# Appendix. Online Appendix

TABLE 3. Glossary of Key Notation

| **Symbol** | **Description** |
|---|---|
| $F(\boldsymbol{m}) := F(\boldsymbol{m}; (\lambda_w, G_w, G_w^{-1})_{w\in\mathbb{W}})$ | Objective of the ground-truth matching problem (GTMP) in (6). |
| $\hat{F}(\boldsymbol{m}) := F(\boldsymbol{m}; (\hat{\lambda}_w, \hat{G}_w, \hat{G}_w^{-1})_{w\in\mathbb{W}})$ | Objective of the data-driven matching problem (DDMP) in (10) |
| $\hat{F}^{\mathrm{KA}}(\boldsymbol{m}) := F(\boldsymbol{m}; (\lambda_w, \hat{G}_w, \hat{G}_w^{-1})_{w\in\mathbb{W}})$ | Objective of the DDMP with known arrival rates (DDMP-KA) in (11). |
| $q_w(p) := q(p; \lambda_w, G_w, G_w^{-1})$ | The invariant queue-length function of type-$w$ agent, where $q(p;\lambda,G,G^{-1})$ is defined in (7). |
| $\hat{q}_w(p) := q(p; \hat{\lambda}_w, \hat{G}_w, \hat{G}_w^{-1})$ | The data-driven invariant queue-length function. |
| $\hat{q}_w^{\mathrm{KA}}(p) := q(p; \lambda_w, \hat{G}_w, \hat{G}_w^{-1})$ | The data-driven invariant queue-length function with known arrival rate. |
| $p_w(\boldsymbol{m}) := p_w(\boldsymbol{m}; \lambda_w)$ | The unmatched-proportion function, where $p_w(\boldsymbol{m};\lambda)$ is defined in (8). |
| $\hat{p}_w(\boldsymbol{m}) := p_w(\boldsymbol{m}; \hat{\lambda}_w)$ | The data-driven unmatched-proportion function. |
| $\mathbb{M} := \mathbb{M}((\lambda_w)_{w\in\mathbb{W}})$ | Feasible region of the GTMP (6) and DDMP-KA (11). |
| $\hat{\mathbb{M}} := \mathbb{M}((\hat{\lambda}_w)_{w\in\mathbb{W}})$ | Feasible region of the DDMP (10). |
| $\boldsymbol{m}^\star \in \operatorname{argsup}_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m})$ | An optimal solution of the GTMP (6). |
| $\hat{\boldsymbol{m}} \in \operatorname{argsup}_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m})$ | An optimal solution of the DDMP (10). |
| $\hat{\boldsymbol{m}}^{\mathrm{KA}} \in \operatorname{argsup}_{\boldsymbol{m}\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m})$ | An optimal solution of the DDMP-KA (11). |
| $\hat{\Delta}_w^G$ | Estimation error related to type-$w$ agent's reneging CDF defined in (12). |
| $\hat{\Delta}_w^{G^{-1}}$ | Moderated uniform estimation error of type-$w$ agent's reneging quantile function, where the moderating function is $1-G_w$, defined in (12). |
| $\hat{\Delta}_{w,f}^{G^{-1}}$ | Moderated uniform estimation error of type-$w$ agent's reneging quantile function, where the moderating function is $f$, defined in (4). |
| $\hat{\alpha}_i^\lambda, i\in\{1,2,3\}$ | Estimation errors related to agent's arrival rates, defined in (32). |
| $\Phi_w(B_w^G)$ | A function related to the finite-sample bound on $\hat{\Delta}_w^G$, defined in Lemma 1. |
| $\Psi_{wi}(B_w^G;\delta,\epsilon), i\in\{1,2,3\}$ | Functions related to the finite-sample bound on $\hat{\Delta}_w^{G^{-1}}$, defined in (22). |
| $\Lambda_w(B_w^\lambda)$ | A function related to the finite-sample bound on $\hat{\alpha}_i^\lambda$, defined in (33). |
| $\chi(\epsilon)$ | A function that quantifies the well-separateness of $\boldsymbol{m}^\star$, defined in (31). |
| $\rho_w(\delta)$ | A function that controls the tail behavior of the reneging time density function, defined in Assumption 1. |
| $y_w(\epsilon)$ | A function that solves $\int_{y_w(\epsilon)}^{H_w} 1-G_w = \epsilon$, defined in (21). |
| $\zeta(\epsilon_1,\epsilon_2,\epsilon_3)$ | A function that maps from $\mathbb{R}_+^3$ to $\mathbb{R}_+$, defined in (34). |
| $\nu_i, i\in\{1,2,3\}$ | Constants defined in Theorems 3 and 4 by $\nu_1 = [4(J+K)\max_{w\in\mathbb{W}}\{c_w\lambda_w\}]^{-1}$, $\nu_2 = [3\sum_{w\in\mathbb{W}}\lambda_w L_F]^{-1}$ and $\nu_3 = [6\sum_{w\in\mathbb{W}} c_w\lambda_w(1/\theta_w + L)]^{-1}$. |

In Appendix A, we provide the proof of Proposition 1, which appears in §3, and discuss the required divergence condition (Remark 7). In Appendix B, we provide omitted proofs for results in §4. In Appendix C, we provide omitted proofs for §6 and §7. In Appendix D, we provide omitted proofs and results for §8. Finally, in Appendix E, we provide omitted proofs and results for §9.

**A. Proof of Proposition 1 in §3 and Remark on the Divergence Condition** In this section, we fix $w\in\mathbb{W}$, omit the subscript $w$, and write $B_w^G$ as $B$ for simplicity.

We first prove (15). Following (14) in the proof of Theorem 1, we have

$$\begin{aligned}\left|\frac{q(p(\boldsymbol{m}))}{\lambda}-\frac{\hat{q}^{\mathrm{KA}}(p(\boldsymbol{m}))}{\lambda}\right| &= \left|\int_{\hat{G}^{-1}(p(\boldsymbol{m}))}^{G^{-1}(p(\boldsymbol{m}))}(1-G(u))du+\int_{0}^{\hat{G}^{-1}(p(\boldsymbol{m}))}\hat{G}(u)-G(u)du\right| \\ &\geq \left|\int_{\hat{G}^{-1}(p(\boldsymbol{m}))}^{G^{-1}(p(\boldsymbol{m}))}(1-G(u))du\right|-\left|\int_{0}^{\hat{G}^{-1}(p(\boldsymbol{m}))}\hat{G}(u)-G(u)du\right| \\ &\geq \left|\int_{\hat{G}^{-1}(p(\boldsymbol{m}))}^{G^{-1}(p(\boldsymbol{m}))}(1-G(u))du\right|-\int_{0}^{\hat{G}^{-1}(p(\boldsymbol{m}))}\left|\hat{G}(u)-G(u)\right|du,\end{aligned} \tag{53}$$

where the first inequality is due to $|a+b|\geq|a|-|b|$, and the second inequality is because $\int f\leq\int|f|$. Taking a supremum over $\boldsymbol{m}\in\mathbb{M}$ in both sides of (53) and using the inequality that $\sup(f-g)\geq\sup(f)-\sup(g)$, we obtain

$$\begin{aligned}&\sup_{\boldsymbol{m}\in\mathbb{M}}\left|\frac{q(p(\boldsymbol{m}))}{\lambda}-\frac{\hat{q}^{\mathrm{KA}}(p(\boldsymbol{m}))}{\lambda}\right| \\ &\geq \sup_{\boldsymbol{m}\in\mathbb{M}}\left|\int_{\hat{G}_j^{-1}(p(\boldsymbol{m}))}^{G^{-1}(p(\boldsymbol{m}))}(1-G(u))du\right|-\sup_{\boldsymbol{m}\in\mathbb{M}}\int_{0}^{\hat{G}^{-1}(p(\boldsymbol{m}))}\left|\hat{G}(u)-G(u)\right|du \\ &= \hat{\Delta}^{G^{-1}}-\hat{\Delta}^{G},\end{aligned} \tag{54}$$

where the identity are due to the facts that taking supremum over $\boldsymbol{m}\in\mathbb{M}$ is equivalent to taking supremum over $p(\boldsymbol{m})\in[0,1]$, that $\hat{G}^{-1}(p(\boldsymbol{m}))\leq H$ for all $\boldsymbol{m}\in\mathbb{M}$, and the definitions of $\hat{\Delta}^{G^{-1}}$ and $\hat{\Delta}^{G}$ in (12).

Next, we prove (16). To this end, it is enough to show that $\hat{\Delta}^{G}/\hat{\Delta}^{G^{-1}}=o_P(1)$ under the stated divergence condition as the sample size $B\to\infty$. For ease of notation, define $A_B^G:=\sqrt{B}\hat{\Delta}^G$ and $A_B^{G^{-1}}:=\sqrt{B}\hat{\Delta}^{G^{-1}}$. Then it suffices to show that $A_B^G/A_B^{G^{-1}}=o_P(1)$ as $B\to\infty$. We proceed in two steps: 1) showing the stochastic boundedness of $A_B^G$, i.e., showing $A_B^G=O_P(1)$ as $B\to\infty$, and 2) showing the stochastic unboundedness of $A_B^{G^{-1}}$, i.e., showing that $A_B^{G^{-1}}$ converges to infinity in probability as $B\to\infty$. With these two steps, it is clear that $A_B^G/A_B^{G^{-1}}=o_P(1)$ as $B\to\infty$, establishing (16).

**Step 1: Stochastic boundedness of $A_B^G$.** By Lemma 1, for any sufficiently small $\epsilon>0$, there always exists a finite constant $M_\epsilon$ such that

$$\mathbb{P}(A_B^G\geq M_\epsilon)=\mathbb{P}(\hat{\Delta}^G\geq B^{-1/2}M_\epsilon)\leq a_1\exp(-a_2M_\epsilon^2)\leq\epsilon$$

for all $B\geq 1$. Therefore, we have $A_B^G=O_P(1)$.

**Step 2: Stochastic unboundedness of $A_B^{G^{-1}}$.** Define $D_B(p):=\sqrt{B}\int_{\hat{G}_j^{-1}(p)}^{G^{-1}(p)}(1-G(u))du$, and so $A_B^{G^{-1}}=\sup_{p\in[0,1]}|D_B(p)|$. We aim to show that for any arbitrarily large constant $M>0$,

$$\lim_{B\to\infty}\mathbb{P}(A_B^{G^{-1}}>M)=1.$$

Since $A_B^{G^{-1}}=\sup_{p\in[0,1]}|D_B(p)|$, we have for any given $p_l\in(0,1)$,

$$\lim_{B\to\infty}\mathbb{P}(A_B^{G^{-1}}>M)\geq\lim_{B\to\infty}\mathbb{P}(|D_B(p_l)|>M).$$

This implies that for any sequence of $p_l\in(0,1)$,

$$\lim_{B\to\infty}\mathbb{P}(A_B^{G^{-1}}>M)\geq\sup_{p_l}\lim_{B\to\infty}\mathbb{P}(|D_B(p_l)|>M) \tag{55}$$

We will show that right-hand side of (55) is one to complete the proof. To this end, we next perform a deeper analysis of the term $D_B(p) = \sqrt{B}\int_{\hat{G}_j^{-1}(p)}^{G^{-1}(p)}(1-G(u))du$ for any fixed $p \in (0,1)$. Consider the function $f(t) := \int_0^t 1-G(u)du$, which satisfies $f'(t) = 1-G(t)$. A Taylor expansion around $t = G^{-1}(p)$ yields

$$\begin{aligned}\frac{D_B(p)}{\sqrt{B}} = \int_{\hat{G}_j^{-1}(p)}^{G^{-1}(p)}(1-G(u))du &= f(G^{-1}(p)) - f(\hat{G}^{-1}(p))\\ &= f'(G^{-1}(p))(G^{-1}(p)-\hat{G}^{-1}(p)) + o\big(G^{-1}(p)-\hat{G}^{-1}(p)\big)\\ &= (1-p)(G^{-1}(p)-\hat{G}^{-1}(p)) + o\big(G^{-1}(p)-\hat{G}^{-1}(p)\big).\end{aligned}$$

Using the Bahadur representation [7] for $G^{-1}(p)$, we have

$$G^{-1}(p) - \hat{G}^{-1}(p) = \frac{\hat{G}(G^{-1}(p)) - p}{g(G^{-1}(p))} + o_P(B^{-1/2}).$$

Combining the above two displays and using $\hat{G}(G^{-1}(p)) = \sum_{b=1}^{B}\mathbb{I}(r_b^G \le G^{-1}(p))/B$ from (9), we obtain

$$D_B(p) = \frac{1-p}{g(G^{-1}(p))}\frac{1}{\sqrt{B}}\sum_{b=1}^{B}\Big(\mathbb{I}\big(r_b^G \le G^{-1}(p)\big) - p\Big) + o_P(1) = \frac{1-p}{g(G^{-1}(p))}\frac{1}{\sqrt{B}}\sum_{b=1}^{B}\psi_b^G + o_P(1),$$

where we have defined i.i.d. random variables $\psi_b^G := \mathbb{I}\big(r_b^G \le G^{-1}(p)\big) - p$ for $b \in [B]$, with mean $\mathbb{E}[\psi_b^G] = 0$ and variance $\sigma_\psi^2 = p(1-p)$. Then, by the central limit theorem, we have

$$D_B(p) \xrightarrow{d} \mathcal{N}(0, \sigma^2(p)) \quad \text{as} \quad B \to \infty, \tag{56}$$

where $\sigma^2(p) := p(1-p)^2/g^2(G^{-1}(p))$ is the asymptotic variance, and $\mathcal{N}(0,\sigma^2(p))$ represents a normal distribution with mean 0 and variance $\sigma^2(p)$. Using (56), it holds for any arbitrarily large constant $M$ that

$$\lim_{B\to\infty}\mathbb{P}(|D_B(p_l)| > M) = \mathbb{P}(|\mathcal{N}(0,\sigma^2(p_l))| > M) = 2\Big(1 - \Phi_{\text{normal}}\Big(\frac{M}{\sigma(p_l)}\Big)\Big),$$

where $\Phi_{\text{normal}}$ is the CDF of a standard normal distribution. By the divergence condition stated in Proposition 1, there exists a sequence of $p_l$ such that the variance $\sigma^2(p_l) \to \infty$. For this specific sequence, we have

$$\sup_{p_l}\lim_{B\to\infty}\mathbb{P}(|D_B(p_l)| > M) = \sup_{p_l}\Big\{2\Big(1-\Phi_{\text{normal}}\Big(\frac{M}{\sigma(p_l)}\Big)\Big)\Big\} = 2\Big(1-\Phi_{\text{normal}}(0)\Big) = 1 \tag{57}$$

In view of (55), this establishes that $\lim_{B\to\infty}\mathbb{P}(A_B^{G^{-1}} > M) = 1$ for any $M$, completing the proof for (16).

