

Quantifying the Tradeoffs Between Cost & Quality For Systems Service Support

Dr. Jess S. Boronico, Management, Monmouth University
Mr. Andre Zirkler, Unixpros Incorporated, Eatontown, New Jersey
Dr. Philip H. Siegel, Accounting, Monmouth University

Abstract

Rapidly increasing computer system's complexity has caused many companies to increase awareness of customer services. Many firms in the computer hardware and/or software industries have devoted increased effort towards customer service through the development of help desk support systems. Help desk support systems development currently face a number of problems, including escalating growth in customer usage, high staff turn-over, maintaining service quality and controlling costs. Industry has responded by providing additional staff training, increasing systems automation, and introducing knowledge based tools. These actions have increased help desk efficiency, but fail to provide a quantitative means for optimizing help desk operations. This paper addresses this problem through the development of a mathematical model which minimizes expected help desk costs by considering both operational costs and social costs. This model provides a basic framework through which policy makers may analyze the effectiveness of capacity decisions as they apply to help desk support systems in a multi-echelon, networked service station.

Introduction

Help desks provide product or service support to customers through the telephone. Help desks have long been used by public utilities and government agencies, but only recently have they been introduced to the computer hardware and software sector (Etchinson, 1994). Help desks consist of a hierarchy of support levels arranged in tiers. There are often two, and sometimes three tiers of support. Tier one consists of call screening agents, who collect information, direct calls and answer questions. Tier two personnel consist of generalists, who have a broad scope of knowledge in an area such as hardware or software technology¹. Call screening agents usually successfully respond to 10% to 15% of incoming calls, with 70% of remaining calls resolved by tier two agents. Those calls not resolved at tier two are passed on

to tier three personnel (Etchinson, 1994). If a call passes through tier three without resolution, it is handled by members of the product development team (whose primary professional activity does not involve responsibility to help desk support systems). As calls progress amongst the tiers, the average response time increases dramatically; tier two calls average 90 seconds, tier three inquiries range from 7 to 10 minutes, and product development team inquiries may require 1 or more hours.

The major problem that help desks face today is rapid growth in the demand for support due to the increased utilization of personal computers. Hunse (1994) reports that between 1993 and 1994, help desk call volumes have increased 25 to 30%. Greater than 80% of customer support

calls concern product usage rather than product defects. Mason (1990) found that since users disregard instruction manuals, support hotlines are increasingly being used as tutorials. Therefore, both the volume and the nature of incoming calls are contributing to the declining ability to provide consistent levels of support.

Even though help desk demand has grown since 1993, Hunse (1994) found that the mean personnel staffing level has decreased by 8% over the same period (from 13 to 12 people) which has led to excessive staff turnover. Mason (1994) concluded that yearly turnover rates for help desk personnel has increased to greater than fifty percent annually, with a mean turnover period approximately 18 months per employee. Bisenuis (1992) suggested that personnel require recovery time to perform non-support related tasks.

The increase in help support services has generated a rapid increase in operational costs. Historically, these product support services have been provided free of charge. Most companies have since abandoned free support policies in favor of service for fee policies. Johnson (1994) warns that the computer industry can either provide quality support, or they can provide free services. It is difficult to provide both. Beame & Whiteside (Editorial Staff, 1994), an outsource support corporation, has addressed the cost/quality issue by developing a cost structure for premium support, in which 96.7 percent of all incoming requests are resolved within two hours. The cost to the consumer for this service is ordinarily fifteen percent of the software package's value.

Due to the increased emphasis on service, "product support" has become a major focus for continued future growth in the computer industry. Quality support provides competitive advantages, and opportunities for recurring revenues (Mason 1990; Frye, 1994). According to Service News (1994), most people switching software products cited unsatisfactory product support and service as a primary determinant. In addition, lack of product support could lead to many PC users switching computer brands due to service dissatisfaction. Companies are finding it increasingly difficult to

differentiate their products from others, and are turning more frequently to customer support to provide a potential competitive edge. However, this rapid growth in support level demanded by clientele has brought about increased pressure on corporations to control rising service costs.

Two proposals focusing on investments in technology have been suggested to reduce help desk service costs. The first solution involves the utilization of software designed to aid in the support function, including automated call distributors. Frye (1994) reports that 39% of the support budget for twenty eight of the Fortune 1,000 firms is apportioned to decision support technologies, such as artificial intelligence and expert systems. The industry has recently experienced a 15% increase in the use of automated call distributors, as well as a 20% increase in the use of expert systems. The second solution involves the use of hypertext and AI systems, which despite their high cost are being used to offset the rising volume of help desk calls (Eskow, 1990). These increases in technological investments fail to address the problems of high turnover rate for support staff and to deal effectively with the cost/quality issue.

Daly (1994) cites the need to introduce and perform quantitative analyses for support operations. Through such analysis both cost control and service support growth can be more effectively. This paper provides a framework through which overall costs for managing a help desk support system can be analyzed. These costs consider both direct labor costs and consumer related costs, through which quality of service can be measured. The end result of this model will be the determination of the level of support to be implemented so that expected system costs (operational and quality) are minimized. The remainder of this paper is organized as follows: in section II the model is developed and solutions are determined; a policy analysis appears in section III; an illustrative example is presented in section IV, and section V provides concluding comments.

Model Development and Solutions

Customer satisfaction can be viewed as a

combination of two factors, the time required for service and the precision provided by the resolution. The time required for service can be partitioned into the time spent waiting for service and the time spent being served. While training and technology can assist in increasing precision and perhaps lowering service time, the determination of the number of support personnel is critical in managing the overall time spent in the system². The task of weighting and balancing operational costs and cost of quality requires an analytical framework which incorporates an element of uncertainty in both the arrival and service processes. The nature of this problem lends itself to a queuing analysis.

In modeling the queuing process of a help desk, certain assumptions must be made regarding arrival and service time mechanisms. If one assumes that arrivals occur at random over a specified period of time then the arrival process follows a Poisson process, which is presumed in the model presented. Other assumptions prohibit balking, renegeing and jockeying between queues. Balking and renegeing involve customers leaving the system while in and prior to joining the queue. Jockeying pertains to the switching queues while holding for service.

When considering the queue discipline and service mechanism, other issues emerge, including priority scheduling rules for incoming arrivals. These rules include those who have paid a premium for service, require special assistance, or are frequent users. Frequently help desks simply utilize a first come, first serve (FCFS) discipline. Larson (1992) refers to FCFS systems as socially just. The utilization of this discipline is well grounded in the literature, and will be assumed in the model presented. The standard assumption of exponential service times will be followed. Analysis of alternate service time distributions can be conducted via simulation, since limited analytical results are available for these distributions.

In evaluating service system configuration, one must consider whether arrivals will be filtered through a single or a multi-channel process. The channel selection decision effects how calls passing

through tier I will queue up at tier II. A single queue system is typified by an ATM machine (single server) or a bank counter (multi-server). An example of a multi-queue system is a supermarket checkout counter. The ordinary assumption of multi-queue systems is that the service mechanisms of parallel channels operate independently.

Generally, single queue systems are considered more efficient than multiple queue systems. For single queue systems, servers are never idle unless the system is empty. This is in contrast to a multiple queue system, in which customers may be queued in one line despite another server being idle³. The issue of single vs. multiple queue system efficiency is a continuing source of debate in the queuing community. While single queue systems are theoretically more efficient, some question whether this advantage is realistic. Young (1982) contends that behavioral issues ordinarily overlooked in mathematical analyses contribute to increased service quality and efficiency in multiple queue systems. He cites the example of the responsibility a server develops when he/she has a dedicated queue.

