm{d}u \mid X_{t-1} = x_n, \alpha\}$$

- Filtered row $r^\top$ propagates over integer gap $\Delta \geq 1$ as

$$r^\top P^{\Delta-1} H(u, \alpha)$$

- Normalize by row sum to obtain observation density and updated probabilities
- Integrating shared loss jointly preserves its dependence across two equations

◂ GMR ◂ Forward recursion

# Permanent effects and initialization

- Every specification uses centered asymmetric Laplace permanent effect
  - tail rates $a, b > 0$
  - variance $a^{-2} + b^{-2}$
- At age 30:

$$X_i(0) \sim \mathcal{N}(0, \sigma_0^2)$$

- Estimate $a, b, \sigma_0$ jointly with all dynamic and transitory parameters
- Propagate initial law to actual entry age; stationarity is not imposed at age 30
- Integrate each individual's whole-history likelihood over same permanent-effect prior

◂ Permanent heterogeneity

# PLV stationary law

For $c_s > 0$, let

$$\lambda_s = \frac{2\kappa}{c_s}, \qquad r_s = \frac{2\kappa\sigma^2}{c_s^2}.$$

Stationary density:

$$\pi(x) \propto \left(1 + \frac{c_s|x|}{\sigma^2}\right)^{r_s - 1} \mathrm{e}^{-\lambda_s|x|}, \quad s = \begin{cases} +, & x \geq 0, \\ -, & x < 0. \end{cases}$$

- Use Gaussian expression on side with $c_s = 0$
- Stationary mean and mode are zero; all polynomial moments are finite
- Exponentiating $X$ gives Pareto-type earnings tails
- For $c_+ > 0$, $\mathrm{E}_\pi \, \mathrm{e}^X < \infty$ if and only if $\lambda_+ > 1$

PLV

# Stationary change symmetry

- Stationary one-dimensional diffusion is reversible
- Consequently, its marginal changes satisfy

$$X(t+h) - X(t) \overset{d}{=} X(t) - X(t+h)$$

- Stationary PLV changes are symmetric even if $c_+ \neq c_-$ and level distribution is skewed
- Nonstationary initialization can produce asymmetric changes at finite ages
- Symmetric marginal changes do not imply equal reversals after large gains and losses

◂ PLV ◂ Reversals

# Covariance matching

- Match fitted OU persistence $\rho$, persistent variance $V$, and transitory variance $T$
- GKOS: rescale persistent innovations and independent transitory shocks
- PLV: rescale state and time,

$$X^*(t) = s_X \widehat{X}(a_\kappa t), \quad s_X = \sqrt{V/\widehat{V}_X}, \quad a_\kappa = \frac{-\log \rho}{\widehat{\kappa}},$$

giving

$$\kappa^* = -\log \rho, \quad \sigma^* = s_X \sqrt{a_\kappa}\, \widehat{\sigma}, \quad c_\pm^* = a_\kappa s_X \widehat{c}_\pm$$

- GMR: match its effective persistent and transitory covariance components while retaining scarring and standardized loss shape

◂ Covariance equivalence ◂ Saving

# Fixed and refining grids

$$s_{i,N}(\theta) = \nabla \ell_{i,N}(\theta), \qquad H_{i,N}(\theta) = \nabla^2 \ell_{i,N}(\theta),$$
$$A_N = -\,\mathrm{E}\, H_{i,N}(\theta_N), \qquad B_N = \mathrm{E}[s_{i,N}(\theta_N) s_{i,N}(\theta_N)^\top].$$

- Fixed grid: unique population maximizer, uniform laws of large numbers, smoothness, interior target, and nonsingular curvature
- Estimate sandwich covariance from individual scores and Hessians
- Refining grid: uniform likelihood approximation gives consistency for continuous-state target
- Continuous-state normality also requires score/Hessian approximation and $\sqrt{I}\,\mathrm{E}[s_{i,N(I)}(\theta_0) - s_i(\theta_0)] \to 0$
- Numerical error must be small relative to sampling error

