Empirical Adventures in Call Centers Emergency Departments, ...

Avishai Mandelbaum
Technion IE&M

with Graduate Students, Research Partners
Technion SEE Center, IBM+Rambam+Technion OCR

WITOR, September 2009

DataMOCCA

Data MOdels for Call Center Analysis

Project Collaborators:

Technion:	Paul Feigin, Avi Mandelbaum Technion SEElab: Valery Trofimov, Ella Nadjharov, Igor Gavako, Katya Kutsy, Polyna Khudyakov, Shimrit Maman, Pablo Liberman Students (PhD, MSc, BSc), RAs
Wharton:	Larry Brown, Noah Gans, Haipeng Shen (N. Carolina), Students, Wharton Financial Institutions Center
Companies:	U.S. Bank, Israeli Telecom, 2 Israeli Banks, Israeli Hospitals, ...

The SEE Center - Project DataMOCCA

Goal: Designing and Implementing a (universal) data-base/data-repository and interface for storing, retrieving, analyzing, displaying and interacting with transaction-based data.

Goal: Designing and Implementing a (universal) data-base/data-repository and interface for storing, retrieving, analyzing, displaying and interacting with transaction-based data.

Enable the Study of:

- Customers (Callers, Patients)	Waiting, Abandonment, Returns
- Servers (Agents, Nurses)	Service Duration, Activity Profile
- Managers (System)	Loads, Queue Lengths, Trends

Call-Center: Hidden Complex Service Network

Call-Centers: "Sweat-Shops of the 21st Century"

A “Good” Hospital in Beijing

DataMOCCA: System Components

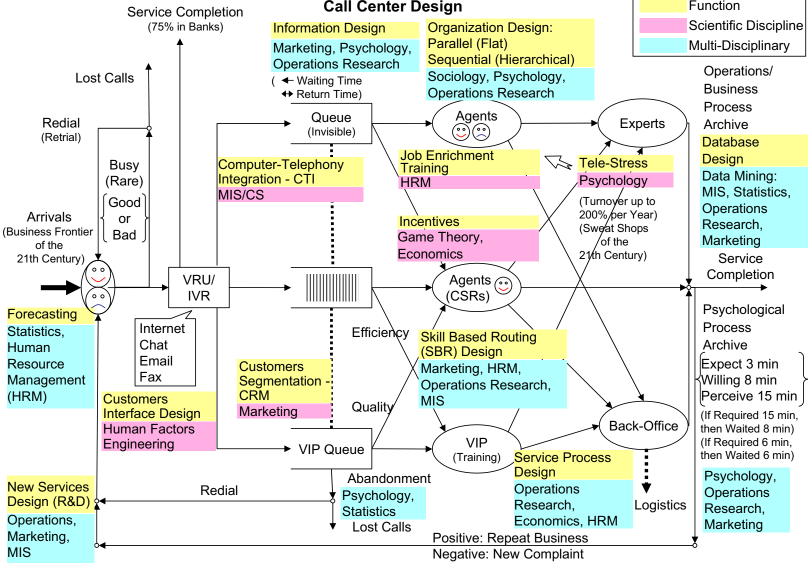
1. Clean Databases: Operational histories of **individual** customers and servers (mostly with IDs).
   - In Call Centers: from IVR to Exit;
   - In Hospitals: from ED to Exit (or just ED).
2. **SEEStat: Online** GUI (friendly, flexible, powerful)
   - Queueing-Science perspective;
   - Operational data (vs. financial, contents or clinical);
   - Flexible customization (e.g. seconds to months);
3. Tools:
   - Online statistics (survival analysis, mixtures, smoothing);
   - Dynamic Graphs (flow-charts, work-flows)
   - Simulators (CC, ED; data-driven).

Current Databases

1. U.S. Bank (**PUBLIC**): 220M calls, 40M agent-calls, 1000 agents, 2.5 years, 7-40GB.
2. Israeli Banks:
   - Small (**PUBLIC**): 350K calls, 15 agents, 1 year. Started it all in 1999 (JASA), now "romancing" again (Medium, with 300 agents);
   - Large (ongoing): 500 agents, 1.5 years, 3-8GB.
3. Israeli Telecom (ongoing): 800 agents, 3.5 years; 5-55GB.
4. Israeli Hospitals:
   - Six ED's (to be made **PUBLIC**);
   - Large (ongoing): 1000 beds, 45 medical units, 75,000 patients hospitalized yearly, 4 years, 7GB.
5. Website (pilot).

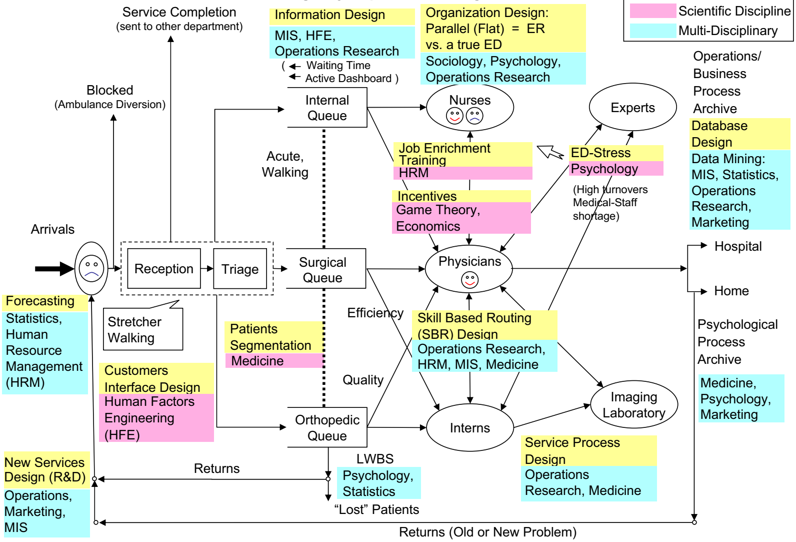
Service Engineering: Multi-Disciplinary Process View

Call Center Design



Service Engineering: Multi-Disciplinary Process View

Emergency Department Design

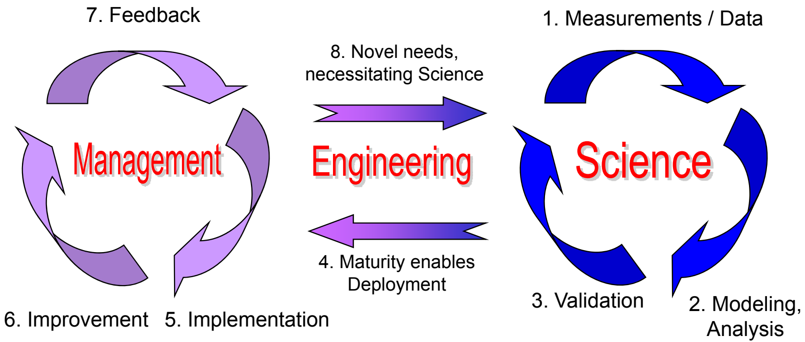


Expanding the Scientific Paradigm (OCR)

- Physics, Biology, . . . : Measure, Model, Experiment, Validate, Refine.
- Human-complexity triggered the above in Transportation, Economics.

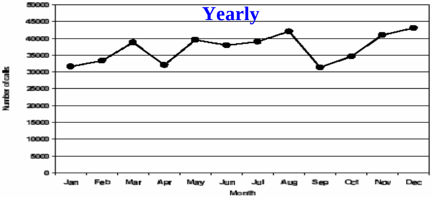
Expanding the Scientific Paradigm (OCR)

- Physics, Biology, . . . : Measure, Model, Experiment, Validate, Refine.
- Human-complexity triggered the above in Transportation, Economics.
- Expand to:

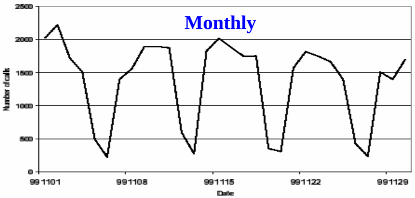


Arrivals to a Call Center (Israel, 1999): Time Scales

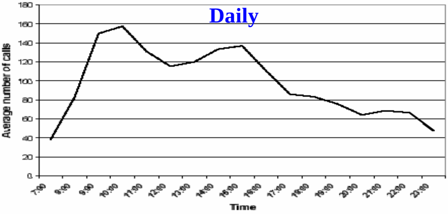
Strategic



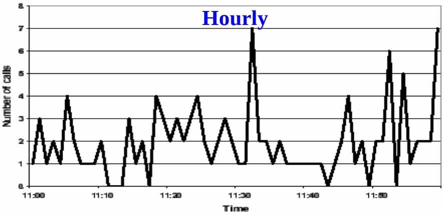
Tactical



Operational (Predictable Var.)