Whitt (1986), on the other hand, disagrees with this viewpoint and illustrates that there exist queue control algorithms which allocate arrivals so as to achieve single queue efficiency. Rothkopf and Rech (1987) contend that queue assignment policies exist that assign new arrivals to the queue with the shortest workload, which can achieve single queue efficiency.

There exist certain operational advantages in utilizing a multiple queue system, even if a simple allocation rule such as random assignment is used. First, by maintaining responsibility of queues at the server level a sense of responsibility is created. Second, there is a greater opportunity to establish a personal relationship with customers. For example, when a customer is forwarded to tier II, the tier I call-screening agent can announce the name of the tier II support person who will be handling their call. Based on this reasoning the model presented here will utilize a multi-channel structure. The assignment pattern for calls being trans-

ferred to tier II will be random since sequential allocation does not change the analysis. The simple allocation rule is used since it is call-screening agents are infrequently required to utilize any sophisticated assignment algorithm when transferring calls⁴.

Finally, the model will utilize a basic tenet of Jackson networks, namely, the input process for any node in the network is based on the output process of the preceding, or feeder node. Since all service times are assumed to be exponential, and the network is closed, the arrival stream at any tier must necessarily follow a Poisson process.

The help desk structure presented is similar to those found in Bisenius (1992) and Deavers-Claspell (1988). A single server at tier I handles all incoming calls, resolves a specified percentage of calls, and forwards the rest to one of two product support groups at tier II. For purposes of analysis, support group A is assumed to represent hardware support specialists. Support group B are software support specialists.

The objective of the model will be to determine the optimal number of servers in each support group at tier II, so that overall expected system costs are minimized. A single call-screening agent will support tier I, while multiple servers, each with his/her own queue will provide tier II support. The cost function includes both operational and social costs. Operational costs are based specifically on the cost for providing service, which is a function of the number of servers required. Costs associated with quality, or social costs, are based on the mean wait time in the system for arrivals, where arrivals may wait at tier I, tier II, or both. Consequently, the resulting objective function is given by the following:

$$\text{Minimize } E\{TC(A,B)\} = E\{C_1 + A \cdot C_A + B \cdot C_B + T \cdot C_S\} \quad (1)$$

where:

C_A = cost ([\$/[hr]) for hardware support personnel (tier II),

C_B = cost ([\$/[hr]) for software support person-

nel (tier II),

C_1 = cost ([\$/[hr]) for the call-screening agent (tier I),

C_S = social cost ([\$/[hr]) per person,

A = the number of hardware support personnel,

B = the number of software support personnel

T = the total time spent waiting for service accumulated over all arrivals

Distribution of the expected value operator is straightforward, however, for purposes of clarity, the determination of $E\{T \cdot C_S\} = C_S E\{T\}$ is provided. Total time spent waiting in the system⁵ is the sum of the waiting times in tiers I and II. Let λ represent the mean arrival rate to the system, Θ represents the percentage of calls which are forward to tier II⁶, and ρ the percentage of calls entering tier II which have been forwarded to a Type A specialist (hardware support personnel). The following mean arrival rates now apply: tier I (λ), tier II-hardware support ($\lambda\Theta\rho$), tier II-software support ($\lambda(1-\Theta)\rho$).

The expected time spent in the system for an arrival, $E\{T\}$, depends upon which tier the call is terminated or resolved. All calls must pass through tier I. The resulting mean time spent in the queue at this initial tier is standard, and given by:

$$W_{Q1} = \frac{\lambda}{\mu_1(\mu_1 - \lambda)} \quad (2)$$

where μ_1 represents the mean service time at tier I. It is assumed that all arrivals to the system are governed by the same service time distribution at tier I. That is, whether a call is resolved or not at tier I, the service time distribution is the same. The concept of a homogeneous service time distribution implies that the call screening agent attempts to resolve all incoming calls. However, if it appears that the call cannot be resolved by the call screening agent, the call is then forwarded to the next tier. To determine the total expected wait time for all arrivals, multiply (2) by the mean arrival rate at tier I, given by λ . The following obtains:

$$E\{T_1\} = \lambda W_{Q1} \tag{3}$$

For calls not resolved at tier I, an additional wait is to be expected at tier II. Recalling that (i) calls sent to tier II are sent to the appropriate area (hardware/software), (ii) calls are assigned randomly to the servers at that station and (iii) there are A hardware specialists and B software specialists at this tier, the resulting arrival rate at each type A and B channel is given by $\Theta\rho\lambda/A$ and $(1-\Theta)\rho\lambda/B$. Let μ_A and μ_B represent the mean service rates at each Type A and B channel⁷. The mean wait time in the queue for an arrival at Stations A and B can then be determined through the substitution of these values, together with the representative arrival rates, into (2). This yields, after simplification, the following:

$$W_{QA} = \frac{\Theta\rho\lambda}{A\mu_A^2 - \Theta\rho\lambda\mu_A} \tag{4}$$

$$W_{QB} = \frac{\Theta(1-\rho)\lambda}{B\mu_B^2 - \Theta(1-\rho)\lambda\mu_B} \tag{5}$$

where W_{QA} and W_{QB} represent the mean wait time for an arrival at stations A and B, in tier II, respectively. As (4) and (5) represent the mean wait time per arrival, the total expected wait time summed over all arrivals at these stations would be determined by multiplying (4) and (5) by their respective arrival rates. This yields the following:

$$E\{T_A\} = \Theta\rho\lambda \cdot W_{QA} \tag{6}$$

$$E\{T_B\} = \Theta(1-\rho)\lambda \cdot W_{QB} \tag{7}$$

The total expected wait time in the system is then found by summing (3), (6) and (7), yielding:

$$E\{T\} = \lambda W_{Q1} + \Theta\rho\lambda \cdot W_{QA} + \Theta(1-\rho)\lambda \cdot W_{QB} \tag{8}$$

Substitution of (8) into (1) yields the following objective function:

$$MinE\{TC(A,B)\} = C_1 + A \cdot C_A + B \cdot C_B + C_s \cdot \lambda \{W_{Q1} + \Theta\rho \cdot W_{QA} + \Theta(1-\rho) \cdot W_{QB}\} \tag{9}$$

In order to determine the optimal number of servers at tier II, it must be shown that (9) is convex. This convexity may be shown by noting that the Hessian matrix for (9), H_f , is positive definite provided that

$$A > \Theta\rho\lambda/\mu_A \text{ and } B > \Theta(1-\rho)\lambda/\mu_B .$$

This poses no complication, since these conditions must be satisfied in order for the queuing system to reach equilibrium and steady state. Consequently, for all feasible values of A and B , (9) is convex (and simultaneously continuous and twice differentiable). Additionally, $\lambda < \mu_1$ if the system is to achieve steady state.

