

Service Outsourcing

Yong-Pin Zhou* and Z. Justin Ren†

February 2, 2010

Abstract

This article reviews the Operations Management (OM) research on service outsourcing, a common practice among today's businesses. We focus on recent literature in three areas: capacity planning and supplier coordination, service outsourcing under information asymmetry, and quality concerns. Additionally, a mathematical framework is presented that can be used to analyze service outsourcing supply chains. We conclude with discussions on some promising future research areas.

1 Introduction

Globalization poses challenges to today's businesses. To keep costs down, many business firms have turned to outsourcing. By definition, outsourcing refers to the act "to procure (as some goods or services needed by a business or organization) under contract with an outside supplier" (Merriam-Webster Dictionary). Almost any business processes can be outsourced: manufacturing, product design, procurement, customer care or other services (See article 4.5.5.3 for an in-depth discussion). Various benefits have been reported that fuel the outsourcing growth. Early outsourcing deals focus on cost savings (Johnson [17], Leavitt [21]), but as Kakabadse and Kakabadse [18] note, recently the goal of outsourcing has shifted from cost reduction to value creation. The value additions found by researchers have included: improved speed or quality of services (Johnson [17]), access to outside talent, innovation, and new technology, strategic alliances with outsourcing partners (Johnson [17], Leavitt [21]), focus on core missions (Brown and Wilson [5]), and greater flexibility (Johnson [17]). For a detailed discussion on advantages and disadvantages of outsourcing, please see article 4.5.5.2 of this book.

With information technology, many types of tasks that are traditionally performed in-house can now be outsourced (Karmarkar and Apte [19]). Indeed, not only is outsourcing here to stay, it is also set to grow. Those projected to grow the fastest in 2008-2009 – financial, legal, supply chain, medical, and news services – are all service functions (Brown and Wilson [6]). A Gartner Group study has estimated that the overall global outsourcing market will grow by 8.1% in 2008 (Gartner [14]). Anand and Aron [3] cite external

*University of Washington, Foster School of Business

†Boston University School of Management

studies that project that the off-shore Business Process Outsourcing (BPO) industry will grow from \$123.6 billion in 2001 to over \$230 billion in 2015. A big portion of the service outsourcing activities are service-related. It is estimated that in the total US GDP, roughly 84% of value-added activities are service-related (Karmarkar and Apte [19]). Service outsourcing, therefore, plays an important role in the current and future US economic activities.

From a research perspective, however, OM research on service outsourcing is less visible, compared with research in manufacturing and physical systems. In this context, we review OM research in the field of service outsourcing, and then discuss some future research opportunities. In particular, we focus on the following three research areas:

(i) **Capacity planning and supplier coordination in service outsourcing.** Probably the most prominent component in service outsourcing is to plan for adequate capacity. This can be difficult considering the uncertainty involved in a service system, and the heterogeneity in customer requirements and server capabilities. In addition, it requires coordination with the service provider, which in turn is dependent upon economic incentives. We will review papers in both areas, i.e., those that analyze capacity systems and aim to improve capacity planning methods, and those focus on incentives and coordination mechanisms.

(ii) **Service outsourcing under information asymmetry.** A frequent problem companies face in service outsourcing is the fact that they lack perfect information about their potential outsourcing partners. Recent researches are fruitful in dealing with such a problem, which we will review in depth.

(iii) **Quality concerns in service outsourcing.** Recently more and more attention has been spent on the potential pitfalls in service outsourcing. Quality concern is frequently cited as one of the greatest issues in outsourcing. We review researches that study this problem and highlight their findings that can help companies avoid the ‘quality trap’ and maintain a high level of service quality when outsourcing.

Before we delve into reviewing individual papers, a note on the distinction between outsourcing and offshoring. Offshoring refers to a change in location of service facilities, and does not necessarily entail sourcing from outside vendors. For example, companies may choose to set up and operate their own customer service centers in some other foreign countries. In comparison, outsourcing means contracting out business processes to outside supplier, which may not be located offshore. Please refer to article 4.5.5.2 of this book for more discussions and examples distinguishing the two terms.

