Economics Bulletin

# Volume 35, Issue 1

# Disability Risk and Hyperbolic Discounting

Sanghoon K Lee
*Hallym University*

## Abstract

This paper presents a model where the ability to enjoy consumption gets impaired by disabilities over time. A continuous-time Markov chain model is constructed to examine how disability risk affects consumers' discounting of future utilities. It is shown that hyperbolic discounting arises for all future periods when disabilities occur independently. When the risks are interdependent, it is shown that hyperbolic discounting arises asymptotically as long as the hazard of disability increases with the set of possible disabilities.

**Citation:** Sanghoon K Lee, (2015) "Disability Risk and Hyperbolic Discounting", *Economics Bulletin*, Volume 35, Issue 1, pages 371-380
**Contact:** Sanghoon K Lee - sanghoonklee@gmail.com.
**Submitted:** June 22, 2014. **Published:** March 11, 2015.

## 1. Introduction

In standard economic models, agents are assumed to discount future payoffs with a constant rate. Numerous studies have found however that this is not the case. Evidence seems to support *hyperbolic discounting*, that is, agents discount the near future at a higher rate than the distant future. [1] This paper aims to provide a new explanation for this empirical regularity.

The model builds on a simple idea that one needs to be healthy in order to enjoy consumption. This is trivially with death: a dead person cannot consume anything. For this reason, discount rate has often been interpreted as the sum of pure time preference and mortality rate.[2] Under this interpretation, discount rate will fall when time preference remains constant as long as mortality rate decreases with time. What is difficult to justify though is why mortality risk should be decreasing rather than increasing over time.

This paper shows that the problem can be resolved by introducing the possibility of partial impairment into the analysis. Accidents occur indeed and some people end up living with disabilities for the rest of their lives. It is not difficult to imagine how the loss of hearing or vision, for example, will diminish the experience of consumption. This implies, however, that rational agents will take this disability risk into account when discounting the future. It might make sense then to discount the near future more heavily than the distant future—people have more to lose from accidents now than in the future when they may have accumulated several disabilities already.

This intuition is formalized in a model where consumers' utility depends on which set of organs are functioning at the time of consumption. Organs are assumed to fail following a multidimensional continuous-time Markov process. The analysis shows that hyperbolic discounting arises when organs fail independently of each other. When the failures are interdependent, hyperbolic discounting is shown to be asymptotic as long as the total failure rate increases with the set of possible disabilities.

Several authors have taken axiomatic approach and provided foundations for hyperbolic discounting. Hayashi (2003) shows that quasi-hyperbolic discounting arises when stationarity assumption is relaxed to quasi-stationarity. Ok and Masatlioglu (2007) present a representation theorem for preferences on the prize-time space. Their general axiomatic framework covers hyperbolic discounting as well as similarity-induced time preference (Rubinstein 2003) and subadditive time preference (Read 2001). It is well known that dynamic inconsistency occurs in non-expected utility models (Machina 1989; Karni and Schmeidler 1991). Halevy (2008) and Saito (2011) derive conditions for quasi-hyperbolic and hyperbolic discounting when the utility function is not linear in the continuation probability.

Within the framework of expected utility theory, Dasgupta and Maskin (2005) showed that hyperbolic discounting arises when there is a possibility that payoffs will be realized early. Becker and Mulligan (1997) present an endogenous model of time preference where agents can enhance their patience through costly investment. Azfar (1999) and Sozou (1998) examined a model where decreasing failure rate occurs due to Bayesian updating of an unknown failure rate. The closest to this paper is Proschan (1963) where a mixture of non-increasing failure rate distributions is shown to have a non-increasing failure rate. Gurland and Sethuraman (1995)

---

[1] See Strotz (1955), Thaler (1981), and Frederick, Loewenstein, and O'Donoghue (2002). For an alternative view, see Read (2001) and Rubinstein (2003). The idea of hyperbolic discounting has been fruitfully applied in various contexts (Phelps and Pollak, 1968; Laibson , 1997; O'Donoghue and Rabin , 1999).

[2] See, for example, Yaari (1965).

extended this result to mixtures of several increasing failure rate distributions. None of these papers, however, examines the case where the failure rate of a component depends on the state of other components in the system.

