Supporting Information

Fibich et al. 10.1073/pnas.1206867109

SI Text

Proof of Theorem 1 and Corollary 3. Because of the differentiability of F, there exist positive constants $\delta(\overline{\mu}_A)$ and $C(\overline{\mu}_A)$, such that for all $\boldsymbol{\mu}$ that satisfy Eq. **3**,

$$\left| F(\boldsymbol{\mu}) - F(\overline{\boldsymbol{\mu}}_A) - \sum_{j=1}^{k} (\mu_j - \overline{\mu}_A) \frac{\partial F}{\partial \mu_j}\Big|_{\overline{\boldsymbol{\mu}}_A} \right| \le C(\overline{\mu}_A) \|\boldsymbol{\mu} - \overline{\boldsymbol{\mu}}_A\|^2.$$

If F is symmetric, then

$$\frac{\partial F}{\partial \mu_i}\Big|_{\overline{\boldsymbol{\mu}}_A} = \frac{\partial F}{\partial \mu_1}\Big|_{\overline{\boldsymbol{\mu}}_A}, \qquad j = 1, \ldots, k. \qquad \textbf{[S1]}$$

Because $\overline{\mu}_A$ is the arithmetic average,

$$\sum_{j=1}^{k} (\mu_j - \overline{\mu}_A) \frac{\partial F}{\partial \mu_j}\Big|_{\overline{\boldsymbol{\mu}}_A} = \frac{\partial F}{\partial \mu_1}\Big|_{\overline{\boldsymbol{\mu}}_A} \sum_{j=1}^{k} (\mu_j - \overline{\mu}_A) = 0.$$

Hence, the result follows.

Note that symmetry was used only to derive Eq. **S1**. Because

$$\frac{\partial F}{\partial \mu_i}\Big|_{\overline{\boldsymbol{\mu}}_A} = \lim_{\eta \to 0} \frac{F(\overline{\boldsymbol{\mu}}_A + \eta \hat{e}_i) - F(\overline{\boldsymbol{\mu}}_A)}{\eta},$$

deriving Eq. **S1** requires only weak symmetry. Therefore, *Corollary 3* follows.

Proof of Lemma 2. We calculate $F(\mu_1, \mu_2)$ explicitly, using the steady-state transition diagram that is shown in Fig. S1. We denote by p_i the steady-state probability for the system to be with i customers and by $p_1^{(1,0)}$ and $p_1^{(0,1)}$ the steady-state probability for the system to be with one customer in servers 1 and 2, respectively. In particular, $p_1 = p_1^{(1,0)} + p_1^{(0,1)}$. Because in steady state the amount of inflow is equal to the amount of outflow, the following equalities hold:

$$\lambda p_0 = \mu_1 p_1^{(1,0)} + \mu_2 p_1^{(0,1)}, \qquad \textbf{[S2a]}$$

$$\frac{\lambda}{2} p_0 + \mu_2 p_2 = (\lambda + \mu_1) p_1^{(1,0)}, \qquad \textbf{[S2b]}$$

$$\frac{\lambda}{2} p_0 + \mu_1 p_2 = (\lambda + \mu_2) p_1^{(0,1)}, \qquad \textbf{[S2c]}$$

$$\lambda p_1^{(1,0)} + \lambda p_1^{(0,1)} + (\mu_1 + \mu_2) p_3 = (\lambda + \mu_1 + \mu_2) p_2, \qquad \textbf{[S2d]}$$

$$\lambda p_n + (\mu_1 + \mu_2) p_{n+2} = (\lambda + \mu_1 + \mu_2) p_{n+1}, \qquad n = 2, 3, \ldots \qquad \textbf{[S2e]}$$

We can view Eqs. **S2a–S2c** as a linear system for the three unknowns $p_0, p_1^{(1,0)}, p_1^{(0,1)}$. Solving this system for p_0 yields

$$p_0 = \frac{2\mu_1 \mu_2}{\lambda^2} p_2.$$

In addition, the solution of Eqs. **S2d** and **S2e** is $p_n = \left(\frac{\lambda}{\mu_1 + \mu_2}\right)^{n-2} p_2 = \rho^{n-2} p_2$ for $n \ge 1$. Substituting the above in

$$1 = \sum_{n=0}^{\infty} p_n = p_0 + \sum_{n=1}^{\infty} \rho^{n-2} p_2 = \left(\frac{2\mu_1\mu_2}{\lambda^2} + \frac{1}{\rho}\frac{1}{1-\rho}\right) p_2$$