# Selected likelihood estimates

| | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| $\rho$ | 0.979 | 0.973 | 0.969 | 0.962 |
| $\sigma_0$ | 0.220 | 0.256 | 0.263 | 0.347 |
| $a$ | 3.237 | 3.596 | 3.130 | 2.981 |
| $b$ | 3.315 | 3.880 | 4.349 | 3.325 |
| $p_\epsilon$ | 0.184 | 0.127 | 0.153 | 0.108 |
| $\delta_\epsilon$ | 0.470 | 0.558 | 0.367 | 0.537 |
| $\sigma_{\epsilon,1}$ | 0.908 | 0.934 | 0.850 | 0.981 |
| $\sigma_{\epsilon,2}$ | 0.025 | 0.069 | 0.029 | 0.025 |
| Free parameters | 9 | 12 | 12 | 11 |

◂ Likelihood comparison

# OU approximation and refinement

Largest implied first-order parameter bias / reference SE

| $N$ | Rouwenhorst | Farmer–Toda | Generator |
|---|---|---|---|
| 33 | 16.42 | 0.003 | 9.38 |
| 65 | 9.30 | 0.000 | 0.85 |
| 129 | 5.10 | 0.000 | 0.32 |
| 257 | 2.71 | 0.000 | 0.16 |
| 513 | 1.41 | 0.000 | 0.08 |

- Evaluate at exact Gaussian OU estimate
- Same Rouwenhorst filtering grid for each method; generator uses four intervals per filtering interval
- Accuracy depends on construction and number of states

Accuracy

# Numerical checks at fitted parameters

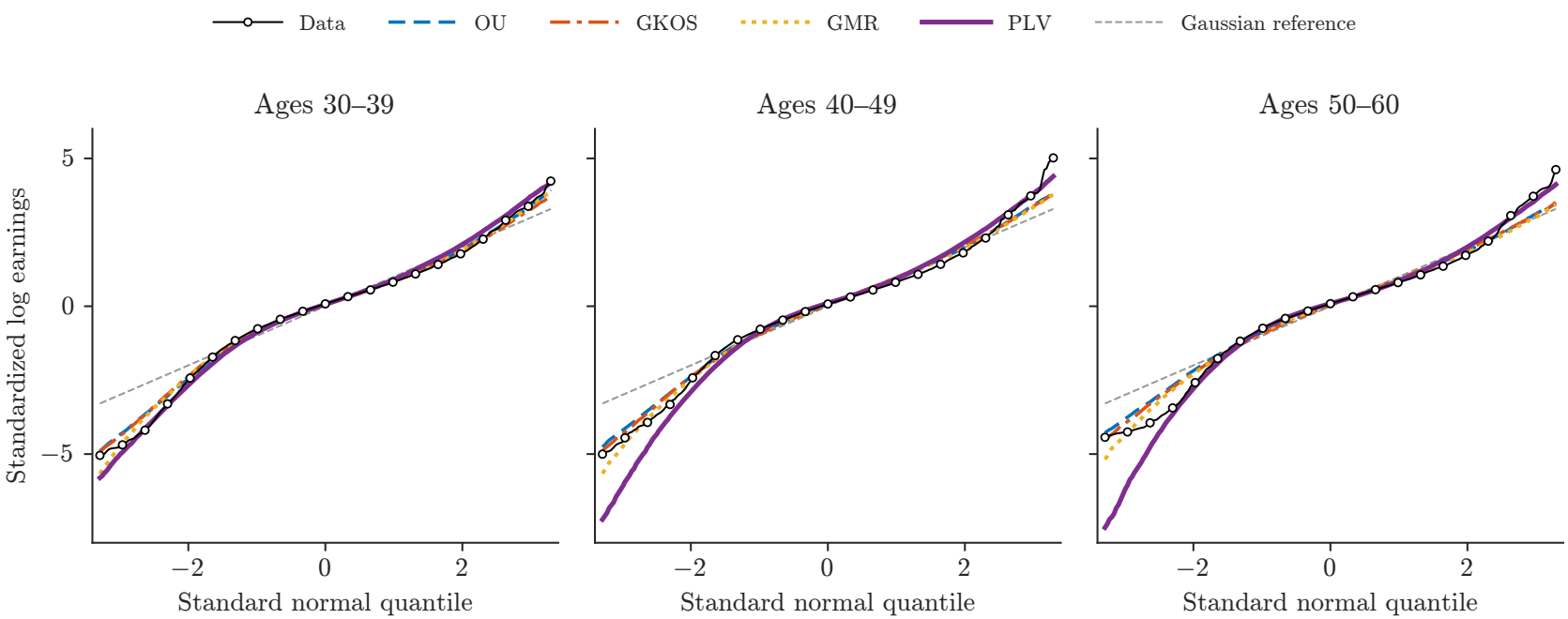
Changes in log likelihood relative to reported value

