Evaluation of appointment scheduling rules
A multi-performance measures approach

Stefan Creemers
IESEG, KU Leuven
France, Belgium

Pieter Colen
KU Leuven
Belgium

Marc Lambrecht
KU Leuven
Belgium
Appointments are part of life

- Car repair
- Accountants
- Dentist
- Barber
- Oral exams
- Internet/cable connecting
- Central heating cleaning
- ...

...
Appointments are frustrating

- Cancellations and no-shows
- Walk-ins and emergencies
- Variability in treatment and arrival times
- Capacity availability
  - Equipement
  - Server

Often service providers respond with overbookings, large appointment timeslots, lateness and no-show fees

(Guapta & Denton, 2008)
Service comes to health care

Good service = fulfilling and surpassing expectations

“Patients’ Bill of Rights” series on ABC World News
Trend to start invoicing doctors

Delay:
5 $: Dr. Timothy Malia of Fairport, New York,
Starbucks cards: Dr. Gwen Hanson of Bellevue, Washington
Movie tickets: Dr. Sharon McCoy George of Irvine, California
Cancellations:
50 $: Dr. Cyrus Peikari of Dallas, Texas
Smaller number of patients (boutique, concierge practices)

Physician: http://www.kevinmd.com/blog/2012/01/doctors-give-cash-patients-running-late.html

Importance of balanced appointment scheduling rules is on the rise

Instead of developing a complex ASR selection tool we want to know which ASR performs well across settings
Appointment scheduling rules research

- **Walk-in systems**
- Open-access (Advanced/online) systems (Murray and Tantau (1999,2000))
  - Customers/patients are scheduled directly when they contact the facility
  - The number of customers/patients to attend is unknown beforehand
  - Capacity chases demand

- **Traditional (offline) system**
- Server/planner has a priori knowledge about the number of patients to attend and possibly their characteristics (referral systems)
  - Demand to suit capacity

(Robinson & Chen 2010, Liu et al. 2010)
Types of ASR: Finding an ASR that works well in a broad range of settings

**Individual ASR**
(Bailey, 1952; Rohleder & Klassen, 2000)
- Number of patients to arrive at start of session
- Delay of first patients to arrive
- Adjustment for service variance

**Block ASR**
(Liu & Liu, 1998; Vanden Bosch & Dietz, 2002)
- Block size
- Adjustment for service variance

**Variable interval (Dome, Early-lateness) ASR**
(Denton & Gupta, 2003; Wang, 1993)
- The pivot patient
- Speed of increasing or decreasing the arrival pace
- Adjustment for service variance

We identified and tested not less than 314 different ASR
Finding an ASR that works well in a broad range of settings

- Selecting the best ASR:
  (e.g. Ho & Lau, 1992, 1999; Klassen & Yoogalingam, 2009)
  - Customer (direct) waiting time
  - Server idle time
  - Service overtime

Cost function: with cost/penalty parameters

Find best/most robust ASR without subjectively imposing a weighting scheme

- Fast and accurate method to obtain performance measures for different ASR
- Data envelopment analysis to identify those ASR that outperform taking into account multiple performance measures
Fast and accurate method to obtain performance measures for different ASR

**Setting:** Schedule N customers in 1 service session; single server
- Customer unpunctuality: \( P(\text{too late}) \), \( P(\text{too early}) \), \( P(\text{no-show}) \), earliness-lateness distributions
- Service session delay
- Stochastic service time and interruptions

**Technique:** Discrete Time Markov Chain
- State-space
  - Time instance when system observed
  - The number of customers in queue
  - Set of customers eligible to arrive but still not arrived
  - State of service system: idle, processing or completion
Fast and accurate method to obtain performance measures for different ASR

- **Technique**: Discrete Time Markov Chain Transitions
  
  Events: eligible to arrive, arrival, no-show, service completion

![Diagram showing cumulative service time distribution with probabilities](chart)

Probabilities: $P(\text{state})$

Performance measures:
- Idle time ($I$)
- Waiting time ($W$)
- Overtime ($O$)
Identify those ASR that outperform taking into account multiple performance measures

- **Setting:** 314 ASR with performance measures $I$, $W$, $O$ for 243 operating environments:
  - $N$, SCV(service), SCV(early), SCV(late), $P$(early), $P$(late), $P$(no-show)

- **Technique:** Data envelopment analysis
  - Composite indicator
    $$Cl_r = v_{or} O_r + v_{ir} I_r + v_{wr} W_r$$
  - Setting weight objectively:
    - input oriented DEA without output \( (Cherchye et al., 2008) \)
    - Avoid zero weights:
      - constrained facet analysis (CFA) \( (Olesen & Petersen, 1996) \)

For each ASR we obtain one score which aggregates multiple-performance measures across different environments
Which type of ASR performs strongly?