gives $p_2 = \left(\frac{2\mu_1\mu_2}{\lambda^2} + \frac{1}{\rho}\frac{1}{1-\rho}\right)^{-1}$. Therefore,

$$F(\mu_1, \mu_2) = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n \rho^{n-2} p_2 = \frac{p_2}{\rho} \sum_{n=0}^{\infty} n \rho^{n-1}$$
$$= \frac{p_2}{\rho} \left(\sum_{n=0}^{\infty} \rho^n\right)' = \frac{p_2}{\rho} \left(\frac{1}{1-\rho}\right)' = \frac{p_2}{\rho} \frac{1}{(1-\rho)^2},$$

and the result follows.

M/M/3 Queue. Consider the case of three heterogeneous servers with average service times μ_1, μ_2, and μ_3. Denote by p_0, $p_1^{(1,0,0)}$, $p_1^{(0,1,0)}$, $p_1^{(0,0,1)}$, $p_2^{(1,1,0)}$, $p_2^{(1,0,1)}$, $p_2^{(0,1,1)}$, and $p_3, p_4, \ldots$, the steady-state probabilities. Thus, for example, $p_2^{(1,0,1)}$ is the steady-state probability that servers 1 and 3 are busy, server 2 is free, and there are no waiting customers in the queue (we denote by p_n, $n \ge 2$ the probability of having n customers in the system). The transition diagram for $k = 3$ servers is given in Fig. S2. The steady-state equations are

$$\lambda p_0 = \mu_1 p_1^{(1,0,0)} + \mu_2 p_1^{(0,1,0)} + \mu_3 p_1^{(0,0,1)},$$

$$\frac{\lambda}{3} p_0 + \mu_2 p_2^{(1,1,0)} + \mu_3 p_2^{(1,0,1)} = (\mu_1 + \lambda) p_1^{(1,0,0)},$$

$$\frac{\lambda}{3} p_0 + \mu_1 p_2^{(1,1,0)} + \mu_3 p_2^{(0,1,1)} = (\mu_2 + \lambda) p_1^{(0,1,0)},$$

$$\frac{\lambda}{3} p_0 + \mu_1 p_2^{(1,0,1)} + \mu_2 p_2^{(0,1,1)} = (\mu_3 + \lambda) p_1^{(0,0,1)},$$

$$\frac{\lambda}{2} p_1^{(1,0,0)} + \frac{\lambda}{2} p_1^{(0,1,0)} + \mu_3 p_3 = (\lambda + \mu_1 + \mu_2) p_2^{(1,1,0)},$$

$$\frac{\lambda}{2} p_1^{(1,0,0)} + \frac{\lambda}{2} p_1^{(0,0,1)} + \mu_2 p_3 = (\lambda + \mu_1 + \mu_3) p_2^{(1,0,1)},$$

$$\frac{\lambda}{2} p_1^{(0,1,0)} + \frac{\lambda}{2} p_1^{(0,0,1)} + \mu_1 p_3 = (\lambda + \mu_2 + \mu_3) p_2^{(0,1,1)},$$
$$= (\lambda + \mu_1 + \mu_2 + \mu_3) p_3,$$

$$\lambda p_n + (\mu_1 + \mu_2 + \mu_3) p_{n+2} = (\lambda + \mu_1 + \mu_2 + \mu_3) p_{n+1}, \qquad n \ge 3,$$

$$\sum_{n=0}^{\infty} p_n = 1.$$

The solution of the last two equations is $p_n = \left(\frac{\lambda}{\mu_1 + \mu_2 + \mu_3}\right)^{n-3} p_3$ for $n \ge 2$. The values of p_0, p_1, p_2 as a function of p_3 can be evaluated explicitly with MAPLE, by solving the first $2^3 - 1 = 7$ linear equations for p_0, $p_1^{(1,0,0)}$, $p_1^{(0,1,0)}$, $p_1^{(0,0,1)}$, $p_2^{(1,1,0)}$, $p_2^{(1,0,1)}$, and

$p_2^{(0,1,1)}$. The resulting expression for $F(\mu_1, \mu_2, \mu_3)$, however, is extremely cumbersome and not informative.

