Welfare Consequences of Doubts about CEX Consumption Dynamics*

Thomas J. Sargent
New York University

Ziyue Yang
Australian National University

April 20, 2026

Abstract

This paper computes welfare costs of risks and of model uncertainties for heterogeneous consumers whose well-being depends on consumption dynamics. We equip each consumer with preferences that express doubts about two components of their model of consumption dynamics: (1) a linear–Gaussian law of motion for quantile-specific consumption, and (2) a Markov chain for cross-percentile transitions. Relative entropy penalty parameters θ_ε and θ_Π determine sets of possible distortions to the probability distribution of shocks to the law of motion for quantile-specific consumption and to the transition matrix across quantiles, respectively. Alternative settings of $\theta_\varepsilon, \theta_\Pi$ activate these components of consumers' specification concerns. We study four cases: no misspecification concerns (Case 0), concerns about dynamics alone (Case 1), mobility alone (Case 2), and both (Case 3). Applied to Consumer Expenditure Survey data (1990–2024), with robustness parameters calibrated via detection-error probabilities, our baseline calibrated compensating-difference measure α exhibits the ordering $\alpha^{(0)} < \alpha^{(1)} < \alpha^{(2)} < \alpha^{(3)}$ with a three- to fourfold amplification of welfare costs with all robustness turned on. Doubts about mobility across quantiles bring larger welfare costs than doubts about within-quantile dynamics.

Keywords: Welfare, heterogeneity, common trends, model uncertainties.
JEL Classification: E3, E6, C7

*We thank Guido Menzio and Fabrizio Perri for helpful suggestions.

1 Introduction

This paper extends representative consumer welfare calculations of Lucas (1987, 2003) by quantifying welfare costs of risks and model uncertainty for consumers exposed to idiosyncratic shocks to percentile-specific consumption growth rates and to Markov transitions across percentiles. We attribute a common approximating statistical model to all consumers. Our title refers to doubts inside our consumers' minds. Following Hansen & Sargent (2008*b*) and Barillas et al. (2009), all consumers harbor doubts about the shared probability model that they use to forecast future consumption and income. They cope with those doubts by forming cautious valuations that take into account worrisome and difficult-to-detect possible model misspecifications. We study how the utility costs of these doubts differ across percentiles and across two household ranking schemes.[1]

Our starting point is work by Obstfeld (1994) and Tallarini (2000), who modify the welfare calculations of Lucas (1987) by improving his specification of a representative consumer's baseline statistical model for consumption growth, and using a version of Epstein-Zin preferences. Sargent et al. (2026*a*) modify Lucas's setup further by estimating a linear–Gaussian model of evolving cross-sections of private income, post-tax income, and consumption from the Consumer Expenditure Survey (CEX, 1990–2024). They represent each percentile in the cross-section as an additive functional (see Hansen (2012)) that is driven by a reduced-rank first-order vector autoregression estimated via an algorithm based on a Dynamic Mode Decomposition (see Schmid (2010), Kutz et al. (2016), Sargent et al. (2026*b*)). Consumers permanently stuck at a quantile evaluate consumption streams via standard expected-utility preferences. Sargent et al. find that serially correlated risks generate much larger welfare costs than do i.i.d. risks, and that benefits (or costs) from participating in the U.S. tax-transfer-and-insurance system dwarf the benefit of elimi-

[1] Following a recommendation of Lucas (2003), we ignore information about the welfare costs of risks in consumption growth that is embedded in asset prices according to formulations of Hansen et al. (1999), Tallarini (2000), and Alvarez & Jermann (2004). Barillas et al. (2009) argued that endowing agents with concerns about model misspecification provides a way to overcome Lucas's objection to using asset prices to calibrate welfare costs of aggregate fluctuations. We leave that for future work.

nating consumption-growth risk. However, their analysis holds each household's percentile rank fixed, and by using expected-utility preferences, they impute no doubts about the statistical model to their percentile-specific representative consumers. The present paper relaxes both restrictions.

Our first extension captures mobility across percentiles via a Markov chain with transition matrix $\Pi \in \mathbb{R}^{M\times M}$ whose (p, j) entry is the probability that a household at percentile p moves to percentile j next period, estimated from consecutive CEX interview pairs under two ranking schemes, by consumption and by private income. Under both ranking schemes, Markov transition matrices exhibit strong diagonal persistence, but the matrix for consumption quantiles is less concentrated on the diagonal, reflecting greater cross-quantile mixing in consumption than in income. Mobility matters because a household in one consumption percentile today may move to another percentile tomorrow, so its value function depends not only on current consumption but on where the Markov transition matrix might send it.

Our second extension activates doubts about the statistical models that consumers use to evaluate utilities of prospective consumption. We equip each agent with what Hansen & Sargent (2001) and Barillas et al. (2009) call multiplier preferences with "multiplier" or "penalty" parameter θ on relative-entropy divergences from a baseline statistical model. Consumers inside our model share the same baseline statistical model, along with us, the model builders.[2] A large penalty parameter θ keeps the set of alternative models close to the approximating baseline model; so as $\theta \to \infty$, a consumer's preferences over consumption plans become ordered by discounted expected utilities. We introduce two multipliers: θ_ε influences the size of the set of distortions to the Gaussian shock distribution (doubts about consumption dynamics within each quantile) and θ_Π influences the size of the set of possible distortions to the transition matrix (doubts about transitions across percentiles). A nested two-stage aggregator implements this specification: the inner stage applies θ_ε for each quantile, and the outer stage applies θ_Π across transition destinations.

[2]Multiplier preferences are closely related to the stochastic differential utility specification of Duffie & Epstein (1992). In both specifications, the decision maker especially dislikes serially correlated risk. Because our consumers' baseline approximating statistical model presents serially correlated risk, we can anticipate that adding concerns about robustness will amplify welfare gains from reducing serially correlated risks relative to those computed by Sargent et al. (2026*a*).

Activating distortions to different components of the likelihood ratio yields four cases. *Case 0* ($\theta_\varepsilon = \theta_\Pi = \infty$, no robustness) recovers the expected-utility value function of Sargent et al. (2026*a*) extended to recognize Markov transitions across quantiles. *Case 1* ($\theta_\varepsilon = \theta < \infty$, $\theta_\Pi = \infty$) allows only the quantile-specific consumption-growth shock distribution to be distorted. Cases 0 and 1 both admit closed-form affine value functions; the Case 1 solution generalizes the scalar formulas of Hansen & Sargent (2008*a*) to M quantiles, with the scalar $(1-\beta)^{-1}$ replaced by a resolvent operator $(I - \beta\Pi)^{-1}$. *Case 2* ($\theta_\varepsilon = \infty$, $\theta_\Pi = \theta < \infty$) allows only the transition matrix to be distorted, and *Case 3* ($\theta_\varepsilon = \theta_\Pi = \theta$) is the unrestricted problem. Cases 2 and 3 involve a log - $\sum$ - exp aggregator that disrupts the affine structure, so we solve them by iterating a Bellman operator on a grid in the low-dimensional DMD state space ($N = 5$ dimensions, $M = 100$ quantiles).

Following Anderson et al. (2003) and Barillas et al. (2009), we calibrate θ_ε and θ_Π via detection-error probabilities (DEPs), targeting a Bayesian detection-error probability $\bar{p}$ of 0.45, so that, to a Bayesian statistician, the baseline approximating and worst-case models are nearly indistinguishable.

Our main quantitative findings are as follows. When we apply our framework to CEX data under consumption-ranked percentiles, Case 0 compensating differences are nearly flat across percentiles at about 0.052% per year. Robustness to dynamics alone (Case 1) roughly doubles this to 0.111%; adding mobility doubts (Case 2) raises it to 0.182–0.187%; full robustness (Case 3) reaches 0.195–0.200%. The worst-case mobility distortion tilts probabilities toward low-consumption quantiles, producing a mild downward slope across quantiles in Cases 2 and 3.

Under private-income ranking, welfare costs are similar in magnitude but the cross-quantile profile tilts in the opposite direction, with compensating differences in Cases 2 and 3 larger at the top of the income distribution. The private-income transition matrix is tightly concentrated on the diagonal, so the worst-case distortion operates by reducing persistence at high income ranks and shifting probability mass downward, which reverses the consumption-ranking pattern. The direction of the tilt results from the off-diagonal structure of the estimated transition matrix.

Across both rankings, robustness raises welfare costs throughout the cross-section.

Doubts about mobility across percentiles contribute more than doubts about within-percentile consumption growth rates. The two sources of doubt together roughly triple compensating differences relative to the no-robustness benchmark.

Related Literature

A large empirical literature uses the Consumer Expenditure Survey to study the joint distribution of income and consumption in the United States. Cutler & Katz (1992) provide an early systematic account showing that consumption inequality rose alongside income inequality during the 1980s. Attanasio & Davis (1996) use synthetic cohort panels from the CEX to show that large relative-wage movements among education groups drove correspondingly large shifts in the distribution of household consumption, implying limited cross-group risk sharing. Krueger & Perri (2003) document diverging paths of income and consumption inequality in the CEX and ask about the welfare consequences of rising US earnings dispersion. Krueger & Perri (2006) argue that the striking gap between income and consumption inequality reflects improved private risk-sharing facilitated by expanded consumer credit markets. Blundell et al. (2008) combine the CEX with the PSID to structurally estimate the degree of insurance against permanent versus transitory income shocks, finding substantial but incomplete insurance, especially for college-educated households. Krueger et al. (2010) survey cross-sectional facts on wages, income, and consumption inequality for the United States and several other countries, providing a standard calibration reference for heterogeneous-agent models. Heathcote, Perri & Violante (2010) construct a unified picture of US inequality from 1967 to 2006 using the CEX, CPS, PSID, and SCF, documenting trends across wages, earnings, income, and consumption for a broad range of household types. Heathcote, Storesletten & Violante (2010) build a quantitative life-cycle model calibrated to CEX consumption data to account simultaneously for trends in wage, earnings, and consumption inequality over 1967–2000. Heathcote et al. (2014) develop a tractable partial-insurance framework calibrated to CEX data that analytically links income-process parameters, insurance coefficients, and the cross-sectional consumption distribution. Heathcote et al. (2017) extend that framework to solve for the optimal degree of

tax progressivity, trading insurance against idiosyncratic risk against disincentive effects on work and skill investment. Attanasio et al. (2014) reassess CEX trends in income, consumption, and leisure inequality from 1980 to 2010, documenting how demographic composition changes and non-monetary returns to leisure affect the inferred rise in consumption inequality. Aguiar & Bils (2015) argue that systematic measurement error in the CEX induced earlier studies to understate the rise in consumption inequality; using a demand-system correction they find that consumption inequality has in fact tracked income inequality closely since 1980. Blundell et al. (2016) extend the partial-insurance framework to incorporate family labor supply, finding that secondary earners provide substantial insurance against permanent income shocks. Finally, Meyer & Sullivan (2023) provide a long-run account from 1963 to 2019 showing that, once CEX under-reporting is addressed, consumption inequality increased substantially and comparably to income inequality.

Although we take the measurement issues raised by Aguiar & Bils (2015), Meyer & Sullivan (2023), and others seriously, we abstract from them in this paper. Our rationale is to establish benchmarks against which to assess the welfare consequences of introducing explicit models of measurement error in income and consumption, which we plan to do in a sequel.

The remainder of the paper is organized as follows. Section 2 specifies the approximating model. Section 3 formulates the multiplier problem and solves the inner minimization (Lemma 3.1). Section 4 derives four cases by restricting the admissible likelihood ratios to isolate robustness to consumption dynamics, mobility, or both (Propositions 4.1–4.3), as well as the no-robustness baseline (Case 0). Section 5 derives closed-form expressions for the robust value function in the linear–Gaussian case, together with certainty equivalents and compensating differences. Section 6 describes the estimation of the approximating model from CEX microdata, the construction of a rank-N reduced model via Dynamic Mode Decomposition, and the calibration of the robustness parameters $(\theta_\varepsilon, \theta_\Pi)$ via detection-error probabilities. We then report welfare results under two alternative household ranking schemes, by consumption (Section 6.7) and by private income (Section 6.8), and show how the choice of grouping variable shapes cross-quantile compensating differences (Sec-

tion 6.9).

2 Consumers' statistical models

This section describes the common approximating statistical model shared by all agents. This baseline model is surrounded by a set of alternative models that are statistically difficult to distinguish from it. Each agent fears that the true data-generating process differs from the approximating model and therefore orders consumption plans with multiplier preferences.

2.1 State space

For each household, the state at date $t \in \mathbb{N}$ consists of three components:

- a quantile index $p_t \in \{0, \ldots, M-1\}$,
- a continuous state vector $\zeta_t \in \mathbb{R}^n$,
- a cross-quantile log-consumption vector $c_t := (c_{0,t}, \ldots, c_{M-1,t})^\top \in \mathbb{R}^M$.

We denote the state space by $\mathsf{S} := \{0, \ldots, M-1\} \times \mathbb{R}^n \times \mathbb{R}^M$.

2.2 The approximating statistical model

The continuous state vector ζ_t and consumption of a quantile-p household evolve according to

$$\zeta_{t+1} = A\zeta_t + C\varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, I_k), \tag{1}$$

$$c_{q,t+1} - c_{q,t} = D_q\zeta_t + F_q\varepsilon_{t+1}, \qquad q = 0, \ldots, M-1, \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{n\times k}$, and, for each quantile q, $D_q \in \mathbb{R}^{1\times n}$ and $F_q \in \mathbb{R}^{1\times k}$ are the q-th rows of matrices $D \in \mathbb{R}^{M\times n}$ and $F \in \mathbb{R}^{M\times k}$.

The quantile state follows a Markov chain with transition matrix $\Pi \in \mathbb{R}^{M\times M}$:

$$\text{Prob}(p_{t+1} = j \mid p_t = i) = \Pi_{ij}, \qquad \sum_{j=0}^{M-1} \Pi_{ij} = 1. \tag{3}$$

At state (p, ζ, c), per-period utility is

$$u(p, c) = (1 - \beta)e_p^\top c = (1 - \beta)c_p, \qquad \beta \in (0, 1), \tag{4}$$

where c_p is the p-th element of the vector c and e_p is the p-th standard basis vector in $\mathbb{R}^M$.

Given the current state (p, ζ, c), the approximating model specifies a joint conditional probability distribution over (p', ε):

$$Q\big(\{j\} \times \mathrm{d}\varepsilon \;\big|\; p, \zeta, c\big) = \Pi_{pj}\varphi(\varepsilon)\, \mathrm{d}\varepsilon, \tag{5}$$

where φ denotes the $N(0, I_k)$ density. The next period state (ζ', c') follows from (ζ, c, ε) via (1)–(2).

3 Multiplier preferences

Following Hansen & Sargent (2008*b*) and Barillas et al. (2009), each consumer distrusts the approximating kernel Q and evaluates continuation values under a worst-case alternative, subject to a penalty $\theta > 0$ on its relative entropy with respect to Q.

Let v_p be the value function of a quantile-p agent and assume that v_p is measurable over $\mathsf{X} := \mathbb{R}^n \times \mathbb{R}^M$ for each p and satisfies the exponential integrability condition

$$\mathbb{E}_Q\left[\exp\left(-\frac{\beta}{\theta}v_{p'}(\zeta', c')\right)\middle| p, \zeta, c\right] < \infty \quad \text{for all } (p, \zeta, c) \text{ and } \theta > 0, \beta \in (0, 1). \tag{6}$$

This condition ensures that the entropic risk operator (13) and the exponential tilts (12) are well defined. It is weaker than boundedness and, in particular, is sat-

isfied by affine functions $v_j(\zeta, c) = \lambda_j^\top \zeta + b_j^\top c + \kappa_j$ whenever the shocks are Gaussian (see Section 5).

3.1 The $\mathcal{T}_\theta$ operator

Fix the current state (p, ζ, c) and let $\mathcal{H}(p, \zeta, c)$ denote the set of admissible likelihood ratios containing all measurable functions $h(p', \varepsilon) \geqslant 0$ satisfying

$$\mathbb{E}_Q \left[h(p', \varepsilon) \middle| p, \zeta, c \right] = 1. \tag{7}$$

The $\mathcal{T}_\theta$ operator applied to the discounted continuation value $\beta v(\cdot)$ is

$$\mathcal{T}_\theta[\beta v](p, \zeta, c) := \min_{h \in \mathcal{H}(p,\zeta,c)} \mathbb{E}_Q \left[h(p', \varepsilon) \Big(\beta v_{p'}(\zeta', c') + \theta \log h(p', \varepsilon) \Big) \middle| p, \zeta, c \right]. \tag{8}$$

This corresponds to the operator T_1 defined in Hansen & Sargent (2008*a*).

