

Replication of ‘Improving Overnight Loan Identification in Payment Systems’ on Simulated Data

Mark Rempel

1 Introduction

This replicates the primary results given in Rempel (2016) on simulated dummy data.¹ The results here are given simply for illustration of the methods/computations used.

1.1 Instructions of use

Run each code chunk segment in order (recommend running each of the command files in section 2 individually directly on a linux/unix server due to computation times etc.). See README and individual scripts for details.

2 Constructing the data

2.1 Baseline Testing

```
> #system(master_command_file)
```

2.2 kNN Refined Variant

```
> #system(kNN_replication_file)
```

2.3 Other Refinements

```
> #system(other_refinements_command_file)
```

¹Note Table 1 omitted since no subsamples to consider with simulated data and Table 3 implements the exact same computations.

3 Output of primary results/figures

3.1 Figure 1

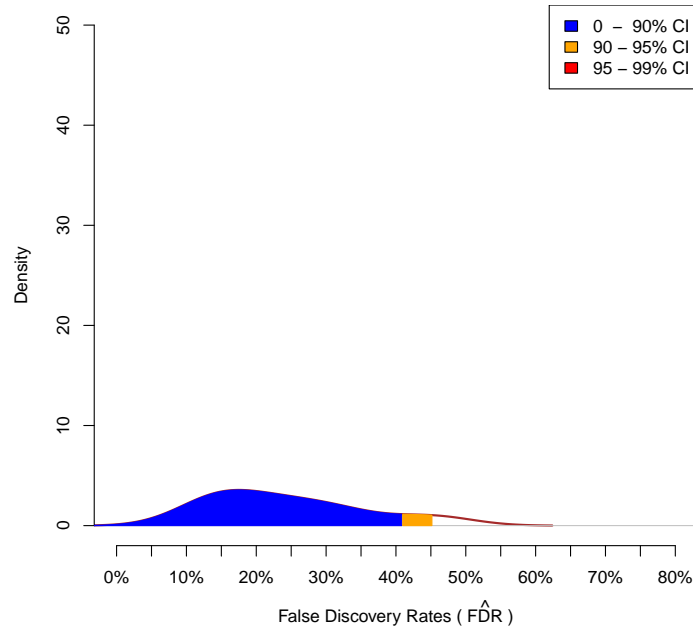


Figure 1: Positive False Discovery Rates - Baseline Algorithm, Simulated Data.
Note the 95% and 99% pFDR are the same for the simulated data.

3.2 Table 2

```
> require(xtable)
> source('typeIerrorfunctions_pfdincluded.R')
> load('FurfineOutput/BaselineOriginalFurfineOutputDummyData.RData')
> furfdat.split <- split(furfdat, as.factor(furfdat[,1]))
> furfdups <- DuplicatePercentagesTable(furfdat.split)
> load('Results/BaselineTypeI_1.RData')
> results <- BaselineTypeI$raw
> Full_dups <- DuplicatePercentagesTable(results)
> DupTable <- cbind(furfdups, Full_dups)
> colnames(DupTable) <- c('Original Sample', 'Permuted Sample')
> xtable(t(DupTable)*100)

% latex table generated in R 3.1.2 by xtable 1.7-4 package
% Sun Nov 01 20:18:21 2015
\begin{table}[ht]
\centering
\begin{tabular}{rrrrrrr}
\hline
& 1\% & 25\% & 50\% & Mean & 75\% & 99\% \\
\hline
Original Sample & 0.00 & 0.00 & 4.55 & 5.50 & 8.31 & 20.42 \\
Permuted Sample & 0.52 & 17.88 & 25.00 & 27.50 & 35.72 & 65.71 \\
\hline
\end{tabular}
\end{table}
```

	1%	25%	50%	Mean	75%	99%
Original Sample	0.00	0.00	4.55	5.50	8.31	20.42
Permuted Sample	0.52	17.88	25.00	27.50	35.72	65.71