## 2. Model

A consumer has $n \geq 2$ organs. A consumer's health is represented by an $n$-dimensional vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in \{0,1\}^n$. Let $x_i = 1$ if the organ $i$ is functioning and $x_i = 0$ otherwise. Zero vector and unit vectors are defined in the usual way:

$$\boldsymbol{0} \equiv (0, \ldots, 0)$$
$$\boldsymbol{e_i} \equiv (0, \ldots, x_i = 1, \ldots, 0).$$

Define $\boldsymbol{x_{-i}}$ such that

$$\boldsymbol{x_{-i}} \equiv \boldsymbol{x} - \boldsymbol{e_i}.$$

Organs fail as time goes on. Let $\boldsymbol{x}(t)$ be the state of a consumer's health at time $t$. $\boldsymbol{x}(t)$ is assumed to be an $n$- dimensional continuous-time Markov chain. The transition probabilities are given by

$$P_{\boldsymbol{xx'}}(t) \equiv P\{\boldsymbol{x}(t+s) = \boldsymbol{x}' |\, \boldsymbol{x}(s) = \boldsymbol{x}\}.$$

Define instantaneous transition rates such that

$$q_{\boldsymbol{xx'}} \equiv \lim_{h \to 0} \frac{P_{\boldsymbol{xx'}}(h)}{h}, \boldsymbol{x} \neq \boldsymbol{x}'.$$

It is assumed that there are only two types of transitions: either one of the functioning organs fails (*injury*) or all of the functioning organs fail at the same time (*death*). The corresponding transition rates are given by

$$q_{\boldsymbol{xx_{-i}}} \equiv q_i(\boldsymbol{x})$$
$$q_{\boldsymbol{x0}} \equiv q_{\boldsymbol{0}}(\boldsymbol{x})$$
$$q_{\boldsymbol{xx'}} = 0 \text{ if } \boldsymbol{x} \neq \boldsymbol{x_{-i}} \text{ and } \boldsymbol{x} \neq \boldsymbol{0}.$$

Implicit here is that failed organs never recover, that is, $P_{\boldsymbol{xx'}}(t) = 0$ for $\boldsymbol{x} < \boldsymbol{x}'$. Let $q(\boldsymbol{x})$ be the *total failure rate* out of state $\boldsymbol{x}$:

$$q(\boldsymbol{x}) \equiv \sum_i q_i(\boldsymbol{x}) + q_{\boldsymbol{0}}(\boldsymbol{x}).$$

The utility from consumption depends on the consumer's health. Let $c_t$ be the rate of consumption at time $t$. Then, the flow utility at time $t$ is given by

$$\boldsymbol{\delta}^T \boldsymbol{x}(t) u(c_t)$$

where $\boldsymbol{\delta}^T \equiv (\delta_1, \ldots, \delta_n) > \boldsymbol{0}$. It shows that organ $j$, if functioning, contributes $\delta_j$ to the ability to enjoy consumption. The weights are normalized so that $\sum_j \delta_j = 1$. For a given initial condition $\boldsymbol{x}(0) = \boldsymbol{x}$, the lifetime expected utility of an agent is then given by

$$EU(\boldsymbol{x}) \equiv E\left[\int_0^\infty e^{-\rho t}\boldsymbol{\delta}^T\boldsymbol{x}(t)u(c_t)dt \middle| \boldsymbol{x}(0) = \boldsymbol{x}\right].$$

Define an "effective discount factor" $\beta_{\boldsymbol{x}}(t)$ and an (expected) "decay factor" $\delta_{\boldsymbol{x}}(t)$ such that:

$$\beta_{\boldsymbol{x}}(t) \equiv E[e^{-\rho t}\boldsymbol{\delta}^T\boldsymbol{x}(t)|\boldsymbol{x}(0) = \boldsymbol{x}]$$
$$\delta_{\boldsymbol{x}}(t) \equiv E[\boldsymbol{\delta}^T\boldsymbol{x}(t)|\boldsymbol{x}(0) = \boldsymbol{x}].$$

Then, the lifetime expected utility can be written as

$$EU(\boldsymbol{x}) = \int_0^\infty \beta_{\boldsymbol{x}}(t)u(c_t)dt = \int_0^\infty e^{-\rho t}\delta_{\boldsymbol{x}}(t)u(c_t)dt.$$

Hyperbolic discounting arises when discount rate decreases as time goes on. Given the discount factor $\beta_{\boldsymbol{x}}(t)$, the corresponding discount rate $r_{\boldsymbol{x}}(t)$ is implicitly given by

$$\beta_{\boldsymbol{x}}(t) = e^{-\int_0^t r_x(u)du}.$$

An agent has hyperbolic discounting in state $\boldsymbol{x}$ if

$$\frac{dr_{\boldsymbol{x}}(t)}{dt} < 0 \Leftrightarrow \frac{d^2(\ln\beta_{\boldsymbol{x}}(t))}{dt^2} > 0.$$

But given the relationship $\beta_{\boldsymbol{x}}(t) = e^{-\rho t}\delta_{\boldsymbol{x}}(t)$, it follows that

$$\frac{d^2(\ln\beta_{\boldsymbol{x}}(t))}{dt^2} = \frac{d^2(\ln\delta_{\boldsymbol{x}}(t))}{dt^2}.$$

Hyperbolic discounting is thus equivalent to hyperbolic "decay" in this model.

