



Home Office

A Mathematical Perspective from the Home Office

Professor Bernard Silverman FRS

Chief Scientific Adviser, Home Office
2015

Professor Bernard Silverman FRS AcSS

Chief Scientific Adviser

Leading scientist appointed (in line with general Government policy on scientific advice) to provide the Home Secretary and ministers with impartial and expert advice.

Head of Home Office Science which has approximately 488 staff.

Responsible for science and research in the Home Office and its agencies.

Member of the cross-government **Chief Scientific Adviser** network.

Responsible for **fostering links to industry and academia**, and internationally.

Facilitates the **science advisory committees**, for example, Drugs, Animals in Science, and DNA Ethics.

Brings own broad expertise to individual projects and topics.

What areas are Home Office Science involved in?



Some recent topics

- **Protection of Freedoms Act:** If someone is arrested but not charged, how long should their DNA profile be retained?
- **The Olympics and subsequently:** Allocating resources to queues at the border
- **Border security and contraband detection:** High energy, multimodal vehicle/freight screening
- **Forensic Science and the Criminal Justice System:** Dealing with different chemistries when moving to 17-point DNA profiles
- **Modern Slavery Strategy:** The scale of Modern Slavery in the UK: measuring a hidden population



Home Office

Protection of Freedoms Act

If someone is arrested but not charged, for how long should their DNA profile be retained?

