RECOVERY RATES IN INVESTMENT-GRADE POOLS OF CREDIT ASSETS: A LARGE DEVIATIONS ANALYSIS

KONSTANTINOS SPILIOPOULOS AND RICHARD B. SOWERS

Abstract. We consider the effect of recovery rates on a pool of credit assets. We allow the recovery rate to depend on the defaults in a general way. Using the theory of large deviations, we study the structure of losses in a pool consisting of a continuum of types. We derive the corresponding rate function and show that it has a natural interpretation as the favored way to rearrange recoveries and losses among the different types. Numerical examples are also provided.

1. Introduction

Understanding the behavior of *large pools of credit assets* is currently a problem of central importance. Banks often hold such large pools and their risk-reward characteristics need to be carefully managed. In many cases, the losses in the pool are (hopefully) *rare* as a consequence of diversification. In a number of papers [Sowa, Sowb], we have used the theory of large deviations to gain some insight into several aspects of rare losses in pools of credit assets. Our interest here is the effect of *recovery*. While a creditor either defaults or doesn't (a Bernoulli random variable), the amount recovered may in fact take a continuum of values. Although many models assume that recovery rate is constant—i.e., a fixed deterministic percentage of the par value, in reality the statistics of the amount recovered should be a bit more complicated. The statistics of the recovered amount should depend on the number of defaults; a large number of defaults corresponds to a bear market, in which case it is more difficult to liquidate the assets of the creditors. Our goal is to understand how to include this effect in the study of rare events in large pools. We would like to look at these rare events via some ideas from statistical mechanics, or more accurately the theory of *large deviations*. Large deviations formalizes the idea that nature prefers "minimum energy" configurations when rare events occur. We would like to see how these ideas can be used in studying the interplay between default rate and recovery rate.

Our work is motivated by the general challenge of understanding the effects of nonlinear interactions between various parts of complicated financial systems. One of the strengths of the theory of large deviations is exactly that it allows one to focus on propagation of rare events in networks. Our interest here is to see how this can be implemented in a model for recovery rates which depend on the default rate.

This work is part of a growing body of literature which applies the theory of large deviations to problems of rare losses in credit assets; cf. [Pha07] and [DDD04]. Our work is most closely related to [LMS09], where the dynamics of a configuration of defaults was studied. There, the recovery rate was assumed to be independent of the defaults. Our work here is explicitly interested in *dependence*. Furthermore, the framework of our efforts is a continuum of "types", whereas [LMS09] was focussed primarily on a model of one type, with a brief treatment of finite types. The case of a continuum of types requires slightly more topological sophistication. We also note that issues of recovery have been considered in the economics literature; see [SH09] and the references therein.

The paper is organized as follows. In Section 2 we introduce our model and establish some notation. In Section 3 we study the "typical" behavior of the loss of our pool; we need to understand this before we can identify what behavior is "atypical". In Section 4, we present some formal sample calculations and examples. These calculations are indicative of the range of possibilities and they illustrate the main results. Furthermore, we state our main result, Theorem 4.5, and present numerical several examples which illustrate some of the main aspects of our analysis. In particular, we use the structure of our work to understand some

Date: June 15, 2010.
R.S. would like to thanks the Departments of Mathematics and Statistics of Stanford University for their hospitality in the Spring of 2010 during a sabbatical stay.

aspects of implied recovery. The proof of Theorem 4.5 is in Sections 5, 6 and 7. Section 8 contains an alternative expression of the rate function, which is a variational formula which optimizes over all possible configuration of recoveries and defaults, and which leads to a Lagrange multiplier approach which can be numerically implemented. Lastly, Section 9 contains the proofs of several necessary technical results.

The model at the heart of our analysis is in fact very stylized. Since our primary interest is the interaction between default rates and recovery rates, our model focusses on this effect, but simplifies a number of other effects. In particular, we assume that the defaults themselves are independent. Notwithstanding this simplification, our model does capture some aspects of empirically observed relations between default and implied recovery (see subsection 4.1). Hopefully, more realistic models (e.g., which include a systemic source of risk) can be analyzed by techniques which are extensions of those developed here.

2. The Model

Let's start by considering a single bond (or "name"). For reference, let's assume that all bonds have par value of \$1. If the bond defaults, the assets underlying the bond are auctioned off and the bondholder recovers r dollars, where $r \in [0,1]$. We will record the default/survival coordinate as an element of $\{0,1\}$, where 1 corresponds to a default and 0 to survival. In case of default, the amount recovered is an element of $[0,1]$. Thus the minimal state space for a single bond is the set $E_\circ \overset{\text{def}}{=} \{0\} \cup (\{1\} \times [0,1])$. Since we want to consider a pool of bonds, the state space in our model will be $E \overset{\text{def}}{=} E_\circ^{\mathbb{N}}$ (as usual, $\mathbb{N} \overset{\text{def}}{=} \{1,2\dots\}$). Giving $\{0,1\}$ the discrete topology and $[0,1]$ its natural topology, we have that $\{0,1\} \times [0,1]$ is of course Polish; thus $E_\circ$, as a subset of $\{0,1\} \times [0,1]$ is also Polish, and thus E is also Polish. We endow E with the natural σ-algebra $\mathscr{F} \overset{\text{def}}{=} \mathscr{B}(\Omega)$.

Let's now define the principal objects of interest associated with *loss* in the pool. The pool suffers a loss when a bond defaults, and the amount of the loss is $\$1 - r$, where r is the recovery amount (in dollars). For $e \in E_\circ$, define

$$\check{\Delta}(e) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } e = (1,r) \\ 0 & \text{if } e = 0 \end{cases} \qquad \text{and} \qquad \check{\ell}(e) \overset{\text{def}}{=} \begin{cases} 1-r & \text{if } e = (1,r) \\ 0 & \text{if } e = 0; \end{cases}$$

then $\check{\Delta}$ is a Bernoulli random variable which records whether the bond has defaulted, and $\check{\ell}$ records the loss. For each $n \in \mathbb{N}$ and $\omega = (\omega_1, \omega_2 \dots)$, define the random variables $\Delta_n(\omega) \overset{\text{def}}{=} \check{\Delta}(\omega_n)$ and $\ell_n(\omega) \overset{\text{def}}{=} \check{\ell}_n(\omega_n)$. The default and loss rates in the pool are then

$$D_N \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N \Delta_n \qquad \text{and} \qquad L_N \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N \ell_n.$$

The only remaining thing to specify is a probability measure on $(E, \mathscr{F})$. For each $N \in \mathbb{N}$, fix $\{p^{N,n} : n \in \{1,2\dots N\}\} \subset [0,1]$. These will be the risk-neutral default probabilities of the names when the pool has N names. We next fix $\{\wp^{N,n} : n \in \{1,2\dots N\}\} \subset C([0,1]; \mathscr{P}[0,1])$; i.e., a collection of probability measures on $[0,1]$ ($[0,1]$ being the range of the recovery[1]) indexed by elements of $[0,1]$ (being the default rate). For each $n \in \mathbb{N}$ and $\omega = (\omega_1, \omega_2 \dots)$, define the coordinate random variable $X_n(\omega) = \omega_n$. For each $N \in \mathbb{N}$, we then fix our risk-neutral probability measure $\mathbb{P}_N \in \mathscr{P}(\Omega)$ (with associated expectation operator $\mathbb{E}_N$) by requiring that

$$\mathbb{E}_N\left[\prod_{n=1}^N f_n(X_n)\right] = \mathbb{E}_N\left[\prod_{n=1}^N \left\{(1-p^{N,n})f_n(0) + p^{N,n}\int_{r\in[0,1]} f_n(1,r)\wp^{N,n}(D_N, dr)\right\}\right]$$

for all $\{f_n\}_{n=1}^N \subset B(E_\circ)$. For each $n \in \mathbb{N}$, this means that under $\mathbb{P}_N$,

- $\{\Delta_n\}_{n=1}^N$ is an independent collection of random variables with $\mathbb{P}_N\{\Delta_N = 1\} = p^{N,n}$.
- for any fixed $A \subset \{1,2\dots N\}$, conditional on the event that $\{n \in \{1,2\dots N\} : \Delta_n = 1\} = A$, $\{X_n\}_{n\in A}$ is an independent collection of random variables and X_n has (conditional) law $\delta_{\{1\}} \times \wp^{N,n}\left(\frac{|A|}{N}, \cdot\right)$.

[1]We shall write $\wp \in C([0,1]; \mathscr{P}[0,1])$ as a map from $[0,1] \times \mathscr{B}[0,1]$ into $[0,1]$ such that for each $D \in [0,1]$, $A \mapsto \wp(D,A)$ is an element of $\mathscr{P}[0,1]$ and such that $D \mapsto \wp(D,\cdot)$ is continuous.

With this probabilistic structure in place, we will clearly want to be able to condition on the default rate so that we can then focus on the recovery rates; define the σ-algebra $\mathscr{D} \overset{\text{def}}{=} \sigma\{\Delta_n : n \in \mathbb{N}\}$.

Our goal is as follows. We firstly want to understand what L_N looks like as $N \to \infty$. In fact, under some regularity assumptions on the $p^{N,n}$'s and $\wp^{N,n}$'s we will be able to identify $\bar{L} \overset{\text{def}}{=} \lim_{N\to\infty} L_N$ (this being a limit in probability). Our second goal is to compute the asymptotics of $\mathbb{P}_N\{L_N \ge l\}$ as $N \to \infty$, particularly for $l > \bar{L}$; then $\{L_n \ge l\}$ is an "atypical" event.

Remark 2.1. *In a sense, we are using the empirical default rate D_N as a "systemic" random variable; we have a sort of "contagion" from this systemic random variable to the recovery rate.*

We conclude this section with two illustrative examples that will guide our calculations in the following sections: a homogeneous pool of assets and a heterogeneous one with only two types.

Example 2.2 (Homogeneous Pool)**.** *Fix $p \in [0,1]$ and $\wp \in C([0,1];\mathscr{P}[0,1])$ and let $p^{N,n} = p$ and $\wp^{N,n} = \wp$ for all $N \in \mathbb{N}$ and $n \in \{1,2\dots N\}$.*

□

Example 2.3 (Heterogeneous Pool)**.** *Fix p_A and p_B in $[0,1]$ and fix $\wp_A$ and $\wp_B$ in $C([0,1];\mathscr{P}[0,1])$. Every third bond will be of type A and have default probability p_A and recovery distribution governed by $\wp_A$, and the remaining bonds will have default probability p_B and recovery distribution governed by $\wp_B$. In other words, $p^{N,n} = p_A$ and $\wp^{N,n} = \wp_A$ if $n \in 3\mathbb{N}$ and $p^{N,n} = p_B$ and $\wp^{N,n} = \wp_B$ if $n \in 3\mathbb{N} + \{1,2\}$. For future reference, let's separate the defaults and into the the two types. Define*

$$D_N^A \overset{\text{def}}{=} \frac{1}{\lfloor N/3 \rfloor} \sum_{\substack{1\le n\le N \\ n\in 3\mathbb{N}}} \Delta_n \qquad \textit{and} \qquad D_N^B \overset{\text{def}}{=} \frac{1}{N - \lfloor N/3 \rfloor} \sum_{\substack{1\le n\le N \\ n\notin 3\mathbb{N}}} \Delta_n.$$

Then

$$D_N = \frac{\lfloor N/3 \rfloor}{N} D_N^A + \frac{N - \lfloor N/3 \rfloor}{N} D_N^B \approx \frac{1}{3} D_N^A + \frac{2}{3} D_N^B.$$

□

3. Typical Events

Let's start our analysis by identifying the "typical" behavior of L_N as $N \to \infty$.

Example 3.1 (Homogeneous Example)**.** *Let's see what L_N looks like in Example 2.2. By the law of large numbers $\lim_{N\to\infty} D_N = p$. Thus in distribution*

$$L_N \approx \frac{1}{N} \sum_{1\le n\le pN} (1 - \xi_n)$$

where the ξ_n's are i.i.d. with distribution $\wp(p,\cdot)$. We should consequently have that

$$\lim_{N\to\infty} L_N = p \int_{r\in[0,1]} (1-r)\wp(p,dr).$$

□

Example 3.2 (Heterogeneous Example)**.** *We can also identify the limit of L_N in Example 2.3. The rate of default among the A bonds is p_A and the rate of default among the B bonds is p_B. Thus $\lim_{N\to\infty} D_N^A = p_A$ and $\lim_{N\to\infty} D_n^B = p_B$, so*

$$\lim_{N\to\infty} D_N = \bar{D} \overset{\text{def}}{=} \frac{p_A}{3} + \frac{2p_B}{3}.$$

Thus we should roughly have

$$L_N \approx \frac{1}{N} \left\{ \sum_{1\le n\le p_A N/3} (1 - \xi_n^A) + \sum_{1\le n\le 2p_B N/3} (1 - \xi_n^B) \right\}$$

where the ξ_n^A's have law $\wp_A(\bar{D},\cdot)$, the ξ_n^B's have distribution $\wp_B(\bar{D},\cdot)$ and the ξ_n^A's and ξ_n^B's are all independent. Consequently, it seems natural that

$$\lim_{N\to\infty} L_N = \frac{p_A}{3}\int_{r\in[0,1]}(1-r)\wp_A(\bar{D},dr) + \frac{2p_B}{3}\int_{r\in[0,1]}(1-r)\wp_B(\bar{D},dr),$$

the first term being the limit of the losses from the type A names, and the second term being the limit of the losses from the type B names.

□

In view of our examples, it seems reasonable that we should be able to describe the average loss in the pool in terms of a frequency count of $\mathsf{p}^{N,n} \stackrel{\text{def}}{=} (p^{N,n},\wp^{N,n})$. We note that $\mathsf{p}^{N,n}$ takes values in the set $\mathsf{X} \stackrel{\text{def}}{=} [0,1]\times C([0,1];\mathscr{P}[0,1])$. Since $\mathscr{P}[0,1]$ is Polish, so is $C([0,1];\mathscr{P}[0,1])$, and thus X is Polish. We will henceforth refer to elements of X as *types*. For each $N\in\mathbb{N}$, we now define $\mathsf{U}_N\in\mathscr{P}(\mathsf{X})$ as

$$\mathsf{U}_N \stackrel{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N \delta_{\mathsf{p}^{N,n}}.$$

A crucial component of our problem is that there is macroscopic "regularity" in type space.

Assumption 3.3. *We assume that $\mathsf{U} \stackrel{\text{def}}{=} \lim_{N\to\infty}\mathsf{U}_N$ exists (in $\mathscr{P}(\mathsf{X})$).*

□

Remark 3.4. *In the case of Example 2.2, we have that $\mathsf{U}=\delta_{(p,\wp)}$, while in the case of Example 2.3, we have that $\mathsf{U}=\frac{1}{3}\delta_{(p_A,\wp_A)}+\frac{2}{3}\delta_{(p_B,\wp_B)}$.*

We can now identify the limiting behavior of L_N. Define

$$\bar{D} \stackrel{\text{def}}{=} \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} p\,\mathsf{U}(d\mathsf{p}) \qquad\text{and}\qquad \bar{L} \stackrel{\text{def}}{=} \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\left\{p\int_{r\in[0,1]}(1-r)\wp(\bar{D},dr)\right\}\mathsf{U}(d\mathsf{p}). \tag{1}$$

To see that these quantities are well-defined, note that

$$(p,\wp)\mapsto p \qquad\text{and}\qquad \nu\mapsto\int_{r\in[0,1]}(1-r)\nu(dr) \tag{2}$$

are continuous mappings from, respectively, X and $\mathscr{P}[0,1]$, to $[0,1]\subset\mathbb{R}$ (the continuity of the second map follows from the obvious fact that $r\mapsto 1-r$ is a continuous map on $[0,1]$). The continuity of the first map of (2) implies that $\bar{D}$ is well-defined. Combining the continuity of both maps of (2), we get that the map

$$(p,\wp)\mapsto p\int_{r\in[0,1]}(1-r)\wp(\bar{D},dr)$$

is a continuous map from X to $[0,1]\subset\mathbb{R}$; thus $\bar{L}$ is also well-defined.