| Refinement | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| $K_\alpha = 256$ | 0.00 | 0.00 | 0.00 | 0.08 |
| $K_\alpha = 512$ | 0.00 | 0.00 | 0.00 | 0.10 |
| Grid twice as fine | 0.00 | 0.01 | 0.40 | 0.30 |
| Range twice as wide | 0.00 | −0.75 | −1.18 | −0.02 |
| Generator twice as fine | — | — | — | −0.23 |

- Gaussian-mixture independent transitory shocks
- Re-evaluate likelihoods at fixed fitted parameters
- Reported baseline uses $K_\alpha = 128$ permanent-effect nodes
- These checks are small relative to likelihood differences among models

◂ Accuracy ◂ Model comparison

# Earnings levels with mixture shocks




- PLV's lower tail remains too heavy after age 40
- Every model understates most extreme upper quantiles after age 40

◂ Fit diagnostics

# Changes and asymmetric reversals

| | Data | OU | GKOS | GMR | PLV |
|---|---|---|---|---|---|
| Change skewness | −0.499 | −0.005 | −0.038 | −0.555 | −0.048 |
| Excess kurtosis | 10.82 | 5.05 | 6.36 | 7.67 | 7.08 |
| *Opposite-sign change over next two years (%)* | | | | | |
| After loss (bottom 10%) | 78.63 | 86.57 | 84.09 | 86.16 | 81.11 |
| After gain (top 10%) | 59.62 | 62.88 | 60.52 | 60.37 | 59.03 |
| Gap (pp) | 19.01 | 23.52 | 23.55 | 25.75 | 21.88 |

- Gaussian-mixture independent transitory shocks
- GMR comes closest to change skewness and kurtosis but ranks third by likelihood

◂ Fit diagnostics ▸ Symmetry

# Fit across eight target sets

| Target set | Count | OU | GKOS | GMR | PLV |
|---|---|---|---|---|---|
| 2y change SD/skew/kurt by age×RE | 27 | .284 | .247 | .259 | .262 |
| 6y change SD/skew/kurt by RE | 9 | .326 | .207 | .198 | .182 |
| 2y/4y mean reversals | 4 | .230 | .167 | .144 | .127 |
| 6y/10y mean reversals | 4 | .157 | .123 | .090 | .094 |
| Change P10/P90 | 6 | .059 | .066 | .057 | .171 |
| Change tail probabilities | 2 | .156 | .069 | .094 | .152 |
| Level variance/covariances | 12 | .156 | .084 | .094 | .194 |
| Level P10/P90 | 6 | .090 | .044 | .047 | .067 |

- Each column: one calibration of that model to all 70 moments
- Entry: RMS over row's moments of $2(m_k - d_k)/(|m_k| + |d_k|)$; $.284 \approx 28\%$ typical miss
- Lower is better; RE: current earnings-residual tercile within age group

◂ Targets ◂ History fit

# Calibration objective and search

$$\min_{\theta} \sum_{k=1}^{70} w_k \left( \frac{2[m_k(\theta) - d_k]}{|m_k(\theta)| + |d_k|} \right)^2$$

- $m_k(\theta), d_k$: model and data moments; weights are equal across sets and within sets
- Screen 64 quasi-random points, MLE, and earlier changes-only calibration
- Run four Nelder–Mead searches, using objective values without derivatives
- Use 40 panels for search, 100 for refinement, and 200 fresh panels for reporting
- PLV uses annual matrix exponentials and 501 / 1,001 / 2,001 nodes, respectively

