Online Appendix
A. Informationally Imperfect Risk Adjustment

The objective in this section is to model the informationally imperfect MA capitation rates, and show that these capitation rates do not reflect the accounting costs of each patient type, i.e., $\exists i \in I, r_i \neq \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*$. To this end, we first define several concepts in this setting:

- **Information Imperfectness** Following Glazer and McGuire (2000), we model this information imperfectness in a Bayesian framework. That is, when the risk adjustment design (47) is informationally imperfect, the characteristic vector $\boldsymbol{x}_i$ in (47) is only an imperfect signal, instead of a sufficient statistic, of the true patient type $i \in I$. Specifically, a type $i \in I$ patient has probability

$$\rho_{ii} =: \mathbb{P}[\boldsymbol{x}_i | i]$$

  reveals a signal with characteristic vector $\boldsymbol{x}_i$, and probability $\rho_{ji} =: \mathbb{P}[x_j|i]\ \forall j \in Z \backslash \{i\}$ reveal other signals with characteristic vectors x_j, where Z is the set of signals and $Z = I$. Lastly, since this signal is imperfect, we must have $\rho_{ii} \in (0,1)\ \ \forall i \in Z\ \forall i \in I$.

- **Signal-Based Capitation Payments** Since CMS can only observe noisy signals of the true patient types, the prediction results of its risk adjustment formula (47), i.e., $\hat{r}_i = \boldsymbol{x}_i^T \boldsymbol{\beta}\ \ \forall i \in Z$, are functions of these imperfect signals. In fact, these prediction results can be explicitly derived and quantified with Bayesian updating of these signals:

$$\hat{r}_i = \boldsymbol{x}_i^T \boldsymbol{\beta} = \sum_{j\in I} \mathbb{P}[j|\boldsymbol{x}_i] \sum_{s\in S} \frac{\lambda_j^s(w_s^*)}{\lambda_j(\boldsymbol{w}^*)} P_s^* \qquad \forall i \in Z, \tag{23}$$

  where $\mathbb{P}[j|\boldsymbol{x}_i] = \frac{\rho_{ij}\lambda_j(\boldsymbol{w}^*)}{\sum_{k\in I}\rho_{ik}\lambda_k(\boldsymbol{w}^*)}\ \forall j \in I$ are the posterior probabilities that an MA enrollee is a type $j \in I$ patient given signal $\boldsymbol{x}_i$. Therefore, when information is imperfect, the capitation payments from CMS, i.e., $\{\hat{r}_i\}_{i\in Z}$, are signal-based, and are different from the capitation rates based on true patient types, i.e., $\{r_i\}_{i\in I}$.

- **Information Asymmetry** When information is imperfect, MA health plans have more granular information on patient types than CMS does. That is, there is information asymmetry between the payer (CMS) and providers (MA health plans). In particular, as commonly assumed in the risk selection literature (c.f. Glazer and McGuire (2000), Braverman and Chassang (2022)), MA health plans always observe the true patient types and can conduct risk selection based on this information.

Finally, we are ready to derive the informationally imperfect MA capitation rates $\{r_i\}_{i\in I}$, and compare them with the accounting costs of each patient type. First, to each set of signal-based capitation based payments from CMS, i.e., $\{\hat{r}_i\}_{i\in Z}$, there corresponds a set of informationally imperfect MA capitation rates based on true patient types, i.e., $\{r_i\}_{i\in I}$, as

$$r_i = \sum_{j\in Z} \rho_{ji}\hat{r}_j \quad \forall i \in I. \tag{24}$$

Specifically, by combining (23) and (24), we can explicitly characterize the informationally imperfect MA capitation rates as

$$r_i = \sum_{j\in Z} \rho_{ji} \sum_{i\in I} \mathbb{P}[i|x_j] \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \quad \forall i \in I. \tag{25}$$

In other words, since MA health plans can observe the true patient types, they are able to derive the informationally imperfect MA capitation rates $\{r_i\}_{i\in I}$, and compare them with the accounting costs of treating each patient type, i.e., $\{\sum_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^*\}_{i\in I}$. The following lemma summarizes this subsection, and shows that MA health plans have incentives to conduct risk selection when MA capitation rates are informationally imperfect.

Lemma 2. *When a risk adjustment design (47) is informationally imperfect, i.e.,* $\rho_{ii} \in (0,1)$ $\forall i \in Z$ $\forall i \in I$*, MA capitation rates would over- or under- compensate for the accounting costs of treating some patient types, i.e.,* $\exists i \in I$ *s.t.* $r_i > \sum_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^*$ *or* $r_i < \sum_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^*$.

Proof of Lemma 2:

To see that informationally imperfect risk adjustment under compensates for the accounting costs of some patient types, it suffices to note that

$$\underset{i\in I}{Max} \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \geq \underset{j\in Z}{Max}\ \hat{r}_j > \underset{i\in I}{Max}\ r_i, \tag{26}$$

where the first inequality is by (23), i.e.,

$$\sum_{i\in I} \mathbb{P}[i|x_j] \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* = \hat{r}_j \quad \forall j \in Z,$$

while the second inequality is by (24), i.e.,

$$\sum_{j\in Z} \rho_{ji}\hat{r}_j = r_i \quad \forall i \in I,$$

and the fact that the patient type signal is imperfect, i.e., $0 < \rho_{ii} < 1$.

Similarly, we can show that informationally imperfect risk adjustment over compensates for the accounting costs of some patient types as

$$\underset{i\in I}{Min} \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \leq \underset{j\in Z}{Min}\ \hat{r}_j < \underset{i\in I}{Min}\ r_i.$$

□

B. Proofs

Proof of Lemma 1:

First, note that the profit function of service s's provider (2) is concave and continuous in the quality level w_s. Specifically, we have for any service $s \in S$, the revenue function $\sum\limits_{i\in I}\lambda_i^s(w_s)P_s$ is concave in w_s, and the total cost function $C_s^f(w_s)+C_s^v(\sum_{i\in I}\lambda_i^s(w_s))$ is strictly convex in w_s. Therefore the profit function $\pi_s(w_s)=\sum\limits_{i\in I}\lambda_i^s(w_s)P_s - C_s^f(w_s) - C_s^v(\sum_{i\in I}\lambda_i^s(w_s))$ is strictly concave in w_s. As such, the game of maximizing (2) w.r.t. w_s by each provider of service s is a concave game, in which at least one Nash equilibrium exists (Cachon and Netessine 2006).