## 3. Analysis

Notice that the expected decay function $\delta_{\boldsymbol{x}}(t)$ is given by

$$\delta_{\boldsymbol{x}}(t) = \sum_{\boldsymbol{x}'} P_{\boldsymbol{x}\boldsymbol{x}'}(t)\,\boldsymbol{\delta}^T\boldsymbol{x}'.$$

By differentiating with respect to time, one obtains

$$\dot{\delta}_{\boldsymbol{x}}(t) \equiv \frac{d\delta_{\boldsymbol{x}}(t)}{dt} = \sum_{\boldsymbol{x}'} \dot{P}_{\boldsymbol{x}\boldsymbol{x}'}(t)\,\boldsymbol{\delta}^T\boldsymbol{x}'$$

where $\dot{P}_{\boldsymbol{x}\boldsymbol{x}'} \equiv \frac{dP_{\boldsymbol{x}\boldsymbol{x}'}(t)}{dt}$.

This expression can be transformed into a system of linear differential equations. Kolmogorov's Backward Equation is given by

$$\dot{P}_{\boldsymbol{x}\boldsymbol{x}'}(t) = \sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}} \{P_{\boldsymbol{y}\boldsymbol{x}'}(t) - P_{\boldsymbol{x}\boldsymbol{x}'}(t)\}.$$

Multiplying both sides by $\boldsymbol{\delta}^T\boldsymbol{x}'$ and summing over $\boldsymbol{x}'$ gives

$$\begin{aligned}\sum_{\boldsymbol{x}'} \dot{P}_{\boldsymbol{x}\boldsymbol{x}'}\,\boldsymbol{\delta}^T\boldsymbol{x}' &= \sum_{\boldsymbol{x}'}\sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}}\{P_{\boldsymbol{y}\boldsymbol{x}'}(t) - P_{\boldsymbol{x}\boldsymbol{x}'}(t)\}\,\boldsymbol{\delta}^T\boldsymbol{x}' \\ &= \sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}} \sum_{\boldsymbol{x}'}\{P_{\boldsymbol{y}\boldsymbol{x}'}(t) - P_{\boldsymbol{x}\boldsymbol{x}'}(t)\}\,\boldsymbol{\delta}^T\boldsymbol{x}' \\ &= \sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}}\{\delta_{\boldsymbol{y}}(t) - \delta_{\boldsymbol{x}}(t)\}.\end{aligned}$$

This shows that $\delta_{\boldsymbol{x}}(t)$ satisfies the following system of first-order linear differential equations:

$$\dot{\delta}_{\boldsymbol{x}}(t) + \sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}}\,\delta_{\boldsymbol{x}}(t) = \sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}}\,\delta_{\boldsymbol{y}}(t).$$

Given the relatively simple transition structure, the solution to the system can be obtained by substitution. A straightforward integration gives

$$\delta_{\boldsymbol{x}}(t) = e^{-q(x)t}\int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{x}\boldsymbol{y}}\,\delta_{\boldsymbol{y}}(u) e^{q(x)u} du + \boldsymbol{\delta}^T\boldsymbol{x} e^{-q(x)t}.$$

Recall that $\delta_{\boldsymbol{0}}(t) \equiv 0$ and $q_{\boldsymbol{x}\boldsymbol{y}} = 0$ unless $\boldsymbol{y} = \boldsymbol{x}_{-\boldsymbol{i}}$ or $\boldsymbol{y} = \boldsymbol{0}$. This implies that $\delta_{\boldsymbol{x}}(t)$ can be solved explicitly if $\delta_{\boldsymbol{x}_{-i}}(t)$ is known for all $i$. Let $\|\cdot\|$ denote the Euclidean norm. For $\|\boldsymbol{x}\| = 1$, that is, with only one functioning organ left, the expected decay has a simple form:

$$\delta_{\boldsymbol{x}}(t) = e^{-q(x)t}\boldsymbol{\delta}^T\boldsymbol{x}.$$

By substituting this for $\delta_{\boldsymbol{y}}(u)$, $\delta_{\boldsymbol{x}}(t)$ can be determined for $\|\boldsymbol{x}\| = 2$. But this implies that sequential substitution will determine all the rest of the expected decay functions for $\|\boldsymbol{x}\| \geq 3$.