2 Literature Review

2.1 Capacity Planning and Supplier Coordination

Capacity planning is a perennial topic in the OM service literature, and it’s no exception in the OM outsourcing literature. In any outsourcing configuration, the capacity (and other things like effort) at both parties need to be planned and coordinated. This is very often done via a contract mechanism. In this part, many papers have to do with contracting issues in the context of call center outsourcing, but they all use

different terminologies. To expedite the discussion, it is necessary to first introduce some common notations.

In a simple one-to-one call center outsourcing setting, two parties are involved: a firm (which we will call the *user*) sends all or some of its customer calls (or emails) to an outside call center (which we will call the *vendor*). The vendor hires and trains its own agents to handle the calls, and each successfully handled call provides a direct benefit to the user firm (direct revenue or customer satisfaction). The contract then determines how much payment should be made from the user to the vendor. At the least a typical contract specifies scope of work, deliverables, terms and conditions, and service level agreement (SLA) (Engelke [10]). Transfer payment can be calculated in a variety of ways. For example, a typical call center outsourcing contract involves at least some forms of pay per time (PPT) or pay per call (PPC) (Gans and Zhou [12]). In the former arrangement, the user pays the vendor based on the agent hours that are spent. This is also referred to as “pay-for-capacity” (Akşin et al. [2]). In the latter arrangement, the payment is based on the number of calls that are handled. This is also referred to as “piecemeal” (Ren and Zhou [27]) or “pay-for-job” (Akşin et al. [2]). In practice, these contracts are often combined with other incentives or penalties based on call resolution (Ren and Zhou [27]), waiting time, or SLA achievement (Hasija et al. [16] and Lee et al. [22]).

Ren and Zhou [27] and Hasija et al. [16] are concerned with the design of a coordinating contract. A contract is said to be *coordinating* if it achieves two goals in the decentralized setting: (1) it induces the vendor to choose system-optimal staffing and effort level (i.e. $s^v = s^I$ and $e^v = e^I$) and (2) it allows for an arbitrary split of system profit between the user and the vendor. With a coordinating contract both parties can first achieve the highest total profit, and then share it in such a way that each is better off than under a non-coordinating contract.

Consider the following model of such a simple one-to-one call center outsourcing supply chain. The vendor is modeled as a multi-server queueing system with customer abandonment. Arrival rate is λ , service rate is μ , and customer patience (i.e. time waiting in the queue until abandonment) has continuously differentiable PDF f and CDF F . The staffing cost rate is c_s , the customer waiting cost rate is c_w , and each time a customer abandons, there is a cost of c_a . Of the served calls, only a portion, p , are satisfactorily resolved. Each resolved call results in a revenue of (representing direct sales, customer satisfaction, etc.), but incurs a loss of goodwill c_g for each resolved call (representing the reduction in both immediate and future sales). The resolution rate p is influenced by the agents directly, but ultimately determined by the vendor’s effort level which includes agent training, infrastructure development, technical support, etc. The vendor’s resolution $p(e)$ is assumed to be concave, increasing in its effort level e . The vendor bears the effort cost at a rate of c_e .

Therefore, the vendor must make two decisions: staffing level and effort level, and the user must determine the appropriate terms for the contract to induce the vendor to make system-optimal decisions.

The vendor’s $G/GI/s+GI$ queueing model is hard to analyze. Even a simpler $M/M/s+M$ system is hard to analyze (see Garnett et al. [13]). Ren and Zhou [27] follow a fluid approximation approach, which is shown to be remarkably accurate by Whitt [29]. In the fluid approximation, if the vendor’s staffing level is s then the system throughput rate, abandonment rate, and steady-state wait time are, respectively:

$T(s) = \min(\lambda, \mu s)$, $L(s) = \lambda - T(s)$, and $W(s) = \frac{L(s)}{f(0)\lambda}$ assuming $f(0) \neq 0$.

In the centralized (i.e. first-best) setting, the objective is to maximize total system profit:

$$\max_{s,e} \pi^I(s, e) = rp(e)T(s) - c_s\mu s - c_e e - c_a L(s) - c_w\lambda W(s) - c_g(1 - p(e))T(s), \quad (1)$$

and the optimal decisions are: $s^I = \lambda/\mu$ and $p'(e^I) = \frac{c_e}{(r+c_g)\lambda}$, if $\frac{c_e}{(r+c_g)\lambda} < p'_0$; $e^I = 0$ otherwise.

