Asymptotically Optimum Estimation of a Probability in Inverse Binomial Sampling under General Loss Functions

L. Mendo[a,*]

[a]*E.T.S. Ing. Telecomunicación, Universidad Politécnica de Madrid, 28040 Madrid, Spain.*
Tel.: +34 91 549 5700. Fax: +34 91 336 7350.

Abstract

The optimum quality that can be asymptotically achieved in the estimation of a probability p using inverse binomial sampling is addressed. A general definition of quality is used in terms of the risk associated with a loss function that satisfies certain assumptions. It is shown that the limit superior of the risk for p asymptotically small has a minimum over all (possibly randomized) estimators. This minimum is achieved by certain non-randomized estimators. The model includes commonly used quality criteria as particular cases. Applications to the non-asymptotic regime are discussed considering specific loss functions, for which minimax estimators are derived.

Keywords: Sequential estimation, Asymptotic properties, Minimax estimators, Inverse binomial sampling.
2000 MSC: 62L12, 62F12.

1. Introduction

The problem of sequentially estimating the probability of success, p, in a sequence of Bernoulli trials arises in many fields of science and engineering. A stopping rule of notable interest, first discussed by Haldane (1945), is *inverse binomial sampling*, which consists in observing the random sequence until a given number r of successes are obtained. The resulting number of trials, N, is a sufficient statistic (Lehmann and Casella, 1998, p. 101), from which p can be estimated. The appeal

*Corresponding author
Email address: lmendo@grc.ssr.upm.es (L. Mendo)

of this rule lies in the useful properties of estimators obtained from it. Namely, previous works have shown that the uniformly minimum variance unbiased estimator, given by (Haldane, 1945)

$$\hat{p} = \frac{r-1}{N-1}, \tag{1}$$

satisfies the following properties. Its normalized mean square error $\mathrm{E}[(\hat{p}-p)^2]/p^2$ has an asymptotic value for $r \geq 3$, namely $1/(r-2)$; and $\mathrm{E}[(\hat{p}-p)^2]/p^2$ is guaranteed to be smaller than this value for any $p \in (0,1)$ (Mikulski and Smith, 1976). Similarly, the normalized mean absolute error $\mathrm{E}[|p-\hat{p}|]/p$ is smaller than its asymptotic value, given by $2(r-1)^{r-2}\exp(-r+1)/(r-2)!$, for any $p \in (0,1)$ and $r \geq 2$ (Mendo, 2009). In addition, given $\mu_1, \mu_2 > 1$ and $r \geq 3$, under certain conditions this estimator, as well as the modified version $\hat{p} = (r-1)/N$, can guarantee that, for p arbitrary, the random interval $[\hat{p}/\mu_1, \hat{p}\mu_2]$ contains the true value p with a confidence level greater than a prescribed value (Mendo and Hernando, 2006, 2008).

The results mentioned apply to specific estimators, defined as functions of the sufficient statistic N. A natural extension is to investigate whether the quality of the estimation can be improved using other estimators. The most general class is that formed by randomized estimators defined in terms of N. This includes non-randomized estimators as a particular case. This problem is addressed by Mendo and Hernando (2010), using the confidence associated with a relative interval as a quality measure. It is shown that the confidence that can be guaranteed for p asymptotically small has a maximum over all estimators. Moreover, non-randomized estimators are given that can guarantee this maximum confidence not only asymptotically, but also for $p \in (0,1)$ arbitrary.

A further generalization is to consider arbitrary estimators with an arbitrary definition of quality. The present paper pursues this direction, focusing on the asymptotic regime. Namely, quality is defined as the risk associated with an arbitrary loss function. The allowed loss functions are restricted only by certain regularity conditions, which are easily satisfied in practice (and which, in particular, hold for all the previously mentioned examples of quality measures). Using this general definition of quality, the asymptotic performance as $p \to 0$ of arbitrary estimators in inverse binomial sampling is analyzed. As will be seen, the quality that can be asymptotically achieved has a maximum over all estimators. Furthermore, this maximum can be accomplished using certain non-randomized estimators, whose form is explicitly given.

Section 2 contains preliminary definitions and observations required for the main results, which are presented in Section 3. Section 4 discusses these results, and considers applications in the non-asymptotic regime. Proofs of all results are given in Appendix A.

2. Preliminaries

The following notation will be used. Let $k^{(i)}$ denote $k(k-1)\cdots(k-i+1)$, for $k \in \mathbb{Z}$, $i \in \mathbb{N}$; and $k^{(0)} = 1$. Given $r \in \mathbb{N}$, the probability function of N, $f(n) = \Pr[N = n]$, is

$$f(n) = \frac{(n-1)^{(r-1)}}{(r-1)!} p^r (1-p)^{n-r}, \quad n \geq r. \tag{2}$$

The upper and lower (not normalized) incomplete gamma functions are respectively denoted as

$$\Gamma(s,u) = \int_u^\infty \tau^{s-1} \exp(-\tau)\,\mathrm{d}\tau, \tag{3}$$

$$\gamma(s,u) = \int_0^u \tau^{s-1} \exp(-\tau)\,\mathrm{d}\tau = \Gamma(s) - \Gamma(s,u). \tag{4}$$

In addition, the functions $\phi(\nu)$ and $\psi(x,\Omega)$ are defined as

$$\phi(\nu) = \frac{\nu^{r-1} \exp(-\nu)}{(r-1)!}, \quad \nu \in \mathbb{R}^+, \tag{5}$$

$$\psi(x,\Omega) = \frac{\Omega^r \exp(-\Omega/x)}{x^{r+1}(r-1)!}, \quad x, \Omega \in \mathbb{R}^+. \tag{6}$$

Given a function h, the one-sided limits $\lim_{x\to a^-} h(x)$ and $\lim_{x\to a^+} h(x)$ are respectively denoted as $h(a-)$ and $h(a+)$. Given two functions $h_1, h_2 : \mathbb{R}^+ \mapsto \mathbb{R}^+ \cup \{0\}$, $h_1(x)$ is $O(h_2(x))$ as $x \to \infty$ (respectively as $x \to 0$) if and only if there exist $a, M \in \mathbb{R}^+$ such that $h_1(x) \leq M h_2(x)$ for all $x \geq a$ (respectively for all $x \leq a$). Similarly, $h_1(x)$ is $\Theta(h_2(x))$ as $x \to \infty$ (respectively as $x \to 0$) if and only if there exist $a, m, M \in \mathbb{R}^+$ such that $m h_2(x) \leq h_1(x) \leq M h_2(x)$ for all $x \geq a$ (respectively for all $x \leq a$).

The quality of an estimator $\hat{p}$ is measured by the *risk* (expected loss) $\eta = \mathrm{E}[L(\hat{p}/p)]$ associated with a non-negative *loss function* $L : \mathbb{R}^+ \mapsto \mathbb{R}^+ \cup \{0\}$, provided that this expectation exists. The function L is defined in terms of $\hat{p}/p$, rather than $\hat{p}$. This is motivated by the fact that a given error value is most meaningful when compared with p, and therefore commonly used quality measures are most often *normalized* ones.

The loss function is assumed to satisfy the following.

Assumption 1. *For any $x_1, x_2 \in \mathbb{R}^+$ with $x_2 > x_1$, L is of bounded variation on $[x_1, x_2]$.*

Assumption 2. *For any* $x_1, x_2 \in \mathbb{R}^+$ *with* $x_2 > x_1$*,* L *has a finite number of discontinuities in* $[x_1, x_2]$.

Assumption 3. *The loss function has the following asymptotic behaviour:*

1. *There exists* $K \in \mathbb{R}$ *such that* $L(x)$ *is* $O(x^K)$ *as* $x \to 0$.
2. *There exists* $K' < r$ *such that* $L(x)$ *is* $O(x^{K'})$ *as* $x \to \infty$.

These restrictions are very mild. Note that the loss function L is not required to be convex, or continuous; however, being of bounded variation implies that its discontinuities can only be jumps or removable discontinuities, i.e. L has left-hand and right-hand limits at every point of its domain, and these limits are finite (Carter and van Brunt, 2000, corollary 2.7.3). All quality measures mentioned in Section 1 can be expressed in terms of functions of $x = \hat{p}/p$ for which Assumptions 1–3 hold. Namely, $L(x) = (x-1)^2$ corresponds to normalized mean square error; $L(x) = |x-1|$ to normalized mean absolute error; and given $\mu_1, \mu_2 > 1$,

$$L(x) = \begin{cases} 0 & \text{if } x \in [1/\mu_2, \mu_1], \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

corresponds to 1 minus the confidence associated with a relative interval $[p/\mu_2, p\mu_1]$.

Since N is a sufficient statistic, for any estimator defined in terms of the observed sequence of Bernoulli variables for which $\mathrm{E}[L(\hat{p}/p)]$ exists, there is a possibly randomized estimator expressed only in terms of N that has the same risk (Lehmann and Casella, 1998, p. 33). Therefore, attention can be restricted to estimators that depend on the observations through N only; however, randomized estimators need to be considered in addition to non-randomized ones.

The set of all functions from $\{r, r+1, r+2, \dots\}$ to $\mathbb{R}^+$ is denoted as $\mathscr{F}$. A *non-randomized estimator* $\hat{p}$ is defined as $\hat{p} = g(N)$, with $g \in \mathscr{F}$. A *randomized estimator* is a positive random variable $\hat{p}$ whose distribution depends on the value of N. The distribution function of $\hat{p}$ conditioned on $N = n$ will be denoted as Π_n. The randomized estimator is completely specified by the functions Π_n, $n \geq r$. Denoting by $\mathscr{F}_{\mathscr{R}}$ the class of all functions from $\{r, r+1, r+2, \dots\}$ to the set of distribution functions, a randomized estimator is defined by a function $G \in \mathscr{F}_{\mathscr{R}}$ that to each n assigns Π_n. Clearly, non-randomized estimators form a subset of the class of randomized estimators. Throughout the paper, when referring to an arbitrary estimator without specifying its type, the general class of randomized estimators (including non-randomized ones) will be meant.

The risk will be explicitly denoted in the sequel as a function of p, that is, $\eta(p)$. For a non-randomized estimator defined by $g \in \mathscr{F}$, the risk $\eta(p)$ is given by

$$\eta(p) = \sum_{n=r}^{\infty} f(n) L(g(n)/p). \tag{8}$$

Depending on L, g and p, this series may be convergent or not; however, boundedness of g is sufficient to ensure that the series converges for all L satisfying Assumptions 1–3 and for all p. In general, for possibly randomized estimators,

$$\eta(p) = \sum_{n=r}^{\infty} f(n) \int_0^{\infty} L(y/p)\, \mathrm{d}\Pi_n(y), \tag{9}$$

where the integral is defined in the Lebesgue-Stieltjes sense. Assumptions 1–3 assure that this integral always exists; however, it may be finite or infinite. Besides, even if it is finite for a given p and for all n, the series in (9) does not necessarily converge for that p. According to this, for an arbitrary estimator and for p given, $\eta(p)$ may be finite or infinite; however, there exist estimators that have a finite risk for all p.

An arbitrary estimator may not have an asymptotic risk, i.e. $\lim_{p\to 0} \eta(p)$ need not exist in general. Therefore, the asymptotic behaviour of an estimator should be characterized by $\limsup_{p\to 0} \eta(p)$. The significance of the limit superior lies in the fact that it is the smallest value such that any greater number is asymptotically an upper bound of $\eta(p)$. That is, given any $\eta_0 > \limsup_{p\to 0} \eta(p)$, there exists $\delta > 0$ such that $\eta(p) < \eta_0$ for all $p < \delta$; and no such δ can be found for $\eta_0 < \limsup_{p\to 0} \eta(p)$.[1]

According to the preceding discussion, a desirable asymptotic property of an estimator is that it achieves a low value of $\limsup_{p\to 0} \eta(p)$. In order to characterize how low this value can be, the infimum of $\limsup_{p\to 0} \eta(p)$ over all estimators should be determined. A related question is whether there is an estimator that can attain this infimum. As will be seen, the answer to this question is affirmative, that is, the infimum is also a minimum. This implies that there exist optimum estimators from the point of view of asymptotic behaviour; moreover, they can be found within the class of non-randomized estimators, as will also be shown. To obtain these results, the following approach will be used. It will be first established that for a certain subclass of non-randomized estimators, $\lim_{p\to 0} \eta(p)$ exists and can be easily computed. Secondly, it will be proved that $\lim_{p\to 0} \eta(p)$ has a minimum value over the referred subclass. Thirdly, this minimum will be shown to coincide with the unrestricted minimum of $\limsup_{p\to 0} \eta(p)$ over the class of arbitrary estimators.

[1] For $\eta_0 = \limsup_{p\to 0} \eta(p)$ the result may hold or not depending on the estimator and loss function; for example, it holds for (1) and normalized mean square error, as mentioned in Section 1, whereas it obviously does not hold for a constant loss function.

3. Main results

For a given loss function L, the set of all functions $g \in \mathscr{F}$ such that $\lim_{p\to 0} \eta(p)$ exists for $\hat{p} = g(n)$ is denoted as $\mathscr{F}_{\mathrm{p}}$. The set of functions $g \in \mathscr{F}$ for which $\lim_{n\to\infty} ng(n)$ exists, is finite and non-zero is denoted as $\mathscr{F}_{\mathrm{n}}$. Observe that the definition of $\mathscr{F}_{\mathrm{p}}$ generalizes that given by Mendo and Hernando (2010), which assumes a specific loss function, namely (7). The result in Theorem 1 to follow establishes that $\mathscr{F}_{\mathrm{n}} \subseteq \mathscr{F}_{\mathrm{p}}$, and explicitly gives $\lim_{p\to 0} \eta(p)$. For any $g \in \mathscr{F}_{\mathrm{n}}$ with $\lim_{n\to\infty} ng(n) = \Omega$, let

$$\bar{\eta} = \int_0^\infty \phi(\nu) L(\Omega/\nu)\, \mathrm{d}\nu. \tag{10}$$

Equivalently, $\bar{\eta}$ can be expressed as

$$\bar{\eta} = \int_0^\infty \psi(x,\Omega) L(x)\, \mathrm{d}x \tag{11}$$

by means of the change of variable $\nu = \Omega/x$ (both expressions are used in the proofs of the results to be presented). By Assumptions 1 and 3, these integrals exist as improper Riemann integrals, and have a finite value. It should be observed (and is exploited in the proofs) that they can also be interpreted as Lebesgue integrals (Apostol, 1974, theorem 10.33).

Theorem 1. *Consider $r \in \mathbb{N}$. For any loss function satisfying Assumptions 1–3, and for any non-randomized estimator defined by a function $g \in \mathscr{F}_{\mathrm{n}}$, the limit $\lim_{p\to 0} \eta(p)$ exists and equals $\bar{\eta}$ given by (10) (or (11)).*

According to this, the asymptotic risk of an estimator defined by any function $g \in \mathscr{F}_{\mathrm{n}}$ depends on this function only through Ω, i.e. only the asymptotic behaviour of g matters. Furthermore, under an additional assumption, it can be shown that the asymptotic risk is a C^1 function of Ω.

Assumption 2'. *L has a finite number of discontinuities in $\mathbb{R}^+$.*

It is evident that Assumption 2' implies Assumption 2. While more restrictive, Assumption 2' is satisfied by a large class of loss functions, including the mentioned examples.

Proposition 1. *Given $r \in \mathbb{N}$, a loss function satisfying Assumptions 1, 2' and 3, and an estimator defined by a function $g \in \mathscr{F}_{\mathrm{n}}$, the asymptotic risk $\bar{\eta}$ is a C^1 function of $\Omega \in \mathbb{R}^+$, with*

$$\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega} = \int_0^\infty \frac{\partial \psi(x,\Omega)}{\partial \Omega} L(x)\, \mathrm{d}x. \tag{12}$$

Denoting by $\bar{\eta}|_r$ the asymptotic risk corresponding to Ω and r given, this derivative can be expressed as

$$\frac{\mathrm{d}\bar{\eta}|_r}{\mathrm{d}\Omega} = \frac{r(\bar{\eta}|_r - \bar{\eta}|_{r+1})}{\Omega}. \tag{13}$$

Within the restricted class of non-randomized estimators defined by $\mathscr{F}_{\mathrm{n}}$, it is natural to search for values of Ω that yield low values of the asymptotic risk $\bar{\eta}$. Depending on the loss function, there may be or not an optimum value of $\Omega \in \mathbb{R}^+$, in the sense of minimizing $\bar{\eta}$. Theorem 2 to follow establishes that, under certain additional hypotheses (represented by Assumption 4), $\bar{\eta}$ indeed has a minimum with respect to Ω.