Second, to establish a one-to-one correspondence between the equilibrium quality level $\{w_s\}_{s\in S}$ and service level reimbursement $\{P_s\}_{s\in S}$, we need to show that the Nash equilibrium of this game is unique. By Cachon and Netessine (2006), a sufficient condition for the uniqueness of Nash equilibrium is that

$$\left|\frac{\partial^2\pi_s(w_s)}{\partial w_s^2}\right| > \sum_{s'\neq s}\left|\frac{\partial^2\pi_s(w_s)}{\partial w_{s'}\partial w_s}\right|, \quad w_s\in[\underline{\mathrm{w}_s},1],\ w_{s'}\in[\underline{\mathrm{w}_{s'}},1]\ \forall s,s'\in S. \qquad (*)$$

We note that (*) holds trivially in our setting because $\frac{\partial^2\pi_s(w_s)}{\partial w_{s'}\partial w_s}=0$ for all $s'\neq s$.

Finally, the one-to-one correspondence between the targeted quality level $\{w_s\}_{s\in S}$ and service reimbursement $\{P_s(w_s)\}_{s\in S}$ is pinned down by (3), which are just the first order conditions of the profit maximization problem (2) of each service provider. □

Proof of Proposition 1:

In the decision problem (9), the action set $\times_{s\in S}[\underline{\mathrm{w}_s},1]$ is compact and convex and the objective function $\pi(\{w_s\}_{s\in S})$ is continuous. Moreover, the objective function, $\pi(\{w_s\}_{s\in S}) = \sum\limits_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{s\in S}C_s^f(w_s) - \sum_{s\in S}C_s^v(\sum_{i\in I}\lambda_i^s(w_s))$, is strictly concave in $(w_1,\ldots,w_s)$, as the revenue function $\sum\limits_{i\in I}\lambda_i(\boldsymbol{w})r_i$ is concave in $(w_1,\ldots,w_{|S|}$, and the total cost function $\sum_{s\in S}C_s^f(w_s)+\sum_{s\in S}C_s^v(\sum_{i\in I}\lambda_i^s(w_s))$ is strictly convex in $(w_1,\ldots,w_{|S|})$. Therefore, an unique solution of (9) exists.

To show that (12) is the necessary and sufficient condition for any capitation rates $\{r_i\}_{i\in I}$ to induce the targeted quality benchmark $\{w_s^*\}_{s\in S}$ in equilibrium, it suffices to observe two facts. First, the equilibrium quality level $\{w_s^{MA}\}_{s\in S}$ is characterized by the first order conditions of the MA health plan's profit maximization problem (9)

$$\sum_{i\in I}\frac{\partial\lambda_i(w^{MA})}{\partial w_s}r_i = \sum_{s'\in S}\frac{\partial\sum_{i\in I}\lambda_i^{s'}(w_s^{MA})}{\partial w_s}\frac{dC_{s'}^v(\sum_{i\in I}\lambda_i^{s'}(w_s^{MA}))}{d(\sum_{i\in I}\lambda_i^{s'}(w_s^{MA}))}+\frac{dC_s^f(w_s^{MA})}{dw_s} \quad \forall s\in S.$$

Therefore, when the targeted quality provision is achieved in equilibrium, we have

$$\sum_{i\in I}\frac{\partial\lambda_i(\boldsymbol{w}^*)}{\partial w_s}r_i = \sum_{s'\in S}\frac{\partial\sum_{i\in I}\lambda_i^{s'}(w_s^*)}{\partial w_s}\frac{dC_{s'}^v(\sum_{i\in I}\lambda_i^{s'}(w_s^*))}{d(\sum_{i\in I}\lambda_i^{s'}(w_s^*))}+\frac{dC_s^f(w_s^*)}{dw_s} \quad \forall s\in S. \qquad (27)$$

Second, we have

$$\sum_{i\in I}\frac{\partial\lambda_i^s(w_s^*)}{\partial w_s}P_s^*=\sum_{s'\in S}\frac{\partial\sum_{i\in I}\lambda_i^{s'}(w_s^*)}{\partial w_s}\frac{dC_{s'}^v(\sum_{i\in I}\lambda_i^{s'}(w_s^*))}{d(\sum_{i\in I}\lambda_i^{s'}(w_s^*))}+\frac{dC_s^f(w_s^*)}{dw_s}\quad\forall s\in S. \tag{28}$$

from (3) in Lemma 1. Combining (27) and (28), we have

$$\sum_{i\in I}\frac{\partial\lambda_i(\boldsymbol{w}^*)}{\partial w_s}r_i=\sum_{i\in I}\frac{\partial\lambda_i^s(w_s^*)}{\partial w_s}P_s^*\quad\forall s\in S,$$

which is exactly (12). □

Proof of Theorem 1:

On one hand, by the nonemptyness of the black swan set S^b characterized by (8), i.e.,

$$S^b:=\{s\in S|P_s^*>\max_{i\in I}\sum\nolimits_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^*\},$$

there exists a service $s\in S^b$ such that

$$P_s^*>\max_{i\in I}\{r_i\}$$

for any set of capitation rates $\{r_i\}_{i\in I}$ which satisfy (7), i.e., $r_i=\sum_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^*\quad\forall i\in I$.