In the decentralized outsourcing supply chain, the user offers the vendor a contract to take all of its calls. The vendor decides its staffing and effort levels, and depending on the service outcome (e.g. T , W , L , etc.), it receives a payment Ψ from the user as specified in the contract. The user receives a revenue r from each served and resolved customer. Because the vendor is invisible to the customers, user also bears all the costs of customer abandonment, customer waiting, and unresolved calls. In such a setting, the vendor and user maximize their respective individual profit π^v and π^u :

$$\begin{aligned} \pi^v(s, e) &= \mathbb{E}(\Psi) - c_s\mu s - c_e e, \\ \pi^u(s, e) &= rp(e)T(s) - c_a L(s) - c_w\lambda W(s) - c_g(1 - \bar{p}(e))T(s) - \mathbb{E}(\Psi). \end{aligned}$$

Ren and Zhou [27] find that a simple PPC contract, where $\Psi = bT(s)$, can coordinate the vendor's staffing level (i.e. $s^v = s^I$), but it fails to coordinate the vendor's effort level. In fact, the vendor will exert no effort (i.e. $e^v = 0$) at all, because it pays for all the effort cost but receives no accordant benefit. A Pay-Per-Call-Resolved (PPCR) contract, where $\Psi = bp(e)T(s)$, also coordinates the vendor's staffing level (i.e. $s^v = s^I$). Although it lifts the vendor's effort level, it still falls short of the system-optimal effort level (i.e. $0 < e^v < e^I$). This is intuitive because while PPCR rewards the vendor for its effort, the benefit resulting from a higher effort level is shared between the user and the vendor. On the other hand, the vendor bears the total cost of effort. This incentive misalignment is similar to the "double marginalization" observed in physical-goods supply chain. To remedy this, Ren and Zhou [27] propose two additional contracts – pay-per-call-resolved plus cost sharing (PPCR+CS) and partnership (PART) – that are shown to coordinate the supply chain (i.e. $s^v = s^I$ and $e^v = e^I$). Under the PPCR+CS contract, $\Psi = bp(e)T(s) + (1 - \alpha)(c_s\mu s + c_e e) - \alpha \left[c_g(1 - p(e))T(s) + c_a L(s) + \frac{c_w}{f(0)}L(s) \right]$. It basically lets the user share the vendor's costs so that overall the vendor's incentive is proportional to the supply chain's. Furthermore, the vendor's share of the total profit is b/r . By varying b , the two parties can achieve a desirable share of the total profit. While PPCR+CS requires a lot of information sharing which may be difficult to do practically, PART requires less information sharing: the vendor pays the user for each customer arrival but gets to keep all the revenue. In addition, the user shares part of the vendor's staffing cost and gives it a fix effort compensation based on the system optimal e^I : $\Psi = rp(e)T(s) - (1 - p(e))c_g T(s) - (1 - \alpha)[rp(e^I) - c_g(1 - p(e^I))]T(s) + (1 - \alpha)c_s\mu s + (1 - \alpha)c_e e^I - \alpha \left(c_a + \frac{c_w}{f(0)} \right) L(s)$. In essence, PART has the flavor of a franchise arrangement.

Instead of the fluid approximation, Hasija et al. [16] use an $M/M/N + M$ queueing system to model the vendor's call center and find that neither a PPC nor a PPT contract can coordinate the vendor's staffing

level. This difference is due to the fact that while it is advantageous to be at 100% utilization level in a fluid system, in a queueing system, system utilization must stay strictly below 100% and customer abandonment is convex increasing in the utilization. However, Hasija et al. [16] show that by adding additional penalty for violating service level agreements the enhanced PPC and PPT contracts can coordinate the outsourcing chain.