Calibration Ranges

# Calibration restrictions

- Standard deviations of transitory components and specified Gaussian/mixture innovation components are at least 0.025
- Annual persistence $\rho \leq 0.99$
- PLV tail restrictions:

$$\frac{2\kappa}{c_+} \geq 1.1, \qquad \frac{2\kappa}{c_-} \geq 0.4$$

- Initial dispersion $\sigma_0 \in [0.025, 1.5]$; permanent-effect rates $a, b \in [1.1, 20]$
- Fix these restrictions before the 70-moment calibration; retain all likelihood estimates

◂ Calibration

# Consumption-equivalent welfare

- Let $\mathcal{W}_m$ be expected lifetime utility from zero assets, averaging over process $m$'s stationary earnings law
- Consumption equivalent solves

$$\frac{u(\mathrm{CE}_m)}{1-\beta} = \mathcal{W}_m, \qquad \mathrm{CE}_m = [(1-\gamma)(1-\beta)\mathcal{W}_m]^{1/(1-\gamma)}$$

- Mean raw earnings are one; total income includes transfer $b = 0.1$
- With zero assets, no borrowing, and $\beta R = 0.9792 < 1$, no-risk consumption benchmark is 1.1
- CE can exceed one while remaining below no-risk benchmark

◂ Saving problem ◂ Welfare results

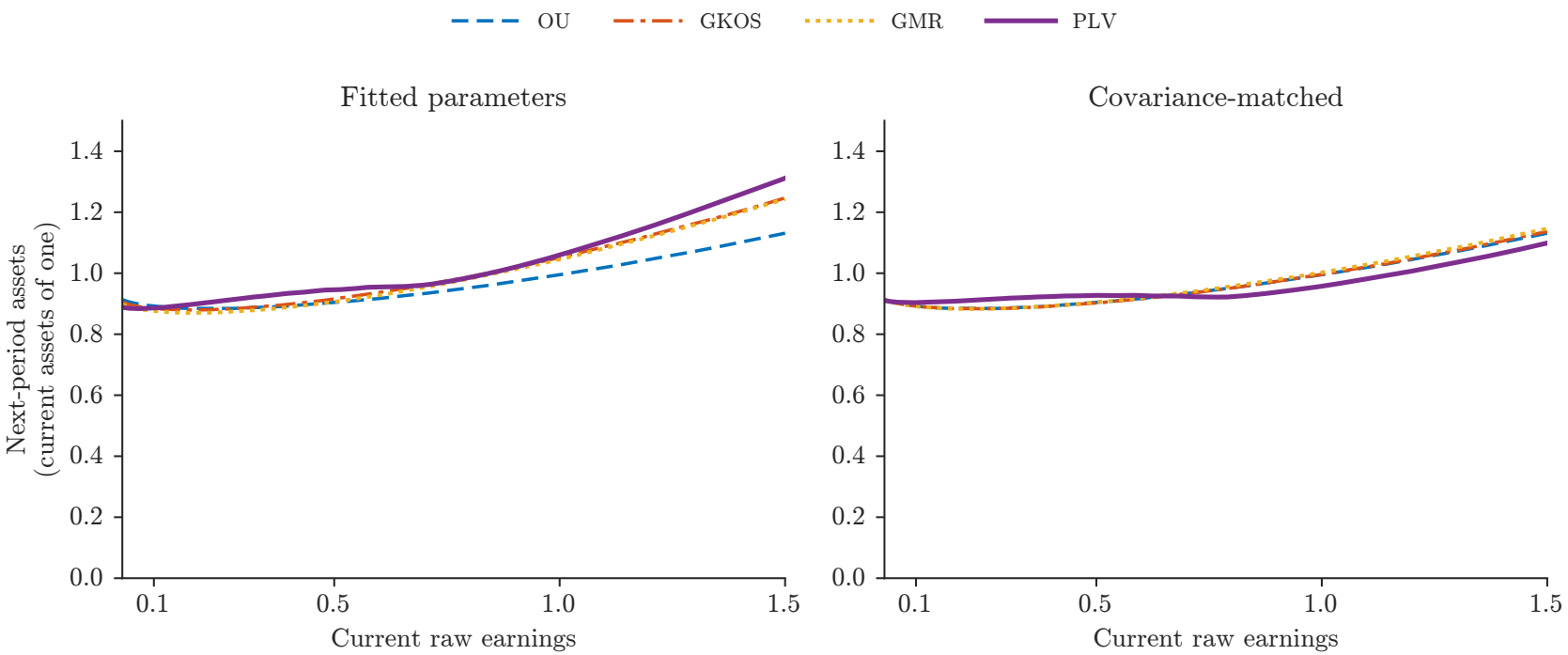
# Calibration to changes only

| | OU | GKOS | GMR | PLV |
|---|---|---|---|---|
| Calibrated objective | 0.534 | 0.0010 | 0.0028 | 0.062 |
| Calibrated LL | −14,476 | −14,361 | −14,087 | −13,538 |
| Mean assets | 4.31 | 18.05 | 13.25 | 5.30 |
| CE | 0.903 | 0.443 | 0.486 | 0.825 |

- 18 moments of two-, four-, and six-year changes; permanent effects cancel
- Hold permanent-effect rates and initial dispersion at likelihood estimates
- GKOS and GMR match these targets closely but fit histories worse than every likelihood-estimated mixture specification

◂ Choice of targets

# Saving policies




Saving comparison

# References


- Altman, R. M. (2007). "Mixed hidden Markov models: An extension of the hidden Markov model to the longitudinal data setting". *Journal of the American Statistical Association* 102.477, 201–210.
- Browning, M., L. P. Hansen, and J. J. Heckman (1999). "Micro data and general equilibrium models". In: *Handbook of Macroeconomics*. Ed. by J. B. Taylor and M. Woodford. Vol. 1. A. Elsevier. Chap. 8, 543–633.