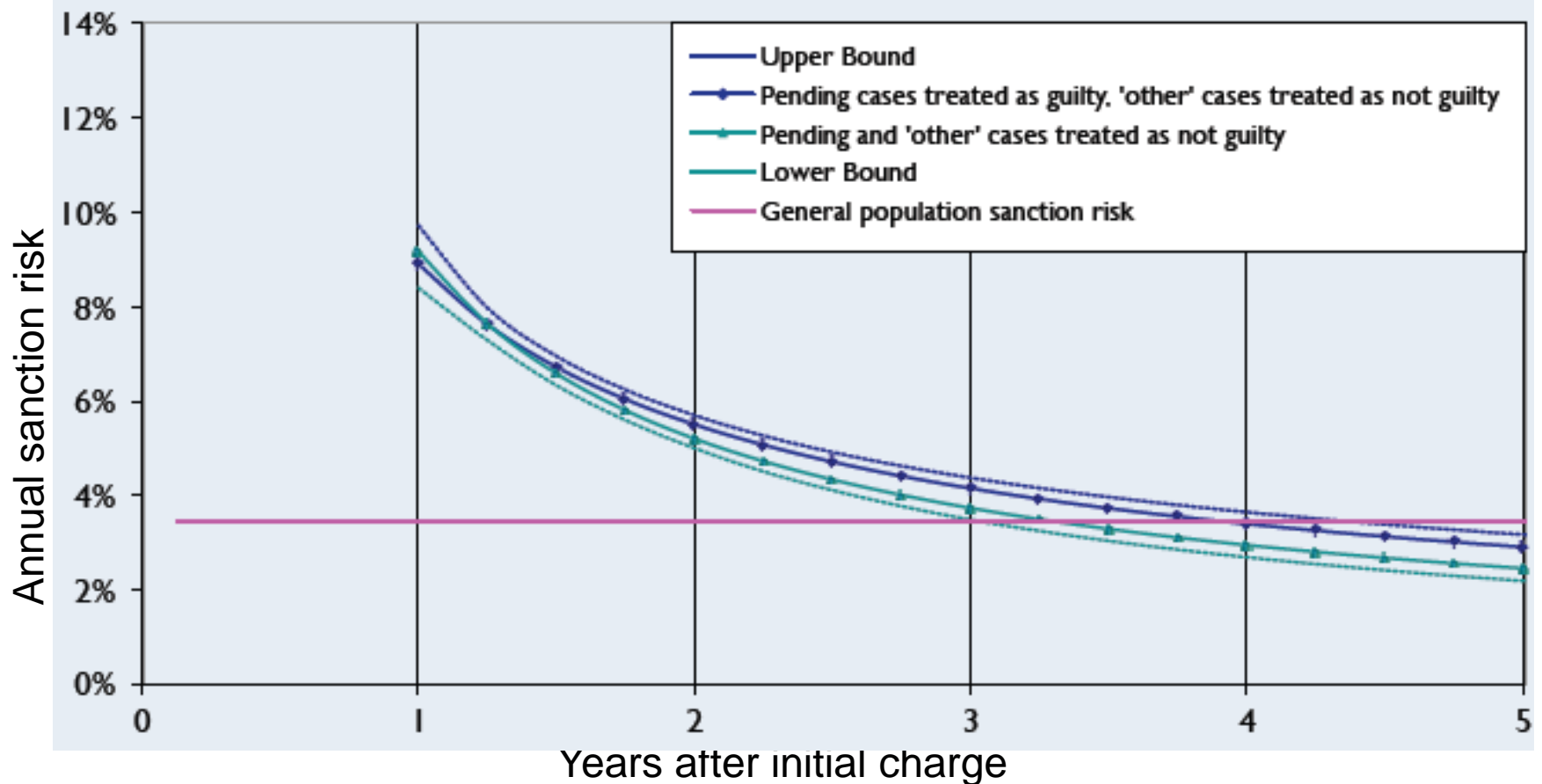
DNA – some background statistics

As at 31st December 2013...

- National DNA database held 4.47m subject samples.
- DNA is important in the detection of both
 - high volume crimes (e.g. burglary) and
 - low volume serious crimes (homicide and rape).
- More than 450,000 crime scenes matched to subjects held on the Database, of which
 - 2,874 murder, manslaughter and attempted murder crime scenes and
 - 6,526 rapes.

Hazard rate curves: offenders with no prior sanction, arrested but not charged with a serious/violence offence

Arrest-to-sanction hazard rates and general population sanction risk for 'CSA+' offence list.



Key findings

- For offenders charged with a serious violent or sexual offence, but not proceeded against, the earliest that offending risk falls to that of the general population is just over **three years**. This is now the retention period specified in law.
- The hierarchy of future offending risk is, in descending order:
 1. Those convicted of a serious violent or sexual offence
 2. Those charged with a violent/sexual offence but not proceeded against
 3. Those charged with other recordable offences but not proceeded against
 4. Those arrested but not charged
 5. The general population
- Key to interpreting the analysis is how the risk of future offending is measured in the general population.



Home Office

Queueing at the border

How should resources be allocated?

The Algorithm

What does the algorithm do?

- It takes a forecast for workload arrival and proposes the number of desks required in 15 minute time slots.
- There are some significant changes to the current approach:
 - tries to minimise queues as well as avoid breaches
 - allows an explicit trade-off between wasted passenger time and wasted staff time