Regulatory (Stochastic)



Arrivals to a Call Center (U.S., 1976): Queueing Science

(E. S. Buffa, M. J. Cosgrove, and B. J. Luce, "An Integrated Work Shift Scheduling System")

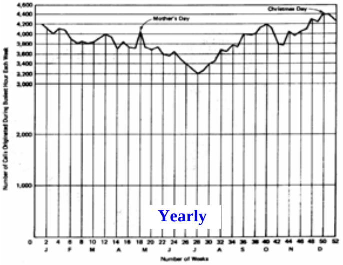


Figure 1 Typical distribution of calls during the busiest hour for each week during a year.

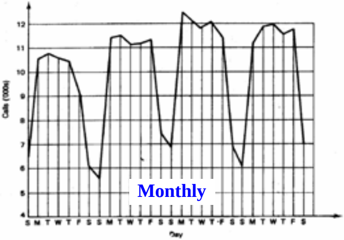


Figure 2 Daily call load for Long Beach, January 1972.

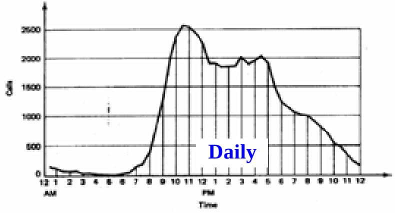


Figure 3 Typical half-hourly call distribution (Bundy D A).

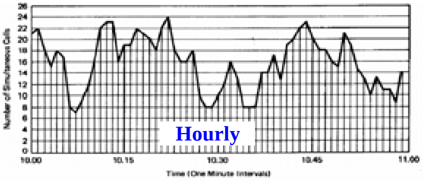
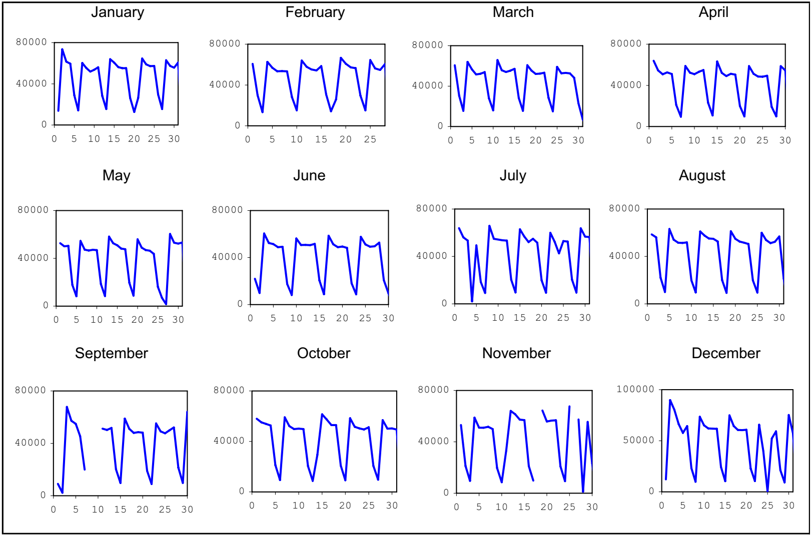


Figure 4 Typical intrahour distribution of calls, 10:00–11:00 A.M.

Monthly Arrivals to Service

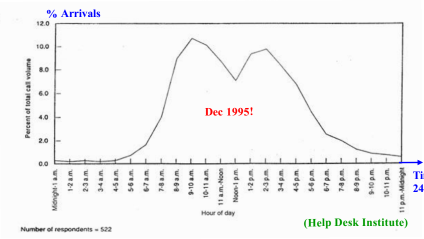
U.S. Bank: Daily Arrival-Rates, over a Month, 2002



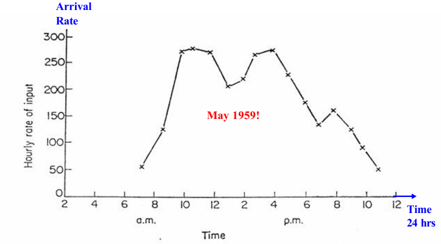
Daily Arrivals to Service: Time-Inhomogeneous (Poisson?)

Intraday Arrival-Rates (per hour) to Call Centers

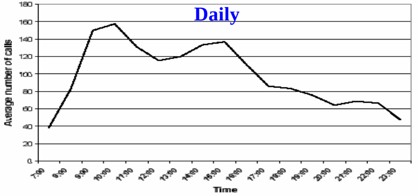
December 1995 (700 U.S. Helpdesks)



May 1959 (England)



November 1999 (Israel)

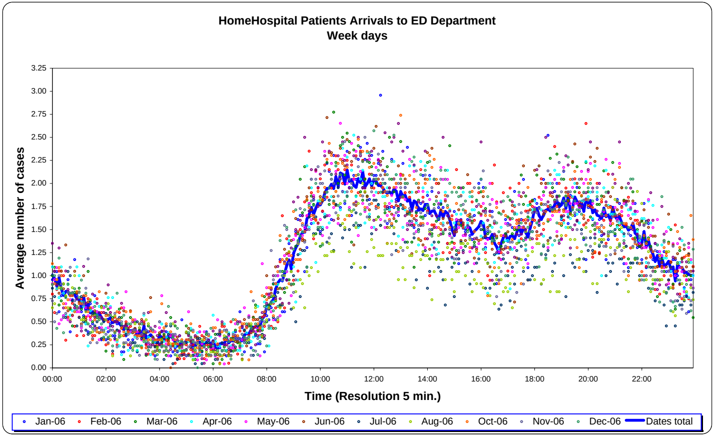


Observation:

Peak Loads at **10:00 & 15:00**

Arrivals to an Emergency Department (ED)

Large Israeli ED, 2006

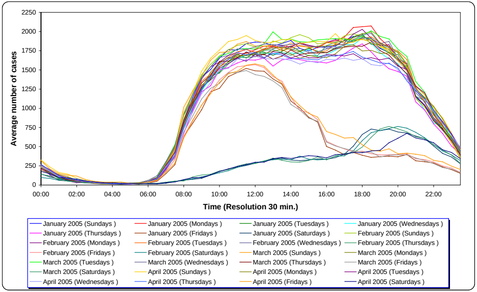


- **Second** peak at 19:00 (vs. 15:00 in call centers).
- How much stochastic variability ?

Intraday Arrival Rates: Does a Day have a Shape ?

Arrival Patterns, Israeli Telecom

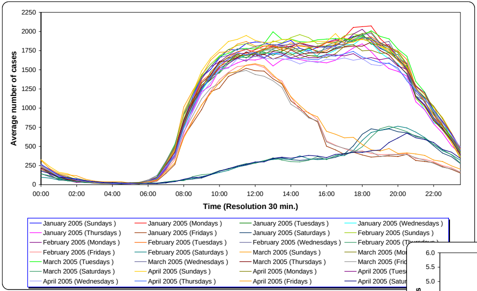
Arrivals, Avg. Weekdays/1-4/2005



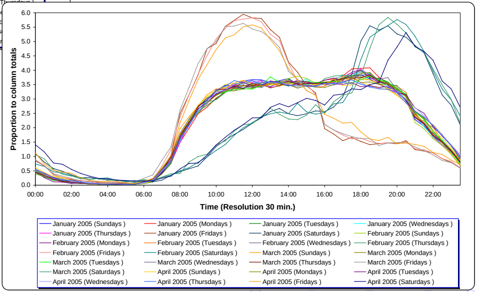
Intraday Arrival Rates: Does a Day have a Shape ?

Arrival Patterns, Israeli Telecom

Arrivals, Avg. Weekdays/1-4/2005



Percent to Total, 30 min.



A (Common) Model for Call Arrivals

Whitt (99'), Brown et. al. (05'), Gans et. al. (09'), and others:

Doubly-stochastic (Cox, Mixed) Poisson with instantaneous rate

$$\mathbf{\Lambda(t) = \lambda(t) \cdot X} \ ,$$

where $\int_0^T \lambda(t)dt = 1$.

- $\boldsymbol{\lambda}\mathbf{(t)}$ = "Shape" of weekday [**Predictable variability**]
- **X** = Total # arrivals [**Unpredictable variability**]

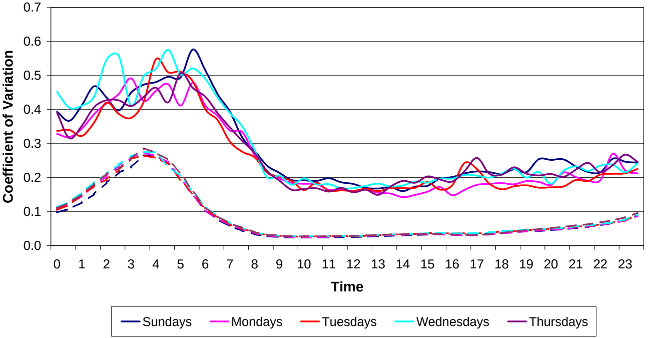
w/ Maman & Zeltyn (09'):
Above assumes **"too-much" stochastic variability!**

Over-Dispersion (Relative to Poisson), Maman et al. ('09)

Israeli-Bank Call-Center

Arrival Counts - Coefficient of Variation (CV), per 30 min.