On the other hand, by the no-arbitrage condition (12), i.e.,

$$\sum_{j\in I}\frac{\frac{\partial\lambda_j(\boldsymbol{w}^*)}{\partial w_s}}{\sum_{i\in I}\frac{\partial\lambda_i^s(w_s^*)}{\partial w_s}}r_j=P_s^*\quad\forall s\in S,$$

and the implication of option demand theory (6), i.e.,

$$\frac{\partial\lambda_i(\boldsymbol{w})}{\partial w_s}\le\frac{\partial\lambda_i^s(w_s)}{\partial w_s}\quad\forall i\in I\;\forall s\in S\;\Rightarrow\;\sum_{j\in I}\frac{\frac{\partial\lambda_j(\boldsymbol{w}^*)}{\partial w_s}}{\sum_{i\in I}\frac{\partial\lambda_i^s(w_s^*)}{\partial w_s}}\le 1\quad\forall s\in S,$$

we have

$$\max_{i\in I}\{r_i\}>\sum_{j\in I}\frac{\frac{\partial\lambda_j(\boldsymbol{w}^*)}{\partial w_s}}{\sum_{i\in I}\frac{\partial\lambda_i^s(w_s^*)}{\partial w_s}}r_j=P_s^*.$$

Therefore, the informationally perfect MA capitation rates $\{r_i\}_{i\in I}$ which satisfy (7) do not satisfy the no-arbitrage condition (12), and thus cannot induce the targeted quality benchmark at the market equilibrium by Proposition 1. □

Proof of Theorem 2:

To prove this Theorem 2, we need to show that the targeted quality benchmark (11) is not achieved in market equilibria with 1. informationally perfect risk adjustment and cross subsidization, 2. informationally imperfect risk adjustment and cross subsidization, 3. informationally imperfect risk adjustment and no cross subsidization, but is achieved in the market equilibrium with 4. informationally perfect risk adjustment and no cross subsidization. We have already shown Case 1 in Theorem 1, and will show the rest of the cases below.

Case 2 (Informationally Imperfect Risk Adjustment and Cross Subsidization): Following the same proof strategy as in Theorem 1, it suffices to show that the informationally imperfect MA capitation rates $\{r_i\}_{i\in I}$ satisfy $P_s^* > \underset{i\in I}{max}\ r_i$ for some $s \in S$. In fact, by the nonemptyness of the black swan set S^b (8), i.e.,

$$S^b := \{s \in S | P_s^* > \underset{i\in I}{max} \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*\},$$

and (26), i.e.,

$$\underset{i\in I}{Max} \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \geq \underset{j\in Z}{Max}\ \hat{r}_j > \underset{i\in I}{Max}\ r_i,$$

of Lemma 2, we have

$$\underset{s\in S}{Max} P_s^* > \underset{i\in I}{Max} \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \geq \underset{j\in Z}{Max}\ \hat{r}_j > \underset{i\in I}{Max}\ r_i.$$

Specifically, the gap between $\underset{s\in S}{Max} P_s^*$ and $\underset{i\in I}{Max} \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*$ gives rise to the risk selection incentives induced by cross subsidization, while the gap between $\underset{i\in I}{Max} \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*$ and $\underset{i\in I}{Max}\ r_i$ causes the risk selection incentives due to imperfect information.

Case 3 (Informationally Imperfect Risk Adjustment and No Cross Subsidization): In this case, the MA health plan profit maximization problem becomes

$$\begin{aligned}
\underset{\{w_s\}_{s\in S}}{Max}\ & \sum_{i\in I} \lambda_i(\boldsymbol{w}) r_i - \sum_{s\in S} \left[C_s^f(w_s) - C_s^v(\sum_{i\in I} \lambda_i^s(w_s)) \right] \\
s.t.\ \ & \underline{\mathrm{w}_s} \leq w_s \leq 1 \quad \forall s \in S \\
& r_i = \sum_{s\in S} \frac{\lambda_i^s(w_s)}{\lambda_i(\boldsymbol{w})} P_s^* \quad \forall i \in I \\
& r_i = \sum_{j\in Z} \rho_{ji} \sum_{i\in I} \mathbb{P}[i|x_j] \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \qquad \forall i \in I,
\end{aligned} \tag{29}$$

where the last constraint is from the definition of informationally imperfect risk adjustment (c.f. (25)). To see that the targeted quality benchmark (11) is not achieved in the equilibrium of (29), we note that, by Lemma 2, the informationally imperfect capitation rates (25), i.e.,

$$r_i = \sum_{j\in Z} \rho_{ji} \sum_{i\in I} \mathbb{P}[i|x_j] \sum\nolimits_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^* \qquad \forall i \in I,$$

are different from accounting treatment costs based on targeted quality provision, i.e., $\{\sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*\}_{i\in I}$. Therefore, the targeted quality provision $\{w_s^*\}_{s\in S}$ does not satisfy no-cross-subsidization constraint $r_i = \sum\limits_{s\in S} \frac{\lambda_i^s(w_s)}{\lambda_i(\boldsymbol{w})} P_s^* \ \ \forall i \in I$, and thus cannot be achieved in the equilibrium of (29).

Case 4 (Informationally Perfect Risk Adjustment and No Cross Subsidization): In this case, the MA health plan profit maximization problem becomes

$$\begin{aligned}
\underset{\{w_s\}_{s\in S}}{Max} \ & \sum_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{s\in S}\left[C_s^f(w_s) - C_s^v(\sum_{i\in I}\lambda_i^s(w_s))\right] \\
s.t. \ & \underline{\mathrm{w}_s} \le w_s \le 1 \quad \forall s\in S \\
& r_i = \sum_{s\in S}\frac{\lambda_i^s(w_s)}{\lambda_i(\boldsymbol{w})}P_s^* \quad \forall i\in I \\
& r_i = \sum\nolimits_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^* \qquad \forall i\in I.
\end{aligned} \tag{30}$$