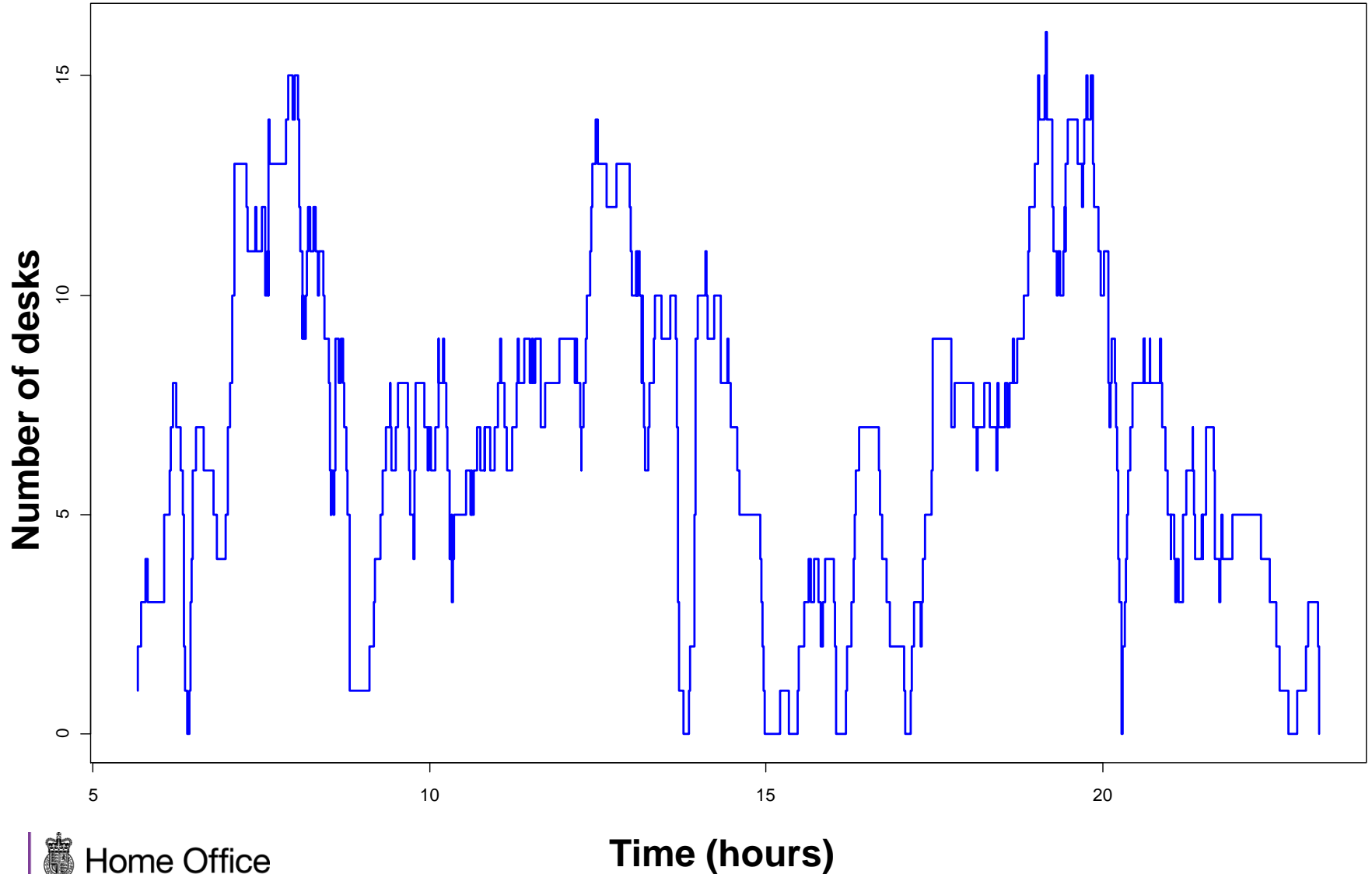
Where can time get wasted in the system?

1. **Passenger** time is wasted whenever they have to queue.
2. **Staff time** is ‘wasted’ whenever they are sitting at a desk when the queue is empty.
3. **Staff time** is also wasted by “**churn**”, i.e. if an officer is repeatedly stood down and then called back to duty as demand fluctuates through the day.

The algorithm can be adjusted to give different weights to the 3 sources of lost time.

