

Discussion Paper Series

RIEB

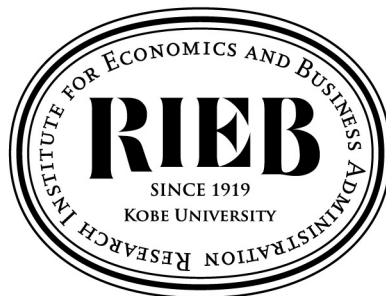
Kobe University

DP2026-09

**Stationary Distributions in Monotone
Markov Models : Theory and Applications**

**Takashi KAMIHIGASHI
John STACHURSKI**

March 24, 2026



Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

STATIONARY DISTRIBUTIONS IN MONOTONE MARKOV MODELS: THEORY AND APPLICATIONS

TAKASHI KAMIHIGASHI

Center for Computational Social Science, Kobe University

JOHN STACHURSKI

National Graduate Institute for Policy Studies

ABSTRACT. Many economic models feature monotone Markov dynamics on state spaces that may be noncompact. Establishing existence, uniqueness, and stability of stationary distributions in such settings has required a patchwork of sufficient conditions, each tailored to specific applications. We provide a single necessary and sufficient condition: a monotone Markov process has a globally stable stationary distribution if and only if it is asymptotically contractive and has a tight trajectory. This characterization covers both compact and noncompact state spaces, discrete and continuous time, and extends to nonlinear Markov operators that depend on aggregate state. We demonstrate the result through applications to wage dynamics, Bayesian learning with belief shocks, and income processes that generate Pareto tails.

1. INTRODUCTION

Many models in economics and finance have monotonicity properties with respect to state variables. A general methodology for studying dynamics of such models in a Markov setting was provided in the frequently cited paper of [Hopenhayn and Prescott \(1992\)](#),

E-mail addresses: tkamihig@rieb.kobe-u.ac.jp, john.stachurski@anu.edu.au.

Key words and phrases. fixed points, semigroups, contraction, global stability.

which in turn extended earlier results by [Stokey and Lucas \(1989\)](#), and [Razin and Yacovlev \(1979\)](#). The monotone mixing condition used in that paper has become a standard approach to establishing stability properties over a large range of applications, including international trade, human capital, business cycles, inequality, intergenerational mobility, and labor markets, using models ranging from overlapping generations and competitive search to mean field games (see, e.g., [Chatterjee and Shukayev \(2010\)](#), [Hidalgo-Cabrillana \(2009\)](#), [Samaniego \(2008\)](#), [Marcet et al. \(2007\)](#), [Antunes and Cavalcanti \(2007\)](#), [Morand and Reffett \(2007\)](#), [Le Grand and Ragot \(2022\)](#), [Light and Weintraub \(2022\)](#), [Balbus et al. \(2025\)](#), and [Kam et al. \(2025\)](#)).

The conditions in [Hopenhayn and Prescott \(1992\)](#) are not universally applicable, however, since they require a compact state space with least and greatest elements. This setup compromises analysis of important features of economic data, such as heavy tails in cross-sectional distributions. In addition, many time series models naturally generate unbounded states. The need to extend [Hopenhayn and Prescott \(1992\)](#) has motivated a substantial literature. Subsequent research has weakened their mixing conditions and replaced compactness of the state space with “tightness” type conditions from the modern literature on Markov process theory, as well as treating partially monotone models (see, e.g., [Kamihigashi and Stachurski \(2012\)](#), [Kamihigashi and Stachurski \(2014\)](#), [Kamihigashi and Stachurski \(2016\)](#), [Foss et al. \(2018\)](#), [Kamihigashi and Stachurski \(2019\)](#), [Foss and Scheutzow \(2024\)](#), [Light \(2024\)](#), and [Light \(2026\)](#)).

In this paper we bring this literature full circle by providing simple necessary and sufficient conditions for global stability—that is, existence, uniqueness, and stability of stationary distributions—in monotone Markov models defined on either compact or non-compact state space. These conditions are built on top of a new fixed point theorem established in the paper. The fixed point theorem shows that, in any complete preordered metric space where the metric satisfies a diagonal property, an order-preserving (i.e., monotone) operator has a unique fixed point if and only if it is asymptotically contractive and generates at least one order-bounded trajectory.

In a Markov setting, we show that asymptotic contractivity holds under weak mixing conditions that can be understood as generalizations of the the well-known monotone mixing condition in [Hopenhayn and Prescott \(1992\)](#). We also show that, when working with the stochastic dominance partial order, existence of an order-bounded trajectory is equivalent to existence of a tight trajectory. This is a weak version of boundedness in probability, which is a standard stability-related condition in the Markov process literature (see, e.g., [Meyn and Tweedie \(2009\)](#); [Hahn et al. \(2024\)](#); [Ma and Toda \(2025\)](#)).

Existence of a tight trajectory is also considerably weaker than assuming, as in [Hopenhayn and Prescott \(1992\)](#), that the state space itself is compact.

In addition, we prove all of these results in an abstract semigroup setting that

- (i) includes both continuous and discrete time, and
- (ii) allows for nonlinear Markov operators as well as linear ones.

Point (i) means that, in the process of establishing our necessary and sufficient conditions, we also extend [Hopenhayn and Prescott \(1992\)](#) and much of the later surrounding literature to handle continuous time models. Point (ii) means that our results can be applied to nonlinear Markov distribution dynamics, where individual Markov updates depend on a distribution and the distribution is determined by the individual Markov laws (see, e.g., [Light \(2026\)](#)).

We demonstrate the scope of our results through several economic applications. In the compact state space setting, we revisit the monotone mixing condition of [Hopenhayn and Prescott \(1992\)](#), providing sharper results including exponential convergence rates and ergodicity, and extending the framework to continuous time. We illustrate with a continuous-time job ladder model of wage dynamics. For noncompact state spaces, we develop applications to piecewise deterministic Markov processes (PDMPs), a class of continuous-time models that combine smooth deterministic dynamics with random jumps. PDMPs arise naturally in models of wealth accumulation or firm dynamics, and are well-suited to capturing the role of jumps and discrete shocks in continuous time.

Our applications include a continuous-time job ladder model of wage dynamics, extending the framework of [Burdett and Mortensen \(1998\)](#) and [Moscarini and Postel-Vinay \(2013\)](#) by allowing state-dependent wage offers; a model of Bayesian learning subject to belief shocks, as arise in models of countercyclical uncertainty and forecast bias ([Orlik and Veldkamp, 2014](#); [Cogley and Sargent, 2008](#)); and two models of income dynamics—a pure jump process and a PDMP with deterministic drift—both of which generate Pareto tails in the stationary distribution, complementing recent work on stationary distributions in continuous-time heterogeneous-agent models ([Gabaix et al., 2016](#)) and on Pareto exponents in multiplicative economies ([Beare and Toda, 2022](#)).

On a theoretical level, the closest papers to ours are [Kamihigashi and Stachurski \(2014\)](#) and [Foss and Scheutzow \(2024\)](#), who prove that global stability holds under a combination of (a) order-theoretic mixing conditions and (b) side conditions that imply existence of a stationary distribution. Under matching assumptions on the underlying

state space, both of these results are special cases of our abstract stability result in Theorem 2.1, which shows that asymptotic contractivity and existence of an order-bounded trajectory are enough for global stability. (The results in Kamihigashi and Stachurski (2014) and Foss and Scheutzow (2024) both imply asymptotic contractivity and order boundedness of at least one trajectory.) At the same time, Theorem 2.1 goes further, admitting operators acting on arbitrary preordered spaces, providing necessary conditions for global stability, and handling continuous time as well as discrete time. While Kamihigashi and Stachurski (2014) is essentially superseded by the results in this paper, Foss and Scheutzow (2024) is complementary, adding valuable sufficient conditions for order boundedness and asymptotic contractivity when time is discrete.

The remainder of the paper is organized as follows. Section 2 develops the abstract fixed point theory for order-preserving semigroups. Section 3 specializes to probability spaces, introducing the Bhattacharya metric and connecting tightness to order boundedness. Section 4 covers stochastic kernels and transition probability functions. Section 5 treats the monotone mixing condition on compact state spaces, with an application to wage dynamics. Section 6 extends to noncompact state spaces in discrete time, with an application to Bayesian learning with belief shocks. Section 7 develops the continuous-time counterpart, with an application to income dynamics. Section 8 develops the PDMP framework and applies it to income dynamics with deterministic drift. Proofs are collected in Section A.

2. STABILITY OF ASYMPTOTIC CONTRACTIONS

We begin with background on preordered metric spaces and the diagonal property. We then introduce semigroups and the notions of asymptotic contractivity and global stability, before stating our main results.

2.1. Background. Let X be a set. If A is any subset of X , then A^c refers to its complement in X . We recall that a binary relation \leq on X is called a *preorder* if it is transitive and reflexive. An antisymmetric preorder on X is called a *partial order* on X . Given a preorder \leq on X , each subset I of X having the form

$$I := [a, b] := \{x \in X : a \leq x \leq b\} \quad (a, b \in X, a \leq b)$$

is called an *order interval* in X . A set $B \subset X$ is called *order-bounded* if there exists an order interval I in X with $B \subset I$. A sequence $(x_i) := (x_i)_{i \in \mathbb{N}} \subset X$ is called *increasing* if $x_i \leq x_{i+1}$ for all i and *decreasing* if $x_{i+1} \leq x_i$ for all i .

A map T from X to itself is called a *self-map* on X . The symbol T^i indicates i compositions of T with itself. A self-map $T : X \rightarrow X$ is called *order-preserving* if $x \leq y$ implies $Tx \leq Ty$. A *fixed point* of T is an element $x^* \in X$ such that $Tx^* = x^*$.

A *preordered metric space* is a triple (X, d, \leq) where (X, d) is a metric space and \leq is a preorder on X . For such a space, the preorder \leq is called *closed* if $(x_i) \subset X$, $(y_i) \subset X$, $x_i \leq y_i$ for all i , $x_i \rightarrow x \in X$ and $y_i \rightarrow y \in X$ implies $x \leq y$. We say that d has the *diagonal property* with respect to \leq if, for all a, b in X ,

$$d(x, y) \leq d(a, b) \quad \text{whenever} \quad x, y \in [a, b]. \quad (2.1)$$

Example 2.1. Let X be a Banach lattice with norm $\|\cdot\|$, partial order \leq , and absolute value function $|\cdot|$. Fix a, b in X and $x, y \in [a, b]$. Since $-x \leq -a$ and $y \leq b$, we have $y - x \leq b - a$. A similar argument gives $x - y \leq b - a$. Hence $|x - y| \leq b - a = |b - a|$. Norms on Banach lattices preserve absolute order, so $\|x - y\| \leq \|a - b\|$.

Example 2.2. Let A be any set, let X be a set of real-valued functions on A , and let \leq be the pointwise order on X . If d has the form $d(x, y) = e(|x - y|)$ for some non-decreasing function e mapping into $\mathbb{R}_+ := [0, \infty)$, then d satisfies the diagonal property. Indeed, if we fix $x, y, a, b \in X$ with $x, y \in [a, b]$, then $a - b \leq x - y \leq b - a$ and hence $|x - y| \leq |b - a|$. Since e is non-decreasing on X , this yields $d(x, y) = e(|x - y|) \leq e(|b - a|) = e(|a - b|) = d(a, b)$.

Examples 2.1 and 2.2 show that the diagonal property also holds in most of the classical spaces and any of their subsets (endowed with the same metric and preorder). Obvious special cases include

- $X = \mathbb{R}^n$, where d is Euclidean distance and \leq is the usual pointwise partial order,
- $X =$ all bounded, real-valued functions on a given set M , where d is the supremum distance and \leq is the pointwise partial order, and
- X is an L_p space, d is the L_p distance, and \leq is the almost everywhere pointwise partial order (see, e.g., [Zaanen \(2012\)](#)).

Figure 1 helps illustrate the diagonal property in the first two cases.

2.2. Semigroups. Let (X, d, \leq) be a preordered metric space and let $(T_t) := (T_t)_{t \in \mathbb{T}}$ be a family of self-maps on X , where \mathbb{T} is either \mathbb{R}_+ or \mathbb{Z}_+ . The case $\mathbb{T} = \mathbb{R}_+$ represents continuous time, while $\mathbb{T} = \mathbb{Z}_+$ represents discrete time. The family (T_t) is called a *semigroup* on X when

$$T_0 = I \quad \text{and} \quad T_{s+t} = T_s \circ T_t \quad \text{for all } s, t \in \mathbb{T}.$$

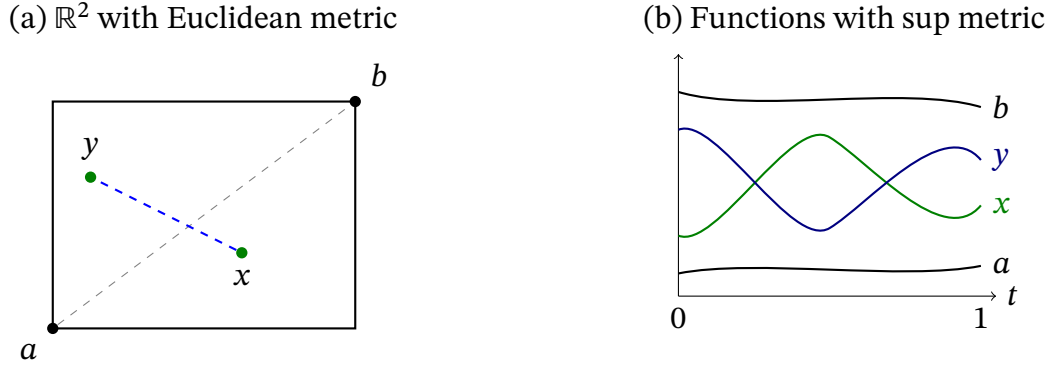


FIGURE 1. Diagonal property: if $x, y \in [a, b]$, then $d(x, y) \leq d(a, b)$

We refer to \mathbb{T} as the *index set*. When time is discrete, the semigroup property implies that T_t is just t compositions of $T := T_1$.

(Note that, unlike some studies of operator semigroups, no algebraic structure is imposed on the underlying set X . Moreover, we do not require that $t \mapsto T_t$ is continuous.)

Given $x \in X$, the map $t \mapsto T_t x$ from $\mathbb{T} \rightarrow X$ is called the *trajectory* of x under the semigroup (T_t) . The trajectory is called *order-bounded* if its range is contained in an order interval; that is, if there exists $a, b \in X$ such that

$$a \leq T_t x \leq b \quad \text{for all } t \in \mathbb{T}. \quad (2.2)$$

We say that $x^* \in X$ is a *stationary point* of (T_t) if x^* is a fixed point of T_t for all $t \in \mathbb{T}$. We say that (T_t) is

- *asymptotically contractive* on X if $d(T_t x, T_t y) \rightarrow 0$ as $t \rightarrow \infty$ for all $x, y \in X$,
- *order-preserving* if T_t is order preserving for all $t \in \mathbb{T}$, and
- *globally stable* if (T_t) has a unique stationary point $x^* \in X$ and

$$d(T_t x, x^*) \rightarrow 0 \text{ as } t \rightarrow \infty \quad \text{for all } x \in X.$$

To simplify terminology, in the case where $\mathbb{T} = \mathbb{Z}_+$ and T_t is the t -th composition of $T := T_1$ for all t , we assign properties listed above directly to T . For example, we say that T is *globally stable* if the semigroup $(T_t)_{t \in \mathbb{Z}_+} = (T^t)_{t \in \mathbb{Z}_+}$ is globally stable.

2.3. Results. We can now state our main results in preordered metric space. Throughout, we take (X, d, \leq) to be such a space. We use the following assumptions.

Assumption 2.1. The metric d is complete and satisfies the diagonal property.

Assumption 2.2. The preorder \leq is closed with respect to d .

In this setting we have the following result, which is the main theoretical contribution of the paper.

Theorem 2.1. *Let (T_t) be an order-preserving semigroup on X . If Assumptions 2.1 and 2.2 hold, then the following statements are equivalent:*

- (i) (T_t) is asymptotically contractive and has at least one order-bounded trajectory.
- (ii) (T_t) is globally stable.

Switching to discrete time, we have the following corollary. The corollary is obvious given Theorem 2.1, but nonetheless worth stating due to its value as a fixed point theorem.

Corollary 2.1. *Let T be an order-preserving self-map on X . If Assumptions 2.1 and 2.2 hold, then the following statements are equivalent:*

- (i) T is asymptotically contractive and has at least one order-bounded trajectory.
- (ii) T is globally stable.

Theorem 2.1 assumes that the preorder \leq is closed. This assumption can be dropped when (T_t) satisfies additional conditions:

Theorem 2.2. *If Assumption 2.1 holds and $x \mapsto T_t x$ is continuous on X for all $t \in \mathbb{T}$, then the equivalence of (i) and (ii) in Theorem 2.1 is also valid.*

The proofs of these and other results can be found in the appendix.