To determine the optimal solution to (9), substitute (2), (4) and (5) into (9), and proceed to differentiate with respect to A and B , setting the resultants to zero. Solving simultaneously yields the following unique pair of critical values, which represent the cost-minimizing solutions to (9):

$$A^* = \frac{\Theta\rho\lambda}{\mu_A} \left(\sqrt{\frac{C_s}{C_A}} + 1 \right) \tag{10}$$

$$B^* = \frac{\Theta(1-\rho)\lambda}{\mu_B} \left(\sqrt{\frac{C_s}{C_B}} + 1 \right) \tag{11}$$

Both (10) and (11) satisfy the conditions to achieve steady state. Substituting (2), (4), (5), (10) and (11) into (9) yields the following minimum expected cost:

$$TC^* = C_1 + \Theta\lambda \left\{ \frac{C_A\rho}{\mu_A} \left(\sqrt{\frac{C_s}{C_A}} + 1 \right) + \frac{C_B(1-\rho)}{\mu_B} \left(\sqrt{\frac{C_s}{C_B}} + 1 \right) \right\} + C_s\lambda \left\{ \frac{\lambda}{\mu_1(\mu_1-\lambda)} + \Theta \left(\frac{\rho}{\mu_A} \sqrt{\frac{C_A}{C_s}} + \frac{(1-\rho)}{\mu_B} \sqrt{\frac{C_B}{C_s}} \right) \right\} \tag{12}$$

This analysis provides a framework through which policy analysis may be conducted. Policy issues are now discussed which deal with management's perspective on help desk support issues.

Policy Analysis

Three particular issues will be discussed: issues of technician costs and service rates, issues of distribution of incoming calls, and average costs and pricing implications. Following this discussion, a numerical example is provided which illustrates the methodology involved in the determination of the optimal solution and the implications discussed in this section.

Policy Issue #1: How do technician costs, accuracy, and proficiency effect total minimum expected costs?

Through observation of (12), it may be demonstrated in any of the cost parameters increase total minimum expected cost. The same effect is realized through either increasing the system mean arrival rate λ or the percentage of calls forwarded to tier II, Θ . Consequently, management needs to weigh the cost/benefit tradeoffs involved in decisions involving increases in accuracy and speed of service at the expense of higher wages. Tradeoffs must also be considered between the proficiency of handling calls at tier I (call screening agent), thus decreasing the rate of calls passed to tier II and the possible expense of higher wages paid to the call screening agent.

Conversely, an increase in any of the mean service rates decreases total minimum expected cost if wages are fixed. Higher proficiency in terms of decreased service time ordinarily implies higher wages (for more skilled employees). Therefore, tradeoffs between the net savings realized by increasing throughput must be weighed against increased wages. These are a direct consequence of upgrading the skill level of hot line personnel.

This policy issue can be illustrated by considering the case of an automated server. Presently, help desks offer such answering services as routing devices for incoming phone calls. The primary drawbacks involved in utilizing the automated server include customer impatience and dissatisfaction with the inability to speak with a "real" individual. The perceived wait time is associated with holding while listening to the automated

message as well as time spent waiting for a technician once the call has been channeled. On the other hand, automated routing devices are cheaper and have a shorter mean service time since no questions are resolved at this station. The automated routing devices decrease the mean queue length at tier I relative to that which develops when a call screening agent is present. Therefore, the tradeoffs between increased congestion at tier II and the increased number of tier II technicians required to meet the increase in demand at this tier must be carefully weighed against the potential savings at the tier I level.

Policy Issue #2: What is the effect on total minimum expected costs when the distribution of calls to stations A and B in tier II is altered?

Recalling that the percentage of tier II calls routed to station A is given by ρ , the conditions can be determined for which increases in ρ will increase total minimum expected costs. Differentiating (12) with respect to ρ , and setting the resultant greater than zero yields the following result: Total minimum expected costs increase as a function of ρ iff:

$$\frac{\sqrt{C_A}}{\mu_A} > \frac{\sqrt{C_B}}{\mu_B} \quad (13)$$

Therefore, increases in the percentage of calls forwarded to station A of those calls unresolved by the call screening agent increase total costs as a function of the service rates and costs of the tier II servers.

Equation (13) can be explained in more general terms by realizing that total expected minimum costs increase if the cost per service ratio is larger for a station A technician than that for a station B technician. This assumes that the hourly cost (C_A) for a technician is greater than 1.0. Since (13) is independent of ρ then total minimum expected costs vary linearly with respect to ρ . Whether minimum expected costs increase or decrease depends on the nature of the relationship given in (13). Although management may be con-

cerned if the percentage of calls moving to station A increases, invoking (13) allows management to determine if increases in demand at that station (when evaluated at optimum) may actually decrease total minimum expected costs.

Policy Issue #3: What is the average cost per call, and how might prices for service be affected?

It can be shown that the cost per call (at optimum) is convex with respect to the arrival rate. To derive this, divide (12) by the arrival rate λ , so the left hand side represents the cost per call. This left hand side is continuous and twice differentiable, and indicates that the second derivative with respect to λ is positive for all $\lambda > \mu_1$. Therefore, the total expected cost per call is convex with respect to the arrival rate. To determine the critical point which minimizes expected total costs, set the derivative equal to zero, which yields the following critical point:

$$\lambda^* = \frac{\mu_1 \sqrt{C_1}}{\sqrt{C_1} + \sqrt{C_s}} \quad (14)$$

Equation (14) has important policy implications. The equation suggests that some degree of scale economies exist in Help Desk services. If the pricing structure for help desk support is based on an average cost pricing mechanism, (14) shows that there exists an optimal arrival rate for which the price imposed on the end user is minimized. In order to achieve this minimum, additional expenditures for promotion and advertising may need to be initiated, and a cost/benefit analysis be undertaken.

Policy Issue #4: How can management effectively deal with staff burn-out?

Management is concerned that higher arrival rates will result in greater utilization of staff. Regardless of the overall mean arrival rate, the optimum solution suggests a fixed server utilization rate and a fixed mean wait time in tier II⁸. Since for any server at station A (station B could be utilized as well), the probability that a particular

technician is idle is given by

$$p_0 = 1 - \phi = \dots = 1 - 1 / \left(\sqrt{C_s / C_A} + 1 \right) \quad (15)$$

This result obtains through substitution of the optimal solution and rearrangement of terms, with $\phi = \lambda_A / A\mu_A$. This expression is independent of the arrival rate to the station, A or B. Similarly, it may also be shown that the mean wait time in any tier II queue is independent of the arrival rate. Substituting the optimal solution (10) into (4), following mean wait time at station A obtains:

$$W_A^* = \sqrt{C_A} / \left(\mu_A \sqrt{C_s} \right) \quad (16)$$

(16) is independent of the mean arrival rate λ . Regardless of the arrival rate, the optimal solution suggests that the mean queue length and idle time at tier II are fixed. Of great concern is the issue of whether this idle time is sufficient to prevent staff burn out. If not, management must alter the optimal solution so as to increase this idle time percentage.

Although management may intuitively feel that increasing the service rate at either station A or B, or both, might increase idle time, this is not the case. Equation (15) indicates that the mean wait time at a tier II service station decreases as the mean service rate increases. Equation (14) suggests that the idle time per server remains unchanged, as (14) is independent of the service rate μ_A . These observations suggest that the only decision variable in the problem which effects idle time is the cost parameter C_A , where decreasing this cost will increase the idle time. Decreasing the wage cost for the purposes of increasing idle time may be a contributing factor to employee dissatisfaction and staff turnover. An alternative strategy may involve moving away from the suggested optimal solution. This optimal solution, while minimizing total expected cost, does not explicitly consider such qualitative issues such as staff burn out or high turnover rate. In such cases, management would need to find a satisficing solution which achieves near-optimal minimum expected cost. An

illustrative example is now provided which demonstrates the optimal solution and highlights some of the policy implications discussed in this section.