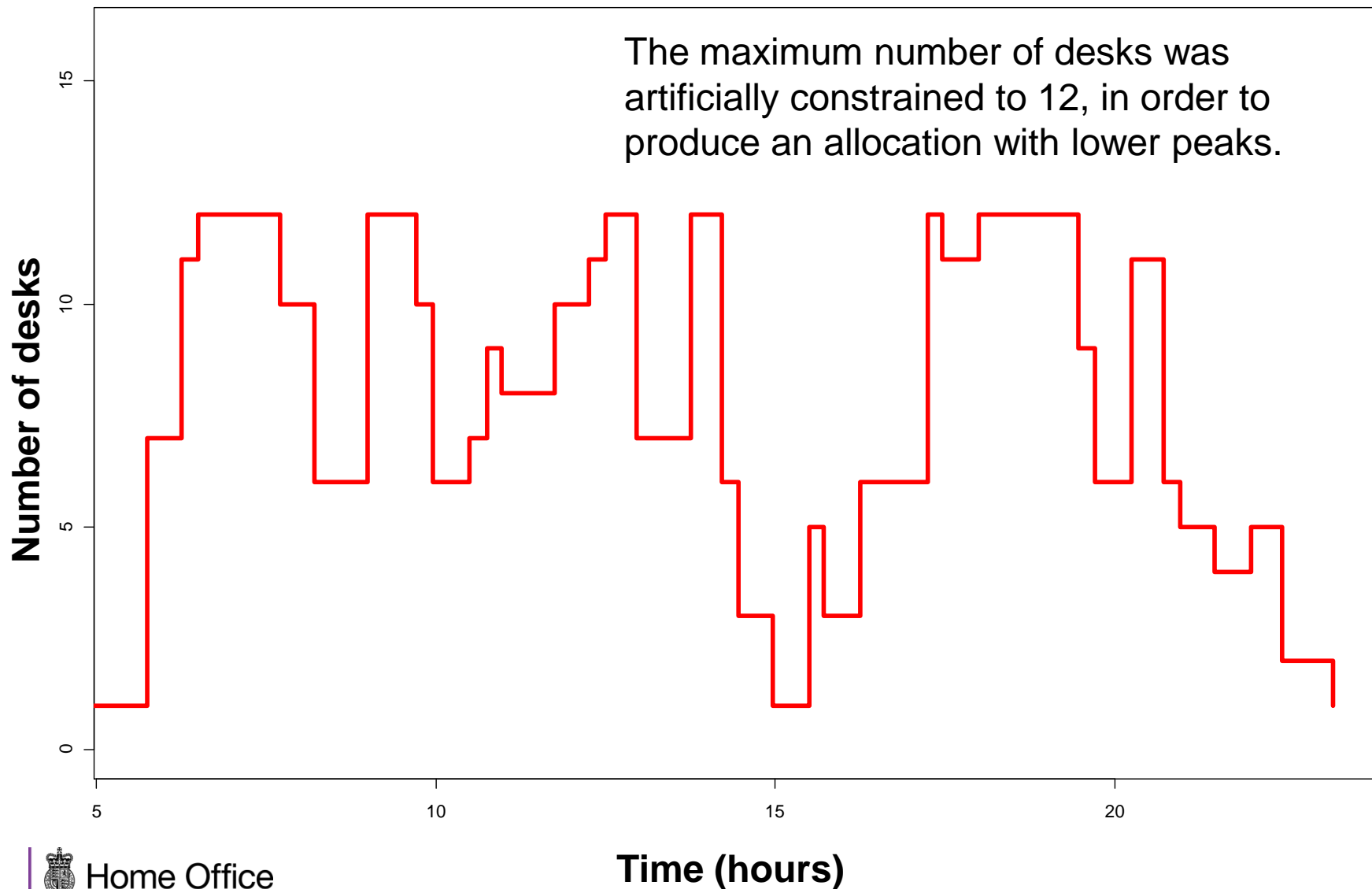
Number of desks open on a particular day...