Lemma 3.5. *For each $\varepsilon>0$, we have that*

$$\lim_{N\to\infty}\mathbb{P}_N\left\{\left|L_N-\bar{L}\right|\ge\varepsilon\right\}=0.$$

□

Proof. For $\wp\in C([0,1];\mathscr{P}[0,1])$ and $D\in[0,1]$, let's first define

$$\Gamma(D,\wp)\stackrel{\text{def}}{=}\int_{r\in[0,1]}(1-r)\wp(D,dr).$$

Again we can use (2) and show, by the same arguments used to show that $\bar{D}$ and $\bar{L}$ of (1) are well-defined, that Γ is well-defined, and furthermore that it is continuous on $[0,1]\times C([0,1];\mathscr{P}[0,1])$. For each $N\in\mathbb{N}$, define

$$L_N^\circ \stackrel{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N\Gamma(D_N,\wp^{N,n})\Delta_n, \qquad \bar{D}_N \stackrel{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N p^{N,n} = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} p\,\mathsf{U}_N(d\mathsf{p})$$

$$\bar{L}_N \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N p^{N,n}\Gamma(\bar{D},\wp^{N,n}) = \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} p\Gamma(\bar{D},\wp)\mathsf{U}_N(d\mathsf{p});$$

note that L°_N is a random variable but $\bar{D}_N$ and $\bar{L}_N$ are deterministic. Note also that by weak convergence, $\lim_{N\to\infty}\bar{L}_N = \bar{L}$ and $\lim_{N\to\infty}\bar{D}_N = \bar{D}$. The claim will follow if we can prove that

(3) $$\lim_{N\to\infty}\mathbb{E}_N\left[|L_N - L^\circ_N|^2\right] = 0 \qquad \text{and} \qquad \lim_{N\to\infty}\mathbb{E}_N\left[|L^\circ_N - \bar{L}_N|^2\right] = 0.$$

To see the first part of (3), we calculate that

$$L_N - L^\circ_N = \frac{1}{N}\sum_{n=1}^N \left(\ell_n - \Gamma(D_N,\wp^{N,n})\right)\Delta_n.$$

Conditioning on $\mathscr{D}$, we have that

$$\mathbb{E}_N\left[\left(L_N - \tilde{L}_N\right)^2\Big|\mathscr{D}\right] = \frac{1}{N^2}\sum_{n=1}^N\left\{\int_{r\in[0,1]}\left((1-r') - \Gamma(D_N,\wp^{N,n})\right)^2\wp^{N,n}(D_N,dr)\right\}\Delta_n^2 \le \frac{1}{N}.$$

This implies the first part of (3).

To see the second part of (3), we write that $L^\circ_N - \bar{L}_N = \mathcal{E}^1_N + \mathcal{E}^2_N$ where

$$\mathcal{E}^1_N = \frac{1}{N}\sum_{n=1}^N\left\{\Gamma(D_N,\wp^{N,n}) - \Gamma(\bar{D}_N,\wp^{N,n})\right\}\Delta_n$$
$$\mathcal{E}^2_N = \frac{1}{N}\sum_{n=1}^N\Gamma(\bar{D}_N,\wp^{N,n})\left\{\Delta_N - p^{N,n}\right\}$$

We first calculate that

$$\mathbb{E}_N\left[\left|\mathcal{E}^2_N\right|^2\right] = \frac{1}{N^2}\sum_{n=1}^N\left(\Gamma(\bar{D}_N,\wp^{N,n})\right)^2 p^{N,n}\left(1-p^{N,n}\right) \le \frac{1}{4N}.$$

Thus $\lim_{N\to\infty}\mathbb{E}_N\left[|\mathcal{E}^2_N|^2\right] = 0$. Similarly,

$$\mathbb{E}_N\left[\left|D_N - \bar{D}_N\right|^2\right] = \frac{1}{N^2}\sum_{n=1}^N p^{N,n}\left(1-p^{N,n}\right) \le \frac{1}{4N}.$$

Thus in particular

$$\lim_{N\to\infty}\mathbb{E}_N\left[\left|D_N - \bar{D}\right|^2\right] = 0.$$

To bound $\mathcal{E}^1_N$, fix $\eta > 0$. By Prohorov's theorem, $\{\mathsf{U}_N;\ N\in\mathbb{N}\}$ is tight, so there is a $K_\eta \subset\subset C([0,1];\mathscr{P}[0,1])$ such that

$$\sup_{n\in\mathbb{N}}\mathsf{U}_N([0,1]\times K^c_\eta) < \eta.$$

Defining

$$\bar{\omega}_\eta(\delta) \overset{\text{def}}{=} \sup_{\substack{\wp\in K_\eta \\ D_1,D_2\in[0,1] \\ |D_1-D_2|<\delta}} |\Gamma(D_1,\wp) - \Gamma(D_2,\wp)|$$

for all $\delta > 0$, compactness of K_η and $[0,1]$ and continuity of Γ imply that $\lim_{\delta\searrow 0}\bar{\omega}_\eta(\delta) = 0$. Thus

$$\begin{aligned}|\mathcal{E}^1_N| &\le \mathbb{E}_N\left[\frac{1}{N}\sum_{n=1}^N\left|\Gamma(D_N,\wp^{N,n}) - \Gamma(\bar{D}_N,\wp^{N,n})\right|\right] \\ &= \mathbb{E}_N\left[\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\left|\Gamma(D_N,\wp^{N,n}) - \Gamma(\bar{D}_N,\wp^{N,n})\right|\mathsf{U}_N(d\mathsf{p})\right] \\ &\le 2\eta + 2\bar{\omega}_\eta(\varkappa) + \tilde{\mathbb{P}}_N\left\{|D_N - \bar{D}| \ge \varkappa\right\}.\end{aligned}$$

Take $N \nearrow \infty$, then let $\varkappa \searrow 0$. Finally let $\eta \searrow 0$. We conclude that indeed $\lim_{N\to\infty} \mathbb{E}_N[|\mathcal{E}_N^1|] = 0$. Collecting things together, the claim follows. □

4. Problem Formulation, Examples and Main Results

Let's now set up our framework for considering atypical behavior of the L_N's; i.e., large deviations. We will in particular focus our discussion on Example 2.3. We shall along the way make a number of definitions which will be crucial in our analysis.

One thing which is clear from Example 2.3 is that we need to keep track of the type associated with each default (but not the types associated with names which do not default). To organize this, let $\mathscr{M}_1(\mathsf{X})$ be the collection of measures ν on $(\mathsf{X}, \mathscr{B}(\mathsf{X}))$ such that $\nu(\mathsf{X}) \le 1$ (i.e., the collection of subprobability measures).

We can topologize $\mathscr{M}_1(\mathsf{X})$ in the usual way. In particular, fix a point $\star$ that is not in X and define $\mathsf{X}^+ \overset{\text{def}}{=} \mathsf{X} \cup \{\star\}$. Give X^+ the standard topology; open sets are those which are open subsets of X (with its original topology) or complements in X^+ of closed subsets of X (again, in the original topology of X). Define a bijection ι from $\mathscr{M}_1(\mathsf{X})$ to $\mathscr{P}(\mathsf{X}^+)$ by setting

$$(\iota\nu)(A) \overset{\text{def}}{=} \nu(A \cap \mathsf{X}) + (1 - \nu(\mathsf{X}))\,\delta_\star(A)$$

for all $A \in \mathscr{B}(\mathsf{X}^+)$. The topology of $\mathscr{M}_1(\mathsf{X})$ is the pullback of the topology of $\mathscr{P}(\mathsf{X}^+)$ and the metric on $\mathscr{M}_1(\mathsf{X})$ is that given by requiring ι to be an isometry.

Since X is Polish, so is X^+, and thus $\mathscr{P}(\mathsf{X}^+)$ is Polish, and thus $\mathscr{M}_1(\mathsf{X})$ is Polish. For each $N \in \mathbb{N}$, define a random element ν_N of $\mathscr{M}_1(\mathsf{X})$ as

$$\nu_N \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^{N} \Delta_n \delta_{\mathsf{p}^{N,n}}$$

Define a sequence $(Z_N)_{N\in\mathbb{N}}$ of $[0,1] \times \mathscr{M}_1(\mathsf{X})$-valued random variables as

$$Z_N \overset{\text{def}}{=} (L_N, \nu_N). \qquad N \in \mathbb{N}$$

Since $[0,1]$ and $\mathscr{M}_1(\mathsf{X})$ are both Polish, $[0,1] \times \mathscr{M}_1(\mathsf{X})$ is also Polish. We seek a large deviations principle for the Z_N's. Since projection maps are continuous, the contraction principle will then imply a large deviations principle for the L_N's. Note furthermore that

$$D_N = \nu_N(\mathsf{X});$$

the map $\nu \mapsto \nu(\mathsf{X})$ is continuous in the topology on $\mathscr{M}_1(\mathsf{X})$, so the recovery statistics depend continuously on ν_N.

Let's now see if we can identify a large deviations principle for the Z_N's in Example 2.3. Namely, we want to find a map $I : [0,1] \times \mathscr{M}_1(\mathsf{X}) \to [0,\infty]$ such that, at least informally,

$$\mathbb{P}_N\{Z_N \in dz^*\} \asymp \exp[-NI(z^*)] \qquad N \to \infty$$

for each fixed $z^* = (\ell^*, \nu^*) \in [0,1] \times \mathscr{M}_1(\mathsf{X})$.

To proceed, note that $\{\mathsf{p}_A, \mathsf{p}_B\}$ is the support of U and the ν_N's. If the support of ν^* contains some other point in X, then $\mathbb{P}_N\{Z_N \in dz^*\} = 0$. In other words, $I(z^*) = \infty$ if $\nu^* \not\ll \mathsf{U}$. Thus let's now assume that indeed $\nu^* \ll \mathsf{U}$. Then we explicitly have that

$$\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}_A) = \frac{\nu^*\{\mathsf{p}_A\}}{1/3} \qquad \text{and} \qquad \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}_B) = \frac{\nu^*\{\mathsf{p}_B\}}{2/3}.$$

Similarly

$$\frac{d\nu_N}{d\mathsf{U}}(\mathsf{p}_A) = \frac{\nu_N\{\mathsf{p}_A\}}{\mathsf{U}\{\mathsf{p}_A\}} = \frac{1}{N/3}\sum_{\substack{1\le n\le N\\ n\in 3\mathbb{N}}} \Delta_n$$

$$\frac{d\nu_N}{d\mathsf{U}}(\mathsf{p}_B) = \frac{\nu_N\{\mathsf{p}_B\}}{\mathsf{U}\{\mathsf{p}_B\}} = \frac{1}{2N/3}\sum_{\substack{1\le n\le N\\ n\not\in 3\mathbb{N}}} \Delta_n.$$

Thus

$$\mathbb{P}_N\{\nu_N \in d\nu^*\} = \mathbb{P}_N\left\{\nu_N\{\mathsf{p}_A\} \approx \nu^*\{\mathsf{p}_A\},\, \nu_N\{\mathsf{p}_B\} \approx \nu^*\{\mathsf{p}_B\}\right\}$$
$$= \mathbb{P}_N\left\{\frac{d\nu_N}{d\mathsf{U}}(\mathsf{p}_A) \approx \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}_A),\, \frac{d\nu_N}{d\mathsf{U}}(\mathsf{p}_B) \approx \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}_B)\right\}.$$

This is now essentially the focus of Sanov's theorem–i.i.d. Bernoulli random variables. For $p \in [0,1]$, define

(4) $$\hbar_p(x) \overset{\text{def}}{=} \begin{cases} x\ln\frac{x}{p} + (1-x)\ln\frac{1-x}{1-p} & \text{for } x \text{ and } p \text{ in } (0,1) \\ \ln\frac{1}{p} & \text{for } x=1,\, p\in(0,1] \\ \ln\frac{1}{1-p} & \text{for } x=0,\, p\in[0,1) \\ \infty & \text{else.} \end{cases}$$

Then

(5) $$\begin{aligned}\mathbb{P}_N\{\nu_N \in d\nu^*\} &\asymp \exp\left[-\frac{N}{3}\hbar_{p_A}\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}_A)\right) - \frac{2N}{3}\hbar_{p_A}\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}_B)\right)\right] \\ &= \exp\left[-N\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \hbar_p\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p})\right].\end{aligned}$$

Let's give a name to the right-hand side of this asymptotic expression.

Definition 4.1. *For $\nu \in \mathscr{M}_1(\mathsf{X})$, define*

$$H(\nu) \overset{\text{def}}{=} \begin{cases} \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p}) & \text{if } \nu \ll \mathsf{U} \\ \infty. & \text{else} \end{cases}$$

□

The functional H will play a central role in our calculations, so we will need to understand it fairly thoroughly. Note that

(6) $$\hbar_0(x) = \begin{cases} 0 & \text{if } x=0 \\ \infty & \text{else} \end{cases} \qquad \text{and} \qquad \hbar_1(x) = \begin{cases} 0 & \text{if } x=1 \\ \infty & \text{else} \end{cases}$$

Thus if $H(\nu) < \infty$, we can restrict the region of integration to get that

$$\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \chi_{\{0,1\}}(p)\hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p}) < \infty$$

so in fact

(7) $$\begin{aligned}\mathsf{U}\left\{\mathsf{p}=(p,\wp)\in\mathsf{X}:\ p=0 \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \neq 0\right\} &= 0 \\ \mathsf{U}\left\{\mathsf{p}=(p,\wp)\in\mathsf{X}:\ p=1 \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \neq 1\right\} &= 0.\end{aligned}$$

Next define

$$\lambda_p(\theta) \overset{\text{def}}{=} \ln\left(pe^\theta + 1 - p\right)$$

for all $\theta \in \mathbb{R}$ and $p \in [0,1]$. Then λ_p and $\hbar_p$ are convex duals; i.e.,

$$\begin{aligned}\hbar_p(x) &= \sup_{\theta\in\mathbb{R}}\{\theta x - \lambda_p(\theta)\} \qquad x\in\mathbb{R} \\ \lambda_p(\theta) &= \sup_{x\in\mathbb{R}}\{\theta x - \hbar_p(x)\}. \qquad \theta\in\mathbb{R}\end{aligned}$$

Lemma 4.2. *We have that*

(8) $$H(\nu) = \sup_{\phi\in C(\mathsf{X})}\left\{\int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu(dp) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right\}$$

for all $\nu \in \mathscr{M}_1(\mathsf{X})$.

□

The proof is given in Section 9. Finally, we have an approximation result.

Lemma 4.3. *Fix $\nu \in \mathscr{M}_1(\mathsf{X})$ such that $H(\nu) < \infty$. Then there is a sequence $(\nu_N)_{n\in\mathbb{N}}$ such that*

$$\lim_{N\in\mathbb{N}} \nu_N = \nu \qquad \text{and} \qquad \lim_{N\to\infty} H(\nu_N) = H(\nu)$$

(and thus $\nu_N \ll \mathsf{U}$ for all $N \in \mathbb{N}$) and such that

$$\mathsf{p} \mapsto \frac{d\nu_N}{d\mathsf{U}}(\mathsf{p}) \qquad \text{and} \qquad \mathsf{p} = (p, \wp) \mapsto \chi_{(0,1)}(p)\hbar'_p\left(\frac{d\nu_N}{d\mathsf{U}}(\mathsf{p})\right)$$

are both well-defined and in $C(\mathsf{X})$ for all $N \in \mathbb{N}$.

Again, the proof is in Section 9.

We now turn to the actual losses. Let's condition on the event that $\nu_N \approx \nu^*$;

$$\mathbb{P}_N\left\{Z_N \in dz^*\right\} = \mathbb{P}\left\{L_N \in d\ell^* \middle| \nu_N \approx \nu^*\right\}\mathbb{P}_N\left\{\nu_N \in d\nu^*\right\}. \tag{9}$$

If $\nu_N \approx \nu^*$, then there are about $N\nu^*\{\mathsf{p}_A\}$ defaults of type A, and $N\nu^*\{\mathsf{p}_B\}$ defaults of type B. Thus in law

$$L_N \approx \frac{1}{N}\left\{\sum_{1\le n\le N\nu^*\{\mathsf{p}_A\}}(1-\xi_n^A) + \sum_{1\le n\le N\nu^*\{\mathsf{p}_B\}}(1-\xi_n^B)\right\} \tag{10}$$

where the ξ_n^A's are i.i.d. with common law $\wp_A(\nu(\mathsf{X}),\cdot)$ and the ξ_n^B's are i.i.d. with common law $\wp_B(\nu(\mathsf{X}),\cdot)$. To understand the behavior of L_N under this measure, we need to construct its moment generating function.