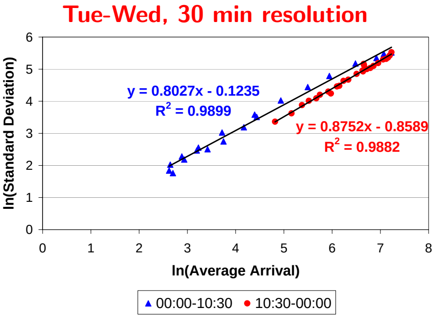
Sampled CV - solid line, **Poisson** CV - dashed line

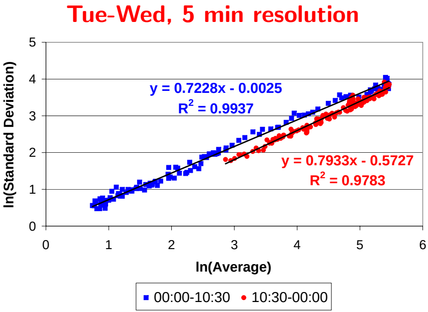


- 263 regular **days**, 4/2007 - 3/2008.
- Poisson CV $= 1/\sqrt{\text{mean arrival-rate}}$.
- Sampled CV's $\gg$ Poisson CV's $\Rightarrow$ **Over-Dispersion**.

Over-Dispersion: Fitting a Regression Model

ln(STD) vs. ln(AVG)





Significant linear relations (Aldor & Feigin):

$$\ln(\mathrm{STD}) = c \cdot \ln(\mathrm{AVG}) + a$$

Over-Dispersion: Random Arrival-Rate Model

The **linear relation** between ln(STD) and ln(AVG) motivates the following model:

Arrivals distributed **Poisson with a Random Rate**

$$\boldsymbol{\Lambda = \lambda + \lambda^{c} \cdot X, \quad 0 \leq c \leq 1 \, ;}$$

- X is a random-variable with $E[X] = 0$, capturing the magnitude of **stochastic deviation** from mean arrival-rate.
- c determines **scale-order** of the over-dispersion:
  $c = 1$, proportional to λ;
  $c = 0$, Poisson-level, same as $0 \leq c \leq 1/2$.

In **call centers**, over-dispersion (per 30 min.) is of order $\boldsymbol{\lambda^{c}}$, $\mathbf{c \approx 0.8 - 0.85}$.

Over-Dispersion: Distribution of X ?

- Fitting a **Gamma Poisson** mixture model to the data:
  Assume a (conjugate) prior Gamma distribution for the arrival rate $\Lambda \stackrel{d}{=} Gamma(a, b)$.
  Then, $Y \stackrel{d}{=} Poiss(\Lambda)$ is Negative Binomial.

- Very good fit of the Gamma Poisson mixture model, to data of the Israeli Call Center, for the majority of time intervals .

- Relation between our c-based model and Gamma-Poisson mixture is established.

- Distribution of X derived, under the Gamma prior assumption:
  X **is asymptotically normal**, as $\lambda \to \infty$.

Over-Dispersion: The QED-c Regime

QED-c Staffing: Under offered-load $\mathbf{R} = \boldsymbol{\lambda} \cdot \mathbf{E[S]}$,

$$\mathbf{n} \;=\; \mathbf{R} \;+\; \boldsymbol{\beta} \cdot \mathbf{R^c}\,, \qquad \mathbf{0.5 < c < 1}$$

Performance measures:

a. Delay probability: $P\{W_q > 0\} \;\sim\; 1 - F(\beta)$

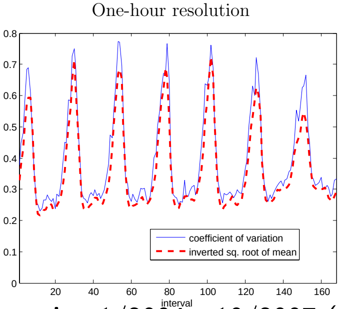
b. Abandonment probability: $P\{Ab\} \;\sim\; \dfrac{E[X-\beta]_+}{n^{1-c}}$

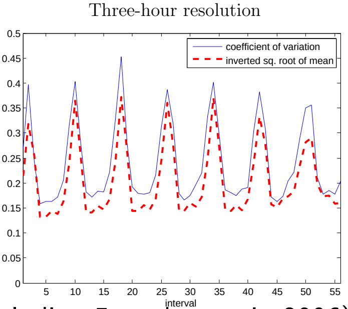
c. Average offered wait: $E[V] \;\sim\; \dfrac{E[X-\beta]_+}{n^{1-c} \cdot g_0}$

d. Average actual wait: $E_{\Lambda,n}[W] \;\sim\; E_{\Lambda,n}[V]$

Israeli-Hospital Emergency-Department

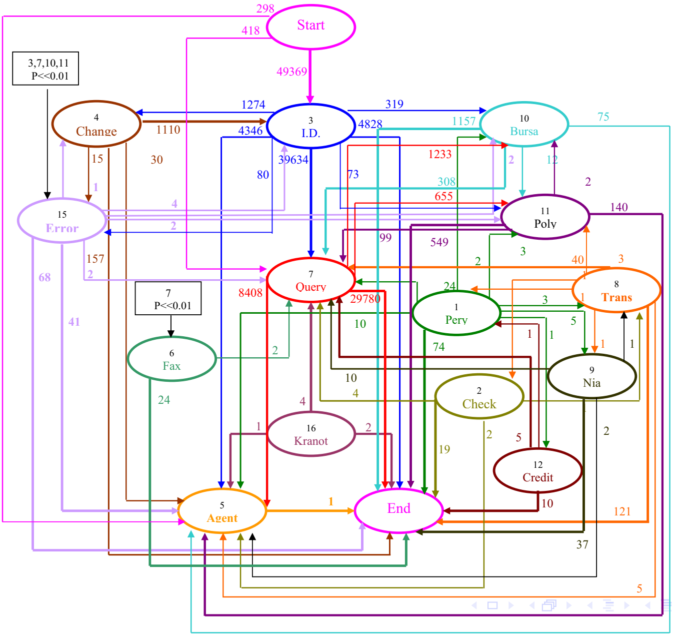
Arrival Counts - Coefficient of Variation, per 1-hr. & 3-hr.





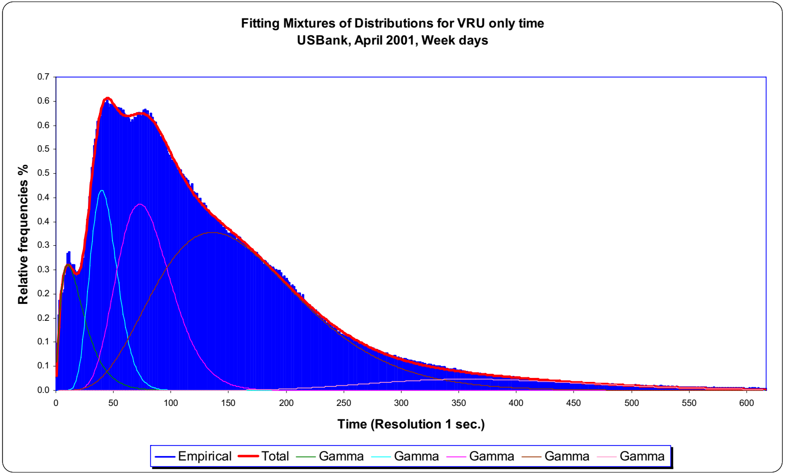
- 194 **weeks**, 1/2004 - 10/2007 (excluding 5 weeks war in 2006).
- Moderate over-dispersion: **$c = 0.5$** reasonable for hourly resolution.
- **ED beds in conventional QED** (**Less** var. than call centers ! ?).

Call Transitions in the IVR - Phase Type



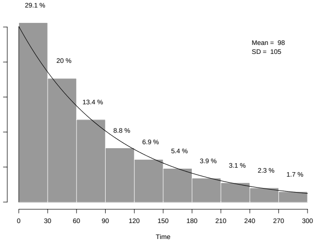
Service Times: Fitting Distribution

Fitting Mixture of 5 Gamma Components

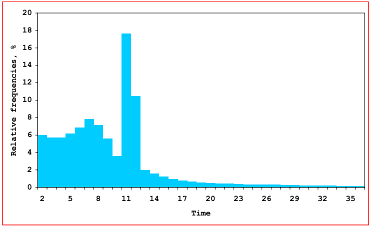


Beyond Averages: Waiting Times in a Call Center

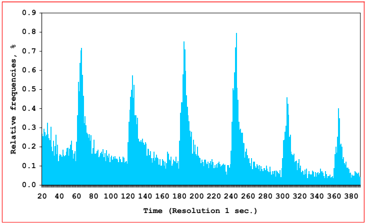
Small Israeli Bank



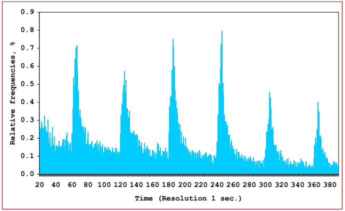
Large U.S. Bank



Medium Israeli Bank



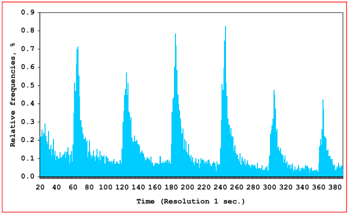
"Waiting-Times" Puzzle at a Large Israeli Bank



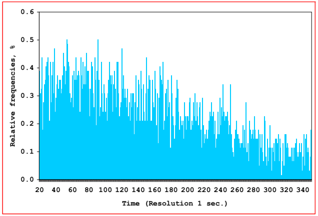
Peaks Every 60 Seconds. **Why?**

- Human: **Voice-announcement** every 60 seconds.
- System: **Priority-upgrade** (unrevealed) every 60 secs (Theory?)