3. STABILITY ON PROBABILITY SPACES

Next we consider the special case where the semigroups under investigation act on a space of probability measures. In other words, we analyze environments where distributions are shifted forward in time under given laws of motion. In this setting, we will be able to relate stability to a compactness-type condition on trajectories (tightness) that is often applied in the analysis of distribution dynamics.

3.1. Background. Let $S = (S, \leq)$ be a partially ordered Polish space, that is, S is a separable and completely metrizable topological space equipped with a closed partial order \leq . The Borel sets of S are denoted by \mathcal{B} . A function $f : S \rightarrow \mathbb{R}$ is called *increasing* if $f(x) \leq f(y)$ whenever $x \leq y$. It is called *decreasing* if $f(y) \leq f(x)$ whenever $x \leq y$. This terminology extends to sets in \mathcal{B} by considering their indicator functions. We let

- bS represent the set of bounded measurable functions from S to \mathbb{R} ,
- cbS represent the set of continuous functions in bS ,
- ibS represent the set of increasing functions in bS ,
- $\mathcal{P}(S)$ be the set of all probability measures on (S, \mathcal{B}) .

Elements of $\mathcal{P}(S)$ are also called *distributions*. For $h \in bS$ and $\phi \in \mathcal{P}(S)$, we set

$$\phi(h) := \int h d\phi.$$

We recall that a set $\Phi \subset \mathcal{P}(S)$ is called *tight* if, for all $\epsilon > 0$, there exists a compact $K \subset S$ with $\phi(K^c) \leq \epsilon$ for all $\phi \in \Phi$.

For any pair $\phi, \psi \in \mathcal{P}(S)$, write

$$\phi \preceq_{sd} \psi \iff \phi(h) \leq \psi(h) \text{ for all } h \in ibS.$$

This is the usual notion of (first order) *stochastic dominance*. Under our assumptions on (S, \leq) , the relation \preceq_{sd} is a partial order on $\mathcal{P}(S)$ (see, e.g., [Kamae and Krengel \(1978\)](#)).

Throughout the paper, we impose the following restriction on S .

Assumption 3.1. A subset of S is compact if and only if it is closed and order bounded.

Here is a simple alternative characterization that is helpful for the proofs.

Lemma 3.1. *Assumption 3.1 holds if and only if order intervals in S are compact and compact subsets of S are order-bounded.*

Proof. Suppose first that Assumption 3.1 holds. Order intervals are closed (since \leq is closed) and order-bounded, hence compact by Assumption 3.1. If K is compact, then K is order-bounded by Assumption 3.1. Conversely, suppose that order intervals are compact and compact sets are order-bounded. If K is compact, then K is closed (since S is Hausdorff) and order-bounded (by assumption). If K is closed and order-bounded, then $K \subseteq [a, b]$ for some $a, b \in S$. Since $[a, b]$ is compact and K is a closed subset, K is compact. \square

Example 3.1. Assumption 3.1 holds when S is a product of intervals in \mathbb{R}^n with the componentwise partial order and Euclidean metric. This includes \mathbb{R}^n , $[0, \infty)^n$, $(0, \infty)^n$, $(0, 1)^n$, $[0, 1]^n$, and mixed cases such as $[0, 1) \times (0, \infty)$.

Assumption 3.1 has two useful implications for the results below. One is completeness of the Bhattacharya metric introduced below. Another is an equivalence between tight and order bounded subsets of $\mathcal{P}(S)$:

Proposition 3.1. $\Lambda \subset \mathcal{P}(S)$ is tight if and only if Λ is order bounded in $(\mathcal{P}(S), \leq_{sd})$.

Below, Proposition 3.1 will be used to connect tightness with global stability.

In this paper, when considering stability of maps over distribution space, we always work with the *Bhattacharya metric* β on $\mathcal{P}(S)$, which is defined by

$$\beta(\phi, \psi) = \sup_{h \in ibS, |h| \leq 1} |\phi(h) - \psi(h)|. \quad (3.1)$$

Kamihigashi and Stachurski (2019) show that β is a complete metric on $\mathcal{P}(S)$ under Assumption 3.1 (see Theorem 4.1).

Lemma 3.2. *The metric β satisfies the diagonal property on $\mathcal{P}(S)$.*

Proof. Fix ϕ, ψ, ℓ and u in $\mathcal{P}(S)$ with $\phi, \psi \in [\ell, u]$. Given $h \in ibS$ with $|h| \leq 1$, we have

$$\phi(h) - \psi(h) \leq u(h) - \ell(h) \quad \text{and} \quad \psi(h) - \phi(h) \leq u(h) - \ell(h).$$

It follows that $|\phi(h) - \psi(h)| \leq u(h) - \ell(h) = |u(h) - \ell(h)| \leq \beta(\ell, u)$. The diagonal property is obtained by taking the sup over $h \in ibS$ with $|h| \leq 1$. \square

Another useful observation is as follows.

Lemma 3.3. *The partial order \leq_{sd} on $\mathcal{P}(S)$ is closed with respect to β .*

Proof. Suppose $(\phi_n), (\psi_n) \subset \mathcal{P}(S)$ with $\phi_n \leq_{sd} \psi_n$ for all n , $\beta(\phi_n, \phi) \rightarrow 0$, and $\beta(\psi_n, \psi) \rightarrow 0$. Fix $h \in ibS$ with $|h| \leq 1$. From convergence in the Bhattacharya metric, we obtain $\phi_n(h) \rightarrow \phi(h)$ and $\psi_n(h) \rightarrow \psi(h)$ as $n \rightarrow \infty$. We also have $\phi_n(h) \leq \psi_n(h)$ for all n . Hence, taking limits, $\phi(h) \leq \psi(h)$. This holds for all $h \in ibS$ with $|h| \leq 1$ and, by scaling, extends to all $h \in ibS$. Therefore $\phi \leq_{sd} \psi$. \square

3.2. Semigroups over Distributions. Let us now consider a semigroup $(T_t)_{t \in \mathbb{T}}$ on $\mathcal{P}(S)$. The semigroup represents a model that shifts distributions forward in time via $t \mapsto T_t \phi$, where ϕ is some initial distribution. In all of what follows, $\mathcal{P}(S)$ is endowed with the stochastic dominance partial order \leq_{sd} and the Bhattacharya metric β . Moreover, $S = (S, \leq)$ is itself a partially ordered Polish space satisfying Assumption 3.1. In this setting, we can state the following result.

Theorem 3.1. *Let (T_t) be an order-preserving semigroup on $\mathcal{P}(S)$. The following statements are equivalent:*

- (i) (T_t) is globally stable on $\mathcal{P}(S)$.
- (ii) (T_t) is asymptotically contractive and has an order-bounded trajectory in $\mathcal{P}(S)$.
- (iii) (T_t) is asymptotically contractive and has a tight trajectory in $\mathcal{P}(S)$.

The statement that (T_t) has a tight trajectory means that there exists at least one $\phi \in \mathcal{P}(S)$ such that $(T_t \phi)_{t \in \mathbb{T}}$ is tight in $\mathcal{P}(S)$.

Specializing to discrete time, we have the following corollary.

Corollary 3.1. *If $T : \mathcal{P}(S) \rightarrow \mathcal{P}(S)$ is order-preserving, then the following statements are equivalent:*

- (i) T is globally stable on $\mathcal{P}(S)$.
- (ii) T is asymptotically contractive and has an order-bounded trajectory in $\mathcal{P}(S)$.
- (iii) T is asymptotically contractive and has a tight trajectory in $\mathcal{P}(S)$.

In summary, Theorem 3.1 and Corollary 3.1 provide necessary and sufficient conditions for global stability of order-preserving semigroups and maps on distribution space: asymptotic contractivity combined with tightness of at least one trajectory.

Often the operators that we consider will be Markov operators generated from transition probability functions. Such operators have other useful properties, such as linearity with respect to distributions. We discuss this case in Section 4. We also note, however, that Theorem 3.1 and Corollary 3.1 can be used to study nonlinear maps over distributions, which occur frequently in economic dynamics (see, e.g., Light (2026)).

4. PROBABILISTIC MARKOV MODELS

In this section we provide background on probabilistic aspects of Markov dynamics, and on the connection between underlying Markov models and the operators and semigroups they generate.

4.1. Stochastic Kernels. A *stochastic kernel on S* is a map P from $S \times \mathcal{B}$ to $[0, 1]$ such that $B \mapsto P(x, B)$ is in $\mathcal{P}(S)$ for each $x \in S$ and $x \mapsto P(x, B)$ is measurable for each $B \in \mathcal{B}$. The kernel P is called *increasing* if $x \leq x'$ implies $P(x, \cdot) \leq_{\text{sd}} P(x', \cdot)$. This is equivalent to both of the following statements:

- $Ph \in \text{ib}S$ whenever $h \in \text{ib}S$.
- $\phi P \leq_{\text{sd}} \psi P$ whenever $\phi \leq_{\text{sd}} \psi$.

Each stochastic kernel P on S generates two operators. The first is the *left Markov operator*, defined via

$$\phi \mapsto \phi P, \quad (\phi P)(B) := \int P(x, B) \phi(dx) \quad (B \in \mathcal{B}). \quad (4.1)$$

The second is the *right Markov operator*, defined by

$$h \mapsto Ph, \quad (Ph)(x) := \int h(y) P(x, dy) \quad (h \in \text{b}S, x \in S). \quad (4.2)$$

For convenience, and in line with much of the literature (see, e.g., [Meyn and Tweedie \(2009\)](#)), the same symbol P is used for both operators, as well as for the stochastic kernel. In addition, we will drop the terms “left” and “right”, assuming that the meaning will typically be clear from context, and adding clarification when necessary.

In (4.1), the map $\phi \mapsto \phi P$ is understood as updating distribution ϕ to the next period. (One easily confirms that $B \mapsto (\phi P)(B)$ is a probability measure on \mathcal{B} , so the mapping is from $\mathcal{P}(S)$ to $\mathcal{P}(S)$.) In (4.2), the value $(Ph)(x)$ is interpreted as the conditional expectation of $h(X)$ when X is the next period state and x is the current state.

For the two operators in (4.1) and (4.2), we have the well-known “duality” relation

$$(\phi P)(h) = \phi(Ph) \text{ for any } \phi \in \mathcal{P}(S) \text{ and any } h \in \text{b}S. \quad (4.3)$$

(This follows directly from the definitions and Fubini’s theorem.) We note that the kernel P is increasing if and only if the Markov operator $\phi \mapsto \phi P$ is order-preserving on $(\mathcal{P}(S), \leq_{\text{sd}})$.

The following result parallels the well-known fact that total variation distance between distributions is not increased after passing them through a Markov kernel. It will be useful for our analysis of continuous time models, and also suggests that asymptotic contractivity should be attainable under some additional conditions, which turn out to be mixing conditions.

Lemma 4.1. *If P is an increasing kernel, then the Markov operator $\phi \mapsto \phi P$ is nonexpansive in the Bhattacharya metric:*

$$\beta(\phi P, \psi P) \leq \beta(\phi, \psi) \quad \text{for all } \phi, \psi \in \mathcal{P}(S). \quad (4.4)$$

Proof. Observe that, for probability measures $\phi, \psi \in \mathcal{P}(S)$, we have

$$\beta(\phi P, \psi P) = \sup_h |(\phi P)(h) - (\psi P)(h)| = \sup_h |(\phi)(Ph) - (\psi)(Ph)|,$$

where the supremum is over $h \in ibS$ with $|h| \leq 1$ and the second equality is by the duality relationship in (4.3). Since P is increasing and h is increasing, Ph is increasing. Since $|h| \leq 1$, we have $|Ph| \leq 1$, so the last expression is dominated by $\beta(\phi, \psi)$. \square

4.2. Transition Probability Functions. Following standard terminology, a *transition probability function* is a family of stochastic kernels $(P_t)_{t \in \mathbb{T}}$ such that

- (i) $P_0(x, \cdot) = \delta_x$ for all $x \in S$, and
- (ii) the *Chapman–Kolmogorov equation* holds:

$$P_{s+t}(x, B) = \int P_t(y, B) P_s(x, dy) \quad (s, t \in \mathbb{T}, x \in S, B \in \mathcal{B}). \quad (4.5)$$

When $(P_t)_{t \in \mathbb{T}}$ is a transition probability function, the family of maps $\phi \mapsto \phi P_t$,

$$(\phi P_t)(B) := \int P_t(x, B) \phi(dx) \quad (B \in \mathcal{B}, t \in \mathbb{T}) \quad (4.6)$$

forms a semigroup on $\mathcal{P}(S)$, with the semigroup property $P_{s+t} = P_s \circ P_t$ following from the Chapman–Kolmogorov relations. In this setting, we say that $(P_t)_{t \in \mathbb{T}}$ is a *Markov semigroup on S* . We use the same symbol $(P_t)_{t \in \mathbb{T}}$ to represent the Markov semigroup and the transition probability function, since both represent the same object.¹

Remark 4.1. Each transition probability function $(P_t)_{t \in \mathbb{T}}$ also defines a family of operators $h \mapsto P_t h$, indexed by $t \in \mathbb{T}$, each P_t mapping bS to bS , through the right Markov operator relation in (4.2). Some authors have this in mind when they refer to the semigroup generated by the transition probability function, rather than the family of maps $\phi \mapsto \phi P_t$ sending $\mathcal{P}(S)$ to $\mathcal{P}(S)$. This is just a matter of convention. Our terminology is similar to Lasota and Mackey (1994) and Rudnicki and Tyran-Kamińska (2017).

¹The transition probability function defines the Markov semigroup on S via (4.6). The semigroup defines the transition probability function via $P_t(x, B) := (\delta_x P_t)(B)$ for all $x \in S$, $t \in \mathbb{T}$ and $B \in \mathcal{B}$.

Example 4.1. In the discrete time case, where we work with a fixed stochastic kernel P on S , we set $P_0 := I$ and $P_t := P^t$ for all $t \in \mathbb{N}$. The family $(P_t)_{t \in \mathbb{Z}_+}$ is a transition probability function.

Example 4.2. Let $S = \mathbb{R}$ and $P(t, x, dy) = p(t, x, y) dy$, where

$$p(t, x, y) := \sqrt{\frac{\theta}{\pi\sigma^2(1 - e^{-2\theta t})}} \exp\left(-\frac{\theta(y - xe^{-\theta t})^2}{\sigma^2(1 - e^{-2\theta t})}\right).$$

This is the transition density function for the Ornstein–Uhlenbeck process

$$dX_t = -\theta X_t dt + \sigma dW_t, \quad X_0 = x, \quad (4.7)$$

where $\theta > 0$ is the mean-reversion rate, $\sigma > 0$ is the volatility parameter, and $(W_t)_{t \geq 0}$ is Brownian motion. The family of maps $(P_t)_{t \geq 0}$ defined by

$$\phi \mapsto \phi P_t \quad \text{with} \quad (\phi P_t)(B) := \int \int_B p(t, x, y) dy \phi(dx)$$

is the Markov semigroup on \mathbb{R} generated by the Ornstein–Uhlenbeck transition probability function. For (4.7), the solution can be expressed as

$$X_t = e^{-\theta t} x + \sigma_t Z, \quad \text{where} \quad \sigma_t := \sigma \sqrt{\frac{1 - e^{-2\theta t}}{2\theta}} \quad (4.8)$$

and Z is a standard normal random variable. Hence, for the right Markov operators, we have

$$(P_t h)(x) = \mathbb{E}[h(e^{-\theta t} x + \sigma_t Z)] \quad (x \in S, h \in bS). \quad (4.9)$$

Returning to the general case, a transition probability function $(P_t)_{t \in \mathbb{T}} = (P_t)$ will be called *increasing* if P_t is an increasing stochastic kernel for all $t \in \mathbb{T}$. For example, (4.9) implies that $P_t h$ is increasing whenever h is increasing, so the Ornstein–Uhlenbeck transition probability function is order-preserving. Evidently, $(P_t)_{t \in \mathbb{T}}$ is an increasing transition probability function if and only if it is order-preserving when viewed as a semigroup on $(\mathcal{P}(S), \leq_{sd})$, as in (4.6).

We say that a transition probability function (P_t) is *bounded in probability* when the family of distributions $(P_t(x, \cdot))_{t \in \mathbb{T}}$ is tight for all $x \in S$. For example, for the Ornstein–Uhlenbeck process, the representation (4.8) and $\sigma_t \leq \bar{\sigma} := \sigma/\sqrt{2\theta}$ for all t give $|X_t| \leq |x| + \bar{\sigma}|Z|$, so the family $(P_t(x, \cdot))_{t \geq 0}$ is tight.

