A Data-Driven Model of an Appointment-Generated Arrival Process at an Endocrinology Outpatient Clinic

Ward Whitt

Department of Industrial Engineering and Operations Research

Columbia University, New York, NY 10027

URL: http://www.columbia.edu/∼ww2040

joint work with

Song-Hee Kim, USC Marshall School of Business

(former doctoral student)

and

Won Chul Cha, Department of Emergency Medicine, Samsung Medical Center, Seoul, Korea

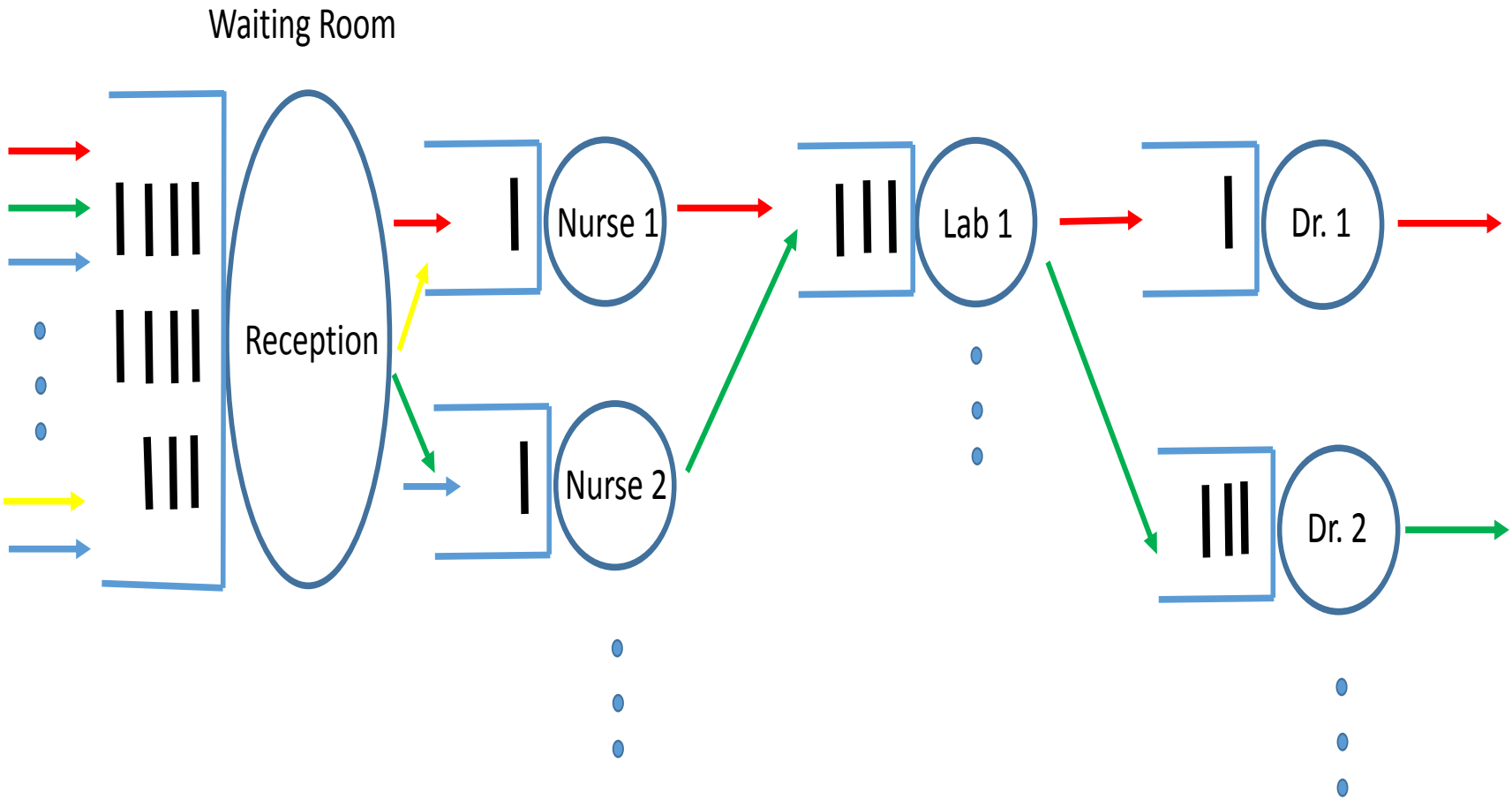
Modeling Outpatient Clinics: A Long History

- Bailey, N. T. J. 1952. A Study of Queues and Appointment Systems in Hospital Outpatient Departments, with Special Reference to Waiting Times. *Journal of the Royal Statistical Society* A14:185-199.
- Welch, J. D., N. T. J. Bailey. 1952. Appointment Systems in Hospital Outpatient Departments, *Lancet* May 31:1105-1108.
- **Fetter, R. B., J. D. Thompson. 1965. The Simulation of Hospital Systems. Operations Research 13(5):689-711.**
- **Jun, J. B., S. H. Jacobson, J. R. Swisher. 1999. Application of Discrete-Event Simulation in Health Care Clinics: A Survey. The Journal of the Operational Research Society 50(2):109-123. (117 refs., 610 citations)**
- Swisher, J. R., S. H. Jacobson, J. B. Jun, O. Balci. 2001. Modeling and Analyzing a Physician Clinic Environment Using Discrete-Event (Visual) Simulation. *Computers and Operations Research* 28: 105-125.
- Chand, S., H. Moskowitz, J. B. Norris, S. Shade, D. R. Willis. 2009. Improving Patient Flow at an Outpatient Clinic: Study of Sources of Variability and Improvement Factors. *Health Care Management Science* 12:325-340.

What Was Done **Before** and What We Have Done **Now**

- Summary of the Literature
  - Common Goal: Improve performance, e.g., reduce congestion.
  - Common Model: Complex multi-class open network of queues
  - Common Analysis Tool: Simulation
- Focus of the Current Work
  - **Probe deeply into a patient arrival process.**
  - Better understand what arrival process model is appropriate.
    - **What is a good stochastic model?**
  - Provide template for better appointment-generated arrival process modeling.

