Appendix - Improving Overnight Loan Identification in Payment Systems

Mark Rempel

June, 2015

A Description of Payments Data

The Canadian RTGS equivalent payments system, the Large Value Transfer System (LVTS), had on average approximately 22,000 transactions totalling $140 billion per day between Q1 2005 and Q4 2011, with the average transaction of roughly $6.5 million.[1] Table 1 presents basic daily summary statistics of the LVTS restricted to payments that would be considered by the Furfine algorithm (i.e. those exceeding $ 1 million). This restriction leads to a consideration of about a quarter of the total daily volume of transactions (on average) as possible loan payments.

Between 2005 and 2010 there were 14 direct LVTS participants (excluding the Bank of Canada) with a new entrant in 2011. From Table 1 it can be seen that on almost all days all the direct participants participate in the LVTS and regularly transact with a majority of the other participants (e.g. the average participant interacts with two-thirds of the other participants on a daily basis). The average concentration of the value of payments made between a given participants' counterparties (measured by the Hirschman-Herfindahl Index applied on the participant basis, that is, the average squared share of payments sent by a participant to each counterparty) is 0.01, while the maximum squared share a participant had with a counterparty on an average day was 0.12.

The time-series path of the volume of transactions and size are provided in Figures 1 and 2 respectively. While the number of transaction is relatively stable over time, the total size grew substantially over the pre-crisis period and demonstrated substantial variability over time. The

[1]Details on the LVTS and how it relates to other RTGS systems is provided in Arjani and McVanel (2006).

outlying nature of payments for days falling on provincial, and U.S. holidays without a Canadian counterpart is evident from both the time-series of the daily volumes and size.

B Furfine Algorithm Data

The variant of the Furfine algorithm referred in the paper as 'Baseline Furfine' is specified as follows:

1. *Transfers involving the Bank of Canada or that are less than $1 million CAD are dropped.*

2. *For a given day d and bank pair $\{i, j\}$, the algorithm constructs a set of potential principal payments x_{ijd} (t denotes the time stamp) in which the payments are in round dollar increments.*[2]

3. *For each potential principal payment x_{ijd}, the algorithm finds all return payments $x_{ji(d+1)}$ from bank j to bank i the following day, constructing a set of duples $L_{ijd}(x_{ijd}) = (x_{ijd}, x_{ji(d+1)})$ that consists of all combinations of inter-day bilateral transactions.*

4. *The implied (annualized) interest rate is then computed for each duple $(x_{ijd}, x_{ji(d+1)})$ and the set of duples is restricted to those with an interest rate within a plausible band about the target overnight rate. Here this band is set as plus or minus 50 basis points.*

5. *This remaining set of duples $\tilde{L}_{ijd}$ (if non-empty) may or may not consist of unique payment legs. In the case where multiple payments on the return day are paired with the same payment on the loan sending day, or vice-versa, (referred to hereon as the non-unique/duplicate loans) the algorithm selects the duple with the closest interest rate to the target.*

Applying the baseline Furfine algorithm to the LVTS data (excluding the outlying payment days and the days preceding them) I identify an average of 270 payment pairs as (Furfine-identified) overnight loans per day. This translates to an average of 6.1% of the eligible LVTS payments are identified as part of a loan each day. Table 2 provides the corresponding daily summary statistics for these identified loans.

Note, the common refinements considered in the paper include checks on the interest rate being rounded to the nearest basis point and that the principal loan payment time stamp be

[2]Many variants (including the Fed's variant-MaPS) require the principal to be in $100,000 increments.

after 4pm. Code implementing each of these variants is available in the Code section of the Appendix.[3]

C K-NN Refined Furfine Algorithm

The proposed refinement for treatment of duplicate loans (i.e. multiple Furfine-identified loans involving an identical payment on one of the legs of the loan) has the following procedure:

1. *Construct the training model:*

    (a) *Obtain a list of true overnight loans (or proxy using uniquely identified Furfine loans)*

    (b) *Generate an equal number of 'faux loans'*

    - *randomly match payments (exceeding $1 million) sent between from the same sending/receiving participants*