Definition 4.4. *For $\nu \in \mathscr{M}_1(\mathsf{X})$, define*

$$\Lambda_\nu(\theta) \overset{def}{=} \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\left\{\ln\int_{r\in[0,1]} e^{\theta(1-r)}\wp(\nu(\mathsf{X}),dr)\right\}\nu(d\mathsf{p}) \qquad \theta\in\mathbb{R}$$

$$\Lambda^*_\nu(\ell) \overset{def}{=} \sup_{\theta\in\mathbb{R}}\left\{\theta\ell - \Lambda_\nu(\theta)\right\} \qquad \ell\in[0,1]$$

□

In our ongoing analysis of Example 2.3, we have that

$$\Lambda^*_\nu(\theta) = \nu^*\{\mathsf{p}_A\}\ln\int_{r\in[0,1]} e^{\theta(1-r)}\wp_A(\nu(\mathsf{X}),dr) + \nu^*\{\mathsf{p}_B\}\ln\int_{r\in[0,1]} e^{\theta(1-r)}\wp_B(\nu(\mathsf{X}),dr)$$

for all $\theta\in\mathbb{R}$; the logarithmic moment generating function of L_N of (10) is approximately $\theta \mapsto N\Lambda_{\nu^*}(\theta/N)$. Thus we should have that

$$\mathbb{P}_N\left\{L_N \in d\ell^* \middle| \nu_N \approx \nu^*\right\} \asymp \exp\left[-N\Lambda^*_\nu(\ell)\right]$$

We should then get the large deviations principle for Z_N by combining this, (5), and (9).

Our main result now makes sense. Recall that if $X_\circ$ is a Polish space and $P_\circ$ is a probability measure on $(X_\circ, \mathscr{B}(X_\circ))$, we say that a collection $(\xi_n)_{n\in\mathbb{N}}$ of $X_\circ$-valued random variables has a large deviations principle with rate function $I_\circ : X \to [0,\infty]$ if

(1) For each $s \ge 0$, the set $\Phi(s) \overset{\text{def}}{=} \{x \in X_\circ : I_\circ(x) \le s\}$ is a compact subset of $X_\circ$.
(2) For every open $G \subset X_\circ$,

$$\varliminf_{n\nearrow\infty} \frac{1}{n}\ln\mathbb{P}_\circ\left\{\xi_n \in G\right\} \ge -\inf_{x\in G} I_\circ(x)$$

(3) For every closed $F \subset X_\circ$,

$$\varlimsup_{n\nearrow\infty} \frac{1}{n}\ln\mathbb{P}_\circ\left\{\xi_n \in F\right\} \le -\inf_{x\in F} I_\circ(x).$$

Theorem 4.5 (Main). *We have that $(Z_N)_{N\in\mathbb{N}}$ has a large deviations principle with rate function*

$$I(z) = \Lambda^*_\nu(\ell) + H(\nu)$$

for all $z = (\ell,\nu) \in [0,1] \times \mathscr{M}_1(\mathsf{X})$. Secondly $(L_N)_{N\in\mathbb{N}}$ has a large deviations principle with rate function

$$I'(\ell) \overset{def}{=} \inf_{\nu\in\mathscr{M}_1(\mathsf{X})} I(\ell,\nu) \tag{11}$$

for all $\ell \in [0,1]$.

□

Proof. Combine together Propositions 5.1, 6.2, and 7.1. This gives the large deviations principle for (Z_N). The large deviations principle for $(L_N)_{N\in\mathbb{N}}$ follows from the contraction principle and the continuity of the map $\nu \mapsto \nu(\mathsf{X})$. □

One way to interpret Theorem 4.5 is that the rate functions I and I' give the correct way to find the "minimum-energy" configurations for atypically large losses to occur. In general, variational problems involving measures can be computationally difficult, so Section 8 addresses some computational issues. In particular, we find an alternate expression which takes advantage of the specific structure of our problem. Define

$$M_\wp(\theta,D) \overset{\text{def}}{=} \ln \int_{r\in[0,1]} e^{\theta(1-r)} \wp(D,dr) \qquad \theta\in\mathbb{R}$$

$$I_\wp(x,D) \overset{\text{def}}{=} \sup_{\theta\in\mathbb{R}} \{\theta x - M_\wp(\theta,D)\} \qquad x\in\mathbb{R}$$

for all $D\in[0,1]$; then

$$\Lambda_\nu(\theta) = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} M_\wp(\theta,\nu(\mathsf{X}))\nu(d\mathsf{p}) \tag{12}$$

for all $\nu\in\mathscr{M}_1(\mathsf{X})$ and $\theta\in\mathbb{R}$. Define next $\mathcal{B} \overset{\text{def}}{=} B(\mathsf{X};[0,1])$.

Theorem 4.6. *For $\ell\in[0,1]$ and $\mathsf{U}\in\mathscr{P}(\mathsf{X})$, set*

$$\begin{aligned} J'(\ell) \overset{def}{=} \inf_{\Phi,\,\Psi\in\mathcal{B}} \Bigg\{ \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left\{ \Phi(\mathsf{p}) I_\wp\left(\Psi(\mathsf{p}), \int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\mathsf{U}(d\mathsf{p})\right) + \hbar_p(\Phi(\mathsf{p}))\right\} \mathsf{U}(d\mathsf{p}) : \\ \int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p}) = \ell \Bigg\}. \end{aligned} \tag{13}$$

We have that $I'(\ell) = J'(\ell)$ for all $\ell\in[0,1]$. An alternate representation of J' is

$$\begin{aligned} J''(\ell) \overset{def}{=} \inf_{D\in[\ell,1]} \inf_{\Phi\in\mathcal{B}} \inf_{\Psi\in\mathcal{B}} \Bigg\{ \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \{\Phi(\mathsf{p}) I_\wp(\Psi(\mathsf{p}),D) + \hbar_p(\Phi(\mathsf{p}))\} \mathsf{U}(d\mathsf{p}) : \\ \int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p}) = \ell,\ \int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\mathsf{U}(d\mathsf{p}) = D \Bigg\}. \end{aligned} \tag{14}$$

□

The proof of this is given in Section 8. The point of the second representation (14) is that the innermost minimization problem (the one with Ψ and D fixed) involves linear constraints; that Φ take values in $[0,1]$ and that two integrals of Φ take specific values. This will be useful in some of our numerical studies in the next section.

4.1. **Numerical Examples.** Let's see what our calculations look like in some specific cases. To focus on the effects of recovery, let's assume a common probability of default of 20%; i.e., $p^{N,n} = 0.2$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots N\}$. We will consider four specific cases, one with fixed recovery rate, a homogenous pool with variable recovery rate and two heterogenous pools with variable recovery rates. In all cases, the expected loss will be 14%, but we will see that the tails (the large deviations principle rate functions) are significantly different. Although our theory has primarily focussed on the rate function in the large deviations principle for $(L_N)_{N\in\mathbb{N}}$, the solution of the variational problem (11) (or equivalently (13) or (14)) gives useful information. In particular, we shall extract some useful information about *implied recovery*.

In **Case 1**, let's assume that the recovery rate is fixed at 30%; i.e., $\wp^{N,n} = \delta_{0.3}$ for all $N \in \mathbb{N}$ and $n \in \{1, 2 \dots N\}$. Setting $\mathsf{p}_1^* \overset{\text{def}}{=} (0.2, \delta_{0.3})$, we here have that $\mathsf{U} = \delta_{\mathsf{p}_1^*}$. We first observe that $H(\nu) < \infty$ if and only if $\nu = \alpha\delta_{\mathsf{p}_1^*}$ for some $\alpha \in [0, 1]$; in this case, $\alpha = \nu(\mathsf{X})$, so $H(\nu) < \infty$ if and only if $\nu = \nu(\mathsf{X})\delta_{\mathsf{p}_1^*}$. For such a ν, $H(\nu) = \hbar_{0.2}(\alpha) = \hbar_{0.2}(\nu(\mathsf{X}))$ and

$$\Lambda_\nu(\theta) = \alpha \ln e^{0.7\theta} = 0.7\theta\nu(\mathsf{X}). \qquad \theta \in \mathbb{R}$$

If $\nu = \nu(\mathsf{X})\delta_{\mathsf{p}_1^*}$, then

$$\Lambda^*_\nu(\ell) = \begin{cases} 0 & \text{if } \ell = 0.7\nu(\mathsf{X}) \\ \infty & \text{else} \end{cases}$$

Collecting things together, we have that $(Z_N)_{N\in\mathbb{N}}$ and $(L_N)_{N\in\mathbb{N}}$ are governed, respectively, by the action functionals.

$$I_1(\ell, \nu) = \begin{cases} \hbar_{0.2}\left(\frac{\ell}{0.7}\right) & \text{if } \nu = \frac{\ell}{0.7}\delta_{\mathsf{p}_1^*} \\ \infty & \text{else} \end{cases}$$

$$I_1'(\ell) = \hbar_{0.2}\left(\frac{\ell}{0.7}\right)$$

We note that $I_1'(\ell)$ is finite only if $0 \le \ell \le 0.7$.

In **Case 2** we consider a homogeneous pool with the recovery rate following a beta distribution. For $\beta > 0$, define

$$\mu_\beta(A) \overset{\text{def}}{=} \beta \int_{r\in A} (1-r)^{\beta-1} dr; \qquad A \in \mathscr{B}[0, 1]$$

this is the law of the beta distribution with parameters 1 and β. As β increases, the amount of mass near 1 decreases. We also have that

$$\int_{r\in[0,1]} r\mu_\beta(dr) = \frac{1}{1+\beta}$$

for all $\beta > 0$ (as β increases, the mean of μ_β decreases); this will allow a number of explicit formulae for the expected recovery (given the default rate).

We want to consider the case that the recovery is in an appropriate sense negatively correlated with the defaults; i.e., that more defaults imply less recovery. This is a documented empirical observation in the financial literature, e.g., see [SH09], [ABRS05] and the references therein. We here assume that the recovery rate has a beta distribution whose parameters depend linearly and monotonically on the empirical default rate. Namely, if the default rate is D, then the recoveries will all have common beta distribution with parameters 1 and

$$f_{\text{aff}}(D) \overset{\text{def}}{=} \frac{1}{0.3 - 0.25(D - 0.2)} - 1;$$

note that $f : [0, 1] \to \mathbb{R}_+$. Define $\wp_{\text{aff}}(D, \cdot) \overset{\text{def}}{=} \mu_{f_{\text{aff}}(D)}$ for all $D \in [0, 1]$. This choice of f_{aff} results in a conditional expected recovery which is affine in D; i.e.,

$$\int_{r\in[0,1]} r\wp_{\text{aff}}(D, dr) = 0.3 - 0.25(D - 0.2).$$

Set $\mathsf{p}_{\text{aff}}^* \overset{\text{def}}{=} (0.2, \wp_{\text{aff}})$; then $\mathsf{U} = \delta_{\mathsf{p}_{\text{aff}}^*}$. According to (1), the typical loss is

$$\bar{L} = 0.2 \times \left(1 - \frac{1}{1 + f_{\text{aff}}(0.2)}\right) = 0.2 \times (1 - 0.3) = 0.14.$$

The action functionals in this case are somewhat similar to those in Case 1. Again we have that

$$H(\nu) = \begin{cases} \hbar_{0.2}(\nu(\mathsf{X})) & \text{if } \nu = \nu(\mathsf{X})\delta_{\mathsf{p}_2^*} \\ \infty. & \text{else} \end{cases}$$

For each $\beta > 0$, define

$$\check{\Lambda}_\beta(\theta) \stackrel{\text{def}}{=} \ln \int_{r\in[0,1]} e^{\theta(1-r)}\mu_\beta(dr) \qquad \theta \in \mathbb{R}$$
$$\check{\Lambda}^*_\beta(\ell) \stackrel{\text{def}}{=} \sup_{\theta\in\mathbb{R}} \left\{\theta\ell - \check{\Lambda}_\beta(\theta)\right\}. \qquad \ell \in \mathbb{R}$$

Then if $\nu = \nu(\mathsf{X})\delta_{\mathsf{p}_2^*}$ where $\nu(\mathsf{X}) > 0$,

$$\Lambda_\nu(\theta) = \nu(\mathsf{X})\check{\Lambda}_{f_{\text{aff}}(\nu(\mathsf{X}))}(\theta) \qquad \theta \in \mathbb{R}$$
$$\Lambda^*_\nu(\ell) = \nu(\mathsf{X})\check{\Lambda}^*_{f_{\text{aff}}(\nu(\mathsf{X}))}\left(\frac{\ell}{\nu(\mathsf{X})}\right). \qquad \ell \in \mathbb{R}$$

Of course if $\nu = 0$, then $\Lambda_\nu \equiv 0$, and

$$\Lambda^*_\nu(\ell) = \begin{cases} 0 & \text{if } \ell = 0 \\ \infty. & \text{else.} \end{cases}$$

Here $(Z_N)_{N\in\mathbb{N}}$ and $(L_N)_{N\in\mathbb{N}}$ are governed, respectively, by the action functionals.

$$(15) \qquad \begin{aligned} I_2(\ell,\nu) &= \begin{cases} \hbar_{0.2}(\nu(\mathsf{X})) + \nu(\mathsf{X})\check{\Lambda}^*_{f_{\text{aff}}(\nu(\mathsf{X}))}\left(\frac{\ell}{\nu(\mathsf{X})}\right) & \text{if } \nu = \nu(\mathsf{X})\delta_{\mathsf{p}_2^*} \\ \infty & \text{else} \end{cases} \\ I'_2(\ell) &= \inf_{D\in(0,1]} \left\{\hbar_{0.2}(D) + D\check{\Lambda}^*_{f_{\text{aff}}(D)}\left(\frac{\ell}{D}\right)\right\}. \end{aligned}$$

In **Case 3**, we replace f_{aff} of Case 2 with one that results in a conditional expected recovery which is quadratic in D; this allows us some insight into the effects of convexity in the conditional expected recovery. We set

$$f_{\text{q}}(D) = \frac{1}{0.3 - 0.25(D-0.2) - 0.1(D-0.2)^2} - 1.$$

Again we have that $f_{\text{q}} : [0,1] \to \mathbb{R}_+$, and we set $\wp_{\text{q}}(D,\cdot) = \mu_{f_{\text{q}}(D)}$, $\mathsf{p}^*_{\text{q}} \stackrel{\text{def}}{=} (0.2, \wp_{\text{q}})$, and have that $\mathsf{U} = \delta_{\mathsf{p}^*_{\text{q}}}$ Here we have that

$$\int_{r\in[0,1]} r\wp_{\text{q}}(D,dr) = 0.3 - 0.25(D-0.2) - 0.1(D-0.2)^2.$$

and we again get that $\bar{L} = 0.14$. The corresponding action functional is I'_3.