Number of desks open at London Heathrow Terminal 4



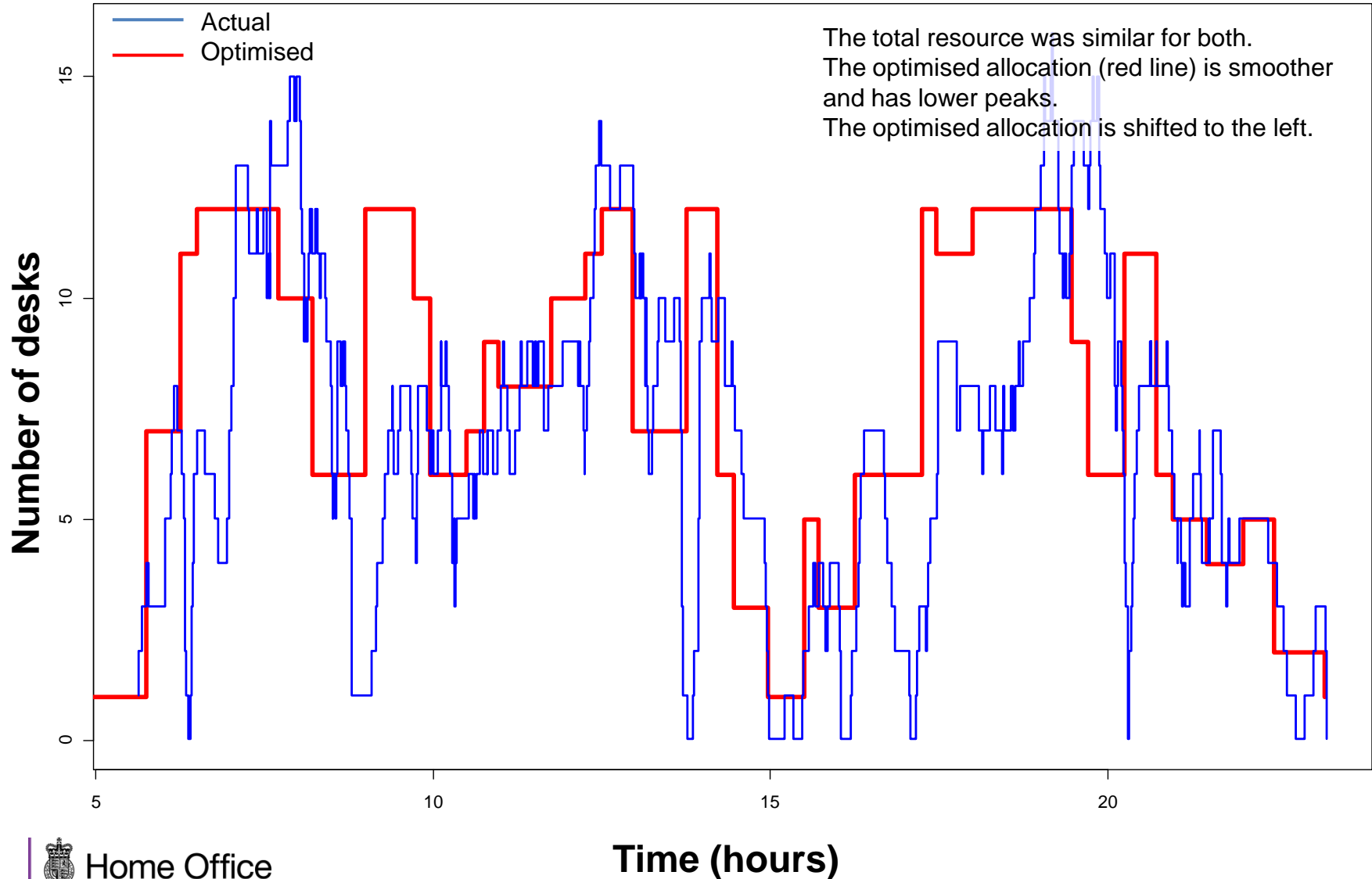
The number of desks proposed by the algorithm