The Bellman equation is

$$v_p(\zeta, c) = (1 - \beta) c_p + \mathcal{T}_\theta[\beta v](p, \zeta, c). \tag{9}$$

Define the Bellman operator Φ_θ on functions $v = (v_0, \ldots, v_{M-1})$ satisfying (9) by

$$(\Phi_\theta v)_p(\zeta, c) := (1 - \beta)\, c_p + \mathcal{T}_\theta[\beta v](p, \zeta, c), \tag{10}$$

so that the Bellman equation (9) can be written as $v = \Phi_\theta v$.

When there is no restriction on what can be taken from $\mathcal{H}(p, \zeta, c)$ (i.e., no restriction apart from $h \geqslant 0$ and the mean-one constraint (7)), the minimizer and the minimized value $\mathcal{T}_\theta$ take the following form.

Lemma 3.1. *The minimizer of*

$$\min_{h \geqslant 0,\, \mathbb{E}_Q[h|p,\zeta,c]=1} \mathbb{E}_Q \left[h \big(\beta v_{p'}(\zeta', c') + \theta \log h \big) \,\middle|\, p, \zeta, c \right] \tag{11}$$

is unique up to $Q(\cdot \mid p, \zeta, c)$-a.s. equivalence and is given on the support of Q by

$$h^*(p', \varepsilon \mid p, \zeta, c) = \frac{\exp\left(-\frac{\beta}{\theta} v_{p'}(\zeta', c')\right)}{\mathbb{E}_Q\left[\exp\left(-\frac{\beta}{\theta} v_{p'}(\zeta', c')\right) \middle| p, \zeta, c\right]}, \tag{12}$$

and the minimized value is

$$\mathcal{T}_\theta[\beta v](p, \zeta, c) = -\theta \log \mathbb{E}_Q\left[\exp\left(-\frac{\beta}{\theta} v_{p'}(\zeta', c')\right) \middle| p, \zeta, c\right]. \tag{13}$$

Proof. For each outcome $(p', \varepsilon) = (j, e)$, treat $h(j, e)$ as an independent scalar variable. Let $Y(j, e) := \beta v_j(\zeta', c')$ for the continuation value at that outcome. Form the Lagrangian with multiplier ν for the constraint $\mathbb{E}_Q[h] = 1$:

$$\mathcal{L}(h, \nu) = \mathbb{E}_Q\left[h(Y + \theta \log h)\right] + \nu\big(\mathbb{E}_Q[h] - 1\big). \tag{14}$$

Differentiating with respect to $h(j, e)$ and using $\frac{d}{dh}(h \log h) = 1 + \log h$:

$$Y + \theta(1 + \log h) + \nu = 0 \quad \Longrightarrow \quad \log h = -\frac{Y}{\theta} - 1 - \frac{\nu}{\theta}.$$

Hence

$$h = \underbrace{\exp\left(-1 - \frac{\nu}{\theta}\right)}_{=:K} \exp\left(-\frac{Y}{\theta}\right), \tag{15}$$

where $K > 0$ is constant in (p', ε). The constraint $\mathbb{E}_Q[h] = 1$ gives

$$K^{-1} = \mathbb{E}_Q\left[\exp\left(-\frac{Y}{\theta}\right)\right]. \tag{16}$$

Substituting $Y = \beta v_{p'}(\zeta', c')$ into (15) and using (16) yields (12).

To obtain (13), substitute h^* in (12) back into the objective and set $Z := \mathbb{E}_Q[\exp(-Y/\theta)]$.

Then

$$\begin{aligned}\mathbb{E}_Q[h^*(Y+\theta\log h^*)] &= \mathbb{E}_Q\left[\frac{e^{-Y/\theta}}{Z}\left(Y+\theta\log\frac{e^{-Y/\theta}}{Z}\right)\right]\\ &= \mathbb{E}_Q\left[\frac{e^{-Y/\theta}}{Z}\left(Y-Y-\theta\log Z\right)\right] = -\theta\log Z.\end{aligned}$$

Uniqueness follows from strict convexity of $h \mapsto h\log h$. On $Q(\cdot \mid p,\zeta,c)$-null outcomes, h does not affect either the objective or the constraint, so uniqueness holds Q-almost surely. □

4 Four cases: restrictions on the likelihood ratio

By restricting the admissible form of the likelihood ratio $h(p',\varepsilon)$, we obtain four cases from the same specification (8). Each case is solved either by the same steps used in proving Lemma 3.1 or by a straightforward extension of it.

4.1 Case 0: No robustness

Since θ is the penalty on distortion, $\theta \to \infty$ corresponds to the case of no distortion. When $\theta \to \infty$, $h \equiv 1$ and the value function satisfies

$$\begin{aligned}v_p(\zeta,c) &= (1-\beta)c_p + \mathbb{E}_Q\left[\beta v_{p'}(\zeta',c')\middle|p,\zeta,c\right]\\ &= (1-\beta)c_p + \beta\sum_{j=0}^{M-1}\Pi_{pj}\mathbb{E}_\varphi\left[v_j(\zeta',c')\middle|\zeta,c\right].\end{aligned} \tag{17}$$

4.2 Case 1: Robustness to consumption dynamics only

Restrict the likelihood ratio $h \in \mathcal{H}$ to the form

$$h(p',\varepsilon) = g_{p'}(\varepsilon), \qquad \mathbb{E}_\varphi[g_j(\varepsilon)] = 1 \quad \text{for each } j = 0,\dots,M-1. \tag{18}$$

Under this restriction, the transition probabilities Π_{pj} are unchanged, but the conditional shock distribution may differ across destination quantiles.

Proposition 4.1. *Under the restriction* (18)*, the value function satisfies*

$$v_p(\zeta, c) = (1-\beta)c_p - \theta \sum_{j=0}^{M-1} \Pi_{pj} \log \mathbb{E}_\varphi \left[\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right) \middle| \zeta, c \right]. \tag{19}$$

The minimizing likelihood ratio for j as next-period quantile is

$$g_j^*(\varepsilon \mid \zeta, c) = \frac{\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right)}{\mathbb{E}_\varphi \left[\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right) \middle| \zeta, c \right]}. \tag{20}$$

Proof. Expanding $\mathbb{E}_Q[\cdot \mid p, \zeta, c]$ using the kernel (5) and substituting $h(j, \varepsilon) = g_j(\varepsilon)$, the objective in (8) becomes

$$\sum_{j=0}^{M-1} \Pi_{pj} \mathbb{E}_\varphi \left[g_j(\varepsilon) \Big(\beta v_j(\zeta', c') + \theta \log g_j(\varepsilon) \Big) \middle| \zeta, c \right]. \tag{21}$$

The constraint $\mathbb{E}_Q[h] = 1$ decomposes as $\sum_j \Pi_{pj} \mathbb{E}_\varphi[g_j] = 1$, which holds whenever $\mathbb{E}_\varphi[g_j] = 1$ for each j. The minimization of (21) therefore separates into M independent problems:

$$\min_{g_j \geqslant 0, \mathbb{E}_\varphi[g_j]=1} \mathbb{E}_\varphi \left[g_j(\varepsilon) \Big(\beta v_j(\zeta', c') + \theta \log g_j(\varepsilon) \Big) \middle| \zeta, c \right]. \tag{22}$$

Each sub-problem has the same structure as (11) with $\mathbb{E}_Q$ replaced by $\mathbb{E}_\varphi$. The Lagrangian for the j-th sub-problem is

$$\mathcal{L}_j(g_j, \nu_j) = \mathbb{E}_\varphi \left[g_j \big(\beta v_j(\zeta', c') + \theta \log g_j \big) \right] + \nu_j \big(\mathbb{E}_\varphi[g_j] - 1 \big).$$

The first-order condition gives $g_j = K_j \exp(-\beta v_j(\zeta', c')/\theta)$, where $K_j > 0$ is constant in ε. The constraint $\mathbb{E}_\varphi[g_j] = 1$ gives $K_j^{-1} = \mathbb{E}_\varphi[\exp(-\beta v_j(\zeta', c')/\theta) \mid \zeta, c]$, yielding (20).

To obtain the minimized value, define $Y_j := \beta v_j(\zeta', c')$ and $Z_j := \mathbb{E}_\varphi[\exp(-Y_j/\theta) \mid$

$\zeta, c]$. Substituting g_j^* back into (22):

$$\mathbb{E}_\varphi\left[g_j^*\left(Y_j + \theta \log g_j^*\right)\right] = \mathbb{E}_\varphi\left[\frac{e^{-Y_j/\theta}}{Z_j}\left(Y_j + \theta \log \frac{e^{-Y_j/\theta}}{Z_j}\right)\right]$$
$$= \mathbb{E}_\varphi\left[\frac{e^{-Y_j/\theta}}{Z_j}\left(-\theta \log Z_j\right)\right] = -\theta \log Z_j.$$

Summing over j with weights Π_{pj} and adding the flow utility gives (19). □

4.3 Case 2: Robustness to mobility only

Restrict the likelihood ratio to the form

$$h(p', \varepsilon) = m(p'), \qquad \sum_{j=0}^{M-1} \Pi_{pj} m(j) = 1. \tag{23}$$

Under this restriction, shocks retain the distribution $\varepsilon \sim N(0, I_k)$, but the transition probabilities are distorted. Define the conditional mean under the approximating model:

$$\bar{v}_j(\zeta, c) := \mathbb{E}_\varphi\left[v_j(\zeta', c') \middle| \zeta, c\right]. \tag{24}$$

Proposition 4.2. *Under the restriction* (23)*, the value function satisfies*

$$v_p(\zeta, c) = (1-\beta)c_p - \theta \log\left(\sum_{j=0}^{M-1} \Pi_{pj} \exp\left(-\frac{\beta}{\theta}\bar{v}_j(\zeta, c)\right)\right). \tag{25}$$

The minimizing likelihood ratio determines the distorted transition probabilities

$$\widetilde{\Pi}_{pj}(\zeta, c) = \frac{\Pi_{pj} \exp\left(-\frac{\beta}{\theta}\bar{v}_j(\zeta, c)\right)}{\sum_{k=0}^{M-1} \Pi_{pk} \exp\left(-\frac{\beta}{\theta}\bar{v}_k(\zeta, c)\right)}, \tag{26}$$

and the shock distribution is unchanged: $\widehat{g}_{pj}(\varepsilon) = 1$.

Proof. Since $h = m(p')$ does not depend on ε, the ε-expectation can be evaluated

first. Expanding $\mathbb{E}_Q[\cdot \mid p, \zeta, c]$ via (5):

$$\mathbb{E}_Q\left[m(p')\Big(\beta v_{p'}(\zeta', c') + \theta \log m(p')\Big)\Big| p, \zeta, c\right] = \sum_{j=0}^{M-1} \Pi_{pj} m(j)\Big(\beta \bar{v}_j(\zeta, c) + \theta \log m(j)\Big),$$

so the problem reduces to

$$\min_{\substack{m(j)\geqslant 0 \\ \sum_j \Pi_{pj} m(j) = 1}} \sum_{j=0}^{M-1} \Pi_{pj} m(j)\Big(\beta \bar{v}_j(\zeta, c) + \theta \log m(j)\Big). \tag{27}$$

The Lagrangian for the constraint $\sum_j \Pi_{pj} m(j) = 1$ is

$$\mathcal{L}(m, \nu) = \sum_{j=0}^{M-1} \Pi_{pj} m(j)\Big(\beta \bar{v}_j + \theta \log m(j)\Big) + \nu\left(\sum_{j=0}^{M-1} \Pi_{pj} m(j) - 1\right).$$

Differentiating with respect to $m(j)$ for each j with $\Pi_{pj} > 0$ (when $\Pi_{pj} = 0$ the objective and constraint are independent of $m(j)$):

$$\Pi_{pj}\Big(\beta \bar{v}_j + \theta(1 + \log m(j)) + \nu\Big) = 0 \quad \Longrightarrow \quad m^*(j) = K \exp\left(-\frac{\beta}{\theta} \bar{v}_j\right),$$

where $K = \exp(-1 - \nu/\theta) > 0$. The constraint gives $K^{-1} = \sum_j \Pi_{pj} \exp(-\beta \bar{v}_j/\theta)$, yielding (26) via $\widetilde{\Pi}_{pj} = \Pi_{pj} m^*(j)$.

For the minimized value, define $Z := \sum_j \Pi_{pj} \exp(-\beta \bar{v}_j/\theta)$. Substituting m^* back into (27):

$$\sum_j \Pi_{pj} m^*(j)\big(\beta \bar{v}_j + \theta \log m^*(j)\big) = \sum_j \Pi_{pj} \frac{e^{-\beta \bar{v}_j/\theta}}{Z}(-\theta \log Z) = -\theta \log Z.$$

Adding the flow utility gives (25). □

4.4 Case 3: Robustness to both mobility and consumption dynamics

The set $\mathcal{H}(p, \zeta, c)$ is unrestricted (apart from $h \geqslant 0$ and the mean-one constraint). Define the exponential-moment term

$$M_j(\zeta, c) := \mathbb{E}_\varphi \left[\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right) \middle| \zeta, c \right]. \tag{28}$$

Proposition 4.3. *When $\mathcal{H}$ is unrestricted, the value function satisfies*

$$v_p(\zeta, c) = (1 - \beta) c_p - \theta \log \left(\sum_{j=0}^{M-1} \Pi_{pj} M_j(\zeta, c) \right), \tag{29}$$

and the minimizing likelihood ratio is

$$h^*(j, \varepsilon \mid p, \zeta, c) = \frac{\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right)}{\sum_{k=0}^{M-1} \Pi_{pk} M_k(\zeta, c)}. \tag{30}$$

Proof. Apply Lemma 3.1 directly. Expanding $\mathbb{E}_Q[\cdot \mid p, \zeta, c]$ via the kernel (5):

$$\mathbb{E}_Q \left[\exp\left(-\frac{\beta}{\theta} v_{p'}(\zeta', c') \right) \middle| p, \zeta, c \right] = \sum_{j=0}^{M-1} \Pi_{pj} M_j(\zeta, c).$$

Substituting into (13) gives (29), and into (12) gives (30). □

The joint likelihood ratio h^* can be factored into a mobility distortion of the transition probabilities and a conditional shock distortion of the density. This factorization helps to clarify the relationship between Case 3 and Cases 1 and 2. In particular, the two components are coupled, so Case 3 is not equivalent to applying Cases 1 and 2 independently.

4.5 Separating penalties for shocks and mobility

To allow separate degrees of distortion for consumption dynamics and mobility, we introduce two penalty parameters $\theta_\varepsilon > 0$ and $\theta_\Pi > 0$ and apply the T operator in two nested stages.

In the first stage, for each destination quantile j, define the robust certainty equivalent

$$(\mathcal{R}_\varepsilon v_j)(\zeta, c) := -\theta_\varepsilon \log \mathbb{E}_\varphi \left[\exp\left(-\frac{\beta}{\theta_\varepsilon} v_j(\zeta', c') \right) \middle| \zeta, c \right]. \tag{31}$$

This applies the T operator with penalty θ_ε to the shock distribution, holding the destination quantile j fixed.

In the second stage, the robust certainty equivalents $\mathcal{R}_\varepsilon v_j$ from the first stage enter the second-stage mobility aggregation:

$$v_p(\zeta, c) = (1-\beta)c_p - \theta_\Pi \log\left(\sum_{j=0}^{M-1} \Pi_{pj} \exp\left(-\frac{(\mathcal{R}_\varepsilon v_j)(\zeta, c)}{\theta_\Pi} \right) \right). \tag{32}$$

The two-stage recursion (32) reduces to:

(a) Case 3 (29) when $\theta_\Pi = \theta_\varepsilon = \theta$;

(b) Case 1 (19) when $\theta_\Pi \to +\infty$ with $\theta_\varepsilon = \theta$;

(c) Case 2 (25) when $\theta_\varepsilon \to +\infty$ with $\theta_\Pi = \theta$;

(d) Case 0 (17) when $\theta_\Pi \to +\infty$ and $\theta_\varepsilon \to +\infty$.