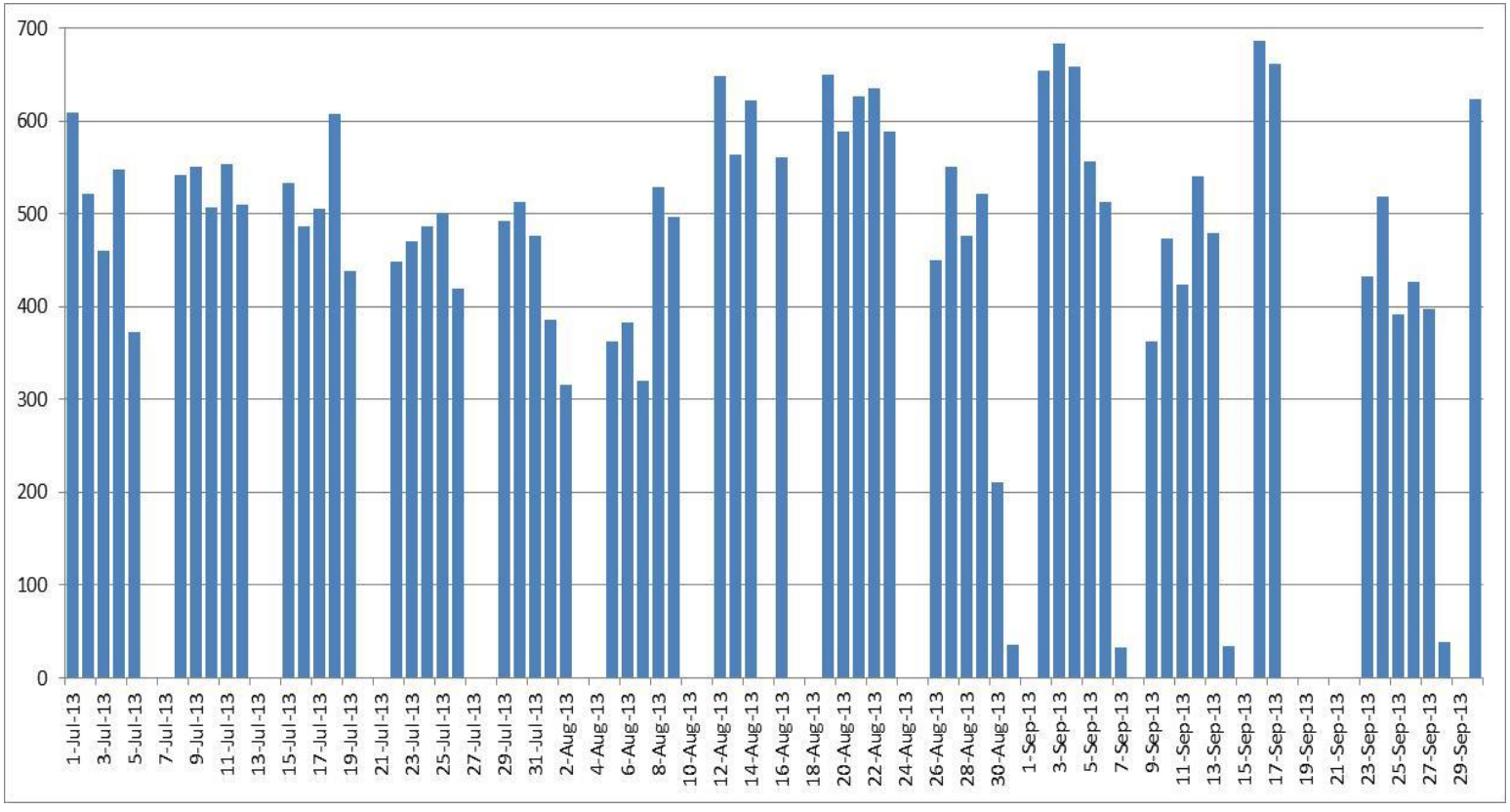
A Clinic as a Network of Queues



Data from an Endocrinology Outpatient Clinic
Samsung Medical Center, Seoul, Korea

- **Endocrinology**: deals with endocrine system (glands that excrete hormones into the blood stream)
- Three months: July 1-September 30, 2013
- Appointments to see one of sixteen doctors
  - day and time when appointment made, when scheduled, and if the patient came
- Outpatient only (within one day)
- 39,253 entries; 8500 cancel, 3000 no shows: **27,800 show up**

Total Daily Arrivals for All 16 Doctors Over 62 Days



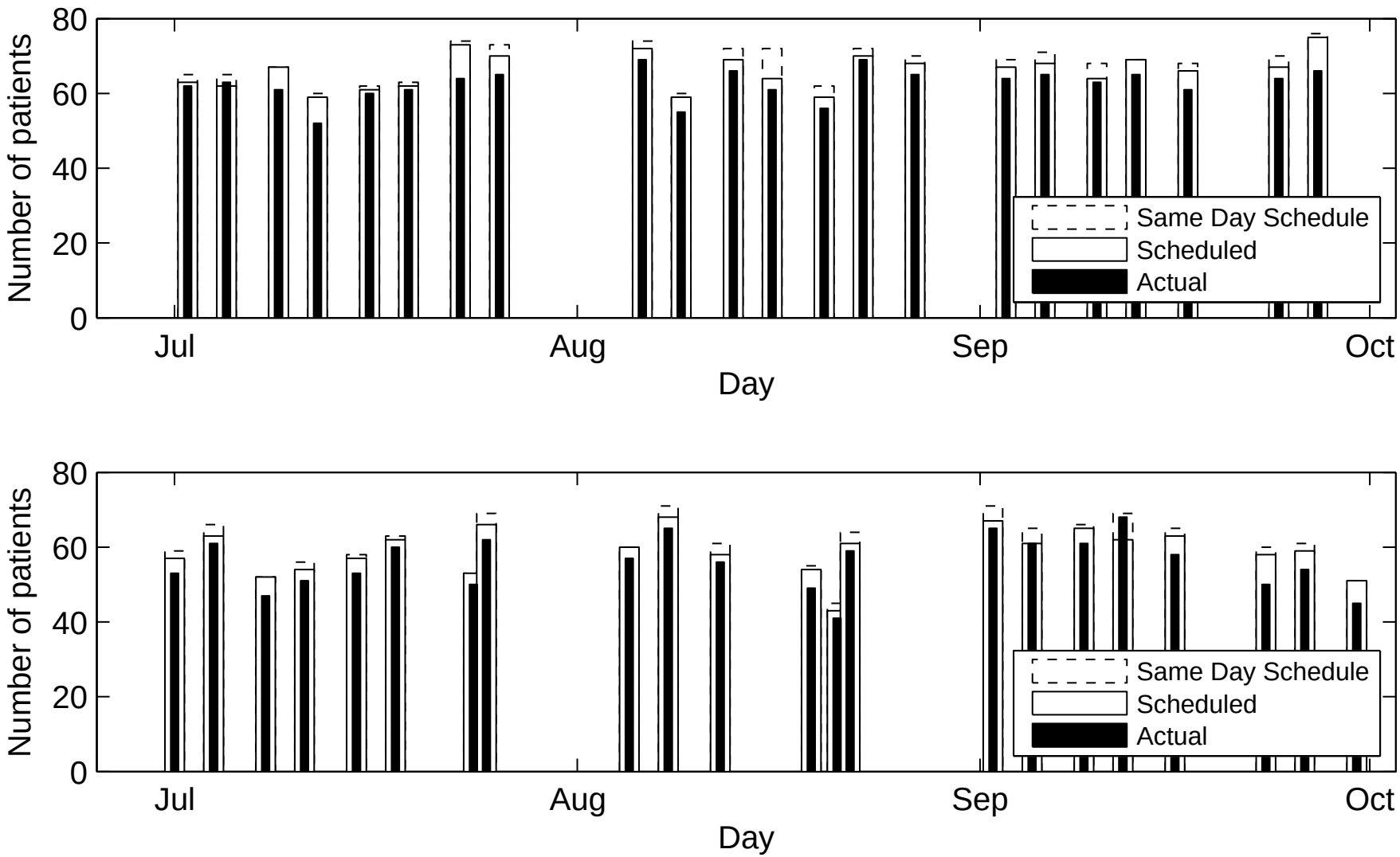
Focus on One Doctor: Doctor 9

22 morning shifts

8:00am-1:00pm

Average of 66 patients per day
(relatively large scale)

22 am shifts and 22 pm shifts for Doctor 9



Sources of Randomness

- **filling schedule over time**
- **adherence to the schedule**
  - **no shows**
  - **extra unscheduled arrivals**
  - **lateness or earliness**