Case 4 involves the beta distribution again. Here, however, we now consider a heterogeneous pool of two types (Example 2.3). We concentrate on the effect of the heterogeneity in the recovery distribution, so as in the previous cases all bonds will have default probability of 20%. For all $D \in [0,1]$, every third bond will have recovery distribution governed by $\wp_{\text{aff}}$ and the remaining bonds will have recovery distribution governed by $\wp_{\text{q}}$. We thus have that $\mathsf{U} = \frac{1}{3}\delta_{\mathsf{p}^*_{\text{aff}}} + \frac{2}{3}\delta_{\mathsf{p}^*_{\text{q}}}$. It is easy to see that again the typical loss is:

$$\bar{L} = 0.2 \times \frac{1}{3} \times \left(1 - \frac{1}{1+f_{\text{aff}}(0.2)}\right) + 0.2 \times \frac{2}{3} \times \left(1 - \frac{1}{1+f_{\text{q}}(0.2)}\right) = 0.14.$$

For notational convenience and in order to illustrate the usage of Theorem 4.6 we use the alternative representation (14). If $\wp(D,\cdot) = \mu_{f(D)}$ for some $f \in C([0,1];\mathbb{R}_+)$, then for all $D \in [0,1]$, we have that $M_\wp(\theta,D) = \check{\Lambda}_{f(D)}(\theta)$ for all $\theta \in \mathbb{R}$ and $I_\wp(\ell,D) = \check{\Lambda}^*_{f(D)}(\ell)$ for all $\ell \in [0,1]$. Since the support of U is exactly $\{\mathsf{p}^*_{\text{aff}}, \mathsf{p}^*_{\text{q}}\}$, we can consider Φ and Ψ in $\mathcal{B}$ of the form

$$\Phi = \phi_A\chi_{\{\mathsf{p}^*_{\text{aff}}\}} + \phi_B\chi_{\{\mathsf{p}^*_{\text{q}}\}} \qquad \text{and} \qquad \Psi = \psi_A\chi_{\{\mathsf{p}^*_{\text{aff}}\}} + \psi_B\chi_{\{\mathsf{p}^*_{\text{q}}\}}.$$

Thus (14) becomes

$$I_4'(\ell) = \inf_{D\in[\ell,1]} \inf_{\psi_A,\psi_B\in[0,1]} \inf_{\substack{\phi_A,\phi_B\in[0,1]\\ \phi_A\psi_A/3+2\phi_A\psi_B/3=\ell\\ \phi_A/3+2\phi_B/3=D}} \left\{\frac{1}{3}\phi_A I_{\wp_{\mathrm{aff}}}(\psi_A,D)+\frac{2}{3}\phi_B I_{\wp_{\mathrm{q}}}(\psi_B,D) + \frac{1}{3}\hbar_{p_A}(\phi_A)+\frac{2}{3}\hbar_{p_B}(\phi_B)\right\}$$

(16)

$$= \inf_{D\in[\ell,1]} \inf_{\psi_A,\psi_B\in[0,1]} \inf_{\substack{\phi_A,\,\phi_B\in[0,1]\\ \phi_A\psi_A/3+2\phi_A\psi_B/3=\ell\\ \phi_A/3+2\phi_B/3=D}} \left\{\frac{1}{3}\phi_A \check{\Lambda}^*_{f_{\mathrm{aff}}(D)}(\psi_A)+\frac{2}{3}\phi_B \check{\Lambda}^*_{f_{\mathrm{q}}(D)}(\psi_B) + \frac{1}{3}\hbar_{p_A}(\phi_A)+\frac{2}{3}\hbar_{p_B}(\phi_B)\right\}$$

Lastly, in **Case 5**, instead of working with an f_{aff} that results in a conditional expected recovery which is affine in D (as in case 2), we assume that f_{aff} is itself affine. In particular, we assume that $f_{\mathrm{aff}}(D) = 4/3+5D$. Other than this difference, the calculations are identical to those in case 2. Again we get that $\bar{L} = 0.14$. The corresponding action functional is I_5'.

In Figure 4.1, we plot the rate functions I_1', I_2', I_3', I_4' and I_5'. We use a Monte Carlo procedure to compute $\check{\Lambda}$ and $\check{\Lambda}^*$. As expected, all action functions are nonnegative and zero at the (common) expected loss of $\bar{L} = 0.14$. We observe that $I_3' \le I_4' \le I_2' \le I_5' \le I_1'$. In particular, the heterogeneous case, which

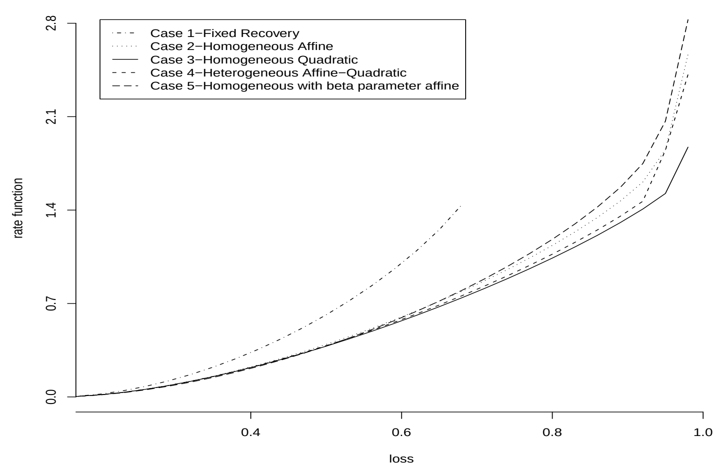

FIGURE 1. Action functionals for fixed recovery, for the homogeneous cases and the heterogeneous case.

is a mixture of an affine conditional expected recovery and a quadratic conditional expected recovery, is in between the two homogeneous cases (cases 2 and 3). We of course should not be surprised that the rate function in Case 1 is larger than that in Cases 2 through 5, there are in general many more configurations which lead to a given overall loss rate.

A second useful insight which we can numerically extract is the "preferred" way which losses stem from defaults versus recovery. For each $\ell \in [0,1]$, let $D_*(\ell)$ be the minimizer[2] in the expression (15) for I_2', I_3' or I_5' or alternately the expression (16) for I_4'. We assume that these minimizers are unique. For $\ell > \bar{L}$ and $\delta > 0$, the Gibbs conditioning principle [DZ98, Section 7.3] implies that we should have that

(17) $$\lim_{N\to\infty} \mathbb{P}_N\left\{|D_N - D_*(\ell)| \ge \delta \middle| L_N \ge \ell\right\} = 0.$$

[2]Case 1 is of course degenerate this sense; for a given loss rate ℓ, the default rate must be very close to $\ell/.7$.

In other words, conditional on the pool suffering losses exceeding rate ℓ, the default rate should converge to $D_*(\ell)$. Using this information, we can then say something about the implied recovery (see [SH09]). We write that

$$\text{Loss=Default} \times \text{(1-Recovery)}$$

to find an effective recovery rate in terms of the loss rate and the default rate. This recovery rate quantifies the fact that losses are due to *both* default and recovery. For atypically large losses in a large pool of credit assets, we should combine this with the Gibbs conditioning calculation of (17). Namely, let's define

$$\mathcal{R}(\ell) = 1 - \frac{\ell}{D_*(\ell)}.$$

This gives us the implied recovery for atypically large losses. Note that in the "typical case" (Section 3), the implied recovery is simply the conditional expectation of the recovery given that the default rate is $\bar{D}$; i.e.,

$$\mathcal{R}(\bar{L}) = \frac{\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} p\left\{\int_{r\in[0,1]} r\wp(\bar{D},dr)\right\}\mathsf{U}(d\mathsf{p})}{\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} p\mathsf{U}(d\mathsf{p})}.$$

Since we have a single default rate of 20% in our examples, $\mathcal{R}(0.14) = .3$ in all of our examples. Under the assumption that D^* is an invertible function, we can also formalize the dependence of recovery on default by letting $\mathcal{R}^* : [0,1] \to [0,1]$ be such that $\mathcal{R}^*(D_*(\ell)) = \mathcal{R}(\ell)$ for all $\ell \in [0,1]$. Figure 4.1 is a plot of $\mathcal{R}^*$ for the cases which we are studying. We observe that, for Cases $2, 3, 4$ and 5, the implied recovery and the optimal defaults are negatively correlated and that the implied recovery is convex as a function of the optimal defaults (see also [SH09] and [ABRS05]). The convexity is clearer in Case 5. Moreover, the graph of the heterogeneous case is between the graph of the homogeneous cases 2 and 3.

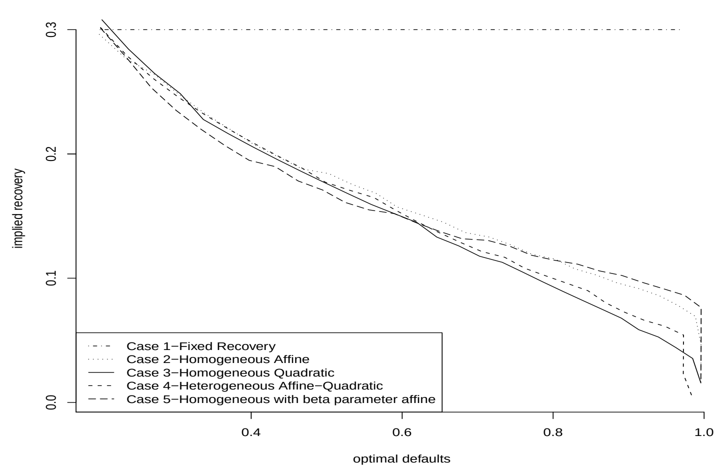

FIGURE 2. Implied recovery versus optimal defaults.

Using the implied recovery $\mathcal{R}^*$, we can shed a bit more light on the significance of rare recovery amounts versus rare default rates. For each $D \in [0,1]$, let's define

$$\mathcal{R}_\circ(D) \stackrel{\text{def}}{=} \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left\{\int_{r\in[0,1]} r\wp(D,dr)\right\}\mathsf{U}(d\mathsf{p}).$$

Thus $\mathcal{R}_\circ$ is the expected recovery rate conditioned on the default rate. In Case 1, $\mathcal{R}_\circ(D) \equiv .3$ for all $D \in [0,1]$, and in all cases $\mathcal{R}_\circ(\bar{D}) = .3$. We note that the *average* recovery rate will in general not coincide with the *optimal* recovery rate. In other words, the most likely recovery rate need not be the average recovery rate for the most likely default rate; $\mathcal{R}_\circ(\ell)$ and $\mathcal{R}^*(\ell)$ will in general be different. The conditional expected recovery $\mathcal{R}_\circ$ takes into account only the structure of rare default rates, but not rare recovery rates. In order to quantify this, let us define the ratio

$$\varrho(D) \stackrel{\text{def}}{=} \frac{\mathcal{R}^*(D)}{\mathcal{R}_\circ(D)}.$$

In Figure 4.1 we plot ϱ for Cases 2, 3, 4 and 5 ($\varrho \equiv 1$ in Case 1).

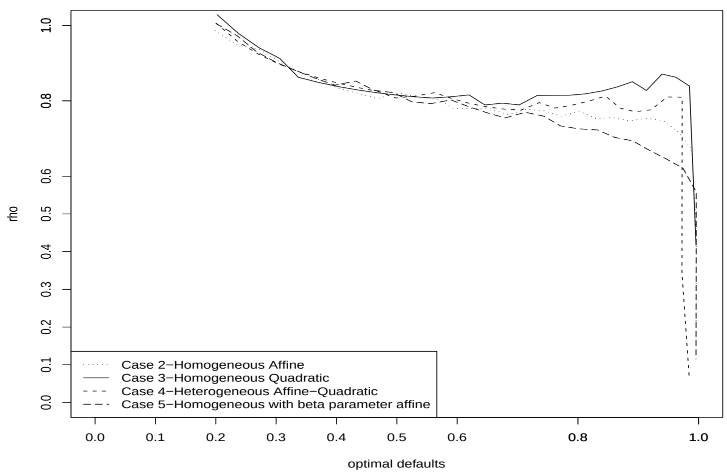

FIGURE 3. ϱ versus optimal defaults.

Remark 4.7. *Implied recovery points out one of the strengths of large deviations. Numerical computation of $D_*(\ell)$ is essentially a free byproduct of the outer minimization in* (14)*. On the other hand, simulation of implied recovery in cases of atypically large defaults would entail sampling rare events, which is very computationally intensive. The various minimizers in* (14) *naturally give information about the precise structure of rare losses.*

5. LOWER SEMICONTINUITY

The first part of the large deviations claim is that the level sets of I are compact. The proof follows along fairly standard lines.

Proposition 5.1 (Compactness of Level Sets)**.** *For each $s \geq 0$, the set*

$$\Phi(s) \stackrel{def}{=} \{z \in [0,1] \times \mathscr{M}_1(\mathsf{X}) : I(z) \leq s\}$$

is a compact subset of $[0,1] \times \mathscr{M}_1(\mathsf{X})$.

□

Proof. We first claim that $\Phi(s)$ is contained in a compact subset of $[0,1] \times \mathscr{M}_1(\mathsf{X})$. Since $[0,1]$ is already compact, it suffices to show that $\Phi_M(s) \stackrel{\text{def}}{=} \{\nu \in \mathscr{M}_1(\mathsf{X}) : H(\nu) \leq s\}$ is a compact subset of $\mathscr{M}_1(\mathsf{X})$. If $\nu \in \Phi_M(s)$, then $\nu \ll \mathsf{U}$ and, since $\hbar_p(x) = \infty$ for $x > 1$, we have that

$$\mathsf{U}\left\{\mathsf{p} \in \mathsf{X} : \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) > 1\right\} = 0,$$

so for any $A \in \mathscr{B}(\mathsf{X})$, $\nu(A) \leq \mathsf{U}(A)$. Since U itself is tight (it is a probability measure on a Polish space), $\Phi_M(s)$ is tight; for every $\varepsilon > 0$, there is a $K_\varepsilon \subset\subset \mathsf{X}$ such that $\nu(\mathsf{X} \setminus K_\varepsilon) < \varepsilon$ for all $\nu \in \Phi_M(s)$. We claim that thus $\iota(\Phi_M(s))$ is also tight. Indeed, fix $\varepsilon > 0$. Letting $\iota_\circ : \mathsf{X} \to \mathsf{X}^+$ be the inclusion map, we have that $\iota_\circ$ is continuous, and thus $\iota_\circ(K_\varepsilon)$ is compact. Since singletons are also compact, $K^* \stackrel{\text{def}}{=} \iota_\circ(K_\varepsilon) \cup \{\star\}$ is a compact subset of X^+. For every $\nu \in \Phi_M(s)$, $(\iota\nu)(\mathsf{X}^+ \setminus K^*) = \nu(\mathsf{X} \setminus K_\varepsilon) < \varepsilon$, so indeed $\iota(\Phi_M(s))$ is tight. Thus $\overline{\Phi_M(s)} \subset\subset \mathscr{P}(\mathsf{X}^+)$ and hence

$$\Phi_M(s) \subset \iota^{-1}\overline{\iota(\Phi_M(s))} \subset\subset \mathscr{M}_1(\mathsf{X})$$

the last claim following since ι is a homeomorphism. Gathering things together, we have that $\Phi(s)$ is indeed contained in a compact subset of $[0,1] \times \mathscr{M}_1(\mathsf{X})$.

We now want to show that $\Phi(s)$ is closed, or equivalently, that $([0,1] \times \mathscr{M}_1(\mathsf{X})) \setminus \Phi(s)$ is open. Using Lemma 4.2, we have that

$$\begin{aligned} &([0,1] \times \mathscr{M}_1(\mathsf{X})) \setminus \Phi(s) \\ &\qquad = \bigcup_{\substack{\theta\in\mathbb{R} \\ \phi\in C(\mathsf{X})}} \left\{(\ell,\nu) \in \mathscr{M}_1(\mathsf{X}) : \theta\ell + \int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu(d\mathsf{p}) - \Lambda_\nu(\theta) > s + \int_{\mathsf{p}=(p,wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right\}. \end{aligned}$$

For each $\theta \in \mathbb{R}$ and $\phi \in C(\mathsf{X})$, then map $(\ell,\nu) \mapsto \theta\ell + \int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu(d\mathsf{p}) - \Lambda_\nu(\theta)$ is continuous, so we have written $([0,1] \times \mathscr{M}_1(\mathsf{X})) \setminus \Phi(s)$ as a union of open sets. □

6. Large Deviations Lower Bound

We next prove the large deviations lower bound. As with most large deviations lower bounds, the idea is to find a measure transformation under which the set of interest becomes "typical". In this case, this measure transformation will come from a combination of Cramer's theorem and Sanov's theorem.