For $\|\boldsymbol{x}\| = 1$, $\delta_{\boldsymbol{x}}(t)$ decays exponentially so that hyperbolic discounting never arises. The question is thus whether hyperbolic discounting ever arises for $\|\boldsymbol{x}\| \geq 2$. A special case of interest is when the organs fail independently. This occurs when $q_i(x) = q(e_i)$ for all $i$ and $x$. In this case, each organ fails with a constant rate, which means that its failure time follows an exponential distribution. Notice that, with normalization, the expected decay factor can be interpreted as a weighted average of organs' survival probabilities. A well-known result in mathematical statistics is that a mixture of exponentials has a decreasing failure rate (Proschan 1963). If organs fail independently, therefore, this suggests that decay will occur at a decreasing rate.

*Proposition 1 (Proschan, 1963).* *If* $q_i(x) = q(e_i)$ *for all* $i$ *and* $x$, *then* $\frac{d^2(\ln \delta_x(t))}{dt^2} \geq 0$ *for all* $t > 0$. *The inequality is strict if* $q(e_i) \neq q(e_j)$ *for* $i \neq j$.

*Proof* See Appendix.

Although it might be a reasonable approximation, the assumption that organs fail independently seems rather restrictive. It looks problematic in particular when several organs have failed already. An organ's failure is then more likely to depend on which set of organs are still functioning at the time. The model developed in this paper is flexible enough to incorporate this kind of state-dependent failure patterns. But allowing for state-dependence makes it more difficult to characterize the behavior of the expected decay factor. The following proposition shows, however, that one can still obtain an *asymptotic* result under relatively mild assumptions.

*Proposition 2.* *If* $q(\cdot)$ *is strictly increasing and* $q(e_i) \neq q(e_j)$ *for* $i \neq j$, *then there exists* $\hat{t} < \infty$ *such that* $\frac{d^2(\ln \delta_x(t))}{dt^2} > 0$ *for all* $t > \hat{t}$.

*Proof* See Appendix.

The monotonicity of the total failure rate $q(\cdot)$ means that injuries occur at a higher rate when there are more functioning organs. If organs fail independently, the total failure rate equals the sum of the individual failure rates. The monotonicity condition is thus satisfied trivially in this case. The condition $q(\boldsymbol{e_i}) \neq q(\boldsymbol{e_j})$ for $i \neq j$ requires that organs have a distinct failure rate. As is the case in Proposition 1, this guarantees that discounting becomes strictly hyperbolic.

## 4. Concluding Remarks

This paper is based on the idea that utilities from consumption depends on the consumers' state of health. In particular, it assumes that consumers may lose some of their abilities to enjoy consumption throughout their lifetime. The author is unaware of any formal empirical study that either directly supports or refutes this assumption. But it seems to be a fact of life that can be confirmed by introspection with relative ease.

Consider a person who got seriously injured from a car accident. It is not uncommon that people with traumatic brain injury lose taste and smell. Suppose that the person made a reservation at an expensive restaurant before the accident. It is not difficult to imagine then how this new disability will affect the utility from consuming the restaurant meal. Or consider someone who lost his or her hearing. The value of a New York Philharmonic ticket, for example, must be different before and after the accident.

Hyperbolic discounting has often been associated with dynamically inconsistent behavior. Failing to quit smoking or saving too little for retirement, for example, has often been deemed as a consequence of hyperbolic discounting. The government then, through taxes and regulations, may increase consumers' welfare by either discouraging smoking or encouraging saving. What this paper shows, however, is that hyperbolic discounting can be compatible with a fully rational and dynamically consistent behavior. This suggests that, contrary to common perception, paternalistic intervention by the government may not be necessary after all.[3]

Research has found several different explanations of hyperbolic discounting. These theories tend to be supported by some empirical evidence but often have conflicting implications as well. It seems possible that hyperbolic discounting is a complex phenomenon with multiple underlying mechanisms. It is hoped that the idea presented in this paper will enrich our collective understanding of the problem.

## References


Azfar, O. (1999) "Rationalizing hyperbolic discounting" *Journal of Economic Behavior and Organization* **38**, 245-52.

Becker, G. and C. Mulligan (1997) "The endogenous determination of time preference" *Quarterly Journal of Economics* **112**, 729-58.

Dasgupta, P. and E. Maskin (2005) "Uncertainty and hyperbolic discounting" *American Economic Review* **95**, 1290-9.