Finally, we prove (17). Note that the upper bound follows directly from Claim 2 and Claim 3 in the proof of Theorem 1, and the upper bound in (16). We next establish the lower bound in (17) under the assumption that $c_w > 0$ for at most one $w \in \mathbb{W}$. By comparing the ground-truth objective function $F$ in (6) and the data-driven objective $\hat{F}^{\text{KA}}$ in (11), we have

$$\sup_{\boldsymbol{m}\in\mathbb{M}}\big|F(\boldsymbol{m}) - \hat{F}^{\text{KA}}(\boldsymbol{m})\big| = \sup_{\boldsymbol{m}\in\mathbb{M}}\Big|\sum_{w\in\mathbb{W}}c_w\big(q_w(p_w(\boldsymbol{m})) - \hat{q}_w^{\text{KA}}(p_w(\boldsymbol{m}))\big)\Big|.$$

Hence, if $c_w = 0$ for all $w \in \mathbb{W}$, the lower bound in (17) trivially holds. It remains to consider the case where $c_w > 0$ for exactly one $w \in \mathbb{W}$. Let $w'$ denote the agent type with positive holding cost $c_{w'} > 0$. Then we have

$$\begin{aligned}\sup_{\boldsymbol{m}\in\mathbb{M}}\big|F(\boldsymbol{m}) - \hat{F}^{\text{KA}}(\boldsymbol{m})\big| &= \sup_{\boldsymbol{m}\in\mathbb{M}}\big|c_{w'}\big(q_{w'}(p_{w'}(\boldsymbol{m})) - \hat{q}_{w'}^{\text{KA}}(p_{w'}(\boldsymbol{m}))\big)\big|\\ &\ge c_{w'}\lambda_{w'}(\hat{\Delta}_{w'}^{G^{-1}} - \hat{\Delta}_{w'}^{G})\\ &= c_{w'}\lambda_{w'}\hat{\Delta}_{w'}^{G^{-1}}(1 - o_P(1))\\ &= \sum\nolimits_{w\in\mathbb{W}}c_w\lambda_w\hat{\Delta}_w^{G^{-1}}(1-o_P(1)),\end{aligned}$$

where the inequality follows from (54), the second-to-last identity is due to $\hat{\Delta}^{G}_{w'}/\hat{\Delta}^{G^{-1}}_{w'} = o_P(1)$ under the divergence condition, and the last identity follows from $c_w = 0$ for all $w \neq w'$. □

REMARK 7. The condition that there exists a sequence of $p_l \in (0,1)$ such that $p_l(1-p_l)^2/g_w^2(G_w^{-1}(p_l))$ diverges to infinity as $l \to \infty$ holds whenever the reneging density $g_w$ decays sufficiently fast to zero near the boundary of $(0, H_w)$. In particular, many standard distribution families satisfy this condition through the lower-tail sequences $p_l \to 0$, including the Weibull and gamma distributions with shape parameter exceeding 2, the beta distribution with first shape parameter exceeding 2, as well as the logistic and lognormal distributions.

**B. Omitted Proofs for Results in Section 4** In this section, we fix $w \in \mathbb{W}$, omit the subscript $w$, and write $B_w^G$ as $B$ for simplicity.

**Proof of Proposition 2.** We first prove part (i). Let $\underline{g} > 0$ be such that $g(x) \geq \underline{g}$ for all $x \in [0, H]$. Assumption 1 is satisfied by $\rho(x) = \underline{g}$ for all $x$. To use (23) of Corollary 1, we need to pick $\delta$ such that $\delta \in [0, \min\{1/2, 1 - G(y(\epsilon)), G(\epsilon)\}]$. We next show that setting $\delta = \delta(\epsilon) := \underline{g}\epsilon/\ln(1/\epsilon)$ suffices for this purpose. Because $H < \infty$ and $g(x) \geq \underline{g}$, we have

$$1 - G(y(\epsilon)) = \int_{y(\epsilon)}^{H} g(u)du \geq (H - y(\epsilon))\underline{g} = \underline{g}\int_{y(\epsilon)}^{H} 1du \geq \underline{g}\int_{y(\epsilon)}^{H} 1 - G(u)du = \underline{g}\epsilon,$$

where the second inequality is because $0 \leq 1 - G(u) \leq 1$, and the last equality is due to the definition of $y(\epsilon)$. Also, we have $G(\epsilon) = \int_0^\epsilon g(u)du \geq \underline{g}\epsilon$. Therefore, by setting $\delta = \delta(\epsilon) := \underline{g}\epsilon/\ln(1/\epsilon)$, the condition on $\delta$ in Corollary 1 (i.e., $\delta \in [0, \min\{1/2, 1 - G(y_f(\epsilon)), G(\epsilon)\}]$) is satisfied for all sufficiently small $\epsilon$. It follows that $\delta(\epsilon) \leq \min\{1/2, 1 - G(y(\epsilon)), G(\epsilon)\}$ for all sufficiently small $\epsilon$.

For any $\eta > 0$, suppose we can show that there exists a finite $M_\eta$ such that setting $\epsilon = M_\eta \ln(B)/\sqrt{B}$ yields $\Psi_i(B; \delta(\epsilon), \epsilon) \leq \eta/3, i \in \{1,2,3\}$ for all sufficiently large $B$. Then (23) of Corollary 1 implies that $\mathbb{P}(\hat{\Delta}^{G^{-1}} \geq \epsilon) \leq \sum_{i\in\{1,2,3\}} \Psi_i(B; \delta(\epsilon), \epsilon) \leq \eta$ for all sufficiently large $B$, and thus $\limsup_B \mathbb{P}(\hat{\Delta}^{G^{-1}} \geq M_\eta \ln(B)/\sqrt{B}) \leq \eta$. In the following, we establish the above claims for $\Psi_i, i \in \{1,2,3\}$ to complete the proof.

We start with $\Psi_3$. Set $\epsilon = M_\eta \ln(B)/\sqrt{B}$. By (22), we have $\ln \Psi_3(B; \delta(\epsilon), \epsilon) = \ln 2 - \frac{1}{2}B\delta^2(\epsilon)$ for all large $B$, because $\delta(\epsilon) \leq \epsilon \underline{g} = \epsilon\rho(\delta(\epsilon))$ for all sufficiently small $\epsilon$. It follows from the definition of $\delta(\epsilon)$ that

$$\ln \Psi_3(B; \delta(\epsilon), \epsilon) = \ln 2 - \frac{1}{2}\underline{g}^2 \frac{M_\eta^2 \ln^2 B}{[\frac{1}{2}\ln B - \ln(\ln B)]^2} \leq \ln 2 - \frac{1}{2}\underline{g}^2 M_\eta^2 \times 3,$$

where the inequality holds because $\frac{\ln^2 B}{[\frac{1}{2}\ln B - \ln(\ln B)]^2} \geq 3$ for all large enough $B$. By setting $M_\eta$ large enough, we can further have $\ln 2 - \frac{1}{2}\underline{g}^2 M_\eta^2 \times 3 \leq \ln(\eta/3)$, ensuring that $\ln \Psi_3(B; \delta(\epsilon), \epsilon) \leq \eta/3$.

Next, we consider $\Psi_1$ and $\Psi_2$. By (22) and Lemma 4, it holds for all $\delta \in (0, 1/2]$ and $B \geq 2/\delta$ (which implies $B\delta + 1 \leq 2B\delta$) that

$$\begin{aligned}
&\Psi_1(B; \delta, \epsilon) \leq \check{\Psi}_1(B; \delta, \epsilon) := 2B\delta\sqrt{B\delta}\exp(2B\delta)\delta^{-B\delta}[G(y_f(\epsilon))]^{B(1-\delta)} \\
&\Psi_2(B; \delta, \epsilon) \leq \check{\Psi}_2(B; \delta, \epsilon) := 2(B - B\delta)\sqrt{B\delta}\exp(2B\delta)\delta^{-B\delta}\left[1 - G(\epsilon)\right]^{B - B\delta - 1}.
\end{aligned} \tag{58}$$

In the following, we assume $B \geq 2/\delta(\epsilon)$ and work with the upper bound $\check{\Psi}_1$ and $\check{\Psi}_2$. Our goal is to show that by setting $M_\eta$ large enough, $\check{\Psi}_i(B; \delta(\epsilon), \epsilon) \leq \eta/3, i \in \{1,2\}$ for all sufficiently large $B$, where $\epsilon = M_\eta \ln(B)/\sqrt{B}$ and $\delta(\epsilon) = \underline{g}\epsilon/\ln(1/\epsilon)$. Taking logarithmic for $\check{\Psi}_1$ yields

$$\begin{aligned}
\ln \check{\Psi}_1(B; \delta(\epsilon), \epsilon) &= B[2\delta(\epsilon) - \delta(\epsilon)\ln\delta(\epsilon) + (1 - \delta(\epsilon))\ln(G(y(\epsilon)))] + \frac{3}{2}\ln B + \frac{3}{2}\ln\delta(\epsilon) \\
&\overset{(a)}{\leq} B\left[2\delta(\epsilon) - \delta(\epsilon)\ln\delta(\epsilon) + (1 - \delta(\epsilon))\ln\left(1 - \underline{g}\epsilon\right)\right] + \frac{3}{2}\ln B \\
&\overset{(b)}{\leq} B\frac{\ln\left(1 - \underline{g}\epsilon\right)}{2} + \frac{3}{2}\ln B \overset{(c)}{\leq} B\frac{-\underline{g}\epsilon}{2} + \frac{3}{2}\ln B \overset{(d)}{\leq} -\frac{\underline{g}\epsilon}{4}B - \frac{3}{2} + \frac{3}{2}\ln\frac{6}{\underline{g}\epsilon},
\end{aligned} \tag{59}$$

where (a) follows from the facts that $G(y(\epsilon)) \le 1 - \underline{g}\epsilon$ and $\delta(\epsilon) \le 1$ for small $\epsilon$; (b) is due to the fact that $2\delta(\epsilon) - \delta(\epsilon)\ln\delta(\epsilon) + (1-\delta(\epsilon))\ln\left(1-\underline{g}\epsilon\right) \le \frac{\ln(1-\underline{g}\epsilon)}{2}$ for all sufficiently small $\epsilon$; (c) uses the inequality $\ln(1-x) \le -x$; and (d) is obtained by constructing a linear function with a slope of $\underline{g}\epsilon/4$ that is tangent to the concave function $B \mapsto \frac{3}{2}\ln B$, leading to the inequality $\frac{3}{2}\ln B \le \underline{g}\epsilon B/4 - \frac{3}{2} + \frac{3}{2}\ln\frac{6}{\underline{g}\epsilon}$ for all $B > 0$. Given (59), a sufficient condition for $\check{\Psi}_1(B;\delta(\epsilon),\epsilon) \le \eta/3$ is $-\frac{\underline{g}\epsilon}{4}B - \frac{3}{2} + \frac{3}{2}\ln\frac{6}{\underline{g}\epsilon} \le \ln(\eta/3)$, which is equivalent to

$$-\frac{\underline{g}M_\eta \ln B}{4\sqrt{B}}B - \frac{3}{2} + \frac{3}{2}\ln\frac{6\sqrt{B}}{\underline{g}M_\eta B} \le \ln\frac{\eta}{3}$$

by the definition of $\epsilon$. For the above condition to hold for large $B$, it suffices to set $M_\eta = 1$.

To analyze $\check{\Psi}_2$, we follow similar arguments used to arrive at (59) to obtain $\ln\check{\Psi}_2(B;\delta(\epsilon),\epsilon) \le -\frac{\underline{g}\epsilon}{4}B - \frac{3}{2} + \frac{3}{2}\ln\frac{6}{\underline{g}\epsilon} - \ln(1-\underline{g}\epsilon)$, where we have also used the facts that $1 - G(\epsilon) \le 1 - \underline{g}\epsilon, \ln(1-\delta(\epsilon)) < 0$ and $\ln\delta(\epsilon) < 0$ for all sufficiently small $\epsilon$. Then, a sufficient condition for $\check{\Psi}_2(B;\delta(\epsilon),\epsilon) \le \eta/3$ is $-\frac{\underline{g}\epsilon}{4}B - \frac{3}{2} + \frac{3}{2}\ln\frac{6}{\underline{g}\epsilon} - \ln(1-\underline{g}\epsilon) \le \ln(\eta/3)$, which is equivalent to

$$-\frac{\underline{g}M_\eta \ln B}{4\sqrt{B}}B - \frac{3}{2} + \frac{3}{2}\ln\frac{6\sqrt{B}}{\underline{g}M_\eta \ln B} - \ln(1-\underline{g}\frac{M_\eta \ln B}{\sqrt{B}}) \le \ln\frac{\eta}{3}.$$

Similarly, it suffices to set $M_\eta = 1$ for the above condition to hold for large $B$.