    (c) *Compute the relevant loan covariates/characteristics*

    i. *Loan duration (in hours)*

    ii. *Rounded Interest Indicator*

    iii. *Repayment difference between actual payment and that implied from rounded interest rate*

    iv. *Principal loan submission time (in number of hours from midnight)*

    v. *Difference of the implied rate from target*

    vi. *Size of principal*

    (d) *Normalize the values within the unique/'faux' loan set independently (this ensures comparability between the obtained levels)*

    (e) *Specify a choice of 'K' nearest neighbours*

    - *Can be determined via cross-validation*
    - *(Recommend considering only odd 'K' in order to guarantee unique majority rule)*

2. *Identify the set of potential loans using steps 1 - 4 of the baseline Furfine algorithm*

[3] In addition to excluding all federal and provincial holidays, federal U.S.-specific holidays are also excluded due to the distinct payment distribution observed on these days.

3. *Take the set of non-uniquely identified loans from the potential loan set, and compute the loan covariates (steps 1c/d)*

4. *Predict whether the potential duplicate loan pairs are true loans using the trained model*

   - *If multiple loans remain that involve the same payment, select the pair with interest rate closest to the target (or other ad-hoc duplicate treatment method)*

Economic differences between the proposed k-NN method and the baseline Furfine algorithm are shown in Figures 5 - 7 (where the comparison (implied bias) is: 'baseline Furfine -refined algorithm). Code on the precise implementation of this procedure is available in the Code section (available at markrempel.org/).

D Estimating False Discovery Rates

D.1 Algorithm

The algorithm for estimating the bounds on the false discovery rate for a sample is as follows:

1. *Randomly sample B^N days from the payment set to be principal loan days ($B^N \approx$ number of days in sample).*

2. *For each day (b^s) sampled, referred to as a lend day, sample B^R non-adjacent days that share the same overnight target rate as the lend day.*

3. *Run the Furfine algorithm on each of these days pairs (treating the pair as if they were adjacent business days) - yielding $D^0_{b^s,b^r}, b^r = 1, .., B^R$ for each lend day b^s.*

4. *Average over the b^r for a given b^s to get $E[D^0]_{b^s}$,*

$$E[D^0]_{b^s} = \frac{1}{B^R} \sum_{b^r=1}^{B^R} D^0_{b^s,b^r}, \tag{1}$$

$$Pr(D^0 > 0)_{b^s} = \frac{1}{B^R} \sum_{b^r=1}^{B^R} \mathbf{1}(D^0_{b^s,b^r} > 0). \tag{2}$$

5. *Take the number of loans identified by the Furfine algorithm with the correct day ordering $D_{b^s,b^{s'}}$ and compute the estimated FDR for that lend day:*

$$\widehat{FDR}_{b^s} = \frac{E[D^0(\Gamma)]_{b^s}}{Pr(D^0 > 0)_{b^s} D(\Gamma)_{b^s,b^{s'}}}.^4 \tag{3}$$

Code implementing this procedure on dummy data is available in the Code section of the online appendix (available at markrempel.org/).

D.2 Methodology Details

As presented in the paper, the Furfine algorithm can be reformulated as a multiple hypothesis test, with each payment on the principal day d (x_i^d) having an associated test statistic,

$$T_i^{dd'} = min\left\{|\text{interest rate}(x_i^d, x_j^{d'}) - \text{ON target}_d| : sender(x_j^{d'}) = receiver(x_i^d), sender(x_i^d) = receiver(x_j^{d'})\right\},$$

where 'ON target$_d$' is the prevailing overnight target rate on that day. With this definition each test statistic T_i gives the closest possible inferred interest rate to the target of payment x_i on day d from the set of candidate (return) payments on day d'. Naturally, the null hypothesis (H_{0i}) for each payment x_i is that it is not the principal payment leg of an overnight loan.

The associated rejection region, Γ_λ, is the set of payments for which the difference in the implied interest rate from target falls below a threshold, λ (set at 50 basis points for the baseline Furfine), and that are consequently classified as candidate overnight loans by the algorithm. Specifically, $\Gamma_\lambda \equiv \{T_i : T_i \leq \lambda\}$ and the daily number of loans identified, or 'discovered', by the Furfine algorithm between day d and d' is $D_{dd'} \equiv \sum_i^m \mathbf{1}\{T_i^{dd'} \in \Gamma_\lambda\}$.