5 Closed forms under the linear–Gaussian model

With Bellman equations for general forms of our four cases in hand, we now specialize to the linear–Gaussian model described in Section 2. We show that the value functions for Cases 0 and 1 have closed-form affine solutions, but that those for Cases 2 and 3 generally do not.

Recall the law of motion primitives: $\zeta' = A\zeta + C\varepsilon$, $c' = c + D\zeta + F\varepsilon$, with $\varepsilon \sim N(0, I_k)$

(equations (1)–(2)). Conjecture that the value function is affine:

$$v_j(\zeta, c) = \lambda_j^\top \zeta + b_j^\top c + \kappa_j, \tag{33}$$

for vectors $\lambda_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}^M$, and scalars κ_j. Substituting the law of motion,

$$v_j(\zeta', c') = \underbrace{\kappa_j + b_j^\top c + (\lambda_j^\top A + b_j^\top D)\zeta}_{=:\mu_j(\zeta,c)} + \underbrace{(\lambda_j^\top C + b_j^\top F)}_{=:s_j^\top} \varepsilon, \tag{34}$$

where $s_j := C^\top \lambda_j + F^\top b_j \in \mathbb{R}^k$ and $\ell_j := \|s_j\|^2$. Since $s_j^\top \varepsilon \sim N(0, \ell_j)$ under φ,

$$\mathbb{E}_\varphi \left[\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right) \middle| \zeta, c \right] = \exp\left(-\frac{\beta}{\theta}\mu_j + \frac{\beta^2}{2\theta^2}\ell_j \right), \tag{35}$$

and hence

$$-\theta \log \mathbb{E}_\varphi \left[\exp\left(-\frac{\beta}{\theta} v_j(\zeta', c') \right) \middle| \zeta, c \right] = \beta\mu_j(\zeta, c) - \frac{\beta^2}{2\theta}\ell_j. \tag{36}$$

Equation (36) shows that the shock-robust certainty equivalent remains affine whenever the candidate value function is affine. We now substitute into each Bellman equation to check whether our conjectured affine form (33) holds and to solve for the coefficients λ_j, b_j, and κ_j.

5.1 Case 0 and Case 1: Affine closed forms

For Case 0, the value function is affine with $\kappa = 0$. Define $U \in \mathbb{R}^{M\times M}$ as the utility-loading matrix whose p-th row is $U_p := (1-\beta)e_p^\top$, so that $u(p, c) = U_p\, c$ and hence $U = (1-\beta)I_M$. The result is summarized in the following proposition.

Proposition 5.1 (Case 0 closed form)**.** *The value function for Case 0 is affine in* (ζ, c) *with* $\kappa = 0$ *and the coefficient matrices* $B := [b_0, \ldots, b_{M-1}]^\top \in \mathbb{R}^{M\times M}$ *and* $\Lambda := [\lambda_0, \ldots, \lambda_{M-1}]^\top \in \mathbb{R}^{M\times n}$ *satisfying*

$$B = (I - \beta\Pi)^{-1} U, \tag{37}$$

$$\Lambda = \beta\Pi\Lambda A + \beta\Pi B D. \tag{38}$$

Proof. Since $\mathbb{E}_\varphi[\varepsilon] = 0$, substituting $\mathbb{E}_\varphi[v_j(\zeta', c') \mid \zeta, c] = \mu_j(\zeta, c)$ into (17):

$$\begin{aligned} v_p(\zeta, c) &= U_p\, c + \beta \sum_j \Pi_{pj} \mu_j(\zeta, c) \\ &= U_p\, c + \beta \sum_j \Pi_{pj} \big(\kappa_j + b_j^\top c + (\lambda_j^\top A + b_j^\top D)\zeta\big). \end{aligned}$$

Collecting the terms that depend on neither ζ nor c on each side gives $\kappa_p = \beta \sum_j \Pi_{pj} \kappa_j$ for every p, and hence $\kappa = \beta \Pi \kappa$. Since Π is a stochastic matrix, $(I - \beta\Pi)$ is invertible, and $(I - \beta\Pi)\kappa = 0$ implies $\kappa = 0$. This proves the first claim.

The right-hand side is affine in (ζ, c). Matching coefficients of c gives $B = U + \beta \Pi B$, hence $B = (I - \beta\Pi)^{-1} U$. Matching ζ gives (38). □

Case 1 also has an affine closed form, but the coefficients differ from Case 0. The following proposition shows that the value function is affine with the same B and Λ as Case 0, but with a nonzero constant term κ that depends on the shock distortion.

Proposition 5.2 (Case 1 closed form)**.** *The value function for Case 1 is affine in (ζ, c) with B and Λ as in Case 0 and*

$$\kappa = -\frac{\beta^2}{2\theta} (I - \beta\Pi)^{-1} \Pi \ell, \tag{39}$$

where $\ell = (\ell_0, \ldots, \ell_{M-1})^\top$ with $\ell_j = \|s_j\|^2$. Conditional on destination j, the worst-case shock is $\varepsilon \sim N(-(\beta/\theta) s_j, I_k)$.

Proof. Substituting (36) into (19):

$$v_p(\zeta, c) = (1 - \beta) c_p + \beta \sum_j \Pi_{pj} \mu_j(\zeta, c) - \frac{\beta^2}{2\theta} \sum_j \Pi_{pj} \ell_j.$$

Matching slopes gives the same B and Λ as Case 0; matching constants gives $(I - \beta\Pi)\kappa = -\frac{\beta^2}{2\theta} \Pi \ell$ and invertibility follows from the same argument in Proposition 5.1. For the worst-case shock, substitute (34) into (20). Since $\mu_j(\zeta, c)$ does not depend

on ε, it appears identically in numerator and denominator and cancels, leaving

$$g_j^*(\varepsilon) \propto \exp\left(-\frac{\beta}{\theta} s_j^\top \varepsilon\right).$$

To find the normalizing constant, use $s_j^\top \varepsilon \sim N(0, \ell_j)$ and the moment generating function identity $\mathbb{E}[\exp(tZ)] = \exp(\frac{1}{2}t^2\ell_j)$ for $Z \sim N(0, \ell_j)$, evaluated at $t = -\beta/\theta$:

$$\mathbb{E}_\varphi\left[\exp\left(-\frac{\beta}{\theta} s_j^\top \varepsilon\right)\right] = \exp\left(\frac{\beta^2}{2\theta^2}\ell_j\right).$$

Dividing gives the normalized form

$$g_j^*(\varepsilon) = \exp\left(-\frac{\beta}{\theta} s_j^\top \varepsilon - \frac{\beta^2}{2\theta^2}\ell_j\right).$$

To see that this equals the Radon-Nikodym derivative $dN(-(\beta/\theta)s_j, I_k)/dN(0, I_k)$, observe that for $P = N(\mu, I_k)$ and $Q = N(0, I_k)$,

$$\frac{dP}{dQ}(\varepsilon) = \exp\left(\mu^\top \varepsilon - \tfrac{1}{2}|\mu|^2\right),$$

which follows by taking the ratio of the two Gaussian densities. Setting $\mu = -(\beta/\theta)s_j$ gives $\mu^\top \varepsilon - \frac{1}{2}|\mu|^2 = -(\beta/\theta)s_j^\top \varepsilon - (\beta^2/2\theta^2)\ell_j$, which matches g_j^* exactly. Hence the worst-case shock distribution is $N(-(\beta/\theta)s_j, I_k)$. □

The last statement in Proposition 5.2 says that an exponential tilt of $N(0, I_k)$ in direction s_j is equivalent to shifting the mean by $-(\beta/\theta)s_j$.

Remark 5.3 (Relation to Hansen & Sargent (2008*a*)). When $M = 1$ (no mobility), $\Pi = [1]$ and $B = (1-\beta)^{-1}(1-\beta) = 1$, so $b_0 = 1$ and $s_0 = C^\top \lambda + F_0^\top$. Then $(I - \beta\Pi)^{-1} = (1-\beta)^{-1}$ and (39) reduces to

$$\kappa = -\frac{\beta^2}{2(1-\beta)\theta}\|C^\top \lambda + F_0^\top\|^2,$$

which matches their scalar multiplier-preference closed form after translating notation. The worst-case mean shift $-(\beta/\theta)s_0 = -(\beta/\theta)(C^\top \lambda + F_0^\top)$ likewise coincides

with their Gaussian exponential tilt. Our result generalizes these scalar formulas to M quantiles so that the scalar κ becomes a vector $\kappa \in \mathbb{R}^M$, and $(1-\beta)^{-1}$ is replaced by the Markov resolvent $(I-\beta\Pi)^{-1}$.

5.2 Case 2 and Case 3: No affine closed forms

In this section, we substitute the affine conjecture (33) into the Case 2 and Case 3 Bellman equations. In each case, for $M > 1$, the log-$\sum$-exp term is generically not affine in (ζ, c), so the conjecture does not close.

5.2.1 Case 2

Substituting $\bar{v}_j = \mu_j$ into (25) and (26) gives

$$v_p(\zeta, c) = (1-\beta)c_p - \theta \log\left(\sum_{j=0}^{M-1} \Pi_{pj} \exp\left(-\frac{\beta}{\theta}\mu_j(\zeta, c)\right)\right), \tag{40}$$

with distorted transition probabilities

$$\widetilde{\Pi}_{pj}(\zeta, c) = \frac{\Pi_{pj} \exp\left(-\frac{\beta}{\theta}\mu_j(\zeta, c)\right)}{\sum_k \Pi_{pk} \exp\left(-\frac{\beta}{\theta}\mu_k(\zeta, c)\right)}. \tag{41}$$

5.2.2 Case 3

Substituting (35) into (29) gives

$$v_p(\zeta, c) = (1-\beta)c_p - \theta \log\left(\sum_{j=0}^{M-1} \Pi_{pj} \exp\left(-\frac{\beta}{\theta}\mu_j(\zeta, c) + \frac{\beta^2}{2\theta^2}\ell_j\right)\right), \tag{42}$$

with distorted transition probabilities

$$\widetilde{\Pi}_{pj}(\zeta, c) = \frac{\Pi_{pj} M_j(\zeta, c)}{\sum_k \Pi_{pk} M_k(\zeta, c)} = \frac{\Pi_{pj} \exp\big(\eta_j(\zeta, c)\big)}{\sum_k \Pi_{pk} \exp\big(\eta_k(\zeta, c)\big)}, \tag{43}$$

where $\eta_j(\zeta, c) := -\frac{\beta}{\theta}\mu_j(\zeta, c) + \frac{\beta^2}{2\theta^2}\ell_j$. The first equality integrates (30) over ε, the second substitutes (35). The distorted mobility is a softmax as in (26) but with μ_j replaced by $\mu_j - (\beta/2\theta)\ell_j$, and the worst-case shock distribution is $N\big(-(\beta/\theta)s_j, I_k\big)$ for destination j, the same shift as in Case 1.

Since the value function is not affine, we cannot solve for the coefficients λ_j, b_j, and κ_j algebraically. Instead, the value function can be computed numerically by iterating Φ_θ. To guarantee convergence, we need to establish that Φ_θ is a contraction mapping on a suitable function space. Let $b\mathsf{S}$ denote the space of bounded measurable functions on S, equipped with the sup norm $\|\cdot\|_\infty$.

Proposition 5.4 (Contraction)**.** *Suppose the per-period payoff is bounded on S. Then Φ_θ is a self-map and a contraction of modulus β on $(b\mathsf{S}, \|\cdot\|_\infty)$.*[3]

Proof. We show that Φ_θ is order preserving and satisfies subadditivity with constant β:

Fix $v, w \in b\mathsf{S}$ and $v \leqslant w$ pointwise. Then $-(\beta/\theta)\, v_{p'}(\zeta', c') \geqslant -(\beta/\theta)\, w_{p'}(\zeta', c')$, so $\exp(-(\beta/\theta)\, v_{p'}) \geqslant \exp(-(\beta/\theta)\, w_{p'})$. Taking $\mathbb{E}_Q$-expectations preserves the inequality, and $x \mapsto -\theta \log x$ is decreasing, so $\mathcal{T}_\theta[\beta v] \leqslant \mathcal{T}_\theta[\beta w]$ and hence $\Phi_\theta v \leqslant \Phi_\theta w$.

Regarding subadditivity, for $a \geqslant 0$,

$$\begin{aligned}
\mathcal{T}_\theta[\beta(v+a)](p,\zeta,c) &= -\theta \log \mathbb{E}_Q\left[\exp\left(-\frac{\beta}{\theta}(v_{p'} + a)\right) \,\middle|\, p, \zeta, c\right] \\
&= -\theta \log\left(e^{-\beta a/\theta}\, \mathbb{E}_Q\left[\exp\left(-\frac{\beta}{\theta} v_{p'}\right) \,\middle|\, p, \zeta, c\right]\right) \\
&= \beta a + \mathcal{T}_\theta[\beta v](p,\zeta,c).
\end{aligned}$$

Hence $(\Phi_\theta(v+a))_p = (1-\beta)\, c_p + \mathcal{T}_\theta[\beta(v+a)] = (\Phi_\theta v)_p + \beta a$. Since equality holds, we have in particular $\Phi_\theta(v+a) \leqslant \Phi_\theta v + \beta a$.

Combining both properties, for any $v, w \in b\mathsf{S}$,

$$\Phi_\theta v = \Phi_\theta(w + v - w) \leqslant \Phi_\theta(w + \|v - w\|_\infty) \leqslant \Phi_\theta w + \beta\, \|v - w\|_\infty,$$

[3]The argument below is a standard application of Blackwell's sufficient conditions (Blackwell 1965). For finite state spaces, see Sargent & Stachurski (2025, Lemma 2.2.4); for the general case, see Stokey et al. (1989, Theorem 3.3).

where the first inequality uses order preservation and $v - w \leqslant \|v - w\|_\infty$, and the second uses the shift inequality. Reversing the roles of v and w gives $\|\Phi_\theta v - \Phi_\theta w\|_\infty \leqslant \beta \|v - w\|_\infty$.

For the self-map property, if $\|v\|_\infty \leqslant B$ then $|\mathcal{T}_\theta[\beta v]| \leqslant \beta B$, so $\|\Phi_\theta v\|_\infty \leqslant (1 - \beta) \sup |c_p| + \beta B < \infty$. □

Under the linear–Gaussian model, $c' = c + D\zeta + F\varepsilon$ with Gaussian innovations, so the continuous state is unbounded and the boundedness hypothesis does not hold on the full state space $\mathsf{S} = \{0, \ldots, M-1\} \times \mathbb{R}^n \times \mathbb{R}^M$. For the numerical implementation (Section 6.6) we work on a truncated finite grid, where the hypothesis is satisfied and the proposition applies.

5.3 Certainty equivalents and compensating differences

This section defines certainty equivalent value functions and compensating differences for all four cases, extending Barillas et al. (2009) and Sargent et al. (2026*a*) to incorporate both cross-quantile mobility ($\Pi \neq I$) and robustness ($\theta < \infty$). Closed-form expressions are available for Cases 0 and 1, where the value function is affine. Cases 2 and 3 require numerical computation.

The consumption growth (2) and state dynamics (1) imply by induction that log consumption at quantile p satisfies, for $h \geqslant 1$,

$$c_{p,t+h} = c_{p,t} + \psi_{p,h}^\top \zeta_t + \sum_{i=0}^{h-1} J_{p,i}\, \varepsilon_{t+h-i}, \tag{44}$$

where the cumulative drift and shock loadings are

$$\psi_{p,h}^\top = \sum_{m=0}^{h-1} D_p A^m, \tag{45}$$

$$J_{p,i} = F_p + \sum_{\ell=0}^{i-1} D_p A^\ell C, \quad i \geqslant 0. \tag{46}$$

In (46), $J_{p,i} \in \mathbb{R}^{1\times k}$ gives the cumulative loading of $c_{p,t+h}$ on the shock ε_{t+h-i} that ar-

rived i periods earlier: F_p is the direct effect, and the sum $\sum_\ell D_p A^\ell C$ captures indirect effects propagated through the state.