Sources of Randomness

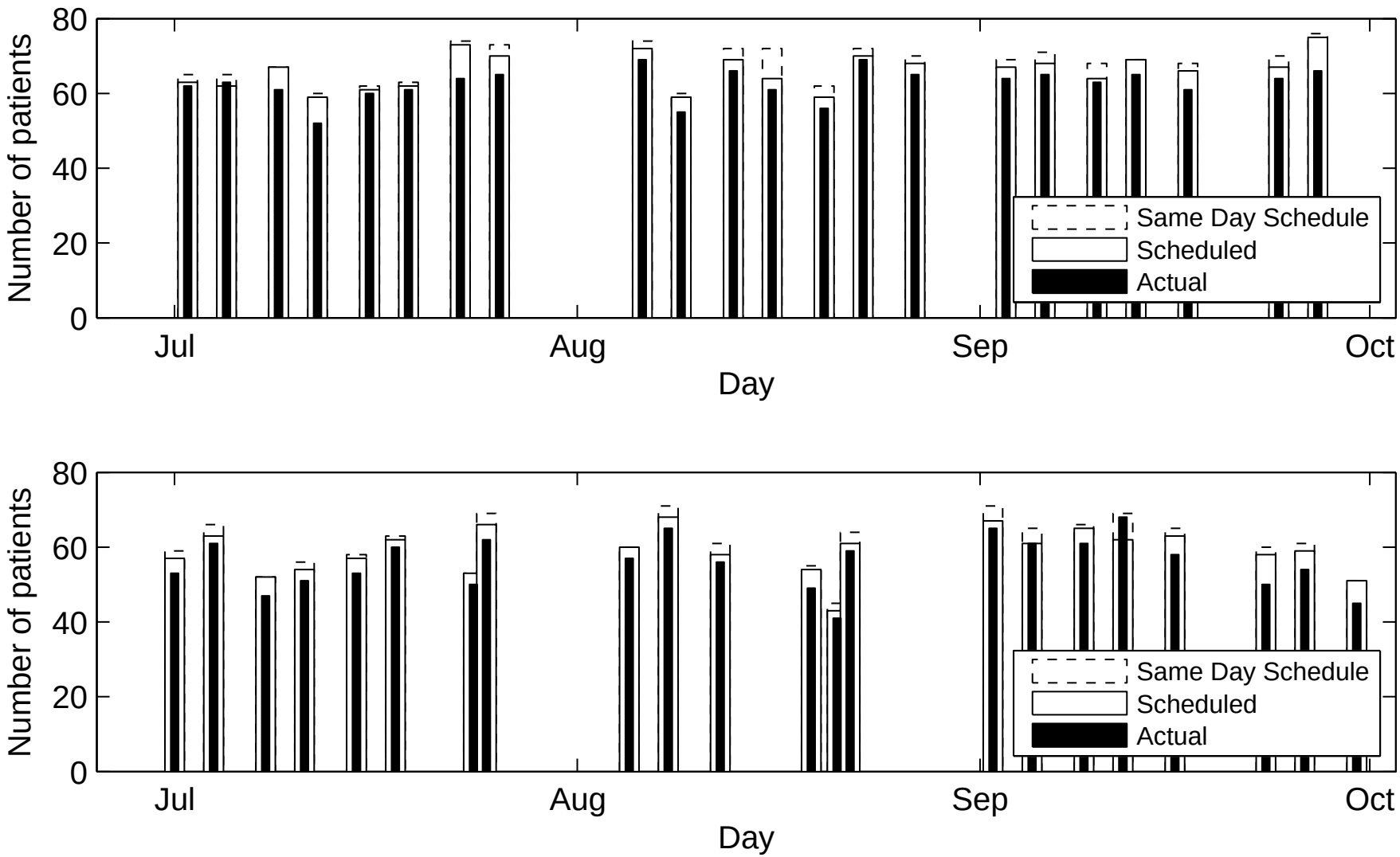
- **filling schedule over time**
- **adherence to the schedule**
  - **no shows**
  - **extra unscheduled arrivals**
  - **lateness or earliness**

- **Our main conclusion for this clinic:**

  The greatest source of randomness is the schedule itself.

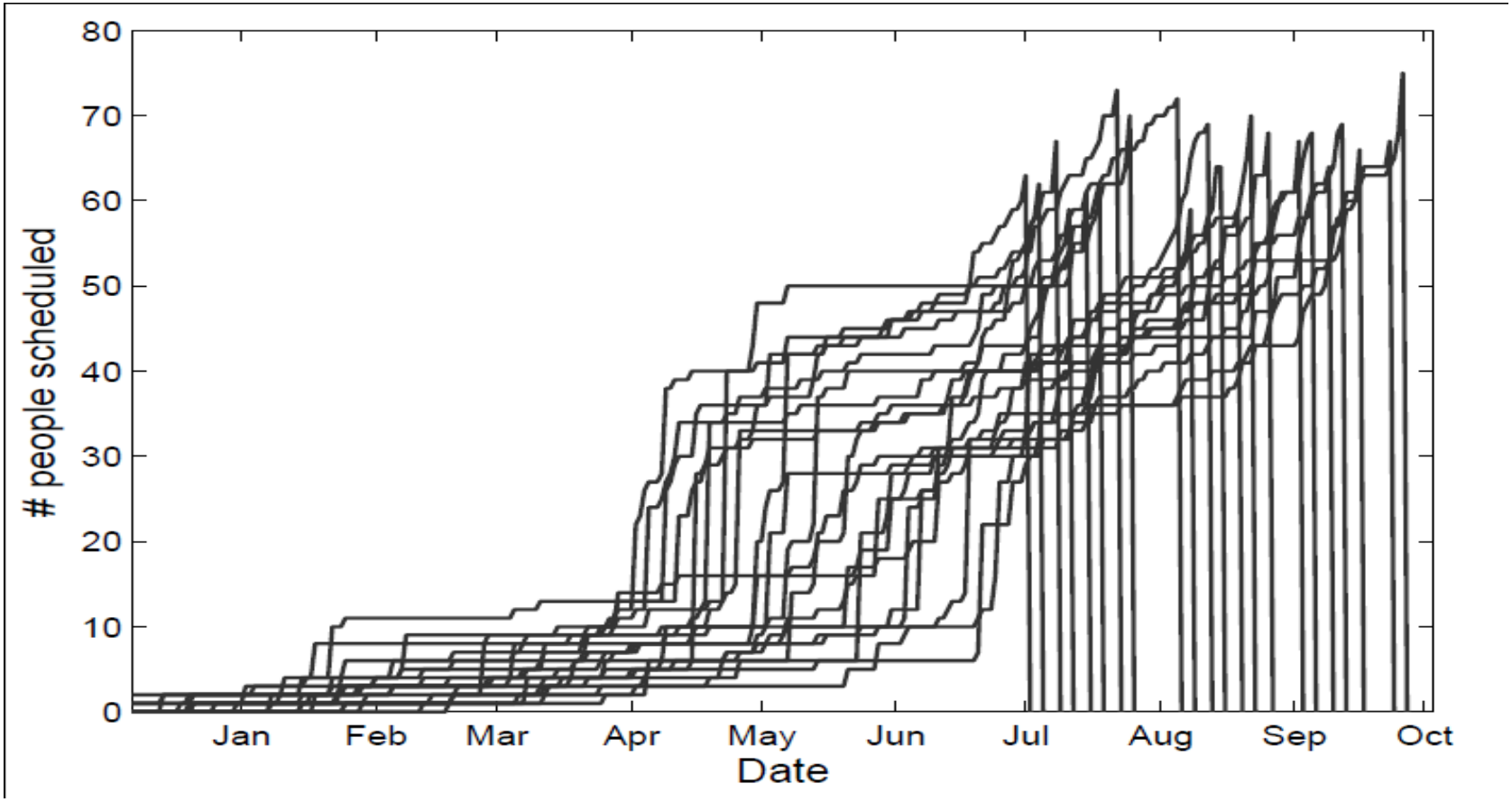
  (The schedule is defined at the end of the previous day.)
- The schedule can be managed!!

