Electronic Companion for "Design of Medical Reimbursement Policy and Effects of Pooling"

Abstract. This electronic companion for "Design of Medical Reimbursement Policy and Effects of Pooling" is organized as follows. Section EC.1 documents problem formulation of pooling systems and provides the auxiliary results, including the optimal solution in full pooling systems and a numerical study on inflow pooling effects. Sections EC.2, EC.3, and EC.4 include the proofs for results in our main manuscript.

Key words: Medical Services, Reimbursement Policy, Pooling, Dynamic Programming, Utility Optimization.

EC.1. Problem Formulation and Auxiliary Results

EC.1.1. Problem Formulation of Pooling Systems

In this section, we provide the formulation for the utility optimization problem in the three pooling systems, as summarized in Table 1. Let $\mathbf{m} = (m^{(1)}, m^{(2)}, ..., m^{(K)})$ be the vector of budget per capita of the K groups. In the non-pooling system, each group operates independently with its own budget. The decision variables are $(\phi_F^{(i)}, \phi_S^{(i)})$, which can vary across groups. The budget constraints are imposed for each group. Thus, the problem can be formulated as

$$U_c^{(NP)}(\mathbf{m}) = \min_{\{\phi_S^{(i)}, \phi_F^{(i)}\}_{i=1}^K} \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[u\big(l_j^{(i)}(C_j^{(i)})\big)\big] \tag{EC.1.1}$$

$$\text{s.t.} \quad \sum_{j \in \{F,S\}} h_j^{(i)} \mathsf{E}\big[\phi_j^{(i)}(C_j^{(i)})\big] = m^{(i)}, \quad \forall i = 1, 2, ..., K, \tag{EC.1.2}$$

where the net cost function $l_j^{(i)}(x) = x - \phi_j^{(i)}(x)$; $w^{(i)} = N^{(i)} / \sum_{j=1}^K N^{(j)}$ represents the population weight of group i. We consider admissible reimbursement policies $(\phi_F^{(i)}, \phi_S^{(i)}) \in \mathcal{C}$ for all groups.

In the full pooling system, we pool the budgets together and use a common reimbursement policy (ϕ_F, ϕ_S) for all groups. The budget constraint is thus based on the total expense and budget of all groups. The optimization problem is formulated as:

$$U_c^{(FP)}(\mathbf{m}) = \min_{(\phi_F, \phi_S) \in \mathcal{C}} \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[u\big(l_j(C_j^{(i)})\big)\big] \tag{EC.1.3}$$

$$\text{s.t.} \quad \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[\phi_j(C_j^{(i)})\big] = \sum_{i=1}^{K} w^{(i)} m^{(i)}. \tag{EC.1.4}$$

Finally, with monetary pooling, the central planner can set different reimbursement policies for each group based on the shared budget. This leads to the following optimization problem:

$$U_c^{(MP)}(\mathbf{m}) = \min_{\{\phi_S^{(i)},\phi_F^{(i)}\}_{i=1}^K} \sum_{i=1}^K \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[u\big(l_j^{(i)}(C_j^{(i)})\big)\big] \tag{EC.1.5}$$

$$\text{s.t.} \quad \sum_{i=1}^K \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[\phi_j^{(i)}(C_j^{(i)})\big] = \sum_{i=1}^K w^{(i)} m^{(i)}. \tag{EC.1.6}$$

When the ratio reimbursement policy is used, we only need to let $\phi_j^{(i)}(c) = r_j^{(i)} \times c$ (for NP and MP) and $\phi_j(c) = r_j \times c$ (for FP) in above formulations. Now, the decision variables become the reimbursement ratio. Accordingly, the net cost function is given by $l_j^{(i)}(c) = (1-r_j^{(i)}) \times c$ (for NP and MP) and $l_j(c) = (1-r_j)\times c$ (for FP). The reimbursement ratios fall in the interval $[0,1]$.

Next, we specify the action space for the three pooling systems in the dynamics setting, as described in Section 4.3. Let $m^{(i)}$ and $s^{(i)}$ denote the spending amount and fund level for group i, respectively. In the non-pooling system, each group operates independently. Thus, the constraint $m^{(i)} \le s^{(i)}$ hold for each group. The reimbursement policies can vary across the groups. Thus, the action space is given by:

$$\mathcal{A}^{(NP)}(\mathbf{s}) := \big\{(m^{(i)}, \phi_F^{(i)}, \phi_S^{(i)})_{i=1}^K : m^{(i)} \le s^{(i)}, \sum_{j\in\{F,S\}} h_j^{(i)} \mathsf{E}\big[\phi_j^{(i)}(C_j^{(i)})\big] = m^{(i)}, \forall i = 1,2,\dots,K\big\}. \tag{EC.1.7}$$

In the full pooling system, groups are constrained to use the same reimbursement policy (ϕ_F, ϕ_S). The medical funds are shared across all groups. The action space is given by

$$\mathcal{A}^{(FP)}(\mathbf{s}) := \big\{(m^{(i)}, \phi_F^{(i)}, \phi_S^{(i)})_{i=1}^K : \sum_{i=1}^K w^{(i)} m^{(i)} \le \sum_{i=1}^K w^{(i)} s^{(i)}, \sum_{i=1}^K \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[\phi_j(C_j^{(i)})\big] = \sum_{i=1}^K w^{(i)} m^{(i)}, \phi_j^{(i_1)} = \phi_j^{(i_2)}, \forall i_1, i_2 = 1,2,\dots,K\big\}. \tag{EC.1.8}$$

The formulation is similar to that in (EC.1.4) for the single-period model. In the monetary pooling system, groups share their medical funds and can use different reimbursement policies. The action space is:

$$\mathcal{A}^{(MP)}(\mathbf{s}) := \big\{(m^{(i)}, \phi_F^{(i)}, \phi_S^{(i)})_{i=1}^K : \sum_{i=1}^K w^{(i)} m^{(i)} \le \sum_{i=1}^K w^{(i)} s^{(i)}, \sum_{i=1}^K \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[\phi_j^{(i)}(C_j^{(i)})\big] = \sum_{i=1}^K w^{(i)} m^{(i)}\big\}. \tag{EC.1.9}$$

For the ratio reimbursement policy, we impose the functional form $\phi_j^{(i)}(c) = r_j^{(i)} \times c$ in above.

EC.1.2. Auxiliary Results

We define $\overline{C}$ and $\underline{C}$ as the upper and lower bounds of all service costs of agents. In the single-group model, they are given by $\overline{C} = \max\{\overline{C}_F, \overline{C}_S\}$ and $\underline{C} = \min\{\underline{C}_F, \underline{C}_S\}$. In the multi-group model, we define $\overline{C} = \max_{i\in\{1,\dots,K\}, j\in\{F,S\}}\{\overline{C}_j^{(i)}\}$ and $\underline{C} = \min_{i\in\{1,\dots,K\}, j\in\{F,S\}}\{\underline{C}_j^{(i)}\}$. Accordingly, we extend the cumulative distribution functions $G_j^{(i)}(x)$ to the interval $[\underline{C}, \overline{C}]$, with zero probability mass assigned to cost levels outside their original support.

EC.1.3. Optimal Reimbursement Policy for Full Pooling System

In this section, we solve the optimal reimbursement policy in the full pooling system. Full pooling combines all groups into a single pooled group with a shared budget and a common reimbursement policy. Thus, we can solve the optimal reimbursement policy using our results for single-group model, with newly defined parameters for the pooled group.

The optimization problem in the full pooling system is given in (EC.1.3), with decision variables (ϕ_F, ϕ_S). For the pooled group, we define the aggregated service incidences as:

$$h_j^{(p)} = \sum_{i=1}^{K} w^{(i)} h_j^{(i)}, \quad \forall j \in \{F, S\}.$$

The per capita budget is given by:

$$m^{(p)} = \sum_{i=1}^{K} w^{(i)} m^{(i)}.$$

Finally, we define the distribution for the aggregated service costs $C_j^{(p)}$ for $j \in \{F, S\}$. Its cumulative distribution function is given by

$$G_j^{(p)}(c) = P\{C_j^{(p)} \le c\} = \frac{1}{h_j^{(p)}} \sum_{i=1}^{K} w^{(i)} h_j^{(i)} G_j^{(i)}(c), \quad \forall j \in \{F, S\}, \tag{EC.1.10}$$

where $G_j^{(i)}$ is the cumulative distribution function of cost $C_j^{(i)}$. With these parameters for the pooled group, we formulate the following problem:

$$U_c^{(FP)}(m^{(p)}) = \min_{(\phi_F, \phi_S)\in\mathcal{C}} \quad h_F^{(p)}\mathsf{E}[u(l_F(C_F^{(p)}))] + h_S^{(p)}\mathsf{E}[u(l_S(C_S^{(p)}))] \tag{EC.1.11}$$

$$\text{s.t.} \quad h_F^{(p)}\mathsf{E}[\phi_F(C_F^{(p)})] + h_S^{(p)}\mathsf{E}[\phi_S(C_S^{(p)})] \le m^{(p)}. \tag{EC.1.12}$$

We show that the optimization problem (EC.1.11) – (EC.1.12) is equivalent to the original problem (EC.1.3) – (EC.1.4) in the full pooling system. Consider a continuous function $f_j(c) \ge 0$ with domain $[\underline{C}, \overline{C}]$ for $j \in \{F, S\}$. Its expectation is given by

$$\mathsf{E}[f_j(C_j^{(p)})] = \int_{\underline{C}}^{\overline{C}} f_j(c) dG_j^{(p)}(c) = \int_{\underline{C}}^{\overline{C}} f_j(c) d\left(\frac{1}{h_j^{(p)}} \sum_{i=1}^{K} w^{(i)} h_j^{(i)} G_j^{(i)}(c)\right)$$

$$= \frac{1}{h_j^{(p)}} \sum_{i=1}^{K} w^{(i)} h_j^{(i)} \int_{\underline{C}}^{\overline{C}} f_j(c) dG_j^{(i)}(c). \tag{EC.1.13}$$

The last equality in (EC.1.13) follows from the linearity of integration with respect to the measure, since the measure $G_j^{(p)}$ is a linear combination of $G_j^{(i)}$ by (EC.1.10).

Summing (EC.1.13) over services F and S, we have:

$$\sum_{j\in\{F,S\}} h_j^{(p)} \mathsf{E}[f_j(C_j^{(p)})] = \sum_{j\in\{F,S\}} \sum_{i=1}^{K} w^{(i)} h_j^{(i)} \int_{\underline{C}}^{\overline{C}} f_j(c) dG_j^{(i)}(c) = \sum_{j\in\{F,S\}} \sum_{i=1}^{K} w^{(i)} h_j^{(i)} \mathsf{E}[f_j(C_j^{(i)})]. \tag{EC.1.14}$$

We specify the function $f_j(c)$ in (EC.1.14) as $u(l_j(c))$ and $\phi_j(c)$, respectively. By direct substitution into (EC.1.14), we verify the equivalence of (EC.1.3)– (EC.1.4) to (EC.1.11)–(EC.1.12), respectively. In addition, the optimization problem (EC.1.11)–(EC.1.12) is identical to that in (4)–(5) with $h_j = h_j^{(p)}$; $C_j = C_j^{(p)}$; and $m = m^{(p)}$. Thus, the optimal reimbursement policies in full pooling system can be solved using the results in Propositions 1 and 2.

EC.1.4. Numerical Study on the Effects of Inflow Pooling

In this section, we conduct a numerical analysis on the effect of inflow pooling in the dynamic setting. Pooling the inflows between groups reduces the volatility in the total inflow as long as these inflows are not perfectly correlated. As shown in Lemma 4, inflow pooling can improve the system performance even the groups are totally the same. Such effect only exist in multiple-periods model.

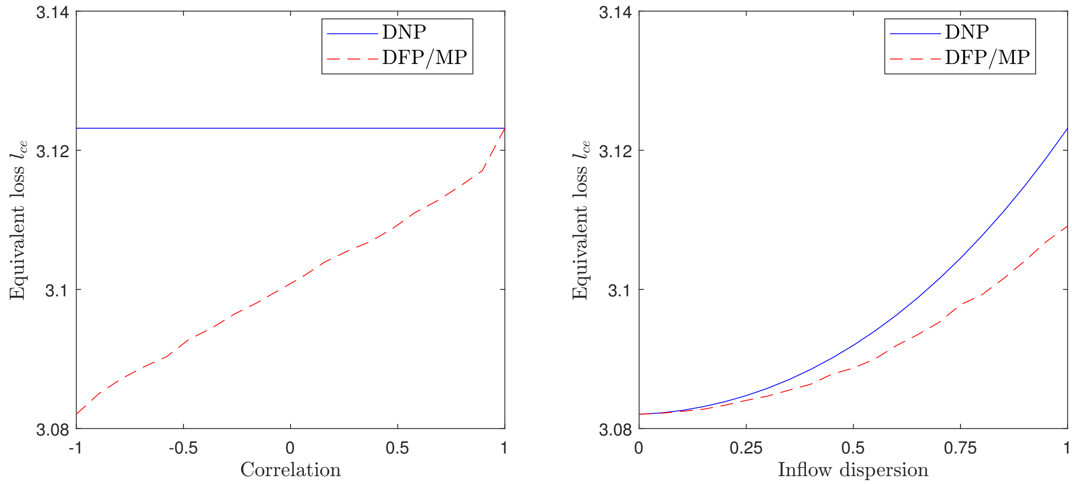
Table 1 Parameters in Numerical Study on Inflow Pooling

γ	r	β	(w_1, w_2)	$(h_F^{(1)}, h_S^{(1)})$	$(h_F^{(2)}, h_S^{(2)})$	$C_S^{(1)}$ and $C_S^{(2)}$	$C_F^{(1)}$ and $C_F^{(2)}$	q	$Corr(q^{(1)}, q^{(2)})$
2	5%	0.95	$(0.5, 0.5)$	$(0.5, 0.2)$	$(0.5, 0.2)$	$U(5, 15)$	$U(2, 6)$	$U(0, 3)$	0.5

In this numerical study, we consider two groups with parameters in Table 1. The two groups have the same population size, services incidence, service cost distribution, and inflow distribution. We further assume the same initial state for the two groups i.e. $s_0^{(1)} = s_0^{(2)} = 2$. We measure the benefit of inflow pooling by comparing the equivalent loss in the non-pooling and full pooling systems. The optimal dynamic policy is used in both systems.

Figure 1 plots the certainty equivalent losses in the non-pooling and full pooling systems. In the left panel, we fix the inflow distribution $q^{(1)}, q^{(2)} \sim U(0, 3)$ and vary the correlation coefficient of inflows, $corr(q^{(1)}, q^{(2)})$, from -1 to 1. When the correlation is -1, the total inflow is deterministic: $q^{(p)} = \mathsf{E}[q] = 1.5$. When the correlation is 1, the total inflow is the same as $q^{(1)}$ and $q^{(2)}$, i.e., $q^{(p)} \sim U(0, 3)$. In the right panel, we vary the dispersion level of the inflow, measured by $(\overline{q} - \underline{q})/(2\mathsf{E}[q])$, while fixing the expectation $\mathsf{E}[q^{(1)}] = \mathsf{E}[q^{(2)}] = 1.5$. The correlation coefficient is set at $corr(q^{(1)}, q^{(2)}) = 0.5$. A larger dispersion means the inflow is more volatile. In each panel, the blue solid (resp. red dashed) line represents the certainty equivalent loss in the non-pooling (resp. full pooling) system. The certainty equivalent loss is defined in Section 5.

Figure 1 Equivalent Loss in Non-pooling and Full Pooling Dynamic Systems



As shown in Figure 1, the inflow pooling leads to smaller certainty equivalent loss. By the left panel, the benefit from inflow pooling is higher when the two inflows are less correlated, and vanishes when the correlation approaches one. The right panel shows that the inflow pooling leads to greater improvement when the dispersion of inflows is higher. This is because pooling brings more smoothing benefit when there is more uncertainty in the inflow levels. In the extreme case where the inflow is constant in each period, the inflow pooling effect vanishes.

EC.2. Proof of Results in Section 2

In this section, we provide proof for the analytical results in Section 2 of the main manuscript, including Lemma 1, and Proposition 1–3.

We first present a lemma on the continuity of expectation operators, which will be used in our subsequent proofs.

Lemma EC.1 *Let C be a random variable with support $[\underline{C}, \overline{C}]$. Suppose that for all $c \in [\underline{C}, \overline{C}]$, the function $f(c,x)$ is continuous in x and satisfies*

$$|f(c,x)| \le \xi(c),$$

where $\xi(c)$ is an integrable function on the domain of c. Then, the function $\mathsf{E}_C[f(C,x)]$ is continuous in x .

Proof of Lemma EC.1: Let G denote the probability measure induced by random variable C. Consider any point x_0 and a sequence $\{x_n\}$ such that $x_n \to x_0$ as $n \to \infty$. To prove continuity, we will show that $\lim_{n\to\infty} \mathsf{E}_C[f(C,x_n)] = \mathsf{E}_C[f(C,x_0)]$. Since $f(c,x)$ is continuous in x for each fixed $c \in [\underline{C}, \overline{C}]$, we have

$$\lim_{n\to\infty} f(c,x_n) = f(c,x_0) \quad \forall c \in [\underline{C}, \overline{C}].$$

In addition, we have $|f(c,x)| \le \xi(c)$ for all x. Since $\xi(c)$ is integrable and does not depend on x, we can apply the Dominated Convergence Theorem to interchange the limit and the expectation as follows:

$$\lim_{n\to\infty} \mathsf{E}[f(C,x_n)] = \lim_{n\to\infty} \int_{\underline{C}}^{\overline{C}} f(c,x_n)dG(c) = \int_{\underline{C}}^{\overline{C}} \lim_{n\to\infty} f(c,x_n)dG(c) = \int_{\underline{C}}^{\overline{C}} f(c,x_0)dG(c) = \mathsf{E}[f(C,x_0)].$$

This yields:

$$\lim_{x\to x_0} \mathsf{E}[f(C,x)] = \mathsf{E}[f(C,x_0)],$$

which implies that $\mathsf{E}_C[f(C,x)]$ is continuous at x_0. This proves that $\mathsf{E}_C[f(C,x)]$ is continuous at any $x \in \mathbb{R}$. Q.E.D.

EC.2.1. Proof of Lemma 1

In this section, we prove Lemma 1, which shows that the optimization problem can be formulated at the aggregate level as in (4). The agents are indexed continuously over the interval $[0,N]$. Agent ι requires service F (S) with probability $p_{\iota F}$ ($p_{\iota S}$). By assumption, each agent requires at most one type of service. For agent ι, his service cost in service j is a random variable $C_{\iota j}$ with cumulative distribution function $G_{\iota j}$. The optimization problem can be formulated as follows:

$$U_{ind}(m) = \min_{(\phi_F,\phi_S)\in\mathcal{C}} \frac{1}{N} \sum_{j\in\{F,S\}} \int_0^N p_{\iota j} \mathsf{E}[u(l(C_{\iota j}))] d\iota \tag{EC.2.1}$$

$$\text{s.t.} \quad \frac{1}{N} \sum_{j\in\{F,S\}} \int_0^N p_{\iota j} \mathsf{E}[\phi(C_{\iota j})] d\iota = m, \tag{EC.2.2}$$

where $l_j(c) = c - \phi_j(c)$ represent the net cost for service $j \in \{F,S\}$. We transform the problem to the aggregate level using the population-based services costs and incidence. We prove that the optimization problem (EC.2.1) – (EC.2.2) is equivalent to the one in (4) – (5) of the main manuscript.

To establish the equivalence, we first introduce service incidence and cost distributions at the aggregate level. Let $p_j := \int_0^N p_{\iota j} d\iota$ represent the total incidence of service j. The aggregate service incidence is given by:

$$h_j := \frac{1}{N} p_j = \frac{1}{N} \int_0^N p_{\iota j} d\iota, \quad \forall j \in \{F,S\}.$$

The cumulative distributions of C_F and C_S are defined as follows:

$$G_j(c) = P(C_j \le c) = \int_0^N \frac{p_{\iota j}}{p_j} P(C_{\iota j} \le c) d\iota = \frac{1}{N h_j} \int_0^N p_{\iota j} G_{\iota j}(c) d\iota, \quad \forall j \in \{F, S\},$$

where ι is the agent index for the CDF $G_{\iota j}(c)$. We first show a key equality for general continuous functions based on definition of h_j and $G_j(c)$. For any continuous function $f_j \ge 0$ defined and bounded on $[\underline{C}, \overline{C}]$, we have

$$h_j \mathsf{E}[f(C_j)] = \frac{p_j}{N} \int_{\underline{C}}^{\overline{C}} f_j(c) dG_j(c) = \frac{p_j}{N} \int_0^N \int_{\underline{C}}^{\overline{C}} \frac{p_{\iota j}}{p_j} f_j(c) dG_{\iota j}(c) d\iota = \frac{1}{N} \int_0^N p_{\iota j} \mathsf{E}[f_j(C_{\iota j})] d\iota.$$

The second equality holds by Chang and Pollard (1997). In particular, if a measure is the sum of multiple measures, then the integral can be calculated as the sum of the integrals over those measures. Summing the above equation for the two services, we get:

$$\sum_{j \in \{F,S\}} h_j \mathsf{E}[f_j(C_j)] = \sum_{j \in \{F,S\}} \frac{1}{N} \int_0^N p_{\iota j} \mathsf{E}[f_j(C_{\iota j})] d\iota. \tag{EC.2.3}$$

Since the feasible reimbursement policies $\phi_j(c) \in \mathcal{C}$ and the utility function $u(c)$ are continuous, the function $u(l_j(c))$ is also continuous. Thus, letting $f_j(c) = u(l_j(c))$ and $f_j(c) = \phi_j(c)$ in (EC.2.3), we can prove the equivalence of (EC.2.1) – (EC.2.2) to (4) – (5) . That is, both the objective function and the constraints are equivalent to their aggregated form. This completes the proof of Lemma 1.

EC.2.2. Proof of Proposition 1

In this section, we give the proof of Proposition 1, which gives the form of optimal reimbursement policy. Since the decision variables in problem (4) are two functions $\phi_j(c)$ for $j \in \{F, S\}$, we first introduce a fundamental lemma base on the calculus of variations with a general measure before proceeding with our proof.

Lemma EC.2 *Let G_j be a Borel measure on $[\underline{C}, \overline{C}]$. If a continuous function $f(c)$ satisfies:*

$$\int_{\underline{C}}^{\overline{C}} f(c) \eta(c) dG_j(c) = 0,$$

for any continuous functions $\eta(c)$, then $f(c) = 0$ almost everywhere in $[\underline{C}, \overline{C}]$ with respect to the measure G_j.