Assumption 4. *The loss function satisfies the following properties:*

1. *There exists $\xi \in \mathbb{R}^+$ such that L is non-increasing on $(0,\xi)$ and*

$$\int_{\xi}^{\infty} \frac{L(\xi) - L(x)}{x^{r+1}}\,\mathrm{d}x > 0. \tag{14}$$

2. *There exists $\xi' \in \mathbb{R}^+$ such that L is non-decreasing on (ξ',∞) and one of these conditions holds:*
   (a) $L(\xi'-) < L(\xi'+)$.
   (b) *There is $t \in \mathbb{N}$ such that L is of class C^t on an interval containing ξ' and*

$$\left.\frac{\mathrm{d}^i L}{\mathrm{d}x^i}\right|_{x=\xi'} = 0 \quad \textit{for } i = 1,2,\ldots,t-1, \tag{15}$$

$$(-1)^{t-1} \left.\frac{\mathrm{d}^t L}{\mathrm{d}x^t}\right|_{x=\xi'} > 0. \tag{16}$$

The next proposition gives a sufficient condition that may help in assessing whether a given loss function satisfies property 1 in Assumption 4.

Proposition 2. *If there exist $A \in \mathbb{R}$ and B, s such that*

$$\lim_{x\to 0} \frac{L(x) - A}{x^s} = B \quad \textit{with } Bs < 0,\ s < r, \tag{17}$$

inequality (14) holds for some $\xi \in \mathbb{R}^+$.

Theorem 2. *Given $r \in \mathbb{N}$ and a loss function satisfying Assumptions 1, 2', 3 and 4, consider the class of non-randomized estimators defined by functions $g \in \mathscr{F}_{\mathrm{n}}$. Denoting $\Omega = \lim_{n\to\infty} ng(n)$, there exists a value of Ω which minimizes the asymptotic risk $\bar{\eta}$ among all $\Omega \in \mathbb{R}^+$.*

This theorem indicates that in the stated conditions, and restricted to the class defined by $\mathscr{F}_\mathrm{n}$, there is an optimum value of Ω from the point of view of asymptotic risk. This optimum is not necessarily unique. In the sequel, η^* will denote the minimum of $\bar{\eta}$ over the class of estimators defined by $\mathscr{F}_\mathrm{n}$, and Ω^* will denote any value of Ω which attains this minimum, that is,

$$\eta^* = \int_0^\infty \phi(\nu) L(\Omega^*/\nu)\,\mathrm{d}\nu. \tag{18}$$

Assumption 4 holds for a wide range of loss functions, and in particular for those corresponding to normalized mean square error, normalized mean absolute error, and confidence associated with a relative interval. It is not difficult, however, to find a loss function for which the assumption does not hold, and for which $\bar{\eta}$ does not have a minimum over the class defined by $\mathscr{F}_\mathrm{n}$. For example, given $A_1, A_2 > 0$, let

$$L(x) = \begin{cases} 0 & \text{if } x \in [1/\mu_2, \mu_1], \\ A_2 & \text{if } x < 1/\mu_2, \\ A_1 & \text{if } x > \mu_1, \end{cases} \tag{19}$$

which is a generalized version of (7). Substituting (19) into (14), it is seen that property 1 in Assumption 4 is satisfied if and only if

$$\frac{A_1}{A_2} < (\mu_1\mu_2)^r, \tag{20}$$

while property 2 holds irrespective of A_1 and A_2. On the other hand, for $\Omega \in \mathbb{R}$, substituting (19) into (10) and computing $\mathrm{d}\bar{\eta}/\mathrm{d}\Omega$ gives

$$\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega} = \frac{\Omega^{r-1}\left(A_1\mu_1^{-r}\exp(-\Omega/\mu_1) - A_2\mu_2^r\exp(-\Omega\mu_2)\right)}{(r-1)!}. \tag{21}$$

This implies that $\bar{\eta}$ has a single minimum over $\Omega \in \mathbb{R}$, located at

$$\Omega = \frac{r\log(\mu_1\mu_2) - \log(A_1/A_2)}{\mu_2 - 1/\mu_1}. \tag{22}$$

This value is positive if and only if (20), or equivalently property 1 in Assumption 4, is satisfied. Thus, if this property does not hold, $\bar{\eta}$ is monotonically increasing for $\Omega \in \mathbb{R}^+$, which implies that there is not an optimum Ω within $\mathbb{R}^+$.

Under the hypotheses of Theorem 2, the optimum value of Ω for the considered r, i.e. Ω^*, satisfies, by Proposition 1,

$$\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega} = 0 \tag{23}$$

(or equivalently, using the notation in the referred proposition, $\bar{\eta}|_r = \bar{\eta}|_{r+1}$). Thus if (23) has only one solution, it must be equal to Ω^*. If there are several solutions, at least one corresponds to the absolute minimum of $\bar{\eta}$, although not necessarily all of them do.

According to Theorem 2, if the loss function satisfies Assumptions 1, 2', 3 and 4, any non-randomized estimator defined by a function $g \in \mathscr{F}_\mathrm{n}$ with $\lim_{n\to\infty} ng(n) = \Omega^*$ minimizes $\limsup_{p\to 0} \eta(p)$ within the restricted class of estimators represented by $\mathscr{F}_\mathrm{n}$; but not necessarily within the class of all non-randomized estimators, or within the general class of possibly randomized estimators. However, under slightly stronger conditions this turns out to be true, as established by the next theorem.

Assumption 3'. *The loss function has the following asymptotic behaviour:*

1. *There exists $K < r$ such that $L(x)$ is $\Theta(x^K)$ as $x \to 0$.*
2. *There exists $K' < r$ such that $L(x)$ is $\Theta(x^{K'})$ as $x \to \infty$.*

Assumption 3' replaces Assumption 3, in the sense that each of the two properties in Assumption 3' implies the corresponding one in Assumption 3. The new conditions are only slightly more restrictive, and are still satisfied by a large set of loss functions, in particular by those previously mentioned as examples.

Theorem 3. *Given $r \in \mathbb{N}$ and any loss function satisfying Assumptions 1, 2', 3' and 4, $\limsup_{p\to 0} \eta(p)$ has a minimum over the general class of estimators defined by $\mathscr{F}_\mathscr{R}$, and this minimum equals η^*.*

Corollary 1. *Under the hypotheses of Theorem 3, any non-randomized estimator defined by a function $g \in \mathscr{F}_\mathrm{n}$ with $\lim_{n\to\infty} ng(n) = \Omega^*$ minimizes $\limsup_{p\to 0} \eta(p)$ among all (possibly randomized) estimators based on inverse binomial sampling.*

Theorem 3 and Corollary 1 show that, under the stated assumptions, an estimator can be found within the class defined by $\mathscr{F}_\mathrm{n}$ that is asymptotically optimum over the general class represented by $\mathscr{F}_\mathscr{R}$.

4. Discussion and applications

Since p is unknown, it is desirable to have an estimator that *guarantees* that the risk is not larger than a given η_0 for p arbitrary, or at least for all p within a certain interval; that is, such that $\eta(p) \leq \eta_0$ for p in some interval (p_1, p_2), with $0 \leq p_1 < p_2 \leq 1$. If $p_1 = 0$, the estimator is said to *asymptotically guarantee* that the risk is not larger than η_0; if, in addition, $p_1 = 1$, it *globally guarantees* that the risk is not larger than η_0.

The results presented in Section 3 generalize the asymptotic analysis by Mendo and Hernando (2010), which considers the specific loss function (7), to arbitrary functions satisfying the indicated assumptions. The importance of these asymptotic results lies not only in the fact that in many applications p is small, but also in the observation that asymptotic behaviour sets a restriction on the risk that can be guaranteed. This restriction is represented by the following proposition (which is a straightforward generalization of Mendo and Hernando (2010, proposition 1)) and its corollary.

Proposition 3. *If an estimator has a risk $\eta(p)$ not larger than a given η_0 for all $p \in (p_1, p_2)$, then necessarily $\limsup_{p \to p_0} \eta(p) \leq \eta_0$ for any $p_0 \in [p_1, p_2]$.*

Corollary 2. *Given $r \in \mathbb{N}$ and a loss function that satisfies Assumptions 1, 2', 3' and 4, for any $\eta_0 < \eta^*$ and $p_2 > 0$, no estimator can guarantee that $\eta(p) \leq \eta_0$ for all $p < p_2$.*

According to the results in Section 3, if Assumptions 1, 2', 3' and 4 are satisfied, any estimator defined by $g \in \mathscr{F}_\mathrm{n}$ with $\lim_{n\to\infty} ng(n) = \Omega^*$ can asymptotically guarantee that the risk is not larger than $\eta^* + \varepsilon$ for any $\varepsilon > 0$, whereas Corollary 2 states that no estimator exists with this property for $\varepsilon < 0$. It remains to be seen if there exist estimators that asymptotically guarantee that $\eta(p) \leq \eta^*$; and, particularly, if this guarantee can be global. The answer to these questions depends on the loss function under consideration. Since a general analysis seems impracticable, a separate study needs to be carried out for each loss function. Several important cases are discussed next, including the loss functions already mentioned as examples.

4.1. Confidence

For the loss function given by (7), $\eta(p)$ equals $1 - c(p)$, where $c(p) = \Pr[p/\mu_2 \leq \hat{p} \leq p\mu_1] = \Pr[\hat{p}/\mu_1 \leq p \leq \hat{p}\mu_2]$ is the *confidence* associated with a relative interval defined by $\mu_1, \mu_2 > 1$. Let $c^* = 1 - \eta^*$, which represents the maximum confidence that could be guaranteed to be exceeded. The analysis by Mendo and Hernando (2010) shows that assuming $r \geq 3$, the inequality $c(p) > c^*$ can indeed be asymptotically guaranteed for any μ_1, μ_2, and globally guaranteed if μ_1, μ_2 satisfy certain conditions.

4.2. Mean absolute error

For $L(x) = |x - 1|$, risk corresponds to *normalized mean absolute error*. Considering an estimator $\hat{p} = g(N)$ with $\lim_{n\to\infty} ng(n) = \Omega$, and for $r \geq 2$, (10) gives the asymptotic risk

$$\bar{\eta} = \int_0^\infty \phi(\nu) \left| \frac{\Omega}{\nu} - 1 \right| \mathrm{d}\nu = \frac{2\,(\Gamma(r,\Omega) - \Omega\Gamma(r-1,\Omega))}{(r-1)!} + \frac{\Omega}{r-1} - 1, \qquad (24)$$

and it is straightforward to show that (23) reduces to $\Gamma(r-1,\Omega) = (r-2)!/2$. This equation has only one solution, which thus corresponds to Ω^*. Interestingly, for $\hat{p} = \Omega^*/(n-1)$ with $r \geq 2$, numerically evaluating $\eta(p)$ suggests that this estimator may globally guarantee $\eta(p) \leq \eta^*$. However, proving this conjecture remains an open problem.

4.3. Mean square error

The function $L(x) = (x-1)^2$ corresponds to *normalized mean square error*. This loss function lends itself easily to non-asymptotic analysis. Considering an estimator $\hat{p} = g(N)$ with $\lim_{n\to\infty} ng(n) = \Omega$, and assuming $r \geq 3$, (10) gives

$$\bar{\eta} = \int_0^\infty \phi(\nu) \left(\frac{\Omega}{\nu} - 1 \right)^2 \mathrm{d}\nu = \frac{\Omega^2}{(r-1)(r-2)} - \frac{2\Omega}{r-1} + 1, \qquad (25)$$

and thus (23) has the single solution $\Omega = r-2$, which is the optimum value for Ω, i.e. Ω^*. From (25) the resulting η^* is $1/(r-1)$. As established by the next proposition, an estimator can be found that globally guarantees that the risk is not larger than η^*, namely

$$\hat{p} = \frac{r-2}{N-1}. \qquad (26)$$

Proposition 4. *Given $r \geq 3$, and for any $p \in (0,1)$, the estimator (26) satisfies*

$$\frac{\mathrm{E}[(\hat{p}-p)^2]}{p^2} < \frac{1}{r-1}. \qquad (27)$$

The following corollary is obtained from Theorem 3 and Proposition 4.

Corollary 3. *For $r \geq 3$, the estimator (26) minimizes $\sup_{p\in(0,1)} \mathrm{E}[(\hat{p}-p)^2]/p^2$ among all (possibly randomized) estimators based on inverse binomial sampling.*

Thus the estimator given by (26) not only minimizes $\limsup_{p\to 0} \mathrm{E}[(\hat{p}-p)^2]/p^2$, but also $\sup_{p\in(0,1)} \mathrm{E}[(\hat{p}-p)^2]/p^2$, i.e. it is minimax with respect to normalized mean square error. Therefore, from the point of view of guaranteeing that the normalized mean square error does not exceed a given value, (26) is optimum among all estimators based on inverse binomial sampling.

Comparing the estimators (1) and (26), the former can only guarantee $\mathrm{E}[(\hat{p} - p)^2]/p^2 < 1/(r-2)$, whereas the latter guarantees $\mathrm{E}[(\hat{p}-p)^2]/p^2 < 1/(r-1)$. This better (in fact, optimum) performance is obtained at the expense of some bias; namely, it is easily seen that (26) gives $\mathrm{E}[\hat{p}]/p = 1 - 1/(r-1)$.

4.4. A generalization of confidence

According to Mendo and Hernando (2010, proposition 3), for the loss function (7), given $\Omega \in \mathbb{R}^+$ and assuming that $r \geq 3$, $\mu_1 \geq \Omega/(r-\sqrt{r})$ and $\mu_2 \geq (r+\sqrt{r}+1)/\Omega$, the estimator

$$\hat{p} = \frac{\Omega}{N+1} \tag{28}$$

globally guarantees that $\eta(p)$ is smaller than its asymptotic value $\bar{\eta}$. Taking into account that, in this case, $\eta(p) = \Pr[\hat{p} < p/\mu_2] + \Pr[\hat{p} > p\mu_1]$ and that the proof given in the cited reference considers the terms $\Pr[\hat{p} < p/\mu_2]$ and $\Pr[\hat{p} > p\mu_1]$ separately, it can be seen that the same result holds for the loss function (19) with $A_1 = 0$ or $A_2 = 0$. Furthermore, the result can be generalized to any loss function that can be approximated as a (possibly infinite) sum of functions of this form. This is the content of the next proposition.

Proposition 5. *Given $r \geq 3$ and $\Omega \in \mathbb{R}^+$, consider a loss function for which Assumptions 2', 3' and 4 hold and that satisfies the following:*

1. *L is constant on an interval $[\upsilon, \upsilon']$, with*

$$\upsilon \leq \frac{\Omega}{r+\sqrt{r}+1}, \quad \upsilon' \geq \frac{\Omega}{r-\sqrt{r}}. \tag{29}$$

2. *L is non-increasing on $(0, \upsilon]$.*
3. *L is non-decreasing on $[\upsilon', \infty)$.*

In these conditions, for any $p \in (0,1)$ the risk $\eta(p)$ of the estimator (28) satisfies $\eta(p) \leq \bar{\eta}$, with $\bar{\eta}$ given by (10) (or (11)).

It is noted that conditions 1–3 of Proposition 5 imply that Assumption 1 necessarily holds, and also imply that $L(\upsilon-) \geq L(\upsilon+)$ and $L(\upsilon'-) \leq L(\upsilon'+)$.

The following result, analogous to Corollary 3, is obtained for the estimator

$$\hat{p} = \frac{\Omega^*}{N+1}. \tag{30}$$

Corollary 4. *Given $r \geq 3$ and a loss function that satisfies Assumptions 1, 2', 3' and 4, let Ω^* be as determined by Theorem 2. If conditions 1–3 in Proposition 5 hold for some υ, υ' with*

$$\upsilon \leq \frac{\Omega^*}{r+\sqrt{r}+1}, \quad \upsilon' \geq \frac{\Omega^*}{r-\sqrt{r}}, \tag{31}$$

the estimator (30) minimizes $\sup_{p \in (0,1)} \eta(p)$ among all (possibly randomized) estimators based on inverse binomial sampling.

This establishes that, under the stated hypotheses, the estimator (30) is minimax, i.e. minimizes the risk that can be globally guaranteed not to be exceeded.

Appendix A. Proofs

The following definitions are necessary:

$$\Phi(p,\nu) = \frac{(1-p)^{\nu/p-r}}{(r-1)!}\prod_{i=1}^{r-1}(\nu - ip), \quad p \in (0,1),\ \nu \in \mathbb{R}^+, \tag{A.1}$$

$$\zeta = \int_{r/\sigma}^{r\sigma} \phi(\nu) L(\Omega/\nu)\,\mathrm{d}\nu, \quad \Omega, \sigma \in \mathbb{R}^+. \tag{A.2}$$

Lemma 1 (Mendo and Hernando (2010, lemma 1))**.** *For any* $\nu \in \mathbb{R}^+$, $0 < \phi(\nu) < 1$.