We next prove part (ii) to show that $\liminf_B \mathbb{P}(\hat{\Delta}^{G^{-1}} \ge \frac{M}{\sqrt{B}}) > 0$ for any $M > 0$. Suppose that $g(\cdot)$ is continuous and positive in a neighborhood of $G^{-1}(x_0)$ for some $x_0 \in (0,1)$. By (12), we have

$$\sqrt{B}\hat{\Delta}^{G^{-1}} \ge \sqrt{B}\left|\int_{\hat{G}^{-1}(x_0)}^{G^{-1}(x_0)} 1 - G(u)du\right| \ge \underbrace{\sqrt{B}\left|\hat{G}^{-1}(x_0) - G^{-1}(x_0)\right|}_{\text{(I)}}\underbrace{\left(1 - \hat{G}^{-1}(x_0) \vee G^{-1}(x_0)\right)}_{\text{(II)}}.$$

Now, as $B \to \infty$, term (I) converges in distribution to an absolute normal distribution by the continuous mapping theorem and standard central limit theorem for point-wise quantile estimation (see e.g. Chapter 21 of [52]), while term (II) converges in probability to $1 - G^{-1}(x_0)$ by the slutsky's theorem. As a result, for any $M > 0$, $\mathbb{P}\left(\sqrt{B}\hat{\Delta}^{G^{-1}} \ge M\right) > 0$ for all large $B$. □

**Proof of Lemma 2.** We need the following lemma whose proof is provided at the end of this subsection.

LEMMA 13. *Let G be a Weibull distribution with shape parameter s and scale parameter κ:*
*(a)* $\hat{\Delta}^{G^{-1}}(x) = O_p(\frac{\kappa}{s}\frac{(\ln\frac{1}{1-x})^{\frac{1}{s}}}{\sqrt{B}})$ *for any* $x \in (0,1)$ *as* $B \to \infty$*;*
*(b)* $\mathbb{P}\left(\hat{\Delta}^{G^{-1}}(0) \ge \frac{\kappa}{2e}B^{-\frac{1}{s}}\right) \ge e^{-(\frac{1}{2})^s} - e^{-B}$*;*
*(c)* $\mathbb{P}\left(\hat{\Delta}^{G^{-1}}(0) \le \kappa\right) \ge 1 - e^{-B}$*;*
*(d)* $\mathbb{P}\left(\hat{\Delta}^{G^{-1}}(1) \ge \frac{\kappa}{s}\frac{(2\ln B)^{\frac{1}{s}-1}}{B^2}\right) \ge (1 - \frac{1}{B^2})^B$ *for all* $s < 1$.
*(e)* $\mathbb{P}\left(\hat{\Delta}^{G^{-1}}(1) \le \frac{\kappa}{se}\right) \ge 1 - (1 - e^{-1})^B$ *for all* $s \ge 1$.

We first prove Lemma 2(i). By Lemma 13(a), for any $\eta > 0$ and $x \in (0,1)$, there exists some $B^1_{\eta,x} > 0$ and $M_{\eta,x}$ such that

$$\mathbb{P}\left(\hat{\Delta}^{G^{-1}}(x) \le \frac{\kappa}{s}\frac{(\ln\frac{1}{1-x})^{\frac{1}{s}}}{\sqrt{B}}M_{\eta,x}\right) \ge 1 - \frac{\eta}{3}$$

for all $B \ge B^1_{\eta,x}$. For a fixed $B$, define events $\Omega_1 := \{\hat{\Delta}^{G^{-1}}(x) \le \frac{\kappa}{s}\frac{(\ln\frac{1}{1-x})^{\frac{1}{s}}}{\sqrt{B}}M_{\eta,x}\}$, $\Omega_2 := \{\hat{\Delta}^{G^{-1}}(1) \le \frac{\kappa}{se}\}$ and $\Omega_3 := \{\hat{\Delta}^{G^{-1}}(0) \ge \frac{\kappa}{2e}B^{-\frac{1}{s}}\}$. For all large enough $s > s^1_{B,\eta,x}$, we have

$$\frac{\kappa}{2e}B^{-\frac{1}{s}} > \frac{\kappa}{s}\frac{(\ln\frac{1}{1-x})^{\frac{1}{s}}}{\sqrt{B}}M_{\eta,x} \quad \text{and} \quad \frac{\kappa}{2e}B^{-\frac{1}{s}} > \frac{\kappa}{se}.$$

Thus for large enough $s$, we have

$$\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(0) > \hat{\Delta}^{G^{-1}}(x)\big) \ge \mathbb{P}(\cap_{i=1}^{3}\Omega_i) \ge \sum_{i=1} 3\mathbb{P}(\Omega_i) - 2 \ge e^{-(\frac{1}{2})^s} - e^{-B} - \frac{\eta}{3} - (1-e^{-1})^B,$$

where the second inequality is the Bonferroni inequality, and last inequality follows from Lemma 13(b)&(e). To ensure that the rightmost term is no less than $1-\eta$, we pick $s > s_\eta^2$ large enough such $e^{-(\frac{1}{2})^s} > 1-\eta/3$, and $B > B_\eta^2$ large enough such that $e^{-B} + (1-e^{-1})^B < \eta/3$. Setting $B_{\eta,x} = \max\{B_{\eta,x}^1, B_\eta^2\}$ and $s_{B,\eta,x} = \max\{s_{B,\eta,x}^1, s_\eta^2\}$ completes the proof Lemma 2(i).

To prove Lemma 2(ii), define events $\Omega_1 := \{\hat{\Delta}^{G^{-1}}(x) \le \frac{\kappa}{s}\frac{(\ln\frac{1}{1-x})^{\frac{1}{s}}}{\sqrt{B}} M_{\eta,x}\}$, $\Omega_2 := \{\hat{\Delta}^{G^{-1}}(0) \le \kappa\}$ and $\Omega_3 := \{\hat{\Delta}^{G^{-1}}(1) \ge \frac{\kappa}{s}\frac{(2\ln B)^{\frac{1}{s}-1}}{B^2}\}$. Let $B_x^3 > 2$ be large enough such that $2\ln B > \ln(1/(1-x))$ for all $B > B_x^3$. Given any $B > B_x^3$, it can be shown that for all sufficiently small $s < s_{B,x}^3 := \frac{\ln(2\ln B) - \ln\frac{1}{1-x}}{\frac{3}{2}\ln B - \ln(2\ln B)}$ and for all $s < s_B^4$,

$$\frac{\kappa}{s}\frac{(2\ln B)^{\frac{1}{s}-1}}{B^2} > \frac{\kappa}{s}\frac{(\ln\frac{1}{1-x})^{\frac{1}{s}}}{\sqrt{B}} M_{\eta,x} \quad \text{and} \quad \frac{\kappa}{s}\frac{(2\ln B)^{\frac{1}{s}-1}}{B^2} > \kappa,$$

respectively. Thus, for all $B > \max\{B_{\eta,x}^1, B_x^3\}$ and all $s < \min\{s_{B,x}^3, s_B^4\}$, we have

$$\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(1) > \hat{\Delta}^{G^{-1}}(x)\big) \ge \sum_{i=1}^{3} \mathbb{P}(\Omega_i) - 2 \ge (1-\frac{1}{B^2})^B - e^{-B} - \frac{\eta}{3},$$

where the last inequality follows from Lemma 13(c)(d). To ensure that this probability is no less than $1-\eta$, we pick $B_\eta^4$ large enough such $(1-\frac{1}{B^2})^B > 1-\eta/3$ and $e^{-B} < \eta/3$ for all $B > B_\eta^4$. Setting $B_{\eta,x} = \max\{B_{\eta,x}^1, B_x^3, B_\eta^4\}$ and $s_{B,x} = \min\{s_{B,x}^3, s_B^4\}$ completes the proof for Lemma 2(ii). □

**Proof of Proposition 3.** We first prove part (i). The fact that $\mathbb{P}(\hat{\Delta}^{G^{-1}}(0) \ge \epsilon) \le \Psi_{f2}(B;0,\epsilon) = 2B\exp(-(B-1)(\epsilon/\kappa)^s)$ follows from the second display of (24) with $\delta = 0$ and that $G(x) = 1-\exp(x/\kappa)^w$. To show that $\hat{\Delta}^{G^{-1}}(0) = \tilde{O}_P(B^{-1/s})$, for each $\eta > 0$, observe that

$$\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(0) \ge M_\eta(B/\ln B)^{-1/s}\big) \le 2B\exp\Big(-(B-1)\frac{1}{\kappa^s}M_\eta^s\frac{1}{B}\ln B\Big) \overset{(a)}{\le} 2\times 2^{-\frac{1}{2}\frac{M_\eta^s}{\kappa^s}+1} \overset{(b)}{\le} \eta,$$

where (a) holds for all $B \ge 2$ such that $(B-1)/B \ge 1/2$ and for all sufficiently large $M_\eta$ such that $-\frac{1}{2}\frac{M_\eta^s}{\kappa^s}+1 < 0$; and (b) holds for all sufficiently large $M_\eta$. We next show that $\liminf_B \mathbb{P}(\hat{\Delta}^{G^{-1}}(0) > B^{-1/s}M) > 0$ for any $M > 0$. Note that $\mathbb{P}(r_1 \le \kappa) = 1-[1-G(\kappa)]^B = 1-e^{-B}$ and $\mathbb{P}(r_1 \ge MeB^{-\frac{1}{s}}) = [1-G(MeB^{-\frac{1}{s}})]^B = e^{-(Me/\kappa)^s}$. Given $r_1 \le \kappa$ and $r_1 \ge MeB^{-\frac{1}{s}}$, we have

$$\hat{\Delta}^{G^{-1}}(0) = \int_0^{r_1} 1-G(u)du \ge \int_0^{r_1} 1-G(r_1)du = e^{-(\frac{r_1}{\kappa})^s}r_1 \ge e^{-1}r_1 \ge MB^{-\frac{1}{s}}.$$

Therefore, by using $\mathbb{P}(\Omega_1 \cap \Omega_2) \ge \mathbb{P}(\Omega_1) + \mathbb{P}(\Omega_2) - 1$, we have for large enough $B$,

$$\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(0) \ge MB^{-\frac{1}{s}}\big) \ge \mathbb{P}\big(\{r_1 \le \kappa\} \cap \{r_1 \ge MeB^{-\frac{1}{s}}\}\big) \ge e^{-(Me/\kappa)^s} - e^{-B} > 0.$$

Next, we prove part (ii). We have $\mathbb{P}(\hat{\Delta}^{G^{-1}}(1) \ge \epsilon) = \mathbb{P}(\int_{r_B}^{\infty} 1-G(u)du \ge \epsilon)$. Then, the fact that $\mathbb{P}(\hat{\Delta}^{G^{-1}}(1) \ge \epsilon) = \Psi_{f1}(B;0,\epsilon) = [G(y(\epsilon))]^B$ follows from (26) with $\delta = 0$ and $f = 1-G$. To show that $\ln[\frac{1}{G(y(\hat{\Delta}^{G^{-1}}(1)))}] = \tilde{O}_P(B^{-1})$, we use the strict monotonicity of the functions $x \mapsto \ln(x)$, $x \mapsto G(x)$ and $x \mapsto y(x)$. For each $\eta > 0$, observe that

$$\begin{aligned}\mathbb{P}\Big(\ln\big[\frac{1}{G(y(\hat{\Delta}^{G^{-1}}(1)))}\big] > \frac{M_\eta}{B}\Big) &= \mathbb{P}\Big(\hat{\Delta}^{G^{-1}}(1) > y^{-1}\Big(G^{-1}\big(e^{-\frac{M_\eta}{B}}\big)\Big)\Big) \\ &\overset{(a)}{\le} \Big[G\Big(y\big(y^{-1}(G^{-1}(e^{-\frac{M_\eta}{B}}))\big)\Big)\Big]^B \\ &\overset{(b)}{\le} \eta,\end{aligned}$$

where (a) follows from $\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(1)\geq\epsilon\big)\leq\big[G(y_f(\epsilon))\big]^B$, and (b) holds for sufficiently large $M_\eta$.

Lastly, we prove part (iii). Assuming that $p\in[\delta,1-\delta]$, we have $\mathbb{P}(\hat{\Delta}^{G^{-1}}(x)\geq\epsilon)\leq\mathbb{P}(\sup_{p\in[\delta,1-\delta]}\hat{\Delta}^{G^{-1}}(x)\geq\epsilon)$. Then, the fact that $\mathbb{P}(\hat{\Delta}^{G^{-1}}(x)\geq\epsilon)\leq\Psi_{f3}(B;\delta,\epsilon)=2\exp\big(-\frac{1}{2}B[(\epsilon\rho(\delta))\wedge\delta]^2\big)$ follows from (30) with $f=1-G$ and $C_{f2}=1$. To show that $\hat{\Delta}^{G^{-1}}(x)=O_P(B^{-1/2})$, observe that for any $\eta>0$, $\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(x)\geq M_\eta B^{-\frac{1}{2}}\big)\leq 2\exp\Big(-\frac{1}{2}B\big[\big(M_\eta B^{-\frac{1}{2}}\rho(\delta)\big)\wedge\delta\big]^2\Big)\leq 2\exp\Big(-\frac{1}{2}M_\eta^2\rho^2(\delta)\Big)\leq\eta$, where the second inequality holds for all sufficiently large $B$ such that $M_\eta B^{-\frac{1}{2}}\leq\delta/\rho(\delta)$, and the last inequality holds for all sufficiently large $M_\eta$. Finally, the fact that $\liminf_B\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(x)\geq B^{-\frac{1}{2}}M\big)>0$ for any $M>0$ follows from Proposition 2, because the Weibull distribution has a density that is continuous and positive. $\square$

**Proof of Lemma 4.** We start by deriving a key inequality (60). Let $\Gamma(x):=\int_0^\infty u^{x-1}e^{-u}du$ denote the usual Gamma function, which satisfies $\Gamma(n+1)=n!$ for $n\in\mathbb{N}$. By the Stirling's bound [1], we have for all $x\geq 0$, $\sqrt{2\pi x}\left(\frac{x}{e}\right)^x\leq\Gamma(x+1)\leq\sqrt{2\pi x}\left(\frac{x}{e}\right)^x e^{\frac{1}{12x}}$. It follows that

$$
\begin{aligned}
\frac{\Gamma(B+1)}{\Gamma(B(1-\delta)+1)\Gamma(B\delta+1)} &\leq \frac{\sqrt{2\pi B}\left(\frac{B}{e}\right)^B e^{\frac{1}{12B}}}{\sqrt{2\pi(B-B\delta)}\left(\frac{B-B\delta}{e}\right)^{B-B\delta}\times\sqrt{2\pi B\delta}\left(\frac{B\delta}{e}\right)^{B\delta}}\\
&=\frac{e^{\frac{1}{12B}}}{\sqrt{2\pi B\delta}}\sqrt{1+\frac{1}{1/\delta-1}}\exp\left(B\ln(1+\frac{1}{1/\delta-1})\right)(\frac{1}{\delta}-1)^{B\delta}\\
&\overset{(a)}{\leq}\frac{e^{\frac{1}{12}}}{\sqrt{\pi B\delta}}\exp\left(\frac{B}{1/\delta-1}\right)\delta^{-B\delta}\overset{(b)}{\leq}\frac{1}{\sqrt{B\delta}}\exp\left(2B\delta\right)\delta^{-B\delta}.
\end{aligned}
\tag{60}
$$

Here, (a) uses the facts that $B\geq 1$, that $\sqrt{1+1/(1/\delta-1)}\leq\sqrt{2}$ for all $\delta\leq 1/2$, and that $\ln(1+x)\leq x$ for all $x\geq 0$; and (b) uses the facts that $B/(1/\delta-1)\leq 2B\delta$ for all $\delta\leq 1/2$ and that $e^{1/12}/\sqrt{\pi}\leq 1$.