Proof of Lemma EC.2: We prove by contradiction. Suppose $f(c)$ is not equal to 0 almost everywhere with respect to the measure G_j. There exists a measurable set $A = \{c \in [\underline{C}, \overline{C}] : \ f(c) \neq 0\} \subseteq [\underline{C}, \overline{C}]$ satisfying $C_A = G_j\{A\} > 0$. We can establish

$$\eta_A(c) = \mathbf{1}_{\{c \in A\}} \frac{f(c)}{C_A}.$$

We show that $\eta_A(c)$ is continuous by the definition of A and continuity of $f(c)$. If c_0 satisfies $f(c_0)=0$, then by continuity of $f(c)$, we have $\lim_{c\to c_0} f(c) = f(c_0) = 0$, thus $\lim_{c\to c_0}\eta_A(c) = 0 = \eta_A(c_0)$. If c_0 satisfies $f(c_0)\neq 0$, by continuity of $f(c)$, there exists a neighborhood $B_{c_0} = (c_0-\varepsilon, c_0+\varepsilon)$ such that $f(c)\neq 0$ for all $c\in B_{c_0}$. Therefore, $\lim_{c\to c_0}\eta_A(c) = \lim_{c\to c_0} f(c)/C_A = f(c_0)/C_A = \eta_A(c_0)$.

Plugging the form of $\eta_A(c)$ into the condition, we have:

$$\int_{\underline{C}}^{\overline{C}} f(c)\eta_A(c)dG_j(c) = \frac{1}{C_A}\int_A f(c)^2 dG_j(c) = 0.$$

Since $f(c)\neq 0$ on A and $C_A>0$, we have $\int_A f(c)^2 dG_j(c) > 0$, which leads to a contradiction. Therefore, the assumption that $f(c)\neq 0$ on a set of positive measure is false. Hence, $f(c)=0$ almost everywhere in $[\underline{C},\overline{C}]$ with respect to the measure G_j. Q.E.D.

We now derive the optimal policy using calculus of variations. Let G_j denote the distribution of cost C_j. The optimization problem in (4) in Lemma 1 can be rewritten in the following integral form:

$$\begin{aligned} U(m) = \min_{(\phi_F,\phi_S)\in\mathcal{C}} \quad & \sum_{j\in\{F,S\}} h_j \int_{\underline{C}}^{\overline{C}} u(c-\phi_j(c))dG_j(c) \\ \text{s.t.} \quad & \sum_{j\in\{F,S\}} h_j \int_{\underline{C}}^{\overline{C}} \phi_j(c)dG_j(c) = m. \end{aligned}$$

Here, $\phi_j\in\mathcal{C}$ indicates that ϕ_j is continuous and $0\leq\phi_j(c)\leq c$. To simplify the proof, we consider a more relaxed version: constraint $0\leq\phi_j(c)\leq c$ holds almost everywhere with respect to the measure $G_j(c)$. However, we will give a solution which still satisfies $0\leq\phi_F(c),\phi_S(c)\leq c$ point-wise even in this relaxation problem. We introduce the auxiliary variables for the inequality constraints as $\boldsymbol{\mu}(c) = (\mu_{0,F}(c),\mu_{0,S}(c),\mu_{1,F}(c),\mu_{1,S}(c))$ and λ for budget constraint. Then, the Lagrangian function for this relaxation optimization problem is given by:

$$\begin{aligned} \mathcal{L}(\phi_F,\phi_S,\lambda,\boldsymbol{\mu}) = & \sum_{j\in\{F,S\}} h_j \int_{\underline{C}}^{\overline{C}} u(c-\phi_j(c))dG_j(c) + \lambda\left(\sum_{j\in\{F,S\}} h_j\int_{\underline{C}}^{\overline{C}}\phi_j(c)dG_j(c) - m\right) \\ & - \sum_{j\in\{F,S\}}\int_{\underline{C}}^{\overline{C}} \mu_{j,0}(c)\phi_j(c)dG_j(c) + \sum_{j\in\{F,S\}}\int_{\underline{C}}^{\overline{C}} \mu_{j,1}(c)(\phi_j(c)-c)dG_j(c), \end{aligned} \tag{EC.2.4}$$

where λ is the multiplier for the budget constraint, and $\mu_{j,0}(c),\mu_{j,1}(c)\geq 0$ are the multipliers for the non-negativity constraints.

As the utility function is strictly convex and the feasible set for ϕ_F and ϕ_S is convex, the problem is convex. We can develop the necessary conditions, i.e., the Euler-Lagrangian equation, for the above system (Luenberger 1997). This leads to the following optimality conditions.

The first condition is the variation of the Lagrangian function to decision variable ϕ_F and ϕ_S: Consider a small perturbation $\eta_j(c)$ to $\phi_j(c)$ for service j, defined as $\phi_j(c)+\varepsilon\eta_j(c)$. Let ϕ_{-j} denote the reimbursement policy for the other service. So, we have

$$\frac{\partial}{\partial \phi_j}\mathcal{L}(\phi_j,\phi_{-j},\lambda,\boldsymbol{\mu})=0 \iff \frac{d}{d\varepsilon}\mathcal{L}(\phi_j+\varepsilon\eta_j,\phi_{-j},\lambda,\boldsymbol{\mu})\bigg|_{\varepsilon=0}=0, \quad \text{for all continuous } \eta_j.$$

Plugging (EC.2.4) into the right hand side, we have

$$\int_{\underline{C}}^{\overline{C}}\left[-h_j u'(c-\phi_j(c))+h_j\lambda-\mu_{j,0}(c)+\mu_{j,1}(c)\right]\eta(c)dG_j(c)=0, \quad \text{for all continuous } \eta_j.$$

By Lemma EC.2, the above equation is equivalent to

$$-h_j u'(c-\phi_j(c))+h_j\lambda-\mu_{j,0}(c)+\mu_{j,1}(c)=0, \quad G_j\text{-a.e. in } [\underline{C},\overline{C}] \text{ for } j\in\{F,S\}. \tag{EC.2.5}$$

The second one is the budget constraint:

$$h_F\int_{\underline{C}}^{\overline{C}}\phi_F(c)dG_F(c)+h_S\int_{\underline{C}}^{\overline{C}}\phi_S(c)dG_S(c)=m. \tag{EC.2.6}$$

Then, we have the complementary slackness condition for the inequality constraint:

$$\mu_{j,0}(c)\phi_j(c)=\mu_{j,1}(c)(c-\phi_j(c))=0, \quad G_j\text{-a.e. in } [\underline{C},\overline{C}] \text{ for } j\in\{F,S\}. \tag{EC.2.7}$$

Finally, we impose the feasible regions on the decision and constraint variables:

$$0\le\phi_j(c)\le c, \quad \mu_{j,0}(c),\ \mu_{j,1}(c)\ge 0, \quad G_j\text{-a.e. in } [\underline{C},\overline{C}] \text{ for } j\in\{F,S\}. \tag{EC.2.8}$$

As we mentioned, by the convexity of objective and linearity of constraints, the solution that satisfies the above conditions is optimal.

We verify that the following solution satisfies the above optimality conditions (EC.2.5) – (EC.2.8). The reimbursement policy is given by:

$$\phi_j^*(c)=\max\{0,c-\tau^*\}; \tag{EC.2.9}$$

and the auxiliary constraint variables are:

$$\lambda=u'(\tau^*),\ \mu_{j,0}(c)=h_j\big(u'(\tau^*)-u'(c)\big)^+,\ \mu_{j,1}(c)=0. \tag{EC.2.10}$$

In this case, the net cost is $l_j(c)=\min\{c,\tau^*\}$.

We now verify the optimality conditions. Plugging the solution (EC.2.9) and (EC.2.10) into (EC.2.5), we have

$$-h_j u'(\min\{c,\tau^*\}) + h_j u'(\tau^*) - h_j(u'(\tau^*) - u'(c))^+ = 0.$$

This follows from the fact that $u'(c)$ is strictly increasing in c. The budget constraint (EC.2.6) holds by the definition of τ^*. We establish the existence of τ^* momentarily. For the complementary slackness condition in (EC.2.7), we have

$$\mu_{j,0}(c)\phi_j(c) = h_j(u'(\tau^*) - u'(c))^+ \max\{0, c-\tau^*\} = 0.$$

Finally, it is obvious that the feasible condition (EC.2.8) holds for $\phi_j^*(c)$, λ, $\mu_{j,0}(c)$ and $\mu_{j,1}(c)$ defined in (EC.2.9) and (EC.2.10).

We now prove the existence of threshold τ^* that satisfies the budget constraint (EC.2.6). For a given level of τ, the total expenditure defined in right-hand-side of (EC.2.6) is given by:

$$B_c(\tau) = \sum_{j\in\{F,S\}} h_j \mathsf{E}[\min\{C_j,\tau\}].$$

Let the support function $\xi(c) = c$, which is integrable and satisfies $\min\{c,\tau\} \le \xi(c)$. By Lemma EC.1, the function $B_c(\tau)$ is continuous in τ. In addition, we can see that $B_c(\tau)$ decreases in τ with $B_c(0) = \bar{m}$ and $B_c(\overline{C}) = 0$. By the Intermediate Value Theorem, there exists a τ^* such that $B_c(\tau^*) = m$ for $0 < m < \bar{m}$. Thus, the solution in (EC.2.9) and (EC.2.10) satisfies all the optimal conditions and is therefore optimal.

EC.2.3. Proof of Proposition 2

In this section, we give the proof of Proposition 2, which gives the optimal ratios. The proof consists of three steps. First, we formulate the optimization problem and derive its KKT conditions. Then, we analyze these conditions in two cases based on the budget level m: when $m \le m_r$ and when $m > m_r$, where m_r is a threshold that will be defined later. In each case, we construct a solution and verify that it satisfies all KKT conditions, thereby proving its optimality.

We assume the cost indexes $b_S > b_F$ without loss of generality. The optimization problem for the ratio policy is formulated as:

$$\begin{aligned} U_r(m) = \min_{0\le r_F, r_S\le 1} \quad & h_F \mathsf{E}[u((1-r_F)C_F)] + h_S \mathsf{E}[u((1-r_S)C_S)] \\ \text{s.t.} \quad & h_F \mathsf{E}[r_F C_F] + h_S \mathsf{E}[r_S C_S] = m. \end{aligned} \tag{EC.2.11}$$

Compared to the optimization problem in (4), the policy for service $j \in \{F,S\}$ in (EC.2.11) is restricted to $\phi_j(c) = r_j \times c$, where r_j is the decision variable.

This optimization problem is a convex problem because: (1) the objective function is convex in r_F and r_S as the function $u(c)$ is convex in c; (2) the constraint is linear in r_F and r_S. Therefore, the Karush–Kuhn–Tucker (KKT) conditions are sufficient and necessary for the optimal solution. We introduce the auxiliary variables $\boldsymbol{\mu} = (\mu_{0,F}, \mu_{0,S}, \mu_{1,F}, \mu_{1,S})$ and λ, associated with the constraints $\phi_F, \phi_S \in \mathcal{C}$ and the budget constraint, respectively. Then, the Lagrangian function is:

$$\mathcal{L}(r_S, r_F, \lambda, \boldsymbol{\mu}) = \sum_{j\in\{F,S\}} h_j \mathsf{E}[u((1-r_j)C_j)] + \lambda\Big[\sum_{j\in\{F,S\}} h_j \mathsf{E}[r_j C_j] - m\Big] + \sum_{j\in\{F,S\}} [\mu_{j,1}(r_j - 1) - \mu_{j,0} r_j].$$

The KKT conditions are listed as follows. The first condition is the gradient with respect to the decision variables:

$$\frac{\partial \mathcal{L}}{\partial r_j} = h_j \mathsf{E}[-C_j u'((1-r_j)C_j)] + \lambda h_j \mathsf{E}[C_j] + \mu_{j,1} - \mu_{j,0} = 0, \quad \forall j \in \{F,S\}. \tag{EC.2.12}$$

The second condition is the budget constraint:

$$h_F \mathsf{E}[r_F C_F] + h_S \mathsf{E}[r_S C_S] = m. \tag{EC.2.13}$$

Next, we have the complementary slackness condition for the inequality constraint:

$$\mu_{j,0} r_j = \mu_{j,1}(r_j - 1) = 0 \quad \forall j \in \{F,S\}. \tag{EC.2.14}$$

Finally, we impose the feasible regions of the decision and constraint variables:

$$\mu_{j,0}, \mu_{j,1} \geq 0, \quad 0 \leq r_j \leq 1, \quad \forall j \in \{F,S\}. \tag{EC.2.15}$$

Since this optimization is a convex problem, the KKT conditions are sufficient and necessary for the optimal solution in problem (EC.2.11).

To simplify the notation, we first define function $g_j(r_j)$ as follow:

$$g_j(r_j) := \frac{\mathsf{E}[C_j u'((1-r_j)C_j)]}{\mathsf{E}[C_j]}, \quad \forall j \in \{F,S\}, \tag{EC.2.16}$$

which represents the marginal benefit of increasing the budget in service j as shown in (12). By applying Lemma EC.1, with $\xi(c) = cu'(c)/\mathsf{E}[C_j]$ dominating $cu'((1-r)c)/\mathsf{E}[C_j]$, the marginal benefit function $g_j(r_j)$ is continuous with respect to r_j. By the strict increasing property of $u'(c)$, the function $g_j(r_j)$ is strictly decreasing in r_j. Thus, the inverse function $(g_j)^{-1}(b)$ exists, which is also continuously and strictly decreasing in b. At $r_j = 0$ and $r_j = 1$, we have $g_j(0) = b_j$ and $g_j(1) = u'(0) = 0$, respectively.

Given the assumption $b_F < b_S$, we define the threshold m_r as:

$$m_r = h_S \mathsf{E}[C_S](g_S)^{-1}(b_F). \tag{EC.2.17}$$

The threshold m_r satisfies $m_r < h_S\mathsf{E}[C_S]$. By definition, m_r represents the minimum budget required to reduce the marginal benefit of service S to that of service F.

We consider two interval $m \in [0, m_r]$ and $m \in (m_r, \bar{m}]$: First, when $0 \leq m \leq m_r$, we prove that the optimal solution is given by:

$$r_F^* = 0, \quad r_S^* = \frac{m}{h_S\mathsf{E}[C_S]}, \tag{EC.2.18}$$

with auxiliary variables

$$\lambda = g_S(r_S^*), \quad \mu_{F,0} = h_F\mathsf{E}[C_F](g_S(r_S^*) - b_F), \quad \mu_{F,1} = \mu_{S,1} = \mu_{S,0} = 0. \tag{EC.2.19}$$

We verify the optimality of the solution by checking the KKT conditions one by one. Plugging the solution (EC.2.18) and (EC.2.19) into (EC.2.12), we have

$$h_F\mathsf{E}[C_F]\left(-b_F + g_S(r_S^*) - (g_S(r_S^*) - b_F)\right) = 0,$$

and

$$h_S\mathsf{E}[C_S]\left(-g_S(r_S^*) + g_S(r_S^*)\right) = 0.$$

The budget constraint (EC.2.13) holds by definition of r_F^* and r_S^* in (EC.2.18) as

$$h_F\mathsf{E}[C_F]r_F^* + h_S\mathsf{E}[C_S]r_S^* = 0 + h_S\mathsf{E}[C_S]\frac{m}{h_S\mathsf{E}[C_S]} = m.$$

The condition (EC.2.14) holds by definition as $\mu_{S,0} = \mu_{S,1} = \mu_{F,1} = 0$ and $\mu_{F,0}r_F^* = \mu_{F,0} \times 0 = 0$. Finally, since $m \leq m_r < h_S\mathsf{E}[C_S]$, by (EC.2.18) , we have $r_S^* < 1$. By the definition of m_r in (EC.2.17) and strict decreasing property of $g_S(r)$, we have

$$g_S(r_S^*) = g_S\left(\frac{m}{h_S\mathsf{E}[C_S]}\right) \geq g_S\left(\frac{m_r}{h_S\mathsf{E}[C_S]}\right) = b_F.$$

Thus, we have $\mu_{F,0} = h_F\mathsf{E}[C_F](g_S(r_S^*) - b_F) \geq 0$. So, condition (EC.2.15) holds by definition in (EC.2.18) and (EC.2.20). We have verified all the conditions (EC.2.12)–(EC.2.15), thus r_F^* and r_S^* defined in (EC.2.18) is optimal when $0 \leq m \leq m_r$.

Second, we consider $m_r < m \leq \bar{m}$: Proposition 2 states that when $m > m_r$, our solution satisfies $g_F(r_F^*) = g_S(r_S^*)$. So we prove the optimal solution is as follows. The auxiliary variables are given by

$$\mu_{F,0} = \mu_{S,0} = \mu_{F,1} = \mu_{S,1} = 0, \quad \lambda = \lambda^*; \tag{EC.2.20}$$

and the optimal ratios are given by:

$$r_S^* = (g_S)^{-1}(\lambda^*), \quad r_F^* = (g_F)^{-1}(\lambda^*). \tag{EC.2.21}$$

The parameter λ^* is solved by (EC.2.13) and (EC.2.21). We will show that $\lambda^* \in [0, b_F]$.

We verify the optimality conditions. Plugging the solution (EC.2.20) and (EC.2.21) into (EC.2.12), we have:

$$\frac{\partial \mathcal{L}}{\partial r_j} = h_j \mathsf{E}\left(g_j(r_j^*) - \lambda^*\right) = 0, \quad \forall j \in \{F, S\}.$$

The budget constraint in (EC.2.13) holds by the definition of λ^*, and we will prove the existence of λ^* and $\lambda^* \in [0, b_F]$ momentarily. Finally, the KKT conditions (EC.2.14) and (EC.2.15) hold by definitions (EC.2.20) and (EC.2.21).

We then prove that there exists λ^* satisfying (EC.2.13) under (EC.2.21). We define the total expenditure $B_r(\lambda)$ as the right-hand-side in (EC.2.13):

$$B_r(\lambda) := h_S \mathsf{E}[C_S](g_S)^{-1}(\lambda) + h_F \mathsf{E}[C_F](g_F)^{-1}(\lambda).$$

We consider the domain $\lambda \in [0, b_F]$. As $g_j^{-1}(\lambda)$ is continuously decreasing in λ, the function $B_r(\lambda)$ is continuously decreasing in λ. We have $B_r(b_F) = m_r$ and $B_r(0) = \bar{m}$. Thus, by the Intermediate Value Theorem, if $m_r < m < \bar{m}$, there exists λ^* satisfying $B_r(\lambda^*) = m$. As $\lambda^* \in [0, b_F]$, we have $r_F^* = (g_F)^{-1}(\lambda^*) \in [0, 1]$ and $r_S^* = (g_S)^{-1}(\lambda^*) \in [0, 1]$. By (EC.2.21), we have $\lambda^* = g_F(r_F^*) = g_S(r_S^*)$. In addition, as $B_r(\lambda)$ is decreasing in λ, by definition of λ^* in (EC.2.13), we have that solution λ^* decreasing in m. So, both the $r_F^* = (g_F)^{-1}(\lambda^*)$ and $r_S^* = (g_S)^{-1}(\lambda^*)$ are non-decreasing in m. We have verified all the conditions (EC.2.12)–(EC.2.15), thus r_F^* and r_S^* satisfy $g_F(r_F^*) = g_S(r_S^*) = \lambda^*$ when $m_r < m \leq \bar{m}$.

In summary, the optimal ratios satisfy: If $m \leq m_r$, the optimal solution is $(r_F^*, r_S^*) = (0, m/h_S \mathsf{E}[C_S])$; If $m > m_r$, optimal solution satisfies $g_S(r_S^*) = g_F(r_F^*)$, as stated in Proposition 2.

We also discuss a special case for the power utility function $u(l) = l^\gamma/\gamma$. Plugging the power utility function into (EC.2.17), the threshold m_r is

$$m_r = h_S \mathsf{E}[C_S](1 - (b_F/b_S)^{\frac{1}{\gamma-1}}). \tag{EC.2.22}$$

If the budget $m \in [0, m_r]$, by (EC.2.18), we have:

$$r_F^* = 0, \quad r_S^* = \frac{m}{h_S \mathsf{E}[C_S]}. \tag{EC.2.23}$$

If the budget $m \in (m_r, \bar{m}]$, by combining (EC.2.13) and (EC.2.18), we have:

$$r_F^* = \frac{1}{H}\left(m b_S^{\frac{1}{\gamma-1}} - h_S \mathsf{E}[C_S](b_S^{\frac{1}{\gamma-1}} - b_F^{\frac{1}{\gamma-1}})\right), \quad r_S^* = \frac{1}{H}\left(m b_F^{\frac{1}{\gamma-1}} + h_F \mathsf{E}[C_F](b_S^{\frac{1}{\gamma-1}} - b_F^{\frac{1}{\gamma-1}})\right), \tag{EC.2.24}$$

where the constant $H := h_S \mathsf{E}[C_S] b_F^{\frac{1}{\gamma-1}} + h_F \mathsf{E}[C_F] b_S^{\frac{1}{\gamma-1}}$. In this case, the optimal ratios can be solved explicitly, which increase piece-wise linearly in the budget level.

EC.2.4. Proof of Proposition 3

In this section, we give the proof of Proposition 3, which indicates decreasing and convexity of $U_r(m)$ and $U_c(m)$. We first give the proof of $U_c(m) \le U_r(m)$. Then, we prove that $U_c(m)$ and $U_r(m)$ are strictly decreasing in m. Finally, we prove that $U_c(m)$ and $U_r(m)$ are strictly convex in m.

To simplify the notation, we let the objective function in (4) be

$$\tilde{U}_c(\phi_F, \phi_S) := \sum_{j \in \{F,S\}} h_j \mathsf{E}[u(C_j - \phi_j(C_j))].$$

Let the objective function in (EC.2.11) be:

$$\tilde{U}_r(r_F, r_S) := h_F \mathsf{E}[u((1-r_F)C_F)] + h_S \mathsf{E}[u((1-r_S)C_S)].$$

(i) $U_c(m) \le U_r(m)$.

The optimal solution in problem (EC.2.11) satisfies $\phi_j(c) = r_j^* c$ and $0 \le r_j^* c \le c$ for $j \in \{F, S\}$. With the same budget constraint, this optimal ratio policy is feasible in problem (4). Therefore, $U_c(m) \le U_r(m)$.

(ii) $U_c(m)$ and $U_r(m)$ are strictly decreasing in $m \in [0, \bar{m})$.