Lemma 2. *Given* $\nu_1, \nu_2 \in \mathbb{R}^+$ *with* $\nu_2 > \nu_1$*, for* $\nu \in [\nu_1, \nu_2]$ *the function* $\Phi(p,\nu)$ *converges uniformly to* $\phi(\nu)$ *as* $p \to 0$.

Proof. The lemma is equivalent to the result that $\Phi(p_k,\nu)$ converges uniformly on $\nu \in [\nu_1,\nu_2]$ for any sequence (p_k) such that $p_k \in (0,1)$, $p_k \to 0$, which is proved by Mendo and Hernando (2010, lemma 3). □

Proof of Theorem 1. The risk $\eta(p)$ tends to $\bar{\eta}$ for $p \to 0$ if and only if $\eta(p_k)$ converges to $\bar{\eta}$ for every sequence (p_k) such that $p_k \in (0,1)$, $p_k \to 0$ (Apostol, 1974, theorem 4.12). Consider an arbitrary sequence of this type. Let $\eta^k = \eta(p_k)$, and let f_k denote the probability function f for $p = p_k$. Defining $\phi_k(\nu) = \Phi(p_k,\nu)$, it is seen from (2) and (A.1) that $f_k(n) = p_k \phi_k(np_k)$.

From property 1 in Assumption 3, there exist $K \in \mathbb{R}$ and $M_\mathrm{L}, x_\mathrm{L} \in \mathbb{R}^+$ such that

$$L(x) < M_\mathrm{L} x^K \quad \text{for } x < x_\mathrm{L}. \tag{A.3}$$

Without loss of generality, it will be assumed that $K < 0$. On the other hand, property 2 implies that there exist $K' < r$ and $M'_\mathrm{L}, x'_\mathrm{L} \in \mathbb{R}^+$ such that

$$L(x) < M'_\mathrm{L} x^{K'} \quad \text{for } x > x'_\mathrm{L}. \tag{A.4}$$

The risk η^k is expressed from (8) as

$$\eta^k = \sum_{n=r}^{\infty} f_k(n) L\left(\frac{g(n)}{p_k}\right). \tag{A.5}$$

Given $\alpha, \beta \in \mathbb{R}^+$ with $\beta > \alpha$, let the set I_k be defined as

$$I_k = \{\lfloor \alpha/p_k \rfloor, \lfloor \alpha/p_k \rfloor + 1, \ldots, \lceil \beta/p_k \rceil\}. \tag{A.6}$$

Under the assumption

$$p_k \leq \frac{\alpha}{r}, \tag{A.7}$$

which implies that $\min I_k = \lfloor \alpha/p_k \rfloor \geq r$, the following definition can be made:

$$\eta_0^k = \sum_{n \in I_k} f_k(n) L\left(\frac{g(n)}{p_k}\right). \tag{A.8}$$

The proof will proceed as follows. With a suitable choice of α and β, and for k sufficiently large, the term η_0^k can be made arbitrarily close to $\bar{\eta}$, as will be seen. On the other hand, the difference $\eta^k - \eta_0^k$ will be decomposed as the sum of three terms, each of which can be made arbitrarily small for sufficiently large k. Adequate bounds will be derived for each of these four terms, and then the bounds will be suitably combined to show that η^k tends to $\bar{\eta}$ as $k \to \infty$.

In the following, np_k will be denoted as $\nu_{n,k}$. Assuming

$$p_k \leq \frac{\alpha}{r+1}, \tag{A.9}$$

(which obviously implies (A.7)), it is easily seen that for $n \in I_k$, $\nu_{n,k}$ is contained in the interval I given as

$$I = \left[\frac{r\alpha}{r+1}, \beta + \frac{\alpha}{r+1}\right]. \tag{A.10}$$

Lemma 2 implies that the sequence of functions (ϕ_k) converges uniformly to ϕ for $\nu \in I$; that is, given $\varepsilon_{\text{unif}} > 0$, there exists k_{unif} such that $|\phi_k(\nu) - \phi(\nu)| < \varepsilon_{\text{unif}}$ for $\nu \in I$, $k \geq k_{\text{unif}}$. Thus $f_k(n) = p_k \phi(\nu_{n,k}) + p_k \theta_{\text{unif},n}$ with $|\theta_{\text{unif},n}| < \varepsilon_{\text{unif}}$ for $n \in I_k$, $k \geq k_{\text{unif}}$. In these conditions, since $\phi(\nu_{n,k}) > 0$ (Lemma 1), (A.8) can be expressed as

$$\eta_0^k = \sum_{n \in I_k} p_k \phi(\nu_{n,k}) \left(1 + \frac{\theta_{\text{unif},n}}{\phi(\nu_{n,k})}\right) L\left(\frac{g(n)}{p_k}\right). \tag{A.11}$$

On the other hand, since $ng(n) \to \Omega$ as $n \to \infty$, given $\varepsilon_{\text{est}} > 0$ there exists $n_{\text{est}} \geq r$ such that $|ng(n) - \Omega| < \varepsilon_{\text{est}}$ for all $n \geq n_{\text{est}}$, i.e. $g(n) = (\Omega + \theta_{\text{est},n})/n$ with $|\theta_{\text{est},n}| < \varepsilon_{\text{est}}$. Therefore, assuming

$$p_k \leq \frac{\alpha}{n_{\text{est}}}, \tag{A.12}$$

which implies that $\min I_k \geq n_{\text{est}}$, (A.11) can be written as

$$\eta_0^k = \sum_{n \in I_k} p_k \phi(\nu_{n,k}) \left(1 + \frac{\theta_{\text{unif},n}}{\phi(\nu_{n,k})}\right) L\left(\frac{\Omega + \theta_{\text{est},n}}{\nu_{n,k}}\right). \tag{A.13}$$

Denoting $m_\phi = \min_{\nu \in I} \phi(\nu)$, which is non-zero because of Lemma 1, it stems from (A.13) that

$$\eta_0^k = \left(1 + \frac{\theta_{\text{unif}}}{m_\phi}\right) \sum_{n \in I_k} p_k \phi(\nu_{n,k}) L\left(\frac{\Omega + \theta_{\text{est},n}}{\nu_{n,k}}\right) \tag{A.14}$$

for some θ_{unif} with $|\theta_{\text{unif}}| < \varepsilon_{\text{unif}}$.

Assuming $\varepsilon_{\text{est}} \leq \Omega/2$, and taking into account (A.9), it follows from (A.6) that for $n \in I_k$, both $\Omega/\nu_{n,k}$ and $(\Omega+\theta_{\text{est},n})/\nu_{n,k}$ are contained in the interval

$$I' = \left[\frac{\Omega}{2(\beta+\alpha/(r+1))}, \frac{3(r+1)\Omega}{2r\alpha}\right]. \tag{A.15}$$

According to Assumption 2, L has a finite number of discontinuities in I'. Let d denote this number. Each of these discontinuities, located at $x_1, \ldots, x_d$, may be either a jump or a removable discontinuity. Let

$$J = \sum_{i=1}^{d} \left(\left| \lim_{x \to x_i^-} L(x) - L(x_i) \right| + \left| \lim_{x \to x_i^+} L(x) - L(x_i) \right| \right). \tag{A.16}$$

Thus J represents the contribution of all discontinuities to the total variation of L on I'.

The function L on the interval I' can be decomposed as the sum of a continuous function L_{c} and a piecewise constant function L_{d}, the latter of which has discontinuities at $x_1, \ldots, x_d$. By the Heine-Cantor theorem (Apostol, 1974, theorem 4.47), L_{c} is uniformly continuous on I'. Since $|\theta_{\text{est},n}| < \varepsilon_{\text{est}}$, it follows that for any $\varepsilon_{\text{cont}} > 0$ there exists δ_{cont} such that $|L_{\text{c}}((\Omega+\theta_{\text{est},n})/\nu_{n,k}) - L_{\text{c}}(\Omega/\nu_{n,k})| < \varepsilon_{\text{cont}}$ for $\varepsilon_{\text{est}} < \delta_{\text{cont}}$, for all $n \in I_k$, and for all k. Regarding L_{d}, let

$$U_k = \left\{ n \in I_k \mid L_{\text{d}}\left(\frac{\Omega+\theta_{\text{est},n}}{\nu_{n,k}}\right) \neq L_{\text{d}}\left(\frac{\Omega}{\nu_{n,k}}\right) \right\}. \tag{A.17}$$

For $n \in I_k \setminus U_k$,

$$\left| L\left(\frac{\Omega+\theta_{\text{est},n}}{\nu_{n,k}}\right) - L\left(\frac{\Omega}{\nu_{n,k}}\right) \right| < \varepsilon_{\text{cont}}. \tag{A.18}$$

For each $n \in U_k$, $|L_{\text{d}}((\Omega+\theta_{\text{est},n})/\nu_{n,k}) - L_{\text{d}}(\Omega/\nu_{n,k})|$ can be at at most J, and thus

$$\left| L\left(\frac{\Omega+\theta_{\text{est},n}}{\nu_{n,k}}\right) - L\left(\frac{\Omega}{\nu_{n,k}}\right) \right| < \varepsilon_{\text{cont}} + J. \tag{A.19}$$

Let χ_k denote the number of elements of U_k divided by that of I_k. Taking into account that the latter is less than $(\beta-\alpha)/p_k + 3 < (\beta-\alpha+3)/p_k$ and that the function ϕ is upper-bounded by 1 (Lemma 1), from (A.18) and (A.19) it follows that, for $\varepsilon_{\text{est}} < \delta_{\text{cont}}$,

$$\left| \sum_{n \in I_k} p_k \phi(\nu_{n,k}) L\left(\frac{\Omega+\theta_{\text{est},n}}{\nu_{n,k}}\right) - \sum_{n \in I_k} p_k \phi(\nu_{n,k}) L\left(\frac{\Omega}{\nu_{n,k}}\right) \right| < (\beta-\alpha+3)(\varepsilon_{\text{cont}} + J\chi_k). \tag{A.20}$$

It is easily seen that $\lim_{k\to\infty}\chi_k$ can be made arbitrarily small by taking ε_{est} sufficiently small. Thus, given $\varepsilon_{\text{disc}}$, there exist δ_{disc}, k_{disc} such that $\chi_k < \varepsilon_{\text{disc}}$ for $\varepsilon_{\text{est}} < \delta_{\text{disc}}$, $k \geq k_{\text{disc}}$. Consequently, for $\varepsilon_{\text{est}} < \min\{\delta_{\text{cont}}, \delta_{\text{disc}}\}$ and $k \geq k_{\text{disc}}$,

$$\left| \sum_{n\in I_k} p_k \phi(v_{n,k}) L\left(\frac{\Omega+\theta_{\text{est},n}}{v_{n,k}}\right) - \sum_{n\in I_k} p_k \phi(v_{n,k}) L\left(\frac{\Omega}{v_{n,k}}\right) \right| < (\beta-\alpha+3)(\varepsilon_{\text{cont}} + J\varepsilon_{\text{disc}}). \quad \text{(A.21)}$$

From (A.14) and (A.21),

$$\eta_0^k = \left(1+\frac{\theta_{\text{unif}}}{m_\phi}\right)\left[\sum_{n\in I_k} p_k \phi(v_{n,k}) L\left(\frac{\Omega}{v_{n,k}}\right) + (\beta-\alpha+3)(\theta_{\text{cont}} + J\theta_{\text{disc}})\right] \quad \text{(A.22)}$$

with $|\theta_{\text{cont}}| < \varepsilon_{\text{cont}}$, $|\theta_{\text{disc}}| < \varepsilon_{\text{disc}}$. The sum over n in (A.22) tends to $\int_\alpha^\beta \phi(v)L(\Omega/v)\,\mathrm{d}v$ as $k\to\infty$. Thus for any $\varepsilon_{\text{int}} > 0$ there exists k_{int} such that for all $k \geq k_{\text{int}}$

$$\left| \sum_{n\in I_k} p_k \phi(v_{n,k}) L\left(\frac{\Omega}{v_{n,k}}\right) - \int_\alpha^\beta \phi(v) L\left(\frac{\Omega}{v}\right) \mathrm{d}v \right| < \varepsilon_{\text{int}}, \quad \text{(A.23)}$$

and therefore (A.22) can be expressed for $k \geq \max\{k_{\text{disc}}, k_{\text{int}}\}$ as

$$\eta_0^k = \left(1+\frac{\theta_{\text{unif}}}{m_\phi}\right)\left[\int_\alpha^\beta \phi(v) L\left(\frac{\Omega}{v}\right) \mathrm{d}v + \theta_{\text{int}} + (\beta-\alpha+3)(\theta_{\text{cont}} + J\theta_{\text{disc}})\right] \quad \text{(A.24)}$$

with $|\theta_{\text{int}}| < \varepsilon_{\text{int}}$. In addition, given any $\varepsilon_{\text{tail}}$, there exist α_{tail}, β_{tail} with $\beta_{\text{tail}} > \alpha_{\text{tail}}$ such that $|\bar\eta - \int_\alpha^\beta \phi(v)L(\Omega/v)\,\mathrm{d}v| < \varepsilon_{\text{tail}}$ for $0 < \alpha \leq \alpha_{\text{tail}}$, $\beta \geq \beta_{\text{tail}}$. Thus, in these conditions,

$$\eta_0^k = \left(1+\frac{\theta_{\text{unif}}}{m_\phi}\right)\left[\bar\eta + \theta_{\text{tail}} + \theta_{\text{int}} + (\beta-\alpha+3)(\theta_{\text{cont}} + J\theta_{\text{disc}})\right]. \quad \text{(A.25)}$$

with $|\theta_{\text{tail}}| < \varepsilon_{\text{tail}}$.

The difference $\eta^k - \eta_0^k$ can be expressed as $\eta_1^k + \eta_2^k + \eta_3^k$, where

$$\eta_1^k = \sum_{n=r}^{n_{\text{est}}-1} f_k(n) L\left(\frac{g(n)}{p_k}\right), \quad \text{(A.26)}$$

$$\eta_2^k = \sum_{n=n_{\text{est}}}^{\lfloor \alpha/p_k \rfloor - 1} f_k(n) L\left(\frac{g(n)}{p_k}\right), \quad \text{(A.27)}$$

$$\eta_3^k = \sum_{\lceil \beta/p_k \rceil + 1}^{\infty} f_k(n) L\left(\frac{g(n)}{p_k}\right). \quad \text{(A.28)}$$

Regarding the term η_1^k, from (2) it is seen that

$$f_k(n) < \frac{n^{r-1} p_k^r}{(r-1)!} \tag{A.29}$$

and therefore

$$0 < \eta_1^k < \sum_{n=r}^{n_{\mathrm{est}}-1} \frac{n^{r-1} p_k^r}{(r-1)!} L\left(\frac{g(n)}{p_k}\right) < \frac{n_{\mathrm{est}}^{r-1} p_k^r}{(r-1)!} \sum_{n=r}^{n_{\mathrm{est}}-1} L\left(\frac{g(n)}{p_k}\right). \tag{A.30}$$

The fact that $\lim_{n\to\infty} n g(n)$ exists and is finite implies that the function g is upper-bounded by some constant M_{g}. For $g(n)/p_k > x'_{\mathrm{L}}$, (A.4) implies that $L(g(n)/p_k) < M'_{\mathrm{L}}(M_{\mathrm{g}}/p_k)^{K'}$. On the other hand, $g(n)/p_k$ in (A.30) is greater than $m_{\mathrm{g}} = \min\{g(r), g(r+1), \dots, g(n_{\mathrm{est}}-1)\}$; and for $g(n)/p_k \in (m_{\mathrm{g}}, x'_{\mathrm{L}}]$, Assumption 1 implies that $L(g(n)/p_k)$ is lower than some value M'_{g}, where both m_{g} and M'_{g} depend on n_{est}. Thus, for the range of values of n in (A.30),

$$L\left(\frac{g(n)}{p_k}\right) < \max\left\{\frac{M'_{\mathrm{L}} M_{\mathrm{g}}^{K'}}{p_k^{K'}}, M'_{\mathrm{g}}\right\} < \frac{\max\{M'_{\mathrm{L}} M_{\mathrm{g}}^{K'}, M'_{\mathrm{g}}\}}{p_k^{K'}}. \tag{A.31}$$

The sum in the right-most part of (A.30) is either empty or else it contains $n_{\mathrm{est}} - r < n_{\mathrm{est}}$ terms. Therefore, using (A.31),

$$0 \le \eta_1^k < \frac{n_{\mathrm{est}}^r \max\{M'_{\mathrm{L}} M_{\mathrm{g}}^{K'}, M'_{\mathrm{g}}\}}{(r-1)!} p_k^{r-K'}. \tag{A.32}$$