In addition, Hasija et al. [16] study the effect of information asymmetry. Specifically, the vendor's effort level is private information, and as a result the user doesn't know whether the vendor has a high or low productivity (i.e. service rate). In such a case, an optimal contract will have to additionally induce the vendor to self-select and reveal its productivity type. The authors show that coordinating contracts in the full-information case may no longer be optimal. Instead, the authors propose several specific combinations of PPT and/or PPC contracts with additional constraints (e.g., average service time limit, SLAs, and /or wait time penalty) and show them to be optimal. In addition to achieving maximum profit for the system, the authors also discuss how profits can be allocated between the user and the vendor under such contracts.

Milner and Olsen [24] demonstrates the importance of selecting the right SLA and setting the right parameter. It studies a call center outsourcing situation where the vendor handles work not only from the user under a contract, but also from other non-contract sources. The vendor call center system is modeled as a multi-server queue and the user requires the vendor to satisfy a constraint on delay percentile (i.e. the percentage of delay that exceeds a limit). To the extent that the vendor has more capacity than necessary, it will also take non-contract work to increase revenue. The problem for the vendor then becomes how to prioritize calls from these two different sources: the vendor wants to minimize average wait time for the non-contract jobs, subject to the delay percentile constraint for the contract jobs. Using the heavy traffic regime, where system load approaches 1, as an approximation to the original system, the author shows that it is optimal to use a two-threshold policy where contract jobs are given priority only if the queue length is neither too long nor too short. This result has implications for both the user and the vendor. For the user, even though the delay percentile constraint in the contract is satisfied, the two-threshold policy could seem unfair because contract jobs will be given lower priority when the queue is long. The authors then suggest that the user should add more constraints to the contract, or impose a convex penalty on excess delay (which makes it optimal for the vendor to use a fairer single-threshold policy). Several other contract forms were also studied and found inadequate.

While these three papers study contract issues in a complete outsourcing setting, where the vendor handles all of the user's customer calls, Akşin et al. [2] study contract issues in a co-sourcing setting where the user outsources only some, but not all, of its calls to the vendor, for strategic (customer segmentation) or operational (call volume uncertainty) reasons (this term is also described in Willcocks and Lacity [30]). Co-sourcing is often promoted as an alternative to complete outsourcing (Fuhrman [11]). Demand at a call center – the call volume – is uncertain and fluctuating in nature, as is the demand at most services. The forecasted call volume pattern can be broken into two parts – the base volume (which is a stable volume

over time) and the fluctuation (anything above the base). It is easy and cost-effective to staff to the base call volume. However, to achieve any reasonable service level, call centers must plan extra staffing to deal with the fluctuation. In a co-sourcing setting, an important question becomes which contract type – contracting the base or contacting the fluctuation – is a better one. Specifically, in the former setting, the user’s would staff just enough to deal with the stable base demand while the vendor would have to carry safety capacity to handle the fluctuation. (Since the vendor is paid by the volume, this contract takes the form of PPC.) In the latter setting, the vendor is responsible for the base and the user handles any fluctuation above the base. The vendor therefore will set a stable staffing level to handle the base and be compensated by PPT.

On the surface the user should prefer to keep the base because of the lower in-house staffing cost it entails. However, the vendor’s optimal charge will be different in the two settings. Overall, Akşin et al. [2] are able to characterize the optimal capacity levels and partially characterize the optimal prices under both contract types. It turns out that because the vendor optimally charges different prices under different contracts, for the user neither contract type dominates. The contract choice depends on the call fluctuation patterns in addition to the revenue and cost parameters.

Two more papers, Gans and Zhou [12] and Koçağa and Ward [20], deal with the co-sourcing arrangement, but they eschew the contracting issues and focus on real-time call routing decisions instead.

Gans and Zhou [12] considers the operational-level coordination issues in a call center co-sourcing arrangement where calls of high value are often kept “in-house” by the user and the low-value calls are outsourced. Since the customer service representatives often receive nested skill training, those handling high-value calls are also capable of handling low-value calls. This presents the user a chance to use idle in-house, high-value representatives to handle low-value calls opportunistically. This helps the user firm to lower the volume of outsourced calls and hence cost. The authors present four coordination schemes differing in their routing sophistication and coordination/communication requirements. Assuming identical service time distribution of low and high-value calls, they show that a relatively simple routing scheme can perform nearly optimally in large systems. This scheme also has the added benefit of not requiring extensive coordination between user and vendor firms on a real-time basis. In small systems, however, it is necessary to use a threshold type routing policy – using in-house high-value representatives to take low-value calls only when idle capacity exceeds a certain threshold.