Consider $m < m' < \bar{m}$. Let ϕ_F^* and ϕ_S^* denote the optimal solution to problem (4) with budget m. We will establish a new policy feasible for problem (4) with budget m' which leads less utility loss than that by ϕ_F^* and ϕ_S^*. The optimal policy ϕ_F^* and ϕ_S^* satisfy:

$$h_F \mathsf{E}[\phi_F^*(C_F)] + h_S \mathsf{E}[\phi_S^*(C_S)] = m.$$

We consider a new policy $\phi_j'(c, \varepsilon)$ based on the optimal policy

$$\phi_j'(c, \varepsilon) := \min\{c, \phi_j^*(c) + \varepsilon\}, \quad \forall c \in [\underline{C}, \overline{C}]. \tag{EC.2.25}$$

Then, for $\varepsilon > 0$ and $c \in [\underline{C}, \overline{C}]$, we have $\phi_j'(c, \varepsilon) \ge \phi_j^*(c)$ (as $c \ge \phi_j^*(c)$ and $\phi_j^*(c) + \varepsilon > \phi_j^*(c)$) and the strict inequality holds at some points because of $m < \bar{m}$ (no fully cover). By the strict increasing property of $u(c)$, we have

$$U_c(m) = \sum_{j \in \{F,S\}} h_j \mathsf{E}[u(C_j - \phi_j^*(C_j))] > \sum_{j \in \{F,S\}} h_j \mathsf{E}[u(C_j - \phi_j'(C_j, \varepsilon))]. \tag{EC.2.26}$$

We then prove there exists such ε making (EC.2.25) feasible to problem (4) with budget m'.

The total expenditure under new policy is

$$B_{c,m}(\varepsilon) := \sum_{j\in\{F,S\}} h_j \mathsf{E}[\phi_j^*(C_j,\varepsilon)].$$

Let the support function $\xi(c) = c$, which is integrable and satisfies $\min\{c,\tau\} \le \xi(c)$. By Lemma EC.1, $B_{c,m}(\varepsilon)$ continuous on ε. Furthermore, $B_{c,m}(c,0) = m$ and $B_{c,m}(c,\overline{C}) = \bar{m}$. So, by the Intermediate Value Theorem, there exists $\varepsilon^* > 0$, satisfying $B_{c,m}(\varepsilon^*) = m'$ for $m < m' < \bar{m}$. Thus, the new policy $\phi_j'(c,\varepsilon^*)$ is feasible for problem (4) with budget m', but might not be optimal:

$$U_c(m') \le \sum_{j\in\{F,S\}} h_j \mathsf{E}[u(C_j - \phi_j'(C_j,\varepsilon^*))] < U_c(m).$$

The second inequality derives from (EC.2.26). Therefore, we conclude $U_c(m') < U_c(m)$ for $m < m' \le \bar{m}$, and thus, $U_c(m)$ is strictly decreasing in m.

For $U_r(m)$ under ratio policy, we follow a similar approach. Consider $m < m' < \bar{m}$. Let r_F^* and r_S^* be the optimal ratios for problem (EC.2.11) with budget m. We construct a new policy $r_j' = (1+\varepsilon)r_j^*$ for $j \in \{F,S\}$. Note that since $r_j^* \in [0,1]$ and $m < \bar{m}$, there exists small enough $\varepsilon > 0$ such that $r_j' \in [0,1]$. The total expenditure under the new policy is

$$B_{r,m}(\varepsilon) := \sum_{j\in\{F,S\}} h_j \mathsf{E}[C_j] r_j' = (1+\varepsilon)m.$$

By the Intermediate Value Theorem, there exists $\varepsilon^* > 0$ such that $B_{r,m}(\varepsilon^*) = m'$. The rest of the proof follows the same argument in the unconstrained case, showing that $U_r(m)$ is strictly decreasing in m.

(iii) $U_c(m)$ and $U_r(m)$ are convex in m.

Consider the budgets m_1 and m_2. Let $(\phi_{F,1}^*, \phi_{S,1}^*)$ and $(\phi_{F,2}^*, \phi_{S,2}^*)$ be the solution in (4) under budget m_1 and m_2 respectively. Let $\theta \in (0,1)$ be given. Then, denote $m' = \theta m_1 + (1-\theta)m_2$ and $\phi_j' = \theta\phi_{j,1}^* + (1-\theta)\phi_{j,2}^*$. Based on the strict convexity of u, utility loss under policies ϕ_j' satisfies

$$\begin{aligned} \tilde{U}_c(\phi_F', \phi_S') &= \sum_{j\in\{F,S\}} h_j \mathsf{E}[u(C_j - \phi_j'(C_j))] \\ &< \theta \sum_{j\in\{F,S\}} h_j \mathsf{E}[u(C_j - \phi_{j,1}^*(C_j))] + (1-\theta) \sum_{j\in\{F,S\}} h_j \mathsf{E}[u(C_j - \phi_{j,2}^*(C_j))] \\ &= \theta \tilde{U}_c(\phi_{F,1}^*, \phi_{S,1}^*) + (1-\theta)\tilde{U}_c(\phi_{F,2}^*, \phi_{S,2}^*). \end{aligned} \tag{EC.2.27}$$

Due to the linearity of budget constraint, ϕ_j' is feasible in (4) with budget m':

$$\sum_{j\in\{F,S\}} h_j \mathsf{E}[\phi_j'(C_j)] = \sum_{j\in\{F,S\}} h_j \mathsf{E}[\theta\phi_{j,1}^*(C_j) + (1-\theta)\phi_{j,2}^*(C_j)] = \theta m_1 + (1-\theta)m_2.$$

Since ϕ'_j is feasible but might not be optimal, with (EC.2.27), we have

$$\begin{aligned}U_c(m') \le \tilde{U}_c(\phi'_F, \phi'_S) &< \theta \tilde{U}_c(\phi^*_{F,1}, \phi^*_{S,1}) + (1-\theta)\tilde{U}_c(\phi^*_{F,2}, \phi^*_{S,2}) \\ &= \theta U_c(m_1) + (1-\theta) U_c(m_2),\end{aligned}$$

which means $U_c(m') < \theta U_c(m_1) + (1-\theta)U_c(m_2)$. Therefore, the function $U_c(m)$ is strictly convex in m.

We then prove the strict convexity of $U_r(m)$ under the ratio policy. Let m_1 and m_2 ($m_1 < m_2$) be given. Denote $(r^*_{F,1}, r^*_{S,1})$ and $(r^*_{F,2}, r^*_{S,2})$ the optimal ratios for budget m_1 and m_2 in (EC.2.11) respectively. Then, for $j = 1, 2$, we have

$$U_r(m_j) = \tilde{U}_r(r^*_{F,j}, r^*_{S,j}). \tag{EC.2.28}$$

Let $\theta \in (0,1)$ be given. Then, denote $m' = \theta m_1 + (1-\theta)m_2$ and $r'_j = \theta r^*_{j,1} + (1-\theta) r^*_{j,2}$. The ratios (r'_F, r'_S) are feasible in problem (EC.2.11) with budget m', because

$$\sum_{j\in\{F,S\}} h_j \mathsf{E}[C_j] r'_j = \theta \sum_{j\in\{F,S\}} h_j \mathsf{E}[C_j] r^*_{j,1} + (1-\theta) \sum_{j\in\{F,S\}} h_j \mathsf{E}[C_j] r^*_{j,2} = m'.$$

From the feasibility of r'_F and r'_S, we have

$$U_r(m') \le \tilde{U}_r(r'_F, r'_S). \tag{EC.2.29}$$

Based on strict convexity of $u(c)$, the utility loss under such ratios satisfies

$$\begin{aligned}\tilde{U}_r(r'_F, r'_S) = \sum_{j\in\{F,S\}} h_j \mathsf{E}[u((1-r'_j)C_j)] &< \sum_{j\in\{F,S\}} h_j \mathsf{E}[\theta u((1-r^*_{j,1})C_j) + (1-\theta) u((1-r^*_{j,2})C_j)] \\ &= \theta \tilde{U}_r(r^*_{F,1}, r^*_{S,1}) + (1-\theta)\tilde{U}_r(r^*_{F,2}, r^*_{S,2}).\end{aligned} \tag{EC.2.30}$$

Using (EC.2.28) and (EC.2.30) in conjunction with (EC.2.29), we get

$$U_r(m') < \theta U_r(m_1) + (1-\theta) U_r(m_2).$$

This completes the proof that $U_r(m)$ is strictly convex in m.

EC.3. Proof of Results in Section 3

In this section, we provide proof for the analytical results in Section 3 of the main manuscript, including Theorem 1, and Proposition 1–3.

EC.3.1. Proof of Proposition 4

In this section, we give a proof of Proposition 4, which gives the optimal reimbursement policies in monetary pooling system. We first show that the optimal solution in the unconstrained problem has a cap structure in part (i) of Proposition 4. Then, we prove that the optimal ratio policy has the structure in part (ii) of Proposition 4.

To simplify the notation, we denote $M := \sum_{i=1}^{K} w^{(i)} m^{(i)}$ the budget and $\bar{M} := \sum_{i=1}^{K} \sum_{j\in\{F,S\}} w^{(i)} \mathsf{E}[C_j^{(i)}]$ the full cover cost.

(i) For the proof of cap policy, we can directly follow the proof in Proposition 1.

Similar to discussion in Proposition 1, we consider a relaxed version of problem (EC.1.5) – (EC.1.6): constraint $0 \le \phi_j^{(i)}(c) \le c$ only need to holds almost everywhere with respect to the measure $G_j^{(i)}(c)$ but not point-wise. However, we will show that, even in this relaxation problem, our solution still satisfies $0 \le \phi_j^{(i)}(c) \le c$ point-wise. Let $\boldsymbol{\mu}(c)$ be an vector auxiliary function for inequality constraint as

$$\boldsymbol{\mu}(c) := \left(\mu_{0,F}^{(i)}(c), \mu_{0,S}^{(i)}(c), \mu_{1,F}^{(i)}(c), \mu_{1,S}^{(i)}(c)\right)_{i=1}^{K}.$$

The Lagrangian function for this relaxed problem is written as:

$$\begin{aligned}\mathcal{L}(\phi_F, \phi_S, \lambda, \boldsymbol{\mu}) = & \sum_{i=1}^{K} \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \int_{\underline{C}}^{\overline{C}} u(c - \phi_j^{(i)}(c)) dG_j^{(i)}(c) \\ & + \lambda \left(\sum_{i=1}^{K} \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \int_{\underline{C}}^{\overline{C}} \phi_j^{(i)}(c) dG_j^{(i)}(c) - M \right) \\ & - \sum_{i=1}^{K} \sum_{j\in\{F,S\}} \int_{\underline{C}}^{\overline{C}} \mu_{j,0}^{(i)}(c) \phi_j^{(i)}(c) dG_j^{(i)}(c) + \sum_{i=1}^{K} \sum_{j\in\{F,S\}} \int_{\underline{C}}^{\overline{C}} \mu_{j,1}^{(i)}(c) (\phi_j^{(i)}(c) - c) dG_j^{(i)}(c).\end{aligned} \tag{EC.3.1}$$

Based on this Lagrangian function, we can establish the optimal conditions as follows: The first one is the variation of the Lagrangian function, i.e., the Euler-Lagrangian equation:

$$-w^{(i)} h_j^{(i)} u'(c - \phi_j^{(i)}(c)) + w^{(i)} h_j^{(i)} \lambda - \mu_{j,0}^{(i)}(c) + \mu_{j,1}^{(i)}(c) = 0, \quad G_j^{(i)}\text{-a.e. in } [\underline{C}, \overline{C}],\ \forall j \in \{F,S\},\ i = 1,2,...,K. \tag{EC.3.2}$$

The second condition is the budget constraint:

$$\sum_{i=1}^{K} w^{(i)} \sum_{j\in\{F,S\}} h_j^{(i)} \int_{\underline{C}}^{\overline{C}} \phi_j^{(i)}(c) dG_j^{(i)}(c) = M. \tag{EC.3.3}$$

Next, we have the complementary slackness conditions for the inequality constraints:

$$\mu_j^{(i)}(c) \phi_j^{(i)}(c) = \mu_j^{(i)}(c)(c - \phi_j^{(i)}(c)) = 0. \quad G_j^{(i)}\text{-a.e. in } [\underline{C}, \overline{C}],\ \forall j \in \{F,S\},\ i = 1,2,...,K. \tag{EC.3.4}$$

Finally, we impose the feasible bounds on the decision and constraint variables:

$$0 \le \phi_j^{(i)}(c) \le c,\ \mu_{j,0}^{(i)}(c), \mu_{j,1}^{(i)}(c) \ge 0, \quad G_j^{(i)}\text{-a.e. in } [\underline{C}, \overline{C}],\ \forall j \in \{F,S\},\ i = 1,2,...,K. \tag{EC.3.5}$$

As the objective function is convex by convexity of $u(c)$ and the feasible set for $\phi_j^{(i)}$ is a convex set, the solution that satisfies the above conditions is optimal.

Our proposed policy is that for $c \in [\underline{C}, \overline{C}]$, $j \in \{F,S\}$, $i = 1,2,...,K$, optimal reimbursement policy is given by:

$$\phi_j^{(i)}(c) = \max\{0, c - \tau^*\}. \tag{EC.3.6}$$

The auxiliary variables are given by:

$$\lambda = u'(\tau^*),\ \mu_{j,1}^{(i)}(c) = 0,\ \mu_{j,0}^{(i)}(c) = w^{(i)} h_j^{(i)} \big(u'(\tau^*) - u'(c)\big)^+. \tag{EC.3.7}$$

We verify each condition one by one. Plugging the solution (EC.3.6) and (EC.3.7) in the first condition (EC.3.2), for every $c \in [\underline{C}, \overline{C}]$, we have

$$-w^{(i)} h_j^{(i)} u'(\min\{c, \tau^*\}) + w^{(i)} h_j^{(i)} u'(\tau^*) - w^{(i)} h_j^{(i)} (u'(\tau^*) - u'(c))^+ = 0.$$

This follows by the increasing property of $u'(c)$ in c. Then, the budget constraint (EC.3.3) holds by the definition of τ^*. We will prove the existence of τ^* momentarily. Conditions (EC.3.4) and (EC.3.5) hold by the definitions of $\mu_{j,0}$, $\mu_{j,1}$ and ϕ_j for service j in (EC.3.6) and (EC.3.7).

We now prove that the threshold τ^* exists. The total expenditure under our proposed policy is

$$B_c^{(MP)}(\tau) := \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}[\min\{C_j^{(i)}, \tau\}].$$

With the integrable support function $\xi(c) = w^{(i)} h_j^{(i)} c > w^{(i)} h_j^{(i)} \min\{c, \tau\}$ and Lemma EC.1, total expenditure $B_c^{(MP)}(\tau)$ is continuous in τ. In addition, the function $B_c^{(MP)}(\tau)$ decreases in τ with $B_c^{(MP)}(\overline{C}) = 0$ and

$$B_c^{(MP)}(0) = \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}[C_j^{(i)}].$$

By the Intermediate Value Theorem, there exists τ^*, which satisfies $B_c^{(MP)}(\tau^*) = M$ for $0 < M < \bar{M}$, which is the budget constraint in EC.3.3. Therefore, the solution in (EC.3.6) exists and is optimal.

(ii) For the proof of optimal ratio policy, We follow the proof idea of Proposition 2. The optimization problem for monetary pooling under ratio policy is given by:

$$\begin{aligned} U_r^{(MP)}(\mathbf{m}) = \min_{\{r_S^{(i)}, r_F^{(i)}\}_{i=1}^K} \quad & \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[u\big((1 - r_j^{(i)}) C_j^{(i)}\big)\big] \\ \text{s.t.} \quad & \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[r_j^{(i)} C_j^{(i)}\big] = M, \\ & 0 \le r_F^{(i)}, r_S^{(i)} \le 1 \quad \forall i = 1,2,..,K. \end{aligned}$$

To simplify the notation, we define the function:

$$g_j^{(i)}(r) := \frac{\mathsf{E}[u'((1-r_j^{(i),*})C_j^{(i)})C_j^{(i)}]}{\mathsf{E}[C_j^{(i)}]},$$

which represents the marginal benefit of increasing the reimbursement ratio in service j for group i. By integrable support function $\xi(c) = cu'(c)/\mathsf{E}[C_j] \geq cu'((1-r_j))/\mathsf{E}[C_j]$ in Lemma EC.1, marginal benefit function $g_j^{(i)}(r)$ is continuous. In addition $g_j^{(i)}(r)$ strictly decreases in r at $[0,1]$. Thus, the inverse function $(g_j^{(i)})^{(-1)}$ exists, which is also continuously decreasing. First, to simplify the notation, let the vector of auxiliary variables for inequality constraint be

$$\boldsymbol{\mu} := \left(\mu_{0,F}^{(i)}, \mu_{0,S}^{(i)}, \mu_{1,F}^{(i)}, \mu_{1,S}^{(i)}\right)_{i=1}^{K}.$$

The Lagrangian function for monetary pooling under ratio policy is

$$\begin{aligned}\mathcal{L}(r_S, r_F, \lambda, \boldsymbol{\mu}) = &\sum_{i=1}^{K} \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[u\big((1-r_j^{(i)})C_j^{(i)}\big)\big] \\ &+\lambda \sum_{i=1}^{K} \sum_{j\in\{F,S\}} \big[w^{(i)} h_j^{(i)} \mathsf{E}\big[r_j C_j^{(i)}\big] - M\big] + \sum_{i=1}^{K} \sum_{j\in\{F,S\}} [\mu_{j,1}^{(i)}(r_j^{(i)}-1) - \mu_{j,0}^{(i)} r_j^{(i)}].\end{aligned}$$

We consider the KKT conditions: The first one is the gradient of decision variables:

$$w^{(i)} h_j^{(i)} \mathsf{E}\big[C_j^{(i)}\big]\big[-g_j^{(i)}(r_j^{(i)}) + \lambda + \frac{1}{w^{(i)} h_j^{(i)} \mathsf{E}\big[C_j^{(i)}\big]}(\mu_{j,1}^{(i)} - \mu_{j,0}^{(i)})\big] = 0, \quad \forall j \in \{F,S\}, \quad i = 1,2,..,K. \tag{EC.3.8}$$

The second one is the budget constraint:

$$\sum_{i=1}^{K} \sum_{j\in\{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[r_j C_j^{(i)}\big] = M, \tag{EC.3.9}$$

Next, we have the complementary slackness conditions for the inequality constraints:

$$\mu_{j,0}^{(i)} r_j^{(i)} = \mu_{j,1}^{(i)}(r_j^{(i)} - 1) = 0, \quad \forall j \in \{F,S\}, \quad i = 1,2,...,K. \tag{EC.3.10}$$

Finally, we impose the feasible regions of the decision and constraint variables:

$$\mu_{j,0}^{(i)}, \mu_{j,1}^{(i)} \geq 0, \quad 1 \geq r_j^{(i)} \geq 0, \quad \forall j \in \{F,S\}, \quad i = 1,2,...,K. \tag{EC.3.11}$$

The optimization problem is convex because the objective function and the constraints are linear. Therefore, the KKT conditions are necessary and sufficient for an optimal solution.

We consider the solution: For service $j \in \{F, S\}$ and $i = 1, 2, ..., K$, auxiliary variables for all services and groups are

$$\lambda = \tilde{b},\ \mu_{j,1}^{(i)} = 0,\ \mu_{j,0}^{(i)} = w^{(i)} h_j^{(i)} \mathsf{E}\big[C_j^{(i)}\big]\big(\tilde{b} - g_j^{(i)}(r_j^{(i),*})\big)^+. \tag{EC.3.12}$$

The optimal ratios are determined by

$$r_j^{(i),*} = \begin{cases} 0, & \text{if } b_j^{(i)} \le \tilde{b}, \\ (g_j^{(i)})^{-1}(\tilde{b}), & \text{if } b_j^{(i)} > \tilde{b}. \end{cases} \tag{EC.3.13}$$

We verify the optimal conditions one by one. Plugging (EC.3.12) and (EC.3.13) into (EC.3.8), we have

$$-g_j^{(i)}(r_j^{(i),*}) + \tilde{b} - (\tilde{b} - g_j(r_j^{(i),*}))^+ = 0, \quad j \in \{F, S\}, \quad i = 1, 2, .., K.$$

Then, the budget constraint (EC.3.9) holds by definition of τ^*. We will prove the existence of $\tilde{b}$ momentarily. Finally, KKT conditions (EC.3.10) and (EC.3.11) holds by (EC.3.12) and (EC.3.13).

We prove the existence of $\tilde{b}$. The total expenditure in reimbursement is

$$B_r^{(MP)}(b) := \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[r_j^{(i),*}(b) C_j^{(i)}\big].$$

We first prove that all $r_j^{(i),*}(b)$ are continuous and (weakly) decreasing in b: (1). If $b_j^{(i)} < b$, then $r_j^{(i),*}(b) = 0$ and it is continuous. (2). If $b_j^{(i)} > b$, then $r_j^{(i),*}(b) = (g_j^{(i)})^{(-1)}(b)$. The function $(g_j^{(i)})^{(-1)}$ is continuous in $r_j^{(i),*}$. (3). If $b_j^{(i)} = b$, then $\lim_{r \to 0} g_j^{(i)}(r) = b_j^{(i)} = b$. Because $r_j^{(i),*}(b)$ continuously (weakly) decreases in b for all $i = 1, 2, ..., K$ and $j \in \{F, S\}$, the function $B_r^{(MP)}(b)$ continuously decreases in b. We have

$$B_r^{(MP)}(\max\{b_j^{(i)}\}) = 0 < M,$$

and

$$B_r^{(MP)}(0) = \sum_{i=1}^{K} \sum_{j \in \{F,S\}} w^{(i)} h_j^{(i)} \mathsf{E}\big[C_j^{(i)}\big] = \bar{M} > M,$$

where $\bar{M} = \sum_{i=1}^{K} w^{(i)} \bar{m}^{(i)}$. By Intermediate Value Theorem, there exists $\tilde{b}$ satisfying $B_r^{(MP)}(\tilde{b}) = M$. So, the $\tilde{b}$ satisfies budget constraint (EC.3.9). Using this threshold $\tilde{b}$, we could get solution (EC.3.12) and (EC.3.13), which satisfy all the KKT conditions. As the problem is convex, this solution is optimal. By the optimal solution in (EC.3.13), we get the optimal reimbursement ratios which satisfy the statement (ii) in Proposition 4.

EC.3.2. Proof of Corollary 1

In this section, we prove the Corollary 1, which gives the optimal ratio policy under power utility function defined in (3). We first give the marginal benefit function $g_j^{(i)}(r)$ under the power utility function, which is used in the solution of optimal ratios in (EC.3.13). By the forms of $g_j^{(i)}(r)$ and Proposition 4, we get the result of Corollary 1. We also give the explicit form of $\tilde{b}$ in Corollary 1.