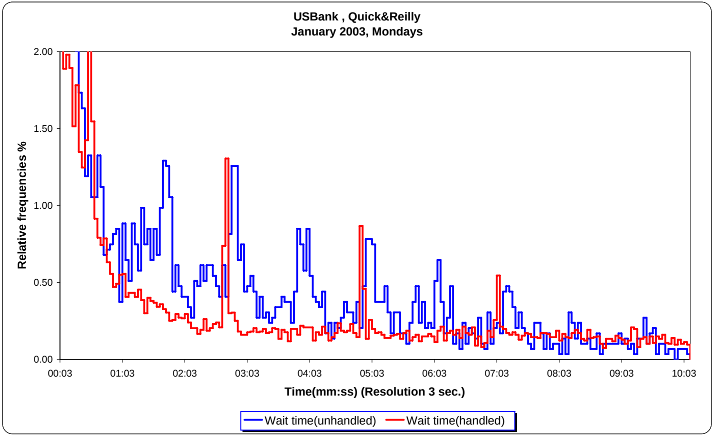
Served Customers



Abandoning Customers

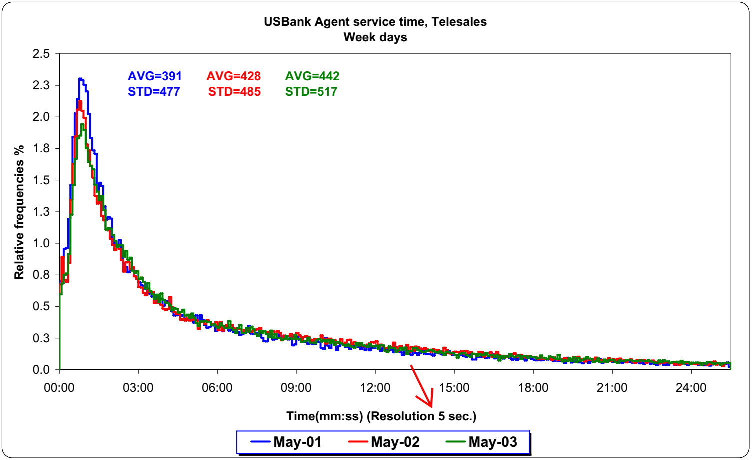


Still a Puzzle at a US Bank



- Different **cycles of peaks** in the waiting times of both served (protocol?) and abandoning (psychology?) customers.
- **No theory** for periodic updates of either priorities or information.

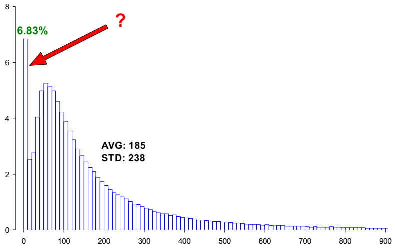
US Bank: Service Time Histograms for Telesales, 2001-3



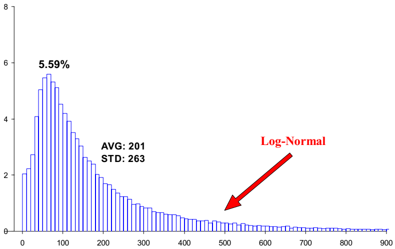
Service Times: Distribution and Psychology

Histogram of Service Times in a Small Israeli Bank

January-October



November-December

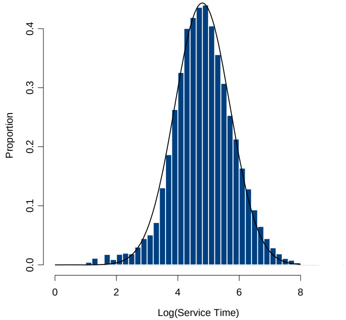


- **Lognormal** service times prevalent in call centers
- **6.8% Short-Services:** Agents' "Abandon" (improve bonus, rest)
- **Distributions**, not only Averages, must be measured.

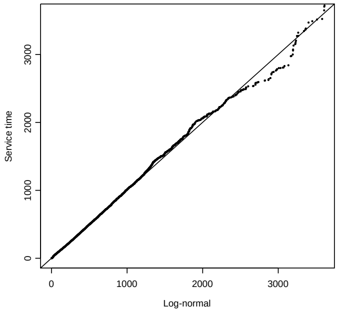
Validating LogNormality of Service Times

Israeli Call Center, Nov-Dec, 1999

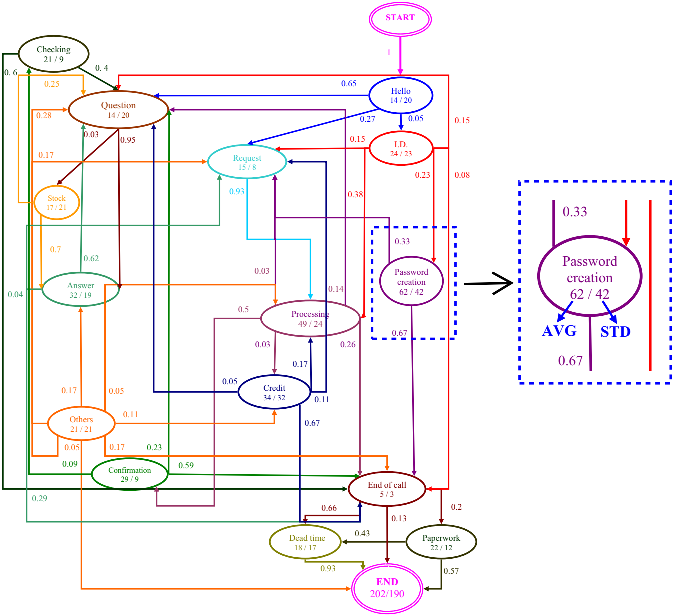
Log(Service Times)