By rewriting the factorials in the lemma using the $\Gamma(x)$, we have $\frac{B!}{(\lceil B(1-\delta)\rceil)!(B-\lceil B(1-\delta)\rceil)!}=\frac{\Gamma(B+1)}{\Gamma(\lceil B(1-\delta)\rceil+1)\Gamma(B-\lceil B(1-\delta)\rceil+1)}$. The right-hand side is upper bounded by $B\delta\frac{\Gamma(B+1)}{\Gamma(B(1-\delta)+1)\Gamma(B\delta+1)}$, where we use the facts that $B(1-\delta)\leq\lceil B(1-\delta)\rceil\leq B(1-\delta)+1$, that $\lceil B(1-\delta)\rceil\geq 1$ and $B-\lceil B(1-\delta)\rceil\geq 1$ for all $B\geq 2/\delta$, that $\Gamma(x)$ is increasing all for $x\geq 2$, and the identity that $\Gamma(x+1)=x\Gamma(x)$ with $x=B\delta$. Furthermore, $B\delta\frac{\Gamma(B+1)}{\Gamma(B(1-\delta)+1)\Gamma(B\delta+1)}\leq\sqrt{B\delta}\exp\left(2B\delta\right)\delta^{-B\delta}$ due to (60). This completes the first part of the proof.

Similarly, we have $(B\delta+2)\frac{B!}{(\lceil B\delta\rceil)!(B-\lceil B\delta\rceil)!}=(B\delta+2)\frac{\Gamma(B+1)}{\Gamma(\lceil B\delta\rceil+1)\Gamma(B-\lceil B\delta\rceil+1)}$. The right-hand side is upper bounded by $(B\delta+2)(B-B\delta)\frac{\Gamma(B+1)}{\Gamma(B\delta+1)\Gamma(B(1-\delta)+1)}$, where we use the facts that $B\delta\leq\lceil B\delta\rceil\leq B\delta+1$, the monotonicity of $\Gamma(x)$, and the identity that $\Gamma(x+1)=x\Gamma(x)$ with $x=B(1-\delta)$. Furthermore, $(B\delta+2)(B-B\delta)\frac{\Gamma(B+1)}{\Gamma(B\delta+1)\Gamma(B(1-\delta)+1)}\leq 2(B-B\delta)\sqrt{B\delta}\exp(2B\delta)\delta^{-B\delta}$ as a result of (60) and the fact that $B\delta+2\leq 2B\delta$ for all $B\geq 2/\delta$. $\square$

**Proof of Lemma 3.** Take $x=\hat{G}^{-1}(p)$ for any $p\in[0,1]$. We have by assumption that $|p-G(\hat{G}^{-1}(p))|\leq\epsilon/2$, that is, $p-\epsilon/2\leq G(\hat{G}^{-1}(p))\leq p+\epsilon/2$. By the monotonicity of $G^{-1}$, it follows that $G^{-1}(p-\epsilon/2)\leq\hat{G}^{-1}(p)\leq G^{-1}(p+\epsilon/2)$ for all $p\in[\epsilon/2+\delta/2,1-\epsilon/2-\delta/2]$, where $\delta\in(0,2-\epsilon)$. The monotonicity also implies that $G^{-1}(p-\epsilon/2)\leq G^{-1}(p)\leq G^{-1}(p+\epsilon/2)$. Therefore,

$$
|\hat{G}^{-1}(p)-G^{-1}(p)|\leq G^{-1}(p+\epsilon/2)-G^{-1}(p-\epsilon/2)
$$

for all $p\in[\epsilon/2+\delta/2,1-\epsilon/2-\delta/2]$. By the inverse function rule, we have $[G^{-1}(p)]'=1/g(G^{-1}(p))$. Hence, Assumption 1 implies that the function $G^{-1}(p)$ is $1/\rho(\delta)$-Lipschitz continuous over the interval $[\delta/2,1-\delta/2]$. As a result, the right hand side of the above display is further upper bounded by $\epsilon/\rho(\delta)$ for all $p\in[\epsilon/2+\delta/2,1-\epsilon/2-\delta/2]$, which completes the proof. $\square$

**Proof of Lemma 13.** We first prove (a). By standard central limit theorem on quantile estimation [52], $\sqrt{B}(\hat{G}^{-1}(p)-G^{-1}(p))\to\text{Normal}\Big(0,\frac{x(1-x)}{g^2(G^{-1}(p))}\Big)$ as $B\to\infty$ for any $x\in(0,1)$. Since $g(G^{-1}(p))=\frac{s}{\kappa}(\ln\frac{1}{1-x})^{1-\frac{1}{s}}(1-x)$ for Weibull distribution, as $B\to\infty$, we have $\sqrt{B}\frac{s}{\kappa}(\ln\frac{1}{1-x})^{-\frac{1}{s}}(\hat{G}^{-1}(p)-G^{-1}(p))\to\text{Normal}\Big(0,\frac{x}{(\ln\frac{1}{1-x})^2(1-x)}\Big)$. Therefore, $\big|\hat{G}^{-1}(p)-G^{-1}(p)\big|=O_p(\frac{\kappa}{\sqrt{B}s}(\ln\frac{1}{1-x})^{\frac{1}{s}})$ by the continuous mapping theorem applied to the function $x \mapsto |x|$. This implies (a) because $\hat{\Delta}^{G^{-1}}(x) = \left|\int_{\hat{G}^{-1}(p)}^{G^{-1}(p)} 1 - G(u)du\right| \leq \left|\hat{G}^{-1}(p) - G^{-1}(p)\right|$.

Now we prove (d). Because $G(x) = 1 - e^{-(x/\kappa)^s}$, we have $\mathbb{P}(r_B \leq (2\ln B)^{\frac{1}{s}}\kappa) = [G((2\ln B)^{\frac{1}{s}}\kappa)]^B = \left(1 - \frac{1}{B^2}\right)^B$. Given that $r_B \leq (2\ln B)^{\frac{1}{s}}\kappa$, we have $\hat{\Delta}^{G^{-1}}(1) = \int_{r_B}^{\infty} 1 - G(u)du \geq \int_{(2\ln B)^{\frac{1}{s}}\kappa}^{\infty} 1 - G(u)du$. Moreover,

$$\int_{(2\ln B)^{\frac{1}{s}}\kappa}^{\infty} 1 - G(u)du = \int_{(2\ln B)^{\frac{1}{s}}\kappa}^{\infty} e^{-(\frac{u}{\kappa})^s}du = \frac{\kappa}{s}\int_{2\ln B}^{\infty} e^{-t}t^{\frac{1}{s}-1}dt > \frac{\kappa}{s}\frac{(2\ln B)^{\frac{1}{s}-1}}{B^2}$$

where the second equality follows from a change of variable $t = (u/\kappa)^s$, and the last inequality is due to $\int_x^{\infty} e^{-t}t^{1/s-1}dt > e^{-x}x^{a-1}$ for all $x > 0$ and $s < 1$ [2].

We now prove (e). Note that $\mathbb{P}(r_B \geq \kappa) = 1 - \mathbb{P}(r_B \leq \kappa) = 1 - [G(\kappa)]^B = [1 - e^{-1}]^B$. Given $r_B \geq \kappa$, we have $\hat{\Delta}^{G^{-1}}(1) = \int_{r_B}^{\infty} 1 - G(u)du \leq \int_{\kappa}^{\infty} 1 - G(u)du$. Moreover,

$$\int_{\kappa}^{\infty} 1 - G(u)du = \frac{\kappa}{s}\int_1^{\infty} e^{-t}t^{\frac{1}{s}-1}dt \leq \frac{\kappa}{s}\int_1^{\infty} e^{-t} = \frac{\kappa}{se},$$

where the first equality follows from $G(x) = 1 - e^{-(x/\kappa)^s}$ and a change of variable $t = (u/\kappa)^s$; and the inequality is due to $t^{\frac{1}{s}-1} \leq 1$ for all $t \geq 1$ and $s \geq 1$.

Now we prove (b) and (c). Note that $\mathbb{P}(r_1 \leq \kappa) = 1 - \mathbb{P}(r_1 \geq \kappa) = 1 - [1 - G(\kappa)]^B = 1 - e^{-B}$ and $\mathbb{P}(r_1 \geq \frac{\kappa}{2}B^{-\frac{1}{s}}) = [1 - G(\frac{\kappa}{2}B^{-\frac{1}{s}})]^B = e^{-(\frac{1}{2})^s}$. Given $r_1 \leq \kappa$ and $r_1 \geq \frac{\kappa}{2}B^{-\frac{1}{s}}$, we have

$$\hat{\Delta}^{G^{-1}}(0) = \int_0^{r_1} 1 - G(u)du \geq \int_0^{r_1} 1 - G(r_1)du = e^{-(\frac{r_1}{\kappa})^s}r_1 \geq e^{-1}r_1 \geq \frac{\kappa}{2e}B^{-\frac{1}{s}}.$$

Therefore, by using $\mathbb{P}(\Omega_1 \cap \Omega_2) \geq \mathbb{P}(\Omega_1) + \mathbb{P}(\Omega_2) - 1$, we have $\mathbb{P}\big(\hat{\Delta}^{G^{-1}}(0) \geq \frac{\kappa}{2e}B^{-\frac{1}{s}}\big) \geq \mathbb{P}\big(\{r_1 \leq \kappa\} \cap \{r_1 \geq \frac{\kappa}{2}B^{-\frac{1}{s}}\}\big) \geq \mathbb{P}\big(\{r_1 \leq \kappa\}\big) + \mathbb{P}\big(\{r_1 \geq \frac{\kappa}{2}B^{-\frac{1}{s}}\}\big) - 1 = e^{-(\frac{1}{2})^s} - e^{-B}$. This completes the proof for (b). To prove (c), note that given $r_1 \leq \kappa$, we have $\hat{\Delta}^{G^{-1}}(0) = \int_0^{r_1} 1 - G(u)du \leq \int_0^{r_1} 1du = r_1 \leq \kappa$. Therefore, $\mathbb{P}(\hat{\Delta}^{G^{-1}}(0) \leq \kappa) \geq \mathbb{P}(r_1 \leq \kappa) = 1 - e^{-B}$. □

## C. Proofs for Results in Sections 6 and 7

**Proof of Corollary 2.** By the first display of Theorem 3 with $\delta_w = \delta$ for all $w \in \mathbb{W}$ and the fact that the functions $\Phi_w(B;\epsilon)$ and $\Psi_{wi}(B;\delta,\epsilon)$ are (ultimately) decreasing in $B$, we have

$$\frac{\ln \mathbb{P}\left(\left|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}^{\text{KA}})\right| \geq \epsilon\right)}{B^G} \leq \frac{\ln\left(\sum_{w\in\mathbb{W}}\Phi_w\left(B^G;\nu_1\epsilon\right) + \sum_{w\in\mathbb{W}}\sum_{i\in\{1,2,3\}}\Psi_{wi}\left(B^G;\delta,\nu_1\epsilon\right)\right)}{B^G}.$$

Letting $C := 4(J+K)$, the right-hand side of the above display is no greater than

$$\max_{w\in\mathbb{W}} \frac{\max\left\{\ln[C\Phi_w(B^G;\nu_1\epsilon)], \ln[C\Psi_{w1}(B^G;\nu_1\epsilon)], \ln[C\Psi_{w2}(B^G;\nu_1\epsilon)], \ln[C\Psi_{w3}(B^G;\nu_1\epsilon)]\right\}}{B^G} \tag{61}$$

due to the inequality that $\ln(a_1 + a_2 + \cdots + a_n) \leq \max\{\ln(na_1), \ln(na_2), \ldots, \ln(na_n)\}$. By the definitions of $\Phi_w$ in (13), $\Psi_{w3}$ in (22), and the asymptotics of $\Psi_{w1}$ and $\Psi_{w2}$ in (20) with $f = 1 - G_w$, we have

$$\begin{aligned}
&\lim_{B\to\infty}\frac{\ln\Psi_{w1}\left(B;\delta,\nu_1\epsilon\right)}{B} \leq -\frac{1}{2}\Upsilon_1(\nu_1\epsilon), \forall w \in \mathbb{W};\\
&\lim_{B\to\infty}\frac{\ln\Psi_{w2}\left(B;\delta,\nu_1\epsilon\right)}{B} \leq -\frac{1}{2}\Upsilon_2(\nu_1\epsilon), \forall w \in \mathbb{W};\\
&\lim_{B\to\infty}\frac{\ln\Psi_{w3}\left(B;\delta,\nu_1\epsilon\right)}{B} \leq -\frac{1}{2}\min\left\{\min_{w\in\mathbb{W}}\rho_w^2(\delta)\nu_1^2\epsilon^2, \delta^2\right\}, \forall w \in \mathbb{W};\\
&\lim_{B\to\infty}\frac{\ln\Phi_w\left(B;\nu_1\epsilon\right)}{B} \leq -\frac{1}{2}a_2\nu_1^2\epsilon^2, \forall w \in \mathbb{W}.
\end{aligned} \tag{62}$$

for all sufficiently small $\delta > 0$. Then, by taking the limit $B^G \to \infty$ in (61) and using (62), we establish the first display of Corollary 2.