By the no cross subsidization constraint, i.e.,

$$r_i = \sum_{s\in S}\frac{\lambda_i^s(w_s)}{\lambda_i(\boldsymbol{w})}P_s^* \quad \forall i\in I,$$

we can rewrite the profit function of the representative MA health as

$$\begin{aligned}
\pi(\{w_s\}_{s\in S}) &= \sum_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{s\in S}C_s^f(w_s) - \sum_{s\in S}C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s)) \\
&= \sum_{i\in I}\lambda_i(\boldsymbol{w})\left[r_i - \sum\nolimits_{s\in S}\frac{\lambda_i^s(w_s)}{\lambda_i(\boldsymbol{w})}P_s^*\right] + \sum_{s\in S}\left[\sum_{i\in I}\lambda_i^s(w_s)P_s^* - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\right] \\
&= \sum_{s\in S}\left[\sum_{i\in I}\lambda_i^s(w_s)P_s^* - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\right].
\end{aligned}$$

As such, the MA health plan profit maximization problem (30) reduces to

$$\begin{aligned}
\underset{\{w_s\}_{s\in S}}{Max} \ & \sum_{s\in S}\left[\sum_{i\in I}\lambda_i^s(w_s)P_s^* - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\right] \\
s.t. \ & \underline{\mathrm{w}_s} \le w_s \le 1 \quad \forall s\in S,
\end{aligned}$$

whose maximizer is the targeted quality vector $\boldsymbol{w}^*$ as shown in Lemma 1. □

Proof of Theorem 3:

To show that the optimizers of

$$\begin{aligned}
V_1 = \underset{\{w_s\}_{s\in S}}{Max} \ & \pi(\{w_s\}_{s\in S}) \\
s.t. \ & \underline{\mathrm{w}_s} \le w_s \le 1 \quad \forall s\in S \\
& \sum_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s)P_s^* = 0
\end{aligned} \tag{31}$$

are $\{w_s^*\}_{s\in S}$, we first note that (31) can be rewritten as

$$V_2 = \underset{\{w_s\}_{s\in S}}{Max} \sum_{s\in S}\left[\sum_{i\in I}\lambda_i^s(w_s)P_s^* - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\right]$$
$$s.t.\quad \underline{\mathrm{w}}_s \le w_s \le 1 \quad \forall s\in S$$
$$\sum_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s)P_s^* = 0 \tag{32}$$

by the fact that

$$\pi(\{w_s\}_{s\in S}) = \sum_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{s\in S}\sum_{i\in I}\lambda_i^s(w_s)P_s^* + \sum_{s\in S}\left[\sum_{i\in I}\lambda_i^s(w_s)P_s^* - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\right].$$

Moreover, by Lemma 1, we know that the targeted quality levels $\{w_s^*\}_{s\in S}$ are the optimizers of

$$V_3 = \underset{\{w_s\}_{s\in S}}{Max} \sum_{s\in S}\left[\sum_{i\in I}\lambda_i^s(w_s)P_s^* - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\right]$$
$$s.t.\quad \underline{\mathrm{w}}_s \le w_s \le 1 \quad \forall s\in S. \tag{33}$$

Therefore, in order to show $V_1 = V_2 = V_3$, we just need to show that the targeted quality levels $\{w_s^*\}_{s\in S}$ satisfy the selection-proof constraint (16), $\sum\limits_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum\limits_{i\in I}\sum\limits_{s\in S}\lambda_i^s(w_s)P_s^* = 0$.

In fact, by evaluating (16) at $\{w_s^*\}_{s\in S}$, we have

$$\sum_{i\in I}\lambda_i(\boldsymbol{w}^*)r_i - \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s^*)P_s^* = \sum_{i\in I}\lambda_i(\boldsymbol{w}^*)\left[r_i - \sum\nolimits_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^*\right]$$

which equals 0 by (7), $r_i = \sum_{s\in S}\frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)}P_s^* \quad \forall i\in I$. □

C. A Model of Medicare Advantage with Medical Loss Ratio Requirements

In §8, we claimed that the existing Medical Loss Ratio mechanism (17) in MA does not change the market equilibrium, and thus was not included in our main analysis. This section conducts a robustness check by including the MLR constraint (17) into the main model. Specifically, we first show that under the health plan demand function (21), the market equilibrium with the MLR constraint is the same as that without the MLR constraint. Furthermore, we would provide an intuitive explanation on why the existing MLR mechanism does not address the risk selection problem in MA.

First, we note that when adding the MLR constraint (17), the representative MA health plan's profit maximization problem (9) becomes

$$\underset{\{w_s,P_s\}_{s\in S}}{Max} \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i - \sum_{s\in S}\left[C_s^f(w_s^{MA}) + C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s^{MA}))\right]$$
$$s.t.\quad \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i - \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s^{MA})P_s = 0$$
$$w_s^{MA} = \underset{w_s\in[\underline{\mathrm{W}}_s,1]}{ArgMax}\{\sum_{i\in I}\lambda_i^s(w_s)P_s - C_s^f(w_s) - C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\} \quad \forall s\in S. \tag{34}$$

To see this, we first show in Proposition 2 that the original problem (9),

$$\begin{aligned} \underset{\{w_s\}_{s\in S}}{Max} \ & \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i - \sum_{s\in S}\Big[C_s^f(w_s^{MA}) + C_s^v(\textstyle\sum_{i\in I}\lambda_i^s(w_s^{MA}))\Big] \\ s.t. \ & \underline{\mathrm{w}_s} \le w_s \le 1 \ \ \forall s\in S, \end{aligned} \tag{35}$$

can also be formulated as

$$\begin{aligned} \underset{\{w_s,P_s\}_{s\in S}}{Max} \ & \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i - \sum_{s\in S}\Big[C_s^f(w_s^{MA}) + C_s^v(\textstyle\sum_{i\in I}\lambda_i^s(w_s^{MA}))\Big] \\ s.t. \ & w_s^{MA} = \underset{w_s\in[\underline{\mathrm{W}_s},1]}{ArgMax}\{\sum_{i\in I}\lambda_i^s(w_s)P_s - C_s^f(w_s) - C_s^v(\textstyle\sum_{i\in I}\lambda_i^s(w_s))\} \ \ \forall s\in S, \end{aligned} \tag{36}$$

which is exactly (34) without the MLR constraint (17). Here, the equivalence between (35) and (36) comes from the one-to-one correspondence between P_s and w_s established in Lemma 1. In other words, without the MLR constraint, the decision variables $\{P_s\}_{s\in S}$ are redundant because for every service quality w_s in $[\underline{\mathrm{w}_s}, 1]$, there is a service reimbursement $P_s(w_s)$ corresponds to it as established in (3).