Regarding η_2^k, the sum in (A.27) is empty for $\alpha/p_k < n_{\mathrm{est}} + 1$. If it is non-empty, since $n \ge n_{\mathrm{est}}$, the term $g(n)/p_k$ can be written as $(\Omega + \theta_{\mathrm{est},n})/\nu_{n,k}$ with $|\theta_{\mathrm{est},n}| < \varepsilon_{\mathrm{est}}$. Therefore, taking into account (A.29),

$$0 \le \eta_2^k < \frac{p_k}{(r-1)!} \sum_{n=n_{\mathrm{est}}}^{\lfloor \alpha/p_k \rfloor - 1} \nu_{n,k}^{r-1} L\left(\frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}}\right). \tag{A.33}$$

Since $\varepsilon_{\mathrm{est}} \le \Omega/2$, it holds that $\Omega/2 < \Omega + \theta_{\mathrm{est},n} < 3\Omega/2$, and thus for the range of values of n in (A.27)

$$\frac{3\Omega}{2\nu_{n,k}} > \frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}} > \frac{\Omega}{2\nu_{n,k}} > \frac{\Omega}{2\alpha}. \tag{A.34}$$

Therefore, assuming $\Omega/(2\alpha) \ge x'_{\mathrm{L}}$, for n within the indicated range it stems from (A.4) that

$$L\left(\frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}}\right) < M'_{\mathrm{L}} \left(\frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}}\right)^{K'} < M'_{\mathrm{L}} \left(\frac{3\Omega}{2\nu_{n,k}}\right)^{K'}. \tag{A.35}$$

Substituting (A.35) into (A.33),

$$0 \leq \eta_2^k < \frac{M_{\mathrm{L}}'(3\Omega/2)^{K'} p_k}{(r-1)!} \sum_{n=n_{\mathrm{est}}}^{\lfloor \alpha/p_k \rfloor - 1} \nu_{n,k}^{r-K'-1} < \frac{M_{\mathrm{L}}'(3\Omega/2)^{K'}}{(r-1)!} \alpha^{r-K'}. \tag{A.36}$$

Consider $\varepsilon_{\mathrm{tail}}' > 0$ arbitrary. Since $K' < r$, defining

$$\alpha_{\mathrm{tail}}' = \left(\frac{(r-1)!\, \varepsilon_{\mathrm{tail}}'}{M_{\mathrm{L}}'(3\Omega/2)^{K'}} \right)^{1/(r-K')} \tag{A.37}$$

it follows from (A.36) that for any $\alpha \leq \alpha_{\mathrm{tail}}'$

$$0 \leq \eta_2^k < \varepsilon_{\mathrm{tail}}'. \tag{A.38}$$

As for η_3^k, taking into account that $(1-p_k)^{1/p_k} < 1/e$, from (2) and (5) it is seen that $f_k(n) < p_k \phi(\nu_{n,k})/(1-p_k)^r$. In addition, (A.12) implies that $n \geq n_{\mathrm{est}}$ for any n within the range in (A.28). Thus

$$0 < \eta_3^k < \frac{p_k}{(1-p_k)^r} \sum_{\lceil \beta/p_k \rceil + 1}^{\infty} \phi(\nu_{n,k}) L\left(\frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}} \right). \tag{A.39}$$

Since $\varepsilon_{\mathrm{est}} \leq \Omega/2$,

$$\frac{\Omega}{2\nu_{n,k}} < \frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}} < \frac{3\Omega}{2\nu_{n,k}} < \frac{3\Omega}{2\beta}. \tag{A.40}$$

Thus, assuming $3\Omega/(2\beta) < x_{\mathrm{L}}$, and taking into account that $K < 0$, it stems that for n within the indicated range

$$L\left(\frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}} \right) < M_{\mathrm{L}} \left(\frac{\Omega + \theta_{\mathrm{est},n}}{\nu_{n,k}} \right)^K < M_{\mathrm{L}} \left(\frac{\Omega}{2\nu_{n,k}} \right)^K. \tag{A.41}$$

If it is additionally assumed that $p_k \leq 1/2$, the factor $1/(1-p_k)^r$ in (A.39) cannot exceed 2^r. Therefore

$$0 < \eta_3^k < \frac{2^{r-K} M_{\mathrm{L}} \Omega^K}{(r-1)!} \sum_{\lceil \beta/p_k \rceil + 1}^{\infty} p_k \nu_{n,k}^{r-K-1} \exp(-\nu_{n,k}). \tag{A.42}$$

The sum in (A.42) tends to $\Gamma(r-K, \beta)$ as $k \to \infty$. Thus, given $\varepsilon_{\mathrm{int}}' > 0$, there exists k_{int}' such that for $k \geq k_{\mathrm{int}}'$

$$0 < \sum_{\lceil \beta/p_k \rceil + 1}^{\infty} p_k \nu_{n,k}^{r-K-1} \exp(-\nu_{n,k}) < \Gamma(r-K, \beta) + \varepsilon_{\mathrm{int}}'. \tag{A.43}$$

In addition, since $\Gamma(r-K,\beta)$ is positive and tends to 0 as $\beta \to \infty$, for any $\varepsilon''_{\text{tail}} > 0$ there exists β''_{tail} such that $0 < \Gamma(r-K,\beta) < \varepsilon''_{\text{tail}}$ for $\beta \geq \beta''_{\text{tail}}$. Therefore (A.42) can be written as

$$0 < \eta_3^k < \frac{2^{r-K} M_{\mathrm{L}} \Omega^K}{(r-1)!} (\varepsilon''_{\text{tail}} + \varepsilon'_{\text{int}}). \tag{A.44}$$

To establish that $\eta^k \to \bar{\eta}$, it suffices to show that for any $\varepsilon_0 > 0$, there exists k_0 such that $|\eta^k - \bar{\eta}| < \varepsilon_0$ for all $k \geq k_0$. With the foregoing results, and taking into account the dependencies between the involved parameters, this is accomplished as follows. Given $\varepsilon_0 > 0$, let

$$\varepsilon_{\text{tail}} = \frac{\varepsilon_0}{9}. \tag{A.45}$$

This determines the values α_{tail} and β_{tail}. Likewise, taking

$$\varepsilon'_{\text{tail}} = \frac{\varepsilon_0}{9} \tag{A.46}$$

determines α'_{tail}, and taking $\varepsilon''_{\text{tail}}$ such that

$$\frac{2^{r-K} M_{\mathrm{L}} \Omega^K}{(r-1)!} \varepsilon''_{\text{tail}} = \frac{\varepsilon_0}{9} \tag{A.47}$$

determines β''_{tail}. The values α and β are selected as

$$\alpha = \min\left\{\alpha_{\text{tail}}, \alpha'_{\text{tail}}, \frac{\Omega}{x'_{\mathrm{L}}}\right\}, \tag{A.48}$$

$$\beta = \max\left\{\beta_{\text{tail}}, \beta''_{\text{tail}}, \frac{3\Omega}{2x_{\mathrm{L}}}\right\}. \tag{A.49}$$

(Note that, since $\beta_{\text{tail}} > \alpha_{\text{tail}}$, (A.48) and (A.49) imply that $\beta > \alpha$.) From α and β, the intervals I and I' are obtained, and the values m_ϕ, d and J can be computed. Taking

$$\varepsilon_{\text{int}} = \frac{\varepsilon_0}{9} \tag{A.50}$$

determines k_{int}. The parameter $\varepsilon_{\text{unif}}$ is selected such that

$$\left(\bar{\eta} + \frac{4\varepsilon_0}{9}\right) \frac{\varepsilon_{\text{unif}}}{m_\phi} = \frac{\varepsilon_0}{9}, \tag{A.51}$$

which determines k_{unif}. Next, $\varepsilon_{\text{cont}}$ is chosen such that

$$(\beta - \alpha + 3)\varepsilon_{\text{cont}} = \frac{\varepsilon_0}{9}, \tag{A.52}$$

from which δ_{cont} is obtained. Taking $\varepsilon_{\text{disc}}$ as

$$\varepsilon_{\text{disc}} = \frac{\varepsilon_{\text{cont}}}{J} \tag{A.53}$$

determines δ_{disc} and k_{disc}. Choosing any ε_{est} smaller than $\min\{\Omega/2, \delta_{\text{cont}}, \delta_{\text{disc}}\}$ determines n_{est}, from which m_{g} and M'_{g} can be obtained. Let k_{est} be such that for all $k \geq k_{\text{est}}$

$$\frac{n_{\text{est}}^r \max\{M'_{\text{L}} M_{\text{g}}^{K'}, M'_{\text{g}}\}}{(r-1)!} p_k^{r-K'} < \frac{\varepsilon_0}{9}. \tag{A.54}$$

Let k'_{est} be chosen such that (A.12) holds for all $k \geq k'_{\text{est}}$, and k_{interv} such that (A.9) holds for all $k \geq k_{\text{interv}}$. The parameter $\varepsilon'_{\text{int}}$ is chosen as

$$\varepsilon'_{\text{int}} = \varepsilon''_{\text{tail}}, \tag{A.55}$$

which determines k'_{int}. Finally, let k_{const} be such that $p_k \leq 1/2$ for all $k \geq k_{\text{const}}$. Taking $k_0 = \max\{k_{\text{int}}, k'_{\text{int}}, k_{\text{unif}}, k_{\text{est}}, k'_{\text{est}}, k_{\text{disc}}, k_{\text{interv}}, k_{\text{const}}\}$, the following inequalities are obtained for $k \geq k_0$. From (A.25), (A.45) and (A.50)–(A.53),

$$|\eta_0^k - \bar{\eta}| < \frac{4\varepsilon_0}{9} + \left(\bar{\eta} + \frac{4\varepsilon_0}{9}\right) \frac{\varepsilon_{\text{unif}}}{m_\phi} = \frac{5\varepsilon_0}{9}; \tag{A.56}$$

from (A.32) and (A.54),

$$0 \leq \eta_1^k < \frac{\varepsilon_0}{9}; \tag{A.57}$$

from (A.38) and (A.46),

$$0 \leq \eta_2^k < \frac{\varepsilon_0}{9}; \tag{A.58}$$

and from (A.44), (A.47) and (A.55),

$$0 < \eta_3^k < \frac{2\varepsilon_0}{9}. \tag{A.59}$$

Inequalities (A.56)–(A.59) imply that $|\eta^k - \bar{\eta}| < \varepsilon_0$ for all $k \geq k_0$, which concludes the proof. □

Proof of Proposition 1. By Assumption 2', let D be the number of discontinuities of L, occurring at points $x_1 < x_2 < \cdots < x_D$. The asymptotic risk $\bar{\eta}$ can be expressed as $\sum_{i=0}^{D} \bar{\eta}_i$ with

$$\bar{\eta}_0 = \int_0^{x_1} \psi(x,\Omega) L(x)\,\mathrm{d}x, \tag{A.60}$$

$$\bar{\eta}_i = \int_{x_i}^{x_{i+1}} \psi(x,\Omega) L(x)\,\mathrm{d}x, \quad i = 1, \ldots, D-1, \tag{A.61}$$

$$\bar{\eta}_D = \int_{x_D}^{\infty} \psi(x,\Omega) L(x)\,\mathrm{d}x. \tag{A.62}$$

Given $i = 1, \ldots, D-1$, let $L_i(x)$ be defined for $x \in [x_i, x_{i+1}]$ as

$$L_i(x) = \begin{cases} L(x), & x_i < x < x_{i+1}, \\ L(x_i+), & x = x_i, \\ L(x_{i+1}-), & x = x_{i+1}, \end{cases} \tag{A.63}$$

and let $T_i(x, \Omega)$ be defined for $x \in [x_i, x_{i+1}]$, $\Omega \in \mathbb{R}^+$ as $T_i(x, \Omega) = \psi(x, \Omega) L_i(x)$. Clearly, the integral in (A.61) does not change if $\psi(x, \Omega) L(x)$ is replaced by $T_i(x, \Omega)$. The function T_i is continuous on $[x_i, x_{i+1}] \times \mathbb{R}^+$, because it is the product of continuous functions. The function $\partial T_i / \partial \Omega$ is similarly seen to be continuous. This implies (Fleming, 1977, corollary to theorem 5.9) that $\bar{\eta}_i$ given by (A.61) is a C^1 function of Ω, with

$$\frac{\mathrm{d}\bar{\eta}_i}{\mathrm{d}\Omega} = \int_{x_i}^{x_{i+1}} \frac{\partial T_i(x, \Omega)}{\partial \Omega} \, \mathrm{d}x = \int_{x_i}^{x_{i+1}} \frac{\partial \psi(x, \Omega)}{\partial \Omega} L(x) \, \mathrm{d}x. \tag{A.64}$$

Regarding $\bar{\eta}_0$, let $T_0(x, \Omega) = \psi(x, \Omega) L(x)$ for $x \in (0, x_{i+1}]$, $\Omega \in \mathbb{R}^+$, and $T_0(0, \Omega) = 0$. It is clear that T_0 is continuous on $(0, x_1] \times \mathbb{R}^+$. In addition, its continuity at any point of the form $(0, \Omega_0)$ can be established as follows. Let Δ be any value such that $0 < \Delta < \Omega_0$. For $\Omega \in (\Omega_0 - \Delta, \Omega_0 + \Delta)$ and $x > 0$, T_0 is bounded as

$$0 \leq T_0(x, \Omega) < \frac{(\Omega_0 + \Delta)^r \exp(-(\Omega_0 - \Delta)/x) L(x)}{x^{r+1}(r-1)!}. \tag{A.65}$$

Property 1 in Assumption 3 implies that the right-hand side of (A.65) tends to 0 as $x \to 0$. Thus there exists $\delta > 0$ such that $0 \leq T_0(x, \Omega) < \varepsilon$ for $0 \leq x < \delta$, $|\Omega - \Omega_0| < \Delta$. This shows that T_0 is continuous at $(0, \Omega_0)$, and thus on $[0, x_1] \times \mathbb{R}^+$. Using analogous arguments, $\partial T_0 / \partial \Omega$ can also be seen to be continuous on $[0, x_1] \times \mathbb{R}^+$. This implies that $\bar{\eta}_0$ is a C^1 function of Ω, and (A.64) holds for $i = 0$ if the lower integration limit is replaced by 0.

As for $\bar{\eta}_D$, let $T(x, \Omega) = \psi(x, \Omega) L(x)$, and consider the function $T(x, \Omega)/\Omega^r$. This function and its partial derivative with respect to Ω are continuous on $(x_D, \infty) \times \mathbb{R}^+$, and satisfy the following bounds:

$$0 < \frac{T(x, \Omega)}{\Omega^r} < \frac{L(x)}{x^{r+1}(r-1)!}, \tag{A.66}$$

$$0 > \frac{\partial (T(x, \Omega)/\Omega^r)}{\partial \Omega} = -\frac{\exp(-\Omega/x) L(x)}{x^{r+2}(r-1)!} > -\frac{L(x)}{x^{r+2}(r-1)!}. \tag{A.67}$$

The right-most parts of (A.66) and (A.67) are integrable on (x_D, ∞), because of property 2 in Assumption 3. This implies (Fleming, 1977, theorem 5.9) that $\bar{\eta}_D / \Omega^r$

is a C^1 function of Ω, and therefore so is $\bar{\eta}_D$; in addition, $\mathrm{d}\bar{\eta}_D/\mathrm{d}\Omega$ satisfies an expression analogous to (A.64) with the integration interval replaced by (x_D, ∞).