22 am shifts and 22 pm shifts for Doctor 9

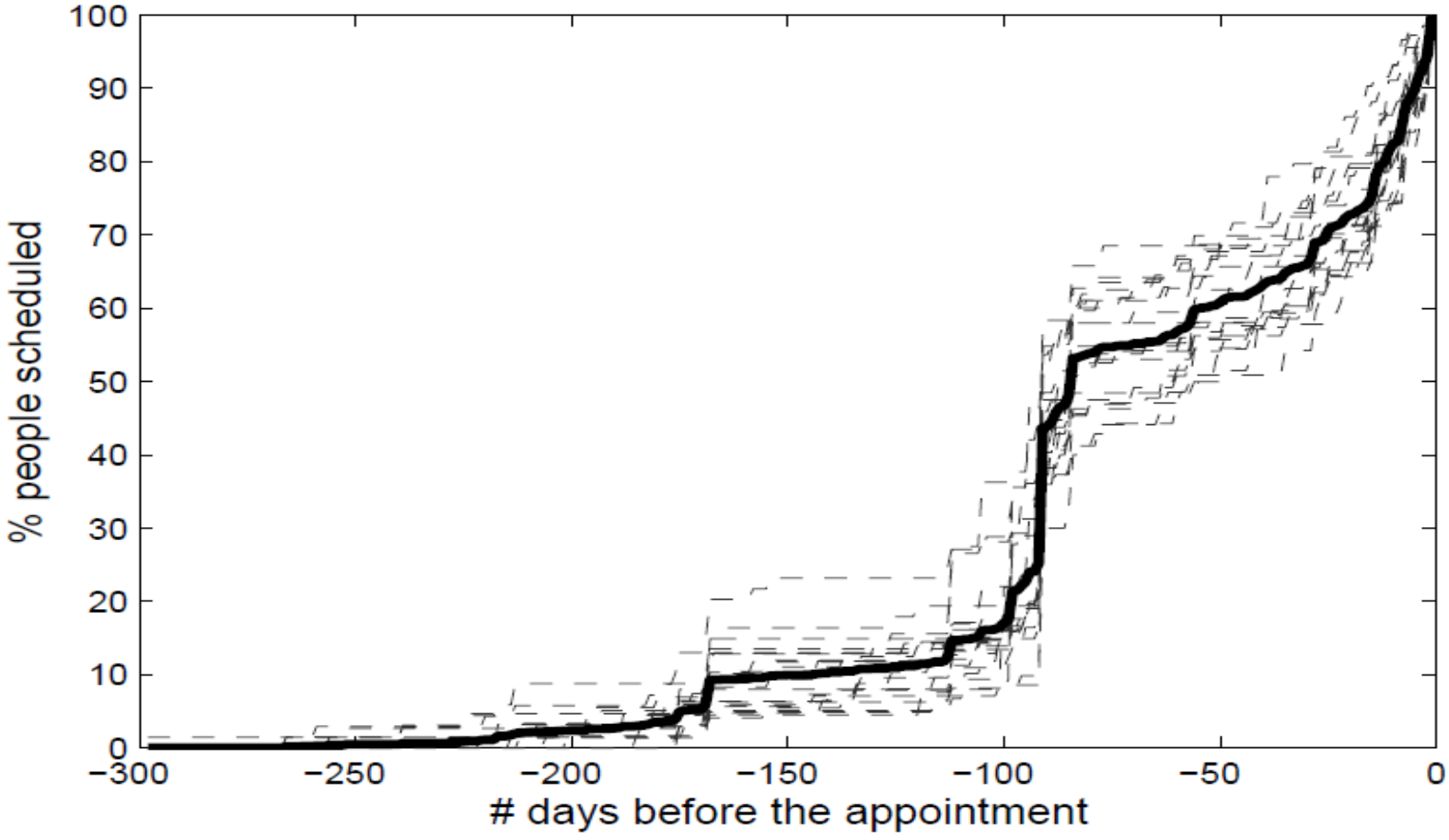


Evolution of the Schedules for Doctor 9: 22 am Shifts

Evolution of the Schedules for Doctor 9: 22 am Shifts



When the Schedule is Made: Number of Days Before



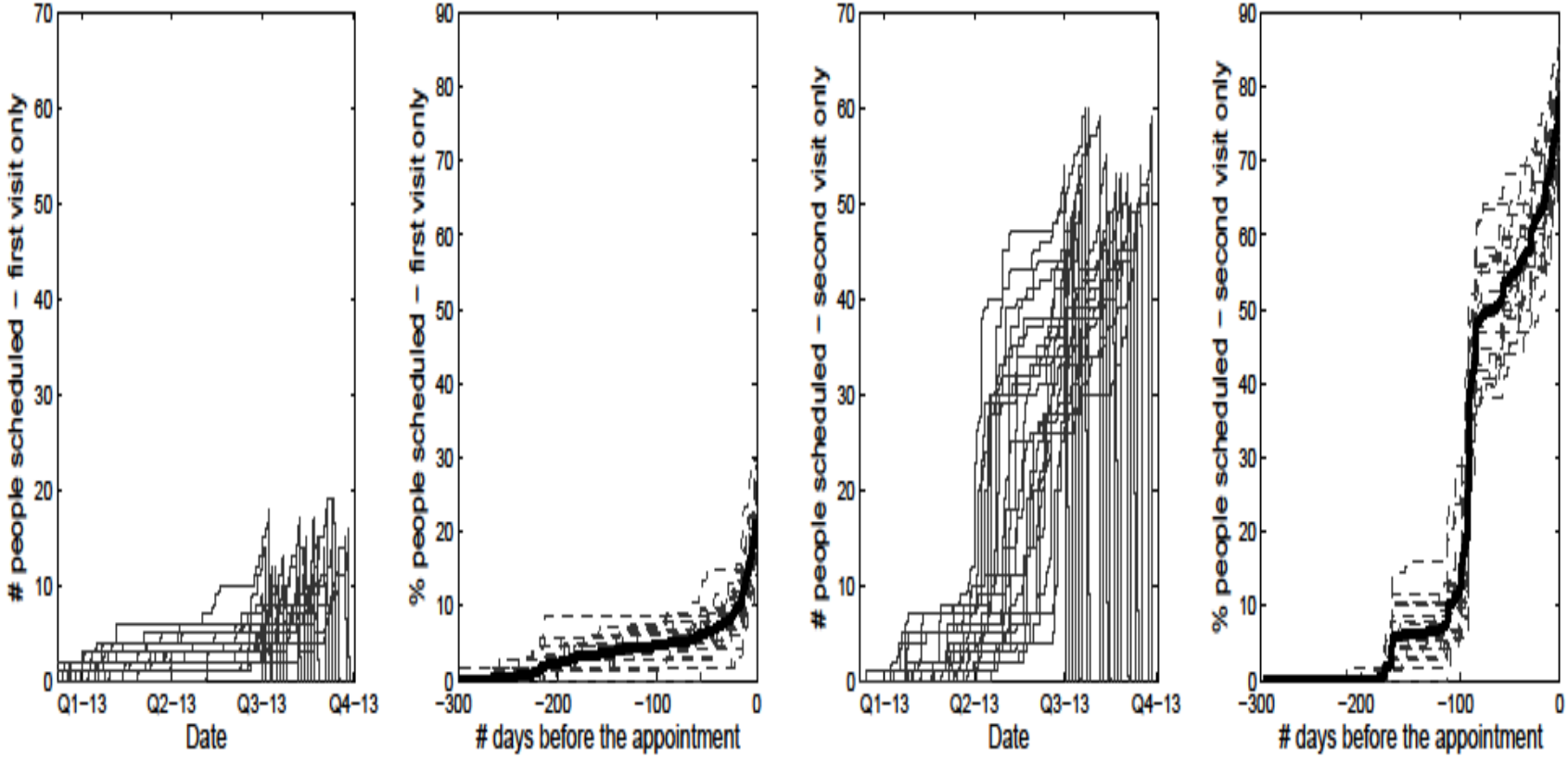
Evolution of Schedule: First Visits vs. Repeat Visits