Number of desks open at London Heathrow Terminal 4



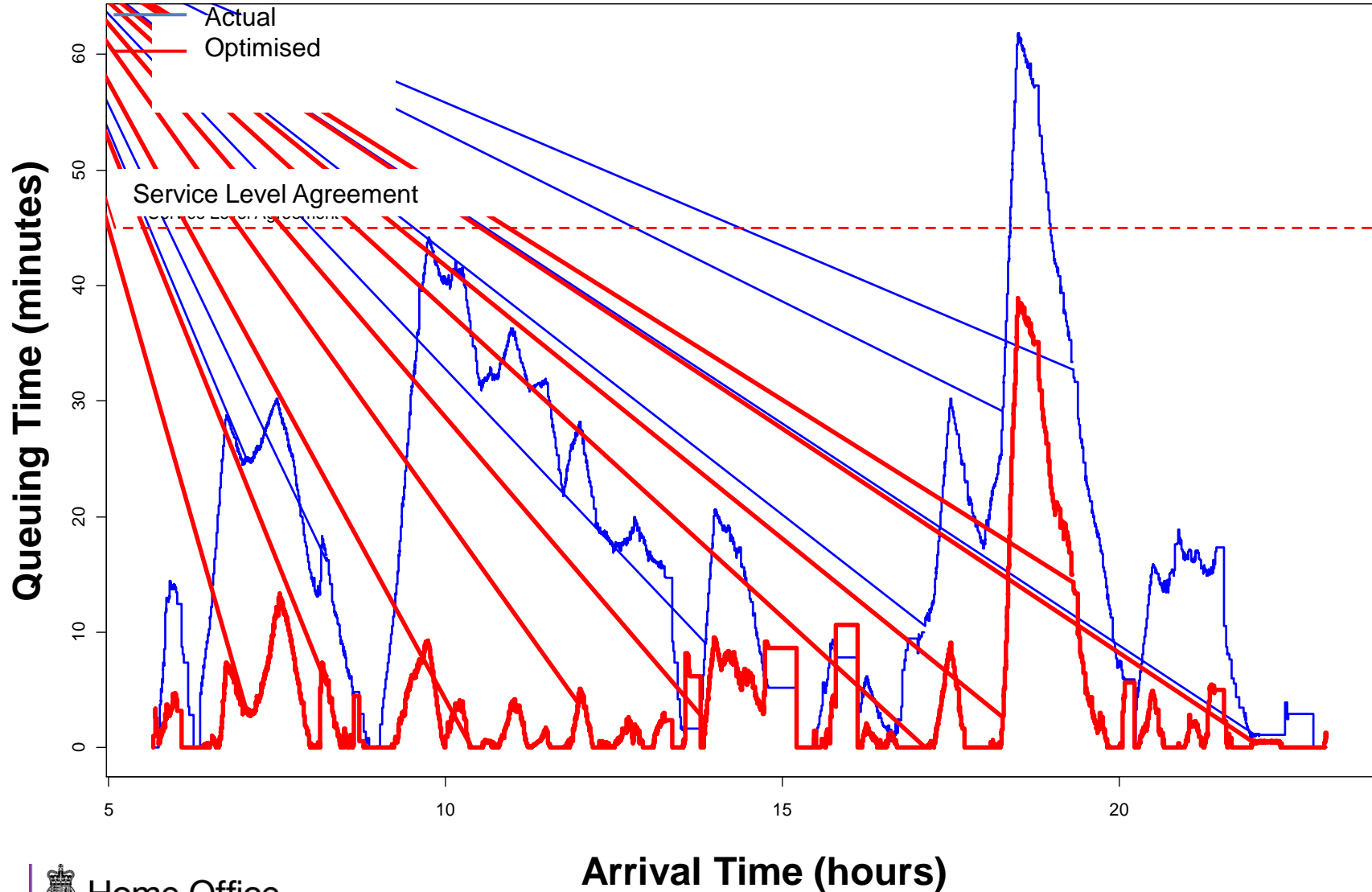
Overlay of actual and potential desk allocations

Number of desks open at London Heathrow Terminal 4



What difference does the new algorithm make?

Number of desks open at London Heathrow Terminal 4





Home Office

Modern Slavery Strategy

Modern Slavery and Multiple Systems Estimation

The National Crime Agency (NCA) Strategic Assessment of the Nature and Scale of Human Trafficking (2013)

- The National Crime Agency (NCA) has a framework for identifying victims of human trafficking and ensuring they receive the appropriate protection and support. In 2013, the NCA Strategic Assessment identified 2,744 unique potential victims of trafficking
- This measure is produced by collating information from various sources:

LA: Local Authority	NG: Non-governmental organisation
PF: Police force	NCA: National Crime Agency
GO: Government Organisation (mostly Home Office agencies e.g. UK Border Force, Gangmasters Licensing Authority)	GP: The general public, through various routes

The dark figure

- Despite all efforts, the Strategic Assessment can only present a partial picture of the size of human trafficking. This is because:
 - Human trafficking is a hidden crime and some victims may be controlled or still in servitude
 - Victims may not come forward due to feelings of fear and shame.
 - Some individuals may not be identified as victims of human trafficking by professionals who encounter them.
 - Some victims may not view themselves as victims of exploitation.
 - The NCA's coverage is not complete: only a limited number of agencies respond to the NCA's intelligence requirement (e.g. in 2013, 37 police forces and 4 NGOs responded to the intelligence requirement) so agencies may be aware of potential victims who are unknown to the Strategic Assessment.
- An approach called **multiple systems estimation (MSE)** can be used to estimate the “dark figure” of potential victims that do not come to the Strategic Assessment's attention.