In Koçağa and Ward [20], the user operates its own call center but when the call center is congested, customers abandon due to long waits. In such cases it is worthwhile to route calls to an outside vendor. A natural tradeoff occurs between call abandonment and outsourcing cost. The paper formulates a Markov Decision Process model and shows that a queue-length based threshold policy is optimal. Because this policy is complicated, the authors also propose a diffusion approximation and calculates the optimal threshold. Numerical analysis reveals that the diffusion approximation is very accurate and optimally utilizing the outsourcing option can yield significant cost savings.

While all the papers discussed so far deal with one-level service processes, Lee et al. [22] study contract

issues in the outsourcing of a two-level service process. When all the customers first arrive, they are handled by first-level agents (who are called “gatekeepers”). Because customers are heterogeneous, the agents first diagnose the difficulty of each request. When the difficulty level is below a threshold k the gatekeepers further proceed to treat these customer requests. This is successful only with probability $F(k)$. Customer requests with difficulty level above k are passed on (along with those unsuccessfully treated by the gatekeepers, called “mistreatment”) to a second-level group of agents (called “experts”). The experts then successfully handle all the requests. Staffing, as well as customer waiting, on both levels incur costs. The mistreatment by the gatekeepers also incurs a further cost. Using a Halfin-Whitt approximation ([15]) to staff both levels, the authors find the optimal first-best staffing levels and the optimal threshold k^* . They then show that 1) when only the expert level process is outsourced, a PPC+WP contract (where WP indicates additional linear penalty for waiting time) coordinates the systems, 2) when only the gatekeeper level process is outsourced, a PPC+PPT+WP contract coordinates the system, and 3) when both levels are outsourced, a PPC+PPT+WP+ST contract (where ST indicates additional linear payment for successful treatments) coordinates the system *only if* system parameters satisfy certain condition.

2.2 Service Outsourcing under Information Asymmetry

Both parties involved in an outsourcing arrangement do not share with each other all of their private information. How does this information asymmetry affect the contract design, and what’s the impact of such information asymmetry? Four papers address this question.

As we previously reviewed, in Hasija et al. [16] the vendor’s effort level is private information not shared with the user. The authors show that to achieve coordination, the contract in this setting is more complicated than that with full information. Additional SLAs or penalties must be added and only particular combinations of PPC and PPT contracts can coordinate.

Akan et al. [1] examine call center outsourcing contract design under a different kind of information asymmetry: the user firm has private information about customer call demand (high or low) while the vendor only has a prior belief of its distribution. When it is the vendor who offers the contract, the authors show that the optimal screening contract offers too low an abandonment probability to the high type user (as compared with the first best) but offers no such distortion in the abandonment probability to the low type user. This result contradicts the “efficiency at top” results in the literature on monopolist screening, and is largely due to the use of queueing systems. When it is the user who offers the contract, the low type will offer the same abandonment probability as in the first best case, but the high type offers an abandonment probability below the first best, and has to pay more for each call. These distortions, however, can be mitigated and even removed by a two-part tariff contract – in addition to the per call charge, the user also makes a fixed payment to the vendor.

Ren and Zhang [26] examine call center outsourcing contract design when the user does not know the the vendor’s capacity cost or quality cost. In many situations, when a vendor’s service quality is high – be it

service speed, or problem resolution rate – it tends to be more expensive as well. In the call center industry, for instance, labor costs are typically lower for call centers located overseas, compared to those located in the US. However, the service quality for domestic call centers is often much higher. In some situations, the two costs may be positively correlated. For example, successful quality improvement initiatives may simultaneously bring down the capacity and quality costs of a firm. Ren and Zhang [26] study service outsourcing where the vendor exhibit such correlations. In their model, a user plans to outsource with an outside vendor, but it has no information about the vendor’s capacity cost or quality cost (i.e., how hard it is for the vendor to reach a certain service level). However, the user does know the correlation between the two costs. Such information is usually obtainable through market research. The authors show that the form of the optimal outsourcing contract depends critically on the relationship between capacity and quality costs. They also study the effectiveness of simple contracts, contracts that are not contingent upon information unknown to the user. They find that those non-contingent contracts perform well when the correlation between capacity cost and quality cost is positive, but they can perform much worse when the correlation is negative.