The preceding results assure that $\mathrm{d}\bar{\eta}/\mathrm{d}\Omega = \sum_{i=0}^{D} \mathrm{d}\bar{\eta}_i/\mathrm{d}\Omega$ is continuous and can be expressed as in (12). The equality (13) readily follows from (6), (11) and (12). □

Lemma 3. *For any $a, c \in \mathbb{R}^+$, $b \in \mathbb{R}$,*

$$\frac{\mathrm{d}}{\mathrm{d}\Omega} \int_a^\infty \frac{\Omega^b \exp(-\Omega/x)}{x^{c+1}} \,\mathrm{d}x = \frac{\Omega^{b-1} \exp(-\Omega/a)}{a^c} + (b-c)\Omega^{b-c-1} \gamma\left(c, \frac{\Omega}{a}\right). \tag{A.68}$$

Proof. Applying the change of variable $x = \Omega/\nu$, the integral in (A.68) can be expressed as

$$\int_a^\infty \frac{\Omega^b \exp(-\Omega/x)}{x^{c+1}} \,\mathrm{d}x = \int_0^{\Omega/a} \Omega^{b-c} \nu^{c-1} \exp(-\nu) \,\mathrm{d}\nu = \Omega^{b-c} \left(c, \frac{\Omega}{a}\right), \tag{A.69}$$

from which (A.68) follows. □

Lemma 4. *For $s \in \mathbb{R}$,*

$$\lim_{u \to 0} \frac{\gamma(s,u)}{u^s} = \frac{1}{s}, \tag{A.70}$$

$$\lim_{u \to \infty} \frac{\Gamma(s,u)}{u^{s-1} \exp(-u)} = 1. \tag{A.71}$$

Proof. These equalities respectively follow from Abramowitz and Stegun (1970, equation 6.5.29) and Abramowitz and Stegun (1970, equation 6.5.32). □

Lemma 5. *The upper incomplete gamma function (3) satisfies for $s, w \in \mathbb{N}$, $\nu \in \mathbb{R}^+$*

$$\Gamma(s,\nu) = \begin{cases} \displaystyle\sum_{k=0}^{s-1} (s-1)^{(s-k-1)} \nu^k \exp(-\nu), & s \geq 1, \\ \displaystyle\sum_{k=s-w}^{s-1} (s-1)^{(s-k-1)} \nu^k \exp(-\nu) + W(\nu), & s \leq 0, \end{cases} \tag{A.72}$$

where $W(\nu)$ is $O(\nu^{s-w-1} \exp(-\nu))$ as $\nu \to \infty$.

Proof. The expression for $s \geq 1$ is equivalent to Abramowitz and Stegun (1970, equation 6.5.13).

For $s \leq 0$, the stated result follows from recursively using the identity (Abramowitz and Stegun, 1970, equation 6.5.21)

$$\Gamma(s,\nu) = (s-1)\Gamma(s-1,\nu) + \nu^{s-1}\exp(-\nu) \tag{A.73}$$

w times and taking into account the equality (A.71) from Lemma 4. □

Lemma 6. *For $t \in \mathbb{N}$, $u \in \mathbb{Z}$,*

$$\sum_{j=0}^{t} \binom{t}{j} j(u-j)^{(i-1)}(-1)^{t-j} = \begin{cases} 0, & i = 1,\dots,t-1, \\ (-1)^{t-1}t!, & i = t. \end{cases} \tag{A.74}$$

Proof. The equality

$$\sum_{j=0}^{t} \binom{t}{j} j^{(k)}(-1)^{t-j} = \begin{cases} 0, & k \neq t, \\ t!, & k = t. \end{cases} \tag{A.75}$$

is easily shown to hold for $k \in \mathbb{N}$ by applying the binomial theorem to $(x-1)^t$, differentiating k times and particularizing for $x = 1$. The term $j(u-j)^{(i-1)}$ in (A.74) can be expressed as $\sum_{k=1}^{i} a_k j^{(k)}$ for appropriate values of the coefficients a_k; furthermore, it is easily seen that a_i equals $(-1)^{i-1}$. Thus

$$\sum_{j=0}^{t} \binom{t}{j} j(u-j)^{(i-1)}(-1)^{t-j} = \sum_{k=1}^{i} \sum_{j=0}^{t} a_k \binom{t}{j} j^{(k)}(-1)^{t-j}. \tag{A.76}$$

If $i \leq t-1$, the inner sum in (A.76) equals 0 for all k within the range specified in the outer sum, because of (A.75). If $i = t$, all values of the index k give a null inner sum except $k = t$, which gives $a_t t! = (-1)^{t-1}t!$. This establishes (A.74). □

Proof of Proposition 2. Assume that (17) holds. Let $\varepsilon = -Bs/(4r)$, which is positive for the allowed values of B and s. From (17), there exists δ such that $|L(x) - A - Bx^s| < \varepsilon x^s$ for all $x \in (0,\delta)$. This implies that for any $\xi \in (0,\delta)$, and for $\xi \leq x < \delta$,

$$L(\xi) - L(x) > B(\xi^s - x^s) - 2\varepsilon x^s = B\xi^s - (B+2\varepsilon)x^s. \tag{A.77}$$

Therefore

$$
\begin{aligned}
&\int_{\xi}^{\infty} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x \\
&= \int_{\xi}^{\delta} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x + \int_{\delta}^{\infty} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x \\
&> B\xi^{s} \int_{\xi}^{\delta} \frac{\mathrm{d}x}{x^{r+1}} - (B+2\varepsilon) \int_{\xi}^{\delta} \frac{\mathrm{d}x}{x^{r-s+1}} + \int_{\delta}^{\infty} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x \\
&= \frac{B}{r\xi^{r-s}} - \frac{B\xi^{s}}{r\delta^{r}} + \frac{B+2\varepsilon}{(r-s)\delta^{r-s}} - \frac{B+2\varepsilon}{(r-s)\xi^{r-s}} + \int_{\delta}^{\infty} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x \\
&> \frac{B}{r\xi^{r-s}} - \frac{B}{r\delta^{r-s}} + \frac{B+2\varepsilon}{(r-s)\delta^{r-s}} - \frac{B+2\varepsilon}{(r-s)\xi^{r-s}} + \int_{\delta}^{\infty} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x
\end{aligned}
\tag{A.78}
$$

Denoting by C the sum of the terms in the right-hand side of (A.78) which do not depend on ξ, i.e. the second, third and fifth, and substituting the value of ε,

$$
\int_{\xi}^{\infty} \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x > -\frac{Bs}{2r(r-s)\xi^{r-s}} + C. \tag{A.79}
$$

Taking into account that $-Bs$ and $r-s$ are positive, and that C is independent of ξ, from (A.79) it is seen that there exists $\xi \in (0,\delta)$ such that (14) holds. □

Lemma 7. *Under the hypotheses of Theorem 2, there exists* Ω_0 *such that* $\mathrm{d}\bar{\eta}/\mathrm{d}\Omega < 0$ *for all* $\Omega \le \Omega_0$.

Proof. Let ξ be as in property 1 in Assumption 4. Since L is non-increasing for all x smaller than ξ, the function ℓ defined as

$$
\ell(x) = \begin{cases} L(x) - L(\xi) & \text{for } 0 < x < \xi \\ 0 & \text{for } x \ge \xi \end{cases} \tag{A.80}
$$

is non-negative and non-increasing. From (10) and (11), $\bar{\eta}$ can be expressed as $\zeta_0 + \zeta_1 + \zeta_2$ with

$$
\zeta_0 = \int_{\Omega/\xi}^{\infty} \phi(\nu) L(\xi)\,\mathrm{d}\nu, \tag{A.81}
$$

$$
\zeta_1 = \int_{\Omega/\xi}^{\infty} \phi(\nu) \ell(\Omega/\nu)\,\mathrm{d}\nu, \tag{A.82}
$$

$$
\zeta_2 = \int_{\xi}^{\infty} \psi(x,\Omega) L(x)\,\mathrm{d}x. \tag{A.83}
$$

Each of these terms can be interpreted as the risk associated with a certain loss function for which Proposition 1 applies.

Since ℓ is non-negative and non-increasing, for ν fixed the integrand in (A.82) is a non-negative, non-increasing function of Ω. This implies that ζ_1 is a non-increasing function of Ω, and thus $\mathrm{d}\zeta_1/\mathrm{d}\Omega \leq 0$.

Regarding the term ζ_0,

$$\frac{\mathrm{d}\zeta_0}{\mathrm{d}\Omega} = -\frac{\Omega^{r-1}\exp(-\Omega/\xi)L(\xi)}{\xi^r(r-1)!}, \tag{A.84}$$

which implies that

$$\lim_{\Omega\to 0}\frac{\mathrm{d}\zeta_0/\mathrm{d}\Omega}{\Omega^{r-1}} = -\frac{L(\xi)}{\xi^r(r-1)!}. \tag{A.85}$$

As for ζ_2, from (A.83) it follows that

$$\frac{\mathrm{d}\zeta_2}{\mathrm{d}\Omega} = \frac{r\Omega^{r-1}}{(r-1)!}\int_\xi^\infty \frac{\exp(-\Omega/x)L(x)}{x^{r+1}}\,\mathrm{d}x - \frac{\Omega^r}{(r-1)!}\int_\xi^\infty \frac{\exp(-\Omega/x)L(x)}{x^{r+2}}\,\mathrm{d}x. \tag{A.86}$$

Interpreting the integrals in (A.86) as Lebesgue integrals, and noting that $\exp(-\Omega/x) < 1$ for $\Omega, x \in \mathbb{R}^+$, Lebesgue's dominated convergence theorem (Apostol, 1974, theorem 10.27) assures that

$$\lim_{\Omega\to 0}\int_\xi^\infty \frac{\exp(-\Omega/x)L(x)}{x^{r+1}}\,\mathrm{d}x = \int_\xi^\infty \frac{L(x)}{x^{r+1}}\,\mathrm{d}x, \tag{A.87}$$

and similarly for the second integral. This implies that the first term in the right-hand side of (A.86) dominates the second for Ω asymptotically small, i.e.

$$\lim_{\Omega\to 0}\frac{\mathrm{d}\zeta_2/\mathrm{d}\Omega}{\Omega^{r-1}} = \frac{r}{(r-1)!}\int_\xi^\infty \frac{L(x)}{x^{r+1}}\,\mathrm{d}x. \tag{A.88}$$

From (A.85) and (A.88),

$$\begin{aligned}\lim_{\Omega\to 0}\frac{\mathrm{d}(\zeta_0+\zeta_2)/\mathrm{d}\Omega}{\Omega^{r-1}} &= -\frac{L(\xi)}{\xi^r(r-1)!} + \frac{r}{(r-1)!}\int_\xi^\infty \frac{L(x)}{x^{r+1}}\,\mathrm{d}x \\ &= -\frac{r}{(r-1)!}\int_\xi^\infty \frac{L(\xi)-L(x)}{x^{r+1}}\,\mathrm{d}x.\end{aligned} \tag{A.89}$$

Combining (A.89) with the inequality (14) from Assumption 4, the limit on the right-hand side of (A.89) is seen to be negative. This implies that there exists Ω_0 such that $\mathrm{d}(\zeta_0+\zeta_2)/\mathrm{d}\Omega < 0$ for $\Omega \leq \Omega_0$. Taking into account that $\mathrm{d}\zeta_1/\mathrm{d}\Omega \leq 0$, it follows that $\mathrm{d}\bar{\eta}/\mathrm{d}\Omega < 0$ for $\Omega \leq \Omega_0$. □

Lemma 8. *Under the hypotheses of Theorem 2, there exists* Ω_0' *such that* $\mathrm{d}\bar{\eta}/\mathrm{d}\Omega > 0$ *for all* $\Omega \geq \Omega_0'$.

Proof. If condition (a) of property 2 in Assumption 4 holds, let H be chosen such that $0 < H < L(\xi'+) - L(\xi'-)$. By definition of $L(\xi'-)$, there exists h such that $L(x) \in (L(\xi'-) - H, L(\xi'-) + H)$ for all $x \in (\xi' - h, \xi')$. If condition (b) holds, it stems that there exists h such that $(-1)^{t-1}\mathrm{d}^t L/\mathrm{d}x^t$ is positive and continuous for $x \in (\xi' - h, \xi')$. Thus, let h be selected as has been indicated.

From property 1 in Assumption 3, there exist $K \in \mathbb{R}$, M_{L} and $x_{\mathrm{L}} < \xi' - h$ such that

$$L(x) < M_{\mathrm{L}} x^K \quad \text{for } x < x_{\mathrm{L}}. \tag{A.90}$$

The asymptotic risk $\bar{\eta}$ can be expressed from (10) and (11) as $\zeta_0' + \zeta_1' + \zeta_2' + \zeta_3' + \zeta_4'$ with

$$\zeta_0' = \int_{\xi'-h}^{\xi'} \psi(x,\Omega) L(x)\,\mathrm{d}x, \tag{A.91}$$

$$\zeta_1' = \int_0^{x_{\mathrm{L}}} \psi(x,\Omega) L(x)\,\mathrm{d}x, \tag{A.92}$$

$$\zeta_2' = \int_{x_{\mathrm{L}}}^{\xi'-h} \psi(x,\Omega) L(x)\,\mathrm{d}x, \tag{A.93}$$

$$\zeta_3' = \int_{\xi'}^{\infty} \psi(x,\Omega) L(\xi'+)\,\mathrm{d}x, \tag{A.94}$$

$$\zeta_4' = \int_0^{\Omega/\xi'} \phi(\nu)\big(L(\Omega/\nu) - L(\xi'+)\big)\,\mathrm{d}\nu. \tag{A.95}$$

Each of these terms corresponds to the risk associated with a certain loss function which satisfies Proposition 1.

By property 2 of Assumption 4, $L(x) - L(\xi'+)$ is non-negative and non-decreasing for $x > \xi'$. An argument analogous to that used for ζ_1 in Lemma 7 shows that the term ζ_4' given by (A.95) is non-decreasing with Ω, and thus

$$\frac{\mathrm{d}\zeta_4'}{\mathrm{d}\Omega} \geq 0. \tag{A.96}$$

According to Lemma 3, $\mathrm{d}\zeta_3'/\mathrm{d}\Omega$ is given by

$$\frac{\mathrm{d}\zeta_3'}{\mathrm{d}\Omega} = \frac{L(\xi'+)\Omega^{r-1}\exp(-\Omega/\xi')}{\xi'^r (r-1)!}. \tag{A.97}$$

Computing

$$\frac{\mathrm{d}\zeta_1'}{\mathrm{d}\Omega} = \int_0^{x_{\mathrm{L}}} \frac{r\Omega^{r-1}\exp(-\Omega/x)}{x^{r+1}(r-1)!} L(x)\,\mathrm{d}x - \int_0^{x_{\mathrm{L}}} \frac{\Omega^r \exp(-\Omega/x)}{x^{r+2}(r-1)!} L(x)\,\mathrm{d}x \tag{A.98}$$

and using (A.90) it stems that

$$\left|\frac{\mathrm{d}\zeta_1'}{\mathrm{d}\Omega}\right| \leq \frac{M_\mathrm{L} r \Omega^{r-1}}{(r-1)!}\int_0^{x_\mathrm{L}} \frac{\exp(-\Omega/x)}{x^{r+1-K}}\,\mathrm{d}x + \frac{M_\mathrm{L}\Omega^r}{(r-1)!}\int_0^{x_\mathrm{L}} \frac{\exp(-\Omega/x)}{x^{r+2-K}}\,\mathrm{d}x. \quad \text{(A.99)}$$

The integrals in (A.99) can be bounded as follows. Let $\lambda = (x_\mathrm{L} + \xi' - h)/(2(\xi' - h))$. It is seen that λ and $1-\lambda$ are lower than 1. Let the function $v_1 : \mathbb{R}^+ \cup \{0\} \mapsto \mathbb{R} \cup \{0\}$ be defined as $v_1(x) = \exp(-\lambda\Omega/x)$ for $x > 0$ and $v_1(0) = 0$. Since $\exp(-\lambda\Omega/x) \to 0$ as $x \to 0$, v_1 is continuous on $[0, x_\mathrm{L}]$. In addition, the function $v_2 : \mathbb{R} \cup \{0\} \mapsto \mathbb{R} \cup \{0\}$ such that

$$v_2(x) = \frac{\exp(-(1-\lambda)\Omega/x)}{x^{r+1-K}} \quad \text{(A.100)}$$

for $x > 0$ and $v_2(0) = 0$ is non-negative and integrable on $[0, x_\mathrm{L}]$. Thus, the mean value theorem (Fleming, 1977, p. 190) can be applied to the first integral in (A.99) to yield:

$$\int_0^{x_\mathrm{L}} \frac{\exp(-\Omega/x)}{x^{r+1-K}}\,\mathrm{d}x = \int_0^{x_\mathrm{L}} v_1(x)v_2(x)\,\mathrm{d}x = v_1(x_\mathrm{m})\int_0^{x_\mathrm{L}} v_2(x)\,\mathrm{d}x \quad \text{(A.101)}$$

for some $x_\mathrm{m} \in [0, x_\mathrm{L}]$. Actually, x_m cannot be 0, because that would give 0 in the right-hand side of (A.101), whereas the left-hand side is greater than 0. Thus $x_\mathrm{m} \in (0, x_\mathrm{L}]$. Similar arguments can be applied to the last integral in (A.101) to obtain

$$\int_0^{x_\mathrm{L}} \frac{\exp(-(1-\lambda)\Omega/x)}{x^{r+1-K}}\,\mathrm{d}x = x_\mathrm{L}\frac{\exp(-(1-\lambda)\Omega/x_\mathrm{m}')}{{x_\mathrm{m}'}^{r+1-K}} \quad \text{(A.102)}$$

with $x_\mathrm{m}' \in (0, x_\mathrm{L}]$. Maximizing the right-hand side of (A.102) with respect to $x_\mathrm{m}' \in \mathbb{R}^+$ gives

$$\int_0^{x_\mathrm{L}} \frac{\exp(-(1-\lambda)\Omega/x)}{x^{r+1-K}}\,\mathrm{d}x \leq x_\mathrm{L}\left(\frac{r+1-K}{1-\lambda}\right)^{r+1-K}\frac{\exp(-(r+1-K))}{\Omega^{r+1-K}}. \quad \text{(A.103)}$$