To establish the second display of Corollary 2, one can use the last display of Theorem 3 to obtain

$$\frac{\ln \mathbb{P}\left(\left|\boldsymbol{m}^\star - \hat{\boldsymbol{m}}^{\mathrm{KA}}\right| \geq \epsilon\right)}{B^G} \leq \frac{\ln\left(\sum_{w\in\mathbb{W}} \Phi_w\left(B^G; \nu_1\chi(\epsilon)\right) + \sum_{w\in\mathbb{W}}\sum_{i\in\{1,2,3\}} \Psi_{wi}\left(B^G; \delta, \nu_1\chi(\epsilon)\right)\right)}{B^G},$$

and then repeating the above arguments with $\chi(\epsilon)$ in place of $\epsilon$. □

**Proof of Corollary 3.** By the first display of Theorem 4 with $\delta_w = \delta$ for all $w \in \mathbb{W}$ and the fact that the functions $\Phi_w(B;\epsilon)$, $\Psi_{wi}(B;\delta,\epsilon)$ and $\Lambda_w(B,\epsilon)$ are (ultimately) decreasing in $B$, we have

$$\frac{\ln \mathbb{P}\left(\left|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}})\right| \geq \epsilon\right)}{B^{G,\lambda}} \leq \frac{\ln\left(\sum_{w\in\mathbb{W}}\left(\Phi_w\left(B^{G,\lambda}; \frac{\nu_1\epsilon}{6}\right) + \sum_{i\in\{1,2,3\}} \Psi_{wi}\left(B^{G,\lambda}; \delta, \frac{\nu_1\epsilon}{6}\right) + \Lambda_w(B^{G,\lambda}, \underline{\epsilon})\right)\right)}{B^{G,\lambda}}.$$

Letting $D := 5(J+K)$, the right-hand side of the above display is no greater than

$$\max_{w\in\mathbb{W}} \frac{\max\left\{\ln[D\Phi_w(B^{G,\lambda}; \frac{\nu_1\epsilon}{6})], \ln[D\Psi_{w1}(B^{G,\lambda}; \frac{\nu_1\epsilon}{6})], \ln[D\Psi_{w2}(B^{G,\lambda}; \frac{\nu_1\epsilon}{6})], \ln[D\Psi_{w3}(B^{G,\lambda}; \frac{\nu_1\epsilon}{6})], \ln[D\Lambda_w(B^{G,\lambda}; \underline{\epsilon})]\right\}}{B^{G,\lambda}} \tag{63}$$

because of the same inequality that led to (61). By the definition of $\Lambda_w$ in (33), we have

$$\lim_{B\to\infty} \frac{\ln \Lambda_w\left(B; \underline{\epsilon}\right)}{B} \leq -\frac{1}{2}\min\{\xi\underline{\epsilon}, \xi^2\underline{\epsilon}^2\}, \forall w \in \mathbb{W}. \tag{64}$$

By taking the limit $B^{G,\lambda} \to \infty$ in (63) and using (62) with $\epsilon \leftarrow \epsilon/6$, and (64), the proof is complete. □

**Proof of Corollary 4.** By the first display of Theorem 5, we have

$$\frac{\ln \mathbb{P}\left(\left|\boldsymbol{m}^\star - \hat{\boldsymbol{m}}\right| \geq \epsilon\right)}{B^{G,\lambda}} \leq \frac{\ln\left(\sum_{w\in\mathbb{W}}\left(\Phi_w\left(B^{G,\lambda}; \frac{\nu_1\chi(\frac{\epsilon}{2})}{12}\right) + \sum_{i\in\{1,2,3\}} \Psi_{wi}\left(B^{G,\lambda}; \delta, \frac{\nu_1\chi(\frac{\epsilon}{2})}{12}\right) + \Lambda_w(B^{G,\lambda}, \underline{\epsilon})\right)\right)}{B^{G,\lambda}}.$$

Then, one can complete the proof by following the same line of arguments in the proof of Corollary 3 upon replacing $\frac{\nu_1\epsilon}{6}$ therein by $\frac{\nu_1\chi(\frac{\epsilon}{2})}{12}$. □

**Proof of Lemma 5.** By the definition of $\hat{\alpha}_i^\lambda$ in (32), the event

$$\Omega_1 := \left\{\left|\frac{\lambda_w}{\hat{\lambda}_w} - 1\right| \leq \epsilon_1, \left|\frac{\hat{\lambda}_w}{\lambda_w} - 1\right| \leq \epsilon_2, \left|\frac{\hat{\lambda}_w}{\lambda_w} - 1\right| \leq \epsilon_3 \quad \text{for all} \quad w \in \mathbb{W}\right\}$$

is a subset of the event $\left\{\left|\hat{\alpha}_1^\lambda - 1\right| \leq \epsilon_1, \left|\hat{\alpha}_2^\lambda - 1\right| \leq \epsilon_2, \left|\hat{\alpha}_3^\lambda - 1\right| \leq \epsilon_3\right\}$. Therefore, it suffices to find a lower bound for the probability of $\Omega_1$ happening. For any $\epsilon$, we have the following event relations:

$$\begin{aligned}\left\{\left|\frac{\hat{\lambda}_w}{\lambda_w} - 1\right| \leq \epsilon\right\} = \left\{[1-\epsilon]^+ \leq \frac{\hat{\lambda}_w}{\lambda_w} \leq 1+\epsilon\right\} &= \left\{\frac{-\epsilon}{1+\epsilon} \leq \frac{\lambda_w}{\hat{\lambda}_w} - 1 \leq \frac{\epsilon}{[1-\epsilon]^+}\right\} \\ &\supseteq \left\{\left|\frac{\hat{\lambda}_w}{\lambda_w} - 1\right| \leq \min\left(\frac{\epsilon}{1+\epsilon}, \frac{\epsilon}{[1-\epsilon]^+}\right)\right\}.\end{aligned}$$

Therefore, we manage to find a subset of $\Omega_1$ as follows:

$$\Omega_1 \supseteq \left\{\left|\frac{\lambda_w}{\hat{\lambda}_w} - 1\right| \leq \min\left(\epsilon_1, \frac{\epsilon_2}{1+\epsilon_2}, \frac{\epsilon_2}{[1-\epsilon_2]^+}, \frac{\epsilon_3}{1+\epsilon_3}, \frac{\epsilon_3}{[1-\epsilon_3]^+}\right) \quad \text{for all} \quad w \in \mathbb{W}\right\}$$

Letting $\tilde{\epsilon} = \zeta(\epsilon_1, \epsilon_2, \epsilon_3)$ with the function $\zeta$ as defined in (34), it follows that

$$\begin{aligned}\mathbb{P}\left(\left|\hat{\alpha}_1^\lambda - 1\right| \leq \epsilon_1, \left|\hat{\alpha}_2^\lambda - 1\right| \leq \epsilon_2, \left|\hat{\alpha}_3^\lambda - 1\right| \leq \epsilon_3\right) \geq \mathbb{P}(\Omega_1) &\geq \mathbb{P}\left(\left\{\left|\frac{\lambda_w}{\hat{\lambda}_w} - 1\right| \leq \tilde{\epsilon} \quad \text{for all} \quad w \in \mathbb{W}\right\}\right) \\ &\geq 1 - \sum_{w\in\mathbb{W}} \mathbb{P}\left(\left|\frac{\lambda_w}{\hat{\lambda}_w} - 1\right| \geq \tilde{\epsilon}\right).\end{aligned}$$

Then, the proof can be completed by upper bounding $\mathbb{P}\left(\left|\frac{\lambda_w}{\hat\lambda_w}-1\right|\geq\tilde\epsilon\right)$ for each $w\in\mathbb{W}$ as follows:

$$\mathbb{P}\left(\left|\frac{\lambda_w}{\hat\lambda_w}-1\right|\geq\tilde\epsilon\right)=\mathbb{P}\left(\left|\frac{1}{\hat\lambda_w}-\frac{1}{\lambda_w}\right|\geq\frac{\tilde\epsilon}{\lambda_w}\right)\overset{(a)}{=}\mathbb{P}\left(\left|\frac{\sum_{b=1}^{B_w^\lambda}r_{wb}}{B_w^\lambda}-\frac{1}{\lambda_w}\right|\geq\frac{\tilde\epsilon}{\lambda_w}\right)\overset{(b)}{\leq}\Lambda_w(B_w^\lambda;\tilde\epsilon),$$

where (a) is by the definition of $\hat\lambda_w$ in (9); and (b) follows from the Berstein's inequality for sub-exponential distribution (see, e.g., Chapter 2.8 of [3]). □

**Proof of Proposition 4.** We first prove claim 1. Without loss of generality, we focus on a fixed $w=j\in\mathbb{J}$. Consider the function $q_j(p)=\lambda_j\int_0^{G_j^{-1}(p)}(1-G_j(u))du$ for $p\in[0,1]$. By the Leibniz integral rule, we have $\frac{dq_j(p)}{dx}=\frac{\lambda_j}{h_j(G_j^{-1}(p))}\leq L\lambda_j$ for all $p\in[0,1]$, where the inequality is due to Assumption 3. Therefore, $q_j(p)$ is $L\lambda_j$-Lipschitz continuous over $p\in[0,1]$, establishing claim 1.

Next, we prove claim 2. Note that for $\boldsymbol{m}\in\mathbb{R}_+^{JK}$, we have $F(\boldsymbol{m})=\sum_{j,k}v_{jk}m_{jk}-\sum_{w\in\mathbb{W}}c_wq_w(p_w(\boldsymbol{m}))$, where $p_w(\boldsymbol{m})=p_w(\boldsymbol{m};\lambda_w)$ for $p_w(\boldsymbol{m};\lambda)$ defined in (8). Consider any $\boldsymbol{m},\boldsymbol{m}'\in\mathbb{R}_+^{JK}$. One can show that $|p_w(\boldsymbol{m})-p_w(\boldsymbol{m}')|\leq|\boldsymbol{m}-\boldsymbol{m}'|/\lambda_w$ for all $w\in\mathbb{W}$. It follows that

$$\begin{aligned}\left|F(\boldsymbol{m})-F(\boldsymbol{m}')\right|&\leq\max_{(j,k)\in\mathcal{E}}v_{jk}|\boldsymbol{m}-\boldsymbol{m}'|+\sum_{w\in\mathbb{W}}c_w\left|q_w(p_w(\boldsymbol{m}))-q_w(p_w(\boldsymbol{m}'))\right|\\&\leq\max_{(j,k)\in\mathcal{E}}v_{jk}|\boldsymbol{m}-\boldsymbol{m}'|+\sum_{w\in\mathbb{W}}c_wL\lambda_w|\boldsymbol{m}-\boldsymbol{m}'|/\lambda_w=L_F|\boldsymbol{m}-\boldsymbol{m}'|.\end{aligned}$$

where the second inequality is due to claim 1, and the last equality is by our definition of $L_F$. □

**Proof of Lemma 6.** By telescoping $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})$ and $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\hat F(\boldsymbol{m})$, we have

$$\begin{aligned}&\left|F(\boldsymbol{m}^\star)-F(\hat{\boldsymbol{m}})\right|\\&\leq\left|\sup_{\boldsymbol{m}\in\mathbb{M}}F(\boldsymbol{m})-\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})\right|+\left|\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})-\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\hat F(\boldsymbol{m})\right|+\left|\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\hat F(\boldsymbol{m})-F(\hat{\boldsymbol{m}})\right|\\&\leq\left|\sup_{\boldsymbol{m}\in\mathbb{M}}F(\boldsymbol{m})-\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})\right|+\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\left|F(\boldsymbol{m})-\hat F(\boldsymbol{m})\right|+\left|\hat F(\hat{\boldsymbol{m}})-F(\hat{\boldsymbol{m}})\right|\\&\leq\left|\sup_{\boldsymbol{m}\in\mathbb{M}}F(\boldsymbol{m})-\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})\right|+2\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\left|F(\boldsymbol{m})-\hat F(\boldsymbol{m})\right|,\end{aligned}$$

where the second inequality follows from the facts that $|\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})-\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\hat F(\boldsymbol{m})|\leq\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}|F(\boldsymbol{m})-\hat F(\boldsymbol{m})|$ and that $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\hat F(\boldsymbol{m})=\hat F(\hat{\boldsymbol{m}})$. □

**Proof of Lemma 7.** Recall that $\boldsymbol{m}^\star$ is an solution of $\sup_{\boldsymbol{m}\in\mathbb{M}}F(\boldsymbol{m})$. Let $\tilde{\boldsymbol{m}}$ be an solution to $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}F(\boldsymbol{m})$. Then it suffices to bound $|F(\boldsymbol{m}^\star)-F(\tilde{\boldsymbol{m}})|$. By our definitions of $\hat\alpha_i^\lambda$ in (32) and the feasible regions $\hat{\mathbb{M}}$ and $\mathbb{M}$ in (10) and (6), one can check that $\hat\alpha_1^\lambda\tilde{\boldsymbol{m}}\in\mathbb{M}$ and $\hat\alpha_2^\lambda\boldsymbol{m}^\star\in\hat{\mathbb{M}}$. Thus, by the optimality of $\boldsymbol{m}^\star$ and $\tilde{\boldsymbol{m}}$, we have $F(\boldsymbol{m}^\star)\geq F(\hat\alpha_1^\lambda\tilde{\boldsymbol{m}})$ and $F(\tilde{\boldsymbol{m}})\geq F(\hat\alpha_2^\lambda\boldsymbol{m}^\star)$. By the Lipschitz continuity of $F$, we have $F(\hat\alpha_1^\lambda\tilde{\boldsymbol{m}})\geq F(\tilde{\boldsymbol{m}})-L_F|\hat\alpha_1^\lambda-1||\tilde{\boldsymbol{m}}|$ and $F(\hat\alpha_2^\lambda\boldsymbol{m}^\star)\geq F(\boldsymbol{m}^\star)-L_F|\hat\alpha_2^\lambda-1||\boldsymbol{m}^\star|$. Combining these yields

$$\left|F(\tilde{\boldsymbol{m}})-F(\boldsymbol{m}^\star)\right|\leq L_F\max\left(|\hat\alpha_1^\lambda-1||\tilde{\boldsymbol{m}}|,|\hat\alpha_2^\lambda-1||\boldsymbol{m}^\star|\right).$$

Now, since $\boldsymbol{m}^\star\in\mathbb{M}$, we have $|\boldsymbol{m}^\star|\leq\sum_{w\in\mathbb{W}}\lambda_w$. Similarly, since $\tilde{\boldsymbol{m}}\in\hat{\mathbb{M}}$, it holds that $|\tilde{\boldsymbol{m}}|\leq\sum_{w\in\mathbb{W}}\hat\lambda_w\leq\hat\alpha_3^\lambda\sum_{w\in\mathbb{W}}\lambda_w$, where the second inequality follows from our definition of $\hat\alpha_3^\lambda$. By replacing $|\boldsymbol{m}^\star|$ and $|\tilde{\boldsymbol{m}}|$ in the above display with theses upper bounds, the proof is complete. □

**Proof of Lemma 8.** First, claim 1 follows from comparing the expressions of the ground-truth $F$ in (6) and the data-driven $\hat F$ in (10).