LogNormal QQPlot



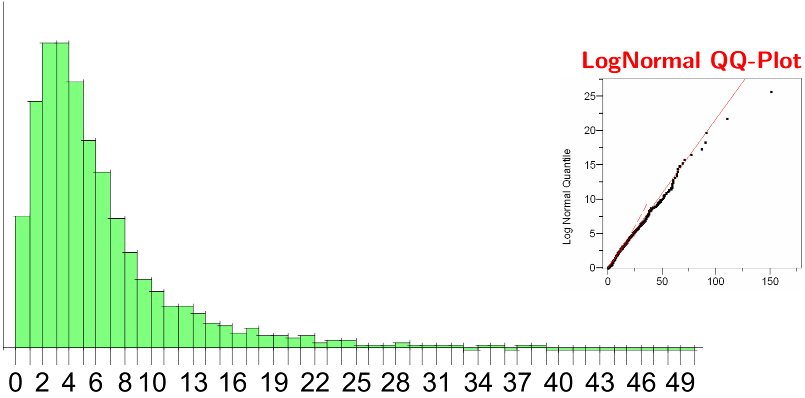
Israeli Bank, Retail Service



Israeli Large Hospital: LOS in IW

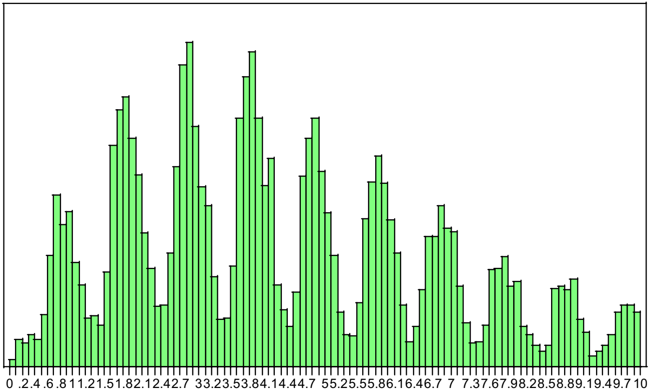
Days Resolution

LogNormal QQ-Plot



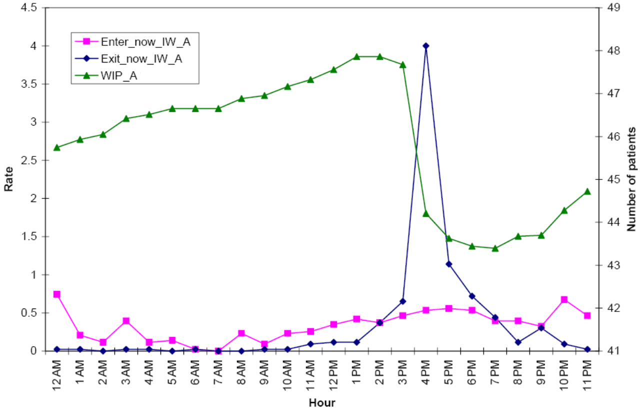
Length of Stay: Resolution Dependence

Hours Resolution: LN = Normal Mixture



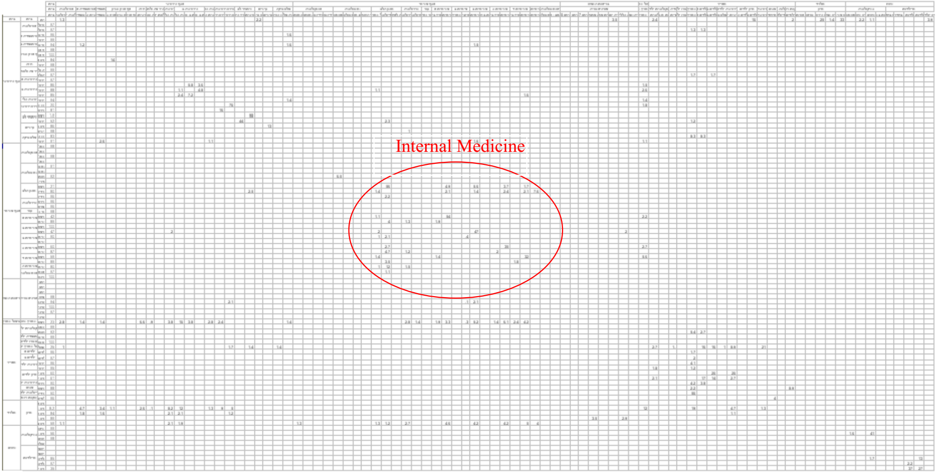
Length of Stay: Resolution Dependence

Internal Ward A: **Arrivals** / **Departures** / **# Patients** , by hour



Ongoing: Empirical Analysis of an ED, IW and Everything In Between, w/ Y. Marmor, Y. Tseytlin, G. Yom-Tov, M. Armony.

90 × 90 Matrix, Sub-Ward Resolution



8×8 Matrix, Division Resolution

Including Arrivals and Releases

	Home	Surgery	Internal	Psychology	Intensive Care	Pediatrics	Emergency Dep.	Gynecology
Home		8.4	3.2	0.1		18.3	**60.3**	9.7
Surgery	90	7.9	1.3		0.7	0.1		
Internal	84.4	1.9	13	0.1	0.5			0.1
Psychology	94.3	1.9	3.8					
Intensive Care	17.2	**40.9**	**38.4**			0.9		2.6
Pediatrics	78.8	0.6				20.6		
Emergency Dep.	69.9	8.9	**19.2**	0.2	0.3	1		0.5
Gynecology	55.3	0.3	0.2		0.1			44.1

Transitions Inside the Hospital

	Surgery	Internal	Psychology	Intensive Care	Pediatrics	Emergency Dep.	Gynecology
Surgery	**78.3**	12.7	0.2	7	1.4		0.4
Internal	12	**83.3**	0.6	3.4	0.2		0.5
Psychology	33.3	66.7					
Intensive Care	**49.5**	**46.4**			1		3.1
Pediatrics	2.6	0.2		0.1	**96.9**		0.2
Emergency Dep.	**29.7**	**63.7**	0.6	0.9	3.4		1.7
Gynecology	0.7	0.4		0.1			**98.8**

- About 50% of transitions between ED and internal wards.
- Most transitions are inside the specific hospitalized unit.

IW Operational Measures, or Efficiency vs. Fairness

Israeli Large Hospital (1/5/06 to 30/10/08, excluding 1-3/07)

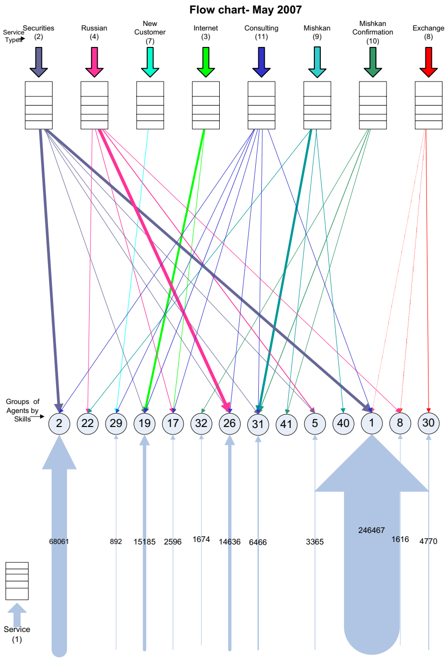
	Ward A	Ward B	Ward C	Ward D
ALOS (days)	6.37	4.47	5.36	5.56
Avg Occupancy Rate	97%	95%	86%	92%
Avg # Patients per Month	206	187	210	210
Standard capacity	45	30	44	42
Avg # Patients /Bed/Month	4.57	6.25	4.77	4.77
Return Rate	15.4%	15.6%	16.2%	14.8%

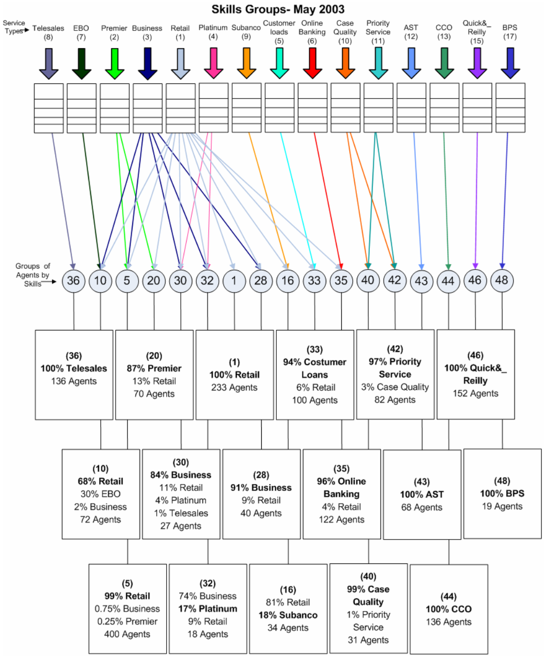
IW Operational Measures, or Efficiency vs. Fairness

Israeli Large Hospital (1/5/06 to 30/10/08, excluding 1-3/07)

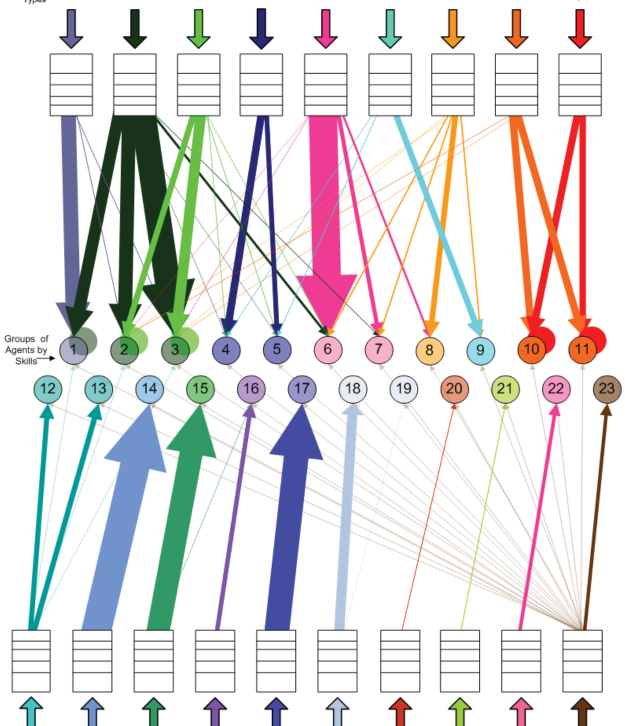
	Ward A	Ward B	Ward C	Ward D
ALOS (days)	6.37	4.47	5.36	5.56
Avg Occupancy Rate	97%	95%	86%	92%
Avg # Patients per Month	206	187	210	210
Standard capacity	45	30	44	42
Avg # Patients /Bed/Month	4.57	6.25	4.77	4.77
Return Rate	15.4%	15.6%	16.2%	14.8%

- The "fastest" + smallest Ward B subject to highest workload: **occupancy, flux**: **unfair**.
- Calls for ED-to-IW routing, which is both efficient and fair (w/ Tseytlin (MSc), Tseytlin & Momcilovic, Tseytlin & Zviran): **exact analysis, QED approximation (natural - hours wait for days service), partial bed-information**.