First Visits

22%

Repeat Visits

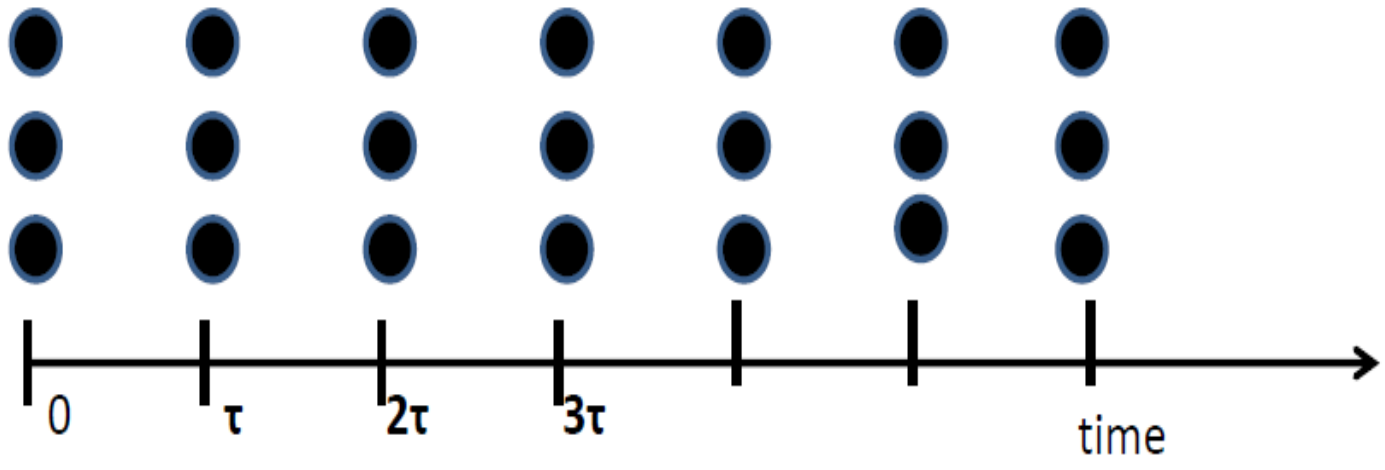
78%



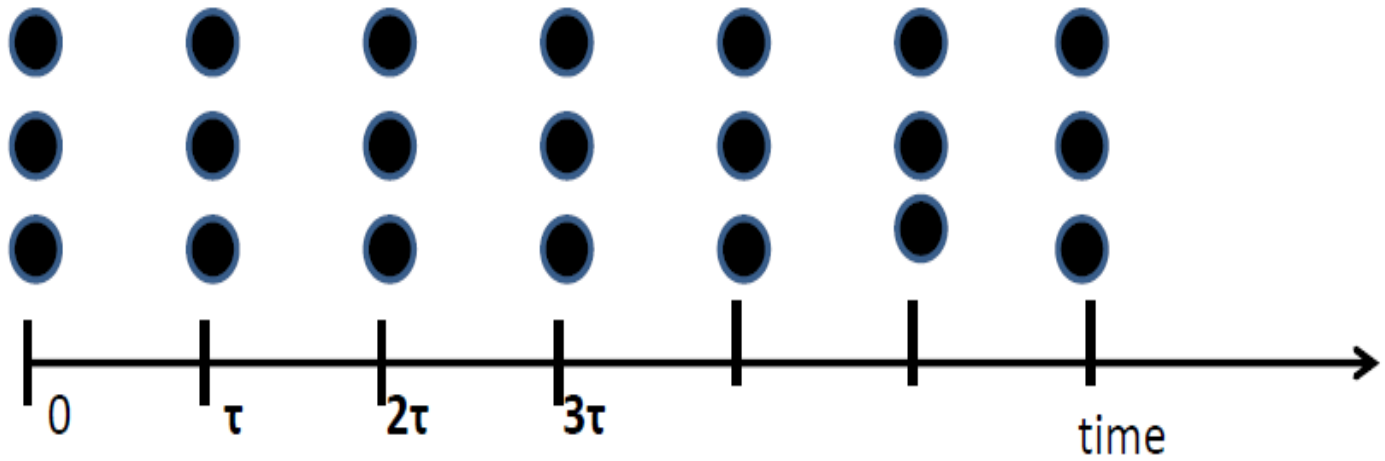
Schedule for Doctor 9 for 22 am Shifts

time slot	22 days in July-October 2013																						Avg	Var	Var/Avg
7:50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:00	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0	0	1	1	0.32	0.23	0.71
8:10	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0.05	1.00
8:20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:50	3	4	5	4	4	4	4	4	4	1	3	2	1	4	2	4	4	2	4	5	4	3	3.41	1.30	0.38
9:00	3	4	2	3	3	2	3	3	3	3	3	2	2	2	2	3	4	3	2	3	4	2	2.77	0.47	0.17
9:10	3	3	3	2	2	2	4	2	2	3	2	3	2	3	3	3	2	2	3	2	3	3	2.59	0.35	0.13
9:20	2	2	4	2	3	2	3	2	2	3	3	3	2	3	2	3	3	3	3	2	3	2	2.59	0.35	0.13
9:30	3	2	3	4	3	3	4	3	3	3	3	3	1	3	2	2	2	2	3	3	3	3	2.77	0.47	0.17
9:40	3	3	3	2	2	2	2	3	3	2	2	3	2	3	2	2	2	2	3	2	2	2	2.36	0.24	0.10
9:50	3	3	3	3	2	3	3	3	3	3	3	2	2	3	3	3	3	3	2	2	3	3	2.77	0.18	0.07
10:00	3	2	3	3	2	3	2	3	2	3	3	3	3	3	3	3	4	4	3	3	3	3	2.91	0.28	0.10
10:10	3	3	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3	3	3	3	3	2.91	0.09	0.03
10:20	2	3	3	3	3	3	3	2	3	3	2	3	3	3	3	2	3	2	3	4	3	3	2.82	0.25	0.09
10:30	3	2	3	3	3	2	4	2	3	2	3	3	3	3	3	2	3	3	2	4	3	3	2.82	0.35	0.12
10:40	3	1	3	3	3	1	3	2	3	2	3	3	2	3	2	1	3	2	3	3	3	2	2.45	0.55	0.22
10:50	2	3	3	3	1	2	3	2	3	3	3	2	3	3	3	3	3	3	2	3	3	3	2.68	0.32	0.12
11:00	3	2	3	2	3	2	3	2	2	4	4	4	2	3	3	3	3	3	3	4	3	4	2.95	0.52	0.18
11:10	3	3	3	1	3	3	3	3	2	3	3	2	3	2	1	3	2	3	3	3	3	3	2.64	0.43	0.16
11:20	2	3	3	3	3	3	3	3	3	3	3	3	3	2	2	3	3	3	3	3	3	4	2.91	0.18	0.06
11:30	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	3	2	2.77	0.18	0.07
11:40	3	2	3	3	2	3	3	3	3	1	2	3	3	2	3	3	3	3	3	3	2	3	2.68	0.32	0.12
11:50	3	3	3	3	3	2	2	3	3	2	3	2	4	3	3	3	2	2	3	3	1	3	2.68	0.42	0.16
12:00	2	3	3	2	3	3	4	3	3	2	3	3	3	3	3	3	3	3	2	2	3	4	2.86	0.31	0.11
12:10	3	3	3	2	3	3	2	3	2	3	3	2	3	3	4	3	1	2	3	2	3	3	2.68	0.42	0.16
12:20	2	4	3	2	3	3	3	3	4	3	3	3	3	2	2	3	1	3	1	4	3	3	2.77	0.66	0.24
12:30	2	1	0	0	0	3	3	3	3	2	2	2	2	3	3	3	2	4	3	1	2	3	2.14	1.27	0.59
12:40	0	0	0	0	0	2	2	4	3	0	3	2	1	2	3	3	4	2	3	0	0	3	1.68	2.13	1.27
12:50	0	0	0	0	0	0	0	1	4	0	0	0	0	3	4	0	2	0	4	0	0	4	1.00	2.67	2.67
13:00	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0.09	0.95
Daily Total	63	62	67	59	61	62	73	70	72	59	69	64	59	70	68	67	68	64	69	66	67	75	66.09	21.32	0.32
[8:50, 12:20] Total	60	61	67	59	60	57	67	60	61	57	63	60	56	62	57	61	60	58	59	65	64	64	60.82	9.77	0.16
All slot avg	2.0	2.0	2.2	1.9	2.0	2.0	2.4	2.3	2.3	1.9	2.2	2.1	1.9	2.3	2.2	2.2	2.2	2.1	2.2	2.1	2.2	2.4	2.07	1.73	0.84
All slot var	1.5	1.9	2.2	1.9	1.8	1.5	1.8	1.3	1.5	1.7	1.5	1.5	1.6	1.5	1.3	1.7	1.8	1.6	1.6	2.2	1.8	1.6	(across all days)		
All slot var/avg	0.7	1.0	1.0	1.0	0.9	0.8	0.8	0.6	0.6	0.9	0.7	0.7	0.8	0.6	0.6	0.8	0.8	0.8	0.7	1.1	0.8	0.7			
[8:50, 12:20] avg	2.7	2.8	3.0	2.7	2.7	2.6	3.0	2.7	2.8	2.6	2.9	2.7	2.5	2.8	2.6	2.8	2.7	2.6	2.7	3.0	2.9	2.9	2.76	0.42	0.15

Ideal Deterministic Framework



Ideal Deterministic Framework



- Roughly see between 8:50am and 12:20pm
- $\nu = 22$ batches of $\beta = 3$ patients arrive every $\tau = 10$ minutes
- daily total: $N = \nu\beta = 66$