We recall that an S -valued stochastic process $(X_t) := (X_t)_{t \in \mathbb{T}}$ supported on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *Markov process* if

$$\mathbb{P}\{X_t \in B \mid \mathcal{F}_s\} = \mathbb{P}\{X_t \in B \mid X_s\} \quad \mathbb{P}\text{-a.s. for all } s, t \in \mathbb{T} \text{ with } s \leq t \text{ and all } B \in \mathcal{B},$$

where $(\mathcal{F}_t)_{t \in \mathbb{T}}$ is the natural filtration generated by (X_t) . A *stopping time* of (\mathcal{F}_t) is a random variable $\tau : \Omega \rightarrow \mathbb{T} \cup \{\infty\}$ such that $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{T}$. We call (X_t) a *strong Markov process* if, in addition, the Markov property holds at all stopping times of (\mathcal{F}_t) : for every (\mathcal{F}_t) -stopping time τ with $\mathbb{P}\{\tau < \infty\} = 1$,

$$\mathbb{P}\{X_{\tau+t} \in B \mid \mathcal{F}_\tau\} = P_t(X_\tau, B) \quad \mathbb{P}\text{-a.s.} \quad \text{for all } t \in \mathbb{T} \text{ and all } B \in \mathcal{B}.$$

Essentially all standard constructions of Markov processes have the strong Markov property, including all discrete time processes and all càdlàg paths under the standard constructions.

Given a transition probability function $(P_t)_{t \in \mathbb{T}}$, we say that a Markov process $(X_t)_{t \in \mathbb{T}}$ is $(P_t)_{t \in \mathbb{T}}$ -*Markov* when

$$\mathbb{P}\{X_t \in B \mid \mathcal{F}_s\} = P_{t-s}(X_s, B) \quad \text{for all nonnegative } s \leq t \text{ and all } B \in \mathcal{B}.$$

For the case of discrete time, when P_t is just the t -composition P^t for some fixed stochastic kernel P , we will simply say that (X_t) is *P-Markov*. In other words, $(X_t)_{t \in \mathbb{Z}_+}$ is *P-Markov* when

$$\mathbb{P}\{X_{t+1} \in B \mid \mathcal{F}_t\} = P(X_t, B) \quad \text{for all } t \in \mathbb{Z}_+ \text{ and all } B \in \mathcal{B}.$$

Given $x \in S$, we write \mathbb{P}_x for the conditional probability $\mathbb{P}\{\cdot \mid X_0 = x\}$.

5. MONOTONE MIXING

Theorem 2 of [Hopenhayn and Prescott \(1992\)](#) shows that, in discrete time, the monotone mixing condition (MMC) is sufficient for global stability of monotone Markov chains. In this section we extend their result by (a) establishing an explicit exponential rate of convergence to the stationary distribution, (b) proving ergodicity for bounded increasing observables, and (c) generalizing from discrete to continuous time.

We state the main result and its proof in Section 5.1, and then apply it to a continuous-time model of wage dynamics in Section 5.2.

5.1. Theory. As before (S, \leq) is a partially ordered Polish space and \leq_{sd} is stochastic dominance. Following [Hopenhayn and Prescott \(1992\)](#), we suppose that S is compact and, in addition, has a greatest element b and a least element a . Note that, under these conditions, Assumption 3.1 is satisfied. Given $x \in S$, we write U_x for all $y \in S$ with $x \leq y$ and D_x for all $y \in S$ with $y \leq x$.

In line with [Hopenhayn and Prescott \(1992\)](#), we say that a stochastic kernel P on S satisfies the *monotone mixing condition* (MMC) if there exists an $\hat{x} \in S$ and an $\epsilon > 0$ such that

$$P(a, U_{\hat{x}}) \geq \epsilon \quad \text{and} \quad P(b, D_{\hat{x}}) \geq \epsilon. \quad (5.1)$$

In other words, a chain starting at a rises above \hat{x} with positive probability, while a chain starting from b falls below \hat{x} with positive probability.

We now present the main result of this section, which extends Theorem 2 of [Hopenhayn and Prescott \(1992\)](#). In the statement, S has the properties listed at the start of Section 5.1.

Theorem 5.1. *Let $(P_t)_{t \in \mathbb{T}}$ be a transition probability function on S . If $(P_t)_{t \in \mathbb{T}}$ is increasing and there exists a positive $u \in \mathbb{T}$ such that P_u satisfies the MMC, then $(P_t)_{t \in \mathbb{T}}$ is globally stable on $\mathcal{P}(S)$ under β and convergence to the unique fixed point ϕ^* is exponential:*

$$\beta(\phi P_t, \phi^*) \leq C e^{-\alpha t} \quad \text{for all } \phi \in \mathcal{P}(S) \text{ and } t \in \mathbb{T}, \quad (5.2)$$

where $C := 2/(1 - \epsilon)$ and $\alpha := \ln(1/(1 - \epsilon))/u$. Moreover, when $\mathbb{T} = \mathbb{Z}_+$, each P -Markov process is monotone ergodic:

$$\mathbb{P} \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} h(X_t) = \phi^*(h) \right\} = 1 \quad \text{for all } h \in \text{ib}S.$$

We note that Assumption 3.1 holds in the setting of Theorem 5.1. The proof can be found in Section A.4.

5.2. Application: wage dynamics. We develop a continuous-time model of wage dynamics based on a job ladder framework with job destruction, extending the class of models studied by [Burdett and Mortensen \(1998\)](#), [Moscarini and Postel-Vinay \(2013\)](#), and [Coles and Mortensen \(2016\)](#) by allowing wage offers to be state-dependent.

Throughout, $S := [0, \bar{w}]$ denotes the state space for wages, with $\bar{w} \in (0, \infty)$. We work with two stochastic kernels on S . The first, denoted Q_u , governs wage draws upon job destruction. The second, Q_e , governs outside offers while employed. We assume throughout that both Q_u and Q_e are increasing, so higher current wages predict higher offers for both employed and unemployed workers.

Let $\delta > 0$ be the job destruction rate and $\lambda > 0$ be the offer arrival rate while employed. Let N_t^d and N_t^e be independent Poisson processes with rates δ and λ respectively. The wage process $(W_t)_{t \geq 0}$ evolves as follows:

- At jump times of N^d , we draw W_t from $Q_u(W_{t-}, \cdot)$
- At jump times of N^e , we draw $W_t^e \sim Q_e(W_{t-}, \cdot)$ and then set $W_t = \max(W_{t-}, W_t^e)$.

Between jumps, W_t remains constant. The max after job arrival means that an employed worker can accept a new offer or retain their current position.

The process $(W_t)_{t \geq 0}$ is a continuous-time pure jump process. Its dynamics can be characterized through the infinitesimal generator \mathcal{A} , which describes the instantaneous expected rate of change of $h(W_t)$ for test functions $h \in bS$. (In discrete time, the analogous object is the one-step conditional expectation operator $Ph - h$.) In this setting, \mathcal{A} acts on $h \in bS$ via

$$(\mathcal{A}h)(w) = \delta \int [h(w') - h(w)] Q_u(w, dw') + \lambda \int [h(w \vee w') - h(w)] Q_e(w, dw').$$

Define the discrete-time *max-chain* $(M_t)_{t \in \mathbb{Z}_+}$ starting from w by $M_0 = w$ and

$$M_{t+1} = \max\{M_t, W_{t+1}\} \quad \text{where} \quad (W_t)_{t \in \mathbb{Z}_+} \text{ is } Q_e\text{-Markov with } W_0 = w.$$

Let Q_m be the kernel associated with the max-chain. We can now rewrite \mathcal{A} more compactly as $\mathcal{A} = \delta(Q_u - I) + \lambda(Q_m - I)$. Some rearranging gives

$$\mathcal{A} = (\delta + \lambda)(K - I) \quad \text{where} \quad K = \frac{\delta}{\delta + \lambda} Q_u + \frac{\lambda}{\delta + \lambda} Q_m.$$

This is the infinitesimal generator of a jump chain with jump kernel K , so the Markov semigroup $(P_t)_{t \geq 0}$ on $[0, \bar{w}]$ admits the representation

$$(P_t f)(w) = \sum_{n=0}^{\infty} e^{-(\delta + \lambda)t} \frac{[(\delta + \lambda)t]^n}{n!} (K^n f)(w). \quad (5.3)$$

(See, e.g., [Siegrist \(2022\)](#), Section 16.20.)

To generate monotone mixing, we impose the following restrictions.

Assumption 5.1. There exists a $\hat{w} \in [0, \bar{w}]$, an $n \in \mathbb{N}$ and an $\epsilon > 0$ such that

$$Q_u^n(\bar{w}, D_{\hat{w}}) \geq \epsilon \quad \text{and} \quad Q_e^n(0, U_{\hat{w}}) \geq \epsilon. \quad (5.4)$$

The assumption requires the existence of an intermediate wage level \hat{w} such that unemployed agents can reach this level and employed agents can fall below this level. Using Assumption 5.1 and Theorem 5.1, we can obtain global stability:

Proposition 5.1. *If Assumption 5.1 holds, then the wage process is globally stable, with unique stationary distribution ϕ^* . Moreover,*

$$\beta(\phi P_t, \phi^*) \leq C e^{-\alpha t} \quad \text{for all } \phi \in \mathcal{P}(S) \text{ and } t \geq 0, \quad (5.5)$$

where

$$C := 2/(1 - \kappa), \quad \alpha := \ln(1/(1 - \kappa)), \quad \text{and} \quad \kappa := \frac{e^{-(\delta+\lambda)}(\delta \wedge \lambda)^n}{n!} \epsilon. \quad (5.6)$$

Proof. First observe that (P_t) is order-preserving. To see this, note that K is a convex combination of the increasing kernels Q_u and Q_m with state-independent weights, hence increasing. Compositions of increasing kernels are increasing, so K^n is increasing for all n . Finally, (5.3) tells us that P_t is a mixture of $(K^n)_{n \geq 0}$ with state-independent Poisson weights, hence increasing.

Next we show that P_t satisfies the MMC for some $t > 0$. Let \hat{w} , n , and ϵ be as in Assumption 5.1. We first claim that $Q_m^n(0, U_{\hat{w}}) \geq Q_e^n(0, U_{\hat{w}})$. To see this, generate $(W_k^e)_{k \in \mathbb{Z}_+}$ from Q_e starting at $W_0^e = 0$, and let $(M_k)_{k \in \mathbb{Z}_+}$ be the max-chain constructed from (W_k^e) , so that $M_k = \max_{j \leq k} W_j^e$. By construction, $M_n \geq W_n^e$, so $W_n^e \geq \hat{w}$ implies $M_n \geq \hat{w}$. Hence $Q_m^n(0, U_{\hat{w}}) \geq Q_e^n(0, U_{\hat{w}})$.

Let n and ϵ be as in Assumption 5.1. The semigroup P_1 assigns positive probability to the event of n jumps by time 1, with the type of each jump (destruction or offer) determined independently. Starting from \bar{w} , the event of n consecutive destructions has probability $e^{-(\delta+\lambda)} \delta^n / n!$, and conditional on this event, the wage follows the Q_u -chain for n steps, falling below \hat{w} with probability at least $Q_u^n(\bar{w}, D_{\hat{w}}) \geq \epsilon$. Similarly, starting from 0, the event of n consecutive offers has probability $e^{-(\delta+\lambda)} \lambda^n / n!$, and conditional on this event, the wage follows the max-chain for n steps, rising above \hat{w} with probability at least $Q_m^n(0, U_{\hat{w}}) \geq Q_e^n(0, U_{\hat{w}}) \geq \epsilon$. Combining these bounds,

$$P_1(\bar{w}, D_{\hat{w}}) \geq \frac{e^{-(\delta+\lambda)} \delta^n}{n!} \epsilon \quad \text{and} \quad P_1(0, U_{\hat{w}}) \geq \frac{e^{-(\delta+\lambda)} \lambda^n}{n!} \epsilon.$$

Both bounds are strictly positive, so P_1 satisfies the MMC with parameter κ as defined in (5.6). Global stability and the exponential bound (5.5) now follow from Theorem 5.1. \square

Figure 2 illustrates a sample path along with the empirical stationary distribution for a wage process evolving on $S = [0, 1]$, with $\delta = 0.1$ and $\lambda = 0.5$. Upon job destruction with current wage W , the new wage is $W' = W \cdot B_u$ where $B_u \sim \text{Beta}(2, 8)$, so the worker retains a random fraction of their previous wage, skewed towards large losses. Upon receiving an outside offer, the offer wage is $W' = 0.5 + 0.5 W \cdot B_e$ where $B_e \sim \text{Beta}(8, 2)$, so offers have a floor of 0.5 and a state-dependent component skewed towards high values (mean 0.8). In this setting, both kernels Q_u and Q_e are increasing: if $w \leq w'$, then $w B_u \leq w' B_u$ and $0.5 + 0.5 w B_e \leq 0.5 + 0.5 w' B_e$ almost surely. In addition, Assumption 5.1 holds

with $\hat{w} = 0.5$ and $n = 1$, since $Q_u(\bar{w}, D_{\hat{w}}) = \mathbb{P}\{B_u \leq 0.5\} \approx 0.98$ and $Q_e(0, U_{\hat{w}}) = \mathbb{P}\{0.5 \geq \hat{w}\} = 1$. Proposition 5.1 then gives global stability with exponential convergence.

In the figure, job destructions are marked in red and accepted offers are marked with green. The empirical stationary distribution is calculated from a simulated path of length 200,000. Convergence of the empirical distribution to the true stationary distribution follows from the ergodicity result in Theorem 5.1.

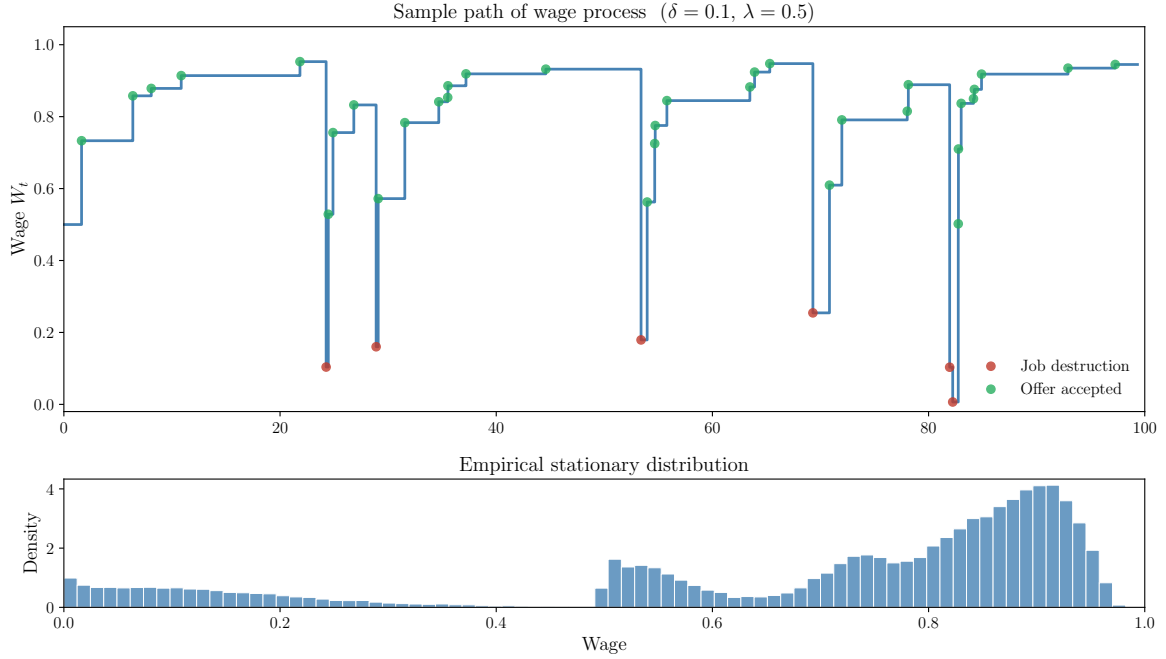


FIGURE 2. Simulation of the wage dynamics model in Section 5.2.

6. NONCOMPACT STATE SPACE, DISCRETE TIME

Let's now drop the assumption that the state is compact, allowing us to include many more useful processes, and to more accurately model properties such as heavy tails in cross-sectional distributions. Dropping compactness requires us to adopt at least some restrictions on divergence, so that probability mass does not vanish towards the “edges” of the state space. Here these restrictions will take the form of tightness conditions, which are, in turn, connected to order boundedness of trajectories via Proposition 3.1. We begin in this section by treating discrete time. Later, in Section 7, we will cover continuous time in a similar environment.

6.1. Discrete Time Theory. As before, we take (S, \leq) to be a partially ordered Polish space and \leq_{sd} to be stochastic dominance on $\mathcal{P}(S)$. The space (S, \leq) is assumed to satisfy Assumption 3.1. A distribution ϕ is called *excessive* for a stochastic kernel P on S if $\phi P \leq_{\text{sd}} \phi$ and *deficient* for P if $\phi \leq_{\text{sd}} \phi P$. As usual, P is called *Feller* if $Ph \in \text{cbS}$ whenever $h \in \text{cbS}$.