	Accept	Reject ($T_i \in \Gamma$)	Total
Not a loan (H_{0i})	W	V	m_0
Overnight Loans (H_{1i})	B	S	m_1
Total	N	D	m

Table 1: Contingency Table for algorithm run on days (d, d').

The contingency table associated to applying the Furfine algorithm to day d (with return payments on day d') is given in Table 1. With multiple hypotheses, the seemingly natural analog

[4] To keep the estimator well behaved, set D = 1 if $D = 0$. See Storey (2002) for greater discussion on this choice.

to the probability of a false positive (α) in a single hypothesis, is the proportion of output that are false positives. That is, $FDR = E\left[\frac{V}{D}\right]$.

A slight issue with this definition is that this rate is only defined if $D > 0$. Consequently, this False Discovery Rate as proposed by Hochberg (1995) is in fact $FDR = E\left[\frac{V}{D}|D > 0\right] \times Pr(D > 0)$. Given this, Storey (2003) argued the appropriate estimator of the rate of false positives is the 'positive False Discovery Rate:' $pFDR = E\left[\frac{V}{D}|D > 0\right]$. This formulation has the attractive property of being directly interpretable as the probability of a false positive for the case of independent statistics and hypotheses (i.e. $pFDR = Pr(H_0|T \in \Gamma)$). It should be noted that, as a minor abuse of notation, the false discovery rate used in the paper is in fact the pFDR. Given the number of payments a day, the distinction between the FDR and pFDR is negligible in our context, but the choice does impact the form of the estimate $\widehat{FDR}$ used. (I will continue this abuse of notation for the remainder of this section.)

Of course, complete independence amongst all payments seems unlikely, especially considering that overnight loans are expected to be predicated in part on the net flow of payments preceding it. Storey and Tibshirani (2001) investigate the properties of the FDR under various forms of dependence and show that, provided the dependence amongst payments does not persist beyond finite clusters of payments,

$$FDR \leq \frac{E[\sum_i^m \mathbf{1}\{T_i^{dd'} \in \Gamma_\lambda, H_{0i}^{dd'}\}|D_{dd'} > 0]}{E[D_{dd'}]}, \forall m \tag{4}$$

and taking $m \to \infty$ the above holds with equality.

This form provides a useful means of conservatively estimating the FDR in practice, where only N, D and m are actually observed. To see this, first note from Table 1 that $E[V, H_{0i}^{dd'}\}] = m \cdot Pr(T \in \Gamma, H_0) = m \cdot Pr(H_0) \cdot Pr(T \in \Gamma|H_0)$. By conditioning on $D > 0$, we must normalize by $\frac{1}{Pr(D>0)}$. Thus, $E[\sum_i^m \mathbf{1}\{T_i^{dd'} \in \Gamma_\lambda, H_{0i}^{dd'}\}|D_{dd'} > 0] = m \cdot Pr(H_0) \cdot Pr(T \in \Gamma|H_0)\frac{1}{Pr(D>0)}$.

To acquire an estimate for $m \cdot Pr(T \in \Gamma|H_0)$, I re-pair day d with another day that is not the next business day (i.e $d_r \neq d'$) generating a series of tests (T^0) that cannot constitute a true overnight loan by definition. Assuming that these non-loan tests T^0 are drawn from the same distribution as the original non-loan payment tests ($T|H_0$), and that the repayment legs culminating in T^0 are independent of those in T, the number of these tests that fall in the rejection region, $D^0 \equiv \sum_i \mathbf{1}(T_i^0 \in \Gamma) = \sum_i^m \mathbf{1}(T_i^0 \in \Gamma | H_{0i})$ times $Pr(H_0)$ in expectation equals $E[V|D > 0] = E[\sum_i^m \mathbf{1}\{T_i^{dd'} \in \Gamma_\lambda, H_{0i}^{dd'}\}|D_{dd'} > 0]$.