Flow chart- May 2007
Service Types
Securities (2)
Russian (4)
New Customer (7)
Internet (3)
Consulting (11)
Mishkan (9)
Mishkan Confirmation (10)
Exchange (8)
Groups of Agents by Skills
2
22
29
19
17
32
26
31
41
5
40
1
8
30
68061
892
15185
2596
1674
14636
6466
3365
246467
1616
4770
Service (1)



Skills Groups- May 2003
Service Types
Telesales (8)
EBO (7)
Premier (2)
Business (3)
Retail (1)
Platinum (4)
Subanco (9)
Customer loads (5)
Online Banking (6)
Case Quality (10)
Priority Service (11)
AST (12)
CCO (13)
Quick&_Reilly (15)
BPS (17)
Groups of Agents by Skills
36
10
5
20
30
32
1
28
16
33
35
40
42
43
44
46
48
(36) 100% Telesales 136 Agents
(20) 87% Premier 13% Retail 70 Agents
(1) 100% Retail 233 Agents
(33) 94% Costumer Loans 6% Retail 100 Agents
(42) 97% Priority Service 3% Case Quality 82 Agents
(46) 100% Quick&_Reilly 152 Agents
(10) 68% Retail 30% EBO 2% Business 72 Agents
(30) 84% Business 11% Retail 4% Platinum 1% Telesales 27 Agents
(28) 91% Business 9% Retail 40 Agents
(35) 96% Online Banking 4% Retail 122 Agents
(43) 100% AST 68 Agents
(48) 100% BPS 19 Agents
(5) 99% Retail 0.75% Business 0.25% Premier 400 Agents
(32) 74% Business 17% Platinum 9% Retail 18 Agents
(16) 81% Retail 18% Subanco 34 Agents
(40) 99% Case Quality 1% Priority Service 31 Agents
(44) 100% CCO 136 Agents


Types
Groups of
Agents by
Skills
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

Service Rate: Class or Pool Dependent?

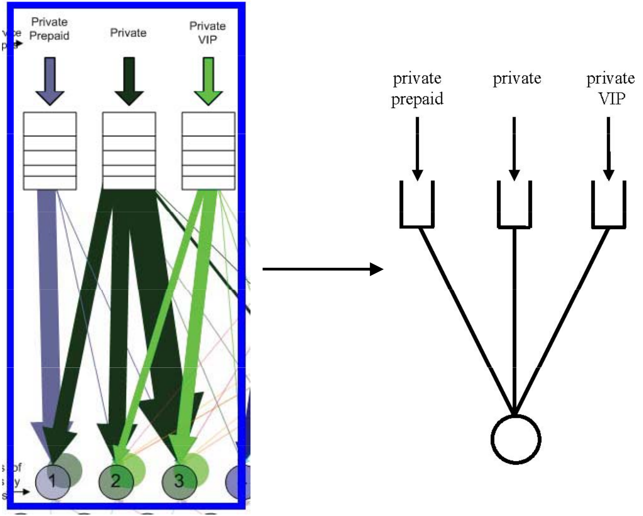
Agents Group\Service Class	Private Prepaid	Private	Private VIP
Private Prepaid	163.1	236.1	
Private - Private VIP (1)		243.5	195.1
Private - Private VIP (2)		244	201.4

$\Rightarrow$ **Class-dependent** service rate

Agents Group\Service Class	Business	Business VIP	Business Preservation
Business (1)	276.9	261.5	
Business (2)	336.7	334.5	
Business VIP	315.9	280.5	
Business Preservation		386.2	634.1

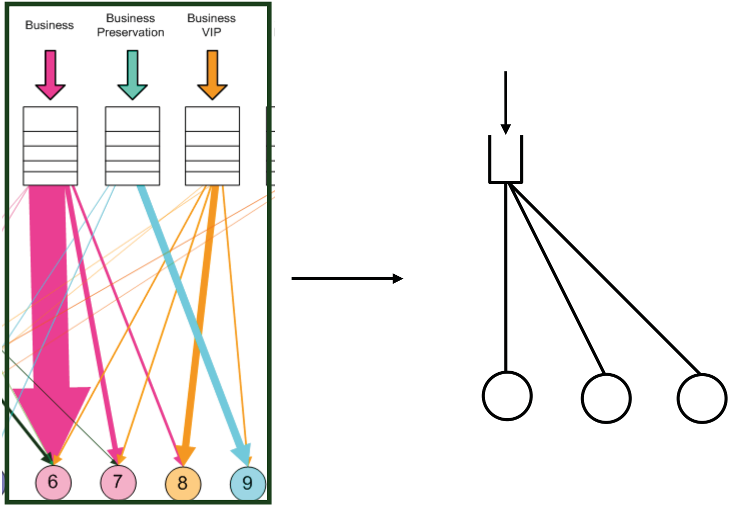
$\Rightarrow$ **Pool-dependent** service rate

Class-Dependent $\approx$ V-Model



Many-Server Approximations: State-Space Collapse

Pool-Dependent $\approx$ Λ-Model



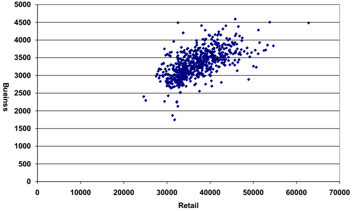
Unpredictable Variability: The Multi-Class Case

Unpredictable variability: $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_I)$

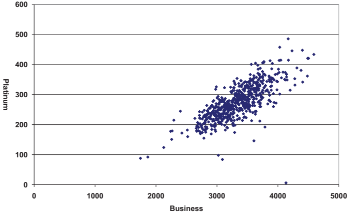
Pairs: $(\mathbf{X}_{Retail}, \mathbf{X}_{Business})$ and $(\mathbf{X}_{Business}, \mathbf{X}_{Platinum})$

US Bank: Correlations, 600 weekdays (Gurvich et al., '09)

Business vs. Retail

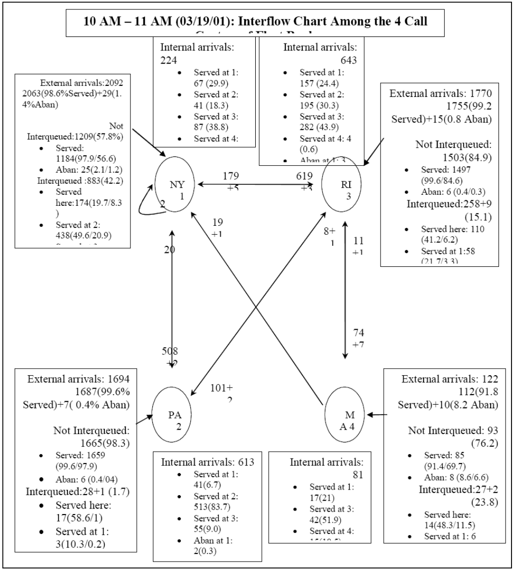


Business vs. Platinum



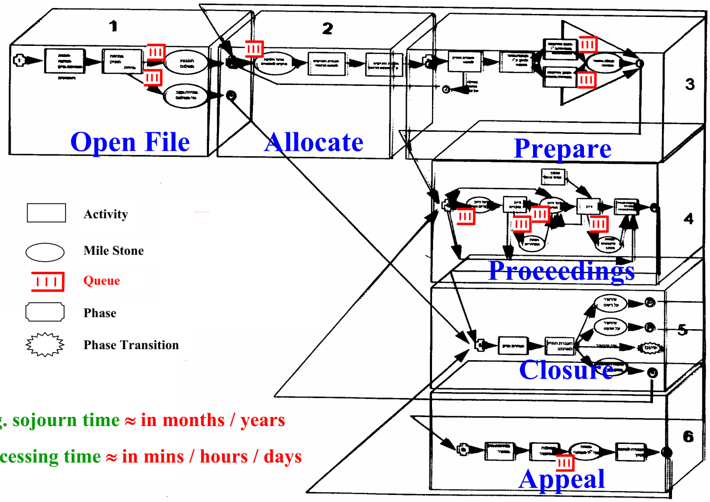
- Positive correlation (not independent)
- Research: Impact on design and control decisions ?