Proof of Theorem 4. Because customers are randomly assigned to the available servers, $F(\mu_1, \ldots, \mu_k)$ is symmetric. To see that F is differentiable in $(\mu_1, \ldots, \mu_k)$, we note that $F = \sum_{n=0}^{\infty} n p_n$, where p_n is the steady-state probability that there are n customers in the system. In addition, $\{p_n\}_{n=1}^{k}$ are the solutions of a linear system with coefficients that depend smoothly on $(\mu_1, \ldots, \mu_k)$, and $p_n = \left(\frac{\lambda}{\mu_1 + \cdots + \mu_k}\right)^{n-k} p_k$ for $n \geq k-1$. This was shown explicitly for the cases $k = 2$ and $k = 3$; the proof for $k > 3$ is similar.

Averaging Principle for Functions (Proof of Eq. 12). Let $F_1, \ldots, F_k$ belong to a function space $\mathcal{F}$, let $\epsilon \in \mathbb{R}$, and let $R : (F_1, \ldots, F_k) \mapsto R[F_1, \ldots, F_k] \in \mathbb{R}$ be a functional. We say that the functional R is differentiable if it is twice differentiable in the sense of Fréchet. (We can also relax this assumption and assume that R is once differentiable in the sense of Fréchet, and the scalar function $\tilde{R}(\epsilon) := R[F_1 = F + \epsilon H_1, \ldots, F_k = F + \epsilon H_k]$ is twice differentiable at and near $\epsilon = 0$, for every $F, H_1, \ldots, H_k \in \mathcal{F}$.) By Taylor expansion,

$$\tilde{R}(\epsilon) = \tilde{R}(0) + \epsilon \sum_{j=1}^{k} \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R\left[(\underbrace{F, \ldots, F}_{\times k}) + \epsilon H_j \hat{e}_j\right] + O(\epsilon^2),$$

where

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} R\Big[(F, \ldots, F) + \epsilon H_j \hat{e}_j\Big] = \frac{\delta R}{\delta F_j}[H_j],$$

and $\frac{\delta R}{\delta F_j}$ is the Fréchet derivative of $R[F_1, \ldots, F_k]$ with respect to F_j. Therefore,

$$\tilde{R}(\epsilon) = \tilde{R}(0) + \epsilon \sum_{j=1}^{k} \frac{\delta R}{\delta F_j}[H_j] + O(\epsilon^2).$$

Because R is symmetric and the Fréchet derivative is a linear operator,

$$\tilde{R}(\epsilon) = \tilde{R}(0) + \epsilon \frac{\delta R}{\delta F_1}\left[\sum_{j=1}^{k} H_j\right] + O(\epsilon^2).$$

Denote $\overline{F} := \frac{1}{k}\sum_{j=1}^{k} F_j$ and $H_j := F_j - \overline{F}$. Then $\sum_{j=1}^{k} H_j = 0$. Hence, $\tilde{R}(\epsilon) = \tilde{R}(0) + O(\epsilon^2)$, which is Eq. **12**.

Proof of Theorem 6. We first prove that F is differentiable. Denote $\delta_{i,i'} = 1$ if individuals i and i' influence each other and $\delta_{i,i'} = 0$ otherwise. For every k, every set of k consumers $\{i_1, i_2, \ldots, i_k\}$, and every increasing sequence of times $0 \leq t_1 \leq \ldots \leq t_k$, denote by $P(i_1, t_1, i_2, t_2, \ldots, i_k, t_k)$ the probability that consumer i_1 adopts the product before time t_1, consumer i_2 adopts the product between times t_1 and t_2, etc., and all consumers who are not in $\{i_1, \ldots, i_k\}$ do not adopt the process by time t_k. Then,

$$P(i_1, t_1) = \left(1 - \exp(-p_{i_1} t_1)\right) \prod_{j \neq i_1} \exp\left(-p_j t_1\right).$$

Similarly,

$$\begin{aligned} &P(i_1, t_1, i_2, t_2, \ldots, i_k, t_k) = \\ &P(i_1, t_1, i_2, t_2, \ldots, i_{k-1}, t_{k-1}) \\ &\times \left(1 - \exp\left(-\left(p_{i_k} + \sum_{m=1}^{k-1} \delta_{i_k, i_m} q_{i_m}\right)(t_k - t_{k-1})\right)\right) \\ &\times \prod_{j \notin \{i_1, \ldots, i_k\}} \exp\left(-\left(p_j + \sum_{m=1}^{k-1} \delta_{j, i_m} q_{i_m}\right)(t_k - t_{k-1})\right). \end{aligned}$$

Hence, the function $P(i_1, t_1, i_2, t_2, \ldots, i_k, t_k)$ is differentiable in $\{p_i, q_i\}$. Finally,

$$\begin{aligned} &E\Big[N\Big(t; \{p_j\}, \{q_j\}\Big)\Big] = \frac{1}{M} \sum_{\pi} \sum_{k=1}^{M} \frac{k}{(M-k)!} \\ &\times \int_{t_1=0}^{t} \int_{t_2=t_1}^{t} \cdots \int_{t_{k-1}=t_{k-2}}^{t} P(i_1, t_1, \ldots, i_{k-1}, t_{k-1}, i_k, t) dt_{k-1} \ldots dt_1, \end{aligned}$$

where π ranges over all permutations on the set of M individuals. Therefore, the differentiability of $E[N(t;\{p_j\},\{q_j\})]$ follows.