5.3.1 Certainty equivalent and compensating difference

Following Obstfeld (1994) and Tallarini (2000), define the risk-free certainty equivalent path for quantile j as the deterministic log-consumption path whose implied consumption levels match the conditional level means of the risky path:

$$\exp(c^{\mathrm{CE}}_{j,t+h}) = \mathbb{E}_\varphi \left[\exp(c_{j,t+h}) \mid \zeta_t, c_t\right], \qquad h \geqslant 0. \tag{47}$$

Since $c_{j,t+h}$ from (44) is conditionally Gaussian with mean $c_{j,t} + \psi_{j,h}^\top \zeta_t$ and variance $\sigma^2_{j,h} := \sum_{i=0}^{h-1} \|J_{j,i}\|^2$, the log-normal moment formula gives

$$c^{\mathrm{CE}}_{j,t+h} = c_{j,t} + \psi_{j,h}^\top \zeta_t + \tfrac{1}{2}\sigma^2_{j,h}, \qquad h \geqslant 1. \tag{48}$$

The CE value function w_p evaluates the non-robust preference recursion on the deterministic CE path (48). It serves as a common benchmark for all four robustness cases. We define w_p under non-robust expected-value preferences ($\theta_\varepsilon = \theta_\Pi = \infty$), so the household evaluates future utility without any concern for model misspecification. On the CE path, consumption is deterministic given (ζ_t, c_t), so the risk-sensitive transformation of next-period continuation values collapses to simple discounting: $-\theta_\varepsilon \log \mathbb{E}_\varphi[\exp(-\beta w_j/\theta_\varepsilon)] = \beta w_j$. The household still transitions across consumption quantiles according to the transition matrix Π. The CE value function is therefore

$$w_p(\zeta, c) = (1-\beta) \sum_{h=0}^{\infty} \beta^h \sum_{j=0}^{M-1} [\Pi^h]_{pj}\, \bar{c}_{j,h}(\zeta, c), \tag{49}$$

where $\bar{c}_{j,h}(\zeta, c) = c_j + \psi_{j,h}^\top \zeta + \frac{1}{2}\sigma^2_{j,h}$ is the horizon-h CE consumption from (48), and $\sigma^2_{j,h} = \sum_{i=0}^{h-1} \|J_{j,i}\|^2$ with $J_{j,i} = F_j + \sum_{\ell=0}^{i-1} D_j A^\ell C$ from (46). Because w_p is a linear function of the CE path, it has a closed-form expression for all four cases (Section 5.3.2).

The compensating difference $\alpha^{(m)} \in \mathbb{R}^M$ is the uniform consumption shift that

equates the risky value to the CE value:

$$v_p^{(m)}(\zeta, c) = w_p(\zeta,\ c - \alpha_p^{(m)}\mathbf{1}), \qquad p = 0, \dots, M-1. \tag{50}$$

5.3.2 Cases 0 and 1

For Cases 0 and 1, the CE recursion is linear in w and both value functions are affine in (ζ, c). The Case 0 value function from Proposition 5.1 is $v_p(\zeta, c) = \lambda_p^\top \zeta + b_p^\top c$. Evaluating the linear CE recursion on the paths (48) gives

$$w_p(\zeta, c) = v_p(\zeta, c) + \frac{1-\beta}{2} \sum_{h=1}^{\infty} \beta^h \left[\Pi^h \sigma_h^2\right]_p, \tag{51}$$

where $\sigma_h^2 := (\sigma_{0,h}^2, \dots, \sigma_{M-1,h}^2)^\top$. We assume the discounted series in (51) is finite; a sufficient condition is stability of the state dynamics, e.g. $\rho(A) < 1$, which makes the loadings in (45)–(46) bounded and $\sigma_{j,h}^2$ grow at most linearly in h.

We first prove a useful lemma: shifting all consumption ranks by a common constant shifts the value function by the same amount.

Lemma 5.5 (Uniform shift)**.** *For any scalar a and any case $m = 0, 1, 2, 3$,*

$$v_p^{(m)}(\zeta,\ c + a\mathbf{1}) = v_p^{(m)}(\zeta, c) + a, \qquad w_p(\zeta,\ c + a\mathbf{1}) = w_p(\zeta, c) + a. \tag{52}$$

Proof. Let Φ denote the operator defined by the right-hand side of (32). We show that Φ maps functions satisfying (52) into functions satisfying (52); since $v = \Phi v$, the fixed point inherits the property.

Let f_j satisfy $f_j(\zeta', c' + a\mathbf{1}) = f_j(\zeta', c') + a$ for all j. Define $\mathcal{R}_\varepsilon f_j$ by (31) evaluated at f_j. Because $f_j(\zeta', c' + a\mathbf{1}) = f_j(\zeta', c') + a$,

$$\begin{aligned}(\mathcal{R}_\varepsilon f_j)(\zeta, c + a\mathbf{1}) &= -\theta_\varepsilon \log \mathbb{E}\left[e^{-\frac{\beta}{\theta_\varepsilon} f_j(\zeta', c' + a\mathbf{1})}\right] \\ &= -\theta_\varepsilon \log\left(e^{-\frac{\beta a}{\theta_\varepsilon}} \mathbb{E}\left[e^{-\frac{\beta}{\theta_\varepsilon} f_j(\zeta', c')}\right]\right) = (\mathcal{R}_\varepsilon f_j)(\zeta, c) + \beta a.\end{aligned}$$

Substituting into (32),

$$\begin{aligned}(\Phi f)_p(\zeta, c + a\mathbf{1}) &= (1-\beta)(c_p + a) - \theta_\Pi \log \sum_j \Pi_{pj} e^{-(\mathcal{R}_\varepsilon f_j)(\zeta, c+a\mathbf{1})/\theta_\Pi} \\ &= (1-\beta)(c_p + a) - \theta_\Pi \log \left(e^{-\beta a/\theta_\Pi} \sum_j \Pi_{pj} e^{-(\mathcal{R}_\varepsilon f_j)(\zeta, c)/\theta_\Pi} \right) \\ &= (1-\beta)(c_p + a) + \beta a - \theta_\Pi \log \sum_j \Pi_{pj} e^{-(\mathcal{R}_\varepsilon f_j)(\zeta, c)/\theta_\Pi} \\ &= (\Phi f)_p(\zeta, c) + a.\end{aligned}$$

For w_p, the same argument applies to the linear recursion (49). The CE path satisfies $c_p^{\mathrm{CE}}(c + a\mathbf{1}) = c_p^{\mathrm{CE}}(c) + a$, since from (48) the variance term $\frac{1}{2}\sigma_{j,h}^2$ does not depend on c. Substituting into (49), $w_p(\zeta, c + a\mathbf{1}) = w_p(\zeta, c) + a$. □

Proposition 5.6 (Compensating difference, Case 0)**.** *The compensating difference* $\alpha^{(0)} \in \mathbb{R}^M$ *is*

$$\alpha^{(0)} = \frac{1-\beta}{2} \sum_{h=1}^{\infty} \beta^h \, \Pi^h \sigma_h^2. \tag{53}$$

When $\Pi = I$, $\alpha_p^{(0)} = \frac{\beta}{2} \sum_{i=0}^{\infty} \beta^i \, \|J_{p,i}\|^2$.

Proof. By Lemma 5.5, $w_p(\zeta, c - \alpha_p^{(0)}\mathbf{1}) = w_p(\zeta, c) - \alpha_p^{(0)}$. Setting $v_p^{(0)} = w_p - \alpha_p^{(0)}$ gives (53). For $\Pi = I$: $\alpha_p^{(0)} = \frac{1-\beta}{2} \sum_{h \geqslant 1} \beta^h \sum_{i=0}^{h-1} \|J_{p,i}\|^2 = \frac{1-\beta}{2} \sum_{i \geqslant 0} \|J_{p,i}\|^2 \sum_{h > i} \beta^h = \frac{1-\beta}{2} \sum_{i \geqslant 0} \|J_{p,i}\|^2 \cdot \frac{\beta^{i+1}}{1-\beta} = \frac{\beta}{2} \sum_{i \geqslant 0} \beta^i \|J_{p,i}\|^2$. □

Remark 5.7. When $\Pi = I$, the compensating difference $\alpha_p^{(0)} = \frac{\beta}{2} \sum_{i \geqslant 0} \beta^i \|J_{p,i}\|^2$ is identical to equation (35) in Sargent et al. (2026*a*). When $M = 1$ with $D = 0$ and $F = \sigma_\varepsilon$ (so that consumption is a random walk), $J_{0,i} = \sigma_\varepsilon$ for all i and $\alpha^{(0)} = \frac{\beta}{2(1-\beta)} \sigma_\varepsilon^2$, which recovers the Obstfeld–Tallarini compensating difference.[4]

[4] Obstfeld (1994) equation (4) gives lifetime utility under the martingale process as $U_0 = \frac{1}{1-\beta}\left[\log C_0 + \frac{\beta}{1-\beta}(\mu - \frac{1}{2}\sigma_\varepsilon^2)\right]$ at $\gamma = 1$, from which the welfare cost of consumption risk is $\frac{\beta}{2(1-\beta)}\sigma_\varepsilon^2$. Tallarini (2000) equation (21) writes the compensating variation for the random-walk model as $\hat{\lambda}_{\mathrm{RW}} = \frac{\sigma_\varepsilon^2}{2} \frac{\beta}{1-\beta} \chi$, which at $\chi = 1$ (unit risk aversion, log utility) gives $\frac{\beta}{2(1-\beta)}\sigma_\varepsilon^2$. See also Sargent et al. (2026*a*) equation (22).

Under Case 1, the value function from Proposition 5.2 is $v_p^{(1)}(\zeta, c) = \lambda_p^\top \zeta + b_p^\top c + \kappa_p$, with the same slopes (B, Λ) as Case 0 and intercept $\kappa = -\frac{\beta^2}{2\theta}(I - \beta\Pi)^{-1}\Pi\ell$ from (39). Setting $v_p^{(1)} = w_p - \alpha_p^{(1)}$ and matching the constant terms gives the compensating difference for Case 1:

$$\alpha^{(1)} = \alpha^{(0)} - \kappa, \tag{54}$$

where $\kappa_p \leqslant 0$ for all p, so $\alpha_p^{(1)} \geqslant \alpha_p^{(0)}$.

5.3.3 Cases 2 and 3

For Cases 2 and 3, the mobility penalty $\theta_\Pi < \infty$ and v_p is no longer affine. The CE value w_p remains the same linear function as in Cases 0 and 1, since it is defined under non-robust preferences.

Denote $v_p^{(m)}$ as the Case m value function ($m = 2, 3$). For a fixed outer mobility penalty θ_Π, Case 3 relates to Case 2 by adding the shock-robust inner aggregation, which lowers the continuation value and hence the current value $v_p^{(3)} \leqslant v_p^{(2)}$. The compensating difference at state (p, ζ, c) is

$$\alpha_p^{(m)}(\zeta, c) = w_p(\zeta, c) - v_p^{(m)}(\zeta, c), \qquad m = 2, 3. \tag{55}$$

Unlike Cases 0 and 1, $\alpha_p^{(m)}$ depends on the state (ζ, c). Moreover, the two terms in (55) shift uniformly with c (Lemma 5.5). Applying the lemma with $a = -\alpha_p^{(m)}$ to the defining equation (50):

$$v_p^{(m)}(\zeta, c) = w_p\left(\zeta,\ c - \alpha_p^{(m)}\mathbf{1}\right) = w_p(\zeta, c) - \alpha_p^{(m)},$$

which gives (55). Once $v^{(m)}$ and w are computed on the same grid, $\alpha_p^{(m)}$ is a pointwise subtraction without the need to find the root of the nonlinear equation (50).

For a fixed outer mobility penalty θ_Π, the ordering between Cases 3 and 2 follows from $v_p^{(3)} \leqslant v_p^{(2)}$. This holds by a monotone-convergence argument: since the shock-robust aggregation satisfies $\mathcal{R}_\varepsilon v_j \leqslant \beta\, \mathbb{E}_\varphi[v_j]$ (Jensen's inequality), the Bellman operator satisfies $\Phi_3 v \leqslant \Phi_2 v$ pointwise for all v. Iterating from $v^{(0)} = 0$, we get $\Phi_3^n 0 \leqslant \Phi_2^n 0$ for all n by induction. Taking the limit gives $v^{(3)} \leqslant v^{(2)}$. Hence Case 3

has a lower continuation value than Case 2:

$$\alpha_p^{(3)}(\zeta, c) - \alpha_p^{(2)}(\zeta, c) = v_p^{(2)}(\zeta, c) - v_p^{(3)}(\zeta, c) \geqslant 0. \tag{56}$$

This is the state-dependent analogue of $\alpha^{(1)} - \alpha^{(0)} = -\kappa \geqslant 0$ from (54).

Lemma 5.5 also implies $\alpha_p^{(m)}(\zeta, c+a\mathbf{1}) = \alpha_p^{(m)}(\zeta, c)$, so the compensating difference is invariant to the common level of c and depends on c only through the demeaned cross-section $\tilde{c} := c - \bar{c}\,\mathbf{1}$, where $\bar{c} := \frac{1}{M}\sum_p c_p$. In general, $\alpha_p^{(m)}$ still depends on the reduced state ζ. In the DMD implementation (Section 6.4), $\tilde{c}_t$ is determined by ζ_t, so $\alpha_p^{(m)}$ reduces to a function of ζ_t alone.

When $\rho(A) < 1$, the reduced state ζ_t is stationary, so we report compensating differences averaged over its stationary distribution:

$$\bar{\alpha}_p^{(m)} := \mathbb{E}_{\mu_x}\left[\alpha_p^{(m)}(\zeta_t)\right], \qquad m = 2, 3, \tag{57}$$

where μ_x denotes the stationary distribution of the reduced continuous state ζ_t. This gives one number per quantile p, making it comparable to the constants $\alpha_p^{(0)}$ and $\alpha_p^{(1)}$. A scalar aggregate is $\bar{\alpha}^{(m)} = \sum_p \pi_p\, \bar{\alpha}_p^{(m)}$, where π is the stationary distribution of Π.

6 Estimation

This section describes how we estimate the approximating model (1)–(3) from Consumer Expenditure Survey (CEX) microdata from 1990–2024, calibrate the robustness parameters $(\theta_\varepsilon, \theta_\Pi)$ via detection-error probabilities, and compute compensating differences for all four cases. We report results under two alternative ranking schemes, based on consumption and private income, to study how the choice of grouping variable shapes the welfare implications of robustness.

6.1 Data and quantile assignments

We use the CEX Interview Survey, which interviews each consumer unit (household) up to five times at quarterly intervals. From each interview we observe private in-

come (pre-tax, pre-transfer), post-tax income (after taxes and transfers), and total consumption expenditure. We use the definitions of these variables from Sargent et al. (2026*a*), which allow us to track them in short panels and estimate transition matrices for each series. All nominal variables are deflated by the Personal Consumption Expenditures Price Index (PCEPI).

Each quarter, households are ranked into $M = 100$ equally weighted bins using the CEX final interview weight (`FINLWT21`). Let $G_\ell^t \in \{0, \ldots, M-1\}$ denote the bin assigned to household ℓ at date t. We consider two alternative ranking variables:

- **Consumption ranking.** Bin membership is determined by total consumption expenditure. The observable state vector uses consumption cross-sections only, yielding a state of dimension $1 + M$.
- **Private-income ranking.** Bin membership is determined by private income alone. Under this scheme, the households assigned to the p-th post-tax-income and consumption cells are those whose private income falls in the p-th private-income bin. This common ranking ensures that a single transition matrix Π governs mobility across all three variables. The observable state vector stacks all three cross-sections, yielding a state of dimension $1 + 3M$.[5]

6.2 Transition matrix

We link consecutive CEX interviews for the same household to form one-quarter panel pairs $(\ell, t, t{+}1)$. The transition matrix is estimated as the weighted relative frequency of households moving from bin p at date t to bin q at date $t{+}1$:

$$\widehat{\Pi}_{pq} = \frac{\sum_\ell w_\ell \, \mathbb{1}\{G_\ell^t = p\} \, \mathbb{1}\{G_\ell^{t+1} = q\}}{\sum_\ell w_\ell \, \mathbb{1}\{G_\ell^t = p\}}, \qquad p, q = 0, \ldots, M-1, \tag{58}$$

where w_ℓ is the interview weight. This estimator imposes the simplifying assumption that transition probabilities are time-invariant and homogeneous across households and survey waves. Figure 1 displays the estimated matrices for three alternative rankings, by private income, post-tax income, and consumption, each with $M = 100$

[5] This differs from Sargent et al. (2026*a*), where each variable is ranked independently.

bins. All three exhibit strong diagonal persistence, but the off-diagonal structure differs markedly across rankings. The consumption matrix (right panel) is less concentrated around the diagonal than the private-income matrix (left panel), reflecting substantially more cross-quantile mixing in consumption than in income.