We next prove claim 2. For each $\boldsymbol{m}\in\hat{\mathbb{M}}$, it holds that

$$\begin{aligned}\left|\hat q_w\left(\hat p_w(\boldsymbol{m})\right)-q_w\left(p_w(\boldsymbol{m})\right)\right|&=\left|\hat q_w\left(\hat p_w(\boldsymbol{m})\right)-q_w\left(\hat p_w(\boldsymbol{m})\right)+q_w\left(\hat p_w(\boldsymbol{m})\right)-q_w\left(p_w(\boldsymbol{m})\right)\right|\\&\overset{(a)}{\leq}\left|\hat q_w\left(\hat p_w(\boldsymbol{m})\right)-q_w\left(\hat p_w(\boldsymbol{m})\right)\right|+L\lambda_w\left|\hat p_w(\boldsymbol{m})-p_w(\boldsymbol{m})\right|\\&\overset{(b)}{\leq}\left|\hat q_w\left(\hat p_w(\boldsymbol{m})\right)-q_w\left(\hat p_w(\boldsymbol{m})\right)\right|+L\lambda_w\left|\frac{\hat\lambda_w}{\lambda_w}-1\right|,\end{aligned}\tag{65}$$

where (a) follows from the triangle inequality and the Lipschitz continuity of $q_w$ established in Proposition 4; and (b) is due to the fact that $|\hat{p}_j(\boldsymbol{m}) - p_j(\boldsymbol{m})| \leq |\sum_{k\in\mathbb{K}} m_{jk}/\lambda_j - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j| \leq |\hat{\lambda}_j/\lambda_j - 1|$, where the second inequality is because $\boldsymbol{m} \in \hat{\mathbb{M}}$ so that $\sum_{k\in\mathbb{K}} m_{jk} \leq \hat{\lambda}_j$ for all $j \in \mathbb{J}$, and that similarly, we have $|\hat{p}_k(\boldsymbol{m}) - p_k(\boldsymbol{m})| \leq |\hat{\lambda}_k/\lambda_k - 1|$ for all $k \in \mathbb{K}$.

We continue to upper bound the first term of the last display of (65) as follows:

$$\begin{aligned}\left|\hat{q}_w\left(\hat{p}_w(\boldsymbol{m})\right) - q_w\left(\hat{p}_w(\boldsymbol{m})\right)\right| &= \left|\hat{\lambda}_w\left[\frac{\hat{q}_w(\hat{p}_w(\boldsymbol{m}))}{\hat{\lambda}_w} - \frac{q_w(\hat{p}_w(\boldsymbol{m}))}{\lambda_w} - \frac{q_w(\hat{p}_w(\boldsymbol{m}))}{\lambda_w}\left(\frac{\lambda_w}{\hat{\lambda}_w} - 1\right)\right]\right| \\ &\overset{(a)}{\leq} \lambda_w\hat{\alpha}_3^\lambda\left|\frac{\hat{q}_w(\hat{p}_w(\boldsymbol{m}))}{\hat{\lambda}_w} - \frac{q_w(\hat{p}_w(\boldsymbol{m}))}{\lambda_w}\right| + \frac{1}{\theta_w}|\lambda_w - \hat{\lambda}_w| \\ &\overset{(b)}{\leq} \lambda_w\hat{\alpha}_3^\lambda\left(\hat{\Delta}_w^G + \hat{\Delta}_w^{G^{-1}}\right) + \lambda_w\frac{1}{\theta_w}\left|1 - \frac{\hat{\lambda}_w}{\lambda_w}\right|,\end{aligned} \tag{66}$$

where $(a)$ follows from the fact that $\hat{\lambda}_w \leq \lambda_w\hat{\alpha}_3^\lambda$ and $q_w(\hat{p}_w(\boldsymbol{m})) \leq \lambda_w\theta_w$; and (b) is obtained by replicating the arguments used to prove claim 3 of Theorem 1, specifically, (14) therein.

Putting together (65) and (66), we have

$$\begin{aligned}\frac{\left|\hat{q}_w\left(\hat{p}_w(\boldsymbol{m})\right) - q_w\left(p_w(\boldsymbol{m})\right)\right|}{\lambda_w} &\leq \hat{\alpha}_3^\lambda\left(\hat{\Delta}_w^G + \hat{\Delta}_w^{G^{-1}}\right) + \left(\frac{1}{\theta_w} + L\right)\left|\frac{\hat{\lambda}_w}{\lambda_w} - 1\right| \\ &\leq \hat{\alpha}_3^\lambda\left(\hat{\Delta}_w^G + \hat{\Delta}_w^{G^{-1}}\right) + \left(\frac{1}{\theta_w} + L\right)\max\left(|\hat{\alpha}_2^\lambda - 1|, |\hat{\alpha}_3^\lambda - 1|\right),\end{aligned}$$

where the second inequality follows from the fact that $|\hat{\lambda}_w/\lambda_w - 1| \leq \max(|\max_{w\in\mathbb{W}}\{\hat{\lambda}_w/\lambda_w\} - 1|, |\min_{w\in\mathbb{W}}\{\hat{\lambda}_w/\lambda_w\} - 1|)$ for all $w \in \mathbb{W}$. □

**Proof of Lemma 9.** By our definition of $\hat{\alpha}_i^\lambda$ in (32) and the fact that $\hat{\boldsymbol{m}} \in \hat{\mathbb{M}}$, we have $\hat{\alpha}_1^\lambda\hat{\boldsymbol{m}} \in \mathbb{M}$. Therefore, we have $|\hat{\boldsymbol{m}} - \hat{\boldsymbol{m}}^\perp| \leq |\hat{\boldsymbol{m}} - \hat{\alpha}_1^\lambda\hat{\boldsymbol{m}}| = |\hat{\boldsymbol{m}}||\hat{\alpha}_1^\lambda - 1|$ because the projection $\hat{\boldsymbol{m}}^\perp$ minimizes the distance. We also have $|\hat{\boldsymbol{m}}| \leq \sum_{w\in\mathbb{W}}\hat{\lambda}_w \leq \hat{\alpha}_3^\lambda\sum_{w\in\mathbb{W}}\lambda_w$, where the first inequality follows from the feasibility of $\hat{\boldsymbol{m}}$ in $\hat{\mathbb{M}}$, and the second inequality is due to (32). The proof is thus complete. □

## D. Omitted Proofs and Results for §8

**Proof of Proposition 5.** We prove by contradiction. Suppose all inequality constraints are non-binding under $\boldsymbol{m}^\star$. Then one can construct an alternative solution $\tilde{\boldsymbol{m}}$ that is feasible to $\mathbb{M}$, while satisfying $\tilde{m}_{jk} \geq m^\star_{jk}$ for all $j \in \mathbb{J}, k \in \mathbb{K}$, and $\tilde{m}_{j'k'} > m^\star_{j'k'}$ for some $j', k'$. It follows that $F(\tilde{\boldsymbol{m}}) > F(\boldsymbol{m}^\star)$ because both the queue-length functions $q_{j'}(\cdot)$ and $q_{k'}(\cdot)$ are strictly decreasing in $m_{j'k'}$, and because $v_{j'k'}\tilde{m}_{j'k'} > v_{j'k'}m^\star_{j'k'}$. This contradicts the optimality of $\boldsymbol{m}^\star$. □

**Proof of Lemma 10.** The proof is immediate by observing that the arguments used to prove Theorem 3 in [4] continue to be valid when the queue-length processes are induced by the matching process in either Definition 3 or Definition 4. □

**Evolution Equation for the Queue-Length Processes $Q_w(\cdot)$.** Recall that $A_w(\cdot)$ and $R_w(\cdot)$ denote the cumulative arrival and reneging processes for type-$w$ agents, respectively, and that $M_{jk}(\cdot)$ represents the cumulative matching process for type-$j$ and type-$k$ agents. Then, by the balance of flow, we have for any $t \geq 0$ that

$$Q_j(t) = Q_j(0) + A_j(t) - R_j(t) - \sum_{k\in\mathbb{K}} M_{jk}(t),\ \forall j \in \mathbb{J}$$
$$Q_k(t) = Q_k(0) + A_k(t) - R_k(t) - \sum_{j\in\mathbb{J}} M_{jk}(t),\ \forall k \in \mathbb{K}.$$

When the matching process $M_{jk}(\cdot)$ follows a discrete-review structure that does not perform any matching during the time interval $\big((i-1)l, il\big)$, where $l > 0$ is the review period, we further have for any $i = 1, 2, 3\ldots$:

$$Q_w\left(il-\right) = Q_w\left((i-1)l\right) + A_w\left(il\right) - A_w((i-1)l) - R_w\left(il\right) + R_w\left((i-1)l\right),\ \forall w \in \mathbb{W}. \tag{67}$$

**Sequence of Matching Processes Induced by the Proposed Estimate-then-Match Policy.** By the estimate-then-match policy defined in (41)–(43), the number of matches made in the $n$th system between type $j \in \mathbb{J}$ demand and type $k \in \mathbb{K}$ supply for $(j,k) = O_h, h = 1, \ldots, |\mathcal{E}|$ at review point $i \in \{1, \ldots, \lfloor t/l^n \rfloor\}$ is given by

$$\mathcal{M}_{ijk}^{\boldsymbol{m},n} := \min\left(\lfloor n m_{jk} l^n \rfloor,\ Q_j^n(il-) - \sum_{k':(j,k') \in \cup_{s=1}^{h-1} O_s} \mathcal{M}_{ijk'}^{\boldsymbol{m},n},\ Q_k^n(il-) - \sum_{j':(j',k) \in \cup_{s=1}^{h-1} O_s} \mathcal{M}_{ij'k}^{\boldsymbol{m},n}\right), \tag{68}$$

where an empty sum is defined to be zero (when $h = 1$). Note that the instantaneous matching rates $m_{jk}$ is scaled by $n$ in (68) because they remain feasible to the DDMP (10) when the estimated arrival rates are scaled by a factor of $n$. For all $(j,k) \in \mathcal{E}$, it follows that the cumulative number of matches until time $t \geq 0$ in the $n$th system is given by

$$M_{jk}^{\boldsymbol{m},n}(t) := \sum_{i=1}^{\lfloor t/l \rfloor} \mathcal{M}_{ijk}^{\boldsymbol{m},n}. \tag{69}$$

Moreover, it is clear that $M_{jk}^{\boldsymbol{m},n}(t) = 0$ for all $t \geq 0$ and $(j,k) \in \mathcal{E}^c$.

**Proofs of Proposition 6.** Our proof closely follow the proof framework used in [4] to establish the C-tightness of their matching-rate-based and priority-ordering policies (Theorems 4 and 5 therein). Although we do not find it necessary to write out the proofs of most results, we find value in stating the relevant lemmas and commenting on their functions in the proof. On a high level, the proof involves four basic steps: 1) establish lower and upper bounds on the number of arrivals during a review period, 2) use the arrival upper bound to establish an upper bound on the amount of reneging during a review period, 3) use the arrival lower bound and the reneging upper bound to derive a lower bound on the queue length right before each review point, and 4) uses the queue length lower bound to establish an lower bound on the number of matches made at each review point, and show that this lower bound approaches the target matching rates $\boldsymbol{m} \in \mathbb{M}$.

The following lemma, which is Lemma A.2 in [4], achieves step 1) above. This lemma holds when the inter-arrival distributions have finite 5th moment, which is implied by Assumption 2.

LEMMA 14 **(Lemma A.2 of [4]).** *Suppose the inter-arrival distributions have finite 5th moment. For any finite constant $a_1 > 0$, there exists finite constants $b > 0$ such that for any $T > 0$,*

$$\mathbb{P}\left(\sup_{i \in \{1,\ldots,\lfloor \frac{T}{l^n} \rfloor\}} \sup_{w \in \mathbb{W}} \left| A_w^n(il^n) - A_w^n((i-1)l^n) - n\lambda_w l^n \right| \leq a_1 n l^n \right) \geq 1 - bn^{-1/6}.$$

Let $\Omega_1^n$ denote the event that for all $i \in \left\{1, \ldots, \left\lfloor \frac{T}{l^n} \right\rfloor\right\}$ and $w \in \mathbb{W}$, $\left|A_w^n(il^n) - A_w^n((i-1)l^n) - n\lambda_w l^n\right| \leq a_1 n l^n$. Lemma 14 implies that $\lim_{n\to\infty} \mathbb{P}(\Omega_1^n) = 1$ because $nl^n \to \infty$.

The following lemma summarizes the first step of the proof of Theorem 4 in [4] to establish an upper bound on the number of reneging during a review period, achieving step 2) above.

LEMMA 15. *Suppose the inter-arrival distributions have finite 5th moment. For any $a_2 > 0$, let $\Omega_2^n$ denote the event that $R_w^n(il^n) - R_w^n((i-1)l^n) \leq a_2 n l^n$ for all $i \in \{1, \ldots, \lfloor T/l^n \rfloor\}, w \in \mathbb{W}$. Then, $\lim_{n\to\infty} \mathbb{P}(\Omega_2^n) = 1$.*

Let $\epsilon > 0$. In the following, we take $a_1 = a_2 = \frac{\epsilon}{4T}$ and consider only $\omega \in \Omega_1^n \cup \Omega_2^n$, where $\lim_{n\to\infty} \mathbb{P}(\Omega_1^n \cap \Omega_2^n) = 1$. Then, by the flow balance equation (67), for $i \in \{1, \ldots, \lfloor T/l^n \rfloor\}$ and $w \in \mathbb{W}$,

$$\begin{aligned} Q_w^n(il^n-) &= Q_w^n((i-1)l^n) + A_w^n(il^n) - A_w^n((i-1)l) - R_w^n(il^n) + R_w^n((i-1)l^n) \\ &\geq nl^n\lambda_w - a_1 n l^n - a_2 n l^n = nl^n(\lambda_w - \frac{\epsilon}{2T}), \end{aligned} \tag{70}$$

which gives a lower bound on the queue length, i.e., achieving step 3) above.