Combining (A.101) and (A.103),

$$\int_0^{x_\mathrm{L}} \frac{\exp(-\Omega/x)}{x^{r+1-K}}\,\mathrm{d}x \leq \frac{x_\mathrm{L}(r+1-K)^{r+1-K}\exp(-(r+1-K+\lambda\Omega/x_\mathrm{m}))}{((1-\lambda)\Omega)^{r+1-K}}. \quad \text{(A.104)}$$

The second integral in (A.99) is bounded analogously:

$$\int_0^{x_\mathrm{L}} \frac{\exp(-\Omega/x)}{x^{r+2-K}}\,\mathrm{d}x \leq \frac{x_\mathrm{L}(r+2-K)^{r+2-K}\exp(-(r+2-K+\lambda\Omega/x_\mathrm{m}''))}{((1-\lambda)\Omega)^{r+2-K}}. \quad \text{(A.105)}$$

with $x''_{\mathrm{m}} \in (0, x_{\mathrm{L}}]$. From (A.99), (A.104) and (A.105),

$$\left|\frac{\mathrm{d}\zeta_1'}{\mathrm{d}\Omega}\right| \leq \frac{M_{\mathrm{L}} x_{\mathrm{L}} r (r+1-K)^{r+1-K} \Omega^{K-2} \exp(-(r+1-K+\lambda\Omega/x_{\mathrm{m}}))}{(1-\lambda)^{r+1-K}(r-1)!} + \frac{M_{\mathrm{L}} x_{\mathrm{L}} (r+2-K)^{r+2-K} \Omega^{K-2} \exp(-(r+2-K+\lambda\Omega/x''_{\mathrm{m}}))}{(1-\lambda)^{r+2-K}(r-1)!}. \tag{A.106}$$

It is easily seen that $x_{\mathrm{m}}/\lambda, x''_{\mathrm{m}}/\lambda < x_{\mathrm{L}} < \xi' - h$. It thus follows from (A.106) that

$$\left|\frac{\mathrm{d}\zeta_1'}{\mathrm{d}\Omega}\right| < Q\Omega^{K-2} \exp(-\Omega/(\xi'-h)) \tag{A.107}$$

where Q is independent of Ω.

For $\mathrm{d}\zeta_2'/\mathrm{d}\Omega$, by Assumption 1, let M be an upper bound of L in the interval $(x_{\mathrm{L}}, \xi'-h)$. An argument based on the mean value theorem can also be applied here; in fact, it is slightly simpler than in the preceding paragraph because in this case the lower integration limit is greater than 0:

$$\begin{aligned}\left|\frac{\mathrm{d}\zeta_2'}{\mathrm{d}\Omega}\right| &\leq \frac{Mr\Omega^{r-1}}{(r-1)!}\int_{x_{\mathrm{L}}}^{\xi'-h} \frac{\exp(-\Omega/x)}{x^{r+1}}\,\mathrm{d}x + \frac{M\Omega^r}{(r-1)!}\int_{x_{\mathrm{L}}}^{\xi'-h} \frac{\exp(-\Omega/x)}{x^{r+2}}\,\mathrm{d}x \\ &= \frac{M\Omega^{r-1}(\xi'-h-x_{\mathrm{L}})}{(r-1)!}\left(\frac{r\exp(-\Omega/x'''_{\mathrm{m}})}{x'''^{\,r+1}_{\mathrm{m}}} + \frac{\Omega\exp(-\Omega/x''''_{\mathrm{m}})}{x''''^{\,r+2}_{\mathrm{m}}}\right)\end{aligned} \tag{A.108}$$

with $x'''_{\mathrm{m}}, x''''_{\mathrm{m}} \in [x_{\mathrm{L}}, \xi'-h]$. Therefore

$$\left|\frac{\mathrm{d}\zeta_2'}{\mathrm{d}\Omega}\right| < \frac{M(\xi'-h-x_{\mathrm{L}})}{x_{\mathrm{L}}{}^{r+1}(r-1)!}\left(r+\frac{\Omega}{x_{\mathrm{L}}}\right)\Omega^{r-1}\exp(-\Omega/(\xi'-h)). \tag{A.109}$$

To compute the derivative of ζ_0', it is necessary to distinguish cases (a) and (b) of property 2 in Assumption 4. In case (a), since $L(x) \in (L(\xi'-)-H, L(\xi'-)+H)$ for all $x \in (\xi'-h, \xi')$, the mean value theorem assures that there is some $\theta \in [L(\xi'-)-H, L(\xi'-)+H]$ such that

$$\begin{aligned}\frac{\mathrm{d}\zeta_0'}{\mathrm{d}\Omega} &= \int_{\xi'-h}^{\xi'} \frac{\partial\psi(x,\Omega)}{\partial\Omega} L(x)\,\mathrm{d}x = \theta\int_{\xi'-h}^{\xi'} \frac{\partial\psi(x,\Omega)}{\partial\Omega}\,\mathrm{d}x \\ &= \frac{\theta}{(r-1)!}\frac{\mathrm{d}}{\mathrm{d}\Omega}\int_{\xi'-h}^{\xi'} \frac{\Omega^r\exp(-\Omega/x)}{x^{r+1}}\,\mathrm{d}x.\end{aligned} \tag{A.110}$$

Applying Lemma 3,

$$\frac{\mathrm{d}\zeta_0'}{\mathrm{d}\Omega} = \frac{\theta\Omega^{r-1}}{(r-1)!}\left(-\frac{\exp(-\Omega/\xi')}{\xi'^r} + \frac{\exp(-\Omega/(\xi'-h))}{(\xi'-h)^r}\right). \tag{A.111}$$

Using (A.96), (A.97), (A.107), (A.109) and (A.111),

$$\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega} \geq \frac{(L(\xi'+)-\theta)\Omega^{r-1}\exp(-\Omega/\xi')}{\xi'^{r}(r-1)!} + O\left(\Omega^{q}\exp(-\Omega/(\xi'-h))\right) \tag{A.112}$$

with $q = \max\{r, K-2\}$. Since $h > 0$ and $\theta \leq L(\xi'-) + H < L(\xi'+)$, from (A.112) it follows that

$$\lim_{\Omega\to\infty}\left(\frac{\exp(\Omega/\xi')}{\Omega^{r-1}}\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega}\right) \geq \frac{L(\xi'+)-\theta}{\xi'^{r}(r-1)!} > 0. \tag{A.113}$$

In case (b), since $\mathrm{d}^{t}L/\mathrm{d}x^{t}$ is continuous on $(\xi'-h, \xi')$, Taylor's theorem (Apostol, 1967, volume 1, theorem 7.6) can be applied to express $L(x)$ for $x \in (\xi'-h, \xi')$ as

$$L(x) = L(\xi'-) + \frac{\theta'(x-\xi')^{t}}{t!} = L(\xi'-) + \frac{\theta'}{t!}\sum_{j=0}^{t}\binom{t}{j}(-\xi')^{t-j}x^{j} \tag{A.114}$$

where θ' is the value of $\mathrm{d}^{t}L/\mathrm{d}x^{t}$ at some point within the interval $(\xi'-h, \xi')$. The choice of h assures that $(-1)^{t-1}\theta'$ is positive. Substituting (A.114) into (A.91), differentiating and making use of Lemma 3 and (4) gives

$$\begin{aligned}\frac{\mathrm{d}\zeta_0'}{\mathrm{d}\Omega} = {} & \frac{L(\xi'-)}{\Omega(r-1)!}\left[-\left(\frac{\Omega}{\xi'}\right)^{r}\exp(-\Omega/\xi') + \left(\frac{\Omega}{\xi'-h}\right)^{r}\exp(-\Omega/(\xi'-h))\right] \\ & + \frac{\theta'}{(r-1)!\,t!}\sum_{j=0}^{t}\binom{t}{j}(-\xi')^{t-j}\Omega^{j-1}\left[-\left(\frac{\Omega}{\xi'}\right)^{r-j}\exp(-\Omega/\xi')\right. \\ & + \left(\frac{\Omega}{\xi'-h}\right)^{r-j}\exp(-\Omega/(\xi'-h)) \\ & \left. + j\left(\Gamma\left(r-j, \frac{\Omega}{\xi'}\right) - \Gamma\left(r-j, \frac{\Omega}{\xi'-h}\right)\right)\right].\end{aligned} \tag{A.115}$$

The identity $\sum_{j=0}^{t}\binom{t}{j}(-1)^{t-j} = 0$ implies that

$$\sum_{j=0}^{t}\binom{t}{j}(-\xi')^{t-j}\Omega^{j-1}\left(\frac{\Omega}{\xi'}\right)^{r-j} = \xi'^{t-r}\Omega^{r-1}\sum_{j=0}^{t}\binom{t}{j}(-1)^{t-j} = 0, \tag{A.116}$$

and thus (A.115) simplifies to

$$\frac{\mathrm{d}\zeta_0'}{\mathrm{d}\Omega} = \frac{L(\xi'-)\Omega^{r-1}}{(r-1)!}\left(-\frac{\exp(-\Omega/\xi')}{\xi'^r} + \frac{\exp(-\Omega/(\xi'-h))}{(\xi'-h)^r}\right)$$
$$+\frac{\Omega^{r-1}\theta'\exp(-\Omega/(\xi'-h))}{(r-1)!t!}\sum_{j=0}^{t}\binom{t}{j}\frac{(-\xi')^{t-j}}{(\xi'-h)^{r-j}}$$
$$+\frac{\theta'}{(r-1)!t!}\sum_{j=0}^{t}\binom{t}{j}j(-\xi')^{t-j}\Omega^{j-1}\left[\Gamma\left(r-j,\frac{\Omega}{\xi'}\right)-\Gamma\left(r-j,\frac{\Omega}{\xi'-h}\right)\right]. \tag{A.117}$$

From Lemma 5, $\Omega^{j-1}\Gamma(r-j,\Omega/\xi')$ for $j \leq r-1$ is given by

$$\Omega^{j-1}\Gamma\left(r-j,\frac{\Omega}{\xi'}\right) = \exp(-\Omega/\xi')\sum_{k=j}^{r-1}\frac{(r-j-1)^{(r-k-1)}\Omega^{k-1}}{\xi'^{k-j}}, \tag{A.118}$$

whereas for $j \geq r$ and for any $w \in \mathbb{N}$

$$\Omega^{j-1}\Gamma\left(r-j,\frac{\Omega}{\xi'}\right) = \exp(-\Omega/\xi')\sum_{k=r-w}^{r-1}\frac{(r-j-1)^{(r-k-1)}\Omega^{k-1}}{\xi'^{k-j}} + O\left(\Omega^{r-w-2}\exp(-\Omega/\xi')\right). \tag{A.119}$$

Replacing ξ' by $\xi'-h$ in (A.118) and (A.119) it is seen that

$$\Omega^{j-1}\Gamma\left(r-j,\frac{\Omega}{\xi'-h}\right) = O\left(\Omega^{r-2}\exp(-\Omega/(\xi'-h))\right). \tag{A.120}$$

Setting $w = t$ in (A.119) and substituting (A.118)–(A.120) into (A.117) yields

$$\frac{\mathrm{d}\zeta_0'}{\mathrm{d}\Omega} = -\frac{L(\xi'-)\Omega^{r-1}\exp(-\Omega/\xi')}{\xi'^r(r-1)!}$$
$$+\frac{\theta'\exp(-\Omega/\xi')}{(r-1)!t!}\left[\sum_{j=0}^{\min\{t,r-1\}}\binom{t}{j}j(-\xi')^{t-j}\sum_{k=j}^{r-1}\frac{(r-j-1)^{(r-k-1)}\Omega^{k-1}}{\xi'^{k-j}}\right.$$
$$\left.+\sum_{j=r}^{t}\binom{t}{j}j(-\xi')^{t-j}\sum_{k=r-t}^{r-1}\frac{(r-j-1)^{(r-k-1)}\Omega^{k-1}}{\xi'^{k-j}}\right]$$
$$+O\left(\Omega^{r-t-2}\exp(-\Omega/\xi')\right) \tag{A.121}$$

(the term $O(\Omega^{r-t-2}\exp(-\Omega/\xi'))$ could be substituted by a lower-order term if $t < r$, but this is unnecessary for the proof). Since $(r-j-1)^{(r-k-1)} = 0$ for $k < j < r$, the summation range of the first sum over k in (A.121) can be extended from $k = j,\ldots,r-1$ to $k = \min\{0, r-t\},\ldots,r-1$. On the other hand, the second sum over j is empty if $t < r$. Thus the second sum over k only appears if $t \geq r$, and in this case $\min\{0, r-t\} = r-t$. Therefore the lower limit in the latter sum can also be expressed as $k = \min\{0, r-t\}$. With these changes, (A.121) is rewritten as

$$\begin{aligned}\frac{\mathrm{d}\zeta'_0}{\mathrm{d}\Omega} = &-\frac{L(\xi'-)\Omega^{r-1}\exp(-\Omega/\xi')}{\xi'^r(r-1)!} + \frac{\xi'^{t-1}\theta'\exp(-\Omega/\xi')}{(r-1)!t!} \\ &\cdot \sum_{k=\min\{0,r-t\}}^{r-1}\left(\frac{\Omega}{\xi'}\right)^{k-1}\sum_{j=0}^{t}\binom{t}{j} j(r-j-1)^{(r-k-1)}(-1)^{t-j} \\ &+ O\left(\Omega^{r-t-2}\exp(-\Omega/\xi')\right).\end{aligned} \tag{A.122}$$

From Lemma 6, the inner sum in (A.122) equals 0 for $k = r-t+1, r-t+2, \ldots, r-1$ and $(-1)^{t-1}t!$ for $k = r-t$. If $t < r$, the terms with index $k = 0, 1, \ldots, r-t-1$ are $O(\Omega^{r-t-2}\exp(-\Omega/\xi'))$. Therefore

$$\begin{aligned}\frac{\mathrm{d}\zeta'_0}{\mathrm{d}\Omega} = &-\frac{L(\xi'-)\Omega^{r-1}\exp(-\Omega/\xi')}{\xi'^r(r-1)!} + \frac{(-1)^{t-1}\xi'^{2t-r}\theta'\Omega^{r-t-1}\exp(-\Omega/\xi')}{(r-1)!} \\ &+ O\left(\Omega^{r-t-2}\exp(-\Omega/\xi')\right).\end{aligned} \tag{A.123}$$

Using (A.96), (A.97), (A.107), (A.109) and (A.123), and considering that $L(\xi'-) = L(\xi'+)$,

$$\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega} \geq \frac{(-1)^{t-1}\xi'^{2t-r}\theta'\Omega^{r-t-1}\exp(-\Omega/\xi')}{(r-1)!} + O\left(\Omega^{r-t-2}\exp(-\Omega/\xi')\right). \tag{A.124}$$

Since $(-1)^{t-1}\theta' > 0$, this implies that

$$\lim_{\Omega\to\infty}\left(\frac{\exp(\Omega/\xi')}{\Omega^{r-t-1}}\frac{\mathrm{d}\bar{\eta}}{\mathrm{d}\Omega}\right) \geq \frac{(-1)^{t-1}\xi'^{2t-r}\theta'}{(r-1)!} > 0. \tag{A.125}$$

As a consequence of (A.113) and (A.125), in either case (a) or (b) of property 2 in Assumption 4, there exists Ω'_0 such that $\mathrm{d}\bar{\eta}/\mathrm{d}\Omega > 0$ for $\Omega \geq \Omega'_0$. □

Proof of Theorem 2. From Lemmas 7 and 8, there exist Ω_0, Ω'_0 such that, denoting by $\bar{\eta}|_\Omega$ the value of $\bar{\eta}$ corresponding to a given Ω,

$$\bar{\eta}|_\Omega > \bar{\eta}|_{\Omega_0} \quad \text{for } \Omega < \Omega_0, \tag{A.126}$$

$$\bar{\eta}|_\Omega > \bar{\eta}|_{\Omega'_0} \quad \text{for } \Omega > \Omega'_0. \tag{A.127}$$

Proposition 1 implies that $\bar{\eta}$ is a continuous function of Ω. Therefore, this function restricted to the interval $[\Omega_0, \Omega_0']$ has an absolute maximum (Apostol, 1974, theorem 4.28). Because of (A.126) and (A.127), this is the absolute maximum of $\bar{\eta}$ over $\mathbb{R}^+$. □

Lemma 9. *Under the hypotheses of Theorem 3, given $\sigma \in \mathbb{R}^+$, ζ as defined by (A.2) is a continuous function of $\Omega \in \mathbb{R}^+$.*

Proof. From Assumptions 1 and 2', L is continuous except possibly at a finite number of points, where it can only have removable discontinuities or jumps. Since removable discontinuities do not have any effect on the integral in (A.2), they can be disregarded. Thus in the following it is assumed that L only has jump discontinuities. Let D be the number of discontinuity points, located at $x_1 < x_2 < \cdots < x_D$. The function L can be decomposed as the sum of L_c and L_d, where L_c is continuous and L_d is piecewise constant with jumps at $x_1, \ldots, x_D$. Accordingly, $\zeta = \zeta_\mathrm{c} + \zeta_\mathrm{d}$, where ζ_c and ζ_d are given as in (A.2) with L replaced by L_c and L_d respectively.