Let P be a stochastic kernel on S . As in [Kamihigashi and Stachurski \(2014\)](#), we say that P is *order reversing* if there exist two independent P -Markov processes $(X_t)_{t \in \mathbb{Z}_+}$ and $(X'_t)_{t \in \mathbb{Z}_+}$ on a common probability space such that, for every $x, x' \in S$ with $x' \leq x$, there exists a $t \in \mathbb{Z}_+$ with $\mathbb{P}_{x, x'}\{X_t \leq X'_t\} > 0$.

Intuitively, order reversing means that no initial ordering is permanent: regardless of how the two chains start, randomness can eventually reverse their relative positions. This generalizes the MMC of [Hopenhayn and Prescott \(1992\)](#), which requires mixing at the least and greatest elements of the state space, to settings where such boundary elements need not exist.

We can now state our main result for this section.

Theorem 6.1. *Let P be an increasing stochastic kernel on S . If P is order reversing, then P is globally stable if and only if P is bounded in probability.*

The key idea behind the proof of Theorem 6.1 is that boundedness in probability, combined with order reversing, provides a form of monotone mixing that leads to asymptotic contractivity. Crucially, the asymptotic contractivity is with respect to a metric defined via monotone test functions (the Bhattacharya metric), so a monotone form of mixing suffices—we do not need anything as strong as irreducibility. Global stability then follows from Theorem 3.1, since boundedness in probability implies the existence of a tight trajectory.

To place Theorem 6.1 in context, it is helpful to compare it to Theorems 1 and 2 of [Kamihigashi and Stachurski \(2014\)](#), where it is shown that an increasing and order reversing stochastic kernel P on S is globally stable whenever

- (KS1) P is bounded in probability and has an excessive or a deficient distribution, or
- (KS2) P is bounded in probability and Feller.

Theorem 6.1 improves both (KS1) and (KS2). In particular, it shows that boundedness in probability is enough for global stability when P is increasing and order reversing. We

do not need to separately check for the existence of an excessive or deficient distribution, as in (KS1), nor do we require the Feller condition, as in (KS2).

Theorem 6.1 implies an ergodicity result: If P is increasing, order reversing, and is bounded in probability, then P has a unique stationary distribution ϕ^* and

$$\mathbb{P} \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} h(X_t) = \phi^*(h) \right\} = 1$$

for all $h \in ibS$, and also for all $h \in cbS$. This follows from Theorem 6.1 and Proposition 4.1 in [Kamihigashi and Stachurski \(2016\)](#).

6.2. Application: Learning with Belief Shocks. In macroeconomic and financial modeling, belief shocks are used to generate persistent macroeconomic effects, such as countercyclical uncertainty, uncertainty spikes, and forecast biases in inflation data (see, e.g., [Orlik and Veldkamp \(2014\)](#), [Cogley and Sargent \(2008\)](#), or [Suda \(2010\)](#)). These shocks often represent exogenous events that degrade accumulated information, or replacement of incumbents by new arrivals. In this section we use Theorem 6.1 to study stability and equilibrium properties in a model of Bayesian learning subject to belief shocks.

In the model, an agent faces uncertainty about a fixed hidden state $\theta \in \{\ell, h\}$. Let $\pi_t \in (0, 1)$ represent the agent's belief that $\theta = h$ at time t . In seeking global stability and ergodicity, the end points of the unit interval are omitted from the state space because both are trivial stationary points. We seek to understand the evolution of beliefs on $(0, 1)$, which represents the dynamics of interest for the applications discussed above. Notice that this precludes use of the methods pioneered by [Hopenhayn and Prescott \(1992\)](#) and extended in Section 5, since the state has no least and greatest element.

The state π_t is primarily driven by Bayesian updating while observing IID signals $(Z_t)_{t \in \mathbb{N}}$ drawn from an unknown distribution. If $\theta = h$, then the sequence of signals is drawn from density f_h . If $\theta = \ell$, then the sequence is drawn from density f_ℓ . Let $L(z) := f_h(z)/f_\ell(z)$ denote the likelihood ratio. Upon observing signal Z_{t+1} , beliefs are updated according to Bayes' rule:

$$\pi_{t+1} = \frac{\pi_t f_h(Z_{t+1})}{\pi_t f_h(Z_{t+1}) + (1 - \pi_t) f_\ell(Z_{t+1})}. \quad (6.1)$$

Rather than studying π_t directly, we work with the log-odds transformation $\eta_t := \ln(\pi_t/(1 - \pi_t))$. Simple algebra shows that, under this transformation, the dynamics for π_t become

$$\eta_{t+1} = \eta_t + \xi_{t+1}, \quad \text{where} \quad \xi_{t+1} := \ln L(Z_{t+1}). \quad (6.2)$$

The state space for $(\eta_t)_{t \in \mathbb{Z}_+}$ is \mathbb{R} .

We introduce occasional belief shocks that reset the agent's posterior according to a stochastic kernel Q on \mathbb{R} . Under these shocks, the belief process $(\eta_t)_{t \in \mathbb{Z}_+}$ evolves on \mathbb{R} according to

$$\eta_{t+1} = I_{t+1} \cdot R_{t+1} + (1 - I_{t+1}) \cdot (\eta_t + \xi_{t+1}), \quad (6.3)$$

where $(I_t)_{t \in \mathbb{N}}$ is the belief reset indicator process and $R_{t+1} \sim Q(\eta_t, \cdot)$ is the reset value. The stochastic kernel Q is assumed to be increasing. Intuitively, this means that beliefs have some positive correlation over the reset. It includes the special case where reset values are IID. The reset indicator process is IID and each I_t is drawn from Bernoulli(ρ).

In what follows, we specialize to $\theta = h$, so that each Z_t is drawn from f_h . (Analysis under $\theta = \ell$ is very similar.) Let P denote the stochastic kernel corresponding to (6.3) under this specialization.

Lemma 6.1. *The stochastic kernel P for the belief process is increasing.*

Proof. Take $\eta_1 \leq \eta_2$. Since Q is increasing, $Q(\eta_1, \cdot) \leq_{\text{sd}} Q(\eta_2, \cdot)$. By Strassen's theorem (see, e.g., Lindvall (2002)), there exists a pair of random variables R_1 and R_2 such that the first component R_1 is drawn from $Q(\eta_1, \cdot)$, the second component R_2 is drawn from $Q(\eta_2, \cdot)$, and $R_1 \leq R_2$ almost surely. Let I and ξ be drawn from their respective distributions, so that

$$\eta'_i := I \cdot R_i + (1 - I) \cdot (\eta_i + \xi) \text{ has distribution } P(\eta_i, \cdot) \text{ for } i = 1, 2.$$

If $I = 0$, then $\eta'_1 = \eta_1 + \xi \leq \eta_2 + \xi = \eta'_2$. If $I = 1$, then $\eta'_1 = R_1$ and $\eta'_2 = R_2$. In both cases, $\eta'_1 \leq \eta'_2$. As a result, for any $h \in \text{ibS}$,

$$(Ph)(\eta_1) = \mathbb{E}h(\eta'_1) \leq \mathbb{E}h(\eta'_2) = (Ph)(\eta_2).$$

This proves that P is increasing. □

We impose the following conditions:

- (A1) The log-likelihood ratio has finite variance: $\sigma^2 = \text{Var}(\ln L(Z)) < \infty$.
- (A2) For all $\varepsilon > 0$, there exists a compact $K \subset \mathbb{R}$ such that $\inf_{x \in \mathbb{R}} Q(x, K) \geq 1 - \varepsilon$.
- (A3) For every $M > 0$, we have $\mathbb{P}\{L(Z) > M\} > 0$ when $Z \sim f_h$.

Assumptions (A1) and (A3) are naturally satisfied in common settings. (For example, if f_h and f_ℓ are Gaussian densities with means $\mu_h > \mu_\ell$ and common variance σ^2 , then $\ln L(z) = (\mu_h - \mu_\ell)(z - \bar{\mu})/\sigma^2$ where $\bar{\mu} = (\mu_h + \mu_\ell)/2$. This has finite variance and is

unbounded above since z has full support.) Assumption (A2) is a tightness condition on the reset kernel and holds trivially when resets are IID.

Lemma 6.2. *Under (A1) and (A2), P is bounded in probability.*

Proof. Fix $x \in \mathbb{R}$ and $\varepsilon > 0$. By (A2), choose $K = [a, b]$ such that $Q(y, K) \geq 1 - \varepsilon/3$ for all y . Choose T large enough that $(1 - \rho)^T < \varepsilon/3$. For $t \geq T$, let N_t denote the number of updates since the last reset prior to t . The probability that no reset has occurred by time t is $(1 - \rho)^t \leq (1 - \rho)^T < \varepsilon/3$. Conditional on a reset occurring, the post-reset state lies in K with probability at least $1 - \varepsilon/3$. Conditional on $N_t = n$ and the last reset landing in K , we have $\eta_t \in [a + S_n, b + S_n]$ where $S_n = \xi_1 + \dots + \xi_n$. The random variable N_t is stochastically dominated by $N \sim \text{Geometric}(\rho)$. By Chebyshev's inequality, $\mathbb{P}\{|S_n| > c\} \leq n\sigma^2/c^2$. Choose c large enough that

$$\sum_{n=0}^{\infty} \rho(1 - \rho)^n \cdot \frac{n\sigma^2}{c^2} = \frac{\sigma^2(1 - \rho)}{c^2\rho} < \frac{\varepsilon}{3}.$$

Then $\mathbb{P}_x\{\eta_t \notin [a - c, b + c]\} < \varepsilon$ for all $t \geq T$. For $t < T$, each $P^t(x, \cdot)$ is a single distribution and hence tight. A finite union of compact sets covers these, completing the proof. \square

Lemma 6.3. *Under (A3), the kernel P is order reversing.*

Proof. Let (η_t) and (η'_t) be independent P -Markov processes with initial conditions $\eta > \eta'$. We will show that $\mathbb{P}\{\eta_1 < \eta'_1\} > 0$. To this end, let (I, R, ξ) and (I', R', ξ') be the independent shocks for the two processes, where $R \sim Q(\eta, \cdot)$ and $R' \sim Q(\eta', \cdot)$. Consider the event $A = \{I = 1, I' = 0, \xi' > R - \eta'\}$. On A , process 1 resets to $\eta_1 = R$ while process 2 updates to $\eta'_1 = \eta' + \xi' > R = \eta_1$, so $\eta_1 < \eta'_1$. Since the two processes are independent, I, R , and ξ' are mutually independent. We have $\mathbb{P}\{I = 1\} = \rho > 0$, $\mathbb{P}\{I' = 0\} = 1 - \rho > 0$, and, for every realization $R = r$, $\mathbb{P}\{\xi' > r - \eta'\} = \mathbb{P}\{L(Z) > e^{r - \eta'}\} > 0$ by (A3). Hence $\mathbb{P}(A) > 0$, proving that P is order reversing. \square

Combining the preceding lemmas with Theorem 6.1, we obtain global stability.

Proposition 6.1. *Under (A1)–(A3), the belief shock kernel P is globally stable on $(\mathcal{P}(\mathbb{R}), \beta)$.*

Proof. The state space $S = \mathbb{R}$ satisfies Assumption 3.1 since order intervals $[a, b]$ are compact and compact sets are bounded, hence order-bounded. By Lemma 6.2, P is bounded in probability. By Lemma 6.1, P is increasing. By Lemma 6.3, P is order reversing. Therefore, by Theorem 6.1, P is globally stable. \square

Figure 3 illustrates a simulation of the belief process in π -space, with Gaussian signals ($\mu_h = 0.3, \mu_\ell = 0, \sigma = 1$), reset probability $\rho = 0.04$, and IID Gaussian resets with mean zero and standard deviation 0.5 in log-odds space. The time series shows the characteristic pattern of gradual learning (upward drift, since the true state is $\theta = h$) interrupted by belief resets that pull the posterior back toward $\pi = 0.5$. The empirical stationary distribution, computed from 200,000 periods, is right-skewed, reflecting the positive drift from Bayesian updating under the true state. Convergence of the empirical distribution to the true stationary distribution follows from the ergodicity result implied by Theorem 6.1.

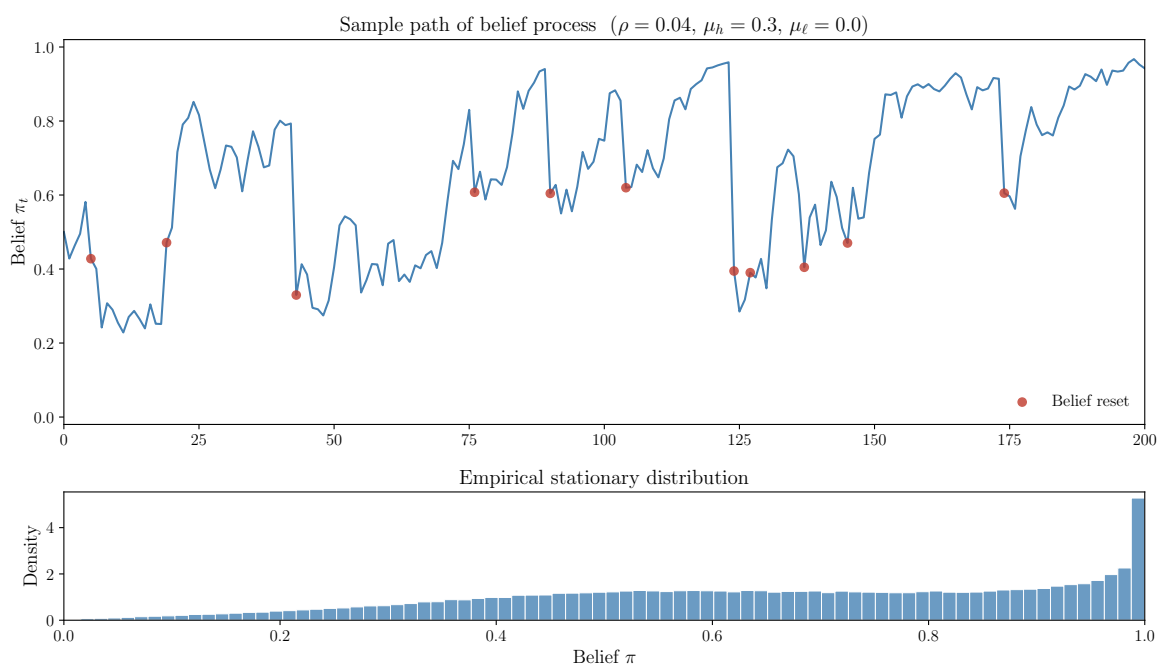


FIGURE 3. Simulation of the belief shock model in Section 6.2.

7. NONCOMPACT STATE SPACE, CONTINUOUS TIME

Now we state a continuous time result that is closely related to Theorem 6.1. As before, we take (S, \leq) to be a partially ordered Polish space satisfying Assumption 3.1 and \leq_{sd} to be stochastic dominance on $\mathcal{P}(S)$. We introduce a continuous-time mixing condition in Section 7.1 and then apply it to a model of income dynamics in Section 7.2.

7.1. Continuous Time Theory. We start with a mixing condition: a transition probability function (P_t) on S will be called *weakly order mixing* if there exists an $(S \times S)$ -valued strong Markov process $((X_t), (X'_t))$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

- (i) (X_t) and (X'_t) are both (P_t) -Markov, and
- (ii) the stopping time $\tau := \inf\{t \geq 0 : X_t \leq X'_t\}$ is \mathcal{F} -measurable with

$$\mathbb{P}_{x,x'}\{\tau < \infty\} = 1 \quad \text{for all } x, x' \in S. \quad (7.1)$$

As with order reversing, the intuition is that randomness eventually induces the order $X_t \leq X'_t$, regardless of initial states x and x' . The key difference between weak order mixing and the order reversing property is that the two processes need not be independent. This additional flexibility is used in the applications below, where coupling jump times (processes jump independently but at the same times) allows us to verify weak order mixing.

The strong Markov and measurability restrictions in the definition of weak order mixing hold in essentially all applications of interest, such as when the process has càdlàg paths under the standard constructions (see, e.g., [Ethier and Kurtz \(2009\)](#)). Thus, the real restriction in weak order mixing is (7.1).

We can now state our main continuous time result for a transition probability function (P_t) on S .

Theorem 7.1. *Let (P_t) be increasing. If (P_t) is weakly order mixing, then (P_t) is globally stable if and only if (P_t) has at least one tight trajectory.*

As in the discrete-time case (Theorem 6.1), global stability is characterized by the combination of mixing properties and tightness. We apply this result to piecewise deterministic Markov processes in Section 8 below.