Proposition 2. *The one-stage decision problem (35) has the same equilibrium quality as the two-stage decision problem (36), where the MA health plan sets service level reimbursement $\{P_s\}_{s\in S}$ in the first stage and decides quality level $\{w_s\}_{s\in S}$ in the second stage.*

However, in the presence of the MLR constraint (17), this one-to-one correspondence may no longer hold because this constraint may not necessarily satisfy for all service quality choice. That is, it is not necessarily true that

$$\forall\{w_s\}_{s\in S} \in \underset{s\in S}{\times}[\underline{\mathrm{w}_s}, 1], \ \ \sum_{i\in I}\lambda_i(\boldsymbol{w})r_i - \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s)P_s(w_s) = 0. \tag{37}$$

Therefore, the decision variables $\{P_s\}_{s\in S}$ are no longer redundant in the presence of the MLR constraint (17). As such, the representative MA health plan's profit maximization problem (9) needs to be formulated as (34) in the presence of the MLR constraint (17).

Second, we show that the profit maximization problem with MLR (34) yields the same equilibrium quality level $\{w_s^{MA}\}_{s\in S}$ as the profit maximization problem without MLR (35). Following the reasoning in the previous paragraph, it suffices to show that the equilibrium quality level $\{w_s^{MA}\}_{s\in S}$ from (35) satisfies the MLR constraint, i.e.,

$$\sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i - \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s^{MA})P_s(w_s^{MA}) = 0. \tag{38}$$

Specifically, Theorem 4 characterizes conditions under which the MLR constraint is redundant and does no change the MA market equilibrium.

Theorem 4. *When the health plan demand function takes the additive form* $\lambda_i(\boldsymbol{w}) = \sum_{s\in S} \Lambda_i \xi_{is} w_s$ *as in (21), the profit maximization problem with MLR (34) yields the same equilibrium quality level* $\{w_s^{MA}\}_{s\in S}$ *as the profit maximization problem without MLR (35).*

In particular, when the health plan demand function takes the additive form, the health plan demand function is equally sensitive to the quality change of service s as the service demand function, i.e., $\frac{\partial \lambda_i(\boldsymbol{w})}{\partial w_s} = \frac{\partial \lambda_i^s(w_s)}{\partial w_s}$ $\forall i \in I$. In this case, maximizing the MA health plan's profit is same as maximizing the profit for all individual service providers.[12] Therefore, the equilibrium revenue of MA health plan $\sum\limits_{i\in I} \lambda_i(\boldsymbol{w}^{MA}) r_i$ always equals the equilibrium revenue of all individual service providers $\sum\limits_{i\in I}\sum\limits_{s\in S} \lambda_i^s(w_s^{MA}) P_s(w_s^{MA})$, which implies that the MLR constraint is always satisfied as in (38).

To summarize, the existing MLR requirement in MA cannot address the risk selection problem because it fails to prevent service level risk selection. Specifically, MA health plans would satisfy MLR as long as they spend all capitation payments on their enrollees' healthcare services. However, different healthcare services have different profit margins. Therefore, MA health plans would still cross subsidize the capitation payments on healthcare services with higher profit margins, which leads to risk selection.

Proof of Proposition 2:

First, we note that (36) is equivalent to

$$\begin{aligned} \underset{\{P_s, w_s^{MA}\}_{s\in S}}{Max} \; & \sum_{i\in I} \lambda_i(\boldsymbol{w}^{MA}) r_i - \sum_{s\in S} \left[C_s^f(w_s^{MA}) + C_s^v\left(\sum\nolimits_{i\in I} \lambda_i^s(w_s^{MA})\right) \right] \\ s.t. \;\; & w_s^{MA} = \underset{w_s \in [\underline{\mathrm{W}}_s, 1]}{ArgMax} \{ \sum_{i\in I} \lambda_i^s(w_s) P_s - C_s^f(w_s) - C_s^v\left(\sum\nolimits_{i\in I} \lambda_i^s(w_s)\right) \} \;\; \forall s \in S \\ & \underline{\mathrm{w}}_s \le w_s^{MA} \le 1 \;\; \forall s \in S, \end{aligned} \tag{39}$$

while (35) can be rewritten as

$$\begin{aligned} \underset{\{w_s^{MA}\}_{s\in S}}{Max} \; & \sum_{i\in I} \lambda_i(\boldsymbol{w}^{MA}) r_i - \sum_{s\in S} \left[C_s^f(w_s^{MA}) + C_s^v\left(\sum\nolimits_{i\in I} \lambda_i^s(w_s^{MA})\right) \right] \\ s.t. \;\; & \underline{\mathrm{w}}_s \le w_s^{MA} \le 1 \;\; \forall s \in S. \end{aligned} \tag{40}$$

Therefore, to show that (35) and (36) lead to the same equilibrium quality level $\{w_s^{MA}\}_{s\in S}$, it suffices to show that for any maximizers $\{w_s^{MA}\}_{s\in S}$ of (40), there exists $\{P_s\}_{s\in S}$ such that they satisfy the following constraints of (39):

$$w_s^{MA} = \underset{w_s \in [\underline{\mathrm{W}}_s, 1]}{ArgMax} \{ \sum_{i\in I} \lambda_i^s(w_s) P_s - C_s^f(w_s) - C_s^v\left(\sum\nolimits_{i\in I} \lambda_i^s(w_s)\right) \} \;\; \forall s \in S.$$