Mark-recapture

- The simplest form of MSE is called “mark-recapture”.
- It was proposed in 1895 as a method of estimating the number of fish in a pond (though it is a much older idea). You catch a number of fish (say 100), mark them in some way, and then release them. Some time later you take a new catch (say another 100) and see how many of the second catch were part of the original first catch. If the overlap between the two catches is 20, for instance, the natural estimate of the whole population size is 500.
- A mark-recapture approach was used in the 2011 census to estimate the undercount, the proportion of the population not “caught” by the initial census, by conducting a survey and assessing the overlap between the census and the survey.



Multiple systems estimation: more than two lists

LA	X					X	X	X								X	X	X	
NG		X				X			X	X	X					X	X	X	X
PF			X				X		X			X	X			X	X		X
GO				X				X		X		X		X	X			X	X
GP					X						X		X	X					
number	54	463	995	695	316	15	19	3	62	19	1	76	11	8	4	1	1	1	???

- The “dark figure” is the number in the last column—which do not appear on any list
- A lot of work went into this table!

Results: the total figure

- The estimated confidence interval for the actual population size is from 10K to 13K, so this suggests that the Strategic Assessment is aware of 20% to 25% of all the potential victims of trafficking in the UK in 2013.
- This must be regarded as a tentative conclusion, because the model is based on assumptions that (while sensible) can't be easily verified and inevitably uses data that has some limitations.

Results: interactions

- Cases reported by a local authority are somewhat more likely also to be reported by an NGO or the police.
 - This may reflect the existence of referral pathways for potential victims between these agencies.
 - It may also reflect joint operations where victims are identified by more than one agency.
- There's a negative interaction (very little overlap) between GP and most other lists, so it suggests that the general public are “fishing in a different pool” in some way.
 - However to check robustness we tried leaving these out altogether and the results were scarcely different

The standard MSE model

- MSE works by fitting a model to the observed data.
- It fits main effects, for example the odds of appearing on list i
- It also allows for interaction effects, where the odds of appearing on both lists i and j are not simply the product of the individual odds. For five lists there will be ten possible interaction effects, but we try to fit a simpler model where only some of these are included
- Once the effects are all estimated, the model provides an estimate (with a standard error) for the total size of the population
- The fitted interactions show which interactions are positive and which negative.

The Poisson log-linear model

- For any collection A of lists, let N_A be the number of cases that appear on the lists in A and no others. Model N_A as $\text{Poisson}(\lambda_A)$ where

$$\log(\lambda_A) = \mu + \sum_{i \in A} \alpha_i + \sum_{\{i,j\} \in A} \beta_{ij}$$

- For example, the number of cases on list 2 only has expected number $\exp(\mu + \alpha_2)$ and the number on lists 3 and 4 but no others has expectation $\exp(\mu + \alpha_3 + \alpha_4 + \beta_{34})$
- The expected value of the dark figure is $\exp(\mu)$ since this corresponds to the case where the collection A is empty, so the estimate and profile likelihood of μ gives the estimate and confidence interval for the dark figure.

Interpreting the model

$$\log(\lambda_A) = \mu + \sum_{i \in A} \alpha_i + \sum_{\{i,j\} \in A} \beta_{ij}$$

- The parameters α_i are negative because the numbers on individual lists are smaller (considerably) than the likely dark figure
- The sign of the parameter β_{ij} indicates whether presence on list i is positively or negatively correlated with presence on list j