- Total Time: $T = (\nu - 1)\tau = 210$ minutes (3.5 hours)

At-Capacity (AC) Days and Overloaded (OL) Days

- OL occurs before and after main time interval 8:50am-12:20pm
- A day with 5 or more scheduled arrivals after 12:20pm is said to be OL.
- In data there are 10 AC Days and 12 OL Days.
- Model: Status of days are IID Bernoulli with $P(OL) = 12/22$.

Schedule for Doctor 9 for 22 am Shifts

time slot	22 days in July-October 2013																						Avg	Var	Var/Avg
7:50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:00	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0	0	1	1	0.32	0.23	0.71
8:10	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0.05	1.00
8:20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00		
8:50	3	4	5	4	4	4	4	4	4	1	3	2	1	4	2	4	4	2	4	5	4	3	3.41	1.30	0.38
9:00	3	4	2	3	3	2	3	3	3	3	3	2	2	2	2	3	4	3	2	3	4	2	2.77	0.47	0.17
9:10	3	3	3	2	2	2	4	2	2	3	2	3	2	3	3	3	2	2	3	2	3	3	2.59	0.35	0.13
9:20	2	2	4	2	3	2	3	2	2	3	3	3	2	3	2	3	3	3	3	2	3	2	2.59	0.35	0.13
9:30	3	2	3	4	3	3	4	3	3	3	3	3	1	3	2	2	2	2	3	3	3	3	2.77	0.47	0.17
9:40	3	3	3	2	2	2	2	3	3	2	2	3	2	3	2	2	2	2	3	2	2	2	2.36	0.24	0.10
9:50	3	3	3	3	2	3	3	3	3	3	3	2	2	3	3	3	3	3	2	2	3	3	2.77	0.18	0.07
10:00	3	2	3	3	2	3	2	3	2	3	3	3	3	3	3	3	4	4	3	3	3	3	2.91	0.28	0.10
10:10	3	3	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3	3	3	3	3	2.91	0.09	0.03
10:20	2	3	3	3	3	3	3	2	3	3	2	3	3	3	3	2	3	2	3	4	3	3	2.82	0.25	0.09
10:30	3	2	3	3	3	2	4	2	3	2	3	3	3	3	3	2	3	3	2	4	3	3	2.82	0.35	0.12
10:40	3	1	3	3	3	1	3	2	3	2	3	3	2	3	2	1	3	2	3	3	3	2	2.45	0.55	0.22
10:50	2	3	3	3	1	2	3	2	3	3	3	2	3	3	3	3	3	3	2	3	3	3	2.68	0.32	0.12
11:00	3	2	3	2	3	2	3	2	2	4	4	4	2	3	3	3	3	3	3	4	3	4	2.95	0.52	0.18
11:10	3	3	3	1	3	3	3	3	2	3	3	2	3	2	1	3	2	3	3	3	3	3	2.64	0.43	0.16
11:20	2	3	3	3	3	3	3	3	3	3	3	3	3	2	2	3	3	3	3	3	3	4	2.91	0.18	0.06
11:30	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	3	2	2.77	0.18	0.07
11:40	3	2	3	3	2	3	3	3	3	1	2	3	3	2	3	3	3	3	3	3	2	3	2.68	0.32	0.12
11:50	3	3	3	3	3	2	2	3	3	2	3	2	4	3	3	3	2	2	3	3	1	3	2.68	0.42	0.16
12:00	2	3	3	2	3	3	4	3	3	2	3	3	3	3	3	3	3	3	2	2	3	4	2.86	0.31	0.11
12:10	3	3	3	2	3	3	2	3	2	3	3	2	3	3	4	3	1	2	3	2	3	3	2.68	0.42	0.16
12:20	2	4	3	2	3	3	3	3	4	3	3	3	3	2	2	3	1	3	1	4	3	3	2.77	0.66	0.24
12:30	2	1	0	0	0	3	3	3	3	2	2	2	2	3	3	3	2	4	3	1	2	3	2.14	1.27	0.59
12:40	0	0	0	0	0	2	2	4	3	0	3	2	1	2	3	3	4	2	3	0	0	3	1.68	2.13	1.27
12:50	0	0	0	0	0	0	0	1	4	0	0	0	0	3	4	0	2	0	4	0	0	4	1.00	2.67	2.67
13:00	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0.09	0.95
Daily Total	63	62	67	59	61	62	73	70	72	59	69	64	59	70	68	67	68	64	69	66	67	75	66.09	21.32	0.32
[8:50, 12:20] Total	60	61	67	59	60	57	67	60	61	57	63	60	56	62	57	61	60	58	59	65	64	64	60.82	9.77	0.16
All slot avg	2.0	2.0	2.2	1.9	2.0	2.0	2.4	2.3	2.3	1.9	2.2	2.1	1.9	2.3	2.2	2.2	2.2	2.1	2.2	2.1	2.2	2.4	2.07	1.73	0.84
All slot var	1.5	1.9	2.2	1.9	1.8	1.5	1.8	1.3	1.5	1.7	1.5	1.5	1.6	1.5	1.3	1.7	1.8	1.6	1.6	2.2	1.8	1.6	(across all days)		
All slot var/avg	0.7	1.0	1.0	1.0	0.9	0.8	0.8	0.6	0.6	0.9	0.7	0.7	0.8	0.6	0.6	0.8	0.8	0.8	0.7	1.1	0.8	0.7			
[8:50, 12:20] avg	2.7	2.8	3.0	2.7	2.7	2.6	3.0	2.7	2.8	2.6	2.9	2.7	2.5	2.8	2.6	2.8	2.7	2.6	2.7	3.0	2.9	2.9	2.76	0.42	0.15