We start with a simplified lower bound where the measure transformation in Cramer's theorem is fairly explicit. For each $\nu \in \mathscr{M}_1(\mathsf{X})$, we make the usual definition [DZ98][Appendix A] that

$$\operatorname{dom}\Lambda^*_\nu \overset{\text{def}}{=} \{\ell \in [0,1] : \Lambda^*_\nu(\ell) < \infty\};$$

this will of course be an interval; ri dom Λ^*_ν will be the relative interior of dom Λ^*_ν.

Proposition 6.1. *Fix an open subset G of $[0,1] \times \mathscr{M}_1(\mathsf{X})$ and $z = (\ell,\nu) \in G$ such that $I(z) < \infty$ and $\ell \in$ ri dom Λ^*_ν. Then*

$$\varliminf_{N\to\infty} \frac{1}{N} \ln \mathbb{P}_N\{Z_N \in G\} \ge -I(z). \tag{18}$$

□

The proof will require a number of tools. Since $\ell \in$ ri dom Λ^*_ν, there is a $\theta \in \mathbb{R}$ such that

$$\Lambda'_\nu(\theta) = \ell \qquad \text{and} \qquad \Lambda^*_\nu(\ell) = \theta\Lambda'_\nu(\theta) - \Lambda_\nu(\theta) \tag{19}$$

(see [DZ98][Appendix A]). Let's now fix a relaxation parameter $\eta > 0$. Then there is an $\eta_1 \in (0,\eta)$ and an open neighborhood $\mathcal{O}_1$ of ν such that $(\ell-\eta_1, \ell+\eta_1) \times \mathcal{O}_1 \subset G$. Using the first equality of (19), we have that $(\Lambda'_\nu(\theta),\nu) = (\ell,\nu) \in (\ell-\eta_1,\ell+\eta_1) \times \mathcal{O}_1$. Since the maps $(\tilde\eta,\tilde\nu) \to (\Lambda'_{\tilde\nu}(\theta)+\tilde\eta,\tilde\nu)$ and $\tilde\nu \mapsto \Lambda_{\tilde\nu}(\theta)$ are continuous, there is an $\eta_2 \in (0,1)$ and an open subset $\mathcal{O}_2$ of $\mathscr{M}_1(\mathsf{X})$ such that

$$\begin{aligned} &\{(\Lambda'_{\tilde\nu}(\theta)+\tilde\eta,\tilde\nu) : \tilde\eta \in (0,\eta_2),\, \tilde\nu \in \mathcal{O}_2\} \subset (\ell-\eta_1,\ell+\eta_1) \times \mathcal{O}_1 \\ &|\Lambda_{\tilde\nu}(\theta) - \Lambda_\nu(\theta)| < \eta \qquad \text{for } \tilde\nu \in \mathcal{O}_2. \end{aligned}$$

We next want to use Lemma 4.3 want to choose a particularly nice element of $\mathcal{O}_2$. Namely, Lemma 4.3 ensures that there is a $\nu^* \in \mathcal{O}_2$ such that $\nu^* \ll \mathsf{U}$ and such that both $\frac{d\nu^*}{d\mathsf{U}}$ and

$$\phi(\mathsf{p}) \overset{\text{def}}{=} \chi_{(0,1)}(p)\hbar'_p\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\right) \qquad \mathsf{p} = (p,\wp) \in \mathsf{X} \tag{20}$$

are in $C(\mathsf{X})$ and such that $|H(\nu^*) - H(\nu)| < \eta$. Let $\mathcal{O}_3$ be an open subset of $\mathcal{O}_2$ which contains ν^* and such that

$$\left|\int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\tilde\nu(d\mathsf{p}) - \int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu^*(d\mathsf{p})\right| < \eta$$

for all $\tilde\nu \in \mathcal{O}_3$.

We can now proceed with our measure change. For each $N \in \mathbb{N}$, define

$$A_1^{(N)} \overset{\text{def}}{=} \theta L_N - \Lambda_{\nu_N}(\theta) \qquad \text{and} \qquad A_2^{(N)} \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N \Delta_n \phi(\mathsf{p}^{N,n}) - \lambda_{p^{N,n}}(\phi(\mathsf{p}^{N,n}))$$

Note that

$$\mathbb{E}_N\left[\exp\left[NA_1^{(N)}\right]\middle|\mathscr{D}\right] = 1 \qquad \text{and} \qquad \mathbb{E}_N\left[\exp\left[NA_2^{(N)}\right]\right] = 1.$$

Define a new probability measure as

$$\tilde{\mathbb{P}}_N(A) \overset{\text{def}}{=} \mathbb{E}_N\left[\chi_A \exp\left[N\left\{A_1^{(N)} + A_2^{(N)}\right\}\right]\right]. \qquad A \in \mathscr{B}(\Omega)$$

This will be the desired measure change.

Define

$$S_N \overset{\text{def}}{=} \left\{\left|L_N - \Lambda'_{\nu_N}(\theta)\right| < \eta_2,\, \nu_N \in \mathcal{O}_3\right\}.$$

On S_N,

$$(L_N, \nu_N) = \left(\Lambda'_{\nu_N}(\theta) + \left\{L_N - \Lambda'_{\nu_N}(\theta)\right\}, \nu_N\right) \in (\ell - \eta_1, \ell + \eta_1) \times \mathcal{O}_1 \subset G$$

so in fact $S_N \subset G$. Thus

$$\mathbb{P}_N\{Z_N \in G\} \geq \tilde{\mathbb{P}}_N\left[\chi_{S_N} \exp\left[-N\left\{A_1^{(N)} + A_2^{(N)}\right\}\right]\right].$$

Let's also assume that N is large enough that

$$\left|\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}_N(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right| < \eta.$$

Collecting our requirements together, we have that

$$A_1^{(N)} \leq \theta\ell + |\theta|\eta_1 - \Lambda_\nu(\theta) + \eta = \Lambda^*_\nu(\ell) + (|\theta| + 1)\,\eta$$

and

$$\begin{aligned} A_2^{(N)} = \int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu_N(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}_N(d\mathsf{p}) \\ \leq \int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu^*(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p}) + 2\eta \\ = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left\{\phi(\mathsf{p})\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}) - \lambda_p(\phi(\mathsf{p}))\right\}\mathsf{U}(d\mathsf{p}) + 2\eta \\ = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \hbar_p\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p}) + 2\eta \leq H(\nu) + 3\eta. \end{aligned}$$

Thus

$$\mathbb{P}_N\{Z_N \in G\} \geq \tilde{\mathbb{P}}_N(S_N)\exp\left[-N\left\{I(z) - (|\theta| + 4)\,\eta\right\}\right]$$

We have used here the calculation that

$$\hbar_p\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\right) = \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\phi(\mathsf{p}) - \lambda_p(\phi(\mathsf{p})) \tag{21}$$

for U-almost all $\mathsf{p} = (p, \wp) \in \mathsf{X}$. This follows from standard convex analysis and the form (20) of ϕ when $p \in (0,1)$. Since $H(\nu^*) < \infty$, (7) implies that, except on a U-negligible set,

$$\hbar_p\left(\frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\right) = \hbar_p(p) = 0 = p \times 0 - \lambda_p(0) = \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\phi(\mathsf{p}) - \lambda_p(\phi(\mathsf{p}))$$

if $\mathsf{p} = (p, \wp) \in \mathsf{X}$ is such that $p \in \{0, 1\}$. In other words, (21) holds except on a U-negligible set.

We now want to show that $\underline{\lim}_{N\to\infty} \tilde{\mathbb{P}}_N(S_N) > 0$, which will in turn follow if $\overline{\lim}_{N\to\infty} \tilde{\mathbb{P}}_N(S_N^c) = 0$. To organize our thoughts, we write that

$$\tilde{\mathbb{P}}_N(S_n^c) \leq \tilde{\mathbb{P}}_N\left\{\nu_N \not\in \mathcal{O}_3\right\} + \tilde{\mathbb{E}}_N\left[\tilde{\mathbb{P}}_N\left\{|L_N - \Lambda'_{\nu_N}(\theta)| \geq \eta_3 \middle| \mathscr{D}\right\}\chi_{\{\nu_N \in \mathcal{O}_3\}}\right]$$

$$\leq \tilde{\mathbb{P}}_N\left\{\nu_N \notin \mathcal{O}_3\right\} + \frac{1}{\eta_3^2}\tilde{\mathbb{E}}_N\left[\tilde{\mathbb{E}}_N\left[\left|L_N - \Lambda'_{\nu_N}(\theta)\right|^2 \middle| \mathscr{D}\right]\chi_{\{\nu_N\in\mathcal{O}_3\}}\right].$$

We can now finish the proof of our initial lower bound.

Proof of Proposition 6.1. Let's understand the law of $\{\ell_n\}_{1\leq n\leq N}$ under $\tilde{\mathbb{P}}_N\{\cdot|\mathscr{D}\}$. For any $\{\psi\}_{1\leq n\leq N}\subset\mathbb{R}$, we have that

$$\mathbb{E}_N\left[\exp\left[\sqrt{-1}\sum_{n=1}^N \psi_n\ell_n + N\theta L_N\right]\middle|\mathscr{D}\right]$$
$$= \prod_{n=1}^N\left\{\Delta_n\int_{r\in[0,1]}\exp\left[\left(\sqrt{-1}\psi_n+\theta\right)(1-r)\right]\wp_{N,n}(\nu_N(\mathsf{X}),dr) + 1 - \Delta_n\right\}.$$

Thus

$$\tilde{\mathbb{E}}_N\left[\exp\left[\sqrt{-1}\sum_{n=1}^N\psi_n\ell_n\right]\middle|\mathscr{D}\right] = \prod_{n=1}^N\left\{\Delta_n\int_{r\in[0,1]}\exp\left[\sqrt{-1}\psi_n(1-r)\right]\tilde{\wp}_{N,n}(\nu_N(\mathsf{X}),dr)+1-\Delta_n\right\}$$

where

$$\tilde{\wp}_{N,n}(D,A) \overset{\text{def}}{=} \frac{\int_{r\in[0,1]\cap A}\exp\left[\theta(1-r)\right]\wp_{N,n}(D,dr)}{\int_{r\in[0,1]}\exp\left[\theta(1-r)\right]\wp_{N,n}(D,dr)} \qquad A\in\mathscr{B}[0,1],\ D\in[0,1]$$

for all $N\in\mathbb{N}$ and $n\in\{1,2\dots N\}$. In other words, the recovery rates for the names which have defaulted are independent with laws given by the $\tilde{\wp}_{N,n}(\nu_N(\mathsf{X}),\cdot)$'s. In particular,

$$\tilde{\mathbb{E}}_N\left[L_N|\mathscr{D}\right] = \frac{1}{N}\sum_{n=1}^N\Delta_n\frac{\int_{r\in[0,1]}(1-r)\exp\left[\theta(1-r)\right]\wp_{N,n}(D,dr)}{\int_{r\in[0,1]}\exp\left[\theta(1-r)\right]\wp_{N,n}(D,dr)} = \Lambda'_{\nu_N}(\theta).$$

Secondly,

$$\tilde{\mathbb{E}}_N\left[\left|L_N-\Lambda_{\nu_N}(\theta)\right|^2\middle|\mathscr{D}\right]$$
$$=\frac{1}{N^2}\sum_{n=1}^N\left\{\int_{r\in[0,1]}(1-r)^2\tilde{\wp}_{N,n}(\nu_N(\mathsf{X}),dr) - \left(\int_{r\in[0,1]}(1-r)^2\tilde{\wp}_{N,n}(\nu_N(\mathsf{X}),dr)\right)^2\right\}\leq\frac{1}{N}.$$

In a similar way, we next need to understand the statistics of the defaults under $\tilde{\mathbb{P}}_N$. For $\{\psi\}_{1\leq n\leq N}\subset\mathbb{R}$,

$$\mathbb{E}_N\left[\exp\left[\sqrt{-1}\sum_{n=1}^N\left\{\psi_n\Delta_n+\phi(\mathsf{p}^{N,n})\Delta_n\right\}\right]\right] = \prod_{n=1}^N\left\{p^{N,n}\exp\left[\sqrt{-1}\psi_n+\phi(\mathsf{p}^{N,n})\right]+1-p^{N,n}\right\}.$$

Thus

$$\tilde{\mathbb{E}}_N\left[\exp\left[\sqrt{-1}\sum_{n=1}^N\psi_n\Delta_n\right]\right] = \prod_{n=1}^N\left\{\tilde{p}^{N,n}\exp\left[\sqrt{-1}\psi_n\right]+1-\tilde{p}^{N,n}\right\} \tag{22}$$

where

$$\tilde{p}^{N,n} = \frac{p^{N,n}e^{\phi(\mathsf{p}^{N,n})}}{p^{N,n}e^{\phi(\mathsf{p}^{N,n})}+1-p^{N,n}} = \lambda'_{p^{N,n}}(\phi(\mathsf{p}^{N,n})) = \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}^{N,n})$$

for all $N\in\mathbb{N}$ and $n\in\{1,2\dots N\}$. In other words, the defaults are independent with probabilities given by the $\tilde{p}^{N,n}$'s. Fix now $\Psi\in C(\mathsf{X})$. Then

$$\int_{\mathsf{p}\in\mathsf{X}}\Psi(\mathsf{p})\nu_N(d\mathsf{p}) - \int_{\mathsf{p}\in\mathsf{X}}\Psi(\mathsf{p})\nu^*(d\mathsf{p}) = \mathcal{E}_1^N+\mathcal{E}_2^N$$

where

$$\mathcal{E}_1^N \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N\left\{\Delta_n-\tilde{p}^{N,n}\right\}\Psi(\mathsf{p}^{N,n})$$

$$\mathcal{E}_2^N \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^N \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p}^{N,n})\Psi(\mathsf{p}^{N,n}) - \int_{\mathsf{p}\in\mathsf{X}} \Psi(\mathsf{p})\nu^*(d\mathsf{p}) = \int_{\mathsf{p}\in\mathsf{X}} \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}_N(d\mathsf{p}) - \int_{\mathsf{p}\in\mathsf{X}} \frac{d\nu^*}{d\mathsf{U}}(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p}).$$

From (22) we have that $\tilde{\mathbb{E}}_N[\mathcal{E}_1^N]=0$; we also have by independence that

$$\tilde{\mathbb{E}}_N\left[\left|\mathcal{E}_1^N\right|^2\right] \le \frac{\sup_{\mathsf{p}\in\mathsf{X}}|\Psi(\mathsf{p})|^2}{N}$$

The requirement that $\frac{d\nu^*}{d\mathsf{U}}\in C(\mathsf{X})$ ensures that $\lim_{N\to\infty}\mathcal{E}_2^N=0$. Combining things together, we have that

$$\lim_{N\to\infty}\tilde{\mathbb{E}}_N\left[\left|\int_{\mathsf{p}\in\mathsf{X}}\Psi(\mathsf{p})\nu_N(d\mathsf{p}) - \int_{\mathsf{p}\in\mathsf{X}}\Psi(\mathsf{p})\nu^*(d\mathsf{p})\right|\right]=0.$$

Since Ψ was an arbitrary element of $C(\mathsf{X})$ and X is Polish, we indeed have (see [Str93])

$$\lim_{N\to\infty}\tilde{\mathbb{P}}_N\{\nu_N\not\in\mathcal{O}_3\}=0.$$

Combining things together, we get the claimed lower bound. □

We can now prove the full lower bound

Proposition 6.2. *Let G be an open subset of $[0,1]\times\mathscr{M}_1(\mathsf{X})$. Then*

$$\underline{\lim}_{N\to\infty}\frac{1}{N}\ln\mathbb{P}_N\{Z_N\in G\}\ge -\inf_{z\in G} I(z)$$

□

Proof. Fix $z=(\ell,\nu)\in G$. If $I(z)<\infty$ and $\ell\in\text{ri dom}\,\Lambda_\nu^*$, then we get (18) from Proposition 6.1. If $I(z)=\infty$, we of course again get (18). Finally, assume that $\ell\in\text{dom}\,\Lambda^*_\nu\setminus\text{ri dom}\,\Lambda^*_\nu$. We use the fact that $\text{dom}\,\Lambda^*_\nu\subset\overline{\text{ri dom}\,\Lambda^*_\nu}$ and convexity of $\ell\mapsto\Lambda^*_\nu(\ell)$. Fix a relaxation parameter $\eta>0$. Then there is an $\ell'\in\text{ri dom}\,\Lambda^*_\nu$ such that $(\ell',\nu)\in G$ and $\Lambda^*_\nu(\ell')<\Lambda^*_\nu(\ell)+\eta$ (see [DZ98][Appendix A]). Using Proposition 6.1, we get that

$$\underline{\lim}_{N\to\infty}\frac{1}{N}\ln\mathbb{P}_N\{Z_N\in G\}\ge -I(\ell',\nu)\ge -I(z)-\eta.$$

Letting $\eta\searrow 0$, we again get (18). Letting z vary over G, we get the claim. □

7. Large Deviations Upper Bound

The heart of the upper bound is an exponential Chebychev inequality. We will mimic, as much as possible, the proof of the upper bound of Cramér's theorem. The main result of this section is

Proposition 7.1. *Fix any closed subset F of $[0,1]\times\mathscr{M}_1(\mathsf{X})$. Then*

$$\overline{\lim}_{N\to\infty}\frac{1}{N}\ln\mathbb{P}_N\{Z_N\in F\}\le -\inf_{z\in F} I(z).$$

□

Not surprisingly, we will first prove the bound for F compact; we will then show enough exponential tightness to get to the full claim.