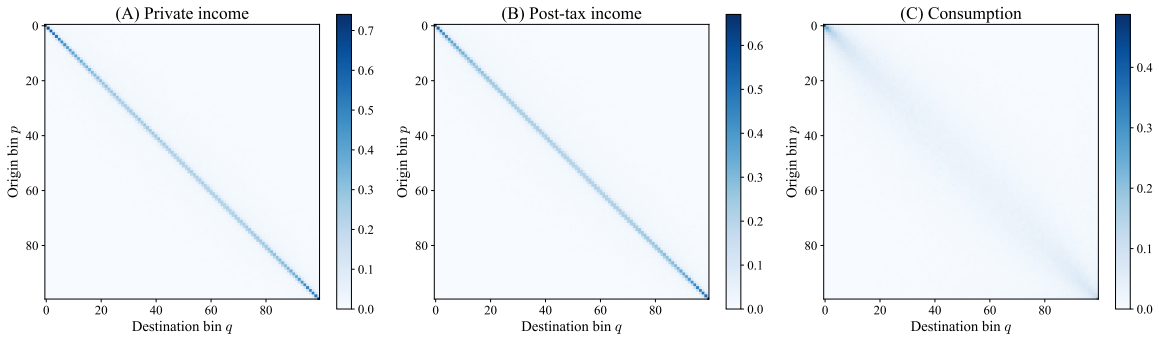


Figure 1: Estimated transition matrices $\widehat{\Pi}_{pq}$ from CEX panel with $M = 100$ bins. Left: private-income rank. Center: post-tax-income rank. Right: consumption rank.

6.3 Additive functional and state construction

For each variable $u \in \{m, d, c\}$ (private income, post-tax income, consumption), let $u_{p,t}$ denote the log Chisini mean of the p-th bin at date t, computed as the weighted average of log levels within the bin. The aggregate log level is $Y_t := M^{-1} \sum_{p=0}^{M-1} u_{p,t}^{\text{rank}}$, where u^{rank} is the ranking variable (consumption or private income), and the demeaned cross-sections are $\widetilde{u}_t := u_t - Y_t \, \mathbf{1}_M$.

The choice of ranking variable determines the composition of the observable state vector:

Under **consumption ranking**, the state stacks aggregate consumption growth with demeaned consumption cross-sections:

$$y_t^{(c)} := \begin{pmatrix} \Delta Y_t - \nu \\ \widetilde{c}_t - \boldsymbol{\nu}^{(c)} \end{pmatrix} \in \mathbb{R}^{1+M}, \tag{59}$$

where ν is the scalar aggregate growth trend and $\boldsymbol{\nu}^{(c)}$ is the time-series mean of the demeaned cross-section. This $(1+M)$-dimensional state captures cross-sectional

consumption dispersion and its aggregate dynamics. It is a direct generalization of the representative-agent ($M = 1$) case studied by Obstfeld (1994) and Tallarini (2000), where the state is one-dimensional and given by aggregate consumption growth alone.

Under **private-income ranking**, following Sargent et al. (2026*a*), we stack the demeaned cross-sections of all three variables together with aggregate income growth:

$$y_t^{(m)} := \begin{pmatrix} \Delta Y_t - \nu \\ \widetilde{m}_t - \boldsymbol{\nu}^{(m)} \\ \widetilde{d}_t - \boldsymbol{\nu}^{(d)} \\ \widetilde{c}_t - \boldsymbol{\nu}^{(c)} \end{pmatrix} \in \mathbb{R}^{1+3M}. \tag{60}$$

The richer $(1+3M)$-dimensional state captures the joint cross-sectional dynamics of income, taxes, and consumption, allowing the reduced-rank estimator to exploit covariation across all three variables.

In both cases we model the demeaned state as a first-order vector autoregression

$$y_{t+1} = B\, y_t + a_{t+1}, \qquad \mathbb{E}[a_{t+1}\, a_{t+1}^\top] = \Omega, \tag{61}$$

following Sargent et al. (2026*a*). This specification implies that the aggregate variable Y_t is an additive functional whose decomposition separates $Y_t - Y_0$ into a deterministic trend νt, a martingale $\sum_{j=1}^{t} H\, a_j$, a stationary component $-g\, y_t$, and a constant $g\, y_0 + Y_0$, with coefficients $g = e_1^\top B(I-B)^{-1}$ and $H = e_1^\top + g$ (Sargent et al. (2026*a*), Proposition 1). The quantile-specific decomposition of Sargent et al. (2026*a*) (Corollary 1) implies that each quantile series $u_{p,t}$ (for $u \in \{m, d, c\}$) shares the same trend and martingale but has a quantile-specific stationary component and constant.

To connect (61) to the state-space form (1)–(2), we follow the augmented-state construction in Sargent et al. (2026*a*) (Section 5.2). Define $\boldsymbol{\zeta}_t = [1;\, y_t]$ and set

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 0 \\ J \end{bmatrix}, \tag{62}$$

where J is the Cholesky factor of Ω, so that $a_{t+1} = J\varepsilon_{t+1}$ with $\varepsilon_{t+1} \sim N(0, I_k)$. The

consumption loadings are

$$D_q = \begin{bmatrix} \nu & e_1^\top B + e_q^{(c)}(B - I) \end{bmatrix}, \qquad F_q = (e_1^\top + e_q^{(c)})\, J, \tag{63}$$

where $e_q^{(c)}$ is a row vector selecting the q-th consumption entry of y_t. These follow from differencing the quantile-specific additive-functional decomposition

$$c_{q,t} = \alpha_q + \nu\, t + \sum_{s=1}^{t} H\, a_s - (g - e_q^{(c)})\, y_t, \tag{64}$$

where $\alpha_q := Y_0 + g\, y_0 + \nu_q^{(c)}$ absorbs the initial level, ν is the common drift, $H := e_1^\top + g$ is the common martingale loading so that each quantile shares the same trend and martingale, and the last term captures the quantile-specific stationary component. Differencing (64) and using $y_{t+1} = By_t + a_{t+1}$ together with $g(I - B) = e_1^\top B$ gives (63). We estimate B using Dynamic Mode Decomposition (Section 6.4), and all welfare calculations are performed in the resulting N-dimensional projected space.

6.4 Dynamic Mode Decomposition

Because the state dimension exceeds the sample size T in the private-income ranking, we consistently estimate B via Dynamic Mode Decomposition (DMD), a dimension-reduction technique that approximates the VAR coefficient matrix via SVD truncation (see, e.g., Schmid (2010), Kutz et al. (2016), Sargent et al. (2026*b*)).

Form the data matrices $Y_0 := [y_1, \ldots, y_{T-1}]$ and $Y_1 := [y_2, \ldots, y_T]$. Take the SVD $Y_0 = U\Sigma V^\top$ and retain the leading N triplets (U_N, Σ_N, V_N). The rank-N estimator is

$$\widetilde{B} := U_N^\top Y_1\, V_N\, \Sigma_N^{-1} \in \mathbb{R}^{N\times N}, \qquad \widehat{B} := U_N\, \widetilde{B}\, U_N^\top. \tag{65}$$

The innovation covariance is $\widehat{\Omega} = (T-1)^{-1}\widehat{E}\,\widehat{E}^\top$, where $\widehat{E} = Y_1 - \widehat{B}\,Y_0$ is the residual matrix. With $X_0 = U_N^\top Y_0 = \Sigma_N V_N^\top$ and $X_1 = U_N^\top Y_1$, we have $\widetilde{B} = X_1 X_0^+$, the least-squares VAR coefficient in the reduced coordinates.[6] We set $N = 5$ throughout; the first

[6]The $N \times N$ estimator $\widetilde{B}$ is the standard DMD algorithm of Schmid (2010), which projects onto the leading left singular vectors of Y_0 and fits dynamics in the reduced coordinates. The full-state

five singular values of Y_0 capture the bulk of the variation in both ranking schemes. Figure 12 in Appendix B plots the variance share and the out-of-sample RMSE for one-step-ahead consumption growth as functions of the truncation rank N for both ranking schemes, confirming that $N = 5$ lies on the plateau where out-of-sample RMSE is minimized under consumption ranking; the private-income minimum is at $N = 6$, but the improvement over $N = 5$ is negligible.

Project the observable onto the leading left singular vectors: $\widehat{x}_t := U_N^\top y_t \in \mathbb{R}^N$. The reduced dynamics are

$$\widehat{x}_{t+1} = \widetilde{B}\,\widehat{x}_t + U_N^\top a_{t+1}, \tag{66}$$

with innovation covariance $Q_x := U_N^\top \widehat{\Omega}\, U_N$. Let L denote the Cholesky factor of Q_x, so that $U_N^\top a_{t+1} = L\,\varepsilon_{t+1}$ with $\varepsilon_{t+1} \sim N(0, I_N)$. The reduced system (66) captures the demeaned continuous-state dynamics. To recover the full state-space form (1)–(2), augment the reduced state with a leading one, $\zeta_t^{(N)} := [1; \widehat{x}_t]$, so that $A_N = \left[\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widetilde{B} \end{smallmatrix}\right]$ and $C_N = [0; L]$.

The reduced consumption growth equation is

$$\Delta c_{q,t+1} = \nu + D_q^{(N)}\,\widehat{x}_t + F_q^{(N)}\,\varepsilon_{t+1}, \tag{67}$$

where $D_q^{(N)}$ and $F_q^{(N)}$ are the loadings on $\widehat{x}_t$ and ε_{t+1} obtained from (63) applied to the reduced coordinates:

$$D_q^{(N)} = e_1^\top U_N \widetilde{B} + e_q^{(c)\top} U_N (\widetilde{B} - I_N), \qquad F_q^{(N)} = (e_1^\top + e_q^{(c)\top})\, U_N L. \tag{68}$$

The aggregate growth in the reduced state is

$$\Delta Y_{t+1} = \nu + e_1^\top U_N \widetilde{B}\,\widehat{x}_t + e_1^\top U_N L\,\varepsilon_{t+1}. \tag{69}$$

Subtracting the aggregate decomposition $Y_t = Y_0 + g\,y_0 + \nu\,t + \sum_{s=1}^{t} H\,a_s - g\,y_t$ from (64) gives $c_{q,t} = Y_t + \nu_q^{(c)} + e_q^{(c)\top} y_t$. Projecting onto the reduced coordinates via $y_t \mapsto U_N\,\widehat{x}_t$, define $\tilde{c}_{q,t} := \nu_q^{(c)} + e_q^{(c)\top} U_N\,\widehat{x}_t$, so that $c_{q,t} = Y_t + \tilde{c}_{q,t}$. By Lemma 5.5, writing $\upsilon_p(Y, x) =$

reconstruction $\widehat{B} = U_N \widetilde{B} U_N^\top$ is the rank-N DMD operator studied in Tu et al. (2014) and Sargent et al. (2026*b*).

$Y + \tilde{v}_p(x)$ reduces the Bellman equation to a recursion in the N-dimensional reduced state alone:

$$\tilde{v}_p(x) = (1-\beta)\,\tilde{c}_p(x) - \theta_\Pi \log \sum_{j=0}^{M-1} \Pi_{pj} \exp\left(-\frac{(\mathcal{R}_\varepsilon \tilde{v}_j)(x)}{\theta_\Pi}\right), \tag{70}$$

where

$$(\mathcal{R}_\varepsilon \tilde{v}_j)(x) = -\theta_\varepsilon \log \mathbb{E}_\varphi\left[\exp\left(-\frac{\beta}{\theta_\varepsilon}\left(\Delta Y(x,\varepsilon) + \tilde{v}_j(x')\right)\right)\right], \qquad x' = \widetilde{B}\,x + L\,\varepsilon. \tag{71}$$

All welfare calculations below are performed in this N-dimensional space.

6.5 Calibrating robustness via detection-error probabilities

We calibrate the robustness parameters θ_ε and θ_Π so that the approximating model Q and a worst-case distortion $\widetilde{Q}$ are statistically difficult to distinguish, following Anderson et al. (2003) and Barillas et al. (2009).

Given a sample of length T, the per-period log-likelihood ratio between the worst-case and approximating models at date $t+1$ is

$$\ell_{t+1} = \underbrace{\log \frac{\widetilde{\Pi}_{p_t,p_{t+1}}(x_t)}{\Pi_{p_t,p_{t+1}}}}_{\text{mobility LLR}} + \underbrace{\log g^*_{p_{t+1}}(\varepsilon_{t+1} \mid x_t)}_{\text{shock LLR}}, \tag{72}$$

where the state-dependent distorted transition probabilities are

$$\widetilde{\Pi}_{pj}(x) = \frac{\Pi_{pj} \exp\left(-\dfrac{(\mathcal{R}_\varepsilon \tilde{v}_j)(x)}{\theta_\Pi}\right)}{\sum_k \Pi_{pk} \exp\left(-\dfrac{(\mathcal{R}_\varepsilon \tilde{v}_k)(x)}{\theta_\Pi}\right)}, \tag{73}$$

and the distorted shock density conditional on destination j is

$$g_j^*(\varepsilon \mid x) = \frac{\exp\left(-\dfrac{\beta}{\theta_\varepsilon}\big(\Delta Y(x,\varepsilon) + \tilde{v}_j(x')\big)\right)}{\mathbb{E}_\varphi\left[\exp\left(-\dfrac{\beta}{\theta_\varepsilon}\big(\Delta Y(x,\tilde{\varepsilon}) + \tilde{v}_j(\tilde{x}')\big)\right)\right]}, \tag{74}$$

with $x' = \widetilde{B}x + L\varepsilon$ and $\tilde{x}' = \widetilde{B}x + L\tilde{\varepsilon}$. The cumulative log-likelihood ratio is $L_T = \sum_{t=0}^{T-1} \ell_{t+1}$. In Case 1, $\widetilde{\Pi} = \Pi$ and only the shock component is active; in Case 2, $g_j^* \equiv 1$ and only the mobility component is active; in Case 3, both components are active. In the affine Case 1 the shock distortion reduces to the Gaussian mean shift $m_j = -(\beta/\theta_\varepsilon)\, s_j$ (Proposition 5.2), but in the nonlinear Cases 2 and 3 the distortions (73)–(74) are state-dependent.

The detection-error probability is

$$\bar{p}(\theta) := \tfrac{1}{2}\big[\Pr(L_T > 0 \mid Q) + \Pr(L_T \leqslant 0 \mid \widetilde{Q})\big]. \tag{75}$$

When $\bar{p}$ is close to $\frac{1}{2}$, an agent with equal priors on Q and $\widetilde{Q}$ selects the wrong model nearly half the time; the distortion is then small enough to be empirically plausible.

Computationally, the detection-error simulation uses the same quadrature approximation as the value function iteration (VFI) that is used to compute the fixed-point problem in Cases 2 and 3. If the Bellman equation uses quadrature weights $\{(\xi_n, \omega_n)\}_{n=1}^{N_q}$, define $M_j(x) = \sum_n \omega_n \exp\big(-(\beta/\theta_\varepsilon)\big(\Delta Y(x,\xi_n) + \tilde{v}_j(x_n')\big)\big)$ and distorted shock weights $\tilde{\omega}_n^{(j)}(x) = \omega_n \exp\big(-(\beta/\theta_\varepsilon)\big(\Delta Y(x,\xi_n) + \tilde{v}_j(x_n')\big)\big)/M_j(x)$. The approximating model is simulated with weights ω_n, the worst-case model with weights $\tilde{\omega}_n^{(j)}(x)$, and the per-period log-likelihood ratio is

$$\ell_{t+1} \approx \log(\widetilde{\Pi}_{p_t,p_{t+1}}(x_t)/\Pi_{p_t,p_{t+1}}) + \log(\tilde{\omega}_{n_t}^{(p_{t+1})}(x_t)/\omega_{n_t}).$$

For each case we solve for the penalty θ such that $\bar{p}(\theta) = \bar{p}_{\text{target}}$ by bisecting over $\log\theta$. Because the worst-case distortions depend on the converged value function, each candidate θ requires solving the Bellman equation before the detection-error probability can be evaluated, i.e., the VFI is nested inside the bisection. The

simulation and convergence details are described in Appendix A. The target is $\bar{p}_{\text{target}} = 0.45$, which is well above the conventional 0.05 threshold for statistical significance. This choice reflects our aim of studying distortions that are difficult to detect empirically.