Frederick, S., Loewenstein, G. and T. O'Donoghue (2002) "Time discounting and time preference: a critical review" *Journal of Economic Literature* **40**, 351-401.

Gurland, J. and J. Sethuraman (1995) "How pooling failure data may reverse increasing failure rates" *Journal of the American Statistical Association* **90**, 1416-23.

Halevy, Y. (2008) "Strotz meets Allais: diminishing impatience and the certainty effect" *American Economic Review* **98**, 1145-62.

Hayashi, T. (2003) "Quasi-stationary cardinal utility and present bias" *Journal of Economic Theory* **112**, 343-52.

Karni, E. and D. Schmeidler (1991) "Atemporal dynamic consistency and expected utility theory" *Journal of Economic Theory* **54**, 401-8.

Laibson, D. (1997) "Golden eggs and hyperbolic discounting" *Quarterly Journal of Economics* **112**, 443-77.

Machina, M. (1989) "Dynamic consistency and non-expected utility models of choice under uncertainty" *Journal of Economic Literature* **27**, 1622-68.

O'Donoghue, T. and M. Rabin (1999) "Doing it now or later" *American Economic Review* **89**, 103-24.



[3] The author thanks the anonymous referee for suggesting this point.

Ok, E. and Y. Masatlioglu (2007) "A theory of (relative) discounting" *Journal of Economic Theory* **137**, 214-45.

Phelps, E. and R. Pollak (1968) "On second best national saving and game equilibrium growth" *Review of Economic Studies* **35**, 185-99.

Proschan, F. (1963) "Theoretical explanation of observed decreasing failure rate" *Technometrics* **5**, 375-83.

Read, D. (2001) "Is time-discounting hyperbolic or subadditive?" *Journal of Risk and Uncertainty* **23**, 5-32.

Rubinstein, A. (2003) '"Economics and psychology"? the case of hyperbolic discounting" *International Economic Review* **44**, 1207-16.

Saito, K. (2011) "Strotz meets Allais: diminishing impatience and the certainty effect: comment" *American Economic Review* **101**, 2271-5.

Sozou, P. (1998) "On hyperbolic discounting and uncertain hazard rates" *Proceedings of the Royal Society of London: Biological Sciences* (Series B) **265**, 2015-20.

Strotz, R. (1955) "Myopia and inconsistency in dynamic utility maximization" *Review of Economic Studies* **23**, 165-80.

Thaler, R. (1981) "Some empirical evidence on dynamic inconsistency" *Economics Letters* **8**, 201-7.

Yaari, M. (1965) "Uncertain lifetime, life insurance, and the theory of the consumer" *Review of Economic Studies* **32**, 137-50.

## Appendix

*Proof of Proposition 1*

Let $q_i(\boldsymbol{x}) = q(\boldsymbol{e_i}) = \hat{q}_i$. The expected decay factor in this case is given by

$$\delta_{\boldsymbol{x}}(t) = \sum_i e^{-\hat{q}_i t}\delta_i.$$

Then,

$$\frac{d^2(\ln \delta_{\boldsymbol{x}}(t))}{dt^2} = \frac{\left(\sum_i \delta_i (\hat{q}_i)^2 e^{-\hat{q}_i t}\right)\left(\sum_i \delta_i e^{-\hat{q}_i t}\right) - \left(\sum_i \delta_i \hat{q}_i e^{-\hat{q}_i t}\right)^2}{\left(\sum_i \delta_i e^{-\hat{q}_i t}\right)^2}$$
$$\propto \sum_{i,i'} \delta_i \delta_{i'} (\hat{q}_i - \hat{q}_{i'})^2 e^{-(\hat{q}_i + \hat{q}_{i'})t} \geq 0.$$

The last inequality is strict if $\hat{q}_i \neq \hat{q}_{i'}$ as claimed. □

*Proof of Proposition 2*

Define $\Omega(\boldsymbol{x})$ be the set of functions $f(t;\boldsymbol{x})$ such that

$$f(t;\boldsymbol{x}) = \sum_{\boldsymbol{y}\le\boldsymbol{x}} b(\boldsymbol{y})e^{-q(\boldsymbol{y})t}.$$

*Lemma* $\delta_{\boldsymbol{x}}(t) \in \Omega(\boldsymbol{x})$.