Finally, Plambeck and Zenios [25] study a dynamic principal-agent model where the principal outsources work to the agent but the agent’s effort is unobservable. Their model can be motivated by the so-called “maintenance problem”. A machine generates profits, but its owner delegates its maintenance to a manager, who has to repair the machine when it breaks down. The owner cannot observe the manager’s effort in repairing the machine, and can only observe if the machine is functioning or not. Therefore the owner’s problem is to design a performance-based contract that will motivate the manager to take actions that maximizes the owner’s profits over a finite horizon. Their model is applicable to similar situations in service outsourcing, where the actions of the outsourced services are not observable, and only their outcomes are observed.

2.3 Quality Concerns in Service Outsourcing

One of the key considerations in service outsourcing is assuring high service quality by the vendor. As we stated in the introduction, increasingly firms are focusing on the quality benefits of outsourcing rather than the cost benefits. Three papers deal with the quality aspect of outsourcing.

As we reviewed before, Ren and Zhou [27] examine how contracts can be used to effectively ensure optimal service quality level by the vendor. They show that simple contracts, even the pay-per-call-resolved contract that takes quality into consideration, do not coordinate the vendor’s quality training effort level. Instead, more complex contract are needed. In this paper, quality is measured by the call resolution rate achieved by the vendor.

Benjaafar et al. [4] study how firms should go about outsourcing to their vendors with service quality considerations. Below, we describe their model in a service outsourcing setting, even though it applies to both service and inventory systems. They compare two scenarios. The first is to allocate demand among a

set of vendors, but the proportion of demand that each vendor gets is an increasing function of the service level of each vendor. Here service level is measured by customer waiting time, abandonment probability etc. They call this approach supplier-allocation (SA) approach. The other, called supplier-selection (SS) approach, uses a probability to select one vendor. The probability of being selected is an increasing function of service level committed by the vendor. They find that often using just one vendor (i.e., the SS approach) is advantageous, but such advantage depends upon the form of cost functions of serving customers incurred by the vendors. Moreover, by selecting an appropriate allocation mechanism, the user can indeed induce a higher service level provided by its vendors.

Cachon and Zhang [8] study a queueing model with one user and two vendors. The user could orchestrate a service competition by allocating demand based on the vendors' performance, which is measured by speed of service. But how? The authors consider a variety of allocation policies, starting from state-independent policies that minimizes the user's lead time given fixed capacities and non-strategic vendors. They then move on to more complex state-dependent policies, and find that the variation in performance across different allocation policies is wide. Moreover, they show that there exists an allocation policy that can induce maximum vendor speed and at the same time minimize the user's lead time. Therefore there need not be a trade-off in incentive and efficiency in outsourcing.

In the ride share model studied by Li et al. [23], service level is measured by percentage of demand that is denied. BostonCoach Inc. is a premier provider of executive sedan, limousine and event transportation services. Not unlike the call centers, it has a relatively inflexible fixed capacity, yet its demand is random and fluctuating: in the Boston area, on a typical day, only about 85% of its demand is booked one day in advance. If BostonCoach projects to have insufficient capacity, it can subcontract with its collaborating companies (thus this is a co-sourcing arrangement), but this has to be done one day ahead. That is, the overall capacity (own capacity plus contracted capacity) must be decided one day ahead. In real time, if, on the day of service, demand is much higher than expected, then requests will be denied, resulting in significant ill-will cost. The authors develop a model that incorporates various demand features (origin-destination, time window, specific vehicle and/or driver request, and trip type) and capacity constraints (driver availability, shift constraints, driver/vehicle pairing, etc.). A math program is devised to minimize total cost and a column generation method is adopted to solved it. Results show that using the proposed method can significantly improve the productivity of existing fleet. This results in reduced percentage of denied services, reduced use of drivers, and reduced use of contract services.