Plugging the power utility function in (3) into the marginal benefit function of increasing the budget in service j defined in (EC.2.16), the marginal benefit is given by:

$$g_j^{(i)}(r_j^{(i)}) = \frac{\mathsf{E}[u'((1-r_j^{(i)})C_j^{(i)})C_j^{(i)}]}{\mathsf{E}[C_j^{(i)}]} = b_j^{(i)}(1-r_j^{(i)})^{\gamma-1}.$$

where the $b_j^{(i)}$ is the cost index in Definition 1. So, the inverse function of $g_j^{(i)}$ is given by:

$$(g_j^{(i)})^{-1}(b) = 1 - \left(\frac{b}{b_j^{(i)}}\right)^{\frac{1}{\gamma-1}}.$$

So, by the optimal ratios in (EC.3.13), the optimal solution satisfies:

$$r_j^{(i),*} = \max\left\{1 - \frac{(\tilde{b})^{\frac{1}{\gamma-1}}}{(b_j^{(i)})^{\frac{1}{\gamma-1}}}, 0\right\}. \tag{EC.3.14}$$

where $\tilde{b}$ is determined by the budget constraint as:

$$\sum_{i=1}^{K}\sum_{j\in\{F,S\}}\left(w^{(i)}h_j^{(i)}\mathsf{E}[C_j^{(i)}]\max\left\{1-\frac{(\tilde{b})^{\frac{1}{\gamma-1}}}{(b_j^{(i)})^{\frac{1}{\gamma-1}}}, 0\right\}\right) = \sum_{i=1}^{K} w^{(i)}m^{(i)}. \tag{EC.3.15}$$

The explicit formulation of threshold $\tilde{b}$ is given by following: The forms of optimal ratios in (EC.3.14) indicate that only the services with a cost index greater than $\tilde{b}$ are reimbursed. So, we can sort the cost index and search for the threshold of service to be reimbursed. Let $b_1, b_2, ..., b_{2K}$ be the sorted cost indexes of all services across groups, where $2K$ is the total number of services since each group i has two services ($j \in \{F, S\}$). These cost indexes satisfy $b_k \geq b_{k+1}$ for all $k = 1, 2, ..., 2K-1$. The corresponding groups and services are $i_1, i_2, ..., i_{2K}$ and $j_1, j_2, ..., j_{2K}$. We define a threshold $\tilde{k}$ as:

$$\tilde{k} := \sup\left\{k : \sum_{l=1}^{k} w^{(i_l)}h_{j_l}^{(i_l)}\mathsf{E}[C_{j_l}^{(i_l)}]\left[1-\left(\frac{b_k}{b_l}\right)^{\frac{1}{\gamma-1}}\right] \leq \sum_{i=1}^{K} w^{(i)}m^{(i)}\right\}. \tag{EC.3.16}$$

The left-hand-side in the inequality in (EC.3.16) is the expenditure of reimbursement if we set the cost index threshold $\tilde{b} = b_k$. Thus, by definition in (EC.3.16), the threshold fall in region $\tilde{b} \in [b_{\tilde{k}}, b_{\tilde{k}+1})$. Thus, the service of sorted index $k > \tilde{k}$ will not be reimbursed. We can transform the budget constraint in (EC.3.15) as follows:

$$\sum_{l=1}^{\tilde{k}} w^{(i_l)}h_{j_l}^{(i_l)}\mathsf{E}[C_{j_l}^{(i_l)}]\left[1-\left(\frac{\tilde{b}}{b_l}\right)^{\frac{1}{\gamma-1}}\right] = \sum_{i=1}^{K} w^{(i)}m^{(i)}.$$

So, $\tilde{b}$ is solved as follows:

$$\tilde{b} = \left(\frac{\sum_{l=1}^{\tilde{k}} w^{(i_l)} h_{j_l}^{(i_l)} \mathsf{E}[C_{j_l}^{(i_l)}] - \sum_{i=1}^{K} w^{(i)} m^{(i)}}{\sum_{l=1}^{\tilde{k}} w^{(i_l)} h_{j_l}^{(i_l)} \mathsf{E}[C_{j_l}^{(i_l)}] \big(b_{j_l}^{(i_l)}\big)^{-\frac{1}{\gamma-1}}} \right)^{\gamma-1},$$

where $\tilde{k}$ is defined in (EC.3.16).

EC.3.3. Proof of Theorem 1

In this section, we give the proof of Theorem 1 which compares the performance of non-pooling (NP), full pooling (FP) and monetary pooling (MP) systems in different assumptions of services costs. We first prove the general inequalities in (ii-b), showing that both NP and FP systems perform no worse than MP. Then, we prove the equivalence between FP and MP in two special cases: for cap policies in (i) and for ratio policies with homogeneous costs in (ii-a).

(ii-b) $U_c^{(NP)} \geq U_c^{(MP)}$ and $U_c^{(FP)} \geq U_c^{(MP)}$.

The solution $\{\phi_S^{(i),*}, \phi_F^{(i),*}\}_{i=1}^{K}$ in non-pooling system (EC.1.1)–(EC.1.2) is feasible in monetary pooling system (EC.1.5)– (EC.1.6) because we could sum the budget constraints in (EC.1.2) for all group to get the budget constraints in (EC.1.6). Due to this feasibility, we have $U_c^{(NP)} \geq U_c^{(MP)}$. Similarly, the optimal policies ϕ_S^* and ϕ_F^* in full pooling system (EC.1.3)–(EC.1.4) are feasible in monetary pooling system (EC.1.5)–(EC.1.6) because budgets in two systems are same and we could let policies in MP be $\phi_F^{(i)} = \phi_F^*$ and $\phi_S^{(i)} = \phi_S^*$ for all $i = 1, 2, ..., K$. Due to this feasibility, we have $U_c^{(FP)} \geq U_c^{(MP)}$. For the ratios policies, the ideas of proof are the same. So, we could also conclude $U_r^{(NP)} \geq U_r^{(MP)}$ and $U_r^{(FP)} \geq U_r^{(MP)}$. Those inequalities are the general situation corresponding the statement (ii-b) in Theorem 1.

(i) $U_c^{(FP)} = U_c^{(MP)} \leq U_c^{(NP)}$ in cap policy.

The optimal cap policy is shown in part (i) of Proposition 4. In the monetary pooling system, the optimal policy is

$$\phi_j^{(i),*}(x) = \max\{x - \tau^*, 0\}, \quad \forall j \in \{F, S\}, i = 1, 2, 3, ..., K, \tag{EC.3.17}$$

where the parameter τ^* is a constant determined by (EC.1.4). The optimal cap policy in (EC.3.17) is feasible in the full pooling system. This is because the reimbursement policies in MP are homogeneous between groups i.e., $\phi_j^{(i_1),*}(x) = \phi_j^{(i_2),*}(x)$ and the budget constraints for FP and MP are the same i.e., the equivalence of (EC.1.4) to (EC.1.6). Therefore, we have $U_c^{(FP)} \leq U_c^{(MP)}$. In former result, we have $U_c^{(MP)} \leq U_c^{(FP)}$. So, we have $U_c^{(MP)} = U_c^{(FP)}$. Combining with discussion in previous of $U_c^{(NP)} \geq U_c^{(MP)}$, we conclude $U_c^{(FP)} = U_c^{(MP)} \leq U_c^{(NP)}$.

(ii-a) $U_r^{(FP)} = U_r^{(MP)} \leq U_r^{(NP)}$ in ratios policy with homogeneous services costs.

The optimal ratios are shown in part (ii) of Proposition 4. With homogeneous service cost assumption, the cost indexes are the same for all groups. Let the cost indexes be b_F and b_S to service F and S for all groups. We define the marginal benefit function, which is the same for all groups by homogeneous assumption, as

$$g_j(r) := \frac{\mathsf{E}[u'((1-r)C_j)C_j]}{\mathsf{E}[C_j]}.$$

which represents the marginal benefit of increasing the reimbursement ratio in service j for group i. By integrable support function $\xi(c) = cu'(c)/\mathsf{E}[C_j] \geq cu'((1-r_j))/\mathsf{E}[C_j]$ in Lemma EC.1, marginal benefit function $g_j(r)$ is continuous. In addition $g_j(r)$ strictly decreases in r at $[0,1]$. Thus, the inverse function $(g_j)^{-1}$ exists, which is also continuously decreasing. By Proposition 4, the optimal reimbursement ratio satisfies

$$r_j^{(i),*} = \begin{cases} 0, & \text{if } b_j \leq \tilde{b}, \\ (g_j)^{-1}(\tilde{b}), & \text{if } b_j > \tilde{b}. \end{cases}$$

This implies that

$$r_j^{(i_1),*} = r_j^{(i_2),*}, \quad \forall i_1, i_2 = 1, 2, ..., K, \text{ and } j \in \{F, S\}.$$

So, all the optimal reimbursement ratios in the monetary pooling system for service j are the same for all groups. Therefore, the optimal reimbursement ratios in the monetary pooling system are also feasible in full pooling. Due to the feasibility, we have that $U_r^{(FP)} \leq U_r^{(MP)}$. The former result shows $U_r^{(MP)} \leq U_r^{(FP)}$. So, we have $U_r^{(FP)} = U_r^{(MP)}$. Combining with the former result $U_r^{(NP)} \geq U_r^{(MP)}$, we conclude $U_r^{(FP)} = U_r^{(MP)} \leq U_r^{(NP)}$.

EC.3.4. Proof of Proposition 5

In this section, we prove the result in Proposition 5, which compares the performance of NP and MP under given parameters. We first present the formulation of utility loss under optimal ratio policy in constant services costs denoted as $U_{r,con}(m)$. Then, we give the value of $U^{(NP)}$ and $U^{(FP)}$ in our setting by using $U_{r,con}(m)$. Finally, we compare the performance $U^{(NP)}$ and $U^{(FP)}$ by difference and discuss the two cases in Proposition 5.

We first give an auxiliary lemma, which presents the utility loss under optimal ratio policy for constant service costs. This is helpful in calculating $U_r^{(FP)}$ and $U_r^{(NP)}$.

Lemma EC.3 *Under power utility loss defined in (3), if the services incidence of a group are (h_F, h_S) and the services costs are constant (c_F, c_S), then the utility loss under optimal ratio policy for the group is*

$$\gamma U_{r,con}(m) = \begin{cases} h_F c_F^\gamma + \frac{1}{h_S^{\gamma-1}}(h_S c_S - m)^\gamma, & \text{if } m \leq h_S(c_S - c_F), \\ \frac{1}{(h_F+h_S)^{\gamma-1}}(h_F c_F + h_S c_S - m)^\gamma, & \text{if } m > h_S(c_S - c_F), \end{cases} \tag{EC.3.18}$$

where m is the budget level of the group.

Proof of Lemma EC.3: The utility loss function has the power utility form $u(l) = l^\gamma/\gamma$ defined in (3). So, the optimal ratio has been solved in (EC.2.22)–(EC.2.24). Plugging the constant services cost c_F and c_S into (EC.2.22)–(EC.2.24), the threshold of reimbursing two services is $m_r = h_S(c_S - c_F)$. The optimal ratios (r_F^*, r_S^*) are given by: If $m \le h_S(c_S - c_F)$, then we have

$$r_F^* = 0, \quad r_S^* = \frac{m}{h_S c_S};$$

and if $m \ge h_S(c_S - c_F)$, then we have

$$r_F^* = \frac{1}{(h_F + h_S)c_F}[m - h_S(c_S - c_F)], \quad r_S^* = \frac{1}{(h_F + h_S)c_S}[m + h_F(c_S - c_F)].$$

Plugging optimal ratios (r_F^*, r_S^*) into objective function in ratio policy in (EC.2.11) with power utility function (3), i.e. $\gamma U_{r,con}(m) = \sum_{j\in\{F,S\}} h_j c_j^\gamma (1 - r_j^*)^\gamma$, we get the result in (EC.3.18). This completes the proof of Lemma EC.3. Q.E.D.

Lemma EC.3 gives the single period utility loss function under constant service cost. Then, we establish the utility loss $U_r^{(NP)}$ and $U_r^{(FP)}$. We first give the utility loss $U_r^{(NP)}$ in a non-pooling system. Recall the setting that both Group 1 and Group 2 have service incidence (h_F, h_S) and population weight $w^{(1)} = w^{(2)} = 0.5$. Their services cost and budget are (c_F, c_S) and m for Group 1 and $(k_c c_F, k_c c_S)$ and $k_m m$ for Group 2. By Lemma EC.3, the utility loss for Group 1 under optimal ratio $U_r^{(1)}$ is given by:

$$U_r^{(1)} = \begin{cases} h_F c_F^\gamma + \frac{1}{h_S^{\gamma-1}}(h_S c_S - m)^\gamma, & \text{if } m \le h_S(c_S - c_F), \\ \frac{1}{(h_F+h_S)^{\gamma-1}}(h_F c_F + h_S c_S - m)^\gamma, & \text{if } m > h_S(c_S - c_F). \end{cases} \tag{EC.3.19}$$

Similarly, the utility loss for Group 2 under optimal ratio $U_r^{(2)}$ is given by:

$$U_r^{(2)} = \begin{cases} h_F k_c^\gamma c_F^\gamma + \frac{1}{h_S^{\gamma-1}}(h_S k_c c_S - k_m m)^\gamma, & \text{if } k_m m \le h_S k_c(c_S - c_F), \\ \frac{1}{(h_F+h_S)^{\gamma-1}}(h_F k_c c_F + h_S k_c c_S - k_m m)^\gamma, & \text{if } k_m m > h_S k_c(c_S - c_F). \end{cases} \tag{EC.3.20}$$

Recall the definition of a non-pooling system: each group operates independently. So the total utility loss in the non-pooling system $U^{(NP)}$ is given by:

$$U_r^{(NP)} = w^{(1)} U_r^{(1)} + w^{(2)} U_r^{(2)} = \frac{1}{2}(U_r^{(1)} + U_r^{(2)}). \tag{EC.3.21}$$

We then give the utility loss $U_r^{(FP)}$ in full pooling system. The optimization problem in full pooling system under ratio policy could be formulated as:

$$U_r^{(FP)} = \min_{0\le r_F, r_S \le 1} \quad \frac{1}{2} \sum_{j\in\{F,S\}} h_j c_j^\gamma (1 + k_c)^\gamma (1 - r_j)^\gamma \tag{EC.3.22}$$

$$\text{s.t.} \quad \frac{1}{2}\sum_{j\in\{F,S\}} r_j\big(h_jc_j + h_jk_cc_j\big) = \frac{1}{2}(1+k_m)m. \tag{EC.3.23}$$

To use Lemma EC.3, we could transform the problem (EC.3.22)–(EC.3.23) into a constant service cost form as follows:

$$U_r^{(FP)} = \min_{0\le r_F, r_S\le 1} \quad \sum_{j\in\{F,S\}} h_j^{(cp)}\big(c_j^{(cp)}\big)^{\gamma}(1-r_j)^{\gamma} \tag{EC.3.24}$$

$$\text{s.t.} \quad \sum_{j\in\{F,S\}} r_j h_j^{(cp)} c_j^{(cp)} = m^{(cp)}. \tag{EC.3.25}$$

where the parameters are given by:

$$h_j^{(cp)} = \frac{1}{2}h_j\left[\frac{(1+k_c)^{\gamma}}{1+k_c^{\gamma}}\right]^{\frac{1}{\gamma-1}}, \quad c_j^{(cp)} = c_j\left(\frac{1+k_c^{\gamma}}{1+k_c}\right)^{\frac{1}{\gamma-1}}, \quad m^{(cp)} = \frac{1}{2}(1+k_m)m, \quad \forall j\in\{F,S\}. \tag{EC.3.26}$$

For any reimbursement ratio (r_S, r_F), we verify the equivalence of transformation as follows: Plugging the (EC.3.26) into objective function (EC.3.24), we have:

$$\sum_{j\in\{F,S\}} h_j^{(cp)}\big(c_j^{(cp)}\big)^{\gamma}(1-r_j)^{\gamma} = \frac{1}{2}\sum_{j\in\{F,S\}}(1-r_j)^{\gamma}(h_jc_j^{\gamma} + h_jk_c^{\gamma}c_j^{\gamma}),$$

where the right-hand side is the objective function defined in (EC.3.22). Similarly, plugging the (EC.3.26) into the budget constraint in (EC.3.25), we have

$$\sum_{j\in\{F,S\}} h_j^{(cp)}c_j^{(cp)}r_j = \frac{1}{2}\sum_{j\in\{F,S\}} r_j\big(h_jc_j + h_jk_cc_j\big),$$

and

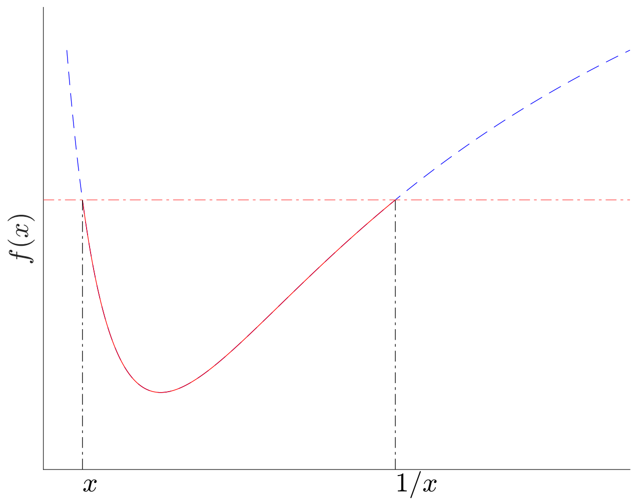
$$m^{(cp)} = \frac{1}{2}(1+k_m)m,$$

where the right-hand sides in the above two equations constitute the budget constraint defined in (EC.3.23). Therefore, we complete the proof of equivalence of problem (EC.3.24)–(EC.3.25) to problem (EC.3.22)–(EC.3.23). For the problem (EC.3.24)–(EC.3.25), by Lemma EC.3 and definition of $h_j^{(cp)}$, $c_j^{(cp)}$ and $m^{(cp)}$ in (EC.3.26), we get $U_r^{(FP)}$ as:

$$\gamma U_r^{(FP)} \tag{EC.3.27}$$
$$= \begin{cases} \frac{1}{2}h_Fc_F^{\gamma}(1+k_c^{\gamma}) + \frac{1}{2h_S^{\gamma-1}}\frac{1+k_c^{\gamma}}{(1+k_c)^{\gamma}}(h_Sc_S(1+k_c)-(1+k_m)m)^{\gamma}, & \text{if } (1+k_m)m \le h_S(1+k_c)(c_S-c_F), \\ \frac{1}{(h_F+h_S)^{\gamma-1}}\frac{1+k_c^{\gamma}}{(1+k_c)^{\gamma}}[(1+k_c)(h_Fc_F+h_Sc_S)-(1+k_m)m]^{\gamma}, & \text{if } (1+k_m)m \ge h_S(1+k_c)(c_S-c_F). \end{cases}$$

We then compare the performance of the non-pooling system and the full pooling system by the difference of $\gamma(U_r^{(NP)} - U_r^{(FP)})$. Before going to comparison, we introduce an auxiliary function $f(x)$ to simplify the

Figure 2 **Schematic Diagram of** $f(x)$



notation:

$$f(x) := \frac{1+x^{\gamma}}{(1+x)^{\gamma}}, \tag{EC.3.28}$$

It is easy to see that $f(x)$ decreases in $(0,1)$ and increases in $(1,+\infty)$. In addition, we have $f(x)=f(1/x)$. We plot the schematic diagram in Figure 2. We discuss the two cases in Proposition 5 one by one.

(i) If $m < h_S(c_S-c_F)$ and $k_m m < h_S k_c(c_S-c_F)$, then $(1+k_m)m \le h_S(1+k_c)(c_S-c_F)$. With auxiliary function $f(x)$ in (EC.3.28), by the $U^{(1)}$, $U^{(2)}$, $U^{(NP)}$, and $U^{(FP)}$ defined in (EC.3.19)– (EC.3.21) and (EC.3.27), we have

$$\begin{aligned}\gamma(U_r^{(NP)}-U_r^{(FP)}) &= \frac{1}{2h_S^{\gamma-1}}\left[(h_Sc_S-m)^{\gamma}+(h_Sk_cc_S-k_mm)^{\gamma}-f(k_c)[(1+k_c)h_Sc_S-(1+k_m)m]^{\gamma}\right]\\ &= \frac{1}{2h_S^{\gamma-1}}[(1+k_c)h_Sc_S-(1+k_m)m]^{\gamma}\left[f\left(\frac{h_Sk_cc_S-k_mm}{h_Sc_S-m}\right)-f(k_c)\right].\end{aligned}$$

To further simplify the notation we let $r_1 := (k_c-k_ma_1)/(1-a_1)$ and $a_1 := m/h_Sc_S \le 1$. So, we have

$$\gamma(U_r^{(NP)}-U_r^{(FP)}) = \frac{[(1+k_c)h_Sc_S-(1+k_m)m]^{\gamma}}{2h_S^{\gamma-1}}\left[f(r_1)-f(k_c)\right]. \tag{EC.3.29}$$

By the case assumption $(1+k_m)m \le h_S(1+k_c)(c_S-c_F)$, the first term in (EC.3.29) is positive. For the last term, we have $f(r_1)-f(k_c)>0$ if and only if $\max\{r_1,1/r_1\}>k_c>\min\{r_1,1/r_1\}$ because $f(x)$ is decreasing in $(0,1)$ and increasing in $(1,+\infty)$ with solution $f(x)-f(1/x)=0$. (see the red solid line in illustrative Figure 2).

So, we compare r_1, k_c, and $1/r_1$. By definition of $r_1=(k_c-k_ma_1)/(1-a_1)$, we have $r_1<k_c$ if and only if $k_c<k_m$. Similar, by definition of r_1, $k_c>1/r_1$ is equivalent to $k_c(k_c-k_ma_1)-(1-a_1)>0$ which has solution at positive region ($k_c > 0$) for

$$k_c > k_1 = \frac{1}{2}[k_m a_1 + \sqrt{k_m^2 a_1 + 4 - 4a_1}] \in (1, k_m).$$

Therefore, we could summarize the result of the comparison as follows:

- If $k_1 < k_m < k_c$, then $1/r_1 < 1 < k_m < k_c < r_1$. So $f(r_1) > f(k_c)$. Therefore $U_r^{(NP)} > U_r^{(FP)}$.
- If $k_1 < k_c < k_m$ then $r_1 < k_c$ and $k_c > 1/r_1$. So $f(r_1) < f(k_c)$. Therefore $U_r^{(NP)} < U_r^{(FP)}$.
- If $k_c < k_1 < k_m$, then $r_1 < k_c < 1/r_1$. So $f(r_1) > f(k_c)$. Therefore $U_r^{(NP)} > U_r^{(FP)}$.