Because the network is translation invariant, F is weakly symmetric in $\{p_j\}$ and in $\{q_j\}$. By this we mean that

If $p_m = \tilde{p}$, $p_j = p$ for all $j \neq m$, and $q_j = q$ for all j, then F is independent of the value of m.

If $q_n = \tilde{q}$, $q_j = q$ for all $j \neq n$, and $p_j = p$ for all j, then F is independent of the value of n.

Therefore, the result follows from a slight modification of the proof of *Theorem 1*.

Proof of Eq. 14. Because F is symmetric, the quadratic term in the Taylor expansion of $F(\mu_1, \ldots, \mu_k)$ around the arithmetic mean is equal to

$$\begin{aligned} &\sum_{i,j=1}^{k} (\mu_i - \overline{\mu}_A)(\mu_j - \overline{\mu}_A) \frac{\partial^2 F}{\partial \mu_i \partial \mu_j}\bigg|_{\overline{\mu}_A} \\ &= \frac{\partial^2 F}{\partial \mu_1 \partial \mu_2}\bigg|_{\overline{\mu}} \sum_{i,j=1, i \neq j}^{k} (\mu_i - \overline{\mu}_A)(\mu_j - \overline{\mu}_A) + \frac{\partial^2 F}{\partial \mu_1 \partial \mu_1}\bigg|_{\overline{\mu}} \sum_{i=1}^{k} (\mu_i - \overline{\mu}_A)^2. \end{aligned}$$

Because $\overline{\mu}_A$ is the arithmetic mean,

$$\sum_{i,j=1}^{k} (\mu_i - \overline{\mu}_A)(\mu_j - \overline{\mu}_A) = \sum_{i=1}^{k} (\mu_i - \overline{\mu}_A) \sum_{j=1}^{k} (\mu_j - \overline{\mu}_A) = 0.$$

Therefore, the result follows.

Proof of Lemma 3. Consider the case where $\mu_i = \overline{\mu} + h$ for $i = 1, \ldots, k$. By Eq. **14**,

$$\frac{1}{2} \sum_{i,j=1}^{k} (\mu_i - \overline{\mu})(\mu_j - \overline{\mu}) \frac{\partial^2 F}{\partial \mu_i \partial \mu_j}\bigg|_{\overline{\mu}} = \frac{1}{2} \frac{\partial^2 F}{\partial \mu_1 \partial \mu_2}\bigg|_{\overline{\mu}} k(k-1)h^2 + \frac{1}{2} \frac{\partial^2 F}{\partial \mu_1 \partial \mu_1}\bigg|_{\overline{\mu}} kh^2.$$

On the other hand, because

$$F(\overline{\mu} + h, \ldots, \overline{\mu} + h) = F_{\text{homog.}}(\overline{\mu} + h),$$

we have

$$\frac{1}{2}\sum_{i,j=1}^{k}\left(\mu_i-\overline{\mu}\right)\left(\mu_j-\overline{\mu}\right)\frac{\partial^2 F}{\partial\mu_i\partial\mu_j}\bigg|_{\overline{\mu}}=\frac{h^2}{2}F''_{\text{homog.}}(\overline{\mu}).$$

Therefore,

$$\frac{1}{2}\frac{\partial^2 F}{\partial\mu_1\partial\mu_2}\bigg|_{\overline{\mu}}k(k-1)h^2+\frac{1}{2}\frac{\partial^2 F}{\partial\mu_1\partial\mu_1}\bigg|_{\overline{\mu}}kh^2=\frac{h^2}{2}F''_{\text{homog.}}(\overline{\mu}).$$

Hence,

$$\frac{\partial^2 F}{\partial\mu_1\partial\mu_2}\bigg|_{\overline{\mu}}=\frac{1}{k-1}\left(\frac{1}{k}F''_{\text{homog.}}(\overline{\mu})-\frac{\partial^2 F}{\partial\mu_1\partial\mu_1}\bigg|_{\overline{\mu}}\right).$$