Illustrative Example and Sensitivity Analysis

This case will focus on a help desk support station with two tiers. A singular call screening agent is available at tier one, while two stations, stations *A* and *B* are available at tier two. One station specializes in hardware support, while the other specializes in software. Calls are handled more efficiently if routed to the correct station, although each station may respond to all incoming calls. The case includes a number of parameters:

- The call screening agent receives a wage of $C_1 = \$20$ per hour; while technicians at tier II receive $C_A = C_B = \$30$ per hour.
- Arrivals enter the system at a mean rate of $\lambda = 80$ per hour.
- The call screening agent resolves 30% of all incoming calls, therefore, $\Theta = .70$.
- For those calls sent to tier II, 60 percent are routed to station *A*, therefore $\rho = .60$.
- Service times for both stations *A* and *B* are exponential with mean $1/\mu_A = 1/\mu_B = 3$ minutes⁹.
- Management assumes that the consumers' cost of waiting is given by $C_s = \$30$ per hour.

Tier II technicians are ordinarily paid higher wages than tier one due to the higher expected level of proficiency.

Table 1 shows optimal solutions as the mean arrival rate varies between 32 and 76 arrivals per hour. Several conclusions may be drawn from this table. The optimal solution suggests both constant server utilization and constant mean wait time at tier II. Management knows, therefore, that the degree of idle time allocated to each tier II server is independent of the arrival rate. As suggested earlier in the paper, increased idle time could lead to a reduction in staff burn out. If the optimal solution suggests that there exists insufficient idle time for tier II servers, then management

policy should provide ways to increase idle time.

The above discussion indicates that management could consider adding additional servers, which would lead towards a "satisficing" solution. Management would need to weigh these additional server costs against the added benefits of increased idle time which impacts turnover and service time¹⁰. Other intangible factors such as loyalty, dedication and interest cannot necessarily be mathematically modeled.

Management may also be interested in reviewing results as they apply to the average cost per arrival. It can be shown, that tier I wait times are convex and monotonically increasing as arrival rate increases. This is due to having only one call screening agent available with a known service time distribution. The significant increase in wait time at the first tier creates disproportionately large increases in the wait time (social) costs as the arrival rate approaches tier I capacity. Capacity is measured by the call screening agent's mean service time. Therefore, total costs increase as the mean arrival rate increases. However, management may not recognize that there exists an optimal mean arrival rate which minimizes the per unit cost of providing service. For the numerical example shown above, this optimal mean arrival rate equals 36 arrivals per hour. As mentioned earlier, if an average cost pricing mechanism were implemented, management would be able to minimize the price appropriated per call if the arrival rate could be maintained at this level.

A significant issue involves the decision of whether to institute an automated server or than a human call screening agent. Automated servers provide no support, but serve only to route calls to the appropriate channels. Service time and costs are (ordinarily) decreased when automated servers are utilized. Two disadvantages to instituting an automated server are increased congestion at tier II, since no calls are resolved at tier I, as well as additional consumer dissatisfaction at the impersonal nature of the automated service. Increased congestion at tier II increases operational costs due to additional servers required at optimal. The automated server case is summarized in Table 2.

Table 1.
Solutions Utilizing Human Call-Screening Agent For Various Arrival Rates

Arrival Rate Total/hour	Server Utilization		Mean Wait Time in Queue				Optimal Solutions			
	Tier I	Tier II	Tier I	Tier II		Servers		Minimum Expected Total Cost	Cost per Arrival	
				A	B	A	B			
32	.60	.52	.55	.01	.05	.05	1.29	.66	\$118.70	\$3.71
36	.55	.52	.55	.01	.05	.05	1.45	.74	132.74	3.68
40	.50	.52	.55	.01	.05	.05	1.61	.82	147.54	3.69
44	.45	.52	.55	.01	.05	.05	1.77	.90	163.34	3.71
48	.40	.52	.55	.02	.05	.05	1.93	.98	180.54	3.76
52	.35	.52	.55	.02	.05	.05	2.09	1.06	199.72	3.84
56	.30	.52	.55	.03	.05	.05	2.25	1.14	221.88	3.96
60	.25	.52	.55	.04	.05	.05	2.41	1.22	248.80	4.14
64	.20	.52	.55	.05	.05	.05	2.57	1.31	284.06	4.44
68	.15	.52	.55	.07	.05	.05	2.73	1.39	335.98	4.94
72	.10	.52	.55	.11	.05	.05	2.89	1.47	429.56	5.97
76	.05	.52	.55	.19	.05	.05	3.05	1.55	689.82	9.08

Table 2.
Solutions Utilizing Automated Call-Screening Agent For Various Arrival Rates

Arrival Rate Total/hour	Server Utilization		Mean Wait Time in Queue				Optimal Solutions			
	Tier I	Tier II	Tier I	Tier II		Servers		Minimum Expected Total Cost	Cost per Arrival	
				A	B	A	B			
40	.25	.52	.55	.002	.05	.05	2.30	1.17	170.42	4.26
44	.27	.52	.55	.002	.05	.05	2.52	1.28	187.38	4.25
48	.30	.52	.55	.003	.05	.05	2.75	1.40	204.42	4.26
52	.32	.52	.55	.003	.05	.05	2.98	1.52	221.35	4.26
56	.35	.52	.55	.003	.05	.05	3.21	1.63	238.78	4.26
60	.37	.52	.55	.004	.05	.05	3.44	1.75	256.13	4.27
64	.40	.52	.55	.004	.05	.05	3.67	1.87	273.60	4.28
68	.43	.52	.55	.005	.05	.05	3.90	1.98	291.23	4.28
72	.45	.52	.55	.005	.05	.05	4.13	2.10	309.01	4.29
76	.47	.52	.55	.006	.05	.05	4.36	2.22	326.98	4.30

Base parameters from Table 1 apply, with the following exceptions: a new, faster mean service time equal to 160 calls/hour (μ_1) at tier I; a lower hourly variable cost of \$4/hr (C_1); and 100 percent of all calls being channeled to tier II ($\Theta = 1$).

The results are similar to those shown in Table 1 are realized. The utilization of tier II technicians and wait times are constant for all mean arrival rates, and equal to the values determined in Table 1. The wait times, as noted earlier, namely, the idle time for a type i tier II server, and the mean wait time in a type i tier II queue, $i = A, B$, are given by the following (generalized here for either type tier II service):

$$P_{0i} = \text{Pr}(\text{type } i \text{ tier II server is idle}) = 1 - \frac{1}{\sqrt{C_s/C_i}} \quad i = A, B \quad (17)$$

$$W_i^* = E(\text{time in type } i \text{ tier II queue}) = \frac{1}{\mu_i} \sqrt{\frac{C_i}{C_s}} \quad i = A, B \quad (18)$$

Both (17) and (18) are independent of the parameters which have been changed between the analyses in Table 1 and Table 2. What has changed are the idle time and mean wait time for tier I. Both tier I idle time and mean wait time are now much lower for any arrival rate, due to the decreased mean service time at tier I. Correspondingly, the number of tier II technicians at the two service areas (A and B) have increased to meet the increased demand for service.