Table 1: Replication of Table 2 on Simulated Data

3.3 Table 3

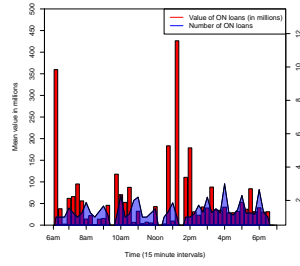
```
> FurfData.filesnames <- c('BaselineOriginalFurfineOutputDummyData.RData', 'EndOfDayFurfineOutputDummyData.RData')
> TypeIresult.list <- c('BaselineTypeI', 'EndOfDayTypeI', 'RoundIntTypeI', 'ExactIntTypeI')
> num_results_files <- 1
> source('AggregateSimulatedResults.R')
> require(xtable)
> xtable(Table3)
```

```
% latex table generated in R 3.1.2 by xtable 1.7-4 package
% Sun Nov 01 20:18:42 2015
\begin{table}[ht]
\centering
\begin{tabular}{rrrrrrrr}
\hline
& Mean(FDR) & x=0.2 & x=0.1 & x=0.05 & x=0.01 & B & \% Full Sample \\
\hline
kNN & 21.21 & 60.00 & 100.00 & 100.00 & 100.00 & 15.00 & 97.57 \\
Baseline & 24.44 & 60.00 & 100.00 & 100.00 & 100.00 & 15.00 & 100.00 \\
End Of Day & 21.41 & 46.67 & 86.67 & 93.33 & 93.33 & 15.00 & 31.46 \\
Exact Int & 5.21 & 0.00 & 0.00 & 73.33 & 73.33 & 15.00 & 51.31 \\
Round Int & 9.81 & 0.00 & 46.67 & 93.33 & 100.00 & 15.00 & 61.80 \\
\hline
\end{tabular}
\end{table}
```

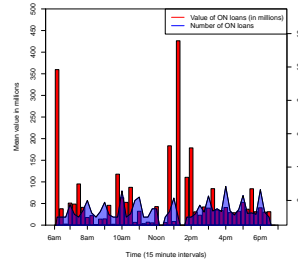
	Mean(FDR)	x=0.2	x=0.1	x=0.05	x=0.01	B	% Full Sample
kNN	21.21	60.00	100.00	100.00	100.00	15.00	97.57
Baseline	24.44	60.00	100.00	100.00	100.00	15.00	100.00
End Of Day	21.41	46.67	86.67	93.33	93.33	15.00	31.46
Exact Int	5.21	0.00	0.00	73.33	73.33	15.00	51.31
Round Int	9.81	0.00	46.67	93.33	100.00	15.00	61.80

Table 2: Replication of Table 3 on Simulated Data 31st

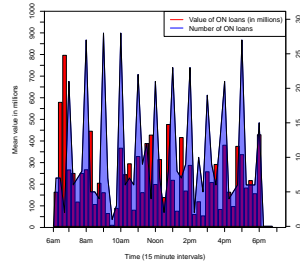
3.4 Time of Day Bias plot



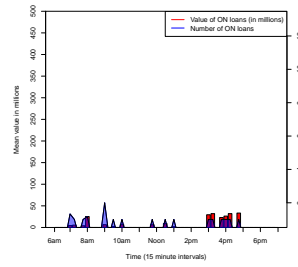
(a) Uniquely Identified Furfine Loans - Simulated Data.



(b) All Furfine Loans Identified - Simulated Data.



(c) Faux Loans - Simulated Data



(d) Figure 2: Non-uniquely identified loans - Simulated Data.

Figure 2: Replication of Figure 2 on Simulated Data.

3.5 Output of Discussion Results

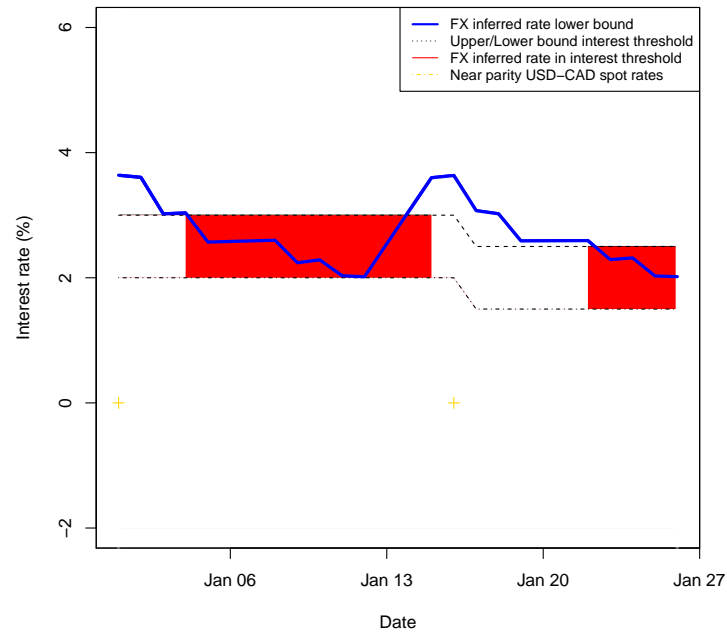


Figure 3: Replication of Figure 3 - Interest/Exchange rate configuration and type I errors on Simulated data.
Note no spurious matches obtained due to FX configuration from the simulated data.