Calculation of α. We illustrate the computation of the coefficient α for a queue with eight servers. Consider then the case of a single server with service time μ_1 and seven servers with service time μ, such that $\rho := \frac{\lambda}{7\mu+\mu_1} < 1$. Denote by $p_{0,n}$ and $p_{1,n}$, $n = 1, \ldots, 6$, the steady-state probabilities that n of the homogeneous servers are busy and that the single heterogeneous server is free or busy, respectively. The equations for the $2 \cdot 8 - 1 = 15$ variables p_0, $p_{0,1}, p_{1,0}, \ldots, p_{1,6}, p_{0,7}$ are

$$\lambda p_{0,0} = \mu p_{0,1} + \mu_1 p_{1,0},$$

$$-p_{0,0}\frac{\lambda}{8} + p_{1,0}(\lambda+\mu_1) - p_{1,1}\mu = 0,$$

$$p_{0,n}(\lambda+n\mu) = p_{0,n-1}\frac{8-n}{9-n}\lambda + p_{1,n}\mu_1 + p_{0,n+1}(n+1)\mu,$$

$$n = 1, \ldots, 6,$$

$$p_{1,n}(\mu_1+\lambda+n\mu) = p_{1,n-1}\lambda + p_{1,n+1}(n+1)\mu + p_{0,n}\frac{\lambda}{8-n},$$

$$n = 1, \ldots, 5,$$

$$p_{0,7}(\lambda+7)\mu) = p_{0,6}\frac{\lambda}{2} + p_7\mu_1\rho,$$

where $p_n = \rho^{n-7}p_7$ for $n \geq 8$, and $\sum_{n=0}^{\infty} p_n = 1$. These equations can be solved with Maple and the solution can be used to calculate $F(\mu_1, \underbrace{\mu, \ldots, \mu}_{\times 7})$ explicitly. (The Maple code is available at www.bgu.ac.il/~ariehg/averagingprinciple.html.) Differentiating this expression twice with respect to μ_1, differentiating $F_{\text{homog.}}$ (Eq. **6**) twice with respect to μ, and using *Lemma 3* yields Eq. **15**. Substituting $\overline{\mu}_A = 5$ and $\lambda = 28$ gives $\alpha \sim 0.00837$. In addition, $\sum_{i=1}^{8}(\mu_i-\overline{\mu})^2 = \epsilon^2\sum_{i=1}^{8}h_i^2 = 71\epsilon^2$. Therefore, $\alpha\sum_{i=1}^{8}(\mu_i-\overline{\mu})^2 \approx 0.594\epsilon^2$.

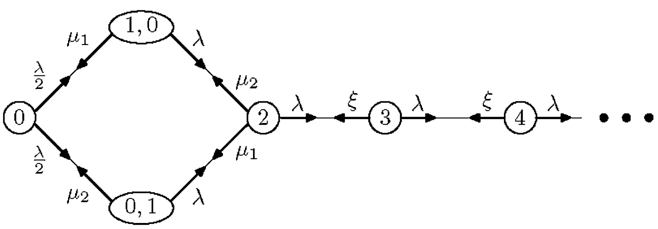

Fig. S1. Transition diagram of a queue with two heterogeneous servers. State "0" corresponds to the situation in which no server is busy. State (1, 0) corresponds to the situation in which server 1 is busy and server 2 is not busy. State (0, 1) corresponds to the situation in which server 1 is not busy and server 2 is busy. State "k" for $k \geq 2$ corresponds to the situation in which both servers are busy and $k - 2$ customers wait in the queue. Here, $\xi = \mu_1 + \mu_2$.

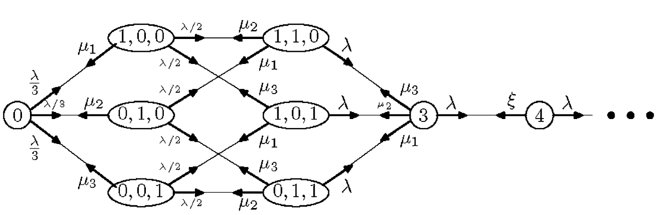

Fig. S2. Same as Fig. S1 with three heterogeneous servers. For example, state (0, 1, 1) corresponds to the situation in which server 1 is not busy and servers 2 and 3 are busy. Here, $\xi = \mu_1 + \mu_2 + \mu_3$.