(ii) If $m < h_S(c_S - c_F)$ and $k_m m < h_S k_c(c_S - c_F)$, then $(1 + k_m)m \le h_S(1 + k_c)(c_S - c_F)$. We follow the calculation in part (i). With auxiliary function $f(x)$, by the $U^{(1)}$, $U^{(2)}$, $U^{(NP)}$, and $U^{(FP)}$ defined in (EC.3.19)– (EC.3.21) and (EC.3.27), we have

$$\gamma U_r^{(NP)} = \frac{1}{2(h_F + h_S)^{\gamma-1}}[(h_F k_c c_F + h_S k_c c_S - k_m m)^\gamma + (h_F c_F + h_S c_S - m)^\gamma];$$

and

$$\gamma U_r^{(FP)} = \frac{1}{2(h_F + h_S)^{\gamma-1}} f(k_c)[(1 + k_c)(h_F c_F + h_S c_S) - (1 + k_m)m]^\gamma.$$

Similarly, we define $r_2 = (1 - k_m a_2)/(1 - a_2)$ and $a_2 = m/(h_S c_S + h_F c_F) \le 1$. Then, the difference in utility between these two pooling systems $\gamma(U^{(NP)} - U^{(FP)})$ is given by:

$$\gamma\left(U_r^{(NP)} - U_r^{(FP)}\right) = \frac{1}{2(h_F + h_S)^{\gamma-1}}[(1 + k_c)(h_F c_F + h_S c_S) - (1 + k_m)m]^\gamma [f(r_2) - f(k_c)]. \quad \text{(EC.3.30)}$$

The equality in (EC.3.30) has the same structure as (EC.3.29), where the first two terms are both positive by case assumption $(1 + k_m)m \le h_S(1 + k_c)(c_S - c_F)$. For the last term, we have $f(r_2) - f(k_c) > 0$ if and only if $\max\{r_2, 1/r_2\} > k_c > \min\{r_2, 1/r_2\}$.

So, by definition of $r_2 = (1 - k_m a_2)/(1 - a_2)$, we have $r_2 < k_c$ if and only if $k_c < k_m$. By definition of r_2, inequality $k_c > 1/r_2$ is equivalent to $k_c(k_c - k_m a_2) - (1 - a_2) > 0$, which has solution at positive region for

$$k_c > k_2 = \frac{1}{2}[k_m a_2 + \sqrt{k_m^2 a_2 + 4 - 4a_2}] \in (1, k_m).$$

We could similarly summarize the analysis like (i): if $k_c < k_2$ or $k_c > k_m$, then $f(r_2) > f(k_c)$, so $U_r^{(NP)} > U_r^{(FP)}$; If $k_2 < k_c < k_m$, then $f(r_2) < f(k_c)$, so $U_r^{(NP)} < U_r^{(FP)}$.

EC.4. Proof of Results in Section 4

In this section, we provide proof for the analytical results in Section 4 of the main manuscript, including Lemma 2 – 4, Theorem 2–4, and Proposition 6.

In the dynamic model, we extend the domain of $U_r(m)$ and $U_c(m)$ to $[0, \infty)$. We let $U_r(m) = U_c(m) = 0$ for $m > \bar{m}$ as it already achieves full cover. In the following proof, we will sometimes use $U(m)$ to denote the single-period utility loss. This means that the proof applies to both the cap and ratio reimbursement policies, i.e., for both $U_r(m)$ and $U_c(m)$.

We first introduce an auxiliary lemma for the derivative of value function $v'(x)$ (if exists) based on the following iteration.

Lemma EC.4 *Consider the following iteration of $v_k(x)$ with initial value $v_0(x) = 0$ for all $x \in \mathcal{S}$,*

$$v_{k+1}(x) = \min_{m\in[0,x]} \left[U(m) + \beta \mathsf{E}_q[v_k((1+r)(x-m)+q)]\right], \quad \forall x \in \mathcal{S}. \tag{EC.4.1}$$

Similar to (EC.4.18), we also define

$$m_k(x) = \inf \arg\min_{0\le m\le x} \left\{U(m) + \beta \mathsf{E}_q[v_k((1+r)(x-m)+q)]\right\}. \tag{EC.4.2}$$

i.e., $m_k(x)$ is the smallest minimizer in (EC.4.1). Then, for each k and x satisfying $m_k(x) > 0$, we have

$$v'_{k+1}(x) = U'(m_k(x)).$$

Thus, for $v(x) = \lim_{k\to\infty} v_k(x)$, we have $v'(x) = U'(m^*(x))$ if x satisfying $m^*(x) > 0$.

Proof of Lemma EC.4: Consider each $k = 1, 2, \ldots$, by (EC.4.1) and (EC.4.2), we have

$$v_{k+1}(x) = U(m_k(x)) + \beta \mathsf{E}_q[v_k((1+r)(x-m_k(x))+q)].$$

The derivative (if exists) satisfies

$$v'_{k+1}(x) = U'(m_k(x))m'_k(x) - (1-m'_k(x))\beta(1+r)\mathsf{E}_q[v'_k((1+r)(x-m_k(x))+q)]. \tag{EC.4.3}$$

First, if $m_k(x) = x$, then we have $v_{k+1}(x) = U'(x) = U'(m_k(x))$. Otherwise, if $0 < m_k(x) < x$, then the spending optimal $m_k(x)$ satisfies the first order condition of (EC.4.1) as follows:

$$U'(m_k(x)) - \beta(1+r)\mathsf{E}_q[v'_k((1+r)(x-m_k(x))+q)] = 0. \tag{EC.4.4}$$

Plugging the first order condition (EC.4.4) into the derivative of $v'_{k+1}(x)$ in (EC.4.3), we have

$$v'_{k+1}(x) = U'(m_k(x))m'_k(x) - (1-m'_k(x))U'(m_k(x)) = U'(m_k(x)).$$

By above two case, $v'_{k+1}(x) = U'(m_k(x))$ holds when state x satisfies $m_k(x) > 0$. Taking the limitation of k, we have $v'(x) = U'(m^*(x))$ holds when state x satisfies $m^*(x) > 0$. Q.E.D.

EC.4.1. Proof of Lemma 2

In this section, we prove Lemma 2, which shows that with a given spending amount for current period, the optimal reimbursement policy in the dynamic model is identical to the single-period model in Propositions 1 and 2. First, we suppose the optimal spending amount to state $m^*(s)$ is given. Then, we consider a relaxed problem where the reimbursement policy can vary in each period, even with the same state variable s. Finally, we show the policy in Propositions 1 and 2 is one of the optimal policies in this relaxed problem and thus is optimal in the original problem.

Let the optimal policy $\pi^*(s) = (m^*(s), \phi^*_F(;s), \phi^*_S(;s))$ for the problem (24) be given. We consider the optimal policy $\phi^*_F(;s)$ and $\phi^*_S(;s)$ under spending amount $m^*(s)$. Under the optimal policy π^*, the transition of states defined in (22) is only determined by the spending amount $m^*(s)$ but not related to the form of $\phi^*_F(;s), \phi^*_S(;s)$. That is, for any initial state s_0, the subsequent states are given by:

$$s_{t+1} = \psi(s_t, m_t; q_{t+1}) = (1+r)(s_t - m^*(s_t)) + q_{t+1}, \quad \forall t \geq 0. \tag{EC.4.5}$$

We consider the optimization problem in (24) for a given path of inflow $\{q_t\}_{t=0}^{\infty}$: Denote $\mathbf{q} = (q_0, q_1, \dots)$ as the sequence of inflow. Then, by (EC.4.5), the state sequence $\{s_t\}_{t=0}^{\infty}$ can be fully determined. The action space (20) and (21) restrict the policy to be the same under the same state s. We consider a relaxed version of the optimization problem, where different policies $\phi_{S,t}$ and $\phi_{F,t}$ can be used in each period, even if the state s_t is the same. On the other hand, we still require that the optimal spending amount is given by $m^*(s)$ in each period. Then, the total discounted utility loss under the relaxed problem $\hat{u}(\mathbf{q})$ is given by:

$$\begin{aligned}
\hat{u}(\mathbf{q}) = \min_{\{(\phi_{S,t}, \phi_{F,t})\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \big(h_F \mathsf{E}[u(l_{F,t}(C_F))] + h_S \mathsf{E}[u(l_{S,t}(C_S))]\big) \\
\text{s.t.} \quad & (\phi_{S,t}, \phi_{F,t}) \in \Phi(m^*(s_t)), \quad t = 0, 1, 2, \dots, \\
& l_{j,t}(x) = x - \phi_{j,t}(x) \quad \forall j \in \{F, S\}, \quad t = 0, 1, 2, \dots .
\end{aligned}$$

As the state sequence $\{s_t\}_{t=0}^{\infty}$ has been determined, this optimization problem can be decomposed into the sum of utility optimization problem in each period as

$$\begin{aligned}
\min_{(\phi_{S,t}, \phi_{F,t}) \in \Phi(m^*(s_t))} \quad & h_F \mathsf{E}[u(l_{F,t}(C_F))] + h_S \mathsf{E}[u(l_{S,t}(C_S))] \\
\text{s.t.} \quad & l_{j,t}(x) = x - \phi_{j,t}(x), \quad \forall j \in \{F, S\}.
\end{aligned}$$

This decomposed problem has been solved in Propositions 1 and 2. Therefore, we have:

$$\hat{u}(\mathbf{q}) = \sum_{t=0}^{\infty} \beta^t U(m^*(s_t)),$$

where the $U(m)$ is the single period utility loss function. Note that optimal policy (ϕ_S^*, ϕ_F^*) in Propositions 1 and 2 is stationary to the budget $m^*(s)$, i.e. we have the same optimal policy under the same budget constraint. In the relaxed problem, the spending amounts are the same $m^*(s)$ for the same state s, thus optimal policies ϕ_S^* and ϕ_F^* for relaxed problem are also stationary to state s. This holds for every path $\{q_t\}_{t=0}^{\infty}$. By the definition of the relaxed problem, we have

$$v(s) \geq \mathsf{E}_{\mathbf{q}}[\hat{u}(\mathbf{q})] = \mathsf{E}_{\mathbf{q}}\big[\sum_{t=0}^{\infty} \beta^t U(m^*(s_t))|s_0 = s\big].$$

This inequality holds because the relaxed problem allows for a larger set of feasible policies than the original problem, potentially leading to a lower (better) objective value. As we have mentioned, the optimal policies ϕ_S^*, ϕ_F^* in Proposition 1 or 2, which are optimal for the relaxed problem, are also feasible for the problem (24) because they are stationary policy. This implies that the lower bound from the relaxed problem can actually be achieved by a feasible policy, and thus ϕ_S^*, ϕ_F^* are also the optimal policy for the original problem. Therefore, we have:

$$\begin{aligned} v(s) &= \mathsf{E}_{\mathbf{q}}\big[\sum_{t=0}^{\infty} \beta^t U(m^*(s_t))|s_0 = s\big] \\ \text{s.t.} &\quad s_{t+1} = (1+r)(s_t - m^*(s_t)) + q_{t+1}, \quad \forall t \geq 0. \end{aligned}$$

This implies that, under a given $m^*(s)$, the optimal reimbursement policies $\phi_S(;s)$ and $\phi_F(;s)$ can be chosen as the solutions given in Propositions 1 and 2 with budget $m^*(s)$.

EC.4.2. Proof of Proposition 6

In this section, we give the proof of Proposition 6, which establishes that the value function $v(s)$ is convex and decreasing in s. We prove it by mathematical induction. Consider the value iteration function $v_k(x)$ with initial value $v_0(x) = 0$ for all $x \in \mathcal{S}$ and updating rule:

$$v_{k+1}(x) = \min_{m \in [0,x]} \left[U(m) + \beta \mathsf{E}_q v_k(\psi(x, m; q))\right], \quad \forall x \in \mathcal{S}. \tag{EC.4.6}$$

By definition, $\lim_{k\to\infty} v_k(x) = v(x)$. So, it suffices to prove that $v_k(x)$ is decreasing and convex in x for every k. Initially, $v_0(x) = 0$ is weakly decreasing and convex in x. Assuming $v_k(x)$ is decreasing and convex, we then show $v_{k+1}(x)$ is also decreasing and convex in x.

(i) $v_{k+1}(x)$ is decreasing in x.

Consider any state x and $x' > x$. Let m^* be optimal spending amount for state x in iteration as follows:

$$m^* \in \arg\min_{m \in [0,x]} \left\{U(m) + \beta \mathsf{E}_q v_k(\psi(x, m; q))\right\}.$$

Then, the new value function $v_{k+1}(x)$ is based on m^*:

$$v_{k+1}(x) = U(m^*) + \beta \mathsf{E}_q v_k(\psi(x, m^*; q)). \tag{EC.4.7}$$

For state x', the spending amount $m^* + (x' - x)$ is also feasible since $m^* + x' - x \leq x + x' - x = x'$ but might not be optimal. Thus, by definition of iteration value function in (EC.4.6), we have

$$v_{k+1}(x') \leq U(m^* + (x' - x)) + \beta \mathsf{E}_q v_k(\psi(x', m^* + (x' - x); q)). \tag{EC.4.8}$$

For the last term in (EC.4.8), by definition of $\psi(x, m; q) = (1 + r)(x - m) + q$ in (22), we have

$$\mathsf{E}_q v_k(\psi(x', m^* + (x' - x); q)) = \mathsf{E}_q[v_k((1 + r)(x - m^*) + q)] = \mathsf{E}_q v_k(\psi(x, m^*; q)). \tag{EC.4.9}$$

Combining (EC.4.7), (EC.4.8) and (EC.4.9), we derive

$$v_{k+1}(x') - v_{k+1}(x) \leq U(m^* + x' - x) - U(m^*) \leq 0.$$

The last inequality holds because $U(m)$ decreases in m by Proposition 3. Thus, we conclude $v_{k+1}(x') \leq v_{k+1}(x)$, which completes the proof that $v_{k+1}(x)$ decreases in x.

(ii) $v_{k+1}(x)$ is convex in x.

To simplify the notation, we define the objective function $h_k(m, x)$ as follows:

$$h_k(m, x) := U(m) + \beta \mathsf{E}_q v_k((1 + r)(x - m) + q), \tag{EC.4.10}$$

which represents the objective function in iteration. Consider state x_1 and x_2 with optimal spending amount m_1^* and m_2^* respectively. We have

$$v_{k+1}(x_j) = h_k(m_j^*, x_j), \quad \forall j = 1, 2. \tag{EC.4.11}$$

Let $\theta \in [0, 1]$ be given. Denote $m' = \theta m_1^* + (1 - \theta) m_2^*$ and $x' = \theta x_1 + (1 - \theta) x_2$. By the convexity of U in Proposition 3, we have

$$U(m') = U(\theta m_1^* + (1 - \theta) m_2^*) \leq \theta U(m_1^*) + (1 - \theta) U(m_2^*). \tag{EC.4.12}$$

Since $\psi(x, m; q) = (1 + r)(x - m) + q$ is linear in m, x and q, we have $\psi(m', x'; q) = \theta \psi(m_1^*, x_1; q) + (1 - \theta) \psi(m_2^*, x_2; q)$. By the induction assumption of convexity of v_k, we have:

$$\mathsf{E}_q[v_k(\psi(m', x'; q))] = \mathsf{E}_q[v_k(\theta \psi(m_1^*, x_1; q) + (1 - \theta) \psi(m_2^*, x_2; q))]$$

$$\leq \theta \mathsf{E}_q[v_k(\psi(m_1^*, x_1; q))] + (1-\theta)\mathsf{E}_q[v_k(\psi(m_2^*, x_2; q))]. \tag{EC.4.13}$$

Summing (EC.4.12) and (EC.4.13), we obtain

$$U(m') + \beta \mathsf{E}_q[v_k(\psi(m', x'; q))] \leq \theta[U(m_1^*) + \mathsf{E}_q[v_k(\psi(m_1^*, x_1; q))]] + (1-\theta)[U(m_2^*) + \mathsf{E}_q[v_k(\psi(m_2^*, x_2; q))]].$$

By (EC.4.10) and (EC.4.11), this inequality is equivalent to the following:

$$\theta h_k(m_1^*, x_1) + (1-\theta) h_k(m_2^*, x_2) \geq h_k(m', x'). \tag{EC.4.14}$$

As $m' = \theta m_1^* + (1-\theta) m_2^* \leq \theta x_1 + (1-\theta) x_2$, the spending amount m' is feasible in state x' but might not be optimal. Thus, we have

$$v_{k+1}(x') \leq h_k(m', x') \leq \theta v_{k+1}(x_1) + (1-\theta) v_{k+1}(x_2).$$

As this inequality holds for every $\theta \in [0,1]$, x_1 and x_2, iteration function $v_{k+1}(x)$ is convex in x.

From (i) and (ii), $v_{k+1}(x)$ is convex and decreasing in x. By mathematical induction, we have $v_k(x)$ is convex and decreasing in x for all the k. Therefore, value function $v(x)$ is convex and decreasing in x.

EC.4.3. Proof of Theorem 2

In this section, we give the proof of Theorem 2. The proof consists of four parts: First, we prove $m^*(x)$ is increasing in x. Then, we establish the continuity of $m^*(x)$ and $v(x)$. Following this, we address the general case of statement (i) and (ii) in Theorem 2, and consider the special case where $\beta(1+r) \leq 1$. Finally, we prove the optimal spending amount satisfies $m^*(\bar{m}) < \bar{m}$.

Before going to the main proof, we establish the following auxiliary inequalities. Let a_1, a_2 and a_m be real numbers satisfying $0 < a_m < a_1 < a_2$. Then we have

$$U(a_2 - a_m) - U(a_2) \leq U(a_1 - a_m) - U(a_1). \tag{EC.4.15}$$

Inequality (EC.4.15) holds as one-period utility function $U(x)$ is convex and decreasing. Similarly we have,

$$\mathsf{E}_q v((1+r)(a_2 - a_m) + q) - \mathsf{E}_q v((1+r)a_2 + q) \leq \mathsf{E}_q v((1+r)(a_1 - a_m) + q) - \mathsf{E}_q v((1+r)a_1 + q). \tag{EC.4.16}$$

Inequality (EC.4.16) holds as $\mathsf{E}_q[v((1+r)x + q)]$ is convex and decreasing in x.

To simplify the notation, we define the objective function in (27) under spending amount m as follows:

$$h(m, x) := U(m) + \beta \mathsf{E}_q[v((1+r)(x-m) + q)]. \tag{EC.4.17}$$

So, optimal spending amount is $m^* \in \arg\min_{0 \le m \le x} h(m,x)$. But, there could be more than one optimal spending amount m^* for minimizing $h(m,x)$. If multiple actions achieve the minimum of $h(m,x)$, we let the optimal spending amount $m^*(x)$ be:

$$m^*(x) = \inf\{m | h(m,x) = \min_{0 \le a \le x} h(a,x)\}. \tag{EC.4.18}$$

This ensures that among all optimal spending amounts, $m^*(x)$ is the smallest one, thereby preserving the monotonicity and continuity properties necessary for the subsequent arguments.

EC.4.3.1. Increasing property of $m^*(x)$: Let x and x' be two states satisfying $x' < x$. Denote $m' = m^*(x')$ as optimal spending amount for state x'. For any action m with $m < m' < x' < x$, we prove that m cannot be the unique optimal solution for state x. Setting $a_1 = x' - m$, $a_2 = x - m$ and $a_m = m' - m$ in (EC.4.16) respectively, we obtain

$$\begin{aligned} &\mathsf{E}_q[v((1+r)(x-m')+q)] - \mathsf{E}_q[v((1+r)(x-m)+q)] \\ \le\ &\mathsf{E}_q[v((1+r)(x'-m')+q)] - \mathsf{E}_q[v((1+r)(x'-m)+q)]. \end{aligned}$$

Since m' is the smallest optimal spending amount for state x' and the spending amount $m < m' < x'$ feasible in state x', we have

$$U(m') + \beta \mathsf{E}_q[v((1+r)(x'-m')+q)] < U(m) + \beta \mathsf{E}_q[v((1+r)(x'-m)+q)].$$

Multiplying the first inequality by β and adding it to the second inequality, the terms $\beta\mathsf{E}_q[v((1+r)(x' - m') + q)]$ and $\beta\mathsf{E}_q[v((1+r)(x'-m)+q)]$ cancel out. Thus, we have

$$h(m,x) = U(m) + \beta \mathsf{E}_q[v((1+r)(x-m)+q)] > U(m') + \beta \mathsf{E}_q[v((1+r)(x-m')+q)] = h(m',x).$$

That implies $h(m,x) > h(m',x)$. As m is taken arbitrarily, no action $m < m'$ can lead to better or the same value $h(m,x)$ than $h(m',x)$ for x. Therefore, we have $m^*(x) \ge m' = m^*(x')$ for $x > x'$. This completes the proof that $m^*(x)$ is (weakly) increasing in x.

EC.4.3.2. Continuity of $m^*(x)$, $v(x)$, and equality $v'_-(x) = v'_+(x)$: Here $v'_-(x)$ and $v'_+(x)$ are the left derivative and right derivative of $v(x)$ respectively. As the function $v(x)$ is convex in x, the left and right derivative $v'_-(x)$ and $v'_+(x)$ exist (Rockafellar 1970). We prove them by mathematical induction and iteration of value function. Consider the value iteration function $v_k(x)$ with initial value $v_0(x) = 0$ for all $x \in \mathcal{S}$. The iteration satisfies the following:

$$v_{k+1}(x) = \min_{m \in [0,x]} \left[U(m) + \beta \mathsf{E}_q v_k(\psi(x,m;q)) \right], \quad \forall x \in \mathcal{S}. \tag{EC.4.19}$$

Similar to (EC.4.18), we also define

$$m_k(x) = \inf \arg\min_{0\le m\le x} \{U(m) + \beta \mathsf{E}_q[v_k((1+r)(x-m)+q)]\}, \tag{EC.4.20}$$

which represents optimal spending amount in iteration k. The limits of them are $\lim_{k\to\infty} v_k(x) = v(x)$ and $\lim_{k\to\infty} m_k(x) = m^*(x)$. So, we prove the continuity of $m_k(x)$, $v_k(x)$, and equality $v'_{k,-}(x) = v'_{k,+}(x)$ for every k. Initially, $v_0(x) = 0$ is continuous in x and $v'_{0,+}(x) = v'_{0,-}(x) = 0$. The corresponding optimal spending policy is $m_0(x) = \inf \arg\min_{0\le m\le x} U(m) = \min\{x, \bar{m}\}$, which is continuous in x. Assuming that $v_k(x)$ and $m_k(x)$ are continuous, and $v'_{k,+}(x) = v'_{k,-}(x)$, we then show $v_{k+1}(x)$, $m_{k+1}(x)$ is also continuous in x, and $v'_{k+1,+}(x) = v'_{k+1,-}(x)$ for all x.