6.6 Welfare calculations

Cases 0 and 1 are computed from the closed-form coefficients in Propositions 5.1 and 5.2. For Cases 2 and 3, the nonlinear Bellman equations (25) and (29) are solved by VFI on a grid in the N-dimensional reduced state space $\widehat{x}$, using the normalized recursion (70). Proposition 5.4 guarantees convergence on the truncated grid. Grid construction, quadrature, and interpolation details are in Appendix A. Throughout the numerical exercises we set the quarterly discount factor to $\beta = 0.99$.

The CE value function w_p from (49) is the same linear function for all four cases, so the compensating difference $\alpha_p^{(m)}(\zeta, c) = w_p(\zeta, c) - v_p^{(m)}(\zeta, c)$ reduces to $\alpha_p^{(m)} = \alpha_p^{(0)} - \left[v_p^{(m)}(\zeta, c) - v_p^{(0)}(\zeta, c)\right]$, where $\alpha_p^{(0)}$ is the closed-form Case 0 compensating difference (53). For Cases 0 and 1 the value functions are affine (Propositions 5.1 and 5.2), so $\alpha_p^{(m)}$ is a constant independent of the state. For Cases 2 and 3 the compensating difference depends on the reduced state ζ in general. Following (57), we report $\bar{\alpha}_p^{(m)} = \mathbb{E}_{\mu_x}[\alpha_p^{(m)}(\zeta_t)]$, averaged over the stationary distribution of the reduced state. We approximate this expectation by weighting the value function on the VFI grid by the stationary Gaussian density of ζ_t.

To convert to an annualized growth equivalent, define

$$\bar{\nu}_p^{(m)} := \frac{400}{\beta\, w_{\text{scale}}}\, \bar{\alpha}_p^{(m)}, \qquad w_{\text{scale}} := e_1^\top (I - \beta A_N)^{-1} e_1, \tag{76}$$

where $A_N := \left[\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widetilde{B} \end{smallmatrix}\right]$ is the augmented transition matrix in the reduced state space and the factor of 400 converts from quarterly log points to annual percentage points. For Cases 0 and 1, $\bar{\alpha}_p^{(m)} = \alpha_p^{(m)}$ is the closed-form constant.

6.7 Results: consumption-based ranking

Under consumption ranking, the observable state (59) has dimension $1 + M = 101$, and the transition matrix Π governs mobility across consumption percentiles. The rank-5 DMD estimator captures the bulk of the variation in the consumption cross-section. The additive functional decomposition of aggregate consumption yields a trend growth rate $\nu = 0.0025$ (1.01% annualized).

Figure 2 plots log levels of private income, post-tax income, and consumption at the 10th, 50th, and 90th percentiles relative to the initial aggregate level Y_0, under consumption ranking. Because consumption is the ranking variable, cross-sectional dispersion is smallest for consumption relative to the other two variables. The 10th–90th percentile spread in consumption is about 1.7 log points, and the plotted percentile means are modestly more volatile at the top and bottom of the distribution. This pattern foreshadows the mild U-shape in the no-mobility compensating differences reported below.

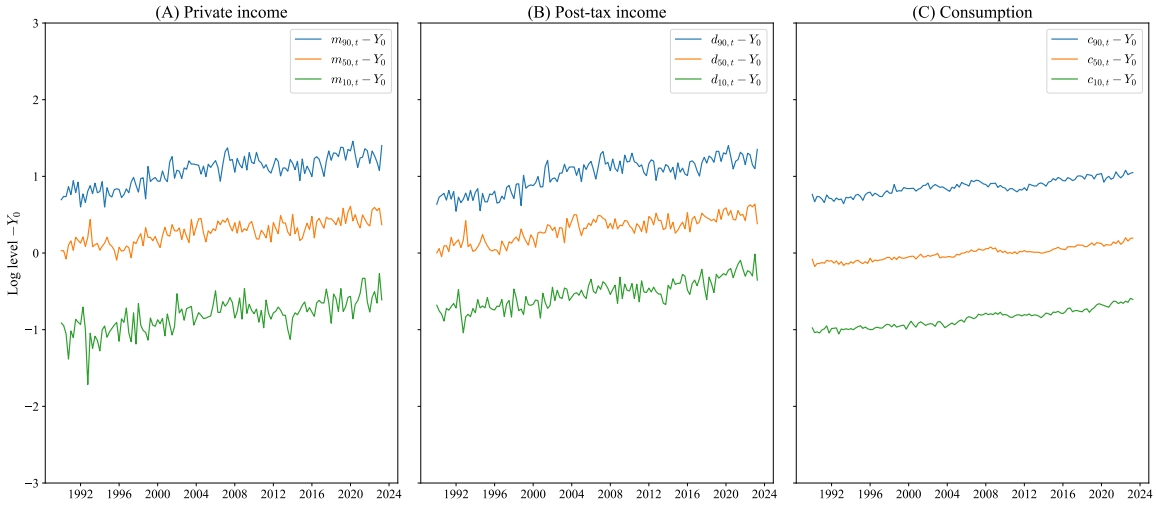


Figure 2: Log levels relative to the aggregate level Y_0 at the 10th, 50th, and 90th percentiles, consumption ranking. (A) Private income, (B) post-tax income, (C) consumption.

When households are ranked by consumption, the no-mobility compensating differences exhibit only a mild U-shape across the distribution (Table 1, "No mobil-

ity" column). Mobility without distortions diversifies risk exposure fairly uniformly across quantiles, reducing compensating differences by 2.0 to 2.3 basis points. The slightly larger effect at the 90th percentile (−0.023) reflects the higher no-mobility compensating difference at the top of the consumption distribution (0.074), and the U-shaped profile becomes nearly flat at about 0.052% per year once mobility is introduced.

Table 1: Effect of mobility on Case 0 compensating differences $\tilde{\nu}_p^{(0)}$ (% per year), consumption ranking.[7]

Percentile	No mobility	With mobility	Effect
10th	0.0730	0.0519	−0.0212
50th	0.0721	0.0516	−0.0204
90th	0.0744	0.0516	−0.0228

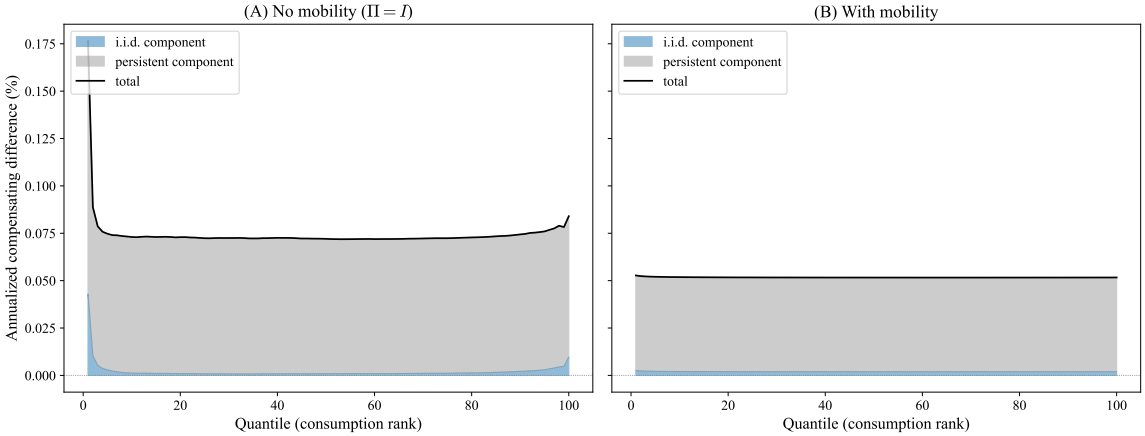


Figure 3: Decomposition of Case 0 compensating differences into i.i.d. and persistent components, consumption ranking. (A) No mobility ($\Pi = I$). (B) With mobility.

Figure 3 decomposes the Case 0 compensating difference into its i.i.d. and persistent components, following the decomposition in Sargent et al. (2026*a*) (Section 5.5). The decomposition splits $\sigma_{j,h}^2 = \|F_j\|^2 + (\sigma_{j,h}^2 - \|F_j\|^2)$ in (53), attributing

[7]Table entries labeled 10th, 50th, 90th correspond to zero-based bin indices $p = 9, 49, 89$ in the model's convention $p \in \{0, \ldots, M-1\}$.

the contemporaneous shock variance $\|F_j\|^2$ to the i.i.d. component and the remainder to the persistent component. Summing the i.i.d. part over all horizons gives $\frac{(1-\beta)\beta}{2}[\Pi(I-\beta\Pi)^{-1}\|F\|^2]_p$, which reduces to $\frac{\beta}{2}\|F_p\|^2$ when $\Pi = I$. The persistent component reflects the accumulated effect of serially correlated risk propagated through the state dynamics. Without mobility, there is a notable U-shape in the compensating differences across percentiles, and the persistent component accounts for the vast majority of the welfare cost at all quantiles. With mobility, the U-shape is replaced by a nearly flat line across percentiles, consistent with Table 1, but the persistent component still accounts for the bulk of the welfare cost.

At the target $\bar{p}_{\text{target}} = 0.45$, the calibrated robustness parameters under consumption ranking are $\theta_\varepsilon = 1.383$ (Case 1), $\theta_\Pi = 0.855$ (Case 2), and $\theta_{\text{joint}} = 1.328$ (Case 3). The larger θ_Π (relative to $\theta_\Pi = 0.692$ under private-income ranking) implies that, in relative-entropy terms, a milder mobility distortion is sufficient to attain the same detection-error probability under consumption ranking.

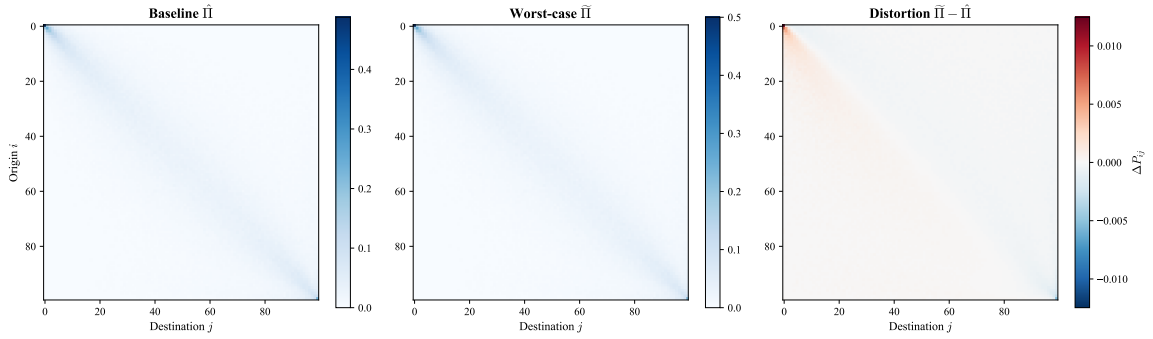


Figure 4: Worst-case transition probabilities under consumption ranking (Case 3, $\bar{p} = 0.45$), evaluated at the stationary-distribution average of the reduced state.

The worst-case transition matrix distortion under Case 3 (Figure 4) shifts mass toward lower-ranked consumption destinations, and the associated stationary distribution (Figure 5) concentrates weight toward the bottom of the consumption distribution.

The right panel of Figure 5 reveals that the diagonal entries of Π and $\widetilde{\Pi}$ are nearly identical, so the distortion leaves row-wise persistence essentially unchanged. The shift in the stationary distribution evident in the left panel originates instead in the

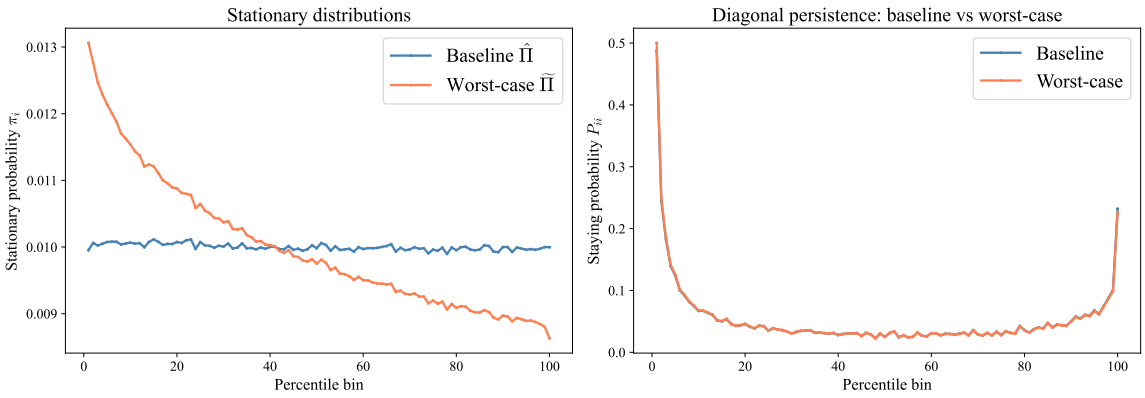


Figure 5: Left panel: Stationary distributions under the approximating and worst-case models, consumption ranking (Case 3). Right panel: Diagonal elements of the approximating and worst-case transition matrices, which are nearly identical. The worst-case stationary distribution is the p-marginal of the invariant law of the simulated worst-case joint process (p_t, x_t).

off-diagonal entries. As Figure 4 shows, every row of the worst-case transition matrix reallocates probability mass toward lower-ranked destinations and away from higher-ranked ones. Although the magnitude of the reallocation is small in each row, the downward tilt appears across all origins. Because the chain mixes slowly ($|\lambda_2(\Pi)| \approx 0.85$, corresponding to a mixing time of roughly six quarters), this systematic downward pressure accumulates over many periods before dissipating, and the stationary distribution $\widetilde{\pi}$ shifts toward the bottom of the consumption distribution. Although this pattern is informative about how introducing robustness distorts the transition matrix, the shift is modest in absolute terms, as the narrow range of the vertical axis in the left panel of Figure 5 indicates. The same pattern of distortion appears under private-income ranking, as we document next.

Table 2 reports annualized compensating differences across all four cases at $\bar{p} = 0.45$. Two features stand out. First, the compensating differences follow the ordering $\alpha_p^{(0)} < \alpha_p^{(1)} < \alpha_p^{(2)} < \alpha_p^{(3)}$ across all quantiles in our baseline calibration. Second, the compensating differences in Cases 2 and 3 decline with quantile: in Case 2, the value falls from 0.187% at the 10th percentile to 0.180% at the 90th, a downward tilt of about 0.6 basis points. Case 3 shows a similar decline from 0.200% to 0.195%.

Table 2: Annualized compensating differences $\bar{\nu}_p^{(m)}$ (% per year), consumption ranking, DEP target $\bar{p} = 0.45$.

Percentile	Case 0	Case 1	Case 2	Case 3
10th	0.0519	0.1111	0.1865	0.1995
50th	0.0516	0.1109	0.1817	0.1963
90th	0.0516	0.1110	0.1802	0.1954

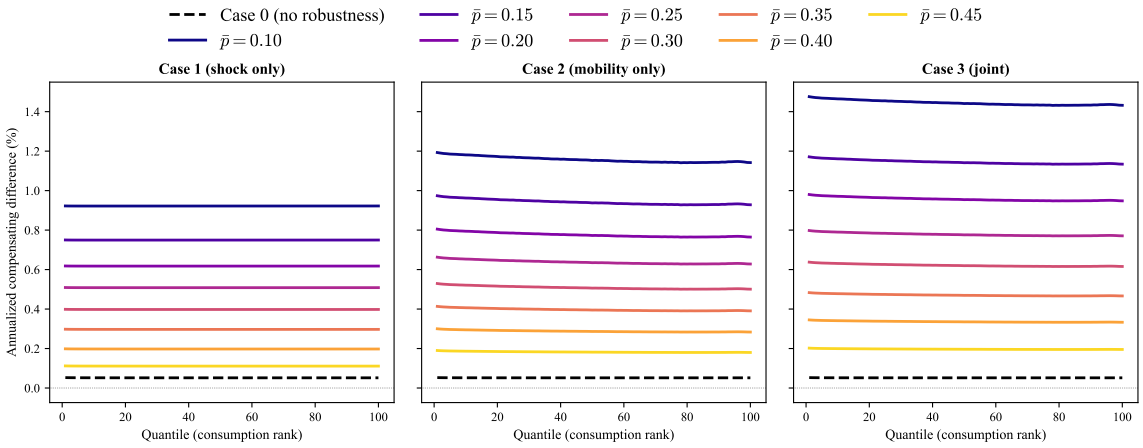


Figure 6: Annualized compensating differences $\bar{\nu}_p^{(m)}$ as a function of the DEP target $\bar{p}$, consumption ranking.