Proposition 7.2. *Fix any compact subset F of $[0,1]\times\mathscr{M}_1(\mathsf{X})$. Then*

$$\overline{\lim}_{N\to\infty}\frac{1}{N}\ln\mathbb{P}_N\{Z_N\in F\}\le -\inf_{z\in F} I(z).$$

□

Proof. To begin, fix $s<\inf_{z\in F}I(z)$. Fix also a relaxation parameter $\eta>0$. For each $(\theta,\phi)\in\mathbb{R}\times C(\mathsf{X})$, define the set

$$\mathcal{O}_{(\theta,\phi)}\overset{\text{def}}{=}\left\{(\ell,\nu)\in[0,1]\times\mathscr{M}_1(\mathsf{X}):\theta\ell+\int_{\mathsf{p}\in\mathsf{X}}\phi(\mathsf{p})\nu(d\mathsf{p})-\Lambda_\nu(\theta)>s+\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right\}$$

(these open sets were used in the proof of Lemma 5.1).

Fix now a $z \in F$. By definition of I and Lemma 4.2, we see that there is a $(\theta_z, \phi_z) \in \mathbb{R} \times C(\mathsf{X})$ such that $z \in \mathcal{O}_{(\theta_z,\phi_z)}$. Since $(\ell, \nu) \mapsto \theta\ell + \int_{\mathsf{p}\in\mathsf{X}} \phi_z(\mathsf{p})\nu(d\mathsf{p}) - \Lambda_\nu(\theta)$ is continuous, there is an open neighborhood $\mathcal{O}^*_z$ of z such that $\mathcal{O}^*_z \subset \mathcal{O}_{(\theta_z,\phi_z)}$ and such that

$$\theta_z\tilde{\ell} + \int_{\mathsf{p}\in\mathsf{X}} \phi_z(\mathsf{p})\tilde{\nu}(d\mathsf{p}) - \Lambda_{\tilde{\nu}}(\theta_z) > s + \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_z(\mathsf{p}))\mathsf{U}(d\mathsf{p})$$

for all $(\tilde{\ell}, \tilde{\nu}) \in \mathcal{O}^*_z$. Thus

$$F \subset \bigcup_{z\in F} \mathcal{O}^*_z,$$

the compactness of F implies that we can extract a finite subset $\mathcal{Z}$ of F such that

$$F \subset \bigcup_{z\in\mathcal{Z}} \mathcal{O}^*_z$$

and thus

$$\mathbb{P}_N\left\{Z_N \in F\right\} \le \sum_{z\in\mathcal{Z}} \mathbb{P}_N\{Z_N \in \mathcal{O}^*_z\}.$$

Fix now $z \in \mathcal{Z}$. We have that

$$\begin{aligned}
\mathbb{P}_N\{Z_N \in \mathcal{O}^*_z\} &\le \mathbb{P}_N\left\{\theta_z L_N + \int_{\mathsf{p}\in\mathsf{X}} \phi_z(\mathsf{p})\nu_N(d\mathsf{p}) > s + \Lambda_{\nu_N}(\theta_z) + \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_z(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right\} \\
&\le e^{-Ns}\mathbb{E}_N\left[\exp\left[N\left\{\theta_z L_N - \Lambda_{\nu_N}(\theta_z)\right\}\right]\exp\left[N\left\{\int_{\mathsf{p}\in\mathsf{X}} \phi_z(\mathsf{p})\nu_N(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_z(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right\}\right]\right] \\
&= e^{-Ns}\exp\left[N\left\{\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_z(\mathsf{p}))\mathsf{U}_N(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_z(\mathsf{p}))\mathsf{U}(d\mathsf{p})\right\}\right]
\end{aligned}$$

We have used here the fact that

$$\mathbb{E}_N[\exp[\theta_z L_N - \Lambda_{\nu_N}(\theta_z)]|\mathscr{D}] = 1$$

and that

$$\mathbb{E}_N\left[\exp\left[N\int_{\mathsf{p}\in\mathsf{X}} \phi_z(\mathsf{p})\nu_N(d\mathsf{p})\right]\right] = \exp\left[N\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_z(\mathsf{p}))\mathsf{U}_N(d\mathsf{p})\right].$$

Letting $N \to \infty$, we get that

$$\overline{\lim_{N\to\infty}} \frac{1}{N}\ln \mathbb{P}_N\{Z_N \in \mathcal{O}^*_z\} \le -s.$$

This gives the claim. □

Let's next show that most of time ν_N is in a compact set.

Proposition 7.3 (Exponential Tightness)**.** *For each $L > 0$ there is a compact subset $\mathcal{K}_L$ of $\mathscr{M}_1(\mathsf{X})$ such that*

$$\overline{\lim_{N\to\infty}} \frac{1}{N}\ln \mathbb{P}_N\{\nu_n \notin \mathcal{K}_L\} \le -L.$$

□

Proof. First note that Assumption 3.3 implies that $\{\mathsf{U}_N\}_{N\in\mathbb{N}}$ is tight. Thus for each $j \in \mathbb{N}$, there is a compact subset K_j of X such that

$$\sup_{N\in\mathbb{N}} \mathsf{U}_N(\mathsf{X}\setminus K_j) \le \frac{1}{(L+j)^2}.$$

We will define

$$\mathcal{K}_L \overset{\text{def}}{=} \overline{\left\{\nu \in \mathscr{M}_1(\mathsf{X}) : \nu(\mathsf{X}\setminus K_j) \le \frac{1}{L+j} \text{ for all } j \in \mathbb{N}\right\}}.$$

Then $\mathcal{K}_L$ is compact, and we have that

$$\mathbb{P}_N\{\nu_N \notin \mathcal{K}_L\} \le \sum_{j=1}^{\infty} \mathbb{P}_N\left\{\nu_N(\mathsf{X}\setminus K_j) \ge \frac{1}{L+j}\right\}$$

$$\leq \sum_{j=1}^{\infty} \mathbb{P}_N \left\{ N(L+j)^2 \nu_N(\mathsf{X} \setminus K_j) \geq N(L+j) \right\}$$

$$\leq \sum_{j=1}^{\infty} \exp\left[-N(L+j)\right] \mathbb{E}_N \left[\exp\left[N(L+j)^2 \nu_N(\mathsf{X} \setminus K_j) \right] \right]$$

We now compute that

$$\mathbb{E}_N \left[\exp\left[(L+j)^2 \nu_N(\mathsf{X} \setminus K_j) \right] \right] = \prod_{n=1}^{N} \mathbb{E}_N \left[\exp\left[(L+j)^2 \sum_{n=1}^{N} \Delta_n \chi_{\mathsf{X}\setminus K_j}(\mathsf{p}^{N,n}) \right] \right]$$

$$= \exp\left[\sum_{n=1}^{N} \lambda_{p^{N,n}} \left((L+j)^2 \chi_{\mathsf{X}\setminus K_j}(\mathsf{p}^{N,n}) \right) \right] \leq \exp\left[N(L+j)^2 \mathsf{U}_N(\mathsf{X} \setminus K_j) \right] \leq e^N.$$

We have used here the calculation that for $\theta > 0$,

$$\lambda_p(\theta) \leq \ln\left(pe^{\theta} + (1-p)e^{\theta} \right) = \theta.$$

Combining things together, we get that

$$\mathbb{P}_N\{\nu_N \notin \mathcal{K}_L\} \leq \sum_{j=1}^{\infty} e^{-N(L+j)} e^N = e^{-NL} \sum_{j=1}^{\infty} e^{-N(j-1)} \leq e^{-NL} \sum_{j=1}^{\infty} e^{-(j-1)} = \frac{e^{-NL}}{1-e^{-1}}.$$

□

We can now get the full upper bound.

Proof of Proposition 7.1. Fix $L > s$. Then

$$\mathbb{P}_N\{Z_N \in F\} \leq \mathbb{P}_N\{Z_N \in F, \, \nu_N \in \mathcal{K}_L\} + \mathbb{P}_N\{\nu_N \notin \mathcal{K}_L\}.$$

We use Proposition 7.3 on the second term. Using Proposition 7.2 on the first term (and note that $[0,1] \times \mathcal{K}_L$ is compact), we get that

$$\overline{\lim_{N\to\infty}} \frac{1}{N} \ln \mathbb{P}_N\{Z_N \in F, \, \nu_N \in \mathcal{K}_L\} \leq - \inf_{\substack{z=(\ell,\nu)\in F \\ \nu \in \mathcal{K}_L}} I(z) \leq - \inf_{z \in F} I(z).$$

□

8. Alternative Expression for the Rate Function

In this section, we discuss the alternative expression for the rate function I' of Theorem 4.5 given by Theorem 4.6. In particular, this alternative representation shows that $I'(\ell)$ has a natural interpretation as the favored way to rearrange recoveries and losses among the different types. In addition to providing intuitive insight, this alternative expression suggests numerical schemes for computing the rate function. We will rigorously verify that the alternative expression is correct, but will be heuristic in our discussion of the numerical schemes.

We defer the proof of Theorem 4.6 to the end of this section and we first study the variational problem (13) using a Lagrange multiplier approach. Even though an explicit expression is usually not available, one can use numerical optimization techniques to calculate the quantities involved. In order to do that, we firstly recall that we can rewrite J' of (13) as a two-stage minimization problem, see expression (14).

This naturally suggests an analysis via a Lagrangian. Define

$$L(\Phi, \Psi, \lambda_1, \lambda_2) = \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \left\{ \Phi(p) I_{\wp}\left(\Psi(\mathsf{p}), D \right) + \hbar_p(\Phi(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p})$$

$$- \lambda_1 \left\{ \int_{\mathsf{p}\in \mathsf{X}} \Phi(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p}) - \ell \right\} - \lambda_2 \left\{ \int_{\mathsf{p}\in \mathsf{X}} \Phi(\mathsf{p})\mathsf{U}(d\mathsf{p}) - D \right\}.$$

Let's assume that Φ^* and Ψ^* are the minimizers. Let's also assume that $I_\wp(\cdot, D)$ is differentiable for all $\mathsf{p} = (p, \wp)$ in the support of U. We should then have that for every η_1 and η_2 in $\mathcal{B}$,

$$\int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \eta_1(\mathsf{p}) \left\{ I_\wp\left(\Psi^*(\mathsf{p}), D\right) + \hbar'_p(\Phi^*(\mathsf{p})) - \lambda_1 \Psi^*(\mathsf{p}) - \lambda_2 \right\} \mathsf{U}(d\mathsf{p}) = 0$$
$$\int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \eta_2(\mathsf{p}) \Phi^*(\mathsf{p}) \left\{ I'_\wp\left(\Psi^*(\mathsf{p}), D\right) - \lambda_1 \right\} \mathsf{U}(d\mathsf{p}) = 0$$

Ignoring any complications which would arise on the set where $\Phi^* = 0$, we should then have that

$$I_\wp\left(\Psi^*(\mathsf{p}), D\right) + \hbar'_p(\Phi^*(\mathsf{p})) = \lambda_1 \Psi^*(\mathsf{p}) + \lambda_2$$
$$I'_\wp\left(\Psi^*(\mathsf{p}), D\right) = \lambda_1$$

for all $\mathsf{p} = (p, \wp) \in \mathsf{X}$. This is a triangular system; the first equation depends on both λ_1 and λ_2, but the second depends only on λ_1. Recalling now (4) and the structure of Legendre-Fenchel transforms, we should have that

$$M'_\wp(\lambda_1, D) = \Psi^*(\mathsf{p})$$
$$\hbar'_p(\Phi^*(\mathsf{p})) = \lambda_2 + \lambda_1 M'_\wp(\lambda_1, D) - I_\wp(\Psi^*(\mathsf{p}), D) = \lambda_2 + M_\wp(\lambda_1, D)$$

for all $\mathsf{p} = (p, \wp) \in \mathsf{X}$. We can then invert this. This leads us to the following. Define

$$\Phi_{\lambda_1,\lambda_2,D}(p, \wp) \overset{\text{def}}{=} \frac{p e^{\lambda_2 + M_\wp(\lambda_1, D)}}{1 - p + p e^{\lambda_2 + M_\wp(\lambda_1, D)}} \qquad \lambda_1, \lambda_2 \in \mathbb{R},\ (p, \wp) \in \mathsf{X}$$
$$\Psi_{\lambda_1,D}(\wp) \overset{\text{def}}{=} M'_\wp(\lambda_1, D) \qquad \lambda_1 \in \mathbb{R},\ \wp \in C([0,1]; \mathscr{P}[0,1])$$

where $(\lambda_1, \lambda_2) = (\lambda_1(\ell, D, \mathsf{U}), \lambda_2(\ell, D, \mathsf{U}))$ is such that

$$\int_{\mathsf{p}\in \mathsf{X}} \Phi_{\lambda_1,\lambda_2,D}(\mathsf{p}) \mathsf{U}(d\mathsf{p}) = D$$
$$\int_{\mathsf{p}\in \mathsf{X}} \Phi_{\lambda_1,\lambda_2,D}(\mathsf{p}) \Psi_{\lambda_1,\lambda_2,D}(\mathsf{p}) \mathsf{U}(d\mathsf{p}) = \ell.$$

We conclude this section with the rigorous proof of the alternate representation.

Proof of Theorem 4.6. First, we prove that $J'(\ell) \geq I'(\ell)$. Consider any Φ and $\Psi \in \mathcal{B}$ such that

$$\int_{\mathsf{p}\in \mathsf{X}} \Phi(\mathsf{p}) \Psi(\mathsf{p}) \mathsf{U}(d\mathsf{p}) = \ell. \tag{23}$$

For any $\theta \in \mathbb{R}$,

$$\begin{aligned}
&\int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \left\{ \Phi(\mathsf{p}) I_\wp\left(\Psi(\mathsf{p}), \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \Phi(\mathsf{p}) \mathsf{U}(d\mathsf{p})\right) + \hbar_p(\Phi(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p}) \\
&\quad = \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \left\{ \Phi(\mathsf{p}) \sup_{\theta'\in\mathbb{R}} \left\{ \theta' \Psi(\mathsf{p}) - M_\wp\left(\theta', \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \Phi(\mathsf{p}) \mathsf{U}(d\mathsf{p})\right) \right\} + \hbar_p(\Phi(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p}) \\
&\quad \geq \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \left\{ \Phi(\mathsf{p}) \left\{ \theta \Psi(\mathsf{p}) - M_\wp\left(\theta, \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \Phi(\mathsf{p}) \mathsf{U}(d\mathsf{p})\right) \right\} + \hbar_p(\Phi(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p}) \\
&\quad = \theta\ell - \int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \left\{ \Phi(\mathsf{p}) M_\wp\left(\theta, \int_{\mathsf{p}\in \mathsf{X}} \Phi(\mathsf{p}) \mathsf{U}(d\mathsf{p})\right) + \hbar_p(\Phi(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p}).
\end{aligned}$$

Define $\nu \in \mathscr{M}_1(\mathsf{X})$ as

$$\nu(A) \overset{\text{def}}{=} \int_{\mathsf{p}\in A} \Phi(\mathsf{p}) \mathsf{U}(d\mathsf{p}); \qquad A \in \mathscr{B}(\mathsf{X})$$

then

$$\int_{\mathsf{p}=(p,\wp)\in \mathsf{X}} \left\{ \Phi(\mathsf{p}) I_\wp\left(\Psi(\mathsf{p}), \int_{\mathsf{p}\in \mathsf{X}} \Phi(\mathsf{p}) \mathsf{U}(d\mathsf{p})\right) + \hbar_p(\Phi(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p})$$

$$\geq \theta\ell - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} M_\wp(\theta,\nu(\mathsf{X}))\nu(d\mathsf{p}) + \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p}).$$

Varying θ, we get that

$$\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left\{\Phi(\mathsf{p})I_\wp\left(\Psi(\mathsf{p}),\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\Phi(\mathsf{p})\mathsf{U}(d\mathsf{p})\right)+\hbar_p(\Phi(\mathsf{p}))\right\}\mathsf{U}(d\mathsf{p}) \geq \Lambda^*_\nu(\ell)+H(\nu)\geq I'(\ell)$$

and then varying Φ and Ψ in $\mathcal{B}$ (such that (23) holds), we get that $J'(\ell)\geq I'(\ell)$.