Fix $\boldsymbol{m} \in \mathbb{M}$. Since Proposition 6 trivially holds for all $(j,k) \in \mathcal{E}^c$, we focus on $(j,k) \in \mathcal{E}$ below. Suppose we can use (70) to show that for large enough $n \in \mathbb{N}$,

$$nl^n \left( m_{jk} - \frac{\epsilon}{2T} \right) \le \mathcal{M}_{ijk}^{\boldsymbol{m},n} \le nl^n m_{jk} \tag{71}$$

for each $(j,k) \in \mathcal{E}$ and $i \in \{1, \ldots, \lfloor T/l^n \rfloor\}$. Then, we have that for large enough $n \in \mathbb{N}$,

$$\frac{M_{jk}^{\boldsymbol{m},n}(t)}{n} = \frac{1}{n} \sum_{i=1}^{\lfloor t/l^n \rfloor} \mathcal{M}_{ijk}^{\boldsymbol{m},n} \ge \lfloor t/l^n \rfloor l^n \left( m_{jk} - \frac{\epsilon}{2T} \right) \ge (t/l^n - 1)\, l^n \left( m_{jk} - \frac{\epsilon}{2T} \right) \ge m_{jk} t - \epsilon,$$

where the last inequality is because $l^n \to 0$ as $n \to \infty$. On the other hand, we have for $n \in \mathbb{N}$

$$\frac{M_{jk}^{\boldsymbol{m},n}(t)}{n} = \frac{1}{n} \sum_{i=1}^{\lfloor t/l^n \rfloor} \mathcal{M}_{ijk}^{\boldsymbol{m},n} \le \lfloor t/l^n \rfloor l^n m_{jk} \le m_{jk} t.$$

Since $\epsilon$ can be arbitrarily small, the above two displays achieve step 4) above and complete the proof.

We next prove (71) to complete the proof. The second inequality of (71) follows directly from the capping mechanism of the proposed policy. Specifically, from (68), we have $\mathcal{M}_{ijk}^{\boldsymbol{m},n} \le \lfloor n m_{jk} l^n \rfloor \le nl^n m_{jk}$. For the first inequality of (71), we have for edge $(j,k) = O_h, h = 1, \ldots, |\mathcal{E}|$ and $i \in \{1, \ldots, \lfloor T/l^n \rfloor\}$,

$$\begin{aligned}
\mathcal{M}_{ijk}^{\boldsymbol{m},n} &= \min\left( \lfloor n m_{jk} l^n \rfloor,\ Q_j^n(il-) - \sum_{k':(j,k') \in \cup_{s=1}^{h-1} O_s} \mathcal{M}_{ijk'}^{\boldsymbol{m},n}, Q_k^n(il-) - \sum_{j':(j',k) \in \cup_{s=1}^{h-1} O_s} \mathcal{M}_{ij'k}^{\boldsymbol{m},n} \right) \\
&\overset{(a)}{\ge} nl^n \min\left( m_{jk} - \frac{1}{nl^n},\ \lambda_j - \frac{\epsilon}{2T} - \sum_{k':(j,k') \in \cup_{s=1}^{h-1} O_s} m_{jk'}, \lambda_k - \frac{\epsilon}{2T} - \sum_{j':(j',k) \in \cup_{s=1}^{h-1} O_s} m_{j'k} \right) \\
&\overset{(b)}{\ge} nl^n \min\left( m_{jk} - \frac{1}{nl^n}, m_{jk} - \frac{\epsilon}{2T} \right) \overset{(c)}{\ge} nl^n \left( m_{jk} - \frac{\epsilon}{2T} \right),
\end{aligned}$$

for all large enough $n \in \mathbb{N}$. Here, (a) follows from the queue length lower bound (70), the matching upper bound that $\mathcal{M}_{ijk}^{\boldsymbol{m},n} \le nl^n m_{jk}$, and the fact that $\lfloor x \rfloor \ge x - 1$; (b) follows from the feasibility of $\boldsymbol{m}$ within $\mathbb{M}$ and the fact that $O_h \cup O_{h'} = \emptyset$ for $h \ne h'$, so that $m_{jk} \le \min(\lambda_j - \sum_{k':(j,k') \in \cup_{s=1}^{h-1} O_s} m_{jk'}, \lambda_k - \sum_{j':(j',k) \in \cup_{s=1}^{h-1} O_s} m_{j'k})$; and (c) is due to $nl^n \to \infty$. □

## E. Omitted Proofs and Results for §9

**E.1. Extending the statistical guarantees for DDMP-KA to $\varepsilon$−optimal solutions** Recall that the statistical guarantees established in Theorem 3 for the DDMP-KA relies on Theorem 1, which provides a sample-path-level upper bound on $\left|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}^{\mathrm{KA}})\right|$. The following result generalizes Theorem 1 to $\varepsilon$−optimal solutions.

PROPOSITION 9. *Fix any $\varepsilon \ge 0$. Given historical reneging samples $\{r_{wb}^G : w \in \mathbb{W}, b \in [B_w^G]\}$, consider any $\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}} \in \{\boldsymbol{m} \in \mathbb{M} : \hat{F}^{\mathrm{KA}}(\boldsymbol{m}) \ge \sup_{\boldsymbol{m}' \in \mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m}') - \varepsilon\}$. It holds that*

$$\left|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})\right| \le 2 \sum_{w \in \mathbb{W}} c_w \lambda_w (\hat{\Delta}_w^G + \hat{\Delta}_w^{G^{-1}}) + \varepsilon,$$

*where $\hat{\Delta}_w^G$ and $\hat{\Delta}_w^{G^{-1}}$ are as defined in (12).*

**Proof.** We closely follow the proof of Theorem 1. Claims 2 and 3 therein still hold. However, Claim 1 needs to be revised as $|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})| \le 2\sup_{\boldsymbol{m}\in\mathbb{M}} |F(\boldsymbol{m}) - \hat{F}^{\mathrm{KA}}(\boldsymbol{m})|$. To see this, observe that

$$\begin{aligned}
|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})| &= |\sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m}) + \sup_{\boldsymbol{m}\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m}) - \hat{F}^{\mathrm{KA}}(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}) + \hat{F}^{\mathrm{KA}}(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})| \\
&\le |\sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m})| + \varepsilon + |\hat{F}^{\mathrm{KA}}(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})| \\
&\le 2\sup_{\boldsymbol{m}\in\mathbb{M}} |F(\boldsymbol{m}) - \hat{F}^{\mathrm{KA}}(\boldsymbol{m})| + \varepsilon,
\end{aligned}$$

where the equality follows from noting that $F(\boldsymbol{m}^\star) = \sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m})$ and telescoping; the first inequality is by the triangle inequality and the $\varepsilon-$optimality of $\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}$; and the last inequality is due to the fact that both $|\sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m})| \le$ and $|\hat{F}^{\mathrm{KA}}(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})|$ are no greater than $\sup_{\boldsymbol{m}\in\mathbb{M}} |F(\boldsymbol{m}) - \hat{F}^{\mathrm{KA}}(\boldsymbol{m})|$. Combining the revised Claim 1 with Claims 2 and 3 leads to the desired result. □

Using Proposition 9, we can generalize Theorem 3 to $\varepsilon-$optimal solutions as follows.

PROPOSITION 10. *Fix any $\varepsilon \ge 0$. Suppose Assumption 1 holds. Then, for any $\epsilon < 1/(\nu_1 \max\{1, \max_{w\in\mathbb{W}} \theta_w\})$, it holds for any $\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}} \in \{\boldsymbol{m}\in\mathbb{M} : \hat{F}^{\mathrm{KA}}(\boldsymbol{m}) \ge \sup_{\boldsymbol{m}'\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m}') - \varepsilon\}$ that*

$$\mathbb{P}\left(\left|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})\right| \ge \epsilon + \varepsilon\right) \le \sum_{w\in\mathbb{W}} \Phi_w\left(B_w^G; \nu_1\epsilon\right) + \sum_{w\in\mathbb{W}} \sum_{i\in\{1,2,3\}} \Psi_{wi}\left(B_w^G; \delta_w, \nu_1\epsilon\right)$$

*for all $\delta_w \in \left[0, \min\left\{\frac{1}{2}, 1 - G_w(y_w(\nu_1\epsilon)), G_w(\nu_1\epsilon)\right\}\right]$ and all $B_w^G \ge 1$, where the functions $\Phi_w$, $y_w$, and $\Psi_{wi}$ and are given in (13), (21), and (22), respectively, and $\nu_1 := [4(J+K)\max_{w\in\mathbb{W}}\{c_w\lambda_w\}]^{-1}$. Furthermore, if*

$$\chi(\epsilon) := F(\boldsymbol{m}^\star) - \sup_{\boldsymbol{m}\in\mathbb{M}:|\boldsymbol{m}-\boldsymbol{m}^\star|\ge\epsilon} F(\boldsymbol{m}) > 0 \quad \text{for all} \quad \epsilon > 0.$$

*Then, for all $\epsilon$ such that $\chi(\epsilon) \in (\varepsilon, \varepsilon + 1/(\nu_1 \max\{1, \max_{w\in\mathbb{W}} \theta_w\})$, it holds for any $\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}} \in \{\boldsymbol{m}\in\mathbb{M} : \hat{F}^{\mathrm{KA}}(\boldsymbol{m}) \ge \sup_{\boldsymbol{m}'\in\mathbb{M}} \hat{F}^{\mathrm{KA}}(\boldsymbol{m}') - \varepsilon\}$ that*

$$\mathbb{P}\left(\left|\boldsymbol{m}^\star - \hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}\right| \ge \epsilon\right) \le \sum_{w\in\mathbb{W}} \Phi_w\left(B_w^G; \nu_1(\chi(\epsilon) - \varepsilon)\right) + \sum_{w\in\mathbb{W}} \sum_{i\in\{1,2,3\}} \Psi_{wi}\left(B_w^G; \delta_w, \nu_1(\chi(\epsilon) - \varepsilon)\right),$$

*for all $\delta_w \in \left[0, \min\left\{\frac{1}{2}, 1 - G_w(y_w(\nu_1(\chi(\epsilon)-\varepsilon))), G_w(\nu_1(\chi(\epsilon)-\varepsilon))\right\}\right]$ and all $B_w^G \ge 1$.*

**Proof.** Proposition 9 implies that $|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}^{\mathrm{KA}})| \le 2\max_{w\in\mathbb{W}}\{c_w\lambda_w\}\sum_{w\in\mathbb{W}}(\hat{\Delta}_w^G + \hat{\Delta}_w^{G^{-1}}) + \varepsilon$. Hence, the event $\{|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}^{\mathrm{KA}})| \ge \epsilon + \varepsilon\}$ implies that at least one of the events $\{\hat{\Delta}_w^G \ge \nu_1\epsilon\}$ and $\{\hat{\Delta}_w^{G^{-1}} \ge \nu_1\epsilon\}, w \in \mathbb{W}$ must happen. Then it follows from the subadditivity of probability measure that

$$\mathbb{P}\left(\left|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}^{\mathrm{KA}})\right| \ge \epsilon + \varepsilon\right) \le \sum_{w\in\mathbb{W}} \mathbb{P}\left(\hat{\Delta}_w^G \ge \nu_1\epsilon\right) + \sum_{w\in\mathbb{W}} \mathbb{P}\left(\hat{\Delta}_w^{G^{-1}} \ge \nu_1\epsilon\right).$$

By applying Lemma 1 and Corollary 1 (i.e., corollary to Theorem 2) to upper bound the right hand side of the above display, we can complete the proof of the optimal value gap. To establish finite-sample bounds on the solution gap $|\boldsymbol{m}^\star - \hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}|$, note that $\mathbb{P}\left(|\boldsymbol{m}^\star - \hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}| \ge \epsilon\right) \le \mathbb{P}\left(|F(\boldsymbol{m}^\star) - F(\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}})| \ge \chi(\epsilon)\right)$ by the definition of $\chi$ and the feasibility of $\hat{\boldsymbol{m}}_\varepsilon^{\mathrm{KA}}$ within $\mathbb{M}$. The result then follows from applying the finite-sample bound on the optimal value gap to control the right-hand side. □

**E.2. Proof of Lemma 12.** By telescoping, we have

$$\begin{aligned}
&\left|F(\hat{\boldsymbol{m}}_\varepsilon) - F(\boldsymbol{m}^\star)\right| \\
&= \left|F(\hat{\boldsymbol{m}}_\varepsilon) - \hat{F}(\hat{\boldsymbol{m}}_\varepsilon) + \hat{F}(\hat{\boldsymbol{m}}_\varepsilon) - \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m}) + \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} F(\boldsymbol{m}) + \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} F(\boldsymbol{m}) - F(\boldsymbol{m}^\star)\right| \\
&\le 2\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \left|F(\boldsymbol{m}) - \hat{F}(\boldsymbol{m})\right| + \left|\hat{F}(\hat{\boldsymbol{m}}_\varepsilon) - \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m})\right| + \left|\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} F(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m})\right| \\
&\le 2\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \left|F(\boldsymbol{m}) - \hat{F}(\boldsymbol{m})\right| + \varepsilon + \left|\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} F(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m})\right|,
\end{aligned}$$

where the first inequality follows from the triangle inequality, the fact that both $\left|F(\hat{\boldsymbol{m}}_\varepsilon) - \hat{F}(\hat{\boldsymbol{m}}_\varepsilon)\right|$ and $\left|\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m}) - \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} F(\boldsymbol{m})\right|$ are no greater than $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}}\left|F(\boldsymbol{m}) - \hat{F}(\boldsymbol{m})\right|$ since $\hat{\boldsymbol{m}}^\kappa \in \hat{\mathbb{M}}$, and the fact that $F(\boldsymbol{m}^\star) = \sup_{\boldsymbol{m}\in\mathbb{M}} F(\boldsymbol{m})$; and the second inequality follows from the definition of $\varepsilon$−optimal solution. □