*Proof* The proof is by induction. For $\|\boldsymbol{x}\| = 2$,

$$\begin{aligned}
\delta_{\boldsymbol{x}}(t) &= e^{-q(\boldsymbol{x})t}\int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}}\,\delta_{\boldsymbol{y}}(u)e^{q(\boldsymbol{x})u}du + \boldsymbol{\delta}^T\boldsymbol{x}e^{-q(\boldsymbol{x})t}\\
&= e^{-q(\boldsymbol{x})t}\int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}}\,e^{-q(\boldsymbol{y})u}\boldsymbol{\delta}^T\boldsymbol{y}e^{q(\boldsymbol{x})u}du + \boldsymbol{\delta}^T\boldsymbol{x}e^{-q(\boldsymbol{x})t}\\
&= e^{-q(\boldsymbol{x})t}\left[\sum_{\boldsymbol{y}}\frac{q_{\boldsymbol{xy}}\boldsymbol{\delta}^T\boldsymbol{y}}{q(\boldsymbol{x})-q(\boldsymbol{y})}e^{\{q(\boldsymbol{x})-q(\boldsymbol{y})\}t} - \sum_{\boldsymbol{y}}\frac{q_{\boldsymbol{xy}}\boldsymbol{\delta}^T\boldsymbol{y}}{q(\boldsymbol{x})-q(\boldsymbol{y})}\right] + \boldsymbol{\delta}^T\boldsymbol{x}e^{-q(\boldsymbol{x})t}\\
&= \sum_{\boldsymbol{y}}\frac{q_{\boldsymbol{xy}}\boldsymbol{\delta}^T\boldsymbol{y}}{q(\boldsymbol{x})-q(\boldsymbol{y})}e^{-q(\boldsymbol{y})t} + \left[\boldsymbol{\delta}^T\boldsymbol{x} - \sum_{\boldsymbol{y}}\frac{q_{\boldsymbol{xy}}\boldsymbol{\delta}^T\boldsymbol{y}}{q(\boldsymbol{x})-q(\boldsymbol{y})}\right]e^{-q(\boldsymbol{x})t}\\
&= \frac{q_{\boldsymbol{xe_i}}\delta_i}{q(\boldsymbol{x})-q(\boldsymbol{e_i})}e^{-q(\boldsymbol{e}_i)t} + \frac{q_{\boldsymbol{xe_j}}\delta_j}{q(\boldsymbol{x})-q(\boldsymbol{e_j})}e^{-q(\boldsymbol{e}_j)t} + \left[\boldsymbol{\delta}^T\boldsymbol{x} - \sum_{\boldsymbol{y}}\frac{q_{\boldsymbol{xy}}\boldsymbol{\delta}^T\boldsymbol{y}}{q(\boldsymbol{x})-q(\boldsymbol{y})}\right]e^{-q(\boldsymbol{x})t}
\end{aligned}$$

where $\boldsymbol{e_i} + \boldsymbol{e_j} = \boldsymbol{x}$. The third equality follows from the assumption that $q(\cdot)$ is strictly increasing. Suppose now that $\delta_{\boldsymbol{x}}(t) \in \Omega(\boldsymbol{x})$ for $\|\boldsymbol{x}\| = k$. But this implies, for $\|\boldsymbol{x}\| = k+1$,

$$\begin{aligned}
\delta_{\boldsymbol{x}}(t) &= e^{-q(\boldsymbol{x})t}\int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}}\,\delta_{\boldsymbol{y}}(u)e^{q(\boldsymbol{x})u}du + \boldsymbol{\delta}^T\boldsymbol{x}e^{-q(\boldsymbol{x})t}\\
&= e^{-q(\boldsymbol{x})t}\int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}}\sum_{\boldsymbol{z}\le\boldsymbol{y}} b(\boldsymbol{z})e^{-q(\boldsymbol{z})u}\,e^{q(\boldsymbol{x})u}du + \boldsymbol{\delta}^T\boldsymbol{x}e^{-q(\boldsymbol{x})t}\\
&= e^{-q(\boldsymbol{x})t}\left[\sum_{\boldsymbol{y}}\sum_{\boldsymbol{z}\le\boldsymbol{y}} q_{xy}\frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}e^{\{q(\boldsymbol{x})-q(\boldsymbol{z})\}t} - \sum_{\boldsymbol{y}}\sum_{\boldsymbol{z}\le\boldsymbol{y}} q_{xy}\frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}\right] + \boldsymbol{\delta}^T\boldsymbol{x}e^{-q(\boldsymbol{x})t}\\
&= \sum_{\boldsymbol{y}<\boldsymbol{x}}\sum_{\boldsymbol{z}\le\boldsymbol{y}} q_{xy}\frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}e^{-q(\boldsymbol{z})t} + \left[\boldsymbol{\delta}^T\boldsymbol{x} - \sum_{\boldsymbol{y}}\sum_{\boldsymbol{z}\le\boldsymbol{y}} q_{xy}\frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}\right]e^{-q(\boldsymbol{x})t}\\
&= \sum_{\boldsymbol{z}<\boldsymbol{x}}\sum_{\boldsymbol{y}<\boldsymbol{x}} q_{xy}\frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}e^{-q(\boldsymbol{z})t} + \left[\boldsymbol{\delta}^T\boldsymbol{x} - \sum_{\boldsymbol{y}}\sum_{\boldsymbol{z}\le\boldsymbol{y}} q_{xy}\frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}\right]e^{-q(\boldsymbol{x})t}\\
&= \sum_{\boldsymbol{z}<\boldsymbol{x}}\hat{b}(\boldsymbol{z})e^{-q(\boldsymbol{z})t} + \tilde{b}(\boldsymbol{x})e^{-q(\boldsymbol{x})t}
\end{aligned}$$