We first show the continuity of $v_{k+1}(x)$: By the definition of v_{k+1} and m_k in (EC.4.19) and (EC.4.20), the optimal spending amount is $m_k(x)$ in iteration, so we have:

$$v_{k+1}(x) = U(m_k(x)) + \beta \mathsf{E}_q[v_k((1+r)(x-m_k(x))+q)]. \tag{EC.4.21}$$

By the continuity of $v_k(x)$ and $m_k(x)$ in induction assumption, $v_{k+1}(x)$ is continuous.

We then prove the continuity of $m_{k+1}(x)$. The optimal spending amount $m_{k+1}(x)$ is given by:

$$m_{k+1}(x) = \inf \arg\min_{0\le m\le x} \{U(m) + \beta \mathsf{E}_q[v_{k+1}((1+r)(x-m)+q)]\}.$$

We define the threshold $\tilde{x}_{k+1}$ as the minimal state where the optimal spending amount achieves the full coverage level $\bar{m}$:

$$\tilde{x}_{k+1} = \inf\{x : m_{k+1}(x) = \bar{m}\}.$$

By the increasing property of $m_{k+1}(x)$, we have $m_{k+1}(x) < \bar{m}$ for $x < \tilde{x}_{k+1}$. The objective function is strictly convex in m when $m \in [0, \bar{m})$ because $U(m)$ is strictly convex on this interval and $v_k(x)$ is convex (as established in part (ii) of Proposition 6). Therefore, the objective function $h_k(m,x)$ admits a unique minimizer m for x in $x \in [0, \tilde{x}_{k+1})$. Since $U(m)$ and $\mathsf{E}_q[v_{k+1}((1+r)(x-m)+q)$ are both continuous in m and q, by Maximum Theorem (Sundaram 1996), $m_{k+1}(x)$ is continuous on interval $[0, \tilde{x}_{k+1})$. Then, for $\tilde{x}_{k+1}$, we have $m^*(x) \ge \bar{m}$. However, for any $m > \bar{m}$, we have $U(m) = U(\bar{m}) = 0$. So, further increasing spending amount does not affect utility loss for $m > \bar{m}$. By monotonicity of v_k (established in Proposition 6), we have

$$U(m) + \beta \mathsf{E}_q[v_k((1+r)(x-m)+q)] \ge U(\bar{m}) + \beta \mathsf{E}_q[v_{k+1}((1+r)(x-\bar{m})+q)], \quad \forall m > \bar{m}.$$

So, $m_{k+1}(x) \le \bar{m}$ for $x > \tilde{x}_k$. By increasing property of $m_{k+1}(x)$, we have $m_{k+1}(x) \ge \bar{m}$ for $x > \tilde{x}_k$. So $m_{k+1}(x) = \bar{m}$ for $x \ge \tilde{x}_k$. Thus, $m_{k+1}(x)$ is continuous for all x.

Finally, we prove that the left and right derivatives $v'_{k+1,-}(x)$ and $v'_{k+1,+}(x)$ satisfy $v'_{k+1,-}(x) = v'_{k+1,+}(x)$ for all x. By convexity of $v_{k+1}(x)$ (proof in part (ii) of Proposition 6), the left and right derivative $v'_{k+1,-}(x)$ and $v'_{k+1,+}(x)$ exists. We will show that $v'_{k+1,-}(x) = v'_{k+1,+}(x)$. For x satisfying $0 < m_k(x) < x$, function $m_k(x)$ is determined by first order condition because the objective function is convex and continuous:

$$U'(m_k(x)) - \beta(1+r)\mathsf{E}_q[v'_k((1+r)(x - m_{k+1}(x)) + q)] = 0.$$

Taking derivative of $v_{k+1}(x)$ in (EC.4.21), we examine the derivatives from the right and left:

$$\begin{aligned} v'_{k+1,+}(x) &= U'(m_k(x))m'_{k,+}(x) + \beta(1+r)(1 - m'_{k,+}(x))\mathsf{E}_q[v'_k((1+r)(x - m_k(x)') + q)], \\ v'_{k+1,-}(x) &= U'(m_k(x))m'_{k,-}(x) + \beta(1+r)(1 - m'_{k,-}(x))\mathsf{E}_q[v'_k((1+r)(x - m_k(x)) + q)]. \end{aligned}$$

Combining the first order condition, the left and right derivative of iteration value function $v'_{k+1,-}(x)$ and $v'_{k+1,+}(x)$ are given by:

$$v'_{k+1,+}(x) = v'_{k+1,-}(x) = U'(m_k(x)).$$

If x satisfies $m_k(x) = x$, then $v_{k+1}(x) = U(x) + \mathsf{E}_q[v_k(q)]$. The derivative is $v'_{k+1,-}(x) = v'_{k+1,+}(x) = U'(x)$. If x satisfies $m_k(x) = 0$, then $v_{k+1}(x) = U(0) + \mathsf{E}_q[v_k((1+r)x + q)]$. By induction assumption $v'_{k,-}(x) = v'_{k,+}(x)$ for all x, the derivative is $v'_{k+1,-}(x) = v'_{k+1,+}(x) = (1+r)\mathsf{E}_q[v'_k((1+r)x + q)]$. Thus, we have $v'_{k+1,-}(x) = v'_{k+1,+}(x)$ for all x.

Therefore, by induction, we conclude that $m^*(x)$ and $v(x)$ are continuous in x. In addition, $v'_-(x) = v'_+(x) = v'(x)$ for all x.

EC.4.3.3. Statement (i) and (ii) in Theorem 2: The two statements can be reformulated as follows: if the optimal spending amount $m^*(x) = x$ or $m^*(x) = 0$, then for $x' < x$, the optimal spending amount also satisfies $m^*(x') = x'$ or $m^*(x') = 0$ respectively. We discuss them one by one.

(1) If the optimal spending amount $m^*(x) = x$, we then prove that, for any state $x' < x$, we have $m^*(x') = x'$. For states $x' < x$, the objective function $h(m, x')$ defined in (EC.4.17) under the spending amount $m = x'$ is given by:

$$h(x', x') = U(x') + \beta\mathsf{E}_q[v(q)]. \tag{EC.4.22}$$

Consider any other spending amount $m' < x'$. We will show that such a spending amount m' leads to greater total discounted utility loss than that under the spending amount x' i.e., $h(x', x') \le h(m', x')$. We utilize the condition $m^*(x) = x$, where $m^*(x)$ is defined as the smallest optimal spending amount for state x. Spending amount $m = m' + (x - x')$ is feasible, but not the optimal for state x as $m' + (x - x') < x' + (x - x') = x$. Substituting $m = m' + (x - x')$ into function $h(m, x)$ defined in (EC.4.17), we have

$$h(m' + (x - x'), x) > h(x, x) = U(x) + \beta\mathsf{E}_q[v(q)]. \tag{EC.4.23}$$

The strict inequality "$>$" holds by assumption $m^*(x) = x$ where $m^*(x)$ is the smallest optimal spending amount defined in (EC.4.18).

We use the auxiliary inequality in (EC.4.15) to further bound $h(m' + (x - x'), x)$. Setting $a_1 = m' + (x - x')$, $a_2 = x$, and $a_m = x - x'$ in (EC.4.15), we get inequality $U(x') - U(x) \leq U(m') - U(m' + x - x')$. Plugging this inequality into $h(m' + (x - x'), x)$, we obtain

$$\begin{aligned} h(m' + (x - x'), x) &= U(m' + (x - x')) + \beta \mathsf{E}_q[v((1+r)(x' - m') + q)] \\ &\leq U(m') + U(x) - U(x') + \beta \mathsf{E}_q[v((1+r)(x' - m') + q)]. \end{aligned} \quad \text{(EC.4.24)}$$

Combining (EC.4.22)–(EC.4.24), we obtain:

$$\begin{aligned} h(x', x') = U(x') + \beta \mathsf{E}_q[v(q)] &< U(x') - U(x) + h(m' + (x - x'), x) \\ &\leq U(m') + \beta \mathsf{E}_q[v((1+r)(x' - m') + q)] = h(m', x'). \end{aligned}$$

That is $h(x', x') < h(m', x')$. Therefore, the spending amount $m = x'$ yields a lower objective value than any other spending amount $m' < x'$ for state x', i.e., $h(x', x') < h(m', x')$. So, we have $m^*(x') = x'$.

(2) If the optimal spending amount satisfies $m^*(x) = 0$ for state x, by increasing of $m^*(x)$, we have $m^*(x') = 0$ for any state $x' < x$.

(2-a) We establish the result $m(x) = x$ for $x \leq \tilde{s}_l$ under the case $\beta(1+r) \leq 1$ using mathematical induction. We will first show $v'(x) < U'(0)$ for $x > 0$, with equality only at $x = 0$. Then, we use this inequality to analyze optimal spending amount to get our result.

We will use mathematical induction to show $v'(x) < U'(0)$. Initially, $v_0(x) = 0$ satisfies the condition. We assume $v_k'(x) > U'(0)$ for all $x > 0$. Similar to (EC.4.17), we define $h_k(m, x)$ as the objective function in each period:

$$h_k(m, x) = U(m) + \beta \mathsf{E}_q[v_k((1+r)(x - m) + q)].$$

Thus, the derivative of iteration function $(h_k)_m'(m, x)$ is

$$(h_k)_m'(m, x) = U'(m) - \beta(1+r)\mathsf{E}[v_k'((1+r)(x - m) + q)].$$

Because $\beta(1+r) \leq 1$, for $x > 0$, we have

$$\begin{aligned} (h_k)_m'(0, x) &= U'(0) - \beta(1+r)\mathsf{E}_q[v_k'((1+r)x + q)] \\ &\leq U'(0) - \mathsf{E}_q[v_k'((1+r)x + q)] < U'(0) - U'(0) = 0. \end{aligned}$$

Thus, $m_k = 0$ is not optimal for $x > 0$ because we can increases spending amount m to decreases the objective value $h_k(m,x)$ at $m = 0$. Then, by Lemma EC.4 with $m_k(x) > 0$, we have

$$v'_{k+1}(x) = U'(m_k(x)) < U'(0).$$

This completes the proof of $\forall x > 0, v'(x) < U'(0)$.

Thus, for value function $v(x)$, by Lemma EC.4, we have following inequality

$$\begin{aligned}(h)'_m(0,x) &= U'(0) - \beta(1+r)\mathsf{E}_q[v'((1+r)x+q)] \\ &\leq U'(0) - \mathsf{E}_q[v'_k((1+r)x+q)] = U'(0) - \mathsf{E}_q[U'(0)] \leq 0.\end{aligned}$$

Thus, for every state $x > 0$, we have $m^*(x) > 0$. Therefore, if $\beta(1+r) \leq 1$, we have $m^*(x) = x$ for state $x < \tilde{s}_l$ which is defined in Theorem 2.

EC.4.3.4. $m^*(\bar{m}) < \bar{m}$ with $P(q < \bar{m}) > 0$: If the inflow q satisfies $P(q < \bar{m}) > 0$, we prove that the optimal spending amount $m^*(x)$ at state $x = \bar{m}$ satisfies $m^*(\bar{m}) < \bar{m}$.

We prove this by showing $h'(\bar{m},\bar{m}) > 0$, which implies that we can decrease the spending amount to decrease objective value when $m = \bar{m}$. We consider the possible spending amount $\bar{m}$ for state $x = \bar{m}$: By $U(m) = 0$ for $m > \bar{m}$, we have

$$h'(\bar{m},\bar{m}) = U'(\bar{m}) - \beta(1+r)\mathsf{E}_q[v'(q)] = -\beta(1+r)\mathsf{E}_q[v'(q)].$$

Firstly, if $m^*(\bar{m}) = 0$, then $m^*(\bar{m}) < \bar{m}$ holds. Otherwise, if $m^*(\bar{m}) > 0$, by former discussion in (2) of Section EC.4.3.2 about continuity, there exists $\varepsilon_1 > 0$, satisfying that $m^*(x) > 0$ holds in $x \in (\bar{m} - \varepsilon_1, \bar{m})$. So, by Lemma EC.4, we have $v'(x) = U'(m^*(x))$ on $(\bar{m} - \varepsilon_1, \bar{m})$. By assumption of $P(q < \bar{m}) > 0$, there exists ε_2 such that $P(q < \bar{m} - \varepsilon_2) > 0$.

Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. We now analyze the term $\mathsf{E}_q[v'(q)]$. Note that $v'(x) \leq 0$ and $v'(x)$ is increasing in x by convexity of $v(x)$, so we have

$$\mathsf{E}_q[v'(q)] = \mathsf{E}_q[v'(q)\mathbf{1}_{\{q \geq \bar{m}-\varepsilon\}} + v'(q)\mathbf{1}_{\{q < \bar{m}-\varepsilon\}}] \leq \mathsf{E}_q[v'(q)\mathbf{1}_{\{q < \bar{m}-\varepsilon\}}] \leq P(q < \bar{m} - \varepsilon)v'(\bar{m} - \varepsilon).$$

The first inequality holds due to $v'(x) \leq 0$. For the last term, by definition of ε, we have $P(q < \bar{m} - \varepsilon_2) > 0$ and $\bar{m} > \bar{m} - \varepsilon > m^*(\bar{m} - \varepsilon) > 0$. By Lemma EC.4, we have:

$$P(q < \bar{m} - \varepsilon)v'(\bar{m} - \varepsilon) = P(q < \bar{m} - \varepsilon)U'(m^*(\bar{m} - \varepsilon)) \; < \; P(q < \bar{m} - \varepsilon)U'(\bar{m}) = 0.$$

The inequality bound by $m^*(\bar{m}-\varepsilon)q < \bar{m}$ by definition of $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. So, we conclude $\mathsf{E}_q[v'(q)] < 0$. We then consider the derivative $h'(\bar{m}, \bar{m})$:

$$(h)'_m(\bar{m}, \bar{m}) = -\beta(1+r)\mathsf{E}_q[v'(q)] > 0.$$

This implies that we can decrease spending amount to decrease the utility. Therefore $m^*(\bar{m}) < \bar{m}$.

By the above proofs, if the optimal spending $m^*(x) = x$ or $m^*(x) = 0$, then for $x' < x$, we have $m^*(x') = x'$ or $m^*(x') = 0$. Then, we can define the threshold:

$$\tilde{s}_l = \sup\{s | m^*(s) = 0 \text{ or } m^*(s) = s\}.$$

Since the only feasible solution in state $s = 0$ is $m = 0$, we have $m^*(0) = 0$. Therefore, $\tilde{s}_l \geq 0$ exists. In addition, with $P(q < \bar{m}) > 0$, we have $m^*(\bar{m}) < \bar{m}$.

EC.4.4. Proof of Lemma 3

In this section, we give the proof of Lemma 3, which give the optimal spending amount $m^*(s)$ under condition $\beta(1+r) = 1$ and constant inflow q. Since inflow q is constant, we could directly transform the dynamic decision problem in (24) into deterministic problem.

We establish the spending amount constraint $m_t \leq s_t$ by following: Let the spending amount $\{m_t\}_{t=0}^{\infty}$ in each period be given. With the constant inflow q, s_t is deterministic by (22). By induction of s_t in definition (22), we could get the explicit form of s_t as:

$$s_t = (s_0 - m_0)(1+r)^t + q + \sum_{l=1}^{t-1}(q - m_l)(1+r)^{t-l}, \quad \forall t > 0,$$

where s_0 is the initial fund level. In addition, constraint $m_t \leq s_t$ is equivalent to $m_t/(1+r)^t \leq s_t/(1+r)^t$. So, by above equation of s_t, the spending amount constraint $m_t \leq s_t$ is equivalent to the following:

$$\frac{m_t}{(1+r)^t} \leq (s_0 - m_0) + \frac{q}{(1+r)^t} + \sum_{l=1}^{t-1}\frac{q - m_l}{(1+r)^l}$$
$$\iff \sum_{l=0}^{t}\frac{m_l}{(1+r)^l} \leq s_0 + \sum_{l=1}^{t}\frac{q}{(1+r)^l}. \tag{EC.4.25}$$

Based on constraint in (EC.4.25), we could formulate our optimization problem as,

$$v(s) = \min_{\{m_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty}\beta^t U(m_t) \tag{EC.4.26}$$
$$\text{s.t.} \quad \sum_{t=0}^{T}\frac{m_t}{(1+r)^t} \leq s_0 + \sum_{t=1}^{T}\frac{q}{(1+r)^t}, \quad \forall T = 0, 1, 2..., \tag{EC.4.27}$$

$$m_t \geq 0, \quad \forall t = 0, 1, 2, \dots,$$

where $U(m)$ is single period utility loss, which is convex decreasing in m.

To get the solution of problem (EC.4.26), we first consider a relaxed version of problem (EC.4.26). Then, we prove the solution in relaxed problem is still feasible in original problem under some condition. In the relaxed version of problem, the constraint (EC.4.26) only works for the infinite period as follows:

$$\hat{v}(s) = \min_{\{m_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(m_t) \tag{EC.4.28}$$

$$\text{s.t.} \quad \sum_{t=0}^{\infty} \frac{m_t}{(1+r)^t} \leq s_0 + \sum_{t=1}^{\infty} \frac{q}{(1+r)^t}, \tag{EC.4.29}$$
$$0 \leq m_t, \quad \forall t = 0, 1, 2\dots.$$

The constraint in (EC.4.29) could be either "$\leq$" or "$=$" because the $U(m)$ is decreasing in m, so, spending all the funds is always optimal. If we transform the constraint (EC.4.29) into "$=$", under $\beta(1+r) = 1$, the relaxed problem in (EC.4.28) is a standard consumption allocation in macroeconomic theory. The Euler equation (first-order condition) for above problem is given by (Sachs and Larrain B. 1993) as:

$$U'(m_t) = \beta(1+r)U'(m_{t+1}) = U'(m_{t+1}).$$

The second equality derives from condition $\beta(1+r) = 1$. Since and $U'(m)$ is decreasing in m, one of the feasible solution is $m_t = m_{t+1}$. Combining this Euler equation $m_t = m_{t+1}$ with constraint (EC.4.29), the optimal solution in relaxed problem (EC.4.28) is given by

$$\hat{m}_t = \frac{r}{1+r} s_0 + \frac{1}{1+r} q. \tag{EC.4.30}$$

However, in the original problem, the constraint is (EC.4.27) for each period but not (EC.4.29) in relax problem. So, we now verify that if $s_0 \geq q$, then the solution in (EC.4.30) satisfies $\hat{m}_t < s_t$, which is equivalent to constraint (EC.4.27): With spending amount in (EC.4.30), we have $s_1 = (1+r)(s_0 - \hat{m}_t) + q = s_0$ and similarly $s_{t+1} = s_t = \dots = s_0$ for each t.

(1) If $s_0 \geq q$, then $\hat{m}_t \leq s_0 = s_t$ by definition (EC.4.30). So, for $s_0 \geq q$, the optimal spending amount in original problem (EC.4.26) is $\hat{m}_t$ defined in (EC.4.30).

(2) If $0 < s_0 < q$, then solution $\hat{m}_0$ is not feasible as $\hat{m}_0 > s_0$ in the zero period $t = 0$. We will first show $m_0 = s_0$ is optimal for $t = 0$. Then, we show that after $t = 0$, the state becomes $s = q$. Finally, we use the result from case (1) $s_0 \geq q$ for subsequent periods. We consider the spending amount $m = s_0$: The derivative of objective function $h'_m(s_0, s_0)$, defined in (EC.4.17), is given by:

$$h'_m(s_0, s_0) = U'(s_0) - \beta(1+r)v'(q).$$

By Theorem 2 with $\beta(1+r) \le 1$, we know that $m_0 > 0$ for $x > 0$. So, with $q > 0$ and Lemma EC.4,

$$U'(s_0) - \beta(1+r)v'(q) = U'(s_0) - U'(m^*(q)).$$

We now consider the spending amount $m^*(q)$: According to the above discussion (i) in the case of $s_0 \ge q$, if the initial state is $s_0 = q \ge q$, then optimal spending amount is $\hat{m}_t = q$ by (EC.4.30). So, we have $m^*(q) = q$, thus

$$U'(s_0) - U'(m^*(q)) = U'(s_0) - U'(q) \le 0.$$

By the condition $0 < s_0 < q$, the equality $U'(s_0) = U'(q)$ only happens at $q > s_0 > \bar{m}$. This is trivial case because we could full cover in each period. In the other case, we have $U'(s_0) - U'(q) < 0$, so $h'_m(s_0, s_0) = U'(s_0) - \beta(1+r)v'(q) < 0$, which implies that $m^(s_0) \ge s_0$. When $t > 0$, the state s_t becomes q, the problem reduce to the case $s_1 = q \le q$. Then, in the following period, optimal spending amount is $m_k = q$, according to the discussion $s_0 \ge q$. We conclude the solution is $m_0 = s_0$ and $m_k = q$ for $k \ge 1$ if $0 < s_0 < q$.

Combining the discussion above (1) and (2), the optimal spending amount is

$$m_t = \min\left\{s_t, \frac{r}{1+r}s_t + \frac{1}{1+r}q\right\}.$$

This solution has a clear interpretation: we spend either the all fund s or the spending amount that would keep the state constant across periods.

EC.4.5. Proof of Theorem 3

In this section, we give the proof of Theorem 3, which implies optimal reimbursement $m^*(x)$ is piece-wise concave in x. We first give the properties we use in proof for single period utility function $U_r(m)$ and $U_c(m)$. Then, we first consider the result for ratio policy in the statement (i). Finally, we can use similar method for cap policy.

Step 1. Properties of $U_r(m)$: We first consider the single period utility loss $U_r(m)$ under ratio policy . The utility function is defined as $u(l) = l^2/2$ with first and second order derivatives $u'(l) = l$ and $u''(l) = 1$, respectively. We introduce the following lemma which introduces the properties we need in proof.

Lemma EC.5 *If the utility loss function is $u(l) = l^2/2$, then $U_r(m)$ has following properties:*

(i) $U_r(m)$ is strictly convex and strictly decreasing on $[0, \bar{m})$, and equals zero on $[\bar{m}, \infty)$.