The downward tilt arises because the worst-case mobility distortion tilts probability mass toward lower-consumption destinations, where continuation values are lowest. Lower-consumption households already occupy the least-favored positions under the approximating model, so robustness amplifies the cost of remaining in or returning to those states. Top-quantile households, by contrast, retain some buffer against reaching the bottom of the distribution even under the distorted transition matrix. In the distortion panel of Figure 4, this appears as positive distortion concentrated in low-destination columns and negative distortion in high-destination columns.

Figure 6 shows how compensating differences vary with the DEP target $\bar{p}$. As $\bar{p}$ falls, welfare costs rise across all cases. At the target $\bar{p} = 0.45$, Case 3 is already about 0.20% per year, and it increases markedly as the target becomes more stringent. The downward tilt in Cases 2 and 3 becomes more pronounced at lower $\bar{p}$, as the mobility distortion grows.

6.8 Results: private-income ranking

Under private-income ranking, the observable state (60) has dimension $1+3M = 301$, stacking aggregate private-income growth with demeaned cross-sections of private income, post-tax income, and consumption.

Figure 7 plots log levels relative to Y_0 at the 10th, 50th, and 90th percentiles under private-income ranking. The spread between the 10th and 90th percentiles is large in private income (about 3 log points), and the bottom income percentile displays substantial temporal volatility of its percentile mean. Consumption (panel C) displays greater temporal volatility of percentile means at each percentile than under consumption ranking (Figure 2), because private-income bins mix households with heterogeneous consumption levels. Another notable feature is that consumption exceeds post-tax-and-transfer income at the bottom of the distribution, consistent with the lowest private-income bins containing a substantial share of retired households who dissave to finance consumption. At the median and top, consumption is lower than post-tax income, reflecting working-age households who save a portion of their income.

The greater temporal volatility of percentile means at the bottom of the distri-

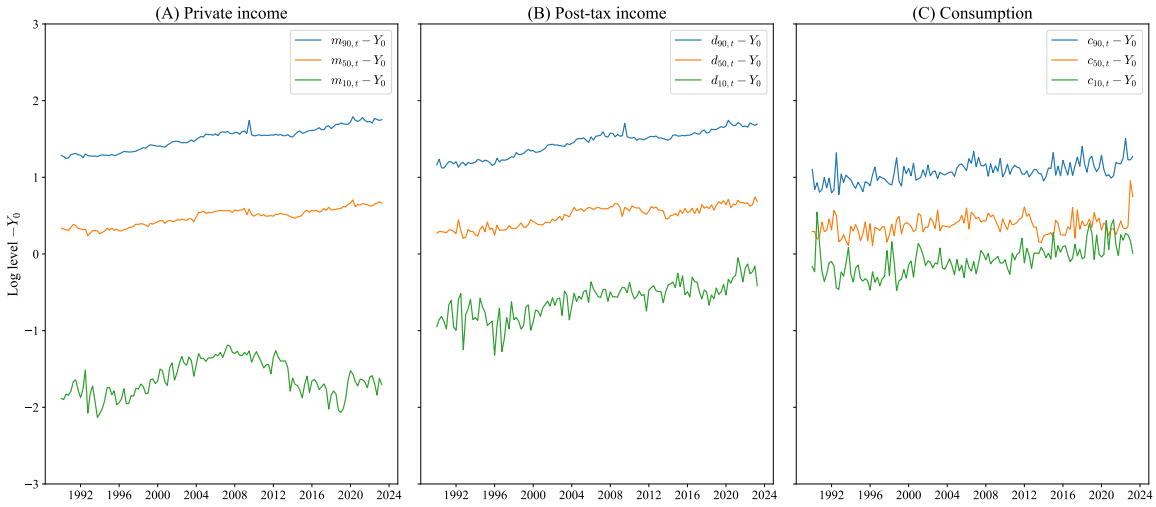


Figure 7: Log levels relative to the initial aggregate level Y_0 at the 10th, 50th, and 90th percentiles, private-income ranking. (A) Private income, (B) post-tax income, (C) consumption.

bution, visible in Figure 7(C), is consistent with the larger compensating differences and more pronounced cross-quantile heterogeneity reported in Table 3. Without mobility across percentiles, the compensating difference at the 10th percentile is 0.117%, falling to 0.076% at the 90th, showing a much steeper gradient than under consumption ranking. Introducing mobility substantially flattens this profile, with the largest reduction occurring at the bottom of the distribution.

Table 3: Effect of mobility on Case 0 compensating differences $\bar{\nu}_p^{(0)}$ (% per year), private-income ranking.

Percentile	No mobility	With mobility	Effect
10th	0.1168	0.0662	−0.0506
50th	0.0854	0.0642	−0.0212
90th	0.0763	0.0631	−0.0132

Figure 8 decomposes the Case 0 compensating difference under private-income ranking into the same i.i.d. and persistent components. In contrast to the consumption case (Figure 3), the U-shape in compensating differences across percentiles is less pronounced when mobility is absent ($\Pi = I$). The U-shape is replaced by a

downward-sloping trend in the share of i.i.d. risk across quantiles, which is consistent with Figure 7(C) showing that the bottom of the distribution has greater temporal volatility of percentile means than the top under private-income ranking. However, once mobility is introduced, the compensating differences become more uniform across quantiles, with slightly higher values at the bottom—analogous to the flattening observed under consumption ranking.

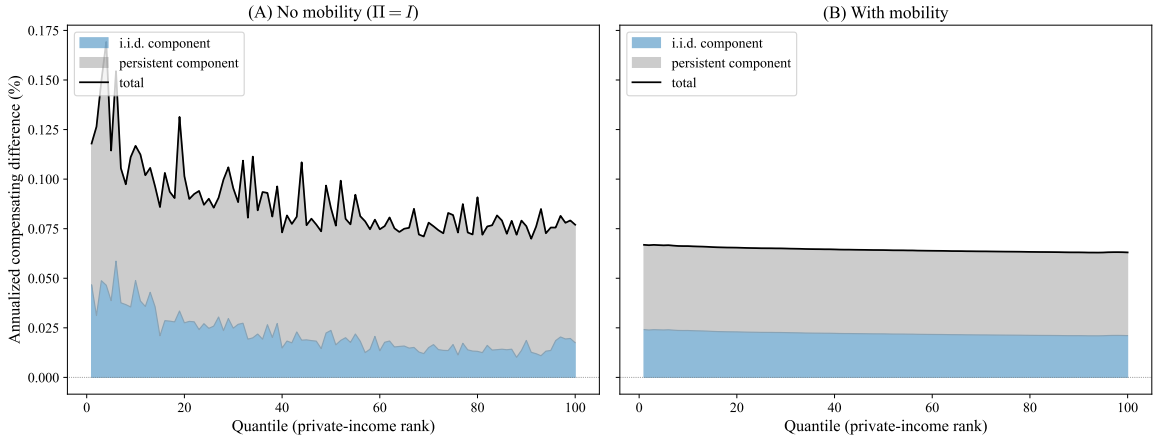


Figure 8: Decomposition of Case 0 compensating differences into i.i.d. and persistent components, private-income ranking. (A) No mobility ($\Pi = I$). (B) With mobility.

At the target $\bar{p} = 0.45$, the calibrated robustness parameters are $\theta_\varepsilon = 1.396$ (Case 1), $\theta_\Pi = 0.692$ (Case 2), and $\theta_{\text{joint}} = 1.200$ (Case 3). Figures 9 and 10 display the worst-case mobility distortion under Case 3 at $\bar{p} = 0.45$. The distorted transition probabilities $\widetilde{\Pi}_{pj}$ shift mass toward lower-ranked destinations, and the implied stationary distribution concentrates more weight at the bottom of the private-income distribution. Compared with the consumption case, the distortion is more tightly concentrated around the diagonal, with a notable transfer of mass from the top ranks to lower ranks.

Table 4 reports annualized compensating differences. In Case 2, the compensating differences increase with quantile, from 0.174% at the 10th percentile to 0.179% at the 90th. In Case 3, the same force is present but weaker at the target, leaving the profile nearly flat with a slight increase at the top. This contrasts with the downward-sloping compensating differences under consumption ranking (Table 2).

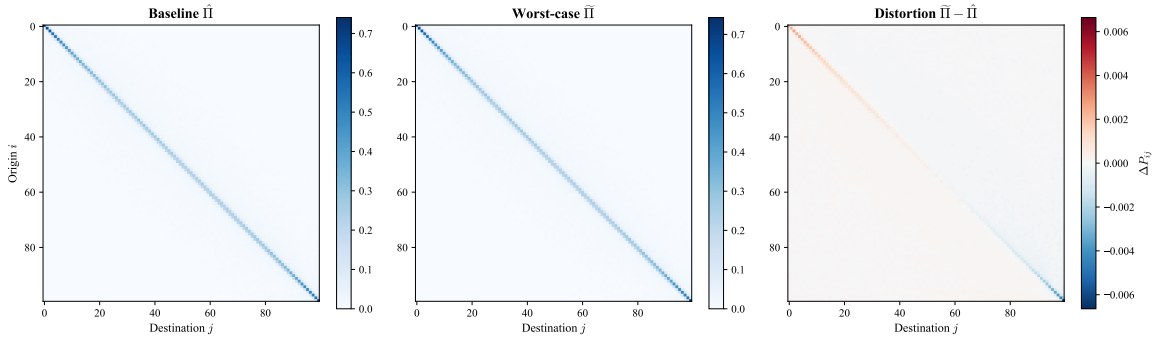


Figure 9: Worst-case transition probabilities under private-income ranking (Case 3, $\bar{p} = 0.45$), evaluated at the stationary-distribution average of the reduced state.

Figure 10: Left panel: Stationary distributions under the approximating and worst-case models, private-income ranking (Case 3). Right panel: Diagonal elements of the approximating and worst-case transition matrices, which are nearly identical. The worst-case stationary distribution is the p-marginal of the invariant law of the simulated worst-case joint process (p_t, x_t).

The upward tilt arises because higher private-income quantiles have more downside exposure under the worst-case mobility distortion. Since the private-income transition matrix is tightly concentrated on the diagonal, the distortion operates mainly by lowering diagonal persistence at high income ranks and shifting that probability mass to lower ranks, while changes far from the diagonal remain small.

Table 4: Annualized compensating differences $\bar{\nu}_p^{(m)}$ (% per year), private-income ranking, DEP target $\bar{p} = 0.45$.

Percentile	Case 0	Case 1	Case 2	Case 3
10th	0.0662	0.1123	0.1738	0.1828
50th	0.0642	0.1102	0.1751	0.1825
90th	0.0631	0.1090	0.1787	0.1840

Figure 11 shows the DEP sensitivity. As $\bar{p}$ falls, the upward tilt in Cases 2 and 3 becomes more pronounced, reaching a spread of roughly 10 basis points between the 10th and 90th percentiles at $\bar{p} = 0.10$. Case 3 exceeds 1% for $\bar{p} \leqslant 0.15$. This contrasts with the consumption case, where the downward tilt becomes more pronounced at lower $\bar{p}$, a pattern that reflects the markedly different structure of the transition matrices across the two rankings.

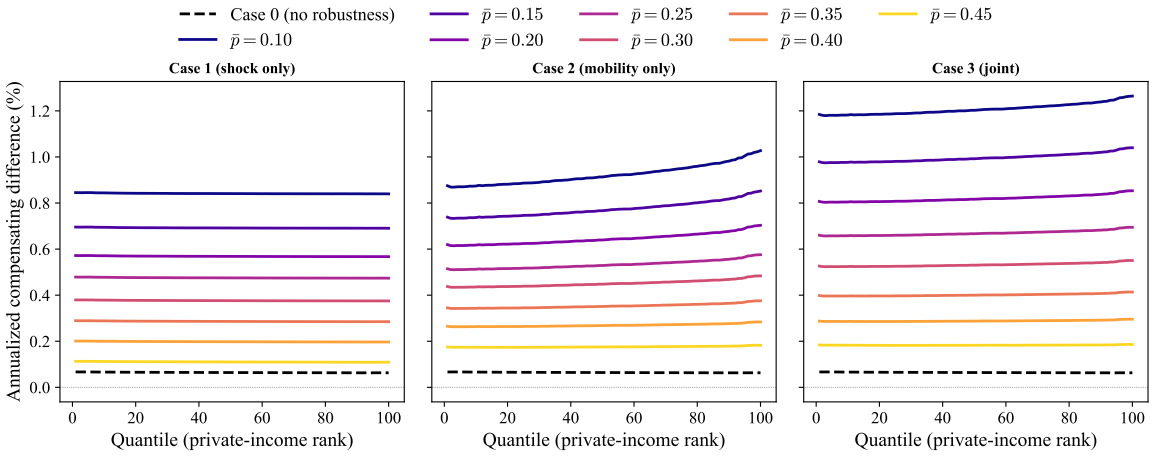


Figure 11: Annualized compensating differences $\bar{\nu}_p^{(m)}$ as a function of the DEP target $\bar{p}$, private-income ranking.

6.9 Role of the ranking variable

The two ranking schemes share one broad qualitative pattern: introducing robustness raises welfare costs across percentiles. However, they differ in the structure of the worst-case transition distortions, the shape of the i.i.d. and persistent components, and the shape of the compensating differences across quantiles. Three mechanisms drive the differences.

First, the conditional variance of consumption within quantiles differs across rankings. Under consumption ranking, within-quantile consumption dispersion is mechanically compressed because households are sorted by consumption itself. This tighter sorting contributes to smaller and more homogeneous values of the time-series variance $\sigma^2_{p,h}$ of future quantile means. Under private-income ranking, households within a given income bin can have widely varying consumption levels owing to differences in tax-transfer benefits and saving behavior, producing larger and more heterogeneous within-bin consumption patterns and thus a larger variance.

Second, the transition matrices have different off-diagonal structure. The consumption transition matrix has more off-diagonal mass than the private-income matrix (Figure 1), reflecting greater cross-quantile mixing under consumption ranking. More off-diagonal mass means that the worst-case mobility distortion can redirect probability more broadly across destinations. The private-income matrix has tighter diagonal persistence, so the corresponding worst-case distortion is concentrated much closer to the diagonal.

Third, the direction of the worst-case mobility tilt reverses across rankings. The distorted transition probabilities (73) tilt toward destinations j with low continuation values. Under consumption ranking, the continuation-value decline is steepest at the bottom, and the robust agent fears being trapped at low consumption quantiles, which generates a downward tilt in compensating differences (bottom quantiles pay more). Under private-income ranking, since the distortion is concentrated near the diagonal, the worst-case tilt mainly erodes persistence at high income ranks and shifts that probability mass to lower ranks, so the robust agent fears losing the buffer of high income and falling to lower ranks, which generates an upward tilt in compensating differences (top quantiles pay more).

These results highlight that the welfare implications of robustness are not invariant to the grouping scheme. The choice of ranking variable shapes both the baseline risk distribution and the direction in which worst-case distortions bite, producing substantively different conclusions about which households bear the largest welfare costs of model uncertainty. At the same time, one qualitative pattern is common to both rankings: introducing robustness raises welfare costs throughout the distribution, and mobility distortions amplify costs most strongly for households that are most exposed to adverse rank changes under the worst-case transition matrix. These results show that the source of model misspecification that the agent fears should be placed on the table when evaluating the welfare implications of robustness.