For any $\Omega' \neq \Omega$, let ζ' denote the right-hand side of (A.2) with Ω replaced by Ω', and let ζ_c' and ζ_d' be defined similarly. For $\varepsilon > 0$ arbitrary, it is necessary to find $\delta > 0$ such that $|\zeta' - \zeta| < \varepsilon$ for $|\Omega' - \Omega| < \delta$. Consider an arbitrary $\delta_0 \in (0, \Omega)$. Since L_c is continuous, by the Heine-Cantor theorem (Apostol, 1974, theorem 4.47) it is uniformly continuous on the interval $[(\Omega - \delta_0)/(r\sigma), (\Omega + \delta_0)\sigma/r]$. This interval contains the values Ω/ν and Ω'/ν for $|\Omega' - \Omega| < \delta_0$, $\nu \in [r/\sigma, r\sigma]$. By virtue of this, defining $\varepsilon_\mathrm{c} = \varepsilon/(2r(\sigma - 1/\sigma))$, let $\delta_\mathrm{c} < \delta_0$ be chosen such that $|L_\mathrm{c}(\Omega'/\nu) - L_\mathrm{c}(\Omega/\nu)| < \varepsilon_\mathrm{c}$ for all $|\Omega' - \Omega| < \delta_\mathrm{c}$, $\nu \in [r/\sigma, r\sigma]$. Taking into account Lemma 1, it follows that

$$|\zeta_\mathrm{c}' - \zeta_\mathrm{c}| \leq \int_{r/\sigma}^{r\sigma} |L(\Omega/\nu) - L(\Omega'/\nu)|\,\mathrm{d}\nu < r\left(\sigma - \frac{1}{\sigma}\right)\varepsilon_\mathrm{c} = \frac{\varepsilon}{2} \quad \text{for } |\Omega' - \Omega| < \delta_\mathrm{c}. \tag{A.128}$$

By construction, there exists an upper bound M_d on $|L_\mathrm{d}(x)|$, $x \in \mathbb{R}^+$. Since $L_\mathrm{d}(\Omega/\nu)$, considered as a function of ν, has jumps at $\Omega/x_1, \ldots, \Omega/x_D$, associated with each discontinuity point Ω/x_i there is an interval of values of ν for which $L_\mathrm{d}(\Omega'/\nu) \neq L_\mathrm{d}(\Omega/\nu)$. The width of this interval is $|\Omega' - \Omega|/x_i \leq |\Omega' - \Omega|/x_1$, and $|L_\mathrm{d}(\Omega'/\nu) - L_\mathrm{d}(\Omega/\nu)| \leq 2M_\mathrm{d}$ for ν within this interval. There are at most D such intervals contained in $[r/\sigma, r\sigma]$, and for any value of ν not belonging to any of these intervals it holds that $L_\mathrm{d}(\Omega'/\nu) = L_\mathrm{d}(\Omega/\nu)$. Using Lemma 1 again, it is seen that $|\zeta_\mathrm{d}' - \zeta_\mathrm{d}| \leq 2DM_\mathrm{d}|\Omega' - \Omega|/x_1$. Thus there exists δ_d such that

$$|\zeta_\mathrm{d}' - \zeta_\mathrm{d}| < \frac{\varepsilon}{2} \quad \text{for } |\Omega' - \Omega| < \delta_\mathrm{d}. \tag{A.129}$$

Taking $\delta = \min\{\delta_c, \delta_d\}$, it follows from (A.128) and (A.129) that

$$|\zeta' - \zeta| \leq |\zeta'_c - \zeta_c| + |\zeta'_d - \zeta_d| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{for } |\Omega' - \Omega| < \delta, \tag{A.130}$$

which shows that ζ is a continuous function of Ω. □

Lemma 10. *Under the hypotheses of Theorem 3, and with ζ defined by (A.2),*

$$\lim_{\sigma\to\infty} \limsup_{\Omega\to 0} \frac{\bar{\eta} - \zeta}{\zeta} = \lim_{\sigma\to\infty} \limsup_{\Omega\to\infty} \frac{\bar{\eta} - \zeta}{\zeta} = 0. \tag{A.131}$$

Proof. According to property 1 in Assumption 3', there exist $K < r$ and $m_L, M_L, x_L \in \mathbb{R}^+$ such that $m_L x^K < L(x) < M_L x^K$ for $x < x_L$, that is,

$$m_L(\Omega/\nu)^K < L(\Omega/\nu) < M_L(\Omega/\nu)^K \quad \text{for } \nu > \Omega/x_L. \tag{A.132}$$

Similarly, property 2 implies that there exist $K' < r$; $m'_L, M'_L \in \mathbb{R}^+$; and $x'_L > x_L$ such that

$$m'_L(\Omega/\nu)^{K'} < L(\Omega/\nu) < M'_L(\Omega/\nu)^{K'} \quad \text{for } \nu < \Omega/x'_L. \tag{A.133}$$

From Assumption 1, L is of bounded variation on $[x_L, x'_L]$, and thus there exists M such that $L(x) \leq M$ for $x \in [x_L, x'_L]$, that is,

$$L(\Omega/\nu) \leq M \quad \text{for } \Omega/x'_L \leq \nu \leq \Omega/x_L. \tag{A.134}$$

The case $\Omega \to 0$ is analyzed first. Given $\sigma \in \mathbb{R}^+$, it will be assumed that $\Omega < r x_L/\sigma$. Under this assumption, any ν within the integration interval in (A.2) exceeds Ω/x_L. Thus, applying (A.132),

$$\zeta > m_L \Omega^K \int_{r/\sigma}^{r\sigma} \frac{\nu^{r-K-1}\exp(-\nu)}{(r-1)!} \mathrm{d}\nu = \frac{m_L \Omega^K (\Gamma(r-K, r/\sigma) - \Gamma(r-K, r\sigma))}{(r-1)!}. \tag{A.135}$$

The difference $\bar{\eta} - \zeta$ can be expressed as[2] $\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4$, where each term is an integral as in (A.2) with the integration interval respectively given as $(0, \Omega/x'_L)$, $(\Omega/x'_L, \Omega/x_L)$, $(\Omega/x_L, r/\sigma)$ and $(r\sigma, \infty)$. In the first case, (A.133) implies that

$$\zeta_1 < \frac{M'_L \Omega^{K'} \gamma(r - K', \Omega/x'_L)}{(r-1)!}, \tag{A.136}$$

[2]Note that this decomposition, and the one to be used for $\Omega \to \infty$, are different from those used in the proofs of Lemmas 7 and 8 respectively, although the same notation is used for simplicity.

and thus

$$\frac{\zeta_1}{\zeta} < \frac{M'_{\mathrm{L}} \Omega^{K'-K} \gamma(r-K', \Omega/x'_{\mathrm{L}})}{m_{\mathrm{L}}(\Gamma(r-K, r/\sigma) - \Gamma(r-K, r\sigma))}. \tag{A.137}$$

Using the equality (A.70) from Lemma 4, and taking into account that $K, K' < r$ by Assumption 3', it is seen that the right-hand side of (A.137) tends to 0 as $\Omega \to 0$. Since ζ_1 and ζ are both positive, this implies that

$$\lim_{\Omega \to 0} \frac{\zeta_1}{\zeta} = 0. \tag{A.138}$$

As for the term ζ_2, using (A.134),

$$\zeta_2 \leq \frac{M(\gamma(r, \Omega/x_{\mathrm{L}}) - \gamma(r, \Omega/x'_{\mathrm{L}}))}{(r-1)!} < \frac{M\gamma(r, \Omega/x_{\mathrm{L}})}{(r-1)!}, \tag{A.139}$$

and thus

$$\frac{\zeta_2}{\zeta} < \frac{M\Omega^{-K} \gamma(r, \Omega/x_{\mathrm{L}})}{m_{\mathrm{L}}(\Gamma(r-K, r/\sigma) - \Gamma(r-K, r\sigma))}. \tag{A.140}$$

Using (A.70) again, and taking into account that $K < r$, it stems that

$$\lim_{\Omega \to 0} \frac{\zeta_2}{\zeta} = 0. \tag{A.141}$$

Regarding the third term, (A.132) holds for all ν within the integration interval, and thus

$$\zeta_3 < \frac{M_{\mathrm{L}} \Omega^K (\gamma(r-K, r/\sigma) - \gamma(r-K, \Omega/x_{\mathrm{L}}))}{(r-1)!} < \frac{M_{\mathrm{L}} \Omega^K \gamma(r-K, r/\sigma)}{(r-1)!}. \tag{A.142}$$

Therefore

$$\frac{\zeta_3}{\zeta} < \frac{M_{\mathrm{L}} \gamma(r-K, r/\sigma)}{m_{\mathrm{L}}(\Gamma(r-K, r/\sigma) - \Gamma(r-K, r\sigma))}. \tag{A.143}$$

Similarly, the fourth term satisfies

$$\zeta_4 < \frac{M_{\mathrm{L}} \Omega^K \Gamma(r-K, r\sigma)}{(r-1)!}, \tag{A.144}$$

and therefore

$$\frac{\zeta_4}{\zeta} < \frac{M_{\mathrm{L}} \Gamma(r-K, r\sigma)}{m_{\mathrm{L}}(\Gamma(r-K, r/\sigma) - \Gamma(r-K, r\sigma))}. \tag{A.145}$$

From (A.138), (A.141), (A.143) and (A.145) it follows that

$$\limsup_{\Omega\to 0}\frac{\bar{\eta}-\zeta}{\zeta}\leq\frac{M_{\mathrm{L}}(\gamma(r-K,r/\sigma)+\Gamma(r-K,r\sigma))}{m_{\mathrm{L}}(\Gamma(r-K,r/\sigma)-\Gamma(r-K,r\sigma))}. \tag{A.146}$$

The right-hand side of (A.146) is seen to converge to 0 as $\sigma\to\infty$, and thus so does the left-hand side. This establishes the first part of the result.

The analysis for $\Omega\to\infty$ is similar. Given $\sigma\in\mathbb{R}^+$, it is assumed that $\Omega > rx'_{\mathrm{L}}\sigma$. The difference $\bar{\eta}-\zeta$ is expressed as $\zeta'_1+\zeta'_2+\zeta'_3+\zeta'_4$, where each term is an integral as in (A.2) with integration intervals respectively given as $(0,r/\sigma)$, $(r\sigma,\Omega/x'_{\mathrm{L}})$, $(\Omega/x'_{\mathrm{L}},\Omega/x_{\mathrm{L}})$ and $(\Omega/x_{\mathrm{L}},\infty)$. Arguments analogous to those used for $\Omega\to 0$ establish that

$$\limsup_{\Omega\to \infty}\frac{\bar{\eta}-\zeta}{\zeta}\leq\frac{M'_{\mathrm{L}}(\gamma(r-K',r/\sigma)+\Gamma(r-K',r\sigma))}{m'_{\mathrm{L}}(\Gamma(r-K',r/\sigma)-\Gamma(r-K',r\sigma))}. \tag{A.147}$$

The right-hand side of (A.147) is seen to converge to 0 as $\sigma\to\infty$, and thus so does the left-hand side. This establishes the second part of the result. □

Lemma 11. *Under the hypotheses of Theorem 3, considering ζ and $\bar{\eta}$ as functions of $\Omega\in\mathbb{R}^+$, $\zeta/\bar{\eta}\to 1$ uniformly on $\mathbb{R}^+$ as $\sigma\to\infty$.*

Proof. The result is equivalent to the statement that for any $\varepsilon>0$ there exists σ_0 such that $|\bar{\eta}/\zeta-1|<\varepsilon$ for all $\Omega\in\mathbb{R}^+$ and for all $\sigma>\sigma_0$. Consider $\varepsilon>0$ arbitrary. Let $R(\sigma)$ and $R'(\sigma)$ respectively denote $\limsup_{\Omega\to 0}(\bar{\eta}-\zeta)/\zeta$ and $\limsup_{\Omega\to\infty}(\bar{\eta}-\zeta)/\zeta$. Since L is a non-negative function, from (A.2) it is seen that ζ is a non-negative, non-decreasing function of σ for any Ω. By Lemma 10, $R(\sigma)$ and $R'(\sigma)$ tend to 0 as $\sigma\to\infty$, and thus there exists σ_1 such that $R(\sigma_1),R'(\sigma_1)\leq\varepsilon/2$. By definition of $R(\sigma)$, there exists Ω_0 such that the following inequality holds (note that the left-hand side is a function of σ and Ω):

$$\frac{\bar{\eta}-\zeta}{\zeta}<R(\sigma_1)+\frac{\varepsilon}{2}\leq\varepsilon\quad\text{for }\Omega<\Omega_0,\ \sigma=\sigma_1. \tag{A.148}$$

The non-decreasing character of ζ with σ implies that (A.148) also holds for $\sigma>\sigma_1$, that is,

$$\frac{\bar{\eta}-\zeta}{\zeta}<\varepsilon\quad\text{for }\Omega<\Omega_0,\ \sigma\geq\sigma_1. \tag{A.149}$$

Analogously, there exists $\Omega'_0>\Omega_0$ such that

$$\frac{\bar{\eta}-\zeta}{\zeta}<\varepsilon\quad\text{for }\Omega>\Omega'_0,\ \sigma\geq\sigma_1. \tag{A.150}$$

According to Lemma 9, for σ fixed, ζ is a continuous function of $\Omega \in [\Omega_0, \Omega_0']$, and therefore it has an absolute minimum on that interval, which will be denoted as $S_1(\sigma)$. The non-negative and non-decreasing character of ζ with σ implies that S_1 is also a non-negative, non-decreasing function. In addition, $S_1(\sigma) > 0$ for all σ greater than a certain value σ_2. This can be seen as follows. By Assumption 3', $L(x)$ is non-zero for all x outside a bounded interval. If σ is sufficiently large, i.e. greater than a certain σ_2, for any $\Omega \in [\Omega_0, \Omega_0']$ the integration interval in (A.2) contains a subinterval where L is non-zero, which gives $\zeta > 0$. Thus $S_1(\sigma) > 0$ for $\sigma > \sigma_2$.

By arguments similar to those in the above paragraph, $\bar{\eta} - \zeta$, considered as a function of Ω, has an absolute maximum on $[\Omega_0, \Omega_0']$; and this maximum, denoted as $S_2(\sigma)$, tends to 0 as $\sigma \to \infty$. Therefore, defining $S(\sigma) = S_2(\sigma)/S_1(\sigma)$ for $\sigma > \sigma_2$,

$$(\bar{\eta} - \zeta)/\zeta \leq S(\sigma) \quad \text{for } \Omega \in [\Omega_0, \Omega_0'],\ \sigma > \sigma_2; \tag{A.151}$$

and $S(\sigma) \to 0$ as $\sigma \to \infty$. Thus, for the considered ε, there exists $\sigma_3 \geq \sigma_2$ such that $S(\sigma) < \varepsilon$ for $\sigma \geq \sigma_3$. Combined with (A.151), this gives

$$(\bar{\eta} - \zeta)/\zeta < \varepsilon \quad \text{for } \Omega \in [\Omega_0, \Omega_0'],\ \sigma \geq \sigma_3. \tag{A.152}$$

From (A.149), (A.150) and (A.152), choosing $\sigma_0 = \max\{\sigma_1, \sigma_3\}$ is sufficient to satisfy $|\bar{\eta}/\zeta - 1| < \varepsilon$ for $\Omega \in \mathbb{R}^+$, $\sigma > \sigma_0$. This completes the proof. □

Proof of Theorem 3. The result will be proved by contradiction. Assume that there exists a possibly randomized estimator $\hat{p}$ with $\limsup_{p\to 0} \eta(p) < \eta^*$. This implies that there exist $\theta < 1$ and a probability p_θ such that the estimator has

$$\eta(p) < \theta\eta^* \quad \text{for all } p < p_\theta. \tag{A.153}$$

For $n = r, r+1, \ldots$, let Π_n denote the distribution function of $\hat{p}$ conditioned on $N = n$.