(ii) $U'_r(m)$ is continuous and piece-wise linear in m. The three linear intervals for m are $[0, m_r)$, $[m_r, \bar{m})$, and $[\bar{m}, \infty)$.

(iii) $U''_r(m)$ is piece-wise constant with two discontinuous points m_r and $\bar{m}$. The left and right derivative $(U_r)''_-(m)$ and $(U_r)''_+(m)$ satisfy $(U_r)''_-(m_r) \ge (U_r)''_+(m_r)$ and $(U_r)''_-(\bar{m}) \ge (U_r)''_+(\bar{m})$ at two discontinuous points m_r and $\bar{m}$. Thus, $U''_r(m)$ is decreasing on $[0, \infty)$

Proof of Lemma EC.5: The first property is given by Proposition 3 and extension definition of $U_r(m)$.

The second property could be proven as follows: Recall the utility loss function $U_r(m) = \sum_j \mathsf{E}[u((1-r_j^*)C_j)]$, where r_F^* and r_S^* are the optimal ratios under the ratio policy. Plugging utility loss $u(l) = l^2/2$, cost index $b_j = \mathsf{E}[C_j^2]/\mathsf{E}[C_j]$, and optimal ratios r_F^* and r_S^* in (EC.2.23) and (EC.2.24) into $U_r(m)$, the single-period utility loss $U_r(m)$ is given by:

$$U_r(m) = \begin{cases} \frac{1}{2}h_F\mathsf{E}[C_F^2] + \frac{\mathsf{E}[C_S^2]}{2h_S(\mathsf{E}[C_S])^2}(h_S\mathsf{E}[C_S]-m)^2, & \text{if } 0 < m < m_r, \\ \frac{1}{2}\frac{\mathsf{E}[C_F^2]\mathsf{E}[C_S^2]}{h_F\mathsf{E}[C_S^2](\mathsf{E}[C_F])^2+h_S\mathsf{E}[C_F^2](\mathsf{E}[C_S])^2}\left(h_F\mathsf{E}[c_F]+h_S\mathsf{E}[c_S]-m\right)^2, & \text{if } m_r < m < \bar{m}, \\ 0, & \text{if } \bar{m} < m. \end{cases} \tag{EC.4.31}$$

Obviously, by definition of $U_r(m)$ in (EC.4.31), single period utility loss $U_r(m)$ is continuous at m_r and $\bar{m}$. The derivative is

$$U_r'(m) = \begin{cases} -\frac{\mathsf{E}[C_S^2]}{h_S(\mathsf{E}[C_S])^2}(h_S\mathsf{E}[C_S]-m), & \text{if } 0 \le m < m_r, \\ -\frac{\mathsf{E}[C_F^2]\mathsf{E}[C_S^2]}{h_F\mathsf{E}[C_S^2](\mathsf{E}[C_F])^2+h_S\mathsf{E}[C_F^2](\mathsf{E}[C_S])^2}\left(h_F\mathsf{E}[C_F]+h_S\mathsf{E}[C`_S]-m\right), & \text{if } m_r < m < \bar{m}, \\ 0, & \text{if } \bar{m} < m. \end{cases}$$

For each interval $[0, m_r)$, $(m_r, \bar{m})$, and $(\bar{m}, \infty)$, the derivative $U_r'(m)$ is linear to m. Thus, $U_r'(m)$ is piecewise linear. Then, we check the continuity of $U_r'(m)$: We take $m = m_r$ into the left and right limits of $U_r'(m)$, where m_r is defined in (EC.2.17), as follows:

$$(U_r)'_-(m_r) = -\frac{\mathsf{E}[C_F^2]}{\mathsf{E}[C_F]} = (U_r)'_+(m_r).$$

Similarly, taking the form $m = \bar{m}$ into the left and right limits of $U_r'(m)$, we have $(U_r)'_-(\bar{m}) = 0 = (U_r)'_+(\bar{m})$. Therefore, $U_r'(m)$ is continuous in m.

The third property could be proven as follows: Similarly, the second order derivative, $U_r''(m)$, is given by:

$$U_r''(m) = \begin{cases} \frac{\mathsf{E}[C_S^2]}{h_S(\mathsf{E}[C_S])^2}, & \text{if } 0 \le m < m_r, \\ \frac{\mathsf{E}[C_F^2]\mathsf{E}[C_S^2]}{h_F\mathsf{E}[C_S^2](\mathsf{E}[C_F])^2+h_S\mathsf{E}[C_F^2](\mathsf{E}[C_S])^2}, & \text{if } m_r < m < \bar{m}, \\ 0, & \text{if } \bar{m} < m. \end{cases}$$

For each interval $[0, m_r)$, $(m_r, \bar{m})$, and $(\bar{m}, \infty)$, the derivative $U_r''(m)$ is constant. Thus, $U_r''(m)$ is piecewise constant. We now compare the constant on each interval: Taking form $m = m_r$ into the left and right limits of $U_r''(m)$, we have

$$(U_r)''_-(m_r) - (U_r)''_+(m_r) = \frac{h_F(\mathsf{E}[C_F])^2(h_F h_S\mathsf{E}[C_S])^2}{(\mathsf{E}[C_F])^2(\mathsf{E}[C_S])^2\mathsf{E}[C_S^2] + h_S^2(\mathsf{E}[C_S])^4\mathsf{E}[C_F^2]} > 0.$$

Similarly, taking form $m = \bar{m}$ into the left and right limits of $U_r''(m)$, we have $(U_r)''_-(\bar{m}) > 0 = (U_r)''_+(\bar{m})$. Thus, $U_r''(m)$ is piece-wise constant with two discontinuous point m_r and $\bar{m}$ satisfying $(U_r)''_-(m_r) \geq (U_r)''_+(m_r)$ and $(U_r)''_-(\bar{m}) \geq (U_r)''_+(\bar{m})$. In general, $U_r''(m)$ is decreasing in m. Q.E.D.

Thus, we have proven the three properties in Lemma EC.5. For cap policy with discrete loss, we have similar properties.

Step 2. Similar structure for cap policy: We prove that cap policy has similar structure.

Lemma EC.6 *If the utility loss function is $u(l) = l^2/2$ and service costs C_S and C_F take random discrete values $c_{(1)} < c_{(2)} < \cdots < c_{(n)}$, then the single period utility loss under cap policy $U_c(m)$ has the following properties:*

(i) $U_c(m)$ is strictly convex and strictly decreasing on $[0, \bar{m})$, and equals zero on $[\bar{m}, \infty)$.

(ii) $U_c'(m)$ is continuous and piece-wise linear in m. The linear intervals for m are $[0, m_{(1)})$, $[m_{(1)}, m_{(2)})$,..., $[m_{(n-1)}, \bar{m})$, and $[\bar{m}, \infty)$, where $m_{(i)}$ satisfies $\tau^*(m_{(i)}) = c_{(n-i)}$.

(iii) $U_c''(m)$ is piece-wise constant with n discontinuous points $m_{(1)}$, $m_{(2)}$,..., $m_{(n-1)}$, and $\bar{m}$. The left and right derivative $(U_c)''_-(m)$ and $(U_c)''_+(m)$ satisfy $(U_c)''_-(m_{(i)}) \geq (U_c)''_+(m_{(i)})$ and $(U_c)''_-(\bar{m}) \geq (U_c)''_+(\bar{m})$ at the discontinuous points. Thus, $U_c''(m)$ is decreasing on $[0, \infty)$.

Proof of Lemma EC.6: Recall the definition of $U_c(m)$ in (4). By Proposition 1, the reimbursement policy is $\phi_j^*(x) = \max\{x - \tau^*, 0\}$. We denote $\tau^*(m)$ as the maximal out-of-pocket threshold when budget is m. The threshold $\tau^*(m)$ is determined by budget constraint in (9). Let j_i denote the corresponding service of the cost $c_{(i)}$. The weight of cost $c_{(i)}$ is $\omega_{(i)} = h_{j_i} P(C_j = c_{(i)})$. Then, the budget constraint in (9) can be rewritten as:

$$\sum_{i=1}^{n} \omega_{(i)} \max\{c_{(i)} - \tau^*(m), 0\} = m. \tag{EC.4.32}$$

Similarly, the objective value i.e., the single period utility loss in (4) can be rewritten as:

$$U_c(m) = \frac{1}{2} \sum_{i=1}^{n} \omega_{(i)} (\min\{c_{(i)}, \tau^*(m)\})^2. \tag{EC.4.33}$$

The first property is given by Proposition 3 and extension definition of $U_c(m)$. The second property can be proven as follows: We first derive the form of $m_{(i)}$ using the inverse function $(\tau^*)^{-1}(c)$. As $\sum_{i=1}^{n} \omega_{(i)} \max\{c_{(i)} - \tau^*(m), 0\}$ in (EC.4.32) is strictly and continuously increasing in τ^*, we have the inverse function $(\tau^*)^{-1}(c)$ exists and satisfies $m_{(i)} = (\tau^*)^{-1}(c_{(n-i)})$. Consider each region $[m_{(i)}, m_{(i+1)})$ with $m_{(0)} = 0$ and $m_{(n)} = \bar{m}$. For the budget $m \in [m_{(i)}, m_{(i+1)})$, the threshold $\tau^*(m) \in [c_{n-i-1}, c_{(n-i)})$ is determined by (EC.4.32) as:

$$\sum_{l=1}^{n} \omega_{(l)} \max\{c_{(l)} - \tau^*(m), 0\} = \sum_{l=n-i}^{n} \omega_{(l)} (c_{(l)} - \tau^*(m)) = m,$$

where the solution is

$$\tau^*(m) = \frac{1}{\sum_{l=n-i}^{n} \omega_{(l)}} \left[\sum_{l=1}^{n} \omega_{(l)} c_{(l)} - m \right]. \tag{EC.4.34}$$

So, plugging this threshold in (EC.4.34) into (EC.4.33), the single period utility loss function under cap policy $U_c(m)$ is given by:

$$U_c(m) = \frac{1}{2} \left[\sum_{l=1}^{n-i-1} \omega_{(l)} c_{(l)}^2 + \sum_{l=n-i}^{n} \omega_{(l)} (\tau^*(m))^2 \right], \quad \text{if } m \in [m_{(i)}, m_{(i+1)}).$$

The derivative is as follows:

$$U_c'(m) = \frac{dU_c}{d\tau^*} \frac{d\tau^*}{dm} = -\tau^*(m) = -\frac{1}{\sum_{l=n-i}^{n} \omega_{(l)}} \left[\sum_{l=1}^{n} \omega_{(l)} c_{(l)} - m \right], \quad \text{if } m \in [m_{(i)}, m_{(i+1)}). \tag{EC.4.35}$$

At the corner point $m_{(i)}$, we have $\lim_{m \to m_{(i)}^-} U_c'(m) = \lim_{m \to m_{(i)}^+} U_c'(m) = -\tau^*(m_{(i)})$. So, $U_c'(m)$ is continuous and linear on $[0, m_{(1)})$, $[m_{(1)}, m_{(2)})$, ..., $[m_{(n-1)}, \bar{m})$, and $[\bar{m}, \infty)$.

The proof of third property is similar. By (EC.4.35), we consider the derivative $U_c''(m)$ as follows:

$$U_c''(m) = \frac{1}{\sum_{l=n-i}^{n} \omega_{(l)}}, \quad \text{if } m \in (m_{(i)}, m_{(i+1)}). \tag{EC.4.36}$$

By (EC.4.36), the derivative $U_c''(m)$ is piece-wise constant. In addition, by (EC.4.36), we have $\lim_{m \to m_{(i)}^-} U_c''(m) \geq \lim_{m \to m_{(i)}^+} U_c''(m)$.

Lemma EC.5 and EC.6 show that $U_c(m)$ and $U_r(m)$ have similar structures. In the following, we prove our main result using single-period utility loss $U_r(m)$ under ratio policy as example. The proof follows similarly for the case when the single-period utility loss is $U_c(m)$.

Step 3. mathematical induction: We employ mathematical induction for the proof. Consider the value iteration function $v_k(x)$ with initial value $v_0(x) = 0$ for all x. The iteration satisfies the following equality:

$$v_{k+1}(x) = \min_{m \in [0,x]} \left[U(m) + \beta \mathsf{E}_q v_k(\psi(x, m; q)) \right], \quad \forall x \in \mathcal{S}, \tag{EC.4.37}$$

which represents the iteration value function. We define $h_k(m, x)$ as follows:

$$h_k(m, x) := U(m) + \beta \mathsf{E}_q [v_k((1+r)(x-m)+q)], \tag{EC.4.38}$$

which represents the objective function to spending amount m for state x in iteration k. We also define $m_k(x)$ as follows:

$$m_k(x) = \inf \arg\min_{0 \leq m \leq x} \left\{ U(m) + \beta \mathsf{E}_q [v_k((1+r)(x-m)+q)] \right\}, \tag{EC.4.39}$$

which represents the optimal spending amount function for state x in iteration k. Their limits are $\lim_{k\to\infty} v_k(x) = v(x)$ and $\lim_{k\to\infty} m_k(x) = m^*(x)$. Our induction assumption for iteration k has the following four parts:

(i) $v_k'(x)$ is concave on $[0,\infty)$.

(ii) $m_k(x)$ is piece-wise concave and $x - m_k(x)$ is non-decreasing.

(iii) $(U_r)''_+(m_k(x))m'_{k,+}(x)$ and $(U_r)''_-(m_k(x))m'_{k,-}(x)$ is decreasing if x satisfies $0 < m_k(x) < x$.

(iv) $(U_r)''_+(m_k(x))m'_{k,+}(x) \le (U_r)''_-(m_k(x))m'_{k,-}(x)$ for state x satisfying $m_k(x) = m_r$ or $m_k(x) = \bar{m}$.

Initially, for $k = 0$, $v_0(x) = 0$ and $v_0'(x) = 0$. The corresponding optimal spending policy is $m_0(x) = \inf \arg\min_{0\le m\le x} U_r(m) = \min\{x, \bar{m}\}$. Conditions (i)–(iv) hold for $k = 0$.

For $k+1$, we proceed as follows. We consider the objective function $h_{k+1}(m,x)$ in (EC.4.38). The derivative of the objective function $h_{k+1}(m,x)$ to spending amount m is

$$(h_{k+1})'_m(m,x) := U_r'(m) - \beta(1+r)\mathsf{E}_q[v'_{k+1}((1+r)(x-m)+q)]. \tag{EC.4.40}$$

The first order condition is $h'_{k+1}(m_{k+1}(x),x) = 0$, which holds if $0 < m_{k+1}(x) < x$.

We first define $\tilde{s}_{l,k} = \sup\{x : m_k(x) = 0 \text{ or } m_k(x) = x\}$ and $\tilde{x}_{l,k} = \inf\{x : m_k(x) = \bar{m}\}$ for all k. We prove that

$$m_k(x) \in \begin{cases} \{0\} \text{ or } \{x\}, & \text{if } x \le \tilde{s}_{l,k}, \\ (0, \min\{x,\bar{m}\}), & \text{if } \tilde{s}_{l,k} < x < \tilde{x}_{l,k}, \\ \{\bar{m}\} & \text{if } \tilde{x}_{l,k} \le x. \end{cases} \tag{EC.4.41}$$

So, in the following proof, we will divide the region $[0,\infty)$ into two or three parts to discuss it one by one.

EC.4.5.1. $m_k(x)$ satisfies equation (EC.4.41) for all k: We first prove the statement: If $x \le \tilde{s}_{l,k}$, then $m_k(x) = x$ or $m_k(x) = 0$; If $x > \tilde{s}_{l,k}$, then $0 < m_k(x) < x$. This statement is equivalent to the following statement: if the optimal spending amount $m_k(x) = x$ or $m_k(x) = 0$, then for $x' < x$, optimal spending amount $m_k(x') = x$ or $m_k(x') = 0$, respectively. The statement is similar to Section EC.4.3.3, so we could follows the proof. We just replacing the $v(x)$, $m^*(x)$, and $h(m,x)$ to $v_k(x)$, $m_k(x)$, and $h_k(m,x)$ in Section EC.4.3.3, respectively . The proof in Section EC.4.3.3 relies on the convexity and the decreasing property of $v(x)$. $v_k(x)$ is also convex and decreasing by the proof of part (i) and (ii) in Section EC.4.2. So, following the proof in Section EC.4.3.3, we have that if the $x \le \tilde{s}_{l,k}$, then $m_k(x) = x$ or $m_k(x) = 0$; If $x > \tilde{s}_{l,k}$, then $0 < m_k(x) < x$.

Then we prove the statement: If the $x \ge \tilde{x}_{l,k}$, then $m_k(x) = \bar{m}$. Firstly, by the increasing property of $m_k(x)$ in x, we have $m_k(x) \ge m_k(\tilde{x}_{l,k}) = \bar{m}$. We then prove $m_k(x) \le \bar{m}$: By the part (i) of Lemma EC.5, for $m > \bar{m}$, we have $U(m) = U(\bar{m}) = 0$. But, $v_{k+1}(x)$ is decreasing by the result of part (i) in Section EC.4.2. So, spending amount $m > \bar{m}$ could not be smallest optimal spending amount in (EC.4.39). Thus, we have $m_k(x) \le \bar{m}$. By above two inequality $m_k(x) \ge \bar{m}$ and $m_k(x) \le \bar{m}$, so $m_k(x) = \bar{m}$.

These two statements complete the proof of equation (EC.4.41). Since these proofs do not rely on the induction assumption, equation (EC.4.41) holds for all $k \geq 0$.

EC.4.5.2. Concavity of $v'_{k+1}(x)$ on $[0,\infty)$: A function is concave if and only if both the left-hand and right-hand derivatives are monotonically decreasing, and the right-hand derivative is less than or equal to the left-hand derivative at each point. Thus, we prove that $v''_{k+1}(x)$ is decreasing in x. We first prove the decreasing property of $v''_{k+1}(x)$ on $[0, \tilde{s}_{l,k})$ and $(\tilde{s}_{l,k}, \infty)$. Then, we prove $(v_{k+1})''_-(\tilde{s}_{l,k}) \geq (v_{k+1})''_+(\tilde{s}_{l,k})$. With those two conditions, $v''_{k+1}(x)$ is decreasing in $x \in [0,\infty)$.

(1) If $x < \tilde{s}_{l,k}$, there are two cases: $m(x) = x$ or $m(x) = 0$ on this interval. We consider each one by one:

(1-a) Case of $m_k(x) = 0$ for all $x < \tilde{s}_{l,k}$: we have $v_{k+1}(x) = \beta \mathsf{E}_q[v_k((1+r)x+q)]$. Then, $v'_{k+1}(x) = \beta(1+r)\mathsf{E}_q[v'_k((1+r)x+q)]$ is concave in x by induction assumption of global concavity $v'_k(x)$. So, $v''_{k+1}(x)$ is decreasing.

(1-b) Case of $m_k(x) = x$ for all $x < \tilde{s}_{l,k}$: We have $v_{k+1}(x) = U_r(x) + \beta \mathsf{E}_q[v_k(q)]$. Then, $v''_{k+1}(x) = (U_r)''(x)$ is decreasing in x as $U''(x)$ is piece-wise constant and $U''_-(x) \geq U''_+(x)$ by part (iii) in Lemma EC.5. So, $v''_{k+1}(x)$ is decreasing.

In both of two cases, we have $v''_{k+1}(x)$ is decreasing in interval $[0, \tilde{s}_{l,k})$.

(2) If $x > \tilde{s}_{l,k}$, then optimal spending amount $0 < m_k(x) < x$. By Lemma EC.4, the derivative of iterated value function $v'_{k+1}(x)$ could be simplified into $v'_{k+1}(x) = U'(m_k(x))$. Since the induction assumption implies the piece-wise concavity of $m_k(x)$, both the left and right derivatives of $m_k(x)$ exist. Thus, for $x > \tilde{s}_{l,k}$, the left derivative $v''_{k+1,-}(x)$ and right derivative $v''_{k+1,+}(x)$ are

$$v''_{k+1,-}(x) = (U_r)''_-(m_k(x))m'_{k,-}(x), \quad v''_{k+1,+}(x) = (U_r)''_+(m_k(x))m'_{k,+}(x).$$

By induction assumption of decreasing property of $(U_r)''_+(m_k(x))m'_{k,+}(x)$ and $(U_r)''_-(m_k(x))m'_{k,-}(x)$, both of $v''_{k+1,-}(x)$ and $v''_{k+1,+}(x)$ are decreasing in x. The induction assumption implies that, $(U_r)''_+(m_k(x))m'_{k,+}(x) \leq (U_r)''_-(m_k(x))m'_{k,-}(x)$ for $m_k(x) = m_r$ at x . So, we have $v''_{k+1,-}(x)$ and $v''_{k+1,+}(x)$ are both decreasing in x, and $v''_{k+1,+}(x) \leq v''_{k+1,-}(x)$ at the discontinuous point. This implies $v''_{k+1}(x)$ (if exists) is decreasing on $(\tilde{s}_{l,k}, \infty)$, and $v''_{k+1,+}(x) \leq v''_{k+1,-}(x)$ at the discontinuous point.

(3) If $x = \tilde{s}_{l,k}$, there are two cases $m_k(x) = x$ or $m_k(x) = 0$. We prove that at $x = \tilde{s}_{l,k}$, the left and right derivative satisfies $(v_{k+1})''_+(\tilde{s}_{l,k}) \leq (v_{k+1})''_-(\tilde{s}_{l,k})$. That is equivalent to $\lim_{x\to\tilde{s}^+_{l,k}} v''_{k+1}(x) \leq \lim_{x\to\tilde{s}^-_{l,k}} v''_{k+1}(x)$.

(3-a) Case of $m_k(\tilde{s}_{l,k}) = \tilde{s}_{l,k}$: By (1) and (2), we have

$$\lim_{x\to\tilde{s}^-_{l,k}} v''_{k+1}(x) = U''(\tilde{s}_{l,k}) \times 1 \geq U''(\tilde{s}_{l,k})m'_{k,+}(x) = \lim_{x\to\tilde{s}^+_{l,k}} v''_{k+1}(x).$$

The inequality holds by induction assumption of decreasing property of $x - m_k(x)$, which implies $m'_{k,+}(x) \leq 1$ and $m'_{k,-}(x) \leq 1$. So, we have $\lim_{x\to\tilde{s}^+_{l,k}} v''_{k+1}(x) \leq \lim_{x\to\tilde{s}^-_{l,k}} v''_{k+1}(x)$ in this case.