Given this, if I could obtain a consistent estimate of $Pr(H_0)$ and $\frac{1}{Pr(D>0)}$ and repeating this process B^R times (sampling with replacement B^R days to be treated as the repayment day to day d) would yield an unbiased and consistent estimate of the righthand side of (4). Unfortunately, this is non-trivial. As $Pr(H_0)$ is believed to be close to 1 given the number of payments generally transacted as compared to the number of overnight loans, I avoid the issue of estimating $Pr(H_0)$ by setting $\widehat{Pr}(H_0) = 1$ (thereby inserting an additional, if slight, upward bias). I estimate $Pr(D > 0)$ by $Pr(D^0 > 0) = \frac{1}{B^R}\sum_{d^r=1}^{B^R} \mathbf{1}(D_{dd^r}^0 > 0)$. Again provided the payment days are independent and identically distributed, this estimate is downward biased (making the reciprocal upward biased) since $E[D^0] = E[V] \leq E[V] + E[S] = E[D]$. Combining all of the above, along with D being observed, yields my estimate of the FDR:

$$\widehat{FDR}_{d_s} = \frac{E[D^0]_{d_s}}{D_{d_s d_s'}} \cdot \frac{1}{Pr(D_{d_s}^0 > 0)}. \tag{5}$$

The re-pairing and sampling of complete days of repayments to estimate the rate of false positives is predicated on 1) the assumption of intra-day dependence amongst payments, 2) that payments are dependent intra-day in only finite clumps, and 3) that there is no inter-day dependence excepting overnight loans.[5] These assumptions simultaneously justify the assignments of the estimated False Discovery Rates as upper bounds on the type I error and, via arguments analogous to Tran et al. (2014), the block-bootstrap procedure to acquire a confidence bound on the upper limits of this rate for the full sample of payments.

D.3 Sensitivity of Results

The selection of B^R and B^N used in the paper were chosen at a level that already pushed the memory capacity of individual objects on the cluster. Naturally, increasing B^R has a corresponding increase to the number of observations used to compute the average false discovery rate and hence the precision of the estimate (provided B^R is less than or equal to the number of days in

[5]One immediate violation of this assumption is term loans, however, any systematic bias is expected to be small, and upwards given that term loans are generally very large and unlikely to be identified by the algorithm except with another term loan.

the same policy rate window). Increasing B^N will also increase the precision of the upper-bound for the sample.

Table 5 demonstrates the robustness of the estimated FDR for the baseline Furfine algorithm to smaller B^N. It shows the average rate varies by 0.5% as we vary B^N from 100 to 1,500. Similarly the right tail of the estimated false discovery rate remains bounded above by 20% at the 99% confidence level across all samples.

The sampling of (false) repayment days for a given principal day within a policy rate window relies on the assumption that the resampled day has distribution of payments that are drawn from a sufficiently similar distribution to the true repayment day. Resampling from the same policy rate window rather than a narrower fixed window about the sample date lessens the likelihood of dependence between the resampled day and the original pair of days, but comes at the cost of a slightly higher risk of significant differences in payment distribution. With that caveat in mind, the upper panel of Table 4 shows the sensitivity of the results on the Baseline Furfine algorithm by computing the same False Discovery Rate statistics on only the subset of sampled repayment days that fall within distinct fixed windows.

The results from these partitions imply some sensitivity to the fixed window, with the average estimated rate spiking to 12% for resampled pairs sharing the same week as each-other. In general, reflecting the higher dependence amongst the sampled dates, the shorter term, fixed windows lead to a wider spread in the estimated distribution of false discovery rates with the 95%/99% bounds no longer falling below 20%, while nearly 12%/3% of the estimated bounds are below 5%/1% (compared with 0.5%/0% in the full sample). The longer-horizon fixed window of resampled payment days that are more than a month apart yield results which most closely resembles the full sample with the right-tail drawn below 20% and the left-tail above 1% at the 99% confidence level.

The lower panel of Table 4 is the same decomposition but for the (k-NN) refined algorithm. While sharing the same qualitative features across windows as the baseline algorithm, the refinement, even within the one week window, displays much less sensitivity to the fixed-window. The estimated average false discovery rate for the worst-case window is at 9% with the 95% confidence bound below 20%. For the remainder of cases the 99% bound is at or below the 20% level.