Individual ASR outperform the other types of ASR
# Top 15 performing ASR

<table>
<thead>
<tr>
<th>Rank</th>
<th>ASR</th>
<th>CI (%)</th>
<th>Weight Sensitivity</th>
<th>Type of ASR</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>99,998</td>
<td>0,195624</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=3 adjustment time var=0</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>99,9976</td>
<td>0,313921</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=4 adjustment time var=0</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>99,9964</td>
<td>0,389075</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=5 adjustment time var=0,05</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>99,9954</td>
<td>0,284273</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=4 adjustment time var=0,05</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>99,9941</td>
<td>0,412453</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=5 adjustment time var=0</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>99,9838</td>
<td>0,163625</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=3 adjustment time var=0,05</td>
</tr>
<tr>
<td>7</td>
<td>228</td>
<td>99,9724</td>
<td>0,066703</td>
<td>Var Interval</td>
<td>pivot patient=15 rate interval increase=2 decrease=1 adjustment time var=0,05</td>
</tr>
<tr>
<td>8</td>
<td>138</td>
<td>99,965</td>
<td>0,171877</td>
<td>Var Interval</td>
<td>pivot patient=5 rate interval increase=1 decrease=1 adjustment time var=0,2</td>
</tr>
<tr>
<td>9</td>
<td>214</td>
<td>99,9453</td>
<td>0,070306</td>
<td>Var Interval</td>
<td>pivot patient=15 rate interval increase=1 decrease=1 adjustment time var=0,1</td>
</tr>
<tr>
<td>10</td>
<td>247</td>
<td>99,9184</td>
<td>0,064002</td>
<td>Var Interval</td>
<td>pivot patient=20 rate interval increase=1 decrease=0 adjustment time var=0,1</td>
</tr>
<tr>
<td>11</td>
<td>31</td>
<td>99,906</td>
<td>0,364031</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=4 adjustment time var=0,1</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>99,8875</td>
<td>0,340854</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=1 adjustment time var=0,3</td>
</tr>
<tr>
<td>13</td>
<td>267</td>
<td>99,8871</td>
<td>0,063741</td>
<td>Var Interval</td>
<td>pivot patient=20 rate interval increase=2 decrease=1 adjustment time var=0,05</td>
</tr>
<tr>
<td>14</td>
<td>9</td>
<td>99,8838</td>
<td>0,076893</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=2 adjustment time var=0,05</td>
</tr>
<tr>
<td>15</td>
<td>176</td>
<td>99,872</td>
<td>0,095337</td>
<td>Var Interval</td>
<td>pivot patient=10 rate interval increase=1 decrease=1 adjustment time var=0,15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>8</td>
<td>99,8136</td>
<td>0,09592</td>
<td>Indiv</td>
<td>first arrival delay=2 Nr of initial patients=2 adjustment time var=0</td>
</tr>
</tbody>
</table>

**Best 7 individual ASR:** no delay in first arrival

**3 or more patient at start of session**

Adjustment for service variance should be zero or very small
Service comes to health care

Limiting customer waiting time is often crucial to satisfy customers

We adapt the DEA model such that customer waiting time has the highest weight in the performance evaluation

\[
\frac{v_w}{v_i} \geq 1 \quad \frac{v_w}{v_o} \geq 1
\]
If customer waiting time has highest importance