Scheduled Batch Sizes in Time Slots

- $B_{s,j}$ is number of scheduled arrivals in time slot j
- During the main time period 8:50am-12:20pm, IID random variables (independence tested)
- with the estimated batch-size distribution:

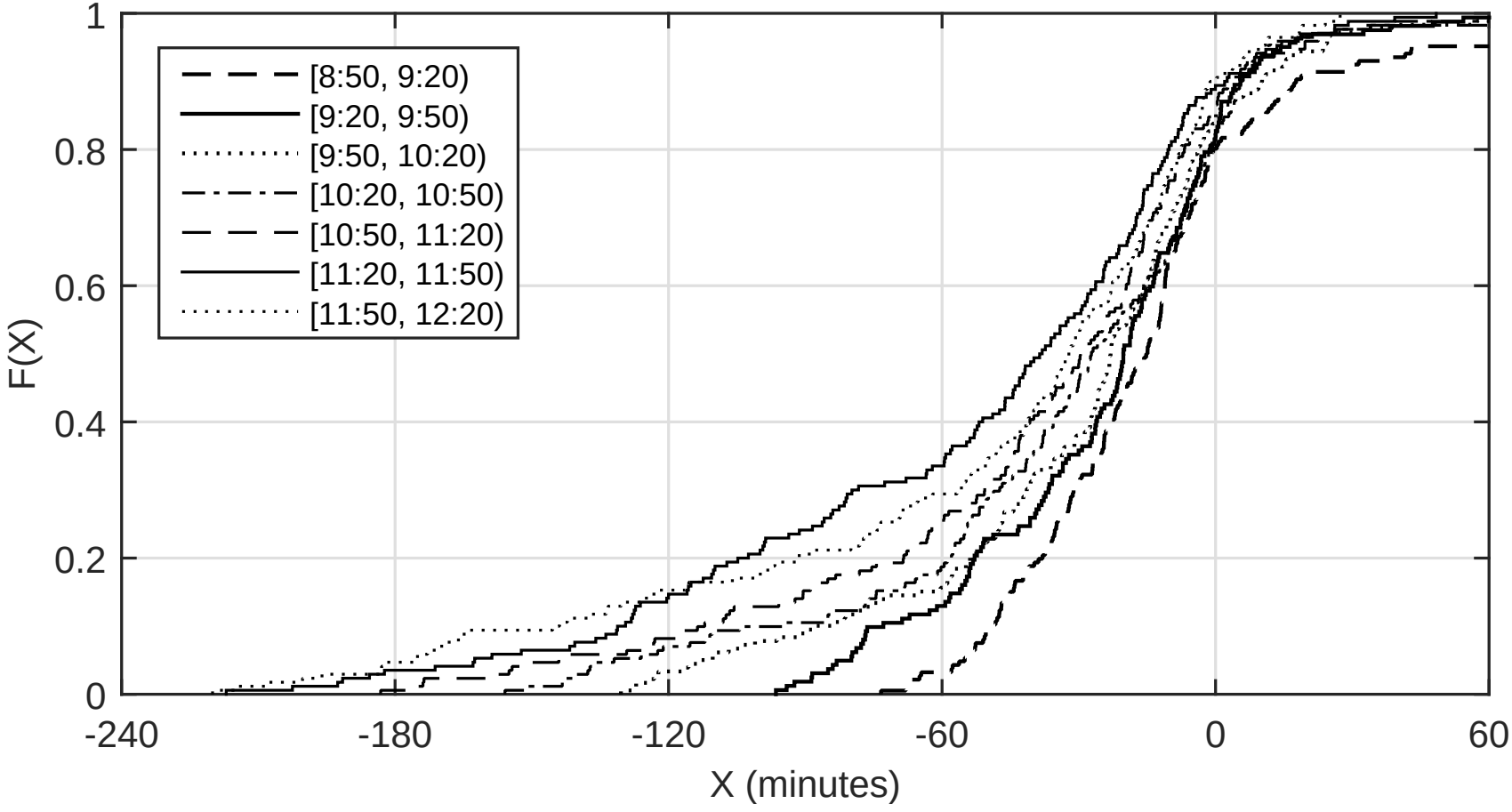
	$\hat{P}(B_{s,j} = k)$				
number k	1	2	3	4	5
10 at-capacity days	0.04	0.25	0.63	0.07	0.01
12 overloaded days	0.02	0.27	0.63	0.08	
All days	0.03	0.26	0.63	0.08	0.004

No-Shows and Unscheduled Arrivals

- Both are relatively rare.
- Model: No-shows IID Bernoullis, i.e., each scheduled arrival is a no show with $P(NoShow) = 0.08$
- Unscheduled arrivals defined as actual arrivals not in schedule. (The schedule is defined at the end of the previous day.)
- On average, 2.18 unscheduled arrivals, of which 1.95 arrive.
- Unscheduled modelled as extra low-rate Poisson process.