where $\hat{b}(\boldsymbol{z}) \equiv \sum_{y<x} q_{xy} \frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}$ and $\tilde{b}(\boldsymbol{x}) \equiv \boldsymbol{\delta}^T \boldsymbol{x} - \sum_{\boldsymbol{y}} \sum_{\boldsymbol{z} \le \boldsymbol{y}} q_{\boldsymbol{xy}} \frac{b(\boldsymbol{z})}{q(\boldsymbol{x})-q(\boldsymbol{z})}$. This shows that $\delta_{\boldsymbol{x}}(t) \in \Omega(\boldsymbol{x})$ for $\|\boldsymbol{x}\| = k+1$. □

Let $\Gamma$ be the set of functions $f(t)$ such that

$$f(t) = c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t} + \cdots + c_m e^{-\lambda_m t}, m \ge 2$$
$$0 < \lambda_1 < \cdots < \lambda_m$$
$$c_1, c_2 > 0.$$

Then,

$$\begin{aligned}
\frac{d^2(\ln f(t))}{dt^2} &= \frac{\left(\sum_i c_i (\lambda_i)^2 e^{-\lambda_i t}\right)\left(\sum_i c_i e^{-\lambda_i t}\right) - \left(\sum_i c_i \lambda_i e^{-\lambda_i t}\right)^2}{\left(\sum_i c_i e^{-\lambda_i t}\right)^2} \\
&\propto \sum_{i,i'} c_i c_{i'} (\lambda_i - \lambda_{i'})^2 e^{-(\lambda_i + \lambda_{i'})t} \\
&= e^{-(\lambda_1+\lambda_2)t} \sum_{i,i'} c_i c_{i'} (\lambda_i - \lambda_{i'})^2 e^{\{\lambda_1+\lambda_2-(\lambda_i+\lambda_{i'})\}t} \\
&\propto \sum_{i,i'} c_i c_{i'} (\lambda_i - \lambda_{i'})^2 e^{\{\lambda_1+\lambda_2-(\lambda_i+\lambda_{i'})\}t} \\
&\equiv D(t).
\end{aligned}$$

Notice that $\lambda_1 + \lambda_2 - (\lambda_i + \lambda_{i'}) < 0$ unless $i = 1$ and $i' = 2$ or $i = 2$ and $i' = 1$. Thus, $\lim_{t\to\infty} D(t) = 2c_1 c_2 (\lambda_1 - \lambda_2)^2 > 0$, and this implies that $\frac{d^2(\ln f(t))}{dt^2}$ becomes positive when $t$ is large enough. The proof is complete if it is shown that $\delta_{\boldsymbol{x}}(t) \in \Gamma$. This is immediate for $\|\boldsymbol{x}\| = 2$ because, from the proof of the lemma,

$$\delta_{\boldsymbol{x}}(t) = \frac{q_{\boldsymbol{x}\boldsymbol{e}_i}\delta_i}{q(\boldsymbol{x}) - q(\boldsymbol{e}_i)} e^{-q(\boldsymbol{e}_i)t} + \frac{q_{\boldsymbol{x}\boldsymbol{e}_j}\delta_j}{q(\boldsymbol{x}) - q(\boldsymbol{e}_j)} e^{-q(\boldsymbol{e}_j)t} + \left[\boldsymbol{\delta}^T \boldsymbol{x} - \sum_{\boldsymbol{y}} \frac{q_{\boldsymbol{xy}} \boldsymbol{\delta}^T \boldsymbol{y}}{q(\boldsymbol{x}) - q(\boldsymbol{y})}\right] e^{-q(\boldsymbol{x})t}$$