To show that $I'(\ell)\geq J'(\ell)$, fix $\nu\in\mathscr{M}_1(\mathsf{X})$ such that that $\nu\ll\mathsf{U}$. We want to show that

$$\Lambda^*_\nu(\ell)+H(\nu)\geq J'(\ell). \tag{24}$$

If $\nu\not\ll\mathsf{U}$, this is trivially true, so we can assume that $\nu\ll\mathsf{U}$. For all $\wp\in C([0,1];\mathscr{P}[0,1])$, define

$$\alpha_-(\wp,D)\stackrel{\text{def}}{=}\inf\{1-r\in[0,1]: r\in\operatorname{supp}\wp(D,\cdot)\}$$
$$\alpha_+(\wp,D)\stackrel{\text{def}}{=}\sup\{1-r\in[0,1]: r\in\operatorname{supp}\wp(D,\cdot)\}.$$

Dominated convergence implies that

$$\lim_{\theta\to-\infty}\Lambda'_\nu(\theta)=\bar\alpha_-\stackrel{\text{def}}{=}\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\alpha_-(\wp,\nu(\mathsf{X}))\nu(d\mathsf{p})$$
$$\lim_{\theta\to\infty}\Lambda'_\nu(\theta)=\bar\alpha_+\stackrel{\text{def}}{=}\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\alpha_+(\wp,\nu(\mathsf{X}))\nu(d\mathsf{p}).$$

From (12) and the monotonicity of moment generating functions, we can see that Λ_ν is nondecreasing; thus $(\bar\alpha_-,\bar\alpha_+)\in\Lambda'_\nu(\mathbb{R})$. This leads to three possible cases.

Case 1: Assume that $\ell\in(\bar\alpha_-,\bar\alpha_+)$, and let $\theta^*\in\mathbb{R}$ be such that $\Lambda'_\nu(\theta^*)=\ell$; i.e.,

$$\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} M'_\wp(\theta^*,\nu(\mathsf{X}))\nu(d\mathsf{p})=\ell \tag{25}$$

Then

$$\begin{aligned}\Lambda^*_\nu(\ell)+H(\nu)&=\sup_{\theta\in\mathbb{R}}\left\{\theta\ell-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}M_\wp(\theta,\nu(\mathsf{X}))\nu(d\mathsf{p})\right\}+\int_{\mathsf{p}=(p,\wp)\in\mathsf{U}}\hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p})\\
&\geq\theta^*\ell-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}M_\wp(\theta^*,\nu(\mathsf{X}))\nu(d\mathsf{p})+\int_{\mathsf{p}=(p,\wp)\in\mathsf{U}}\hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p})\\
&=\theta^*\Lambda'_\nu(\theta^*)-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}M_\wp(\theta^*,\nu(\mathsf{X}))\nu(d\mathsf{p})+\int_{\mathsf{p}=(p,\wp)\in\mathsf{U}}\hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p})\\
&=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\left(\theta^*M'_\wp(\theta^*,\nu(\mathsf{X}))-M_\wp(\theta^*,\nu(\mathsf{X}))\right)\nu(d\mathsf{p})+\int_{\mathsf{p}=(p,\wp)\in\mathsf{U}}\hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p})\\
&=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}I_\wp(M'_\wp(\theta^*,\nu(\mathsf{X})),\nu(\mathsf{X}))\nu(d\mathsf{p})+\int_{\mathsf{p}=(p,\wp)\in\mathsf{U}}\hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)\mathsf{U}(d\mathsf{p}).\end{aligned}$$

Define now $\Phi(\mathsf{p})\stackrel{\text{def}}{=}\frac{d\nu}{d\mathsf{U}}(\mathsf{p})$ and $\Psi(\mathsf{p})\stackrel{\text{def}}{=}M'_\wp(\theta^*,\nu(\mathsf{X}))$. Then (25) is exactly that $\int_{\mathsf{p}\in\mathsf{X}}\Phi(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p})=\ell$. Thus

$$\Lambda^*_\nu(\ell)+H(\nu)\geq\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\left[\Phi(\mathsf{p})I_\wp\left(\Psi(\mathsf{p}),\int_{\mathsf{p}\in\mathsf{X}}\Phi(\mathsf{p})\mathsf{U}(d\mathsf{p})\right)+\hbar_p(\Phi(\mathsf{p}))\right]\mathsf{U}(d\mathsf{p})\geq J'(\ell).$$

Case 2: Assume next that $\ell\in[\bar\alpha_+,1]$. For every $\wp\in C([0,1];\mathscr{P}[0,1])$, define

$$\mathcal{E}^\wp_+(\theta)\stackrel{\text{def}}{=}M_\wp(\theta,\nu(\mathsf{X})-\theta\alpha_+(\wp,\nu(\mathsf{X}))=\ln\int_{r\in[0,1]}e^{-\theta(\alpha_+(\wp,\nu(\mathsf{X}))-(1-r))}\wp(\nu(\mathsf{X}),dr).$$

for all $\theta\in\mathbb{R}$; thus

$$M_\wp(\theta,\nu(\mathsf{X}))=\theta\alpha_+(\wp,\nu(\mathsf{X}))+\mathcal{E}^\wp_+(\theta)\qquad\wp\in C([0,1];\mathscr{P}[0,1])$$
$$\Lambda_\nu(\theta)=\theta\bar\alpha_++\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}^\wp_+(\theta)\nu(d\mathsf{p})$$

for all $\theta\in\mathbb{R}$. For all $\wp\in C([0,1];\mathscr{P}[0,1])$ and $(1-r)\in\operatorname{supp}\wp(\nu(\mathsf{X}),\cdot)$, the mapping $\theta\mapsto e^{-\theta(\alpha_+(\wp,\nu(\mathsf{X}))-(1-r))}$ is decreasing and maps $[0,\infty)$ into $(0,1]$. Monotone convergence implies that

$$\lim_{\theta\to\infty}\mathcal{E}_+^{\wp}(\theta)=\ln\wp\{1-\alpha_+(\wp,\nu(\mathsf{X}))\}.$$

If $\ell>\bar\alpha_+$, then we can use the fact that $\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_+^{\wp}(\theta)\nu(d\mathsf{p})\le 0$ for all $\theta>0$ to see that

$$\Lambda_\nu^*(\ell)\ge\varlimsup_{\theta\to\infty}\left\{\theta(\ell-\bar\alpha_+)-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_+^{\wp}(\theta)\nu(d\mathsf{p})\right\}\ge\varlimsup_{\theta\to\infty}\{\theta(\ell-\bar\alpha_+)\}=\infty\ge J'(\ell).$$

If $\ell=\bar\alpha_+$, then by the monotonicity of the $\mathcal{E}_+^{\wp}$'s,

$$I_\wp(\alpha_+(\wp,\nu(\mathsf{X}),\nu(\mathsf{X})))=\sup_{\theta\in\mathbb{R}}\left\{-\mathcal{E}_+^{\wp}(\theta)\right\}=\varlimsup_{\theta\to\infty}\left\{-\mathcal{E}_+^{\wp}(\theta)\right\}=\ln\left(\frac{1}{\wp\{1-\alpha_+(\wp,\nu(\mathsf{X}))\}}\right)\tag{26}$$

for all $\wp\in C([0,1];\mathscr{P}[0,1])$, and

$$\begin{aligned}\Lambda_\nu^*(\ell)&=\sup_{\theta\in\mathbb{R}}\left\{-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_+^{\wp}(\theta)\nu(d\mathsf{p})\right\}=\varlimsup_{\theta\to\infty}\left\{-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_+^{\wp}(\theta)\nu(d\mathsf{p})\right\}\\&=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\ln\left(\frac{1}{\wp\{1-\alpha_+(\wp,\nu(\mathsf{X}))\}}\right)\nu(d\mathsf{p})=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}I_\wp(\alpha_+(\wp,\nu(\mathsf{X})))\nu(d\mathsf{p}).\end{aligned}$$

Defining $\Phi(\mathsf{p})\overset{\text{def}}{=}\frac{d\nu}{d\mathsf{U}}(\mathsf{p})$ and $\Psi(\mathsf{p})=\alpha_+(\wp,\nu(\mathsf{X}))$, we have that

$$\int_{\mathsf{p}\in\mathsf{X}}\Phi(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p})=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\alpha_+(\wp,\nu(\mathsf{X}))\nu(d\mathsf{p})=\ell.$$

Collecting things together, we see that if $\ell=\bar\alpha_+$, we again get (24).

Case 3: We finally assume that $\ell\in[0,\bar\alpha_-]$. The calculations are very similar to those of Case 2. For every $\wp\in C([0,1];\mathscr{P}[0,1])$, define

$$\mathcal{E}_-^{\wp}(\theta)\overset{\text{def}}{=}M_\wp(\theta,\nu(\mathsf{X})-\theta\alpha_-(\wp,\nu(\mathsf{X}))=\ln\int_{r\in[0,1]}e^{\theta((1-r)-\alpha_-(\wp,\nu(\mathsf{X})))}\wp(\nu(\mathsf{X}),dr).$$

for all $\theta\in\mathbb{R}$, so that

$$M_\wp(\theta,\nu(\mathsf{X}))=\theta\alpha_-(\wp,\nu(\mathsf{X}))+\mathcal{E}_-^{\wp}(\theta)\wp\in C([0,1];\mathscr{P}[0,1])$$
$$\Lambda_\nu(\theta)=\theta\bar\alpha_-+\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_-^{\wp}(\theta)\nu(d\mathsf{p})$$

for all $\theta\in\mathbb{R}$. For all $\wp\in C([0,1];\mathscr{P}[0,1])$ and $(1-r)\in\operatorname{supp}\wp(\nu(\mathsf{X}),\cdot)$, the mapping $\theta\mapsto e^{\theta((1-r)-\alpha_-(\wp,\nu(\mathsf{X})))}$ is increasing and maps $(-\infty,0]$ into $(0,1]$. Monotone convergence implies that

$$\lim_{\theta\to-\infty}\mathcal{E}_-^{\wp}(\theta)=\ln\wp\{1-\alpha_-(\wp,\nu(\mathsf{X}))\}.$$

If $\ell<\bar\alpha_-$, then we can use the fact that $\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_-^{\wp}(\theta)\nu(d\mathsf{p})\ge 0$ for all $\theta<0$ to see that

$$\Lambda_\nu^*(\ell)\ge\varlimsup_{\theta\to-\infty}\left\{\theta(\ell-\bar\alpha_-)-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_-^{\wp}(\theta)\nu(d\mathsf{p})\right\}\ge\varlimsup_{\theta\to-\infty}\{\theta(\ell-\bar\alpha_-)\}=\infty\ge J'(\ell).$$

If $\ell=\bar\alpha_-$, then by the monotonicity of the $\mathcal{E}_-^{\wp}$'s,

$$I_\wp(\alpha_-(\wp,\nu(\mathsf{X}),\nu(\mathsf{X})))=\sup_{\theta\in\mathbb{R}}\left\{-\mathcal{E}_-^{\wp}(\theta)\right\}=\lim_{\theta\to-\infty}\left\{-\mathcal{E}_-^{\wp}(\theta)\right\}=\ln\left(\frac{1}{\wp\{1-\alpha_-(\wp,\nu(\mathsf{X}))\}}\right)$$

for all $\wp\in C([0,1];\mathscr{P}[0,1])$, and

$$\begin{aligned}\Lambda_\nu^*(\ell)&=\sup_{\theta\in\mathbb{R}}\left\{-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_-^{\wp}(\theta)\nu(d\mathsf{p})\right\}=\lim_{\theta\to-\infty}\left\{-\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\mathcal{E}_-^{\wp}(\theta)\nu(d\mathsf{p})\right\}\\&=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}\ln\left(\frac{1}{\wp\{1-\alpha_-(\wp,\nu(\mathsf{X}))\}}\right)\nu(d\mathsf{p})=\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}}I_\wp(\alpha_-(\wp,\nu(\mathsf{X})))\nu(d\mathsf{p}).\end{aligned}$$

Defining $\Phi(\mathsf{p}) \stackrel{\text{def}}{=} \frac{d\nu}{d\mathsf{U}}(\mathsf{p})$ and $\Psi(\mathsf{p}) = \alpha_-(\wp, \nu(\mathsf{X}))$, we have that

$$\int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\Psi(\mathsf{p})\mathsf{U}(d\mathsf{p}) = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \alpha_-(\wp,\nu(\mathsf{X}))\nu(d\mathsf{p}) = \ell$$

again implying (24).

Collecting things together, we have (13). We get (14) by definng $D \stackrel{\text{def}}{=} \int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\mathsf{U}(d\mathsf{p})$. Note that since Φ and Ψ both take values in $[0,1]$,

$$\int_{\mathsf{p}\in\mathsf{X}} \Psi(\mathsf{p})\Phi(\mathsf{p})\mathsf{U}(d\mathsf{p}) \le \int_{\mathsf{p}\in\mathsf{X}} \Phi(\mathsf{p})\mathsf{U}(d\mathsf{p}).$$

This allows us to restrict the minimization in D to the interval $[\ell, 1]$. □

9. Detailed structure of H

We here want to understand some of the detailed behavior of H more clearly. Specifically, we want to prove Lemmas 4.3 and 4.2.

Fix $\nu \in \mathscr{M}_1(\mathsf{X})$ such that $H(\nu) < \infty$. The main technical challenges in both proofs is to stay away from the singularities in $\hbar_p$ and $\hbar_p'$. Note that

$$\hbar_p'(x) = \ln\left(\frac{x}{1-x}\frac{1-p}{p}\right), \qquad x, p \in (0,1)$$

and keeping (6) in mind, we thus need to be careful near $p \in \{0,1\}$, and for $(x,p) \in \{0,1\} \times (0,1)$.