E Supplementary Figures/Tables

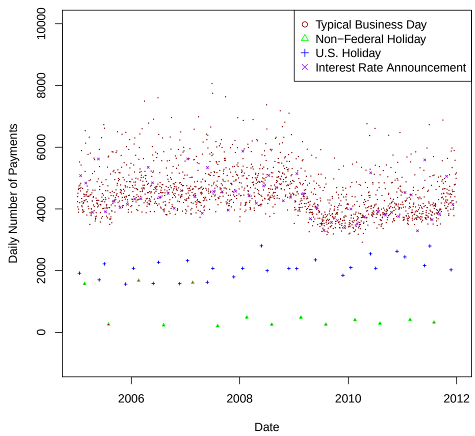

Figure 1: Daily Number of Payments - LVTS transaction exceeding a million, Q1 2005 - Q4 2011

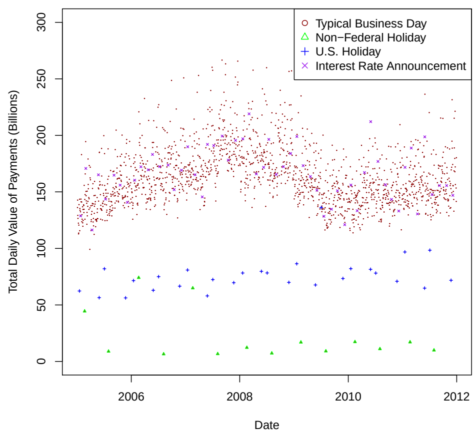

Figure 2: Daily Value of Payments (in Billions) - LVTS transactions exceeding a million, Q1 2005 - Q4 2011

	Min	25%	Median	Mean	75%	Max
Number of Transactions	206	3961	4386	4447	4868	8065
Number of Counterparties (Min)	0	5	5	5.13	6	8
Number of Counterparties (Mean)	4.14	10.13	10.36	10.33	10.64	11.64
Number of Counterparties (Max)	9	14	14	14	14	15
Min payment share	0.00	0.00	0.00	0.00	0.00	0.00
Mean payment share	0.01	0.01	0.01	0.01	0.01	0.02
Max payment share	0.06	0.11	0.12	0.11	0.12	0.15
Mean Value (Millions)	23.71	33.87	36.34	36.38	38.90	53.07
Total Value (Millions)	6497	143900	158900	160800	176900	266600

Table 2: Daily Summary Statistics, LVTS payments data (exceeding $1 million CAD), Q1 2005 - Q4 2011.
Rows 2 - 4 provide the minimum, mean and maximum number of different receivers of payments from a given participant. Rows 5 - 6 provide the minimum, mean and maximum payment value shares (squared share of total value of payments sent by day), the mean of which corresponds to the computation of the Hirschman-Herfindahl Index applied to a given participant rather than the market as a whole. Row 7 and 8 provide the average and total value of payments greater than a million transacted per day.

	Min	25%	Median	Mean	75%	Max
Number of Payments	2821.00	3844.00	4261.00	4372.00	4741.00	7898.00
Number of Furfine Loans	129.00	249.00	270.00	275.20	294.00	434.00
Proportion Uniquely Id'd	0.64	0.80	0.83	0.82	0.85	0.92
Furfine Loans Daily Value (Billions)	5.44	11.88	13.49	13.83	15.66	22.60
Mean Furfine Daily Value (Billions)	0.02	0.04	0.05	0.05	0.06	0.08
Effective Furfine Rate (%)	0.17	0.95	2.52	2.29	4.02	4.67
Effective Furfine Spread (bps)	-16.25	-3.65	-0.47	-0.62	2.01	17.29
Std. Dev. Furfine Spread (bps)	5.00	8.22	10.26	11.39	13.92	26.76

Table 3: Furfine Algorithm Summary Output, 2005 - 2011.
Note the data excludes federal/provincial and U.S. holidays as well as the day immediately preceding each of them.

Figure 3: Daily Number of Overnight Loans - (Baseline) Furfine Output, Q1 2005 - Q4 2011.