Fitting the model

- Use the R package `Rcapture`. The routine `closedp.mX` allows for an arbitrary design matrix giving specified interactions.
- Starting with a model fitting main effects only, stepwise add the two-factor interaction which makes the biggest improvement in the AIC, until the AIC starts to increase again or until the new interaction isn't statistically significant.
- Some individuals may appear on lists in groups (for example as a result of an operation which finds several victims) so the accuracy of estimation may be optimistically assessed by the standard Poisson model.
- Overall there may also be other reasons why the accuracy is optimistic.

```

Deviance Residuals:
      Min       1Q   Median       3Q      Max
-1.96919  -0.39529  -0.06652   0.14906   1.33782

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  9.05591    0.09305   97.320 < 2e-16 ***
LA           -5.08848    0.15254  -33.357 < 2e-16 ***
NG           -2.90507    0.09507  -30.558 < 2e-16 ***
PF           -2.14852    0.08809  -24.389 < 2e-16 ***
GO           -2.52177    0.09129  -27.624 < 2e-16 ***
GP           -3.30533    0.10827  -30.530 < 2e-16 ***
LA*NG        1.52395    0.27625    5.517 3.46e-08 ***
NG*GP       -2.92170    1.00582   -2.905 0.003675 **
PF*GP       -1.24675    0.31883   -3.910 9.21e-05 ***
LA*PF        0.92243    0.26209    3.519 0.000432 ***
GO*GP       -1.19052    0.36926   -3.224 0.001264 **
NG*GO       -0.55335    0.22399   -2.470 0.013495 *

---
Null deviance: 9559.514  on 30  degrees of freedom
Residual deviance: 16.351  on 19  degrees of freedom

```

R output for the fitted model

Robustness studies

- Vary data for model, for example:
- Use six-list data where “police” and “NCA” are separated
- Omit general public list on the grounds that matching with other lists may be problematic
- Omit the fairly small local authority list as well
- The overall result for the estimate of the dark figure is robust against all these variations

Source data	Estimate of total population	Standard error	Lower confidence limit	Upper confidence limit	Two factor interactions fitted
Six lists	11418	809	9982	13181	PF*NCA LA*NG NG*GP LA*PF PF*GP GO*GP NG*GO
Combine NCA into PF	11313	802	9889	13063	LA*NG NG*GP PF*GP LA*PF GO*GP NG*GO
Combine NCA into PF; omit GP	11015	805	9587	12771	LA*NG LA*PF NG*GO
Combine NCA into PF; omit GP, LA	10951	807	9520	12714	NG*GO
Omit GP	11048	808	9615	12810	PF*NCA LA*NG LA*PF NG*GO
Omit GP, LA	10951	808	9547	12697	PF*NCA NG*GO

UK is home to 13,000 slaves: Home Office says number is four times higher than previously thought

- The shocking figures include women trafficked for prostitution or service
- Others are brought in to work in factories, fields or even fishing boats
- The Home Office confirmed it underestimated the scale of the problem
- The most common victims are from Romania, Poland, Albania and Nigeria

By MAIL ON SUNDAY REPORTER

PUBLISHED: 01:55, 30 November 2014 | UPDATED: 15:56, 30 November 2014



Up to 13,000 people in Britain are being held in conditions of slavery, four times the number previously thought, the Home Office has said.

In what is said to be the first scientific estimate of the scale of modern slavery in the UK, the Home Office has said the number of victims last year was between 10,000 and 13,000.

They include women forced into prostitution, domestic staff and workers in fields, factories and fishing boats.

Scroll down for video



Home Secretary Theresa May, pictured, described the scale of human trafficking as 'shocking'

Data from the National Crime Agency's Human Trafficking Centre had previously put the number of slavery victims in 2013 at 2,744.

The new estimate is based on a statistical analysis by the Home Office chief scientific adviser, Professor Bernard Silverman, which aims for the first time to calculate the 'dark figure' of victim numbers who are not reported to the law enforcement agencies.

'Modern slavery is very often deeply hidden and so it is a great challenge to assess its scale,' he said.

'The data collected is inevitably incomplete and, in addition, has to be very carefully handled because of its sensitivity.'

Read more:

<http://www.dailymail.co.uk/news/article-2854601/UK-home-13-000-slaves-Home-Office-says-number-four-times-higher-previously-thought.html#ixzz3L2PU4VAD>

Reflections: science and policy

- Accuracy of reporting: did the paper get it right?
- Thinking about how it will be received
- Explaining a model and a confidence interval
- The Minister explained it better than I could
- Importance of relationships
- Science really does inform policy

Thank
you!