US Bank


10 AM – 11 AM (03/19/01): Interflow Chart Among the 4 Call
External arrivals:2092
2063(98.6%Served)+29(1.4%Aban)
Not Interqueued:1209(57.8%)
Served: 1184(97.9/56.6)
Aban: 25(2.1/1.2)
Interqueued :883(42.2)
Served here:174(19.7/8.3)
Served at 2: 438(49.6/20.9)
Internal arrivals: 224
Served at 1: 67 (29.9)
Served at 2: 41 (18.3)
Served at 3: 87 (38.8)
Served at 4:
Internal arrivals: 643
Served at 1: 157 (24.4)
Served at 2: 195 (30.3)
Served at 3: 282 (43.9)
Served at 4: 4 (0.6)
External arrivals: 1770
1755(99.2 Served)+15(0.8 Aban)
Not Interqueued: 1503(84.9)
Served: 1497 (99.6/84.6)
Aban: 6 (0.4/0.3)
Interqueued:258+9 (15.1)
Served here: 110 (41.2/6.2)
Served at 1:58
NY 1
RI 3
PA 2
M A 4
179
619
19 +1
20
8+ 1
11 +1
74 +7
508
101+
External arrivals: 1694
1687(99.6% Served)+7(0.4% Aban)
Not Interqueued: 1665(98.3)
Served: 1659 (99.6/97.9)
Aban: 6 (0.4/04)
Interqueued:28+1 (1.7)
Served here: 17(58.6/1)
Served at 1: 3(10.3/0.2)
Internal arrivals: 613
Served at 1: 41(6.7)
Served at 2: 513(83.7)
Served at 3: 55(9.0)
Aban at 1: 2(0.3)
Internal arrivals: 81
Served at 1: 17(21)
Served at 3: 42(51.9)
Served at 4:
External arrivals: 122
112(91.8 Served)+10(8.2 Aban)
Not Interqueued: 93 (76.2)
Served: 85 (91.4/69.7)
Aban: 8 (8.6/6.6)
Interqueued:27+2 (23.8)
Served here: 14(48.3/11.5)
Served at 1: 6

Conceptual Model: The "Production of Justice"

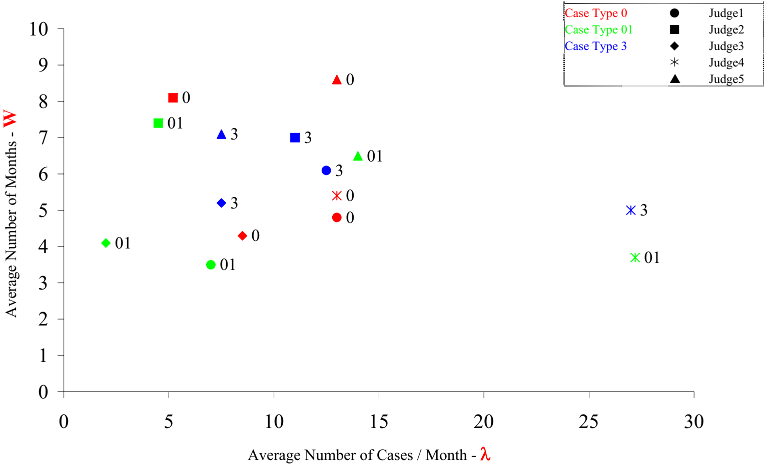
The Labor-Court Process in Haifa, Israel

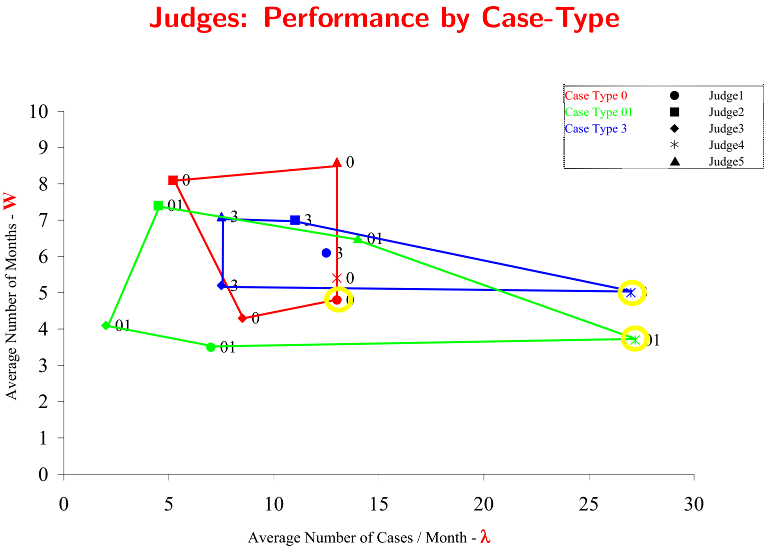


Avg. sojourn time ≈ in months / years

Processing time ≈ in mins / hours / days

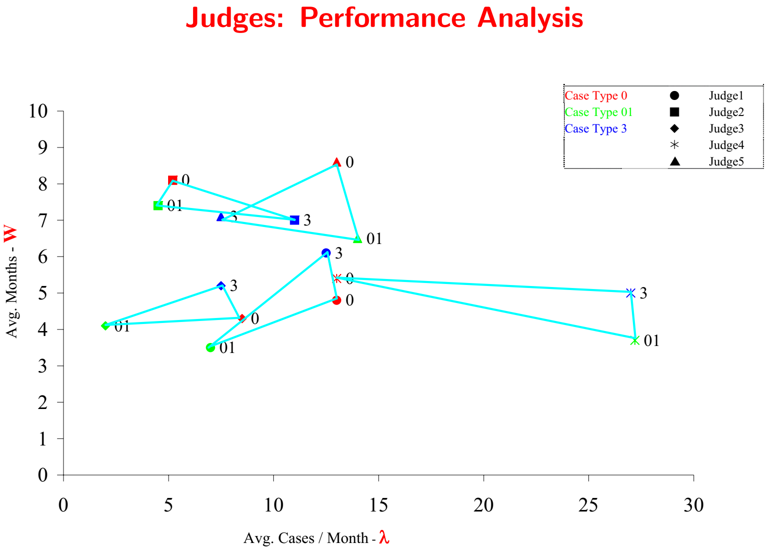
Judges: Operational Performance - Base Case



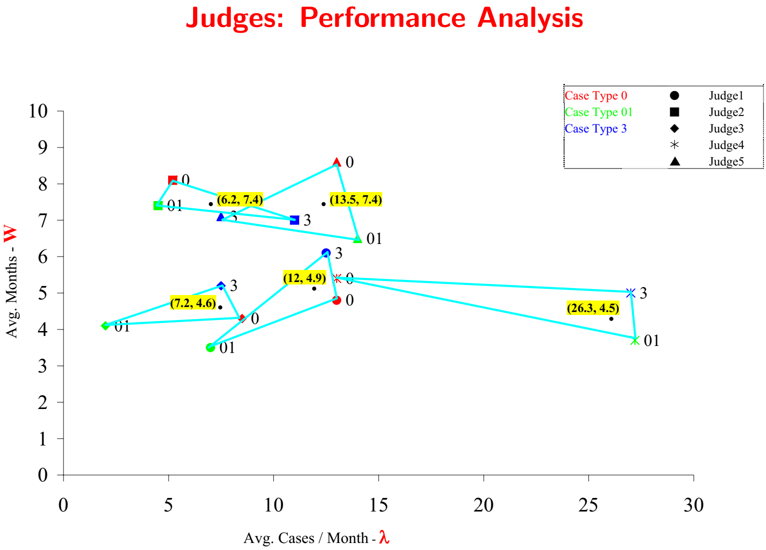


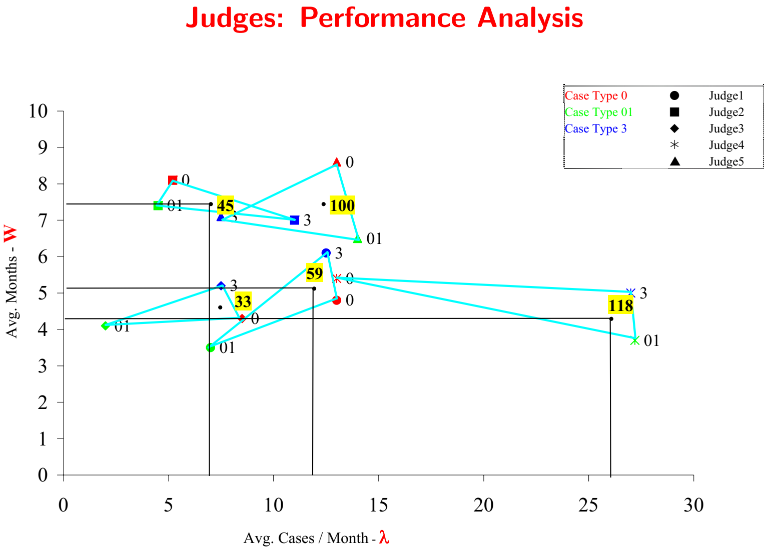
Judges: Performance by Case-Type
Case Type 0
Case Type 01
Case Type 3
Judge1
Judge2
Judge3
Judge4
Judge5
Average Number of Months - W
Average Number of Cases / Month - λ
0
5
10
15
20
25
30
1
2
3
4
6
7
8
9
01

Analytical Model: Little's Law in Court (III)



Analytical Model: Little's Law in Court (IV)

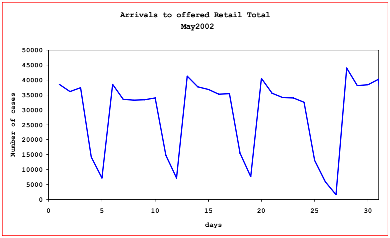



Judges: Performance Analysis
Case Type 0
Case Type 01
Case Type 3
Judge1
Judge2
Judge3
Judge4
Judge5
Avg. Months - W
Avg. Cases / Month - λ
0
1
2
3
4
5
6
7
8
9
10
0
5
10
15
20
25
30
45
100
59
33
118
01
3