This above analysis leads us to consider the policy question of whether the automated server should be utilized. It is necessary to weigh the decreased wait time cost and lower tier I operational cost against the increased server cost at tier II. For arrival rates up to (greater than) 60 per hour, the call screening agent yields a lower (higher) expected minimum cost. This policy is intuitive since the call screening agent has a lower capacity (due to the higher mean service time) and would find the queue growing exceedingly long as the arrival rate approaches the mean service time. This disproportionate increase causes social costs (measured by wait time) to increase rapidly as the arrival rate approaches the mean tier I service time

(or threshold in order to achieve steady state). This leads to the optimal solution of utilizing an automated server, despite the increased operational costs at tier II. The overall minimum expected costs between the two policies analyzed in Tables I and II can be further illustrated in Figure 1:

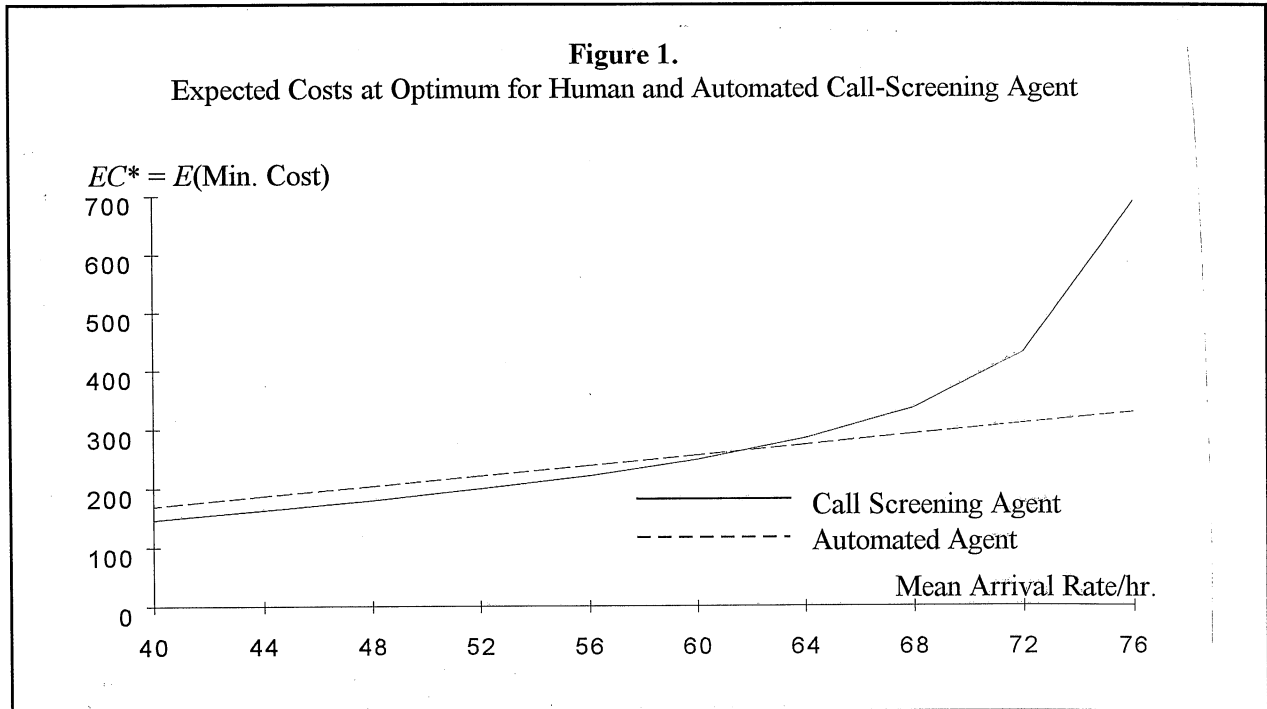
The per unit cost for arrivals attains a minimum at a unique optimal arrival rate. The minimum cost per call is \$4.25, achieved at an arrival rate of 44 calls per hour. This per unit cost is higher than that achieved with the call screening agent and is determined at a higher optimal arrival rate. Consequently, if price is of concern, and an average cost pricing policy is utilized, lower prices can be passed on to the consumer through the utilization of a call screening agent. The arrival rate, however, must be set at a lower level than when the automated server is employed.

Another capacity policy which may be considered involves using an automated server for periods of peak demand (where total minimum expected costs are lower for the automated server) and a call screening agent at other times. Finally, since cost per call increases to infinity as the mean arrival rate approaches the mean service, this suggests that outsourcing may be a viable alternative for small help desks.

Conclusions

Help desks, which provide computer/software support, have experienced significant growth in the past decade. This growth has led to several problems, including increased operational costs and support staff burn-out. The growing cost of support has led many corporations to reconsider the policy of providing support free of charge. At the same time, support has become a major factor in consumers' product preferences. Consequently, help desks are faced with the problem of achieving a balance between cost of and quality of support, while at the same time addressing problems of personnel turn over and staff burnout.

This paper has presented a queuing theo-



retic model that can be used to aid in the determination of policy regarding capacity analysis for help desks which have multi-echelon networked service stations. Numerical results were utilized to contrast and compare two alternate strategies; automated tier I service versus a human call screening agent. The model could suggest that a combination of utilizing an automated server during peak demand and a call screening agent at other times minimizes total expected costs.

The objective of the mathematical model, which was to minimize total expected cost included both operational and social costs in order to evaluate the trade-off between support costs and quality. The model recommends an optimal capacity staffing policy, as well as ancillary figures regarding utilization rates and wait times. These figures may be utilized when inputting aspects of staff burn out into the analysis. This, in turn, enables decision makers to evaluate solutions which satisfice, as opposed to optimize.

The intention of this model is to provide a decision support tool which assists in the determination of optimal policy regarding capacity planning. This model does not capture all the elements contained within the sphere of influence for help

desk support analysis. There also exist many extensions for the model herein not considered. Some of these include peak load and congestion effects, multiple tier I servers, and alternate routing algorithms. Future research could incorporate many of these aspects of the problem into the mathematical analysis.

The utilization of this mathematical model combined with qualitative policy tools form a dynamic and powerful combination. These decision making tools can be used to assist management in providing quality service in an environment where user support has become a significant product characteristic. This type of analysis can assist a firm in remaining competitive in a period of rapid technological change.

Suggestions for Future Research

While the model presented here captures many of the basic operating characteristics of help desks, the incorporation of additional features would enhance the applicability of the model. For example, the determination of a more representative service time distribution would enhance the model, as would a more flexible routing procedure which allows for calls to be rerouted from a par-

ticular tier two service station to another. Alternatively, or additionally, the incorrect routing of a given percentage of calls can be considered. These calls would not be rerouted, but handled by the agent assigned. Since this agent is not the primary agent ordinarily handling this type of call, a resulting increase in service time would obtain. Although the mathematics involved in these instances may become intractable, simulation might be useful as an alternative optimization tool.

The incorporation of unsatisfied callers making return calls might also be considered, as well as the incorporation of consumer satisfaction into the quality cost component.

The final suggestion involves the development of a computer supported decision support system to handle the computational aspect involved in performing extensive sensitivity analysis on problem parameters. This would allow management to understand the impact of simultaneous changes in service time and costs. For example, the impact of additional training resulting in higher wages with associated smaller mean service times can be investigated. □

Endnotes

1. Tier three personnel, when present, are specialists, who are qualified to respond to questions within their area of expertise.
2. While noting that precision is an important attribute in determining customer satisfaction, we will assume that all calls, when completed, have attained the same level of precision. Modeling the precision dimension of the problem at hand would provide an interesting extension of the basic model.
3. If jockeying is allowed, one may assume that customers will filter to the empty queue in an attempt to balance either the queue lengths, or perhaps the perceived wait time. This situation raises the question of etiquette, since jockeying back and forth may create a degree of "irritation" within the queue.
4. In some cases automated AI routines assign calls in such a way so as to achieve higher efficiency. Also, some help desks institute

single queue, rather than multi-queue configurations. While these alternate cases are not considered, they are viable realizations which should be recognized.