7.2. Application: Income Dynamics Part I. In the model of [Gabaix et al. \(2016\)](#), individual log income follows a diffusion with constant drift, punctuated by jumps and random resets. This model is rather stylized. In practice, most workers' incomes exhibit neither continuous drift nor positive quadratic variation (e.g., influence from Brownian motion). Instead, most real-world income processes remain constant for nonzero time intervals and then jump due to promotions, layoffs, retirement, etc. Here we study a model of income that mimics these dynamics. (An additional advantage of the model is ease of simulation: no Euler–Maruyama style discretization is required.) We analyze stationarity and asymptotics using Theorem 7.1. Then we examine tail properties.

Let X_t represent log income at time t . In our model, *income shocks* arrive at rate $\lambda_1 > 0$ and represent raises, promotions, or better job offers. At such an event, log income jumps from x to $x + \eta$, where $\eta > 0$ is an IID draw from a distribution μ on $(0, \infty)$. Resets arrive at rate $\lambda_2 > 0$ and represent job loss, health shocks, retirement, or career disruptions: the state jumps from x to $h(x) + \zeta$, where ζ is an IID draw from a *reset distribution* ν on \mathbb{R} and $h \in \text{ib}\mathbb{R}$ is a *reset function*. The assumption that h is increasing represents the idea that high-wage workers are likely to be better off after career disruptions than low wage workers. Pure resets can be obtained by setting h equal to some constant.

By the superposition property of Poisson processes, the two independent shock streams can be merged into a single Poisson process with rate $\lambda := \lambda_1 + \lambda_2$. At each jump, a reset occurs with probability $p := \lambda_2/\lambda$ and an income shock with probability $q := \lambda_1/\lambda$. Income is constant between jumps, so the process satisfies $X_t = Z_{N_t}$, where $(N_t)_{t \geq 0}$ is a Poisson process with parameter λ that counts the number of jumps by time t , and $(Z_n)_{n \in \mathbb{Z}_+}$ is the *embedded chain*, defined by

$$Z_{n+1} = \begin{cases} Z_n + \eta_{n+1} & \text{with probability } q \text{ (income shock),} \\ h(Z_n) + \zeta_{n+1} & \text{with probability } p \text{ (reset).} \end{cases} \quad (7.2)$$

Here $(\eta_n)_{n \in \mathbb{N}}$ and $(\zeta_n)_{n \in \mathbb{N}}$ are independent and IID, drawn from μ and ν respectively. We can now state the following stability result.

Proposition 7.1. *If the support of ν is larger than the range of h , then the log income process is globally stable: there exists a unique stationary distribution ϕ^* on \mathbb{R} and*

$$\lim_{t \rightarrow \infty} \beta(\phi P_t, \phi^*) = 0 \quad \text{for all } \phi \in \mathcal{P}(\mathbb{R}). \quad (7.3)$$

Proposition 7.1 obviously implies that income itself is globally stable, with unique stationary density ψ^* on $(0, \infty)$ given by $\psi^*(y) = \pi(\ln y)/y$, where π is the density of ϕ^* .

To clarify the assumptions, let $\underline{h} = \inf h$ and $\bar{h} = \sup h$. The condition on ν says that there exist a, b in the support of ν with $b - a > \bar{h} - \underline{h}$. If h is constant (pure resets), this just means that the reset shock is nondegenerate. If the support of ν is all of \mathbb{R} , then the condition is always satisfied, since h is assumed to be bounded. (For this case the boundedness assumption on h can potentially be weakened via a drift condition on $x \mapsto |h(x)|$. We leave this analysis to future research.)

Proof of Proposition 7.1. Let $(P_t)_{t \geq 0}$ be the transition probability function of the jump process $(X_t)_{t \geq 0}$. We verify that (P_t) is increasing, weakly order mixing, and has a tight trajectory. Global stability then follows from Theorem 7.1.

(P_t) is increasing. Fix $t > 0$. On a common probability space, we construct two copies of the income process from initial conditions $x \leq x'$ using the same jump times, the same shock-type indicators (income vs. reset), and the same shock values. For the embedded chains, $Z_0 = x \leq x' = Z'_0$. If $Z_n \leq Z'_n$ and the $(n+1)$ -th jump is an income shock, then

$$Z_{n+1} = Z_n + \eta_{n+1} \leq Z'_n + \eta_{n+1} = Z'_{n+1},$$

since $\eta_{n+1} > 0$. If it is a reset, then

$$Z_{n+1} = h(Z_n) + \zeta_{n+1} \leq h(Z'_n) + \zeta_{n+1} = Z'_{n+1},$$

since h is increasing. By induction $Z_n \leq Z'_n$ for all n , and therefore $X_t = Z_{N_t} \leq Z'_{N_t} = X'_t$. It follows directly that $P_t(x, \cdot) \leq_{\text{sd}} P_t(x', \cdot)$.

(P_t) is weakly order mixing. To prove this fact, we construct two copies using the same jump times and the same shock-type indicators, but independent shock values: (η_n, ζ_n) for the first copy and (η'_n, ζ'_n) for the second. Each marginal is (P_t) -Markov. The joint process $((X_t), (X'_t))$ is a pure jump process on \mathbb{R}^2 with constant total rate λ and hence strong Markov (Rudnicki and Tyran-Kamińska, 2017, Theorem 2.4). Since $\{(x, x') \in \mathbb{R}^2 : x \leq x'\}$ is closed, $\tau := \inf\{t \geq 0 : X_t \leq X'_t\}$ is a measurable stopping time.

Let A_n be the event that the n -th jump is a reset and $\zeta'_n - \zeta_n \geq \bar{h} - \underline{h}$. On A_n , we have

$$Z_n = h(Z_{n-1}) + \zeta_n \leq \bar{h} + \zeta_n \leq \underline{h} + \zeta'_n \leq h(Z'_{n-1}) + \zeta'_n = Z'_n.$$

The last inequality means that $\tau \leq T_n$ on A_n , where T_n is the n -th jump time. Since the shock type and shock values at step n are independent of the filtration \mathcal{F}_{n-1} ,

$$\mathbb{P}\{A_n^c \mid \mathcal{F}_{n-1}\} \leq 1 - p\delta,$$

where $\delta := \mathbb{P}\{\zeta' - \zeta \geq \bar{h} - \underline{h}\}$. Hence $\mathbb{P}\{\tau > T_n\} \leq (1 - p\delta)^n$. By our assumption on the support of ν , we have $\delta > 0$ and hence $\mathbb{P}\{\tau > T_n\} \rightarrow 0$ as $n \rightarrow \infty$. As $T_n \rightarrow \infty$ almost surely, it follows that $\mathbb{P}_{x, x'}\{\tau < \infty\} = 1$.

(P_t) has a tight trajectory. Fix $x_0 \in \mathbb{R}$ and $\epsilon > 0$. Since the shock types and values (η_n, ζ_n) are independent of the jump times, the embedded chain (Z_n) is independent of the counting process (N_t) . Therefore

$$\mathbb{P}_{x_0}\{X_t \notin D\} = \sum_{n=0}^{\infty} \mathbb{P}\{N_t = n\} \mathbb{P}_{x_0}\{Z_n \notin D\} \leq \sup_{n \geq 0} \mathbb{P}_{x_0}\{Z_n \notin D\},$$

so it suffices to find, for each $\epsilon > 0$, a compact D with $\mathbb{P}_{x_0}\{Z_n \notin D\} < \epsilon$ for every $n \geq 0$.

Let $R \sim \text{Geom}(p)$ on $\{0, 1, 2, \dots\}$ denote the number of income shocks in a generic cycle between consecutive resets, and let $W := \sum_{i=1}^R \eta_i$ be the total income accumulated in

one such cycle. Before the first reset (i.e., for $n \leq R$), every jump is an income shock, so $Z_n = x_0 + \eta_1 + \dots + \eta_n \leq x_0 + W$. At each reset step k , the post-reset state satisfies $\underline{h} + \zeta_k \leq Z_k \leq \bar{h} + \zeta_k$. Between that reset and the next, the chain only accumulates positive income shocks, so for any n falling in the same cycle,

$$\underline{h} + \zeta_k \leq Z_n \leq \bar{h} + \zeta_k + W',$$

where W' is an independent copy of W . Since (ζ_n) and (η_n) are IID sequences, the distributions of ζ_k and W' do not depend on the cycle. It follows that the sequence of distributions $(P^n(x_0, \cdot))_{n \in \mathbb{N}}$ of the embedded chain is order bounded in $(\mathcal{P}(\mathbb{R}), \leq_{\text{sd}})$, and hence tight by Proposition 3.1. Global stability now follows from Theorem 7.1. \square

Figure 4 illustrates a simulation of income $Y_t := \exp(X_t)$ with $h \equiv 0$, $\mu = \text{Exp}(20)$ (mean raise of 5% in log income), $\lambda_1 = 1.0$, $\lambda_2 = 0.1$, and IID Gaussian reset shocks $\zeta \sim N(0, 0.09)$. Income shocks are frequent, producing a staircase pattern of gradual accumulation depleted by sharp resets. The empirical stationary distribution, computed from 200,000 time units, is right-skewed with a heavy tail.

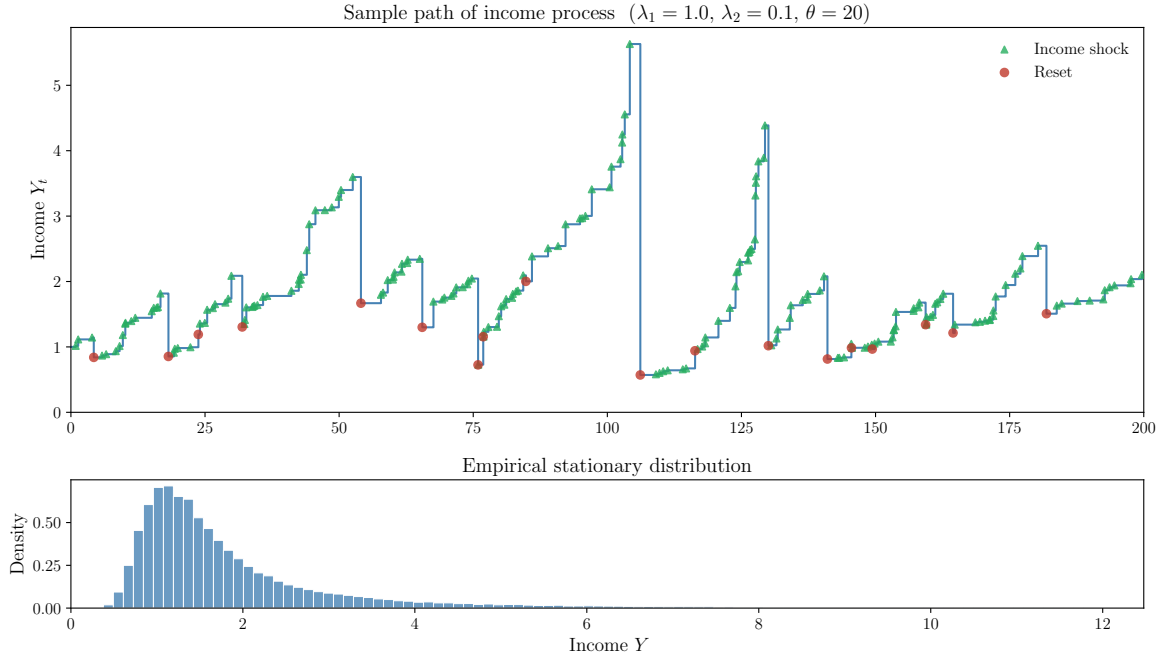


FIGURE 4. Simulation of the pure jump income model in Section 7.2.

The next result specializes the model to obtain a closed-form stationary distribution and a Pareto tail. Income shocks are exponentially distributed and resets return the agent to a fixed entry level.

Corollary 7.1. *Let $h \equiv x_0$, $\nu = \delta_0$, and $\mu = \text{Exp}(\theta)$ for some $\theta > 0$. In this setting, for the income level $Y = e^X$, the tail of the unique stationary distribution is Pareto: for $y > y_0 := e^{x_0}$,*

$$\mathbb{P}(Y > y) = q \left(\frac{y}{y_0} \right)^{-\alpha} \quad \text{with} \quad \alpha := \frac{\lambda_2 \theta}{\lambda_1 + \lambda_2}. \quad (7.4)$$

Proof. Since income is constant between jumps and the jump rate λ is independent of the state, the stationary distribution of the continuous-time process (X_t) coincides with that of the embedded chain (Z_n) . The chain regenerates at each reset: the post-reset state is always x_0 . Between consecutive resets, the chain accumulates M income shocks before the next reset, where $M \sim \text{Geom}(p)$ on $\{0, 1, 2, \dots\}$. In stationarity, the state is $x_0 + S_M$ where $S_M = \eta_1 + \dots + \eta_M$.

With probability p , $M = 0$ and $S_M = 0$, giving an atom at x_0 . For $M \geq 1$, conditional on $M = m$, the sum $S_m \sim \text{Gamma}(m, \theta)$. Summing over m , the density of S_M on $(0, \infty)$ is

$$f_{S_M}(s) = \sum_{m=1}^{\infty} pq^m \cdot \frac{\theta^m s^{m-1}}{(m-1)!} e^{-\theta s} = pq\theta e^{-\theta s} \sum_{k=0}^{\infty} \frac{(q\theta s)^k}{k!} = pq\theta e^{-p\theta s}$$

for $s > 0$. Since ϕ^* is the distribution of $x_0 + S_M$, it follows that the density of ϕ^* on (x_0, ∞) is $\pi(x) := pq\theta e^{-p\theta(x-x_0)}$. Thus, for $y \geq y_0$,

$$\mathbb{P}(Y > y) = \int_{\ln y}^{\infty} pq\theta e^{-p\theta(x-x_0)} dx = q e^{-p\theta(\ln y - x_0)} = q \left(\frac{y}{y_0} \right)^{-\alpha}. \quad \square$$

8. PIECEWISE DETERMINISTIC MARKOV PROCESSES

In this section we develop a general stability result for piecewise deterministic Markov processes (PDMPs) that can be applied off the shelf in a range of settings. The theory is presented in Section 8.1 and applied to a model of income dynamics with deterministic drift in Section 8.2.

8.1. Theory. Let (S, \leq) be a partially ordered Polish space satisfying Assumption 3.1. A *piecewise deterministic Markov process* (PDMP) on S is generated by the following characteristics:

- (i) a *semi-flow* $\Phi : S \times \mathbb{R}_+ \rightarrow S$,
- (ii) a *jump intensity* $\lambda \in (0, \infty)$,
- (iii) a *jump shock space* U and shock distribution μ on U , and
- (iv) a *jump function* $F : S \times U \rightarrow S$.

Here U is a Polish space and the jump function F is assumed to be measurable. The map Φ is called a *semi-flow* because it satisfies $\Phi(x, 0) = x$ and $\Phi(\Phi(x, s), t) = \Phi(x, s + t)$ for all $x \in S$ and $s, t \geq 0$. We assume that Φ is continuous. Informally, the process evolves as follows:

- (i) Follow the deterministic flow: $X_t = \Phi(X_0, t)$ until the first jump.
- (ii) At the first jump time T , draw ξ from μ and set $X_T = F(X_{T-}, \xi)$.
- (iii) Repeat from the new state X_T .

In most applications, the semi-flow is the solution to an ODE. For example, if $S = \mathbb{R}$ and the deterministic dynamics have the linear form $\dot{x} = ax$, then $\Phi(x, t) = x \exp(at)$.

To formalize the construction of the associated Markov process $(X_t)_{t \geq 0}$, we first collect the following components:

- (i) an S -valued random element X_0 with distribution $\psi \in \mathcal{P}(S)$,
- (ii) a real-valued IID sequence $(E_n)_{n \in \mathbb{N}}$ with common distribution $\text{Exp}(\lambda)$, and
- (iii) a U -valued IID sequence $(\xi_n)_{n \in \mathbb{N}}$ with common distribution μ ,

All random elements live on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We first build the *embedded chain* $(Z_n)_{n \in \mathbb{Z}_+}$ via

$$Z_0 = X_0 \quad \text{and} \quad Z_{n+1} = F(\Phi(Z_n, E_{n+1}), \xi_{n+1}).$$

We define the jump times $T_0 := 0$ and $T_n := \sum_{i=1}^n E_i$, while $N_t := \max\{n \geq 0 : T_n \leq t\}$ counts the number of jumps up until time t . Finally, we construct $(X_t)_{t \geq 0}$ via

$$X_t = \Phi(Z_{N_t}, t - T_{N_t}) \quad t \geq 0.$$

A more intuitive representation is

$$X_t = \Phi(Z_n, t - T_n) \quad \text{for } t \in [T_n, T_{n+1}) \text{ and } n \in \mathbb{Z}_+.$$

Sample paths are càdlàg, with smooth deterministic segments broken by jumps. The process (X_t) obeys the strong Markov property (Rudnicki and Tyran-Kamińska, 2017, Theorem 2.4). We work with the following assumptions.