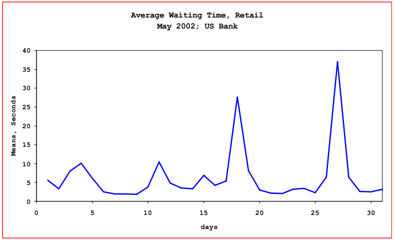
Little's Law, $L = \lambda \cdot W$

US Bank: Retail calls, May 2002

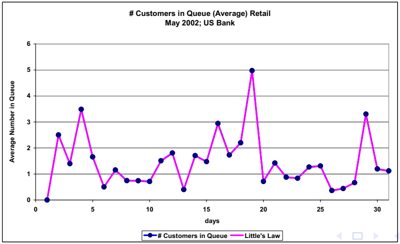
λ, Throughput Rate



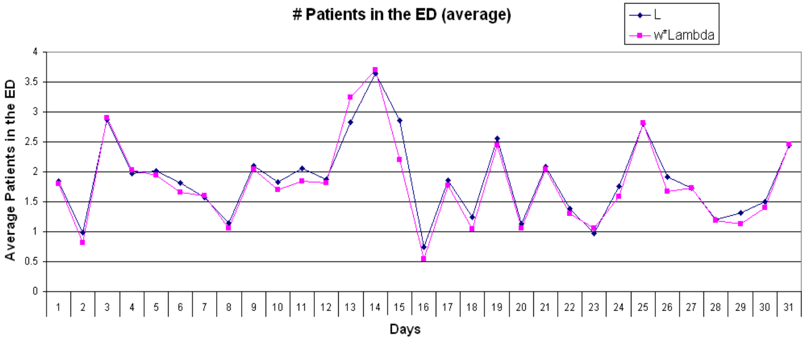
W, Average Waiting Time



L, Average Queue Length



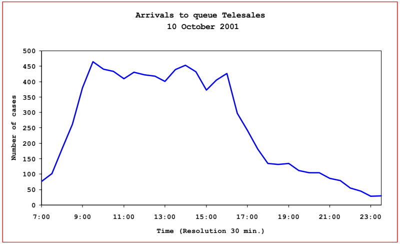
Israeli ED, October 1999, Day Resolution



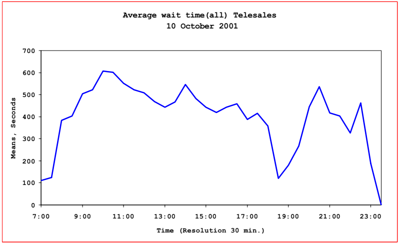
Little's Law, $L = \lambda \cdot W$

US Bank: Telesales Calls, October 10, 2001

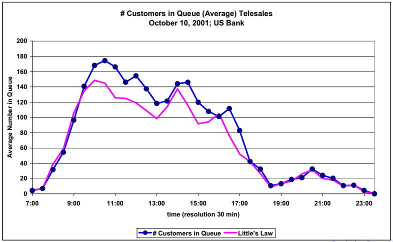
λ, Throughput Rate



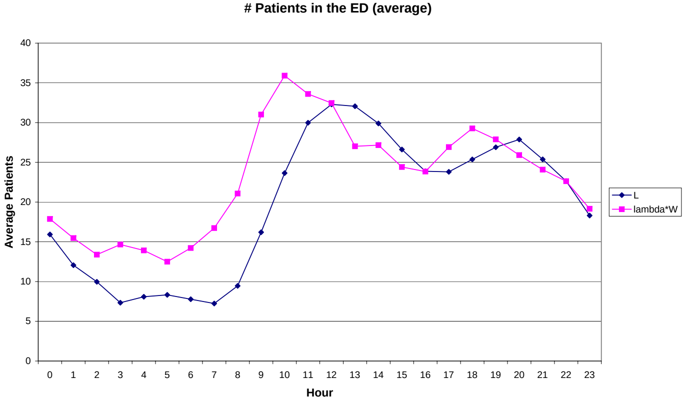
W, Average Waiting Time



L, Average Queue Length



Israeli ED, Hour Resolution



Workload and Offered-Load

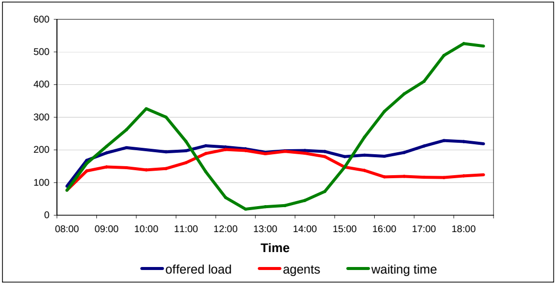
- **Workload**: Stochastic process, representing the amount of work present at time t, under the assumptions of **infinitely many resources** (service commences immediately upon arrival).
- **Offered-Load**: Function of time $t \geq 0$, representing the average of the workload at time t.

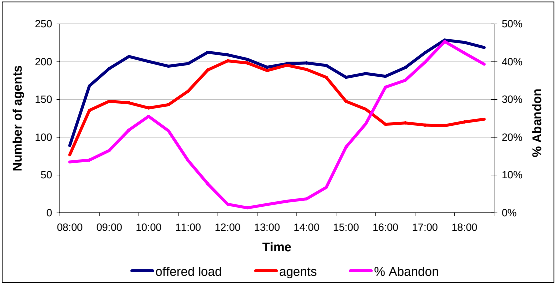
The Offered-Load, $R(t)$, determines staffing level via c-staffing ($c = 0.5$ is conventional square-root staffing):

$$N(t) = R(t) + \beta \cdot [R(t)]^c$$

Offered-Load vs. # Agents

Israeli Cable Company,
Retail Service,
January 2009





For the $M_t/GI/N_t + GI$ queue, the **offered-load** $R = \{R(t),\ t \geq 0\}$, has the following representations:

$$R(t) = E[L(t)] = \int_{-\infty}^{t} \lambda(u) \cdot P(S > t - u) du = E\Big[A(t) - A(t - S)\Big] =$$
$$= E\left[\int_{t-S}^{t} \lambda(u) du\right] = E[\lambda(t - S_e)] \cdot E[S]\,,$$

where
$A = \{A(t),\ t \geq 0\}$ is the Arrival process;
S is a generic service time;
S_e is a generic excess (residual) service.

In stationary models, where $\lambda(t) \equiv \lambda$, the offered-load $R(t)$ is the familiar $\lambda \cdot E[S]$ (or λ/μ), measured in Erlangs.

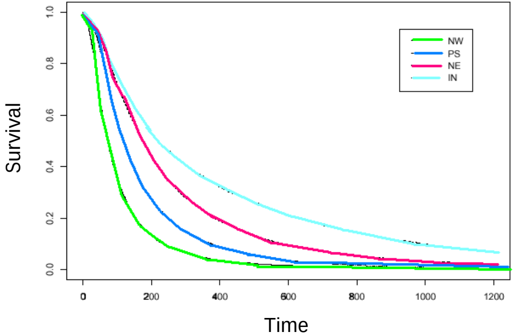
Imputing Service Times of Abandoning Customers

In calculating the offered-load, one must account for service-times of abandoning customers.

A prevalent assumptions is that service times and (im)patience times are independent. Experience suggests that this assumption is often violated.
For example, it is not unreasonable that customers who anticipate longer service times, will be willing to wait more for service before abandoning.

Service Times: Stochastic Order

Small Israeli Bank: Survival Functions by Type



Service times stochastic order: $\mathbf{S_{NW}} \overset{st}{<} \mathbf{S_{PS}} \overset{st}{<} \mathbf{S_{NE}} \overset{st}{<} \mathbf{S_{IN}}$

Patience times stochastic order: $\tau_{NW} \overset{st}{<} \tau_{IN} \overset{st}{<} \tau_{PS} \overset{st}{<} \tau_{NE}$

Relationship Between Service-Time and (Im)Patience

Ongoing research (w/ M. Reich, Y. Ritov) develops a procedure for calculating the function $E(S|\tau = w)$:

1. Introduce $g(w) = E(S|\tau > W = w)$, which is the mean service time of those who waited exactly w units of time **and were served**. Then calculate g via the non-linear regression:

   $$S_i = g(W_i) + \varepsilon_i \ ,$$

   where i indexing served customers.

2. Calculate $E(S|\tau = w)$ via the (established) relation

   $$E(S|\tau = w) = g(w) - \frac{g'(w)}{h_\tau(w)} \ ,$$

   where $h_\tau(w)$ is the hazard-rate function of (im)patience, to be estimated via un-censoring.

Finally, extend the above to calculate the distribution of S, given w, which is then used to impute service-times for calculating the offered-load.