3 Conclusion and Future Research Directions

In this article, we aim to give an overview of the current status of OM research in service outsourcing. Currently the number of papers in this area is not large but we expect the area to grow over time.

Several areas are promising for future research:

- Outsourcing vendor selection is a difficult but important topic. Due to the uniqueness of services being outsourced and the difficulty of measuring and monitoring various aspects of service quality, switching vendors is a process that involves significant cost and time. The current service outsourcing contracting papers have focused on one-shot games. Clearly, it would be interesting and helpful to study the use of repeated-game models and other more complex contract structures (for an example of the use of repeated game in a supply chain context, see Plambeck and Taylor [28]).
- Most services are complex, multiple-stages processes. Outsourcing decisions do not have to be all-or-nothing. Two directions are worth pursuing: 1) The user can choose to outsource some stages of the service process while keeping the rest in house, but the current papers, with the exception of Lee et al. [22], are all concerned with the outsourcing of a single-stage process. The choice of what to outsource, how to outsource, and how to coordinate the stages that are outsourced with those that aren't, are all interesting areas for future research. 2) Even after the user has decided what to outsource, there is still the question of how much to outsource. Co-outsourcing, as we review above, is a useful arrangement. More research is needed to shed light on its effective use.
- Other types of information asymmetry, such as private demand forecasts, are worth studying. In service outsourcing, the vendor often has to make capacity decision with only a rough estimate of future demand, while the user may have better information. Hence, depending on specific situations the user may have an incentive to over-forecast demand to the vendor, causing costly capacity over-investment. How to induce truthful behavior from the user and achieve the maximum system efficiency is an interesting topic. Similar problems have been studied in the context of supply chain management. Interested readers are referred to Cachon and Lariviere [7].
- On a more technical level, the difference in the results obtained by Ren and Zhou [27] and Hasiija et al. [16] are due to the different queueing models they choose (the fluid model in Whitt [29] and the diffusion model in Garnett et al. [13]). Therefore, much work still needs to be done to understand the impact of using different queueing models and approximation methods. Interested readers are referred to Ward Whitt's and Avi Mandelbaum's research webpages¹.

Since there already exists a large literature on Information Technology Outsourcing (ITO), it is also beneficial for OM researchers to learn about the other aspects of outsourcing that we have not covered in this article. Dibbern et al. [9] gives an excellent and comprehensive review of the IS outsourcing literature. We have touched upon several relevant topics in this article, but interested readers are highly encouraged to read [9]. Although they only review the ITO literature, we believe the methodologies and frameworks apply to other outsourcing settings, including service outsourcing.

¹<http://www.columbia.edu/~ww2040/allpapers.html> and <http://iew3.technion.ac.il/serveng/References/references.html> respectively.

Finally, we fully expect research in the areas of personal, legal, medical, and other professional service outsourcing to grow rapidly, along with the increasing use of outsourcing in these areas in practice.

References

- [1] Mustafa Akan, Baris Ata, and Martin A. Lariviere. Asymmetric information and economies-of-scale in service contracting. Working paper, COSM-07-001, available at <http://www.kellogg.northwestern.edu/faculty/lariviere/research/CallCenterContract.htm>, 2007.
- [2] O. Zeynep Akşin, Francis de Véricourt, and Fikri Karaesmen. Call center outsourcing contract analysis and choice. *Management Science*, 54(2):354–368, 2008.
- [3] K. Anand and R. Aron. Quality, incentives and inspection regimes in offshore service production: Theory and evidence. Working paper, available at http://digital.mit.edu/wise2006/papers/7B-2_Anand-Aron.pdf.
- [4] S. Benjaafar, E. Elahi, and K. Donohue. Outsourcing via service quality competition. *Management Science*, 53(2):241–259, 2007.
- [5] Douglas Brown and Scott Wilson. *The Black Book of Outsourcing*. John Wiley & Sons, Inc., New Jersey, 2005.
- [6] Douglas Brown and Scott Wilson. *2008 Black Book of Outsourcing: State of the Industry Report*. Brown-Wilson Group, Inc., 2008.
- [7] G. Cachon and M. Lariviere. Contracting to assure supply: how to share demand forecasts in a supply chain. *Management Science*, 47(5):629–46, 2001.
- [8] G. P. Cachon and F. Zhang. Obtaining fast service in a queueing system via performance-based allocation of demand. *Management Science*, 53(3):408–420, 2007.
- [9] Jens Dibbern, Tim Goles, Rudy Hirschheim, and Bandula Jayatilaka. Information systems outsourcing: A survey and analysis of the literature. *The Database for Advances in Information Systems*, 35(4), 2004.
- [10] William D. Engelke. Writing an outsourcing contract. Available at <http://www.hsv.com/writers/engel/outsou3.html>.
- [11] F. Fuhrman. Co-sourcing: A winning alternative to outsourcing call center operations. *Call Center Solutions*, 17(12):100–103, 1999.
- [12] Noah Gans and Yong-Pin Zhou. Call-routing schemes for call-center outsourcing. *Manufacturing and Service Operations Management*, 9(1):33–50, 2007.