5. The reader should note that waiting time costs accrue only for the time spent waiting for service, and does not include the time spent being serviced by the call-screening agent and/or tier II personnel.
6. i.e. those calls which cannot be resolved by the call screening agent at tier I.
7. It is assumed that all Type A (Type B) servers are governed by independent identically distributed service times.
8. Although the tier II wait time remains constant as the mean arrival rate increases, the mean wait time in tier I increases. Consequently, increasing in the overall mean arrival rate increases total wait time to which social cost applies.
9. Differentiating the service times and costs for tier II servers has no effect on the comparative relationships to be demonstrated in the numerical example.
10. An interesting extension to the analysis would involve defining service rate as a function of idle time. Extensions of this nature are not uncommon, for example, mean service time is occasionally considered a function of queue length.

References

1. Editorial Staff. "Beam & Whiteside adds new fee-based premium services," *Service News*, Vol. 14, No. 10, pp. 29, 10/94.
2. Editorial Staff. Interview with Coleman Sisson (Customer Service Manager, Powersoft), *Service News*, Vol. 14, No. 10, pp. 26,29, 10/94.
3. Editorial Staff. "Research Points to Support's Role," *Service News*, Vol. 14, No. 10, pp. 4, 10/94.
4. Bisenius, B., "Implementing a Help Desk at a Small Liberal Arts College," *ACM SIGUCCS XX*, pp. 37-41, 1992.
5. Daly, C., "Getting Past Call Tracking," *Service News*, Vol. 14, No. 6, pp. 21,23, 6/94.

6. Deavers-Claspell, P., "How to Build a Help Desk that Floats," *ACM SIGUCCS XVI*, pp. 219-223, 1988.
7. Eskow, D., "Firms Unite to Give Help Desks the Help They Need," *PC Week*, 2/90.
8. Etchinson, J., "How to Determine the Best Skill Level For Your Staff," *LifeRaft*, Publication of the Help Desk Institute, pp. 4-7, 9-10/94.
9. Frye, C., "To Better Handle Beefs, Companies Beef Up Support Operations," *Client/Server Computing*, pp. 46-51, 4/94.
10. Goitein, M., "Sounding Board - Waiting Patiently," *The New England Journal of Medicine*, pp. 604-608, August 1990.
11. Grassmann, W., "Finding the Right Number of Servers in Real-World Queuing Systems," *Interfaces*, Vol.18, No.2, 1988.
12. Gross, D. & Harris, C., *Fundamentals of Queuing Theory*, John Wiley & Sons, New York, 1985.
13. Hunse, T., "Evolving Toward the Multi Application Help Desk," *Service News*, Vol. 14, No. 8, pp. 23,26, 8/94.
14. Johnson, K., "PC Software Support Takes Center Stage," *Service News*, Vol. 14, No. 3, pp. 37-40, 3/94.
15. Larson, R., "Perspectives on Queues: Social Justice and the Psychology of Queuing," *Operations Research*, Vol. 35, No. 6, pp. 895-905, 1987.
16. Mason, S., "The Application of Expert Systems in Customer Support," *IEEE International Engineering Management Conference*, pp. 125-128, 1990.
17. Peters, S., "Setting Up and Automating a Help Desk: the First Year of Operation," *ACM SIGUCCS*, 1992.
18. Pinedo, M. & Wolff, R., "A Comparison Between Tandem Queues with Dependent and Independent Service Times," *Operations Research*, Vol. 30, No. 3, pp. 464-479, 1982.
19. Rothkopf, M., & Rech, P., "Perspectives on Queues: Combining Queues is Not Always Beneficial," *Operations Research*, Vol. 35, No. 6, pp. 906-909, 1987.
20. Stallings, W., "A Practical Guide to Queuing Analysis," *Byte*, pp. 309-316, 2/91.
21. Whitt, W., "Understanding the Efficiency of Multi-Server Service Systems," *Management Science*, Vol. 38, No. 5, 1992.
22. Whitt, W., "Deciding Which Queue to Join: Some Counterexamples," *Operations Research*, Vol. 34, No. 1, pp. 55-62, 1986.
23. Young, R., "Why Single-Line Teller Queues May Not Work," *ABU Banking Journal*, pp. 32-33, 5/82.

Purchasing Power Parity And The Canadian Fixed And Float Since 1962

Dr. Hamid Baghestani, The Economics Institute and University of Colorado at Boulder
Mr. Mark Gohr, University of Colorado at Denver

Abstract

The empirical validity of PPP as a long run constraint between the U.S. and Canada is examined for 1962.3-1994.4. While PPP appears to hold over the entire sample period, the ECM estimates fail to show parameter stability due to pooling the data from the fixed (1962.3-1970.1) and floating (1970.2-1994.4) exchange rate regimes. The ECM estimates for 1970.2-1994.4, however, show stability and further indicate that (1) U.S. prices do not respond to deviations from PPP, and (2) the real exchange rate follows a nonstationary process.

Introduction

In current theories of international economics, the purchasing power parity (PPP) is often hypothesized to constrain the movements of domestic and foreign price levels and exchange rates, at least, in the long-run. Given that the policy conclusions of such theoretical models critically depend on the PPP hypothesis, then the question is whether PPP as a long-run constraint holds empirically. The findings of the existing empirical studies are at best mixed. Among others, the studies by McNown and Wallace (1990) and Taylor (1988) generally report results not supportive of PPP. The studies by Frankel (1986), McNown and Wallace (1989), and Taylor and McMahon (1988), however, provide empirical evidence supportive of PPP. The studies reporting results favorable to PPP require either long time periods (often one hundred years or more), or large differences in price movements between country pairs. In order to show that such requirements are not necessary, Choudhry, McNown, and Wallace (1991) present empirical evidence in favor of PPP between the United States and Canada for the

1950-1961 period of floating exchange rate regime.

The present paper adds to Choudhry, McNown, and Wallace (1991) by further examining the empirical validity of PPP as a long-run constraint between the United States and Canada for the period from 1962 to the present. Noting that this period includes both fixed and floating exchange rate regimes, we want to see (i) if a change in exchange rate regime affects the dynamics of the U.S. price level and the exchange rate adjusted Canadian price level, (ii) in case PPP holds, how these price levels respond to deviations from PPP, and (iii) whether PPP holds in strict form, or the real exchange rate follows a stationary process.

Using quarterly data, two sample periods have been examined here: the first sample period is 1962.3-1994.4 which includes both the fixed (1962.3-1970.1) and floating (1970.2-1994.4) exchange rate regimes; and the second sample period is 1970.2-1994.4 which covers only the recent

floating exchange rate regime.

Methodology and empirical results

Following Choudhry, McNown, and Wallace (1991), PPP is expressed in terms of the "Law of One Price" in order to also accommodate the fixed exchange rate period of 1962.3-1970.1; that is,

$$Pd_t = a + b XPf_t + u_t \quad (1)$$

where $XPf_t = (X_t + Pf_t)$. Pd_t and Pf_t are the logarithms of domestic and foreign prices in period t ; X_t is the logarithm of exchange rate, the domestic currency price of foreign exchange in period t ; and u_t represents deviations from PPP. To be further consistent with Choudhry, McNown, and Wallace (1991), Pd_t is defined as the U.S. consumer price index (CPI), and XPf_t is defined as the Canadian CPI adjusted for the exchange rate movements. Strict PPP requires $b = 1$, indicating that the real exchange rate, given by $(Pd_t - XPf_t)$, must follow a stationary process.