(3-b) Case of $m_k(\tilde{s}_{l,k})=0$: By applying first order condition at $\tilde{s}_{l,k}$,

$$(h_k)'_m(m_k(\tilde{s}_{l,k}),\tilde{s}_{l,k}) = U'(m_k(\tilde{s}_{l,k})) - \beta(1+r)\mathsf{E}_q[v'_k((1+r)\tilde{s}_{l,k}+q)].$$

Then, the right derivative of above equation satisfies

$$U''(m_k(\tilde{s}_{l,k}))(m_k)'_+(0) = \beta(1+r)^2(1-(m_k)'_+(0))\mathsf{E}_q[v''_k((1+r)\tilde{s}_{l,k}+q)].$$

By the continuous assumption q, the function $\mathsf{E}_q[(v_k)''_-((1+r)\tilde{s}_{l,k}+q)] = \mathsf{E}_q[(v_k)''_+((1+r)\tilde{s}_{l,k}+q)]$. Plugging this equation into limits of $v''_{k+1}(x)$, the limit is:

$$\begin{aligned}\lim_{x\to\tilde{s}^-_{l,k}} v''_{k+1}(x) &= \beta(1+r)^2\mathsf{E}_q[v''_k((1+r)x+q)]\\ &= (U_r)''(0)\frac{m'_{k,+}(\tilde{s}_{l,k})}{1-m'_{k,+}(\tilde{s}_{l,k})} \geq (U_r)''(0)m'_{k,+}(\tilde{s}_{l,k}) = \lim_{x\to\tilde{s}^+_{l,k}} v''_{k+1}(x).\end{aligned}$$

Hence, by (3-a) and (3-b), we conclude $(v_{k+1})''_-(\tilde{s}_{l,k}) \geq (v_{k+1})''_+(\tilde{s}_{l,k})$.

The above results (1)–(3) show that $v''_{k+1}(x)$ is decreasing on both $[0,\tilde{s}_{l,k})$ and $(\tilde{s}_{l,k},\infty)$, with $v''_{k+1,+}(x) \leq v''_{k+1,-}(x)$ at all discontinuous points. Therefore, $v'_{k+1}(x)$ is concave on $[0,\infty)$.

EC.4.5.3. Piece-wise concavity of $m_{k+1}(x)$ and non-decreasing property of $x - m_{k+1}(x)$: We first prove the piece-wise concavity of $m_{k+1}(x)$. By Section EC.4.5.1, we could define $\tilde{s}_{l,k+1} := \sup\{x : m_{k+1}(x) = 0 \text{ or } m_{k+1}(x) = x\}$ and $\tilde{x}_{l,k+1} := \inf\{x | m_{k+1}(x) = \bar{m}\}$. So, we can discuss the concavity in each region: $[0,\tilde{s}_{l,k+1}]$, $(\tilde{x}_{l,k+1},\infty)$, and $(\tilde{s}_{l,k+1},\tilde{x}_{l,k+1}]$ one by one:

(1) If $x \leq \tilde{s}_{l,k+1}$, then we have $m_{k+1}(x) = x$ or $m_{k+1}(x) = 0$ on $[0,\tilde{s}_{l,k+1}]$. Therefore, $m_{k+1}(x)$ is concave in this region.

(2) If $x > \tilde{x}_{l,k+1}$, then by Section EC.4.5.1, the optimal spending amount $m_{k+1}(x) = \bar{m}$. So, $m_{k+1}(x) = \bar{m}$ is concave in this region.

(3) If $\tilde{s}_{l,k+1} < x \leq \tilde{x}_{l,k+1}$, then we have $x < \tilde{x}_{l,k+1}$, which implies we have $m_{k+1}(x) \in [0,\bar{m})$ by Section EC.4.5.1. The optimal spending amount $m_{k+1}(x)$ has unique solution by the strict convexity of objective function $h_{k+1}(m,x)$ in (EC.4.38) in this region. We examine the following intervals where $U''_r(m)$ is continuous: In ratio policy, the intervals are $M_l = [0,m_r)$ and $M_r = (m_r,\bar{m})$ for the spending amount. By the continuous and increasing property of $m_{k+1}(x)$, the corresponding intervals for the state are $S_l = (\tilde{s}_{l,k+1}, m^{-1}_{k+1}(m_r))$ and $S_r = (m^{-1}_{k+1}(m_r),\tilde{x}_{l,k+1})$. Within M_l or M_r, $U''_r(m)$ is constant.

We first prove $m_{k+1}(x)$ is concave on S_l. The corresponding co-domain for $m_{k+1}(x)$ is M_l . The proof for M_r and S_r is the same. Let $\theta \in (0,1)$ be given. Let states $x \in S_l$ and $y \in S_l$ be given. Denote $m_x = m_{k+1}(x)$ and $m_y = m_{k+1}(y)$, so $m_x, m_y \in M_l$. Denote the convex combination $z = \theta x + (1-\theta)y$ and

$m_z = \theta m_x + (1-\theta)m_y$. We have $z \in S_l$ and $m_z \in M_l$. By part (ii) in Lemma EC.5, the function $U_r'(m)$ is linear in M_l, thus we have

$$U_r'(m_z) = \theta U_r'(m_x) + (1-\theta)U_r'(m_y). \tag{EC.4.42}$$

As $\tilde{s}_{l,k+1} < x \le \tilde{x}_{l,k+1}$, we have $0 < m_x < x$. By (EC.4.40) and first-order condition, we have:

$$(h_{k+1})_m'(m_x, x) = U_r'(m_x) - \beta(1+r)\mathsf{E}_q[v_{k+1}'((1+r)(x-m_x)+q)] = 0; \tag{EC.4.43}$$

and similar

$$(h_{k+1})_m'(m_y, y) = U_r'(m_y) - \beta(1+r)\mathsf{E}_q[v_{k+1}'((1+r)(y-m_y)+q)] = 0. \tag{EC.4.44}$$

By concavity of v_{k+1}', we have

$$\mathsf{E}_q[v_{k+1}'((1+r)(z-m_z)+q)] \le \mathsf{E}_q[\theta v_{k+1}'((1+r)(x-m_x)+q)] + \mathsf{E}_q[(1-\theta)v_{k+1}'((1+r)(y-m_y)+q)]. \tag{EC.4.45}$$

Plugging (EC.4.42) – (EC.4.45) into $h'(m_z, z)$, we obtain,

$$(h_{k+1})_m'(m_z, z) \le \theta(h_{k+1})_m'(m_x, x) + (1-\theta)(h_{k+1})_m'(m_y, y) = 0.$$

Since the objective function $h_{k+1}(m,x)$ is strictly convex in m in this region as $x \le \tilde{x}_{l,k+1}$ so $m(x) < \bar{m}$, this inequality implies that increasing the spending amount reduces the loss. Therefore, we have $m_{k+1}(\theta x + (1-\theta)y) \ge \theta m_{k+1}(x) + (1-\theta)m_{k+1}(y)$. Hence, $m_{k+1}(x)$ is concave in S_l. This process could be repeated for interval S_r.

From cases (1)–(3), we have that $m_{k+1}(x)$ is piece-wise concave.

We then prove that $x - m_{k+1}(x)$ is non-decreasing in x. Obviously, it holds for $x \le \tilde{s}_{l,k+1}$, $m_{k+1}(x) = 0$ or $m_{k+1}(x) = x$. Then, for any $x' > x > \tilde{s}_{l,k+1}$, we consider the spending amount $m_{k+1}(x) + (x'-x)$, under which the remaining fund to next period under is $x' - m_{k+1}(x) - (x'-x) = x - m_{k+1}(x)$, the same as $m_{k+1}(x)$ in x. Thus, we consider $(h_{k+1})_m'(m_{k+1}(x) + (x'-x), x')$. In addition, since $x > \tilde{s}_{l,k+1}$, we have $0 < m_{k+1}(x) < x$, so $m_{k+1}(x)$ satisfies first order condition as follows,

$$(h_{k+1})_m'(m_{k+1}(x), x) = U_r'(m_{k+1}(x)) - \beta(1+r)\mathsf{E}_q[v_{k+1}'((1+r)(x-m_{k+1}(x))+q)] = 0.$$

Plugging this the first order condition into derivative $(h_{k+1})_m'(m_{k+1}(x) + (x'-x), x')$, we have:

$$\begin{aligned}(h_{k+1})_m'(m_{k+1}(x) + (x'-x), x') &= U_r'(m_{k+1}(x) + (x'-x)) - \beta(1+r)\mathsf{E}_q[v_{k+1}'((1+r)(x-m_{k+1}(x))+q)] \\ &= U'(m_{k+1}(x) + (x'-x)) - U_r'(m_{k+1}(x)).\end{aligned}$$

The last equation $U_r'(m_{k+1}(x)+(x'-x)) - U_r'(m_{k+1}(x)) > 0$ for $x < \tilde{x}_{l,k+1}$ i.e., $m_{k+1}(x) \le \bar{m}$ because $x' > x$ and decreasing property of $U_r'(m)$ by (ii) in Lemma EC.5. So, we conclude $(h_{k+1})_m'(m_{k+1}(x) + (x'-x), x') > 0$. This implies that decreasing the spending amount in $m = m_{k+1}(x)+(x'-x)$ reduces the objective value $h_{k+1}(m,x')$. So, for $x < \tilde{x}_{l,k+1}$, we have $m_{k+1}(x') < m_{k+1}(x)+(x'-x)$. For $x \ge \tilde{x}_{l,k+1}$, we have $m_{k+1}(x) = \bar{m}$. Above all, $x - m_{k+1}(x)$ is non-decreasing.

EC.4.5.4. Decreasing property of $U''(m_{k+1}(x))m_{k+1}\text{`}(x)$ for $x > \tilde{s}_{l,k+1}$ where $0 < m_{k+1}(x) < x$: For state $x > \tilde{s}_{l,k+1}$, we have $0 < m_{k+1}(x) < x$. Then, the spending amount $m_{k+1}(x)$ satisfies first order condition:

$$\begin{aligned} & (h_{k+1})_m'(m_{k+1}(x), x) = 0 \\ \iff \quad & U_r'(m_{k+1}(x)) = \beta(1+r)\mathsf{E}_q[v_{k+1}'((1+r)(x-m_{k+1}(x))+q)]. \end{aligned}$$

Then, we take the right derivative of the first order condition:

$$(U_r)_+''(m_{k+1}(x))(m_{k+1})_+'(x) = \beta(1+r)^2(1-(m_{k+1})_+'(x))\mathsf{E}_q[v_{k+1}''((1+r)(x-m_{k+1}(x))+q)].$$

We have similar equation for the left derivative:

$$(U_r)_-''(m_{k+1}(x))(m_{k+1})_-'(x) = \beta(1+r)^2(1-(m_{k+1})_-'(x))\mathsf{E}_q[v_{k+1}''((1+r)(x-m_{k+1}(x))+q)].$$

For above two equations, we merge the terms about $(m_{k+1})_+'(x)$ and $(m_{k+1})_-'(x)$ and multiply $(U_r)_+''(m_{k+1}(x))$ and $(U_r)_-''(m_{k+1}(x))$ on both sides, respectively. By rearranging the terms and solving for $(m_{k+1})_+'(x)$ and $(m_{k+1})_-'(x)$ respectively, we obtain:

$$(U_r)_+''(m_{k+1}(x))(m_{k+1})_+'(x) = \left(\frac{1}{\beta(1+r)^2\mathsf{E}_q[(v_{k+1})_+''((1+r)(x-m_{k+1}(x))+q)]} + \frac{1}{(U_r)_+''(m_{k+1}(x))}\right)^{-1}; \tag{EC.4.46}$$

and

$$(U_r)_-''(m_{k+1}(x))(m_{k+1})_-'(x) = \left(\frac{1}{\beta(1+r)^2\mathsf{E}_q[(v_{k+1})_-''((1+r)(x-m_{k+1}(x))+q)]} + \frac{1}{(U_r)_-''(m_{k+1}(x))}\right)^{-1}. \tag{EC.4.47}$$

We check the monotonicity of $(U_r)_+''(m_{k+1}(x))(m_{k+1})_+'(x)$ and $(U_r)_-''(m_{k+1}(x))(m_{k+1})_-'(x)$ by the monotonicity of right-hand-side of (EC.4.46) and (EC.4.47).

- By (iii) in Lemma EC.5, $U_r''(m)$ is piece-wise constant and decreasing in x.
- Since $x - m_{k+1}(x)$ is non-decreasing in x and $v''(x)$ is decreasing in x, we have that $\mathsf{E}_q[(v_{k+1})_+''((1+r)(x-m_{k+1}(x))+q)]$ is decreasing in x.

So, by above discussion, both two functions $(U_r)''_+(m_{k+1}(x))(m_{k+1})'_+(x)$ and $(U_r)''_-(m_{k+1}(x))(m_{k+1})'_-(x)$ are decreasing in x. At the discontinuous point x satisfying $m_{k+1}(x) = m_r$ or $m_{k+1}(x) = \bar{m}$, by (iii) in Lemma EC.5, we have $(U_r)''_+(m_{k+1}(x)) \le (U_r)''_-(m_{k+1}(x))$. So, we have $(U_r)''_+(m_{k+1}(x))(m_{k+1})'_+(x) \le (U_r)''_-(m_{k+1}(x))(m_{k+1})'_-(x)$ at the discontinuous point. This completes our analysis of the decreasing property of $U''(m_{k+1}(x))m'_{k+1}(x)$.

With the result in Section EC.4.5.1 – EC.4.5.4, by mathematical induction, we know $m^*(x)$ piece-wise concave in x. Then, we use a similar analysis in Section EC.4.5.4 for $m^*(x)$:

$$(m^*)'_-(x) = \frac{\beta(1+r)^2\mathsf{E}_q[(v)''_-((1+r)(x-m^*(x))+q)]}{\beta(1+r)^2\mathsf{E}_q v''_-((1+r)(x-m^*(x))+q)+(U_r)''_-(m^*(x))}, \tag{EC.4.48}$$

$$(m^*)'_+(x) = \frac{\beta(1+r)^2\mathsf{E}_q[(v)''_+((1+r)(x-m^*(x))+q)]}{\beta(1+r)^2\mathsf{E}_q[v''_+((1+r)(x-m^*(x))+q)+(U_r)''_+(m^*(x))]}. \tag{EC.4.49}$$

The discontinuous point occurs in two reasons:

• Discontinuity due to $v''(x)$. At these discontinuous points, we have $(U_r)''_-(m) \ge (U_r)''_+(m)$ We have that $m'(x)$ is increasing in $\mathsf{E}_q[v''((1+r)(x-m(x))+q)]$. With that the function $\mathsf{E}_q[v''((1+r)(x-m(x))+q)]$ at discontinuous point satisfies $v''_-(x) \ge v''_+(x)$, By (EC.4.48) and (EC.4.49), we have that $m^*_-(x) \ge m^*_+(x)$ holds at the discontinuous point incurred by $v''(x)$, which will not break the concavity of $m(x)$.

• Discontinuity due to $U''_r(m^*(x))$. This discontinuous point x satisfies $(U_r)''_-(m) \ge (U_r)''_+(m)$ at this discontinuous point m_r and $\bar{m}$. By (EC.4.48) and ((EC.4.49)), this breaks the concavity as it leads to $(m^*)'_-(x) \le (m^*)'_+(x)$. But $U''_r(m)$ only jumps at these points and is constant in subsequent intervals.

Therefore, considering the above two cases, the discontinuous point, breaking the concavity, only occurs at the discontinuous point due to $U''_r(m)$. They are the state x satisfying $x = \inf\{x|m^*(x) = m_r\}$ and $x = \inf\{x|m^*(x) = \bar{m}\}$. At the discontinuous point, we have $(U_r)''_+(x) \le (U_r)''_-(x)$, so $m^*_+(x) \ge m^*_-(x)$.

For the cap policy, we can follows the proof to get the result of statement (ii) in Theorem 3.

EC.4.6. Proof of Theorem 4

In this section, we give the proof of Theorem 4, which compares the performance of non-pooling (NP), full pooling (FP) and monetary pooling (MP) systems in different assumptions of services costs in dynamic model. Based on the definition of value function $v(s)$ in (24), the value function is determined by these parameters: discount factor β, inflow q, interest rate r, and single-period utility function $U(m)$. If all these parameters or functions are identical, the value function are the same. Consequently, under identical value functions, the total discounted utility loss are the same if initial fund levels are the same.

(i) MP$\succeq$FP and MP$\succeq$NP: From equations (EC.1.7), (EC.1.8) and (EC.1.9), for each period and state vector $\mathbf{s}$, $\mathcal{A}^{(FP)}(\mathbf{s}) \subseteq \mathcal{A}^{(MP)}(\mathbf{s})$ and $\mathcal{A}^{(NP)}(\mathbf{s}) \subseteq \mathcal{A}^{(MP)}(\mathbf{s})$. Therefore, for any inflow vector, actions feasible in NP and FP systems are also feasible in MP. This implies $v^{(MP)}(\mathbf{s}) \le v^{(NP)}(\mathbf{s})$ and $v^{(MP)}(\mathbf{s}) \le v^{(FP)}(\mathbf{s})$.

(ii) FP=MP under cap policy or ratio policy with homogeneous cost: By Theorem 1, the utility loss $U_c^{(MP)} = U_c^{(FP)}$ and $U_r^{(MP)} = U_r^{(FP)}$ under this case. So, as we discussed, MP and FP systems have totally the same discount β, inflow q, rate r for the pooled group, and single period utility loss function $U(m)$. So we know $v^{(MP)}(\mathbf{s}) = v^{(FP)}(\mathbf{s})$.

As mentioned above, FP achieves the same performance as MP if (i).cap policy is used; (ii). ratio policy is used, but the costs are homogeneous. Similar, FP$\succeq$NP.

EC.4.7. Proof of Lemma 4

In this section, we give the proof of Lemma 4. Recall that K homogeneous groups have the same incidence rate $(h_F^{(i)}, h_S^{(i)})$, distribution of service costs $(C_F^{(i)}, C_S^{(i)})$, distributions of inflow $\{q_t^{(i)}\}_{t=0}^{\infty}$, and initial fund level $s_0^{(i)}$. The services costs satisfy the homogeneous assumption. By Theorem 1, the single period utility loss satisfies $U^{(MP)} = U^{(FP)} \leq U^{(NP)}$. Given identical $(h_F^{(i)}, h_S^{(i)})$ and $(C_F^{(i)}, C_S^{(i)})$ across groups, the service incidence for the pooled Group is $h_j^{(p)} = \sum_{i=1}^{K} w^{(i)} h_j^{(i)} = h_j^{(1)}$. By (EC.1.10), the pooled group's service costs satisfy $C_j^{(p)} = C_j^{(1)}$. From the definitions of $h_j^{(p)}$, $C_j^{(p)}$ and $U^{(FP)}$ in (EC.1.11), we have $U^{(FP)} = U^{(1)} = U^{(2)} = ... = U^{(K)} = U^{(NP)}$, where $U^{(i)}$ denotes the utility loss under the optimal policy for Group i. Therefore, $U^{(FP)} = U^{(NP)} = U^{(MP)}$.

For the total discounted costs $v^{(NP)}$, $v^{(FP)}$, and $v^{(MP)}$, we consider two cases:

(i) $\{q^{(i)}\}_{i=1,2,...,K}$ are perfectly correlated. For each $i, j = 1, 2, ..., K$,we have

$$corr(q^{(i)}, q^{(j)}) = \frac{Cov(q^{(i)}, q^{(j)})}{\sqrt{Var(q^{(i)})Var(q^{(j)})}} = 1.$$

By Cauchy-Schwarz inequality (Pishro-Nik 2014), we have

$$Cov(q^{(i)}, q^{(j)}) \leq \sqrt{Var(q^{(i)})Var(q^{(j)})},$$

with equality if and only if $q^{(i)} = \alpha_{ij} q^{(j)}$ for some constant α_{ij}. Perfectly correlation implies that the equality in Cauchy-Schwarz inequality holds, Since the inflow in groups $q^{(i)}$ and $q^{(j)}$ have the same distribution, this leads to $\alpha_{ij} = 1$ and $q^{(i)} = q^{(j)}$. Thus, inflow for all group satisfies $q^{(1)} = q^{(2)} = ... = q^{(K)}$. The pooled Group inflow is

$$q^{(p)} = \sum_{i=1}^{K} w^{(i)} q^{(i)} = q^{(1)},$$

where $w^{(i)} = N^{(i)} / \sum N^{(i)}$ is the population weight of groups which satisfies $\sum_{i=1}^{K} w^{(i)} = 1$. By (24), the value function is given by:

$$v(s) := \min_{\pi} \mathsf{E}_{\mathbf{q}} \left[\sum_{t=0}^{\infty} \beta^t \, \tilde{u}(s_t, \pi(s_t)) | \, s_0 = s \right].$$

The value function only depends on the parameter discount β, inflow q, rate r, and single period utility function $\tilde{u}$. Since these parameters are the same for each Group including the pooled Group, we have $v^{(NP)} = v^{(1)} = v^{(2)} = ... = v^{(K)} = v^{(FP)} = v^{(MP)}$.

(ii) $\{q^{(i)}\}_{i=1,2,...,K}$ are not perfectly correlated. We prove that $v^{(NP)} > v^{(FP)}$ can occur using the case $K = 2$. Consider $q^{(1)}, q^{(2)} \sim U(0,3)$ and the correlation between inflows of two groups is $corr(q^{(1)}, q^{(2)}) = -1$. Then, for the pooled Group, the inflow is $q^{(p)} = 1.5$, constant in each period. Let parameter be $(c_F, c_S) = (5, 20)$; $(h_F, h_S) = (0.1, 0.05)$; and $s_0^{(1)} = s_0^{(2)} = 1.5$. So, the full cover spending amount in each period is $\bar{m} = 1.5$ for all Group 1, Group 2, and pooled Group. Thus the total discounted utility loss for pooled Group $v^{(FP)} = 0$ in a full pooling system. In a non-pooling system, the total utility loss is positive because there is a 0.5 probability of incurring a loss in period 1, which implies $v^{(NP)} > 0 = v^{(FP)}$. Therefore, FP could be better than NP even when all the parameters of groups are the same. The more detailed example can be seen in Section EC.1.4.

References

Chang JT, Pollard D (1997) Conditioning as disintegration. *Statistica Neerlandica* 51(3):287–317.

Luenberger DG (1997) *Optimization by Vector Space Methods* (John Wiley & Sons).

Pishro-Nik H (2014) *Introduction to Probability, Statistics, and Random Processes* (LLC Blue Bell, PA, USA: Kappa Research).

Rockafellar RT (1970) *Convex Analysis* (Princeton: Princeton University Press).

Sachs J, Larrain B F (1993) *Macroeconomics in the Global Economy* (New York: Harvester Wheatsheaf).

Sundaram RK (1996) *A First Course in Optimization Theory* (Cambridge University Press).