Assumption 8.1 (PDMP Monotonicity). The primitives are such that

- (i) For all $t \geq 0$, the function $x \mapsto \Phi(x, t)$ is order preserving on S .
- (ii) For all $u \in U$, the function $x \mapsto F(x, u)$ is order preserving on S .

In the next assumption, ξ and ξ' are independent draws from μ .

Assumption 8.2 (PDMP Mixing). There exist measurable functions $f_1, f_2 : U \rightarrow S$ such that

$$f_1(\xi) \leq F(x, \xi) \leq f_2(\xi) \text{ for all } x \in S \text{ and } \mathbb{P}\{f_2(\xi) \leq f_1(\xi')\} > 0.$$

The transition probability function corresponding to the PDMP $(X_t)_{t \geq 0}$ is given by

$$P_t(x, B) = \mathbb{P}\{X_t \in B \mid X_0 = x\} \quad (t \in \mathbb{R}_+, x \in S, B \in \mathcal{B}).$$

Under the stated monotonicity and mixing conditions, we can prove the following result.

Theorem 8.1. *If Assumptions 8.1 and 8.2 hold, then the PDMP is globally stable, with a unique stationary distribution ϕ^* and*

$$\lim_{t \rightarrow \infty} \beta(\phi P_t, \phi^*) = 0 \quad \text{for all } \phi \in \mathcal{P}(S). \quad (8.1)$$

The proof is based on Theorem 7.1 and can be found in Section A.7. The strategy is to verify the three hypotheses of Theorem 7.1. Monotonicity is established by coupling two copies of the PDMP with the same jump times and the same shocks: if one copy starts above the other, then Assumption 8.1 ensures that it stays above through both the deterministic flow and the jumps. Boundedness in probability follows from the pathwise bounds in Assumption 8.2, which prevent post-jump states from escaping to infinity, while the exponential inter-arrival times prevent the deterministic flow from carrying the process too far between jumps. The most interesting step is weak order mixing, which uses a different coupling: the two copies share jump times but receive independent shocks. The condition $\mathbb{P}\{f_2(\xi) \leq f_1(\xi')\} > 0$ in Assumption 8.2 then guarantees that, at each shared jump, there is a positive probability that the relative ordering of the two processes is reversed, regardless of their current states.

Now we turn to applications.

8.2. Application: Income Dynamics Part II. In Section 7.2 we analyzed income dynamics through a pure jump process. We now consider a model of income dynamics more closely aligned with Gabaix et al. (2016), where log income has a continuous drift component. Unlike Gabaix et al. (2016), we continue to avoid the assumption of positive quadratic variation (e.g., Brownian motion), which is difficult to motivate from data. To make the proofs slightly simpler, we drop the positive income shocks in Section 7.2 (while retaining the resets).

In our model, log income drifts deterministically between jumps, following the ODE $\dot{x} = g(x)$ for some locally Lipschitz function $g : \mathbb{R} \rightarrow \mathbb{R}$. At jump times, which occur at rate λ , the state jumps from x to $F(x, \zeta)$, where ζ is an independent draw from ν .

More formally, we study a PDMP on $S = \mathbb{R}$ with the following primitives:

- (i) a semi-flow $\Phi : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ generated by the ODE $\dot{x} = g(x)$,
- (ii) a constant jump intensity $\lambda > 0$,
- (iii) a shock distribution ν on \mathbb{R} , and
- (iv) a jump function $F(x, z) = h(x) + z$, where $h \in ibS$.

Between jumps, $X_t = \Phi(Z_{N_t}, t - T_{N_t})$, where $(E_n)_{n \in \mathbb{N}}$ are IID $\text{Exp}(\lambda)$ inter-arrival times and N_t counts jumps up to time t (as in Section 8). At each jump time, the state moves to $h(X_{T^-}) + \zeta$. The embedded chain is $Z_{n+1} = h(\Phi(Z_n, E_{n+1})) + \zeta_{n+1}$.

The next proposition shows global stability under essentially the same assumptions as Proposition 7.1.

Proposition 8.1. *If the support of ν is larger than the range of h , then the log income process is globally stable: there exists a unique stationary distribution ϕ^* and*

$$\lim_{t \rightarrow \infty} \beta(\phi P_t, \phi^*) = 0 \quad \text{for all } \phi \in \mathcal{P}(\mathbb{R}). \quad (8.2)$$

Proof. We verify Assumptions 8.1 and 8.2. Then we apply Theorem 8.1.

Assumption 8.1(i): the semiflow is order preserving. Since g is locally Lipschitz, the Picard–Lindelöf theorem gives uniqueness of solutions. Fix $x_1 < x_2$ and $t > 0$. Seeking a contradiction, suppose that $\Phi(x_1, s) \geq \Phi(x_2, s)$ for some $s > 0$. Since $\Phi(x_1, 0) < \Phi(x_2, 0)$ and Φ is continuous, the intermediate value theorem yields $t_0 \in (0, s]$ with $\Phi(x_1, t_0) = \Phi(x_2, t_0)$. Both trajectories pass through the same state at time t_0 and the ODE is autonomous, so uniqueness forces them to agree for all time, contradicting $x_1 \neq x_2$. Hence $x \mapsto \Phi(x, t)$ is strictly increasing for every $t \geq 0$.

Assumption 8.1(ii): the jump function is order preserving. For each $\zeta \in \mathbb{R}$, the map $x \mapsto F(x, \zeta) = h(x) + \zeta$ is increasing since h is assumed to be increasing.

Assumption 8.2: bounding functions and coupling. Set $f_1(\zeta) = \underline{h} + \zeta$ and $f_2(\zeta) = \bar{h} + \zeta$, where $\underline{h} := \inf h$ and $\bar{h} := \sup h$. By construction, $f_1(\zeta) \leq F(x, \zeta) \leq f_2(\zeta)$ for all $x \in \mathbb{R}$. For the coupling condition, let ζ, ζ' be independent draws from ν . Then

$$\mathbb{P}\{f_2(\zeta) \leq f_1(\zeta')\} = \mathbb{P}\{\zeta' - \zeta \geq \bar{h} - \underline{h}\} > 0,$$

where the strict inequality follows from our assumption on the support of ν . Global stability now follows from Theorem 8.1. \square

If we specialize to constant drift rate g and a pure reset, we can easily obtain a Pareto tail, in a result analogous to Corollary 7.1. The tail exponent is λ/g , and hence entirely determined by the first-order balance between deterministic growth and random resetting. Further details are available from the authors on request.

Figure 5 illustrates a simulation with constant drift $g(x) = \mu = 0.05$, jump rate $\lambda = 0.15$, pure reset $h \equiv 0$, and IID Gaussian reset shocks $\zeta \sim N(0, 0.09)$. The time series of income $Y_t = e^{X_t}$ displays characteristic PDMP dynamics: smooth exponential growth between jumps, interrupted by sharp resets. The empirical stationary distribution is right-skewed with a Pareto tail of exponent $\lambda/\mu = 3$.

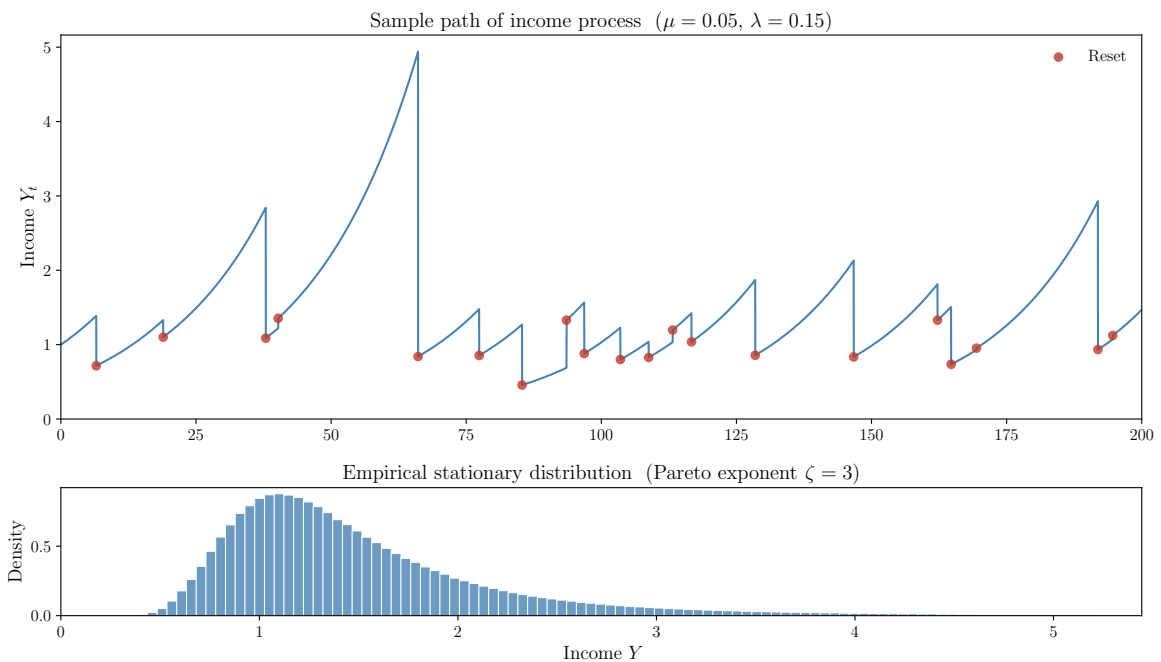


FIGURE 5. Simulation of the income dynamics model in Section 8.2.

9. CONCLUSION

We have shown that an order-preserving semigroup on a complete preordered metric space is globally stable if and only if it is asymptotically contractive and has at least one order-bounded trajectory. When applied to monotone Markov models on partially ordered Polish spaces, this yields a necessary and sufficient condition for the existence, uniqueness, and global stability of stationary distributions: asymptotic contractivity combined with tightness of at least one trajectory.

These results unify and extend earlier work by [Hopenhayn and Prescott \(1992\)](#), [Kamihigashi and Stachurski \(2014\)](#), and other authors, weakening assumptions, introducing new methods, and removing extraneous side conditions. The abstract semigroup framework covers both discrete and continuous time, as well as nonlinear Markov operators. Our applications to wage dynamics, Bayesian learning with belief shocks, and income dynamics with Pareto tails illustrate the breadth of the approach.

Many directions remain open. The mixing conditions used in this paper could potentially be weakened further, for example by allowing state-dependent jump intensities in the PDMP framework. Our Pareto tail result for income is specialized and may hold under weaker assumptions. The relationship between tightness and order boundedness established here may also prove useful in other settings involving monotone economic dynamics.

APPENDIX A. PROOFS

This appendix collects remaining proofs.

A.1. Proof of Theorem 2.1. Throughout this section, Assumptions 2.1 and 2.2 are in force.

A.1.1. *(ii) implies (i).* Let (T_t) be globally stable with stationary point $x^* \in X$. Then the sequence given by $T_t x^*$ for all $t \in \mathbb{T}$ is an order-bounded trajectory of (T_t) . Moreover, picking arbitrary $x, y \in X$ and applying the triangle inequality,

$$d(T_t x, T_t y) \leq d(T_t x, x^*) + d(x^*, T_t y) \rightarrow 0 \quad (t \rightarrow \infty),$$

where the convergence holds by global stability. Hence (T_t) is asymptotically contractive.

A.1.2. *(i) implies (ii).* Now we assume that (T_t) is asymptotically contractive on X with at least one order-bounded trajectory. Due to the last property, we can take $a, b, \bar{x} \in X$ such that (2.2) holds.

Lemma A.1. *For all $\epsilon > 0$, there exists a $\tau \in \mathbb{T}$ such that*

$$s, t \in \mathbb{T} \text{ with } \tau \leq s, t \quad \implies \quad d(T_s \bar{x}, T_t \bar{x}) < \epsilon. \quad (\text{A.1})$$

Proof. Fix $\epsilon > 0$ and then use asymptotic contractivity to choose $\tau \in \mathbb{T}$ such that $d(T_\tau a, T_\tau b) < \epsilon$. For $s, t \geq \tau$ we have

$$d(T_s \bar{x}, T_t \bar{x}) = d(T_\tau T_{s-\tau} \bar{x}, T_\tau T_{t-\tau} \bar{x}) \leq d(T_\tau a, T_\tau b) < \epsilon.$$

The first inequality is due to the diagonal property of d , combined with the fact that $T_{s-\tau} \bar{x}$ and $T_{t-\tau} \bar{x}$ both lie in $[a, b]$, which in turn implies that

$$T_\tau a \leq T_\tau T_{s-\tau} \bar{x} \leq T_\tau b \quad \text{and} \quad T_\tau a \leq T_\tau T_{t-\tau} \bar{x} \leq T_\tau b. \quad \square$$

Lemma A.2. *There exists an $x^* \in X$ such that*

$$d(T_t \bar{x}, x^*) \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty. \quad (\text{A.2})$$

Proof. Let $(t_n)_{n \in \mathbb{N}}$ be an increasing sequence in \mathbb{T} with $t_n \rightarrow \infty$ as $n \rightarrow \infty$. Consider the sequence defined by $x_n := T_{t_n} \bar{x}$. This sequence is Cauchy because, given $\epsilon > 0$, we can choose $\tau \in \mathbb{T}$ such that (A.1) holds and then $d(x_n, x_m) < \epsilon$ for all n, m such that $t_n, t_m \geq \tau$. Since X is complete, it follows that (x_n) converges to a limit x^* . That is,

$$d(T_{t_n} \bar{x}, x^*) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty. \quad (\text{A.3})$$

We claim that this can be extended to (A.2). To see this, fix $\epsilon > 0$ and choose τ as in (A.1). For $t \geq \tau$ and $n \in \mathbb{N}$ such that $t_n \geq \tau$, we have

$$d(T_t \bar{x}, x^*) \leq d(T_t \bar{x}, T_{t_n} \bar{x}) + d(T_{t_n} \bar{x}, x^*) < \epsilon + d(T_{t_n} \bar{x}, x^*)$$

Applying (A.3) and taking the limit in n completes the argument. \square

Lemma A.3. *The point x^* in Lemma A.2 is stationary for (T_t) .*

Proof. Fix $s, t \in \mathbb{T}$. From the existence of an order-bounded trajectory we have $a \leq T_t \bar{x} \leq b$. From the order-preserving property of T_s we get $T_s a \leq T_{t+s} \bar{x} \leq T_s b$. Since \leq is closed, taking the limit with respect to t and applying Lemma A.2 yields

$$T_s a \leq x^* \leq T_s b \quad \text{for all } s \in \mathbb{T}. \quad (\text{A.4})$$

In addition, from (A.4) and the order-preserving property of T_t ,

$$T_{u+t} a \leq T_t x^* \leq T_{u+t} b \quad \text{for all } u, t \in \mathbb{T}. \quad (\text{A.5})$$

Combining (A.4) and (A.5) with the diagonal property (2.1), we get

$$d(x^*, T_t x^*) \leq d(T_{u+t} a, T_{u+t} b) \quad \text{for all } u \in \mathbb{T}. \quad (\text{A.6})$$

Taking the limit in u and using asymptotic contractivity yields $T_t x^* = x^*$. \square

Proof of Theorem 2.1. We have already shown that x^* is stationary for (T_t) in Lemma A.3. Regarding convergence, fix $x \in X$. Since (T_t) is asymptotically contractive and x^* is stationary, we have $d(T_t x, x^*) = d(T_t x, T_t x^*) \rightarrow 0$ as $t \rightarrow \infty$. This concludes the proof of Theorem 2.1. \square

A.2. Proof of Theorem 2.2. We only used closedness of \leq in Lemma A.3, when we proved that the limit point x^* of the order-bounded trajectory $(T_t \bar{x})$ is a fixed point of each T_t . We now prove this without closedness by fixing $t \in \mathbb{T}$ and applying the triangle inequality to obtain

$$d(x^*, T_t x^*) \leq d(x^*, T_{u+t} \bar{x}) + d(T_{u+t} \bar{x}, T_t x^*).$$

The first term converges to zero in u by Lemma A.2. If we can show that the second term also converges to zero in u then we are done. This holds because T_t is continuous and $T_u \bar{x} \rightarrow x^*$, so, as $u \rightarrow \infty$,

$$d(T_{u+t} \bar{x}, T_t x^*) = d(T_t T_u \bar{x}, T_t x^*) \rightarrow d(T_t x^*, T_t x^*) = 0.$$

A.3. Proof of Theorem 3.1. Let the conditions of Theorem 3.1 hold. Because S is a partially ordered Polish space and Assumption 3.1 is in force, $(\mathcal{P}(S), \beta)$ is a complete metric space (Kamihigashi and Stachurski, 2019, Theorem 4.1) and \leq_{sd} is closed (Kamae and Krengel, 1978). We begin with the following lemma.