- Crawley, E., M. B. Holm, and H. Tretvoll (2026). "A parsimonious model of idiosyncratic income". *International Economic Review* 67.2, 533–548.
- Farmer, L. E. (2021). "The discretization filter: A simple way to estimate nonlinear state space models". *Quantitative Economics* 12.1, 41–76.

# References


- Farmer, L. E. and A. A. Toda (2017). "Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments". *Quantitative Economics* 8.2, 651–683.
- Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2007). "Estimating macroeconomic models: A likelihood approach". *Review of Economic Studies* 74.4, 1059–1087.
- Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). "What do data on millions of U.S. workers reveal about lifecycle earnings dynamics?" *Econometrica* 89.5, 2303–2339.
- Guvenen, F., A. McKay, and C. Ryan (2026). "A tractable income process for business cycle analysis". *Review of Economic Dynamics* 61, 101345.
- Hansen, L. P. and J. J. Heckman (1996). "The empirical foundations of calibration". *Journal of Economic Perspectives* 10.1, 87–104.

# References


- Hnatkovska, V., V. Marmer, and Y. Tang (2012). "Comparison of misspecified calibrated models: The minimum distance approach". *Journal of Econometrics* 169.1, 131–138.
- Kitagawa, G. (1987). "Non-Gaussian state-space modeling of nonstationary time series". *Journal of the American Statistical Association* 82.400, 1032–1041.
- Kydland, F. E. and E. C. Prescott (1982). "Time to build and aggregate fluctuations". *Econometrica* 50.6, 1345–1370.
- Kydland, F. E. and E. C. Prescott (1996). "The computational experiment: An econometric tool". *Journal of Economic Perspectives* 10.1, 69–85.
- Langrock, R., I. L. MacDonald, and W. Zucchini (2012). "Some nonstandard stochastic volatility models and their estimation using structured hidden Markov models". *Journal of Empirical Finance* 19.1, 147–161.

# References


- Vuong, Q. H. (1989). "Likelihood ratio tests for model selection and non-nested hypotheses". *Econometrica* 57.2, 307–333.
- Watson, M. W. (1993). "Measures of fit for calibrated models". *Journal of Political Economy* 101.6, 1011–1041.
- White, H. (1982). "Maximum likelihood estimation of misspecified models". *Econometrica* 50.1, 1–25.