7 Concluding remarks

Our work extends the representative-agent welfare-cost analysis of Obstfeld (1994), Tallarini (2000), and Barillas et al. (2009) to a cross-section of consumers with cross-quantile mobility. Relative to Barillas et al. (2009), who study doubts in a representative-agent endowment economy, we study the distributional consequences of doubts for heterogeneous consumers. We extend the welfare-cost framework of Sargent et al. (2026*a*) along two dimensions that together transform both the magnitude and cross-sectional distribution of welfare costs. The first extension brings mobility across quantiles, modeled as a time-invariant Markov chain estimated from consecutive CEX interview pairs under two alternative ranking schemes. The second extension equips each consumer with multiplier preferences that express doubts about two distinct statistical components of a shared approximating model, the Gaussian shock distribution governing within-quantile consumption dynamics and the transition matrix governing cross-quantile mobility. Restricting or removing these two components independently generates four analytically tractable cases; Cases 0 and 1 admit closed-form affine value functions, while Cases 2 and 3 require value function iteration. The quantitative results in our baseline calibrations reveal a strict ordering of welfare costs across cases, with a roughly three- to fourfold amplification from the expected-utility benchmark to the fully robust specification, and a pattern

of cross-quantile heterogeneity whose direction is determined by the off-diagonal structure of the estimated transition matrix.

Turning to the main quantitative findings, under consumption ranking at $\bar{p} = 0.45$, compensating differences rise from 0.052% per year with no robustness (Case 0) to 0.111% with dynamics robustness alone (Case 1), to 0.182–0.187% with mobility robustness alone (Case 2), and to 0.195–0.200% with full robustness (Case 3). The tables exhibit the ordering $\alpha_p^{(0)} < \alpha_p^{(1)} < \alpha_p^{(2)} < \alpha_p^{(3)}$ at every quantile in our numerical experiment. Mobility doubts alone (Case 2) contribute more to welfare costs than dynamics doubts alone (Case 1), and the amplification from Case 0 to Case 3 is roughly fourfold under consumption ranking and about threefold under private-income ranking, growing sharply as the detection-error probability target falls. Introducing mobility into the expected-utility benchmark reduces compensating differences relative to the no-mobility counterfactual, because transitioning to higher-consumption quantiles partly diversifies idiosyncratic risk; this diversification effect is larger and more heterogeneous under private-income ranking.

The main new qualitative finding concerns the slope of the compensating differences in Cases 2 and 3 across quantiles, which reverses across the two rankings. Under consumption ranking the gradient is mildly downward sloping, while under private-income ranking it is upward sloping, with compensating differences larger at the top of the distribution than at the bottom. This reversal is a direct consequence of the off-diagonal structure of the transition matrix. The consumption transition matrix has substantial off-diagonal mass, so the worst-case distortion shifts probability broadly toward lower destinations and households at the bottom bear the highest costs. The private-income matrix is tightly concentrated on the diagonal, so the worst-case distortion mainly erodes persistence at high income ranks and households at the top face the greatest additional downside exposure. The direction of the worst-case tilt is therefore a product of the transition matrix structure, identifying which part of the distribution bears the largest welfare costs of mobility uncertainty.

Three broader conclusions follow. First, both components of model uncertainty, doubts about consumption dynamics and doubts about mobility, contribute substantially and independently to welfare costs, so omitting either understates the welfare

implications of robustness by a substantial amount. Second, welfare assessments are sensitive to the statistical model through which the microdata are filtered. The ranking variable shapes both the baseline risk distribution and the direction in which worst-case distortions are felt, producing reversals in the cross-sectional gradient of compensating differences that would be missed by any representative-agent analysis. Third, the estimated transition matrix is a central input driving the cross-sectional pattern of welfare heterogeneity introduced by mobility doubts, and households whose position in the income or consumption distribution exposes them to less favorable row distributions in the transition matrix tend to bear larger welfare costs.

As noted in the introduction, the CEX consumption data underlying our analysis are subject to systematic measurement errors whose effects on inequality trends have been documented by Aguiar & Bils (2015), Meyer & Sullivan (2023), and others. Our calculations take the reported data at face value and should therefore be read as benchmarks constructed under the assumption that the statistical model is correctly specified for the observed series. Explicitly modeling measurement error in consumption and income, and asking how such errors propagate through the estimated dynamics, the transition matrix, and ultimately the welfare costs of robustness, is a worthwhile challenge that we plan to confront in a sequel to this paper currently underway.

A Numerical implementation details

For Cases 2 and 3, we solve the nonlinear Bellman equations by value function iteration on a regular grid in the N-dimensional reduced state space $\widehat{x}$. The grid spans ±4 standard deviations of the stationary distribution of $\widehat{x}$ in each coordinate, with 10 equally spaced points per dimension (10^5 grid points total for $N = 5$). Expectations over $\varepsilon \sim N(0, I_N)$ are approximated using $2N+1$ sigma points, and continuation values at off-grid points are evaluated by multilinear interpolation. We initialize $v^{(0)} \equiv 0$ and iterate $v^{(i+1)} = \Phi_\theta v^{(i)}$ until $\|v^{(i+1)} - v^{(i)}\|_\infty < 10^{-6}$.

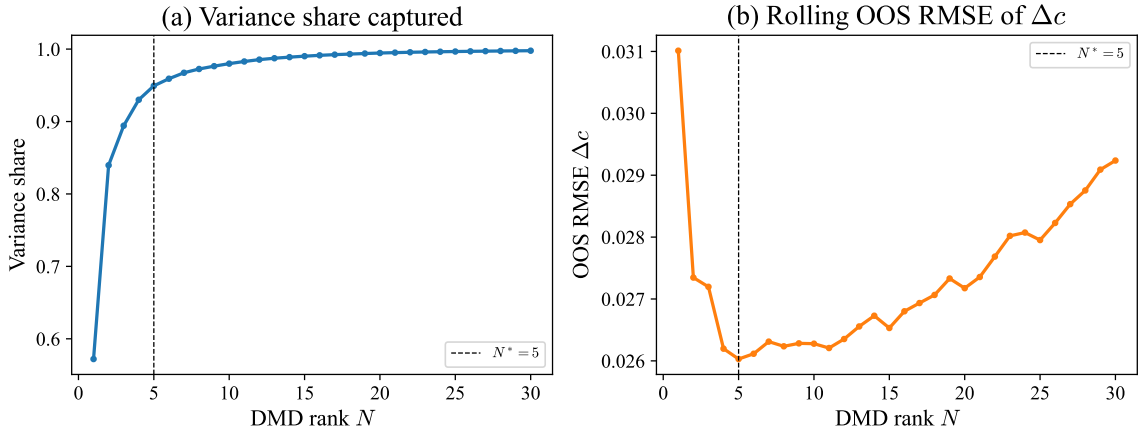
For the DEP calibration, we bisect over $\log\theta$ to find the penalty parameter that achieves the target detection-error probability $\bar{p}_{\text{target}}$. Given the converged worst-case distortion, we simulate $N_{\text{sim}} = 2{,}000$ paths of length $T = T_{\text{data}} = 137$ under both Q (transitions via Π, shocks drawn with quadrature weights ω_n) and $\widetilde{Q}$ (transitions via $\widetilde{\Pi}_{pj}(x_t)$, shocks drawn with the distorted weights $\tilde{\omega}_n^{(j)}(x_t)$ from Section 6.5). We accumulate L_T along each path with a burn-in of 50 periods. The empirical detection-error probability is $\bar{p} = \frac{1}{2}(p_A + p_B)$, where p_A is the fraction of baseline paths with $L_T > 0$ (Type 1 error) and p_B is the fraction of worst-case paths with $L_T \leqslant 0$ (Type 2 error). If $\bar{p} < \bar{p}_{\text{target}}$, the bisection increases θ (reducing the distortion); otherwise it decreases θ.

B DMD rank sensitivity

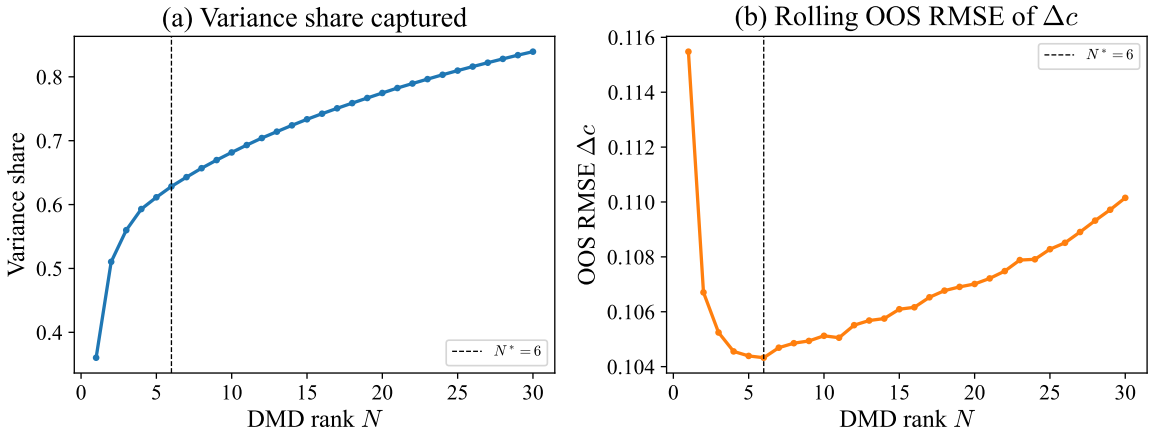
Figure 12 plots two diagnostics against the DMD truncation rank N: (a) the fraction of variance in Y_0 explained by the first N singular values, and (b) the out-of-sample RMSE for one-step-ahead consumption growth $\Delta c_{q,t+1}$ in (67).

The out-of-sample RMSE is computed by an expanding-window forecast. Starting from a minimum training window of $s_{\min} = \lfloor T/2 \rfloor$ periods, we estimate the rank-N DMD operator $\widehat{B}^{(s)}$ on periods $0, \ldots, s-1$ and predict the next-period state $\widehat{y}_{s+1} = \widehat{B}^{(s)} y_s$. We then expand the training window by one period and repeat until $s = T-1$. The RMSE is the root-mean-square of the prediction errors on the consumption rows of y over all $T - s_{\min}$ test periods.

Under consumption ranking, the RMSE is minimized at $N = 5$. Under private-income ranking, the minimum falls at $N = 6$, but the improvement over $N = 5$ is negligible. We therefore set $N = 5$ for both rankings throughout the paper.



(a) Consumption ranking.



(b) Private-income ranking.

Figure 12: DMD rank sensitivity. Variance share, out-of-sample RMSE of consumption growth as functions of the truncation rank N.

References

Aguiar, M. & Bils, M. (2015), 'Has consumption inequality mirrored income inequality?', *American Economic Review* **105**(9), 2725–2756.

Alvarez, F. & Jermann, U. J. (2004), 'Using asset prices to measure the cost of business cycles', *Journal of Political Economy* **112**(6), 1223–1256.

Anderson, E. W., Hansen, L. P. & Sargent, T. J. (2003), 'A quartet of semigroups for model specification, robustness, prices of risk, and model detection', *Journal of the European Economic Association* **1**(1), 68–123.

Attanasio, O. & Davis, S. J. (1996), 'Relative wage movements and the distribution of consumption', *Journal of Political Economy* **104**(6), 1227–1262.

Attanasio, O., Hurst, E. & Pistaferri, L. (2014), The evolution of income, consumption, and leisure inequality in the United States, 1980–2010, *in* C. D. Carroll, T. F. Crossley & J. Sabelhaus, eds, 'Improving the Measurement of Consumer Expenditures', Vol. 74 of *Studies in Income and Wealth*, University of Chicago Press, pp. 100–140.

Barillas, F., Hansen, L. P. & Sargent, T. J. (2009), 'Doubts or variability?', *Journal of Economic Theory* **144**(6), 2388–2418.

Blackwell, D. (1965), 'Discounted dynamic programming', *The Annals of Mathematical Statistics* **36**(1), 226–235.

Blundell, R., Pistaferri, L. & Preston, I. (2008), 'Consumption inequality and partial insurance', *American Economic Review* **98**(5), 1887–1921.

Blundell, R., Pistaferri, L. & Saporta-Eksten, I. (2016), 'Consumption inequality and family labor supply', *American Economic Review* **106**(2), 387–435.

Cutler, D. M. & Katz, L. F. (1992), 'Rising inequality? Changes in the distribution of income and consumption in the 1980s', *The American Economic Review* **82**(2), 546–551.

Duffie, D. & Epstein, L. G. (1992), 'Stochastic differential utility', *Econometrica* **60**(2), 353–394.

Hansen, L. P. (2012), 'Dynamic valuation decomposition within stochastic economies', *Econometrica* **80**(3), 911–967.

Hansen, L. P. & Sargent, T. J. (2001), 'Robust control and model uncertainty', *American Economic Review* **91**(2), 60–66.

Hansen, L. P. & Sargent, T. J. (2008*a*), 'Fragile beliefs and the price of model uncertainty', `http://www.tomsargent.com/research/longrunrisk_tom_14.pdf`. Online manuscript.

Hansen, L. P. & Sargent, T. J. (2008*b*), *Robustness*, Princeton University Press.

Hansen, L. P., Sargent, T. J. & Tallarini, T. D. J. (1999), 'Robust permanent income and pricing', *The Review of Economic Studies* **66**(4), 873–907.

Heathcote, J., Perri, F. & Violante, G. L. (2010), 'Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006', *Review of Economic Dynamics* **13**(1), 15–51.

Heathcote, J., Storesletten, K. & Violante, G. L. (2010), 'The macroeconomic implications of rising wage inequality in the United States', *Journal of Political Economy* **118**(4), 681–722.

Heathcote, J., Storesletten, K. & Violante, G. L. (2014), 'Consumption and labor supply with partial insurance: An analytical framework', *American Economic Review* **104**(7), 2075–2126.

Heathcote, J., Storesletten, K. & Violante, G. L. (2017), 'Optimal tax progressivity: An analytical framework', *Quarterly Journal of Economics* **132**(4), 1693–1754.

Krueger, D. & Perri, F. (2003), 'On the welfare consequences of the increase in inequality in the United States', *NBER Macroeconomics Annual* **18**, 83–121.

Krueger, D. & Perri, F. (2006), 'Does income inequality lead to consumption inequality? Evidence and theory', *Review of Economic Studies* **73**(1), 163–193.

Krueger, D., Perri, F., Pistaferri, L. & Violante, G. L. (2010), 'Cross-sectional facts for macroeconomists', *Review of Economic Dynamics* **13**(1), 1–14.

Kutz, J. N., Brunton, S. L., Brunton, B. W. & Proctor, J. L. (2016), *Dynamic mode decomposition: data-driven modeling of complex systems*, SIAM.

Lucas, R. E. J. (1987), *Models of business cycles*, Oxford Blackwell.

Lucas, R. E. J. (2003), 'Macroeconomic priorities', *American Economic Review* **93**(1), 1–14.

Meyer, B. D. & Sullivan, J. X. (2023), 'Consumption and income inequality in the united states since the 1960s', *Journal of Political Economy* **131**(2), 247–284.

Obstfeld, M. (1994), 'Evaluating risky consumption paths: The role of intertemporal substitutability', *European Economic Review* **38**(7), 1471–1486.

Sargent, T. J., Selvakumar, Y. J. & Yang, Z. (2026*a*), 'Aggregate shocks and cross-section dynamics: Quantifying redistribution and insurance in US household data'. Forthcoming in Journal of Political Economy: Macroeconomics.

Sargent, T. J., Selvakumar, Y. J. & Yang, Z. (2026*b*), 'Dynamic mode decompositions and vector autoregressions'. Forthcoming in International Economic Review.

Sargent, T. J. & Stachurski, J. (2025), *Dynamic Programming: Finite States*, Cambridge University Press.

Schmid, P. J. (2010), 'Dynamic mode decomposition of numerical and experimental data', *Journal of Fluid Mechanics* **656**, 5–28.

Stokey, N. L., Lucas, Jr., R. E. & Prescott, E. C. (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press.

Tallarini, T. D. J. (2000), 'Risk-sensitive real business cycles', *Journal of monetary Economics* **45**(3), 507–532.

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L. & Kutz, J. N. (2014), 'On dynamic mode decomposition: theory and applications', *Journal of Computational Dynamics* **1**(2), 391–421.