In general, for PPP to hold, equation (1) must prove to be a cointegrating vector. As a precondition for cointegration, the individual series on Pd_t and XPf_t are tested for a common order of integration. Based on the augmented Dickey-Fuller

(ADF) tests (Dickey and Fuller, 1979) with the critical values from MacKinnon (1991), Table 1 presents evidence that each series is integrated of order one for both sample periods under consideration; that is, each series is $I(1)$ in the notation of Engle and Granger, 1987. The equations for the tests on the levels include an intercept and time trend term, to allow a deterministic trend under the alternative. In testing for a unit root in the first differences, only the intercept is included. All ADF test equations also include augmented lags, where the adequacy of the lag length (indicated in Table 1) is checked with tests for serial correlation using the Lagrange multiplier χ^2 -statistic.

Following Engle and Granger (1987), the time series on Pd_t and XPf_t are said to be cointegrated, if u_t is stationary, $u_t \sim I(0)$. In the same manner, for the two series to be cointegrated, v_t , the error term from the reverse cointegrating equation, should be stationary. Table 2 reports the ordinary least squares (OLS) regression estimates of both normalizations and the ADF test statistics on the respective residual series. The residual series from each normalization appear to be stationary, with the ADF test statistics compared to MacKinnon's (1991) critical values. In other words, Pd_t and XPf_t appear to be cointegrated for both sample periods, and that the series seem to possess an equilibrium relation, consistent with the PPP hypothesis,

to which they converge in the long-run. The ADF test equations for cointegration exclude both the intercept and time trend, since such deterministic components fail to be statistically significant. These test equations, however, include augmented lags, where the adequacy of the lag length (indicated in Table 2) is checked with tests for serial

Table 1
Unit root testing

	ADF (Lags)			
	1962.3 - 1994.4		1970.2 - 1994.4	
Pd	-1.60	(6)	-0.70	(8)
XPf	-2.16	(12)	-1.88	(12)
DPd	-2.71 ^c	(6)	-2.58 ^c	(7)
DXPf	-3.61 ^a	(12)	-3.24 ^b	(12)

Notes: Pd is the U.S. CPI (in logarithms), XPf is the exchange rate adjusted Canadian CPI (in logarithms). The number of augmented lags included in the ADF test equation is in parenthesis. a,b,c denotes significance at the 1%, 5%, and 10% levels, respectively.

correlation using the Lagrange multiplier χ^2 -statistic.

	ADF (Lags)
1962.3-1994.4:	
$u = Pd - 0.034 - 1.047 XPf$	-3.26 ^b (8)
$v = XPf - 0.019 - 0.941 Pd$	-3.38 ^c (8)
1970.2-1994.4:	
$u = Pd + 0.123 - 1.084 XPf$	-3.32 ^c (8)
$v = XPf - 0.223 - 0.896 Pd$	-3.16 ^c (8)

Notes: \hat{u} is the OLS residual series from equation (1) which is normalized for Pd. \hat{v} is the OLS residual series from equation (1) when renormalized for XPf.

Following the Granger representation theorem (Engle and Granger 1987), for the cointegrated series Pd_t and XPf_t , there exists an ECM of the form

$$\Delta Pd_t = \alpha + \sum_i \beta_i \Delta Pd_{t-i} + \sum_i \gamma_i \Delta XPf_{t-i} - \lambda_1 \hat{u}_{t-1} + \varepsilon_{1t} \quad (2)$$

$$\Delta XPf_t = a + \sum_i b_i \Delta XPf_{t-i} + \sum_i c_i \Delta Pd_{t-i} - \lambda_2 \hat{v}_{t-1} + \varepsilon_{2t} \quad (3)$$

The ECM embodies both the short-run dynamics and the long-run equilibrium relation of the series, or PPP. Equation (2) specifies the short-run convergence process of Pd_t to PPP with convergence being assured when λ_1 is between zero and one. Similarly, Equation (3) specifies the short-run convergence process of XPf_t to PPP with convergence being assured when λ_2 is between zero and one.

The ECM, which initially includes eight lag differences for both Pd_t and XPf_t , is estimated using OLS. Excluding the insignificant lag differences, the ECM is then reestimated, with the results for both sample periods reported in Table 3. Based on the cusum of squares test of stability developed by Brown, Durbin, and Evans (1975), the

ECM estimated for 1962.3-1994.4 fails to be stable in terms of parameters at the 1% level of significance.¹ This may be due to pooling the data from the fixed (1962.3-1970.1) and floating (1970.2-1994.4) exchange rate regimes. Such an argument may be reinforced by the evidence that the 1970.2-1994.4 ECM estimates show stability in terms of parameters at the 10% or lower level of significance.² Further evidence suggests that Pd_t and XPf_t are not cointegrated for 1962.3-1970.1, and therefore, PPP fails to hold for this period of fixed exchange rate regime.³ This is not surprising, since during this period the Canadian dollar at U.S. \$0.925 was substantially undervalued for every year except 1966. Originally, the Bank of Canada lent its support to the fixed exchange rate with higher interest rates and the Canadian government by increasing foreign exchange reserves through foreign loans (Binhammer and McDonough, 1990, p. 54). For 1963-1967, however, Canada was required to meet foreign exchange reserve ceilings imposed by the U.S.-Canada trade agreement. Such requirements resulted in the immobilization of the Canadian monetary policy which adversely affected the Canadian price level and its long-run equilibrium relation with the U.S. price level.

Turning back to the 1970.2-1994.4 estimates of the ECM in Table 3, we see that in the short-run the U.S. price level does not respond to departures from PPP. This may be due to the fact that the U.S. is a large country compared to Canada. The Canadian price level adjusted for exchange rate, however, responds to departures from PPP (measured by $v_t = XPf_t - .223 - .896Pd_t$ from Table 2) with a parameter estimate of 0.095. This indicates that 9.5% of any positive deviation in PPP is corrected within a quarter by a fall in the U.S. value of the Canadian price level.

To examine whether strict PPP holds, equation (3) is reestimated by replacing v_{t-1} with XPf_{t-1} and Pd_{t-1} . Based on the Wald test, the null hypothesis that the parameters on XPf_{t-1} and Pd_{t-1} are identical (in absolute value) is rejected; the calculated chi-squared test statistic is found to be 3.84 with a P-value of 0.05. Such evidence leads us to conclude that b in (1) significantly differs