[12] However, some service $s \in S$ can still be more attractive than other service $s' \in S$ for certain patients, i.e., $\exists s, s' \in S$ $s.t.$ $\frac{\partial \lambda_i^s(w_s)}{\partial w_s} > \frac{\partial \lambda_i^{s'}(w_{s'})}{\partial w_{s'}}$ $\forall i \in I$. Therefore, this additive functional form does not imply that all services are equally attractive. Particularly, MA health plans can still cherry pick type $i \in I$ patients from Medicare by offering a more attractive service $s \in S$ instead of a less attractive service $s' \in S$, i.e., $\frac{\partial \lambda_i(\boldsymbol{w})}{\partial w_s} = \frac{\partial \lambda_i^s(w_s)}{\partial w_s} > \frac{\partial \lambda_i^{s'}(w_{s'})}{\partial w_{s'}} = \frac{\partial \lambda_i(\boldsymbol{w})}{\partial w_{s'}}$.

Finally, the existence of such $\{P_s\}_{s\in S}$ is shown in Lemma 1. $\square$

Proof of Theorem 4:

To show that the profit maximization problem with MLR (34),

$$\begin{aligned}
\underset{\{w_s,P_s\}_{s\in S}}{Max}\ & \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i-\sum_{s\in S}\left[C_s^f(w_s^{MA})+C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s^{MA}))\right]\\
s.t.\ & \sum_{i\in I}\lambda_i(\boldsymbol{w})r_i-\sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s)P_s=0\\
& w_s^{MA}=\underset{w_s\in[\underline{\mathrm{w}}_s,1]}{ArgMax}\{\sum_{i\in I}\lambda_i^s(w_s)P_s-C_s^f(w_s)-C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\}\ \ \forall s\in S,
\end{aligned}$$

yields the same equilibrium quality level as the profit maximization problem without MLR (35),

$$\begin{aligned}
\underset{\{w_s\}_{s\in S}}{Max}\ & \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i-\sum_{s\in S}\left[C_s^f(w_s^{MA})+C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s^{MA}))\right]\\
s.t.\ & \underline{\mathrm{w}}_s\le w_s\le 1\ \ \forall s\in S
\end{aligned}$$

$\Leftrightarrow$

$$\begin{aligned}
\underset{\{w_s,P_s\}_{s\in S}}{Max}\ & \sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i-\sum_{s\in S}\left[C_s^f(w_s^{MA})+C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s^{MA}))\right]\\
s.t.\ & w_s^{MA}=\underset{w_s\in[\underline{\mathrm{w}}_s,1]}{ArgMax}\{\sum_{i\in I}\lambda_i^s(w_s)P_s-C_s^f(w_s)-C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\}\ \ \forall s\in S,
\end{aligned}$$

it suffices to show that the equilibrium quality level $\{w_s^{MA}\}_{s\in S}$ and the corresponding service reimbursements $\{P_s^{MA}\}_{s\in S}$ of (35) satisfy the MLR constraint, i.e.,

$$\sum_{i\in I}\lambda_i(\boldsymbol{w}^{MA})r_i-\sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s^{MA})P_s^{MA}=0. \tag{41}$$

To see this, we first note that since $\{w_s^{MA}\}_{s\in S}$ and $\{P_s^{MA}\}_{s\in S}$ satisfy

$$w_s^{MA}=\underset{w_s\in[\underline{\mathrm{w}}_s,1]}{ArgMax}\{\sum_{i\in I}\lambda_i^s(w_s)P_s^{MA}-C_s^f(w_s)-C_s^v(\sum\nolimits_{i\in I}\lambda_i^s(w_s))\}\ \ \forall s\in S,$$

we have

$$\sum_{i\in I}\left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\right)\Big|_{w_s=w_s^{MA}}P_s^{MA}=\left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\frac{dC_s^v(\sum_{i\in I}\lambda_i^s(w_s))}{d(\sum_{i\in I}\lambda_i^s(w_s))}+\frac{dC_s^f(w_s)}{dw_s}\right)\Big|_{w_s=w_s^{MA}}\quad\forall s\in S, \tag{42}$$

when $w_s>\underline{\mathrm{w}}_s$ and

$$\sum_{i\in I}\left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\right)\Big|_{w_s=w_s^{MA}}P_s^{MA}=\left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\frac{dC_s^v(\sum_{i\in I}\lambda_i^s(w_s))}{d(\sum_{i\in I}\lambda_i^s(w_s))}+\frac{dC_s^f(w_s)}{dw_s}\right)\Big|_{w_s=w_s^{MA}}-L_s\quad\forall s\in S \tag{43}$$

for some Lagrangian multiplier $L_s > 0$ when $w_s = \underline{w}_s$.

Furthermore, since $\{w_s^{MA}\}_{s\in S}$ and $\{P_s^{MA}\}_{s\in S}$ are the optimizers of (35), they satisfy the following first order condition

$$\sum_{i\in I}\left(\frac{\partial\sum_{i\in I}\lambda_i(\boldsymbol{w})}{\partial w_s}\right)\Big|_{w_s=w_s^{MA}} r_i = \left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\frac{dC_s^v(\sum_{i\in I}\lambda_i^s(w_s))}{d(\sum_{i\in I}\lambda_i^s(w_s))} + \frac{dC_s^f(w_s)}{dw_s}\right)\Big|_{w_s=w_s^{MA}} \quad \forall s\in S, \tag{44}$$

when $w_s > \underline{w}_s$ and

$$\sum_{i\in I}\left(\frac{\partial\sum_{i\in I}\lambda_i^s(\boldsymbol{w})}{\partial w_s}\right)\Big|_{w_s=w_s^{MA}} r_i = \left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\frac{dC_s^v(\sum_{i\in I}\lambda_i^s(w_s))}{d(\sum_{i\in I}\lambda_i^s(w_s))} + \frac{dC_s^f(w_s)}{dw_s}\right)\Big|_{w_s=w_s^{MA}} - L_s \quad \forall s\in S \tag{45}$$

for some Lagrangian multiplier $L_s > 0$ when $w_s = \underline{w}_s$.