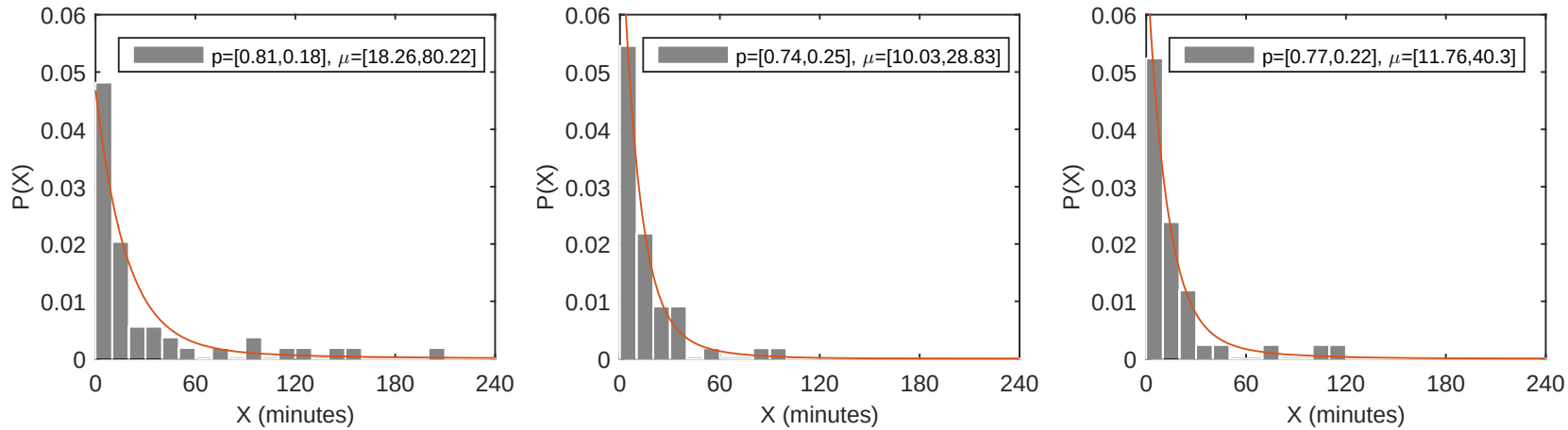
Pattern of Arrivals Over the Day

Lateness Empirical CDF's in 7 Half Hours: stochastically ordered!



Lateness Histograms and Hyperexponential Fits

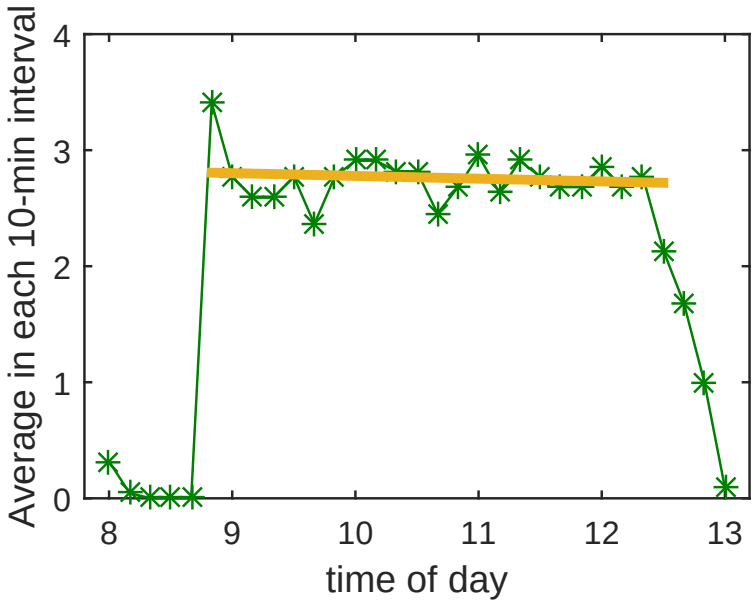
- fit $P(X > 0)$, $P(X \geq -x | X < 0)$ and $P(X \leq x | X > 0)$ to data
- Parametric models for conditional lateness cdf's



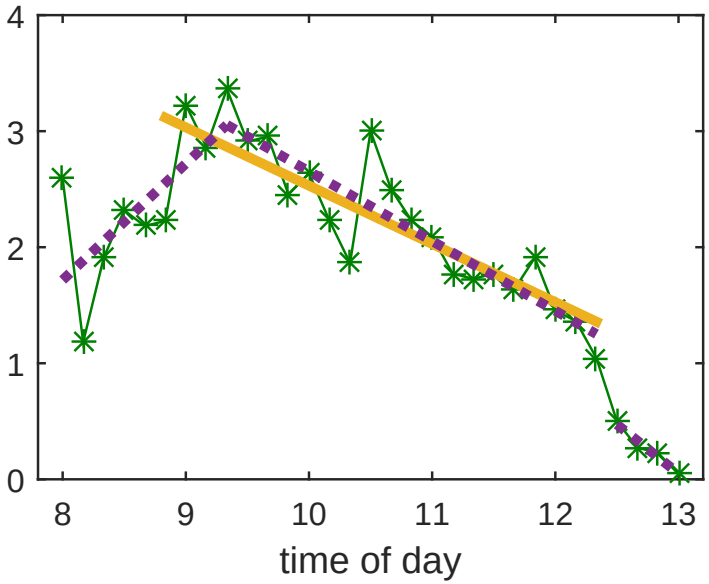
Average Number of Arrivals Per Time Slot

- Excess earliness affects the overall arrival rate.

Scheduled



Actual



Summary: Data Analysis and Model Construction

- Randomness in the Schedule
  - At-capacity days versus overloaded days (with extra scheduled outside main time interval)
  - Batch size distribution in the main interval (same for AC days and OL days)
  - Batch size distributions outside the main interval
  - Dispersion of Daily Total (variance/mean) = 0.3
    - Dispersion of daily totals same as for actual arrivals.
- Adherence to the Schedule
  - No-shows are low-probability events.
  - Extra unscheduled arrivals are rare.
  - Significant deviations due to lateness and, mostly, earliness.
    - Altered arrival rate over the day.

Thank you!!

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:
    - NHPP fits within day, but **overdispersion** over days.

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:
    - NHPP fits within day, but **overdispersion** over days.
  - For clinic with arrivals by appointments:

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:
    - NHPP fits within day, but **overdispersion** over days.
  - For clinic with arrivals by appointments:
    - NHPP fits within day, but **underdispersion** over days.

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:
    - NHPP fits within day, but **overdispersion** over days.
  - For clinic with arrivals by appointments:
    - NHPP fits within day, but **underdispersion** over days.
  - From more careful analysis of clinic now:

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:
    - NHPP fits within day, but **overdispersion** over days.
  - For clinic with arrivals by appointments:
    - NHPP fits within day, but **underdispersion** over days.
  - From more careful analysis of clinic now:
    - Two kinds of days: at-capacity days and overloaded days

Research Project for Service Systems: What Arrival Process Model is Appropriate?

- Candidate: Nonhomogeneous Poisson process (NHPP)
- Perform statistical tests of NHPP with arrival data.
- Previous Answers for Systems Studied:
  - For call centers and hospital emergency departments:
    - NHPP fits within day, but **overdispersion** over days.
  - For clinic with arrivals by appointments:
    - NHPP fits within day, but **underdispersion** over days.
  - From more careful analysis of clinic now:
    - Two kinds of days: at-capacity days and overloaded days
    - Lateness and earliness affect the arrival rate.

References: Previous and Current Talk

- Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. 2005. Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective. *Journal of the American Statistical Association* 100:3650.
- Kim, S.-H. and Whitt, W. 2014. Choosing Arrival Process Models for Service Systems: Tests of a Nonhomogeneous Poisson Process, *Naval Research Logistics* 61(1):66-90.
- Kim, S.-H. and Whitt, W. 2014. Are Call Center and Hospital Arrivals Well Modeled by Nonhomogeneous Poisson Processes?, *Manufacturing and Service Operations Management* 16(3):464-480.
- Kim, S.-H., P. Vel, W. Whitt and W. C. Cha. 2015. Poisson and non-Poisson properties in appointment-generated arrival processes: the case of an endocrinology clinic, *OR Letters* 43:247-251.
- Kim, S.-H., W. Whitt and W. C. Cha. 2015. A data-driven model of an appointment-generated arrival processes at an endocrinology clinic. working paper. (detailed stochastic model)