Fix now $\nu \in \mathscr{M}_1(\mathsf{X})$ such that $H(\nu) < \infty$. To start, let's note some implications of the assumption that $H(\nu) < \infty$. Clearly $\nu \ll \mathsf{U}$. Secondly,

$$\mathsf{U}\left\{\mathsf{p} \in \mathsf{X} : \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) > 1\right\} = 0. \tag{27}$$

Let's now do the following. Fix $N \in \mathbb{N}$. Define

$$\xi_N(\mathsf{p}) \stackrel{\text{def}}{=} \begin{cases} p & \text{if } p \notin \left[\frac{1}{N}, 1-\frac{1}{N}\right] \\ \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) & \text{if } p \in \left(\frac{1}{N}, 1-\frac{1}{N}\right) \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \in \left(\frac{1}{N}, 1-\frac{1}{N}\right) \\ \frac{1}{N} & \text{if } p \in \left(\frac{1}{N}, 1-\frac{1}{N}\right) \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \le \frac{1}{N} \\ 1-\frac{1}{N} & \text{if } p \in \left(\frac{1}{N}, 1-\frac{1}{N}\right) \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \ge 1-\frac{1}{N} \end{cases} \tag{28}$$

Clearly $0 \le \xi_N \le 1$, so we can define $\nu_N \in \mathscr{M}_1(\mathsf{X})$ as

$$\nu_N(A) \stackrel{\text{def}}{=} \int_{\mathsf{p}\in A} \xi_N(\mathsf{p})\mathsf{U}(d\mathsf{p}). \qquad A \in \mathscr{B}(\mathsf{X})$$

In light of (27) and (7), $\lim_{N\to\infty} \xi_N = \frac{d\nu}{d\mathsf{U}}$ U-a.s., so it follows that $\lim_{N\to\infty} \nu_N = \nu$. We next compute that

$$\hbar_p(\xi_N(\mathsf{p})) = \begin{cases} 0 & \text{if } p \notin \left[\frac{1}{N}, 1-\frac{1}{N}\right] \\ \hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right) & \text{if } \frac{1}{N} \le p \le 1-\frac{1}{N} \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \in \left(\frac{1}{N}, 1-\frac{1}{N}\right) \\ \hbar_p\left(\frac{1}{N}\right) & \text{if } \frac{1}{N} \le p \le 1-\frac{1}{N} \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \le \frac{1}{N} \\ \hbar_p\left(1-\frac{1}{N}\right) & \text{if } \frac{1}{N} \le p \le 1-\frac{1}{N} \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \ge 1-\frac{1}{N} \end{cases}$$

Using again (27) and (7), we have that $\lim_{N\to\infty} \hbar_p(\xi_N(\mathsf{p})) = \hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)$ for U-almost-all $\mathsf{p} = (p,\wp) \in \mathsf{X}$. If $p \in \left[\frac{1}{N}, 1-\frac{1}{N}\right]$, then $\hbar_p$ is increasing on $[p,1] \supset \left[1-\frac{1}{N}, 1\right]$ and decreasing on $[0,p] \supset \left[0, \frac{1}{N}\right]$. Thus $\hbar_p(\xi_N(\mathsf{p})) \le \hbar_p\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right)$ for U-almost-all $\mathsf{p} = (p,\wp) \in \mathsf{X}$. Dominated convergence thus implies that $\lim_{N\to\infty} H(\nu_N) = H(\nu)$.

Proof of Lemma 4.3. Fix $N \in N$; we want to approximate ξ_N by "nice" elements of $C(\mathsf{X})$. Note that

$$\xi_N(\mathsf{p}) = p\chi_{[0,1]\setminus[N^{-1},1-N^{-1}]}(p) + \chi_{[N^{-1},1-N^{-1}]}(p)\xi_N(\mathsf{p}).$$

Since U is regular (recall that X is Polish), we can approximate $\mathsf{p} = (p,\wp) \mapsto \chi_{[N^{-1},1-N^{-1}]}(p)\xi_N(\mathsf{p})$ by elements of $C(\mathsf{X})$. From (28), we have that $N^{-1} \le \xi_N(\mathsf{p}) \le 1-N^{-1}$ if $\mathsf{p} = (p,\wp) \in \mathsf{X}$ is such that

$N^{-1} \le p \le 1-N^{-1}$, so we can truncate these approximations at N^{-1} and $1-N^{-1}$ without any loss. Namely, there is a sequence $(\tilde{\xi}^1_\varepsilon)_{\varepsilon>0}$ in $C(\mathsf{X})$ such that

$$N^{-1} \le \tilde{\xi}^1_\varepsilon \le 1-N^{-1}$$
$$\lim_{\varepsilon\searrow 0}\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left|\tilde{\xi}^1_\varepsilon(\mathsf{p}) - \chi_{[N^{-1},1-N^{-1}]}(p)\xi_N(\mathsf{p})\right| \mathsf{U}(d\mathsf{p}) = 0. \tag{29}$$

For each $\varepsilon>0$, let $\varphi_\varepsilon \in C(\mathbb{R};[0,1])$ be such that $\varphi_\varepsilon(u)=1$ if $u\in[N^{-1},1-N^{-1}]$ and $\varphi_\varepsilon(u)=0$ if $u\in[0,1]\setminus[N^{-1}-\varepsilon,1-N^{-1}+\varepsilon]$. For each $\varepsilon>0$, define

$$\tilde{\xi}^2_\varepsilon(\mathsf{p}) \overset{\text{def}}{=} p\left\{1-\varphi_\varepsilon(p)\right\} + \tilde{\xi}^1_\varepsilon(\mathsf{p})\varphi_\varepsilon(p)$$

for all $\mathsf{p}=(p,\wp)\in\mathsf{X}$. Then $\tilde{\xi}^2_\varepsilon\in C(\mathsf{X})$ for all $\varepsilon>0$. We also have that

$$\begin{aligned}\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} &\left|\tilde{\xi}^2_\varepsilon(\mathsf{p}) - \xi_N(\mathsf{p})\right| \mathsf{U}(d\mathsf{p}) \\ &\le \mathsf{U}\left\{\mathsf{p}=(p,\wp)\in\mathsf{X} : p\in[N^{-1}+\varepsilon,1-N^{-1}+\varepsilon]\setminus[N^{-1},1-N^{-1}]\right\} \\ &\qquad + \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \chi_{[N^{-1},1-N^{-1}]}(p)\left|\tilde{\xi}^1_\varepsilon(\mathsf{p}) - \xi_N(\mathsf{p})\right| \mathsf{U}(d\mathsf{p}).\end{aligned}$$

Dominated convergence and (29) then ensure that

$$\lim_{\varepsilon\to 0}\int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left|\tilde{\xi}^2_\varepsilon(\mathsf{p}) - \xi_N(\mathsf{p})\right| \mathsf{U}(d\mathsf{p}) = 0. \tag{30}$$

Clearly $0\le\tilde{\xi}^2_\varepsilon\le 1$, so we can define $\nu_{N,\varepsilon}\in\mathscr{M}_1(\mathsf{X})$ as

$$\nu_{N,\varepsilon}(A) \overset{\text{def}}{=} \int_{\mathsf{p}\in A} \tilde{\xi}^2_\varepsilon(\mathsf{p})\mathsf{U}(d\mathsf{p}). \qquad A\in\mathscr{B}(\mathsf{X})$$

Thanks to (30), we have that $\lim_{\varepsilon\to 0}\nu_{N,\varepsilon}=\nu_N$. Note next that for $\mathsf{p}=(p,\wp)\in\mathsf{X}$ such that $p\in[0,1]\setminus[N^{-1}-\varepsilon,1-N^{-1}+\varepsilon]$,

$$\hbar_p(\tilde{\xi}^2_\varepsilon(\mathsf{p})) - \hbar_p(\xi_N(\mathsf{p})) = \hbar_p(p)-\hbar_p(p) = 0.$$

If $\mathsf{p}=(p,\wp)\in\mathsf{X}$ is such that $p\in[N^{-1}-\varepsilon,1-N^{-1}+\varepsilon]$, then

$$N^{-1}-\varepsilon \le \tilde{\xi}^2_\varepsilon(\mathsf{p}) \le 1-N^{-1}+\varepsilon, \tag{31}$$

so if $\varepsilon<1/(2N)$,

$$\left|\hbar_p(\tilde{\xi}^2_\varepsilon(\mathsf{p})) - \hbar_p(\xi_N(\mathsf{p}))\right| \le \varkappa\left|\tilde{\xi}^2_\varepsilon(\mathsf{p}) - \xi_N(\mathsf{p})\right|$$

where

$$\varkappa \overset{\text{def}}{=} \sup\left\{|\hbar_p(x)| : \frac{1}{2N}\le x\le 1-\frac{1}{2N} \text{ and } \frac{1}{2N}\le p\le 1-\frac{1}{2N}\right\}.$$

Thus if $\varepsilon<\frac{1}{2N}$,

$$\left|\hbar_p(\tilde{\xi}^2_\varepsilon(\mathsf{p})) - \hbar_p(\xi_N(\mathsf{p}))\right| \le \varkappa\left|\tilde{\xi}^2_\varepsilon(\mathsf{p}) - \xi_N(\mathsf{p})\right|$$

for all $\mathsf{p}=(p,\wp)\in\mathsf{X}$. Thanks to (30), we thus have that $\lim_{\varepsilon\to 0}H(\nu_{N,\varepsilon})=H(\nu)$.

We finally note that $\mathsf{p}=(p,\wp)\mapsto\hbar'_p(\tilde{\xi}^2_\varepsilon(\mathsf{p}))$ is continuous on $\{\mathsf{p}=(p,\wp)\in\mathsf{X} : p\in(N^{-1}-\varepsilon,1-N^{-1}+\varepsilon)\}$ ((31) ensures that $\tilde{\xi}_2$ takes values in $(0,1)$ in this case). On $\{\mathsf{p}=(p,\wp)\in\mathsf{X} : p\in(0,1)\setminus(N^{-1}-\varepsilon,1-N^{-1}+\varepsilon)\}$, we have that $\hbar'_p(\tilde{\xi}^2_\varepsilon(\mathsf{p}))=\hbar'_p(p)=0$. This finishes the proof. □

Proof of Lemma 4.2. Assume first that ν is not absolutely continuous with respect to U. Then there is an $A\in\mathscr{B}(\mathsf{X})$ such that $\nu(A)>0$ and $\mathsf{U}(A)=0$. Since X is Polish, ν is regular; i.e.,

$$\nu(A) = \sup\left\{\nu(F) : F\subset A,\ F \text{ closed}\right\}.$$

Thus there is a closed subset F of A such that $\nu(F)>0$. Fix also now $c>0$. For each $n\in\mathbb{N}$, define

$$\varphi_n(\mathsf{p}) \overset{\text{def}}{=} c\exp\left[-n\,\mathrm{dist}(\mathsf{p},F)\right] \qquad \mathsf{p}\in\mathsf{X}$$

where $\operatorname{dist}(\mathsf{p}, F)$ is the distance (in X) from x to F. Then $0 \le \varphi_n \le c$ for all $n \in \mathbb{N}$, and $\varphi_n \searrow c\chi_F$. Since $\theta \mapsto \lambda_p(\theta)$ is nondecreasing and continuous for each $p \in [0,1]$, we also have that $\lambda_p(\varphi_n(\mathsf{p})) \searrow \lambda_p(c\chi_F(\mathsf{p}))$ for all $\mathsf{p} = (p, \wp) \in \mathsf{X}$. Thus

$$\sup_{\phi \in C(\mathsf{X})} \left\{ \int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu(dp) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p}) \right\}$$
$$\ge \varlimsup_{n\to\infty} \left\{ \int_{\mathsf{p}\in\mathsf{X}} \phi_n(\mathsf{p})\nu(dp) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi_n(\mathsf{p}))\mathsf{U}(dp) \right\} = c\nu(F).$$

Let $c \nearrow \infty$ to see that the right-hand side of (8) is infinite.

Assume next that $\nu \ll \mathsf{U}$. We use the fact that $\hbar_p$ and λ_p are convex duals of each other. For any $\phi \in C(\mathsf{X})$,

$$\int_{\mathsf{p}\in\mathsf{X}} \phi(\mathsf{p})\nu(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p})$$
$$= \int_{\mathsf{p}\in\mathsf{X}} \inf_{x\in\mathbb{R}} \left\{ \phi(\mathsf{p}) \left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p}) - x \right) + \hbar_p(x) \right\} \mathsf{U}(d\mathsf{p}) \le \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \hbar_p \left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \right) \mathsf{U}(d\mathsf{p}).$$

To show the reverse inequality, let's write that

$$H(\nu) = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \sup_{\theta\in\mathbb{R}} \left\{ \theta \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) - \lambda_p(\theta) \right\} \mathsf{U}(d\mathsf{p}) = \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lim_{N\to\infty} F_N(\mathsf{p})\mathsf{U}(d\mathsf{p})$$

where

$$F_N(\mathsf{p}) = \sup_{|\theta|\le N} \left\{ \theta \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) - \lambda_p(\theta) \right\} \tag{32}$$

for all $N \in \mathbb{N}$ and $\mathsf{p} = (p, \wp) \in \mathsf{X}$. We can explicitly solve this minimization problem; for $N \in \mathbb{N}$ and $\mathsf{p} = (p, \wp) \in \mathsf{X}$, define

$$\phi_N(\mathsf{p}) = \begin{cases} \hbar_p'\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right) & \text{if } p \in (0,1),\ \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) \in (0,1), \text{ and } -N \le \hbar_p'\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p})\right) \le N \\ 0 & \text{if } p \in \{0,1\} \text{ and } \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) = p \\ N \operatorname{sgn}\left(\frac{d\nu}{d\mathsf{U}}(\mathsf{p}) - p\right) & \text{else} \end{cases}$$

(where sgn is the standard signum function). Then

$$F_N(\mathsf{p}) = \phi_N(\mathsf{p}) \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) - \lambda_p(\phi_N(\mathsf{p}))$$

for all $\mathsf{p} = (p, \wp) \in \mathsf{X}$ and $N \in \mathbb{N}$. Clearly F_N and ϕ_N are measurable, and $\phi_N \in B(\mathsf{X})$. From (32), we also see that F_N is nondecreasing in N. Thus by monotone convergence

$$H(\nu) = \lim_{N\to\infty} \int_{\mathsf{p}\in\mathsf{X}} F_N(\mathsf{p})\mathsf{U}(d\mathsf{p}) = \lim_{N\to\infty} \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \left\{ \phi_N(\mathsf{p}) \frac{d\nu}{d\mathsf{U}}(\mathsf{p}) - \lambda_p(\phi_N(\mathsf{p})) \right\} \mathsf{U}(d\mathsf{p})$$
$$\le \sup_{\varphi\in B(\mathsf{X})} \left\{ \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \phi(\mathsf{p})\nu(d\mathsf{p}) - \int_{\mathsf{p}=(p,\wp)\in\mathsf{X}} \lambda_p(\phi(\mathsf{p}))\mathsf{U}(d\mathsf{p}) \right\}.$$

Since X is Polish, ν and U are regular; and thus we can approximate elements of $B(\mathsf{X})$ by elements of $C(\mathsf{X})$, completing the proof. □

References

[ABRS05] Edward I. Altman, Brooks Brady, Andrea Resti, and Andrea Sironi. The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications. *The Journal of Business*, vol. 78, no. 6, 2203–2227, 2005.

[DDD04] Amir Dembo, Jean-Dominique Deuschel, and Darrell Duffie. Large portfolio losses. *Finance Stoch.*, 8(1):3–16, 2004.

[DZ98] Amir Dembo and Ofer Zeitouni. Large deviations techniques and applications. Springer-Verlag, New York, 1998.

[LMS09] Vincent Leijdekker, Michel Mandjes, and Peter Spreij. Sample-path large deviations in credit risk, `arXiv:0909.5610v1 [math.PR]`, 2009.

[Pha07] Huyên Pham. Some applications and methods of large deviations in finance and insurance. In *Paris-Princeton Lectures on Mathematical Finance 2004*, volume 1919 of *Lecture Notes in Math.*, pages 191–244. Springer, Berlin, 2007.

[SH09] Das Sanjiv and Paul Hanouna. Implied recovery. *Journal of Economic Dynamics and Control*, 33(11):1837–1857, 2009.

[Sowa] Richard B. Sowers. Exact pricing asymptotics of investment-grade tranches of synthetic cdo's part i: A large homogeneous pool. *International Journal of Theoretical and Applied Finance*. to appear.

[Sowb] Richard B. Sowers. Large correlation structures in pools of defaults: A large deviations analysis. submitted.

[Str93] Daniel W. Stroock. *Probability theory, an analytic view*. Cambridge University Press, Cambridge, 1993.

Division of Applied Mathematics, Brown University, Providence, RI 02912
E-mail address: `kspiliop@dam.brown.edu`

Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL 61801
E-mail address: `r-sowers@illinois.edu`