By (42)-(45), we have

$$\sum_{i\in I}\left(\frac{\partial\sum_{i\in I}\lambda_i^s(w_s)}{\partial w_s}\right)\Big|_{w_s=w_s^{MA}} P_s^{MA} = \sum_{i\in I}\left(\frac{\partial\sum_{i\in I}\lambda_i(\boldsymbol{w})}{\partial w_s}\right)\Big|_{w_s=w_s^{MA}} r_i \quad \forall s\in S. \tag{46}$$

In addition, when the health plan demand function takes the additive form $\lambda_i(\boldsymbol{w}) = \sum_{s\in S}\Lambda_i\xi_{is}w_s$, (46) implies that

$$\sum_{i\in I}\sum_{s\in S}\Lambda_i\xi_{is}r_i = \sum_{i\in I}\sum_{s\in S}\Lambda_i\xi_{is}P_s^{MA}.$$

Therefore, we have

$$\sum_{i\in I}\lambda_i(w^{MA})r_i = \sum_{i\in I}\sum_{s\in S}\Lambda_i\xi_{is}w_s^{MA}r_i = \sum_{i\in I}\sum_{s\in S}\Lambda_i\xi_{is}w_s^{MA}P_s^{MA} = \sum_{i\in I}\sum_{s\in S}\lambda_i^s(w_s^{MA})P_s^{MA},$$

as in (41). □

D. CMS-HCC Risk Adjustment

CMS pays MA health plans a risk-adjusted capitation rate per patient per year for each enrolled Medicare beneficiary in the MA health plans. This risk-adjusted capitation rate is calculated using the CMS-HCC risk adjustment model. This section first explains the specification and estimation of this model (§D.1). Subsequently, we illustrate how CMS uses the CMS-HCC risk adjustment model to calculate the capitation rate for each Medicare beneficiary enrolling in Medicare Advantage (§D.2). Lastly , we explain how these institutional details of the CMS-HCC Risk Adjustment model are reflected in our theoretical model in the main text (§D.3).

D.1. Model Specification and Estimation

According to §2.9.3 of Pope et al. (2011), "the specification of the CMS-HCC model is a linear regression in which expenditures are predicted by diagnoses (CMS-HCCs) and demographics". That is, the CMS-HCC risk adjustment model has the following specification:

$$R_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + \epsilon_i, \tag{47}$$

where R_i is the healthcare spending on a type i Medicare beneficiary; $\boldsymbol{x}_i$ is a vector of individual risk factors capturing this Medicare beneficiary's diagnoses and demographics (e.g., age and gender).

Here, the Medicare beneficiary's diagnoses are represented by the CMS Hierarchical Condition Categories (CMS-HCCs), a diagnostic classification system. As illustrated by Figure 2, CMS-HCCs are hierarchical condition categories aggregations of the ICD-9-CM codes, an international disease classification system. Specifically, the CMS-HCCs first aggregate 14,000 ICD-9-CM diagnosis codes into 805 diagnostic groups (DXGs), each of which represents a well-specified medical condition. DXGs are further aggregated into 189 Condition Categories (CCs), each of which represents a class of diseases that are related clinically and with respect to cost. Subsequently, hierarchies are imposed on CCs to represent the most severe manifestation among related diseases. Lastly, the CMS-HCC V12 model includes the 70 HCCs (out of the 189 HCCs) that best explain medical expenditures in TM. Further details about the CMS Hierarchical Condition Categories can be found in §2.5 of Pope et al. (2011).

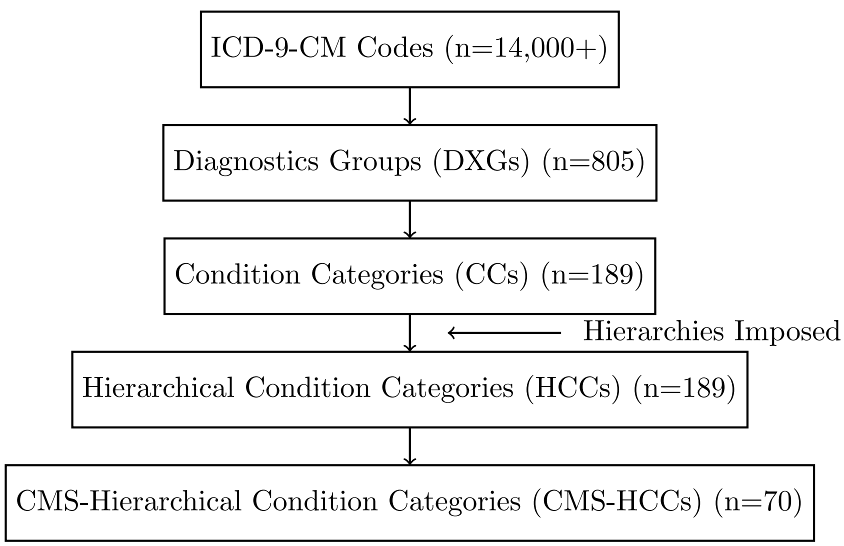

Figure 2 **Hierarchical Condition Categories aggregations of ICD-9-CM codes, version 12 CMS-HCC model (Pope et al. 2011)**

CMS estimates the CMS-HCC risk adjustment model (47) using claims data from TM. Specifically, for each TM enrollee, the claims data contain his/her healthcare cost (R_i) and disease diagnoses and demographics ($\boldsymbol{x}_i$) information. First, to correct for potential measurement errors in the claims data, CMS first preprocesses the data with a set of FFS adjusters (Centers for Medicare & Medicaid Services 2012). Subsequently, through an Ordinary Least Squares (OLS) estimation framework, CMS obtain estimates $\hat{\boldsymbol{\beta}}$ for the model coefficient $\boldsymbol{\beta}$.