- [13] O. Garnett, A. Mandelbaum, and M.Reiman. Designing a call center with impatient customers. *Manufacturing & Service Operations Management*, 4(3):208–227, 2002.
- [14] Gartner Group. Gartner says worldwide outsourcing market to grow 8.1 percent in 2008. Available at <http://www.gartner.com/it/page.jsp?id=578307>.
- [15] S. Halfin and W. Whitt. Heavy-traffic limits for queues with many exponential servers. *Operations Research*, 29(3):567–588, May–June 1981.
- [16] Sameer Hasija, Edieal J. Pinker, and Robert A. Shumsky. Call center outsourcing contracts under information asymmetry. *Management Science*, 54(4):793–807, 2008.
- [17] Mary M. Dickens Johnson. Current trends of outsourcing practice in government and business: Causes, case studies and logic. *Journal of Public Procurement*, 2:248–268, 2008.
- [18] N. Kakabadse and A. Kakabadse. Critical review – outsourcing: A paradigm shift. *Journal of Management Development*, 19(8):670–728, 2000.
- [19] U.S. Karmarkar and U.M. Apte. Operations management in the information economy: Information products, processes, and chains. *Journal of Operations Management*, 25(2):438–453, 2007.
- [20] Yaşar Levent Koçağa and Amy Ward. Dynamic outsourcing for large-scale service systems. Working paper, 2009.
- [21] Neal Leavitt. The changing world of outsourcing. *Computer - IEEE Computer Society*, 40(12):13–16, 2007.
- [22] Hsiao-Hui Lee, Edieal Pinker, and Robert Shumsky. Outsourcing a two-level service process. University of Rochestre, Available at http://www.hsiao-huilee.com/My_Research.html, 2009.
- [23] Yihua Li, Xiubin Wang, and Teresa M. Adams. Ride service outsourcing for profit maximization. *Transportation Research Part E*, 45:138–148, 2009.
- [24] Joseph M. Milner and Tava Lennon Olsen. Service-level agreements in call centers:perils and prescriptions. *Management Science*, 54(2):238–252, February 2008.
- [25] E.L. Plambeck and S.A. Zenios. Performance-based incentives in a dynamic principal-agent model. *Manufacturing & Service Operations Management*, 2:240–263, 2000.
- [26] Z. Justin Ren and Fuqiang Zhang. Service outsourcing: Capacity, quality and correlated costs. Working paper, available at <http://apps.olin.wustl.edu/faculty/zhang/>, 2009.
- [27] Z. Justin Ren and Yong-Pin Zhou. Call center outsourcing: Coordinating staffing level and service quality. *Management Science*, 54(2):369–383, 2008.

- [28] T.A. Taylor and E. L. Plambeck. Supply chain relationships and contracts: The impact of repeated interaction on capacity investment and procurement. *Management Science*, 53(10):1577–1593, 2007.
- [29] W. Whitt. Fluid models for many-server queues with abandonments. *Operations Research*, 54(1):37–54, 2006.
- [30] Leslie P. Willcocks and Mary C. Lacity. *Strategic Sourcing of Information Systems: Perspectives and Practices*. John Wiley, 1998.