Lemma A.4. *If $\Lambda \subset \mathcal{P}(S)$ is tight, then there exists an increasing sequence of order intervals $(I_n)_{n \in \mathbb{Z}_+}$ such that $\sup_{\phi \in \Lambda} \phi(I_n^c) \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. Let (ϵ_n) be a positive real sequence with $\epsilon_n \downarrow 0$. By tightness, we can choose compact set K_0 with $\sup_{\phi \in \Lambda} \phi(K_0^c) \leq \epsilon_0$. By Assumption 3.1, compact sets are order bounded, so we can also take an order interval I_0 containing K_0 with $\sup_{\phi \in \Lambda} \phi(I_0^c) \leq \epsilon_0$. Next, we choose a compact set K_1 with $\sup_{\phi \in \Lambda} \phi(K_1^c) \leq \epsilon_1$, and then, using Assumption 3.1 again, an order interval I_1 such that $I_0 \cup K_1 \subset I_1$. This order interval obeys $I_0 \subset I_1$ and $\sup_{\phi \in \Lambda} \phi(I_1^c) \leq \epsilon_1$. Continuing in this way produces an increasing sequence of order intervals $(I_n)_{n \in \mathbb{Z}_+}$ with $\sup_{\phi \in \Lambda} \phi(I_n^c) \rightarrow 0$. \square

Proof of Proposition 3.1. Suppose first that Λ is tight. We show that there exists a pair of distributions $\ell, u \in \mathcal{P}(S)$ such that $\ell \leq \phi \leq u$ for all $\phi \in \Lambda$. By Lemma A.4, we can take an increasing sequence of order intervals $(I_n)_{n \in \mathbb{Z}_+}$ and a positive real sequence (ϵ_n) with $\epsilon_n \downarrow 0$ and $\sup_{\phi \in \Lambda} \phi(I_n^c) \leq \epsilon_n$ for all $n \in \mathbb{Z}_+$. Let $I_n = [x_n, y_n]$. Since (I_n) is

increasing, the sequence (x_n) is decreasing and (y_n) is increasing. For the distributions ℓ and u , we set

$$\ell := \sum_{n=0}^{\infty} \alpha_n \delta_{x_n} \quad \text{and} \quad u := \sum_{n=0}^{\infty} \alpha_n \delta_{y_n}, \quad \text{where} \quad \alpha_n := \epsilon_{n-1} - \epsilon_n.$$

In the definition above, $\epsilon_{-1} := 1$ and δ_z is the probability measure concentrated on z . One easily confirms that both ℓ and u are elements of $\mathcal{P}(S)$.

Now let J be any increasing set in \mathcal{B} . We claim that $\phi(J) \leq u(J)$ for all $\phi \in \Lambda$. To see that this is so, fix $\phi \in \Lambda$ and suppose first that $y_m \in J^c$ for all m . If this is so, then $\phi(J) = 0$. Indeed, fixing $m \in \mathbb{Z}_+$, the property $y_m \in J^c$ implies $J \subset I_m^c$. (Otherwise there is an $x \in J \cap I_m$ and, since J is increasing and $x \leq y_m$, we have $y_m \in J$.) This means that $\phi(J) \leq \phi(I_m^c) \leq \epsilon_m$. Taking the limit in m yields $\phi(J) = 0$. Hence $\phi(J) \leq u(J)$.

Now consider the other case, where there exists an integer m satisfying

$$m = \min\{n \in \mathbb{Z}_+ : y_n \in J\}.$$

Observe that $J \cap I_{m-1}$ is empty. (If x is in both, then, since $x \leq y_{m-1}$ and $x \in J$, we have $y_{m-1} \in J$, which contradicts the definition of m .) As a result,

$$\phi(J) = \phi(J \cap I_{m-1}) + \phi(J \cap I_{m-1}^c) = \phi(J \cap I_{m-1}^c) \leq \phi(I_{m-1}^c) \leq \epsilon_{m-1}.$$

At the same time, since (y_n) is increasing,

$$u(J) = \epsilon_{m-1} - \epsilon_m + \epsilon_m - \epsilon_{m+1} + \epsilon_{m+1} - \epsilon_{m+2} + \cdots = \epsilon_{m-1}.$$

As a consequence, we have $\phi(J) \leq u(J)$. This shows that $\phi \leq_{\text{sd}} u$. A similar argument shows that $\ell \leq_{\text{sd}} \phi$ also holds. Hence Λ is order bounded.

Conversely, suppose Λ is order bounded, so there exist $\ell, u \in \mathcal{P}(S)$ with $\ell \leq_{\text{sd}} \phi \leq_{\text{sd}} u$ for all $\phi \in \Lambda$. Fix $\epsilon > 0$. Since ℓ and u are distributions on a Polish space, they are tight, so we can choose compact sets K_ℓ and K_u with $\ell(K_\ell^c) \leq \epsilon/2$ and $u(K_u^c) \leq \epsilon/2$. By Assumption 3.1, the compact set $K_\ell \cup K_u$ is order bounded, so there exist $a, b \in S$ with $K_\ell \cup K_u \subset [a, b]$. By Assumption 3.1 again, $[a, b]$ is compact.

Let $U_a := \{x \in S : x \geq a\}$ and $D_b := \{x \in S : x \leq b\}$. Then $[a, b] = U_a \cap D_b$ and $[a, b]^c = U_a^c \cup D_b^c$. Fix $\phi \in \Lambda$. Since U_a is increasing, we have $\ell(U_a) \leq \phi(U_a)$, giving $\phi(U_a^c) \leq \ell(U_a^c)$. Similarly, since D_b^c is increasing and $\phi \leq_{\text{sd}} u$, we have $\phi(D_b^c) \leq u(D_b^c)$. Now $K_\ell \subset [a, b]$ implies $U_a^c \subset [a, b]^c \subset K_\ell^c$, so $\ell(U_a^c) \leq \epsilon/2$. Likewise, $u(D_b^c) \leq \epsilon/2$. Therefore,

$$\phi([a, b]^c) \leq \phi(U_a^c) + \phi(D_b^c) \leq \ell(U_a^c) + u(D_b^c) \leq \epsilon.$$

Since $[a, b]$ is compact and this holds for all $\phi \in \Lambda$, we conclude that Λ is tight. \square

Proof of Theorem 3.1. Let the conditions of Theorem 3.1 hold.

((ii) \Leftrightarrow (iii)) This equivalence is immediate from Proposition 3.1.

((i) \implies (ii)) If (T_t) is globally stable, then (T_t) is asymptotically contractive by Theorem 2.1. Moreover, when ψ^* is the unique stationary distribution, the constant trajectory $(T_t\psi^*) = (\psi^*)$ is trivially order bounded.

((ii) \implies (i)) Suppose that (T_t) is asymptotically contractive and has an order-bounded trajectory. The space $(\mathcal{P}(S), \beta)$ is complete and β satisfies the diagonal property by Lemma 3.2. Moreover, \leq_{sd} is closed with respect to β by Lemma 3.3. Hence (T_t) is globally stable by Theorem 2.1. \square

A.4. Proof of Theorem 5.1. Let the conditions of Theorem 5.1 hold. We write $[\cdot]$ for the floor function.

Lemma A.5. *If the conditions of Theorem 5.1 hold, then*

$$(1 - \epsilon)\delta_a + \epsilon\delta_{\bar{x}} \leq_{sd} P_u(a, \cdot) \quad \text{and} \quad P_u(b, \cdot) \leq_{sd} (1 - \epsilon)\delta_b + \epsilon\delta_{\bar{x}} \quad (\text{A.7})$$

Proof. See the proof of Eq. (2) in Theorem 2 of Hopenhayn and Prescott (1992). \square

Lemma A.6. *If the conditions of Theorem 5.1 hold, then*

$$\beta(\phi P_u^n, \psi P_u^n) \leq 2(1 - \epsilon)^n \quad \text{for all } \phi, \psi \in \mathcal{P}(S) \text{ and } n \in \mathbb{Z}_+. \quad (\text{A.8})$$

Proof. We set $P = P_u$ for brevity. Fix $h \in ibS$ with $|h| \leq 1$. Define $d_n = P^n h(a)$ and $e_n = P^n h(b)$. Since P is increasing and h is increasing, $P^{n-1}h$ is also increasing for all $n \geq 1$. Applying the bounds from Lemma A.5 to the increasing function $P^{n-1}h$ yields

$$\begin{aligned} d_n &= (P(P^{n-1}h))(a) \geq (1 - \epsilon)d_{n-1} + \epsilon(P^{n-1}h)(\bar{x}) \\ e_n &= (P(P^{n-1}h))(b) \leq (1 - \epsilon)e_{n-1} + \epsilon(P^{n-1}h)(\bar{x}) \end{aligned}$$

Subtracting one inequality from the other, we get $e_n - d_n \leq (1 - \epsilon)(e_{n-1} - d_{n-1})$, and hence

$$e_n - d_n \leq (1 - \epsilon)^n(e_0 - d_0) = (1 - \epsilon)^n(h(b) - h(a)) \leq 2(1 - \epsilon)^n. \quad (\text{A.9})$$

Since a is the bottom and b is the top of S , we have $(P^n h)(a) \leq (P^n h)(x) \leq (P^n h)(b)$ for all $x \in S$, so, integrating with respect to $\phi \in \mathcal{P}(S)$,

$$(P^n h)(a) \leq \phi(P^n h) \leq (P^n h)(b).$$

Therefore, for any two probability measures ϕ and ψ , we have

$$|\phi(P^n h) - \psi(P^n h)| \leq (P^n h)(b) - (P^n h)(a) \leq 2(1 - \epsilon)^n.$$

Taking the supremum over all $h \in ibS$ with $|h| \leq 1$ proves the claim in the lemma. \square

Proof of Theorem 5.1. Let the stated conditions hold. We first claim that $(P_t)_{t \in \mathbb{T}}$ is asymptotically contractive. For $\mathbb{T} = [0, \infty)$, fix $t \in (0, \infty)$, write $t = nu + r$ where $n = \lfloor t/u \rfloor$ and $0 \leq r < u$. By the semigroup property, we have $P_t = P_{nu+r} = P_{nu}P_r = P_u^n P_r$, so

$$\beta(\phi P_t, \psi P_t) = \beta(\phi P_u^n P_r, \psi P_u^n P_r) \leq \beta(\phi P_u^n, \psi P_u^n),$$

where the inequality is by the nonexpansiveness in Lemma 4.1. Applying the contraction bound in Lemma A.6, we obtain

$$\beta(\phi P_u^n, \psi P_u^n) \leq 2(1 - \epsilon)^n = 2(1 - \epsilon)^{\lfloor t/u \rfloor}.$$

As $t \rightarrow \infty$, $\lfloor t/u \rfloor \rightarrow \infty$, so $(P_t)_{t \geq 0}$ is asymptotically contractive, as claimed.

For $\mathbb{T} = \mathbb{Z}_+$, the proof is similar. We write $t = nu + r$ where $n = \lfloor t/u \rfloor$ and r is an integer obeying $0 \leq r < u$. Since $P_t = P_u^n P_r$, we can again use Lemmas 4.1 and A.6 to obtain

$$\beta(\phi P_t, \psi P_t) \leq \beta(\phi P_u^n, \psi P_u^n) \leq 2(1 - \epsilon)^n \rightarrow 0$$

as $t \rightarrow \infty$. Hence $(P_t)_{t \in \mathbb{Z}_+}$ is asymptotically contractive.

To complete the proof of global stability, we use Theorem 3.1. For any increasing $h \in ibS$ and any $\phi \in \mathcal{P}(S)$, we have $h(a) \leq \phi(h) \leq h(b)$, and hence $\delta_a \leq_{\text{sd}} \phi \leq_{\text{sd}} \delta_b$. In particular, for any $\phi \in \mathcal{P}(S)$ and any $t \in \mathbb{T}$, we have $\delta_a \leq_{\text{sd}} \phi P_t \leq_{\text{sd}} \delta_b$. Thus every trajectory of (P_t) is order bounded. Since (P_t) is asymptotically contractive, increasing, and has an order-bounded trajectory, Theorem 3.1 implies that (P_t) is globally stable.

For the exponential bound (5.2), let ϕ^* be the unique fixed point. Since $\lfloor t/u \rfloor \geq t/u - 1$, we have $(1 - \epsilon)^{\lfloor t/u \rfloor} \leq (1 - \epsilon)^{t/u - 1}$. Writing $(1 - \epsilon)^{t/u} = e^{-\alpha t}$ where $\alpha = \ln(1/(1 - \epsilon))/u > 0$, we obtain

$$\beta(\phi P_t, \psi P_t) \leq \frac{2}{1 - \epsilon} e^{-\alpha t}.$$

Setting $\psi = \phi^*$ and using $\phi^* P_t = \phi^*$ yields the bound in (5.2).

The monotone ergodicity result in Theorem 5.1 follows from Theorem 3.1 and Proposition 4.1 of Kamihigashi and Stachurski (2016). \square

A.5. Order-Theoretic Mixing and Asymptotic Contractivity. In this section we explore the connection between order mixing properties and asymptotic contractivity in the Bhattacharya metric. Here is one key result.

Proposition A.1. *Let (P_t) be an increasing transition probability function on S . If (P_t) is weakly order mixing, then (P_t) is asymptotically contractive on $\mathcal{P}(S)$ with respect to the metric β .*

Proof of Proposition A.1. Fix $x, x' \in S$ and $h \in ibS$ with $|h| \leq 1$. Let $((X_t), (X'_t))$ be as in the definition of weakly order mixing and τ be as defined in (7.1). We have

$$\delta_x P_t h - \delta_{x'} P_t h = \mathbb{E}_x h(X_t) - \mathbb{E}_{x'} h(X'_t) = \mathbb{E}_x \mathbb{E}_\tau h(X_t) - \mathbb{E}_{x'} \mathbb{E}_\tau h(X'_t)$$

By construction, (X_t) and (X'_t) are defined on the same probability space, so we can rewrite this as

$$\delta_x P_t h - \delta_{x'} P_t h = \mathbb{E}_{x, x'} [\mathbb{E}_\tau h(X_t) - \mathbb{E}_\tau h(X'_t)] = A + B,$$

where

$$\begin{aligned} A &:= \mathbb{E}_{x, x'} [\mathbb{E}_\tau h(X_t) - \mathbb{E}_\tau h(X'_t)] \{\tau \leq t\} \\ B &:= \mathbb{E}_{x, x'} [\mathbb{E}_\tau h(X_t) - \mathbb{E}_\tau h(X'_t)] \{\tau > t\}. \end{aligned}$$

Using the strong Markov property and the fact that τ is a finite stopping time, we can write A as

$$A = \mathbb{E}_{x, x'} [(P_{t-\tau} h)(X_\tau) - (P_{t-\tau} h)(X'_\tau)] \{\tau \leq t\}$$

Since h is increasing, $X_\tau \leq X'_\tau$ and (P_t) is increasing, we must have $A \leq 0$. Moreover,

$$B \leq \mathbb{E}_{x, x'} |\mathbb{E}_\tau h(X_t) - \mathbb{E}_\tau h(X'_t)| \{\tau > t\} \leq 2\mathbb{P}_{x, x'} \{\tau > t\}.$$

As a result,

$$\delta_x P_t h - \delta_{x'} P_t h \leq 2\mathbb{P}_{x, x'} \{\tau > t\}.$$

Since the one-sided bound holds for all $x, x' \in S$, applying it to the pair (x', x) gives $\delta_{x'} P_t h - \delta_x P_t h \leq 2\mathbb{P}_{x', x} \{\tau > t\}$, and hence

$$|\delta_x P_t h - \delta_{x'} P_t h| \leq 2 \max(\mathbb{P}_{x, x'} \{\tau > t\}, \mathbb{P}_{x', x} \{\tau > t\}).$$

Since $h \in ibS$ with $|h| \leq 1$ was arbitrary, we can take the supremum to get

$$\beta(\delta_x P_t, \delta_{x'} P_t) \leq 2 \max(\mathbb{P}_{x, x'} \{\tau > t\}, \mathbb{P}_{x', x} \{\tau > t\}) \quad \text{for all } x, x' \in S. \quad (\text{A.10})$$

We wish to extend this to arbitrary (nondegenerate) initial conditions. To this end, fix $\phi, \psi \in \mathcal{P}(S)$. Let (X, X') be any coupling with marginals ϕ and ψ . For $h \in ibS$ with $|h| \leq 1$, we have

$$(\phi P_t)(h) - (\psi P_t)(h) = \mathbb{E}[(P_t h)(X)] - \mathbb{E}[(P_t h)(X')] = \mathbb{E}[(\delta_X P_t)(h) - (\delta_{X'} P_t)(h)].$$

Taking absolute values and using (A.10), we get

$$|(\phi P_t)(h) - (\psi P_t)(h)| \leq \mathbb{E}[\beta(\delta_X P_t, \delta_{X'} P_t)].$$

The right-hand side does not depend on h , so taking the supremum over $h \in ibS$ with $|h| \leq 1$ gives

$$\beta(\phi P_t, \psi P_t) \leq \mathbb{E}[\beta(\delta_X P_t, \delta_{X'} P_t)].$$

By (A.10),

$$\beta(\phi P_t, \psi P_t) \leq 2\mathbb{E}[\max(\mathbb{P}_{X, X'}\{\tau > t\}, \mathbb{P}_{X', X}\{\tau > t\})]. \quad (\text{A.11})$$

Since $\mathbb{P}_{x, x'}\{\tau < \infty\} = 1$ by (7.1) for all $x, x' \in S$, both $\mathbb{P}_{x, x'}\{\tau > t\}$ and $\mathbb{P}_{x', x}\{\tau > t\}$ converge to 0 as $t \rightarrow \infty$ for each (x, x') . By the dominated convergence theorem, the right-hand side of (A.11) converges to 0. Since $\phi, \psi \in \mathcal{P}(S)$ were arbitrary, (P_t) is asymptotically contractive. \square

A.6. Proof of Theorem 6.1 and Theorem 7.1. First we prove the discrete time result in Theorem 6.1.