D.2. Model Implementation

As discussed in the Introduction, the Social Security Act requires that the capitation payment for an Medicare beneficiary enrolled in MA should reflect the cost of a similar Medicare beneficiary in TM. In practice, this capitation payment for a type i MA enrollee is calculated as

$$r_i = \alpha_i \times Benchmark, \tag{48}$$

where α_i is the risk score of a type i MA enrollee; $Benchmark$ is the benchmark payment of an average (standard risk) Medicare beneficiary (Medicare Payment Advisory Commission 2012).[13] This section explains how CMS estimates these risk scores $\{\alpha_i\}_{i\in I}$ and the benchmark payment.

First, the benchmark payment $Benchmark$ is determined by the average TM spending per Medicare beneficiary.[14] That is, if the TM enrollment for each patient type $j \in I$ is $\lambda_j(\boldsymbol{w}^*)$, the benchmark payment is

$$Benchmark = \sum_{j\in I} \frac{\lambda_j(\boldsymbol{w}^*)}{\sum_{j\in I}\lambda_j(\boldsymbol{w}^*)}\mathbb{E}[R_j]. \tag{49}$$

In other words, $Benchmark$ represents the healthcare cost of a standard risk TM enrollee.

Second, the risk scores $\{\alpha_i\}_{i\in I}$ are determined by the CMS-HCC risk adjustment model (47). Specifically, for a type i MA enrollee with individual characteristics $\boldsymbol{x}_i$ (i.e., age, gender, pre-existing conditions), the CMS first predicts his/her healthcare cost using the CMS-HCC risk adjustment model (47), i.e., $\mathbb{E}[R_i] = \boldsymbol{x}_i^T\hat{\boldsymbol{\beta}}$. Subsequently, CMS converts these predicted costs $\{\mathbb{E}[R_i]\}_{i\in I}$ to risk scores $\{\alpha_i\}_{i\in I}$, i.e.,

$$\alpha_i = \frac{\mathbb{E}[R_i]}{\sum\limits_{j\in I}\frac{\lambda_j(\boldsymbol{w}^*)}{\sum_{j\in I}\lambda_j(\boldsymbol{w}^*)}\mathbb{E}[R_j]}, \tag{50}$$

so that "payment adjustments can be made relative to the average Medicare beneficiary".[15] Here, $\sum\limits_{j\in I}\frac{\lambda_j(\boldsymbol{w}^*)}{\sum_{j\in I}\lambda_j(\boldsymbol{w}^*)}\mathbb{E}[R_j]$ is the average TM spending per Medicare beneficiary across all patient types $j\in I$. As such, the risk score of an average Medicare beneficiary is normalized to 1 (Better Medicare Alliance 2018).

[13] See Page 95 in Medicare Payment Advisory Commission (2012): "Health plans that participate in the Medicare Advantage (MA) program receive monthly capitated payments for each Medicare enrollee. Each capitated payment is the product of two general parts: a base rate, which reflects the payment if an MA enrollee has the health status of the national average beneficiary, and a risk score, which indicates how costly the enrollee is expected to be relative to the national average beneficiary".

[14] See Medicare Payment Advisory Commission (2021): "The local MA benchmarks are determined under statutory formulas whereby county-level rates vary depending on average FFS spending per Medicare beneficiary".

[15] See §2.1 of Centers for Medicare & Medicaid Services (2018b): "The predicted costs from the risk adjustment models are then converted to relative risk factors so that payment adjustments can be made relative to the average Medicare beneficiary".

D.3. Link to Our Theoretical Model

This section illustrates how our theoretical model in the main text captures these institutional details of CMS-HCC risk adjustment. Specifically, we explain how data used in the CMS-HCC risk adjustment formula, i.e., R_i, $\mathbb{E}[R_i]$, r_i, $\hat{\boldsymbol{\beta}}$ are generated in our theoretical model, as summarized in Table 1. Particularly, we demonstrate that the capitation rates (7) derived in our theoretical model reflects the capitation payment calculation (48) in practice.

We first explain how our theoretical model generates $\mathbb{E}[R_i]$, which is CMS's expected spending on a type i patient in TM. By definition, this statistics can be represented as

$$\mathbb{E}[R_i] = \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*,$$

where the numerator and denominator stand for the total TM spending and the total enrollment of type i Medicare beneficiaries in TM, respectively. The service demand $\{\lambda_i^s(w_s^*)\}_{s\in S, i\in I}$ and service prices $\{P_s^*\}_{s\in S}$ were derived in §4, while the type i Medicare beneficiaries enrollment in TM $\lambda_i(\boldsymbol{w}^*)$ was derived in §5.

Lastly, we explain the link between our capitation rate characterization (7) and the capitation payment calculation (48) in practice. Specifically, by (49) and (50), the capitation payment calculation (48) reduces to

$$r_i = \mathbb{E}[R_i],$$

which can be defined as the expected TM spending per Medicare beneficiary in risk group i. Therefore, (7), i.e., $r_i = \mathbb{E}[R_i] = \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*$, directly follows from this definition.

Definition	Data	Data Generating Process
CMS's spending on a type i patient in Traditional Medicare	R_i	$R_i := \boldsymbol{x}_i^T \boldsymbol{\beta} + \epsilon_i$
CMS's expected spending on a type i patient in Traditional Medicare	$\mathbb{E}[R_i]$	$\mathbb{E}[R_i] := \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*$
MA capitation rate for a type i patient	r_i	$r_i := \mathbb{E}[R_i] = \sum_{s\in S} \frac{\lambda_i^s(w_s^*)}{\lambda_i(\boldsymbol{w}^*)} P_s^*$
CMS-HCC risk adjustment model parameter	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\beta}}$ s.t. $\boldsymbol{x}_i^T \hat{\boldsymbol{\beta}} = \mathbb{E}[R_i] = r_i$

Table 1 **CMS-HCC Risk Adjustment Formula: Data and Data Generating Process**