<table>
<thead>
<tr>
<th>Rank</th>
<th>ASR</th>
<th>CI (%)</th>
<th>Weight Sensitivity</th>
<th>Type of ASR</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>228</td>
<td>99,97242</td>
<td>0,050776</td>
<td>Var Interval</td>
<td>pivot patient=15 rate interval increase=2 decrease=1 adjustment time var=0,15</td>
</tr>
<tr>
<td>2</td>
<td>138</td>
<td>99,96501</td>
<td>0,114976</td>
<td>Var Interval</td>
<td>pivot patient=5 rate interval increase=1 decrease=1 adjustment time var=0,2</td>
</tr>
<tr>
<td>3</td>
<td>214</td>
<td>99,94535</td>
<td>0,053632</td>
<td>Var Interval</td>
<td>pivot patient=15 rate interval increase=1 decrease=1 adjustment time var=0,1</td>
</tr>
<tr>
<td>4</td>
<td>247</td>
<td>99,91838</td>
<td>0,054437</td>
<td>Var Interval</td>
<td>pivot patient=20 rate interval increase=1 decrease=0 adjustment time var=0,1</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>99,88746</td>
<td>0,273575</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=1 adjustment time var=0,3</td>
</tr>
<tr>
<td>6</td>
<td>267</td>
<td>99,88713</td>
<td>0,053451</td>
<td>Var Interval</td>
<td>pivot patient=20 rate interval increase=2 decrease=1 adjustment time var=0,05</td>
</tr>
<tr>
<td>7</td>
<td>176</td>
<td>99,87204</td>
<td>0,073426</td>
<td>Var Interval</td>
<td>pivot patient=10 rate interval increase=1 decrease=1 adjustment time var=0,15</td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>99,85562</td>
<td>0,192716</td>
<td>Indiv</td>
<td>first arrival delay=0,5 Nr of initial patient=2 adjustment time var=0,3</td>
</tr>
<tr>
<td>9</td>
<td>175</td>
<td>99,84572</td>
<td>0,06281</td>
<td>VariInterval</td>
<td>pivot patient=10 rate interval increase=1 decrease=1 adjustment time var=0,1</td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>99,8433</td>
<td>0,150799</td>
<td>Indiv</td>
<td>first arrival delay=0,5 Nr of initial patient=2 adjustment time var=0,25</td>
</tr>
<tr>
<td>11</td>
<td>139</td>
<td>99,83446</td>
<td>0,138516</td>
<td>Var Interval</td>
<td>pivot patient=5 rate interval increase=1 decrease=1 adjustment time var=0,25</td>
</tr>
<tr>
<td>12</td>
<td>68</td>
<td>99,83392</td>
<td>0,114452</td>
<td>Indiv</td>
<td>first arrival delay=0,5 Nr of initial patient=2 adjustment time var=0,2</td>
</tr>
<tr>
<td>13</td>
<td>208</td>
<td>99,82541</td>
<td>0,051386</td>
<td>Var Interval</td>
<td>pivot patient=15 rate interval increase=1 decrease=0 adjustment time var=0,1</td>
</tr>
<tr>
<td>14</td>
<td>152</td>
<td>99,7837</td>
<td>0,087551</td>
<td>Var Interval</td>
<td>pivot patient=5 rate interval increase=2 decrease=1 adjustment time var=0,15</td>
</tr>
<tr>
<td>15</td>
<td>137</td>
<td>99,78285</td>
<td>0,094437</td>
<td>Var Interval</td>
<td>pivot patient=5 rate interval increase=1 decrease=1 adjustment time var=0,15</td>
</tr>
</tbody>
</table>

Variable interval ASR take over as best performing ASR

Top four Var interval ASR: perform very strong across different weighting schemes
Which environmental factor influence the performance the most?

Environmental effects studied
- $P$(late); $P$(early)
- $SCV$(late); $SCV$(early)
- $N$
- $P$(no-show)
- $SCV$(service time)

We identify no-shows as the environment factor with the most detrimental effect on performance
Which environmental factor influence the performance the most?

**Environmental effects studied**

- $P$(late); $P$(early)
- $SCV$(late); $SCV$(early)
- $N$
- $P$(no-show)
- $SCV$(service time)

We identify no-shows as the environment factor with the most detrimental effect on performance.
No-shows are not only detrimental but also prevalent in healthcare

“No-show rates at medical centers can vary from as little as 3% to as much as 80% depending on the type of center and demographic information of the patients of the medical Center” (Alaeddini et al., 2011)

Quoted average values vary e.g. 42% (Lacy et al., 2004), 10-25% (Pinedo, 2012)

US national level study revealed that one third of all family practice residencies had over 21% no-shows of appointments (Hixon et al. 1999)

Adapted ASR to deal with no-shows (Cayirli et al., 2012)

Which (standard) ASR perform strongly under high amount of no-shows?
**Under high no-shows**