Proof of Theorem 6.1. Let P be any stochastic kernel on S . If P is globally stable on $\mathcal{P}(S)$, then P is bounded in probability, since the sequence of distributions $t \mapsto \delta_x P^t = P^t(x, \cdot)$ is convergent and hence tight. (By Prokhorov's theorem, every weakly convergent sequence of probability measures on a Polish space is tight.)

Now suppose that P is increasing, order reversing, and bounded in probability. By Lemma A.5 from Kamihigashi and Stachurski (2014), there exists an $(S \times S)$ -valued Markov process $((X_t), (X'_t))_{t \in \mathbb{Z}_+}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $(X_t)_{t \in \mathbb{Z}_+}$ and $(X'_t)_{t \in \mathbb{Z}_+}$ are independent, both of these processes are P -Markov, and the stopping time $\tau := \inf\{t \geq 0 : X_t \leq X'_t\}$ obeys $\mathbb{P}_{x, x'}\{\tau < \infty\} = 1$ for all $x, x' \in S$. Since \leq is closed, $\{(x, x') \in S \times S : x \leq x'\}$ is measurable, so τ is a stopping time. The strong Markov property always holds in discrete time (see, e.g., Meyn and Tweedie (2009)). As a result, the associated discrete time semigroup is weakly order mixing. Proposition A.1 now implies that $(P_t)_{t \in \mathbb{Z}_+}$ is asymptotically contractive on $\mathcal{P}(S)$ with respect to the metric β . By definition, this means that the operator P is asymptotically contractive on $\mathcal{P}(S)$. Global stability now follows from Corollary 3.1 (iii), given that boundedness in probability implies existence of a tight trajectory. \square

Now we prove the continuous time result in Theorem 7.1.

Proof of Theorem 7.1. Since (P_t) is increasing and weakly order mixing, Proposition A.1 implies that (P_t) is asymptotically contractive on $(\mathcal{P}(S), \beta)$. If (P_t) is globally stable, then the trajectory $t \mapsto \phi P_t$ is convergent for every $\phi \in \mathcal{P}(S)$, and hence tight (since every

weakly convergent sequence of probability measures on a Polish space is tight). Conversely, if (P_t) has at least one tight trajectory, then global stability follows from Theorem 3.1, since (P_t) is order-preserving and asymptotically contractive. \square

A.7. Proof of Theorem 8.1. We prove this theorem using a sequence of lemmas.

Lemma A.7. *If Assumption 8.1 holds, then (P_t) is increasing.*

Proof. Fix $x, x' \in S$ with $x \leq x'$. On a common probability space, construct two PDMPs (X_t) and (X'_t) starting from x and x' respectively, using the same jump times (T_n) and the same shocks (ξ_n) . The embedded chains satisfy $Z_0 = x \leq x' = Z'_0$. By induction: if $Z_n \leq Z'_n$, then $\Phi(Z_n, E_{n+1}) \leq \Phi(Z'_n, E_{n+1})$ by Assumption 8.1 (i) and

$$Z_{n+1} = F(\Phi(Z_n, E_{n+1}), \xi_{n+1}) \leq F(\Phi(Z'_n, E_{n+1}), \xi_{n+1}) = Z'_{n+1}$$

by Assumption 8.1 (ii). Hence $Z_n \leq Z'_n$ for all n . For any $t \geq 0$, applying Assumption 8.1 (i) gives

$$X_t = \Phi(Z_{N_t}, t - T_{N_t}) \leq \Phi(Z'_{N_t}, t - T_{N_t}) = X'_t.$$

Therefore, for any $h \in \text{ibS}$, $P_t h(x) = \mathbb{E}_x h(X_t) \leq \mathbb{E}_{x'} h(X'_t) = P_t h(x')$. \square

Lemma A.8. *If Assumptions 8.1 and 8.2 hold, then (P_t) is bounded in probability.*

The idea behind the proof is that the pathwise bounds in Assumption 8.2 prevent the post-jump states from escaping to infinity, while the inherently exponential inter-arrival times prevent the deterministic flow from carrying the process too far between jumps.

Proof. Fix $x_0 \in S$ and $\varepsilon > 0$. We need to find a compact set D such that $\mathbb{P}_{x_0}\{X_t \in D\} \geq 1 - \varepsilon$ for all $t \geq 0$. By Assumption 8.2, $f_1(\xi_n) \leq Z_n \leq f_2(\xi_n)$ for all $n \geq 1$, so each post-jump state lies in $[f_1(\xi_n), f_2(\xi_n)]$. Choose a compact $K \subset S$ with $\mathbb{P}\{f_i(\xi) \in K\} \geq 1 - \varepsilon/6$ for $i = 1, 2$. By Assumption 3.1, $K \subset [a, b]$ for some $a, b \in S$, and $C := [a, b]$ is compact. Whenever both $f_1(\xi_n)$ and $f_2(\xi_n)$ fall in K , the post-jump state Z_n lies in $K \subset C$.

Next, we show that Z_{N_t} , the state at the most recent jump before time t , lies in C with high probability. Formally, ξ_n is independent of (E_1, \dots, E_{n+1}) and $\{N_t = n\}$ is determined by (E_1, \dots, E_{n+1}) , so

$$\mathbb{P}_{x_0}\{Z_{N_t} \notin C, N_t \geq 1\} = \sum_{n=1}^{\infty} \mathbb{P}\{N_t = n, Z_n \notin C\} \leq \frac{\varepsilon}{3} \sum_{n=1}^{\infty} \mathbb{P}\{N_t = n\} = \frac{\varepsilon}{3}.$$

Between consecutive jumps, the process follows the flow Φ , so $X_t = \Phi(Z_{N_t}, R_t)$ where $R_t := t - T_{N_t}$ is the time elapsed since the last jump. The event $\{R_t > M\}$ requires no

Poisson arrivals in the interval $(t - M, t]$, so $\mathbb{P}\{R_t > M\} \leq e^{-\lambda M}$. Choose $M > 0$ with $e^{-\lambda M} < \varepsilon/3$. Before the first jump ($N_t = 0$), the process follows its initial trajectory: $X_t = \Phi(x_0, t)$. The probability of no jump by time t is $\mathbb{P}\{N_t = 0\} = e^{-\lambda t}$, which is small for large t . Choose $T > 0$ with $e^{-\lambda T} < \varepsilon/3$.

Since Φ is continuous,

$$D := \Phi(C \times [0, M]) \cup \Phi(\{x_0\} \times [0, T])$$

is compact. (The first component captures the flow starting from any post-jump state in C and running for at most M time units. The second captures the initial trajectory from x_0 up to time T , covering the possibility that no jump has yet occurred.) Consider first the case $t \leq T$: if $N_t = 0$, then $X_t = \Phi(x_0, t) \in D$. For any such t : on the event $\{N_t \geq 1, Z_{N_t} \in C, R_t \leq M\}$, we have $X_t = \Phi(Z_{N_t}, R_t) \in D$. The three complementary events each have small probability:

$$\mathbb{P}_{x_0}\{X_t \notin D\} \leq \mathbb{P}\{N_t = 0\} + \mathbb{P}\{Z_{N_t} \notin C, N_t \geq 1\} + \mathbb{P}\{R_t > M\} \leq e^{-\lambda t} + \frac{\varepsilon}{3} + e^{-\lambda M}.$$

For $t > T$ each term is at most $\varepsilon/3$, giving a total bound of ε . For $t \leq T$ the first term covers the event $\{N_t = 0\}$ (on which $X_t \in D$ by construction), so $\mathbb{P}_{x_0}\{X_t \notin D\} \leq \varepsilon/3 + \varepsilon/3 < \varepsilon$. \square

Lemma A.9. *If Assumptions 8.1 and 8.2 hold, then (P_t) is weakly order mixing.*

Proof. Construct two PDMPs (X_t) and (X'_t) on a common probability space using the same Poisson jump times (T_n) but independent shock sequences (ξ_n) and (ξ'_n) , each IID with common distribution μ . Each marginal process is (P_t) -Markov, and the joint process $((X_t), (X'_t))$ is itself a PDMP on $S \times S$ (with the same jump rate λ , joint flow $(\Phi(x, t), \Phi(x', t))$, and independent shocks), hence strong Markov.

Fix $x, x' \in S$ and write $A_n = \{Z_n \leq Z'_n\}$ for the event that the two embedded chains become ordered at the n -th shared jump. On A_n , the semi-flow identity gives $X_{T_n} = \Phi(Z_n, 0) = Z_n \leq Z'_n = \Phi(Z'_n, 0) = X'_{T_n}$. Hence $\tau \leq T_n$, where $\tau := \inf\{t \geq 0 : X_t \leq X'_t\}$, and it suffices to show that A_n occurs for some n with probability one.

Let f_1 and f_2 be as in Assumption 8.2 and set $\delta := \mathbb{P}\{f_2(\xi) \leq f_1(\xi')\} > 0$. On the event $\{f_2(\xi_n) \leq f_1(\xi'_n)\}$, we have $Z_n = F(\Phi(Z_{n-1}, E_n), \xi_n) \leq f_2(\xi_n) \leq f_1(\xi'_n) \leq F(\Phi(Z'_{n-1}, E_n), \xi'_n) = Z'_n$, so A_n occurs regardless of the pre-jump states. Since (ξ_n, ξ'_n) is independent of all other randomness, $\mathbb{P}\{A_n^c \mid \mathcal{F}_{n-1}\} \leq 1 - \delta$ for every n . Therefore

$$\mathbb{P}_{x, x'}\{\tau > T_k\} \leq \mathbb{P}\{A_1^c \cap \dots \cap A_k^c\} \leq (1 - \delta)^k.$$

Since $T_k \rightarrow \infty$ a.s., $\mathbb{P}_{x, x'}\{\tau < \infty\} = 1$. \square

Proof of Theorem 8.1. By Lemma A.7, (P_t) is increasing. By Lemma A.8, (P_t) is bounded in probability, which implies the existence of a tight trajectory. By Lemma A.9, (P_t) is weakly order mixing. Global stability now follows from Theorem 7.1. \square

REFERENCES

- Antonio Antunes and Tiago Cavalcanti. Start up costs, limited enforcement, and the hidden economy. *European Economic Review*, 51:203–224, 2007.
- Łukasz Balbus, Anna Jaśkiewicz, Andrzej S Nowak, and Łukasz Woźny. Markov perfect equilibria in stochastic growth models with quasi-hyperbolic discounting and risk-sensitive preferences. *Dynamic Games and Applications*, pages 1–19, 2025.
- Brendan K. Beare and Alexis Akira Toda. Determination of Pareto exponents in economic models driven by Markov multiplicative processes. *Econometrica*, 90(4):1811–1833, 2022.
- Kenneth Burdett and Dale T. Mortensen. Wage differentials, employer size, and unemployment. *International Economic Review*, 39(2):257–273, 1998.
- Partha Chatterjee and Malik Shukayev. A stochastic dynamic model of trade and growth: Convergence and diversification. mimeo, National University of Singapore, 2010.
- Timothy Cogley and Thomas J Sargent. Drifting risk-adjusted inflation targets. *Journal of Monetary Economics*, 55(5):740–757, 2008.
- Melvyn Coles and Dale T. Mortensen. Equilibrium labor turnover, firm growth, and unemployment. *Econometrica*, 84(1):347–363, 2016.
- Stewart N Ethier and Thomas G Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 2009.
- Sergey Foss and Michael Scheutzow. Compressibility and stochastic stability of monotone markov chain, 2024. URL <https://arxiv.org/abs/2403.15259>.
- Sergey Foss, Vsevolod Shneer, Jonathan P. Thomas, and Tim Worrall. Stochastic stability of monotone economies in regenerative environments. *Journal of Economic Theory*, 173:334–360, 2018. doi: 10.1016/j.jet.2017.12.007.
- Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. The dynamics of inequality. *Econometrica*, 84(6):2071–2111, 2016.
- Jinyong Hahn, Guido Kuersteiner, and Maurizio Mazzocco. Central limit theory for combined cross section and time series with an application to aggregate productivity shocks. *Econometric Theory*, 40(1):162–212, 2024.
- Ana Hidalgo-Cabrillana. Endogenous capital market imperfections, human capital, and intergenerational mobility. *Journal of Development Economics*, 90:285–298, 2009.

- Hugo A Hopenhayn and Edward C Prescott. Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica*, pages 1387–1406, 1992.
- Timothy Kam, Tina Kao, and Junsang Lee. Inflation, inequality, and welfare in a competitive search model. *Macroeconomic Dynamics*, 29:e110, 2025.
- T. Kamae and U. Krengel. Stochastic partial ordering. *The Annals of Probability*, 6(6): 1044–1049, 1978. doi: 10.1214/aop/1176995392.
- Takashi Kamihigashi and John Stachurski. An order-theoretic mixing condition for monotone Markov chains. *Statistics and Probability Letters*, 82:262–267, 2012.
- Takashi Kamihigashi and John Stachurski. Stochastic stability in monotone economies. *Theoretical Economics*, 9(2):383–407, 2014. doi: 10.3982/TE1367.
- Takashi Kamihigashi and John Stachurski. Seeking ergodicity in dynamic economies. *Journal of Economic Theory*, 163:900–924, 2016. doi: 10.1016/j.jet.2016.03.006.
- Takashi Kamihigashi and John Stachurski. A unified stability theory for classical and monotone Markov chains. *Journal of Applied Probability*, 56(1):1–22, 2019.
- Andrzej Lasota and Michael C. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, volume 97 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2nd edition, 1994.
- François Le Grand and Xavier Ragot. Managing inequality over business cycles: Optimal policies with heterogeneous agents and aggregate shocks. *International Economic Review*, 63(1):511–540, 2022.
- Bar Light. A note on the stability of monotone Markov chains. *Operations Research Letters*, 57, 2024.
- Bar Light. Invariant distributions in nonlinear Markov chains with aggregators: Theory, computation, and applications. *Operations Research*, in press, 2026.
- Bar Light and Gabriel Y Weintraub. Mean field equilibrium: Uniqueness, existence, and comparative statics. *Operations Research*, 70(1):585–605, 2022.
- Torgny Lindvall. *Lectures on the coupling method*. Courier Corporation, 2002.
- Qingyin Ma and Alexis Akira Toda. Optimal savings with preference for wealth. *arXiv preprint arXiv:2509.12195*, 2025.
- Albert Marcet, Francesc Obiols-Homs, and Philippe Weil. Incomplete markets, labor supply and capital accumulation. *Journal of Monetary Economics*, 54:2621–2635, 2007.
- Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, 2009.
- Olivier F. Morand and Kevin L. Reffett. Stationary Markovian equilibrium in overlapping generation models with stochastic nonclassical production and Markov shocks. *Journal of Mathematical Economics*, 43:501–522, 2007.

- Giuseppe Moscarini and Fabien Postel-Vinay. Stochastic search equilibrium. *Review of Economic Studies*, 80(4):1545–1581, 2013.
- Anna Orlik and Laura Veldkamp. Understanding uncertainty shocks and the role of black swans. *NBER Working Paper No. 20445*, 2014.
- Assaf Razin and Joseph A. Yahav. On stochastic models of economic growth. *International Economic Review*, 20:599–604, 1979.
- Ryszard Rudnicki and Marta Tyran-Kamińska. *Piecewise deterministic processes in biological models*, volume 1. Springer, 2017.
- Roberto M. Samaniego. Can technical change exacerbate the effects of labor market sclerosis? *Journal of Economic Dynamics and Control*, 32:497–528, 2008.
- Kyle Siegrist. Probability, mathematical statistics, and stochastic processes, 2022. Section 16.20: Chains Subordinate to the Poisson Process.
- Nancy L Stokey and Robert E Lucas. *Recursive methods in dynamic economics*. Harvard University Press, 1989.
- Jacek Suda. Belief shocks and the macroeconomy. Banque de France Working Paper, 2010.
- Adriaan C Zaanen. *Introduction to operator theory in Riesz spaces*. Springer, 2012.