**E.3. Proof of Proposition 7.** Since any feasible solution $\boldsymbol{m} \in \hat{\mathbb{M}}$ to the DDMP (10) must satisfy $m_{jk} = 0$ for all $(j,k) \notin \mathcal{E}$, these variables can be treated as fixed constants. Hence, the optimization can be restricted to $\{m_{jk} : (j,k) \in \mathcal{E}\}$. Under this observation, the constraint $\boldsymbol{m} \in \hat{\mathbb{M}}$ can be equivalently rewritten as (49b). For any fixed $j \in \mathbb{J}$ and $\boldsymbol{m} \in \hat{\mathbb{M}}$, we have $\hat{p}_j(\boldsymbol{m}) = [(1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j) \wedge 1]^+ = 1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j$. We introduce binary variables $z_{jb}, b \in \{0\} \cup [B_j^G]$ to indicate which of the intervals $(\frac{b-1}{B_j^G}, \frac{b}{B_j^G}]$ that $1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j$ belongs to. Specifically, $z_{j0} = 1$ indicates that $1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j = 0$, while for $b > 0$, $z_{jb} = 1$ indicates that $1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j \in (\frac{b-1}{B_j^G}, \frac{b}{B_j^G}]$. This logical relation is enforced by constraints (49c), (49d) and (49g). Given that $z_{jb} = 1$, we have $(\hat{G}_j)^{-1}(1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j) = r_{jb}^G$ by definition of the quantile function in (9), where we have created an artificial sample $r_{j0}^G := r_{j1}^G$. Thus, given $z_{jb} = 1$, by the definition of $\hat{q}_w(p)$ below (10), we have $\hat{q}_j(\hat{p}_j(\boldsymbol{m})) = \hat{\lambda}_j \int_0^{r_{jb}^G} (1 - \hat{G}_j(u))du = \hat{I}_{jb}$, where the second equality follows from directly calculating the integration and the fact that $\hat{G}_j$ defined in (9) is a step function that jumps upward by $1/B_j^G$ at each sample point $r_{jb}^G$. Therefore, we can express $\hat{q}_j(\hat{p}_j(\boldsymbol{m}))$ as the summation $\sum_{b\in[B_j^G]} \hat{I}_{jb} z_{jb}$. Replicating the above arguments for all $k \in \mathbb{K}$ leads to constraints (49e) and (49f), and we can express $\hat{q}_k(\hat{p}_k(\boldsymbol{m}))$ as the summation $\sum_{b\in[B_k^G]} \hat{I}_{kb} z_{kb}$. Finally, substituting these expressions of $\hat{q}_w(\hat{p}_w(\boldsymbol{m}))$ into the objective function $\hat{F}$ of the DDMP (10) yields the objective function (49a). □

**E.4. Proof of Proposition 8.** Throughout the proof, we treat $\{m_{jk} : j \in \mathbb{J}, k \in \mathbb{K}\}$ as the decision variables in (49) and the DDMP-Lin, with the additional constraints $m_{jk} = 0$ whenever $(j,k) \notin \mathcal{E}$. First, we prove that $0 \le \sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m}) - \hat{F}(\hat{\boldsymbol{m}}^\kappa)$. We do this by showing that the feasible region of the decision variables $\{m_{jk} : j \in \mathbb{J}, k \in \mathbb{K}\}$ in the DDMP-Lin is a subset of those in formulation (49). To see this, note that in (49), $1 - \sum_{k\in\mathbb{K}} m_{jk}/\hat{\lambda}_j$ is allowed to take value in the whole interval $[0,1]$, but in the DDLP-Lin, due to the positivity of $\kappa$, it can only take value in $[0,1] \backslash \{(b/B_j^G, b/B_j^G + \kappa) : b = 0, \ldots, B_j^G - 1\}$, which is a subset of $[0,1]$. A similar argument applies to $1 - \sum_{j\in\mathbb{J}} m_{jk}/\hat{\lambda}_k, k \in \mathbb{K}$. Therefore, $\hat{F}(\hat{\boldsymbol{m}}^\kappa)$ is no greater than the optimal value of (49), which equals $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m})$ by Proposition 7.

Second, we prove the upper bound using a constructive argument. Specifically, for any $\varepsilon$−optimal solution $\hat{\boldsymbol{m}}_\varepsilon$ of (49), we will construct a solution $\tilde{\boldsymbol{m}}_\varepsilon$ such that 1) $\tilde{\boldsymbol{m}}_\varepsilon$ is feasible for the DDMP-Lin, and 2) $\hat{F}(\hat{\boldsymbol{m}}_\varepsilon) - \hat{F}(\tilde{\boldsymbol{m}}_\varepsilon) \le 2(J+K)\hat{\lambda}_{\max}(\sum_{j\in\mathbb{J},k\in\mathbb{K}} v_{jk})\kappa + \sum_{w\in\mathbb{W}} c_w \hat{\lambda}_w \sup_{b\in[B_w^G]}(r_{w,b}^G - r_{w,b-1}^G)$. Moreover, by the definition of $\varepsilon$−optimal solution and the optimality of $\hat{\boldsymbol{m}}^\kappa$ in the DDLP-Lin, we have $\sup_{\boldsymbol{m}\in\hat{\mathbb{M}}} \hat{F}(\boldsymbol{m}) - \hat{F}(\hat{\boldsymbol{m}}^\kappa) \le \hat{F}(\hat{\boldsymbol{m}}_\varepsilon) - \hat{F}(\tilde{\boldsymbol{m}}_\varepsilon) + \varepsilon$ for any $\varepsilon > 0$. Then, proof can be completed by taking the limit as $\varepsilon \downarrow 0$.

In the following, we describe the construction of $\tilde{\boldsymbol{m}}_\varepsilon$ for any given $\hat{\boldsymbol{m}}_\varepsilon$. For notational simplicity, we omit the subscript $\varepsilon$. The construction algorithm operates in three steps: (1) Initialize $\tilde{m}_{jk} = \hat{m}_{jk}$ for all $j \in \mathbb{J}$ and $k \in \mathbb{K}$; (2) For each $j \in \mathbb{J}$, if $\tilde{f}_j := 1 - \sum_{k\in\mathbb{K}} \tilde{m}_{jk}/\hat{\lambda}_j \in (b/B_j^G, b/B_j^G + \kappa)$ for some $b = 0, \ldots, B_j^G - 1$, update $\tilde{m}_{jk} \leftarrow [\tilde{m}_{jk} - \hat{\lambda}_j\kappa]^+$ for all $k \in \mathbb{K}$; and for each $k \in \mathbb{K}$, if $\tilde{f}_k := 1 - \sum_{j\in\mathbb{J}} \tilde{m}_{jk}/\hat{\lambda}_k \in (b/B_k^G, b/B_k^G + \kappa)$ for some $b = 0, \ldots, B_k^G - 1$, update $\tilde{m}_{jk} \leftarrow [\tilde{m}_{jk} - \hat{\lambda}_k\kappa]^+$ for all $j \in \mathbb{J}$; (3) Repeat step (2) if any update on $\tilde{\boldsymbol{m}}$ has occurred, and the algorithm terminates if otherwise.

Next, we show that this algorithm terminates within at most $(J+K)$ iterations of step (2) and returns a feasible solution $\tilde{\boldsymbol{m}}$ for DDMP-Lin. Note that in each iteration of step (2), the values of $\tilde{m}_{jk}$ decreases by no more than $2\hat{\lambda}_{\max}\kappa$ for each $j$ and $k$. As a result, the quantities $\tilde{f}_j$ and $\tilde{f}_k$ increases by no more than $2\hat{\lambda}_{\max}\kappa \times \max\{J,K\}/\hat{\lambda}_{\min}$ in each iteration of step (2). In the statement of the theorem, we have set $\kappa$ sufficiently small such that $\kappa + (J+K) \times 2\hat{\lambda}_{\max}\kappa \max\{J,K\}/\hat{\lambda}_{\min} \le 1/B_{\max}^G$. This ensures that if $\tilde{f}_j \in (b/B_j^G, b/B_j^G + \kappa)$ for some $b = 0, \ldots, B_j^G - 1$ at some iteration of step (2), $\tilde{f}_j$ will increase and exit this interval after that iteration (by the way we decrease $\tilde{m}_{jk}, k \in \mathbb{K}$), and importantly, $\tilde{f}_j$ will remain below $(b+1)/B_j^G$ in all subsequent $(J+K)$ iterations (by our choice of a sufficiently small $\kappa$). This means that for this particular $j \in \mathbb{J}$, no further update on $\tilde{m}_{jk}, k \in \mathbb{K}$ will occur in all subsequent $(J+K)$ iterations, and the constraint imposed on $\tilde{f}_j$ in the DDMP-Lin is satisfied because $\tilde{f}_j \notin (b/B^G_j, b/B^G_j + \kappa)$ for all $b = 0, \ldots, B^G_j - 1$. Since the same arguments apply to each $\tilde{f}_k, k \in \mathbb{K}$, the algorithm must terminate within $(J+K)$ executions of step (2). At termination,

we have $(J+K) \times 2\hat{\lambda}_{\max}\kappa \le \tilde{m}_{jk} - \hat{m}_{jk} \le 0$ for all $j$ and $k$. This gives the first component of the upper bound. Moreover, we have $1/B^G_j \ge (1 - \sum_{k\in\mathbb{K}} \tilde{m}_{jk}/\hat{\lambda}_j) - (1 - \sum_{k\in\mathbb{K}} \hat{m}_{jk}/\hat{\lambda}_j) \ge 0$ since $\tilde{f}_j$ remains below $(b+1)/B^G_j$. This implies that $\hat{q}_j(1 - \sum_{k\in\mathbb{K}} \tilde{m}_{jk}/\hat{\lambda}_j) - \hat{q}_j(1 - \sum_{k\in\mathbb{K}} \hat{m}_{jk}/\hat{\lambda}_j) \le \sup_{b\in\{1,\ldots,B^G_w-1\}} \big(\hat{I}_{j,b} - \hat{I}_{j,b-1}\big)$ by the role of $\hat{I}_{j,b}$ in the proof of Proposition 7. From (50), we have $\hat{I}_{j,b} - \hat{I}_{j,b-1} = \hat{\lambda}_j\big(B - b + 1\big)(r_{j,b} - r_{j,b-1})/B^G_j \le \hat{\lambda}_j(r_{j,b} - r_{j,b-1})$. This yields the second component of the upper bound as the same reasoning applies for $\hat{q}_k(1 - \sum_{j\in\mathbb{J}} \tilde{m}_{jk}/\hat{\lambda}_k) - \hat{q}_k(1 - \sum_{j\in\mathbb{J}} \hat{m}_{jk}/\hat{\lambda}_k)$. □

**E.5. MILP Solution Time** We conduct numerical experiments to evaluate the solution time of DDMP-Lin (Definition 7) under varying network sizes $(J \times K)$, reneging sample sizes $B^G_w$, and network parameters $(c_w, v_{jk}, \hat{\lambda}_w)$, assuming a fully connected network $\mathcal{E} = \mathbb{J} \times \mathbb{K}$. Specifically, we consider $(J,K) \in \{(2,1),(2,2),(3,3),(4,4),(5,5)\}$ and $B^G_w \in \{100, 250, 500, 1000, 1500\}$. For each combination of network size and sample size, we run five independent replications. In each replication, we (i) generate reneging samples $r^G_{wb} : b \in [B^G_w]$ by sampling from a Weibull-4 distribution with mean 1 for each $w \in \mathbb{W}$; (ii) sample $c_w$ and $\hat{\lambda}_w$ independently from uniform distributions on $[1,5]$ and $[1,10]$, respectively, for each $w \in \mathbb{W}$; and (iii) sample $v_{jk}$ independently from a uniform distribution on $[1,5]$ for each $(j,k) \in \mathbb{J} \times \mathbb{K}$. Each generated instance is solved using Gurobi 12.0.2 on a personal MacBook equipped with an Apple M2 Pro chip and 16 GB of memory, and the corresponding runtime is recorded. Instances not solved within 3600 seconds are terminated and assigned a runtime of 3600 seconds. Table 4 reports the median runtime, with the 25% and 75% quantiles shown in parentheses. The results show that instances with up to 12,000 binary variables can typically be solved within seven minutes. However, some larger instances (e.g., $J \times K = 5 \times 5$) cannot be solved within one hour. Developing more efficient algorithms to accelerate the MILP solution process is therefore an important direction for future research.

TABLE 4. Runtime of solving the DDMP-Lin (Definition 7) under different network sizes ($J \times K$), reneging sample sizes ($B_w^G$), and randomly generated network parameters ($c_w, v_{jk}, \hat{\lambda}_w$). The third column reports the median solution time (in seconds) over five replications, with the 25% and 75% quantiles shown in parentheses; runs exceeding 3600 seconds are truncated at 3600 seconds.

| Network Size: $J \times K$ | Sample Size: $B_w^G$ | Runtime in Seconds: Median (25%, 75%) | Number of Binary Variables | Number of Continuous Variables | Number of Constraints |
|---|---|---|---|---|---|
| $2 \times 1$ | 100 | 0.0438 (0.0402, 0.0492) | 303 | 2 | 12 |
| | 250 | 0.117 (0.116, 0.131) | 753 | | |
| | 500 | 0.612 (0.564, 1.049) | 1,503 | | |
| | 1000 | 0.330 (0.293, 0.472) | 3,003 | | |
| | 1500 | 0.488 (0.467, 0.503) | 4,503 | | |
| $2 \times 2$ | 100 | 0.0944 (0.0664, 0.0948) | 404 | 4 | 16 |
| | 250 | 0.683 (0.322, 0.786) | 1,004 | | |
| | 500 | 2.190 (1.998, 2.398) | 2,004 | | |
| | 1000 | 3.460 (1.555, 7.262) | 4,004 | | |
| | 1500 | 5.787 (3.717, 10.284) | 6,004 | | |
| $3 \times 3$ | 100 | 0.374 (0.313, 0.496) | 606 | 9 | 24 |
| | 250 | 0.828 (0.792, 0.847) | 1,506 | | |
| | 500 | 10.042 (3.397, 11.161) | 3,006 | | |
| | 1000 | 23.936 (9.220, 66.666) | 6,006 | | |
| | 1500 | 20.766 (18.725, 33.720) | 9,006 | | |
| $4 \times 4$ | 100 | 0.422 (0.261, 0.461) | 808 | 16 | 32 |
| | 250 | 4.134 (2.385, 4.773) | 2,008 | | |
| | 500 | 13.819 (11.509, 85.959) | 4,008 | | |
| | 1000 | 288.237 (14.983, 462.870) | 8,008 | | |
| | 1500 | 250.693 (54.387, 347.499) | 12,008 | | |
| $5 \times 5$ | 100 | 1.249 (0.352, 2.660) | 1,010 | 25 | 40 |
| | 250 | 27.308 (19.290, 42.157) | 2,510 | | |
| | 500 | 205.237 (149.912, 271.616) | 5,010 | | |
| | 1000 | 3600 (3600, 3600) | 10,010 | | |
| | 1500 | 3600 (609.986, 3600) | 15,010 | | |

**F. Author Contributions** The authors are listed in alphabetical order. All authors contributed to the development of the proofs and the writing of the manuscript.

**References**

[1] Gubner JA (2021) The Gamma function and Stirling's formula. URL https://gubner.ece.wisc.edu/ notes/GammaFunctionStirling.pdf, accessed: 2024-10-25.

[2] Natalini P, Palumbo B (2000) Inequalities for the incomplete Gamma function. Math. Inequal. Appl 3(1):69–77.

[3] Vershynin R (2018) High-Dimensional Probability: An Introduction with Applications in Data Science, volume 47 (Cambridge University Press).