<table>
<thead>
<tr>
<th>Rank</th>
<th>ASR</th>
<th>CI (%)</th>
<th>Weight Sensitivity</th>
<th>Type of ASR</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
<td>99.9998</td>
<td>0.3921</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=5 adjustment time var=0</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>99.9996</td>
<td>0.2892</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=4 adjustment time var=0</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>99.9994</td>
<td>0.1732</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=3 adjustment time var=0</td>
</tr>
<tr>
<td>4</td>
<td>285</td>
<td>99.9978</td>
<td>0.0518</td>
<td>Var Interval</td>
<td>pivot patient =25 rate interval increase=1 decrease=0 adjustment time var=0,05</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>99.9971</td>
<td>0.3687</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=5 adjustment time var=0,05</td>
</tr>
<tr>
<td>6</td>
<td>175</td>
<td>99.9966</td>
<td>0.0700</td>
<td>Var Interval</td>
<td>pivot patient=10 rate interval increase=1 decrease=1 adjustment time var=0,1</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>99.9891</td>
<td>0.3217</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=1 adjustment time var=0,3</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>99.9862</td>
<td>0.2610</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=4 adjustment time var=0,05</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>99.9709</td>
<td>0.0839</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=2 adjustment time var=0</td>
</tr>
<tr>
<td>10</td>
<td>291</td>
<td>99.9696</td>
<td>0.0520</td>
<td>Var Interval</td>
<td>pivot patient =25 rate interval increase=1 decrease=1 adjustment time var=0,05</td>
</tr>
<tr>
<td>11</td>
<td>189</td>
<td>99.9644</td>
<td>0.0567</td>
<td>Var Interval</td>
<td>pivot patient=10 rate interval increase=2 decrease=1 adjustment time var=0,05</td>
</tr>
<tr>
<td>12</td>
<td>138</td>
<td>99.9601</td>
<td>0.1645</td>
<td>Var Interval</td>
<td>pivot patient =5 rate interval increase=1 decrease=1 adjustment time var=0,2</td>
</tr>
<tr>
<td>13</td>
<td>151</td>
<td>99.9578</td>
<td>0.1154</td>
<td>Var Interval</td>
<td>pivot patient =5 rate interval increase=2 decrease=1 adjustment time var=0,1</td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>99.9559</td>
<td>0.1447</td>
<td>Indiv</td>
<td>first arrival delay=0 Nr of initial patient=3 adjustment time var=0,05</td>
</tr>
<tr>
<td>15</td>
<td>228</td>
<td>99.9487</td>
<td>0.0564</td>
<td>Var Interval</td>
<td>pivot patient =15 rate interval increase=2 decrease=1 adjustment time var=0,05</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>214</td>
<td>99.8833</td>
<td>0.0609</td>
<td>Var Interval</td>
<td>pivot patient =15 rate interval increase=1 decrease=1 adjustment time var=0,1</td>
</tr>
<tr>
<td>31</td>
<td>247</td>
<td>99.8066</td>
<td>0.0543</td>
<td>Var Interval</td>
<td>pivot patient =20 rate interval increase=1 decrease=0 adjustment time var=0,1</td>
</tr>
</tbody>
</table>

Bailey-Welch rules perform strong

Also the robust Var interval ASR perform strongly but are beaten by more simple rules
Findings

We develop a fast and accurate analytical model based on DTMC and DEA to assess the performance of ASR.

314 ASR were compared across 243 environmental setting based on three performance measures (overtime, idle time and waiting time).

- **Individual ASR** perform strongly especially w.r.t. overtime and idle time: thanks to arrival of buffer patients at the start of the session.
- **Var Interval ASR** outperform if customer waiting time is the most critical performance criteria: thanks to the adapted arrival pace.
- **Block ASR** perform badly and should be avoided.
- No-shows have the most severe impact on performance of the studied environmental parameters.
Thank you
Example of service session

3 customers to schedule with 6h session
Different type of ASR

\[ A_i = ia\mu^{-1}, \forall i < l \]
\[ A_i = A_{i-1} + \mu^{-1} + h\sigma_i, \forall i \geq l \]

\[ A_i = 0, \forall i < b \]
\[ A_{nb} = A_{(n-1)b} + b\mu^{-1} + h\sqrt{b\sigma_n}, \forall n: 1 \leq n < \frac{N}{b} \]
\[ A_{nb+i} = A_{nb}, \forall i: 1 \leq i < b \]

1. \[ A_0 = 0, A_i = A_{i-1} + \mu^{-1}, \forall i < N \]
2. \[ A_i = A_i - r_1(z - i)h\sigma_i, \forall i: 1 \leq 1 \leq z \]
\[ A_i = A_i - r_2(z - i)h\sigma_i, \forall i: z < 1 < N \]
Which environmental factor influence the performance the most?
Which environmental factor influence the performance the most?