By Lemma 11, let σ be selected such that

$$\int_{r/\sigma}^{r\sigma} \phi(\nu) L(\Omega/\nu)\,\mathrm{d}\nu > \sqrt[3]{\theta} \int_0^\infty \phi(\nu) L(\Omega/\nu)\,\mathrm{d}\nu \quad \text{for all } \Omega \in \mathbb{R}^+. \tag{A.154}$$

In particular, this implies that

$$\int_{r/\sigma}^{r\sigma} \phi(\nu) L(\Omega/\nu)\,\mathrm{d}\nu > \sqrt[3]{\theta}\eta^* \quad \text{for all } \Omega \in \mathbb{R}^+. \tag{A.155}$$

Given ν_1, ν_2 with $\nu_2 > \nu_1 > 0$, according to Lemma 2, $\Phi(p,\nu) \to \phi(\nu)$ uniformly on $[\nu_1, \nu_2]$ as $p \to 0$. By virtue of this, let $p_1 < p_\theta$ be such that

$$|\Phi(p,\nu) - \phi(\nu)| < (1 - \sqrt[3]{\theta})\phi(\nu) \quad \text{for all } p < p_1,\ \nu \in [r/\sigma, r\sigma]. \tag{A.156}$$

Let $u = \lceil r\sigma/p_1 \rceil$. Taking into account that $\lim_{w\to\infty}(\sum_{n=1}^{w} 1/n - \log w) = \gamma$, where γ is the Euler-Mascheroni constant (Abramowitz and Stegun, 1970, equation 6.1.3), it is easy to see that

$$\lim_{p\to 0}\left(\sum_{n=u}^{\lfloor r/(\sigma p)\rfloor} \frac{1}{n} - \log\frac{p_1}{p}\right) = \gamma + \log\frac{r}{\sigma p_1} - \sum_{n=1}^{u-1}\frac{1}{n}. \tag{A.157}$$

This implies that there exist $\lambda > 0$ and p_0' such that

$$\sum_{n=u}^{\lfloor r/(\sigma p)\rfloor} \frac{1}{n} - \log\frac{p_1}{p} > -\lambda \quad \text{for all } p \le p_0'. \tag{A.158}$$

Let λ and p_0' be chosen such that (A.158) holds, and let p_0'' be defined by the equation

$$\log\frac{p_1}{p_0''} = \frac{\lambda}{1-\sqrt[3]{\theta}}. \tag{A.159}$$

Since $\lambda > 0$ and $\theta < 1$, it follows that $p_0'' < p_1$.

Let $p_0 = \min\{p_0', p_0''\}$. For a given n, the measure associated with the distribution function Π_n is obviously finite, and thus sigma-finite. This implies (Billingsley, 1995, theorem 18.3) that for each n the integral in (9), considered as a function of p, is measurable with respect to Lebesgue measure. In addition, since $p_1 < p_\theta$, it stems from (A.153) that the series in (9) converges for $p \le p_1$. This assures (Billingsley, 1995, theorem 13.4(ii)) that $\eta(p)$ restricted to $p \le p_1$ is measurable. Therefore, the integral

$$X = \int_{p_0}^{p_1} \frac{\eta(p)}{p}\,\mathrm{d}p \tag{A.160}$$

exists in the Lebesgue sense, and according to (A.153) it satisfies

$$X < \theta\eta^* \int_{p_0}^{p_1}\frac{\mathrm{d}p}{p} = \theta\eta^*\log\frac{p_1}{p_0}. \tag{A.161}$$

Substituting (9) into (A.160),

$$X = \int_{p_0}^{p_1} \frac{1}{p}\sum_{n=r}^{\infty} f(n)\left(\int_0^\infty L(y/p)\,\mathrm{d}\Pi_n(y)\right)\mathrm{d}p. \tag{A.162}$$

Defining $v = \lfloor r/(\sigma p_0)\rfloor$, it is clear from (A.162) that

$$X > \sum_{n=u}^{v}\int_{p_0}^{p_1}\left(\int_0^\infty \frac{f(n)L(y/p)}{p}\,\mathrm{d}\Pi_n(y)\right)\mathrm{d}p. \tag{A.163}$$

Since both measures in (A.163) are sigma-finite, and both the inner and outer integrals are finite, the order of integration can be reversed (Billingsley, 1995, theorem 18.3), which gives

$$X > \sum_{n=u}^{v} \int_0^\infty \left(\int_{p_0}^{p_1} \frac{f(n)L(y/p)}{p} \,\mathrm{d}p \right) \mathrm{d}\Pi_n(y). \tag{A.164}$$

Making the change of variable $\nu = np$ in the inner integral and taking into account that $f(n)/p = \Phi(p, np)$, (A.164) becomes

$$X > \sum_{n=u}^{v} \frac{1}{n} \int_0^\infty \left(\int_{np_0}^{np_1} \Phi(\nu/n, \nu) L(ny/\nu) \,\mathrm{d}\nu \right) \mathrm{d}\Pi_n(y). \tag{A.165}$$

For $u \leq n \leq v$ it holds that $np_0 \leq r/\sigma$ and $np_1 \geq r\sigma$. Therefore

$$X > \sum_{n=u}^{v} \frac{1}{n} \int_0^\infty \left(\int_{r/\sigma}^{r\sigma} \Phi(\nu/n, \nu) L(ny/\nu) \,\mathrm{d}\nu \right) \mathrm{d}\Pi_n(y). \tag{A.166}$$

For $\nu \in [r/\sigma, r\sigma]$ and $u \leq n \leq v$ it holds that $\nu/n < p_1$. Thus (A.156) gives $\Phi(\nu/n, \nu) > \sqrt[3]{\theta}\phi(\nu)$. Substituting into (A.166),

$$X > \sqrt[3]{\theta} \sum_{n=u}^{v} \frac{1}{n} \int_0^\infty \left(\int_{r/\sigma}^{r\sigma} \phi(\nu) L(ny/\nu) \,\mathrm{d}\nu \right) \mathrm{d}\Pi_n(y). \tag{A.167}$$

From (A.155), the inner integral in (A.167) exceeds $\sqrt[3]{\theta}\eta^*$, and thus

$$X > \theta^{2/3} \eta^* \sum_{n=u}^{v} \frac{1}{n}. \tag{A.168}$$

Since $p_0 \leq p_0'$ and $p_0 \leq p_0'' < p_1$, (A.158) and (A.159) give

$$\sum_{n=u}^{v} \frac{1}{n} > -\lambda + \log \frac{p_1}{p_0} \geq \log \frac{p_1}{p_0} \left(1 - \frac{\lambda}{\log(p_1/p_0'')} \right) = \sqrt[3]{\theta} \log \frac{p_1}{p_0}. \tag{A.169}$$

Substituting into (A.168),

$$X > \theta \eta^* \log \frac{p_1}{p_0}, \tag{A.170}$$

in contradiction with (A.161). This establishes the result. □

Proof of Proposition 3. The proof is analogous to that of Mendo and Hernando (2010, proposition 1). □

Proof of Proposition 4. For the considered estimator,

$$\frac{\mathrm{E}[(\hat{p}-p)^2]}{p^2} = \frac{(r-2)^2}{p^2}\mathrm{E}\left[\frac{1}{(N-1)^2}\right] - \frac{2(r-2)}{p}\mathrm{E}\left[\frac{1}{N-1}\right] + 1. \tag{A.171}$$

The equality

$$\mathrm{E}\left[\frac{1}{N-1}\right] = \frac{p}{r-1} \tag{A.172}$$

directly stems from the fact that (1) is unbiased. On the other hand, according to Mikulski and Smith (1976), for $p \in (0,1)$

$$\mathrm{Var}\left[\frac{r-1}{N-1}\right] \leq \frac{p^2(1-p)}{r-2} < \frac{p^2}{r-2}. \tag{A.173}$$

From (A.172) and (A.173),

$$\mathrm{E}\left[\frac{1}{(N-1)^2}\right] = \left(\mathrm{E}\left[\frac{1}{N-1}\right]\right)^2 + \mathrm{Var}\left[\frac{1}{N-1}\right] < \frac{p^2}{(r-1)(r-2)}. \tag{A.174}$$

Substituting (A.172) and (A.174) into (A.171), the desired result (27) is obtained. □

Lemma 12. *Given $r \geq 3$ and $\Omega \in \mathbb{R}^+$, considering the loss function (19) with $A_1 = 0$, $A_2 > 0$, if $\mu_2 \geq (r+\sqrt{r}+1)/\Omega$ the risk of the estimator (28) satisfies $\eta(p) < \bar{\eta}$ for any $p \in (0,1)$. Similarly, for the loss function (19) with $A_1 > 0$, $A_2 = 0$, if $\mu_1 \geq \Omega/(r-\sqrt{r})$ the inequality $\eta(p) < \bar{\eta}$ holds for any $p \in (0,1)$.*

Proof. The stated results follow from the arguments used in the proof of Mendo and Hernando (2010, proposition 3). □

Proof of Proposition 5. The result will be proved by approximating the loss function as a sum of terms of the form (19) with $A_1, A_2 \geq 0$ and using Lemma 12. It may be assumed without loss of generality that $L(x) = 0$ for $x \in [\upsilon, \upsilon']$, because if $L(x) = C$ within that interval, defining $L'(x) = L(x) - C$ the risk corresponding to L is expressed as C plus the risk resulting from the loss function L', which satisfies the hypotheses of the proposition.

Let $\varepsilon > 0$, and suppose for the moment that L is unbounded on the interval $(0, \upsilon)$. This implies that for any $i \in \mathbb{N}$, the set $V_{\varepsilon,i} = \{x \in (0,\upsilon) \mid L(x) \geq i\varepsilon\}$ is non-empty. In fact, since L is non-increasing on $(0,\upsilon)$, $V_{\varepsilon,i}$ is an interval. Let $x_{\varepsilon,i}$ be defined as the supremum of $V_{\varepsilon,i}$, and let

$$\ell_{\varepsilon,i}(x) = \begin{cases} \varepsilon & \text{if } x \leq x_{\varepsilon,i}, \\ 0 & \text{otherwise.} \end{cases} \tag{A.175}$$

If L is bounded on $(0,\upsilon)$, the sets $V_{\varepsilon,i}$ are empty for i greater than a certain value. In this case, the corresponding $\ell_{\varepsilon,i}$ functions are defined as the null function. In a similar manner, for L unbounded on (υ',∞), let $V'_{\varepsilon,i} = \{x \in (\upsilon',\infty) \mid L(x) \geq i\varepsilon\}$, which is again non-empty interval; let $x'_{\varepsilon,i}$ be its infimum, and

$$\ell'_{\varepsilon,i}(x) = \begin{cases} \varepsilon & \text{if } x \geq x'_{\varepsilon,i}, \\ 0 & \text{otherwise.} \end{cases} \tag{A.176}$$

If L is bounded on (υ',∞), for i greater than a certain value the sets $V'_{\varepsilon,i}$ are empty, and the corresponding $\ell'_{\varepsilon,i}$ are defined as null. Let $L_{\varepsilon,i}(x) = \ell_{\varepsilon,i}(x) + \ell'_{\varepsilon,i}(x)$ and $L_\varepsilon(x) = \sum_{i=1}^{\infty} L_{\varepsilon,i}(x)$. By construction, for all $x \in \mathbb{R}^+$,

$$0 \leq L(x) - L_\varepsilon(x) \leq \varepsilon. \tag{A.177}$$

Each function $L_{\varepsilon,i}$ satisfies Assumptions 1–3, and therefore a risk can be defined considering $L_{\varepsilon,i}$ as the loss function. This risk will be denoted as $\eta_{\varepsilon,i}(p)$. The function L_ε also satisfies Assumptions 1–3. Let $\eta_\varepsilon(p)$ denote its corresponding risk,

$$\eta_\varepsilon(p) = \sum_{n=r}^{\infty} f(n) L_\varepsilon(g(n)/p) = \sum_{n=r}^{\infty}\sum_{i=1}^{\infty} f(n) L_{\varepsilon,i}(g(n)/p) \tag{A.178}$$

For each n, the inner series in (A.178) converges absolutely; namely, to $f(n)L_\varepsilon(g(n)/p)$. In addition, from (A.177) it is seen that $L_\varepsilon(g(n)/p) \leq L(g(n)/p)$, and this implies that the outer series in (A.178) is also absolutely convergent. This allows interchanging the sums over n and i (Apostol, 1974, theorem 8.43), which gives

$$\eta_\varepsilon(p) = \sum_{i=1}^{\infty} \eta_{\varepsilon,i}(p). \tag{A.179}$$

Theorem 1 assures that $\eta_{\varepsilon,i}(p)$ has an asymptotic value $\bar{\eta}_{\varepsilon,i}$, given by

$$\bar{\eta}_{\varepsilon,i} = \int_0^\infty \phi(\nu) L_{\varepsilon,i}(\Omega/\nu)\,\mathrm{d}\nu, \tag{A.180}$$

Similarly, $\eta_\varepsilon(p)$ has an asymptotic value

$$\bar{\eta}_\varepsilon = \int_0^\infty \phi(\nu) \sum_{i=1}^{\infty} L_{\varepsilon,i}(\Omega/\nu)\,\mathrm{d}\nu. \tag{A.181}$$

Since $L_{\varepsilon,i}$ is a nonnegative function for all i, the monotone convergence theorem (Athreya and Lahiri, 2006, theorem 2.3.4) implies that the sum and integral signs in (A.181) commute, and thus

$$\bar{\eta}_\varepsilon = \sum_{i=1}^{\infty} \bar{\eta}_{\varepsilon,i}. \tag{A.182}$$

From Lemma 12, $\eta_{\varepsilon,i}(p) < \bar{\eta}_{\varepsilon,i}$. Combined with (A.179) and (A.182), this gives

$$\eta_\varepsilon(p) < \bar{\eta}_\varepsilon. \tag{A.183}$$

On the other hand, from (A.177) it stems that

$$0 \leq \eta(p) - \eta_\varepsilon(p) \leq \varepsilon, \tag{A.184}$$

which in turn implies

$$0 \leq \bar{\eta} - \bar{\eta}_\varepsilon \leq \varepsilon. \tag{A.185}$$

From (A.183)–(A.185),

$$\eta(p) \leq \eta_\varepsilon(p) + \varepsilon < \bar{\eta}_\varepsilon + \varepsilon < \bar{\eta} + \varepsilon. \tag{A.186}$$

Since (A.186) holds for ε arbitrary, the desired inequality $\eta(p) \leq \bar{\eta}$ follows. □

References

Abramowitz, M., Stegun, I. A. (Eds.), 1970. Handbook of Mathematical Functions, ninth Edition. Dover.

Apostol, T. M., 1967. Calculus, 2nd Edition. Vol. 1. John Wiley and Sons.

Apostol, T. M., 1974. Mathematical Analysis, 2nd Edition. Addison-Wesley.

Athreya, K. B., Lahiri, S. N., 2006. Measure Theory and Probability Theory, 2nd Edition. Springer.

Billingsley, P., 1995. Probability and Measure, 3rd Edition. John Wiley and Sons.

Carter, M., van Brunt, B., 2000. The Lebesgue-Stieltjes Integral: A Practical Introduction. Springer.

Fleming, W., 1977. Functions of Several Variables, 2nd Edition. Springer-Verlag.

Haldane, J. B. S., 1945. On a method of estimating frequencies. Biometrika 33 (3), 222–225.

Lehmann, E. L., Casella, G., 1998. Theory of Point Estimation, 2nd Edition. Springer.

Mendo, L., November 2009. Estimation of a probability with guaranteed normalized mean absolute error. IEEE Communications Letters 13 (11), 817–819.

Mendo, L., Hernando, J. M., February 2006. A simple sequential stopping rule for Monte Carlo simulation. IEEE Transactions on Communications 54 (2), 231–241.

Mendo, L., Hernando, J. M., November 2008. Improved sequential stopping rule for Monte Carlo simulation. IEEE Transactions on Communications 56 (11), 1761–1764.

Mendo, L., Hernando, J. M., May 2010. Estimation of a probability with optimum guaranteed confidence in inverse binomial sampling. Bernoulli 16 (2), 493–513.

Mikulski, P. W., Smith, P. J., 1976. A variance bound for unbiased estimation in inverse sampling. Biometrika 63 (1), 216–217.