given that $q(\boldsymbol{e}_i) \neq q(\boldsymbol{e}_j)$ and $q(\cdot)$ is strictly increasing. Suppose now that $\delta_{\boldsymbol{x}}(t) \in \Gamma$ for $\|\boldsymbol{x}\| = k$, i.e.,

$$\delta_{\boldsymbol{x}}(t) = c_1(\boldsymbol{x}) e^{-\lambda_1(\boldsymbol{x})t} + c_2(\boldsymbol{x}) e^{-\lambda_2(\boldsymbol{x})t} + \cdots + c_m(\boldsymbol{x}) e^{-\lambda_m(\boldsymbol{x})t}$$

where $\lambda_1(\boldsymbol{x}) < \cdots < \lambda_m(\boldsymbol{x})$ and $c_1(\boldsymbol{x}), c_2(\boldsymbol{x}) > 0$. Then, for $\|\boldsymbol{x}\| = k+1$,

$$\begin{aligned}
\delta_{\boldsymbol{x}}(t) &= e^{-q(\boldsymbol{x})t} \int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}}\, \delta_{\boldsymbol{y}}(u) e^{q(\boldsymbol{x})u} du + \boldsymbol{\delta}^T \boldsymbol{x} e^{-q(\boldsymbol{x})t} \\
&= e^{-q(\boldsymbol{x})t} \int_0^t \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}} \left\{c_1(\boldsymbol{y}) e^{-\lambda_1(\boldsymbol{y})u} + \cdots + c_k(\boldsymbol{y}) e^{-\lambda_k(\boldsymbol{y})u}\right\} e^{q(\boldsymbol{x})u} du + \boldsymbol{\delta}^T \boldsymbol{x} e^{-q(\boldsymbol{x})t} \\
&= e^{-q(\boldsymbol{x})t} \left[\sum_{\boldsymbol{y}} q_{\boldsymbol{xy}} \left\{\sum_i \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})} e^{\{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})\}t} - \sum_i \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})}\right\}\right] + \boldsymbol{\delta}^T \boldsymbol{x} e^{-q(\boldsymbol{x})t} \\
&= \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}} \left\{\sum_i \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})} e^{-\lambda_i(\boldsymbol{y})t}\right\} + \left[\boldsymbol{\delta}^T \boldsymbol{x} - \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}} \left\{\sum_i \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})}\right\}\right] e^{-q(\boldsymbol{x})t}
\end{aligned}$$

Let

$$\Lambda \equiv \left\{\lambda_i(\boldsymbol{y}) \middle| q_{\boldsymbol{xy}} > 0, \boldsymbol{y} \neq \boldsymbol{0}\right\}$$

and $\lambda_{(q)}$ be its $q^{\text{th}}$ smallest element. Define

$$I_j = \begin{cases} 1 & \text{if } \lambda_i(\boldsymbol{y}) = \lambda_{(j)} \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\delta_{\boldsymbol{x}}(t)$ can be rewritten as

$$\delta_{\boldsymbol{x}}(t) = \sum_j d_j e^{-\lambda_{(j)} t} + \left[\boldsymbol{\delta}^T \boldsymbol{x} - \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}} \left\{\sum_i \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})}\right\}\right] e^{-q(x)t}.$$

where

$$d_j = \sum_{\boldsymbol{y}} \sum_i q_{\boldsymbol{xy}} \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})} I_j.$$

Given that $c_1(\boldsymbol{x}), c_2(\boldsymbol{x}) > 0$ and $\lambda_1(\boldsymbol{x}) < \lambda_2(\boldsymbol{x})$,

$$d_1 = \sum_{\boldsymbol{y}} \sum_i q_{\boldsymbol{xy}} \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})} I_1 = \sum_{\boldsymbol{y}} q_{\boldsymbol{xy}} \frac{c_1(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_1(\boldsymbol{y})} I_1 > 0$$

$$d_2 = \sum_{\boldsymbol{y}} \sum_i q_{\boldsymbol{xy}} \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})} I_2 = \sum_{\boldsymbol{y}} \sum_{i=1}^{2} q_{\boldsymbol{xy}} \frac{c_i(\boldsymbol{y})}{q(\boldsymbol{x}) - \lambda_i(\boldsymbol{y})} I_2 > 0.$$

Lastly, notice that $\delta_{\boldsymbol{x}}(t) \in \Omega(\boldsymbol{x})$ , which implies $\lambda_{(j)} = q(\boldsymbol{y})$ for some $\boldsymbol{y} < x$. For all $j$, therefore, $\lambda_{(j)} < q(\boldsymbol{x})$. This implies that $\delta_x(t) \in \Gamma$ for $\|x\| = k + 1$ as desired. □