Table 3
Error correction model estimates

1962.3 - 1994.4			
Equation (2)		Equation (3)	
$\alpha = .002^a$ (2.8)		$a = -.001$ (.44)	
$\Sigma\beta_i = .907^a$ (15.8)	$i=1-3,6$	$\Sigma b_i = .976^a$ (5.1)	$i=1,3,6-8$
$\Sigma\gamma_i = -.077^b$ (2.2)	$i=1-2,6$	$\Sigma c_i = .128$ (.50)	$i=1,8$
$\lambda_1 = -.012^b$ (1.9)		$\lambda_2 = -.104^a$ (3.5)	
$R^2 = .77^a, \chi^2_{12} = 9.8, S_r^m = .16^c$		$R^2 = .29^a, \chi^2_{12} = 10.2, S_r^m = .23^a$	
1970.2 - 1994.4			
$\alpha = .003^b$ (2.5)		$a = .001$ (.05)	
$\Sigma\beta_i = .830^a$ (3.6)	$i=1-8$	$\Sigma b_i = .735^a$ (4.5)	$i=3,6-7$
$\Sigma\gamma_i = -.028^c$ (1.7)	$i=1-2,4,6$	$\Sigma c_i = .161$ (.49)	$i=1,8$
$\lambda_1 = -.004$ (.57)		$\lambda_2 = -.095^a$ (2.9)	
$R^2 = .78^a, \chi^2_{12} = 9.1, S_r^m = .13$		$R^2 = .30^a, \chi^2_{12} = 3.1, S_r^m = .14$	
<p>Notes: χ^2_{12} is the Lagrange multiplier chi-squared statistic, which detects serial correlation up to twelve lags; S_r^m is the maximum absolute deviation of the calculated cusum of squares test statistic, S_r, from its expected value (see footnote 1); and t-statistics (in absolute value) are given in parentheses.</p>			

from one due to perhaps transportation costs as shown by Taylor (1988). That is, PPP does hold *but* not in strict form, implying that the real exchange rate, defined as $[Pd_t - (X_t + Pf_t)]$, follows a nonstationary process. This conclusion was also supported when testing for a unit root in the real exchange rate series. For example, the calculated ADF test statistic for the real exchange rate series in levels is found to be -1.98 which is not significant at any reasonable level. This test statistic for the real exchange rate series in the first differences is -2.67 which is significant at the 1% level. Such


findings suggest that the real exchange rate for 1970.2-1994.4 is I(1).

Conclusions

Based on our empirical findings, the U.S. price level and the exchange rate adjusted Canadian price level appear to be cointegrated for both sample periods of 1962.3-1994.4 and 1970.2-1994.4. The ECM estimated for 1962.3-1994.4, however, fails to be stable in terms of parameters due to pooling the data from the fixed and floating

exchange rate regimes. That is, once the fixed exchange rate period was excluded, then the ECM estimates for the 1970.2-1994.4 period of floating exchange rate regime show stability in terms of parameters and further indicate that the U.S. value of the Canadian prices respond to correct deviations from PPP. The U.S. price level, however, does not respond to deviations from PPP, perhaps because the U.S. is a large country compared to Canada, and that movements in the U.S. price level are (weakly) exogenous to the Canadian variables. Our findings also suggest that, for 1970.2-1994.4, PPP holds *but* not in strict form, perhaps due to transportation costs as shown by Taylor (1988). Such evidence implies that the real exchange rate follows a nonstationary process. This conclusion differs from Choudhry, McNown, and Wallace (1991) who found the real exchange rate to be stationary for the 1950s.

Suggestions for future research

An important issue, often raised, is whether or not government interventions in the foreign exchange market can distort PPP. During the floating exchange rate regime, both the Canadian and U.S. governments, at times, influenced the exchange rate through market interventions. Future research may, therefore, investigate the robustness of our conclusions to such interventions. 

Footnotes

1. The cusum of squares test of stability is based on a standardized one-step-ahead prediction of errors, w_t , from recursive regressions and is designed to test the null hypothesis that the regression parameters vector of the recursive regressions are identical. The cusum of squares test statistic is

$$S_r = \left(\sum_{t=k+1}^r w_t^2 \right) / \left(\sum_{t=k+1}^T w_t^2 \right), \quad r = k+1, \dots, T$$

where k is the number of regression parameters, and T is the number of observations of the last estimation period. The expected value of S_r under the null hypothesis of parameter stability is $(r-k)/(T-k)$. One rejects this null hypothesis if the absolute de-

viation of S_r (for all r) from its expected value is above the critical values given in Durbin (1969). For the 1962.3-1994.4 estimates of both equations (2) and (3) in Table 3, the maximum absolute deviation of S_r^m from its expected value, denoted S_r^m , is significant or above the critical value, leading to the rejection of the null hypothesis of parameter stability.

2. For the 1970.2-1994.4 estimates of both equations (2) and (3) in Table 3, the maximum absolute deviation of S_r from its expected value is insignificant or below the critical value, leading to the acceptance of the null hypothesis of parameter stability. The ECM estimates for 1970.2-1994.4, in Table 3, also pass a series of diagnostic tests including serial correlation (based on the inspection of the autocorrelation functions of the residuals as well as the reported insignificant Lagrange multiplier χ^2 -statistics at the 10% or lower level) and omitted variables such as a time trend and other lags.
3. For 1962.3-1970.1, we have found $u = Pd - 0.515 - 0.893 XPf$ and $v = XPf + 0.525 - 1.104 Pd$. The ADF test statistics for u and v are -1.57 and -1.48 , respectively. These test statistics are not significant, leading to the conclusion that Pd and XPf are not cointegrated for 1962.1-1970.1. It is noted that each ADF test equation also includes six augmented lags, with the adequacy of the lag length checked with tests for serial correlation using the Lagrange multiplier χ^2 -statistic.

References

1. Binhammer, H. and L.C. McDonough, "The Canadian-U.S. Dollar Exchange Rate: History and Policy Issues," *Research in International Business and Finance*, Vol. 8, pp. 51-68, 1990.
2. Brown, R.L., J. Durbin, and J.M. Evans, "Techniques for Testing the Constancy of Regression Relationship over Time," *Jour-*

- nal of the Royal Statistical Society*, Vol. 37, Series B, pp. 149-192, 1975.
3. Choudhry, T., R. McNown, and M. Wallace, "Purchasing Power Parity and the Canadian Float in the 1950s," *Review of Economics and Statistics*, Vol. 73, pp. 558-563, 1991.
4. Dickey, D.A. and W.A. Fuller, "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, Vol. 74, pp. 427-431, 1979.
5. Durbin, J., "Tests for Serial Correlation in Regression Analysis Based on the Peridogram of Least Squares Relationship," *Biometrika*, Vol. 56, pp. 1-15, 1969.
6. Engle, R.F. and C.W.J. Granger, "Cointegration and Error Correction: Representation Estimation and Testing," *Econometrica*, Vol. 55, pp. 251-76, 1987.
7. Frankel, J.A., "International Capital Mobility and Crowding-out in the U.S. Economy: Imperfect Integration of Financial or of Goods Markets?" in R.W. Hafer (ed.), *How Open is the U.S. Economy?* Lexington Books, Lexington, 1986.
8. MacKinnon, J.G., "Critical Values for Cointegration Tests," in *Long-run Economic Relationships*, edited by R.F. Engle and C.W.J. Granger, Oxford: Oxford Press, pp. 267-276, 1991.
9. McNown, R. and M. Wallace, "National Price Levels, Purchasing Power Parity, and Cointegration: A Test of Four High Inflation Economies," *Journal of International Money and Finance*, Vol. 8, pp. 533-545, 1989.
10. _____, "National Price Levels as Cointegrated Systems: A Test of Purchasing Power Parity Among Four Industrialized Countries," *Applied Economics*, Vol. 22, pp. 1729-1737, 1990.
11. Taylor, M.P., "An Empirical Examination of Long Run Purchasing Power Parity Using Cointegration Techniques," *Applied Economics*, Vol. 20, pp. 1369-1381, 1988.
12. Taylor, M.P., and P.C. McMahon, "Long Run Purchasing Power Parity in the 1920s," *European Economic Review*, Vol. 32, pp. 179-197, 1988.