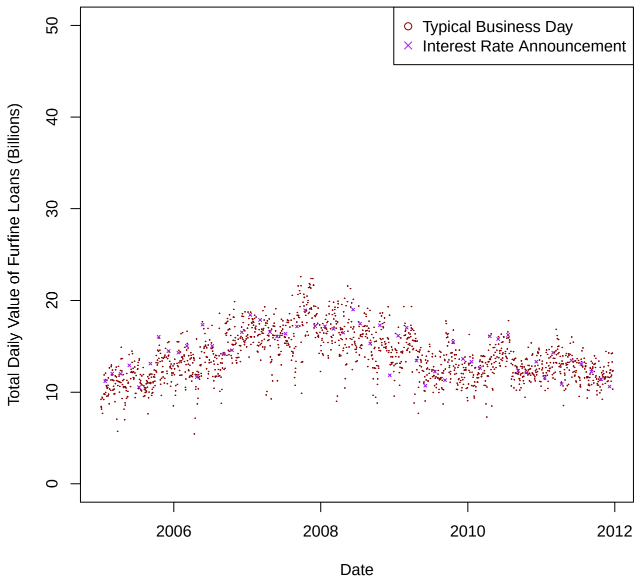

Figure 4: Total Daily Value of Overnight Loans - (Baseline) Furfine Output, Q1 2005 - Q4 2011.

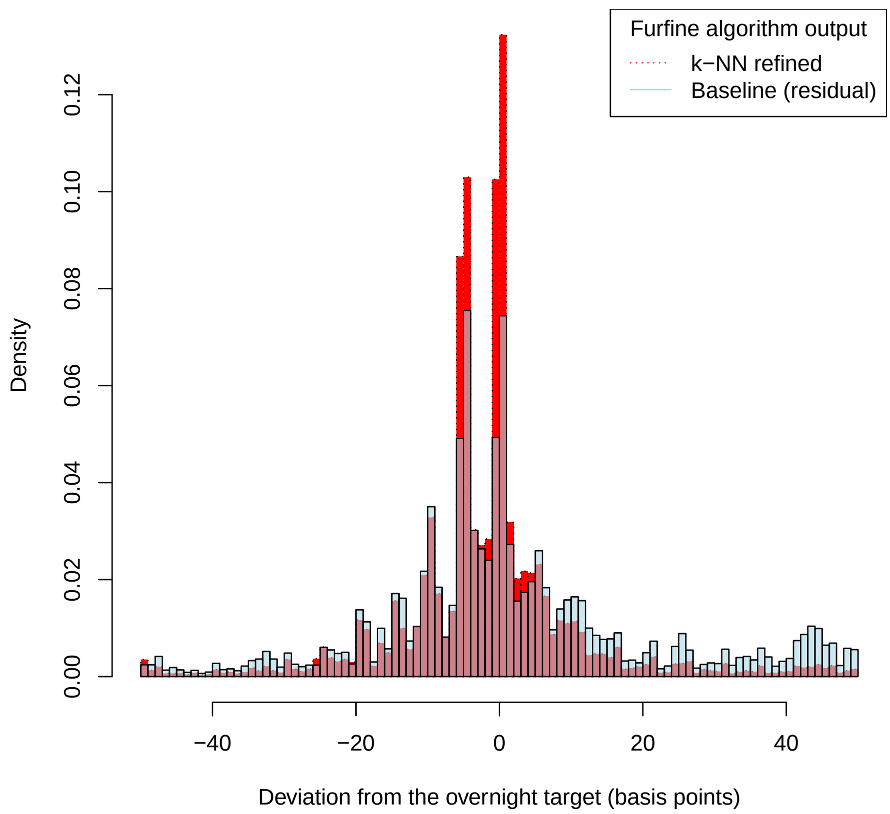

Figure 5: Implied Bias in Inferred Interest Rates, Q1 2005 - Q4 2011.
Implied bias of the Baseline Furfine algorithm output vs. refinement (Baseline - kNN refinement). The blue bars depict the interest rates of the baseline algorithm that correspond to loans not shared by the refined variant.

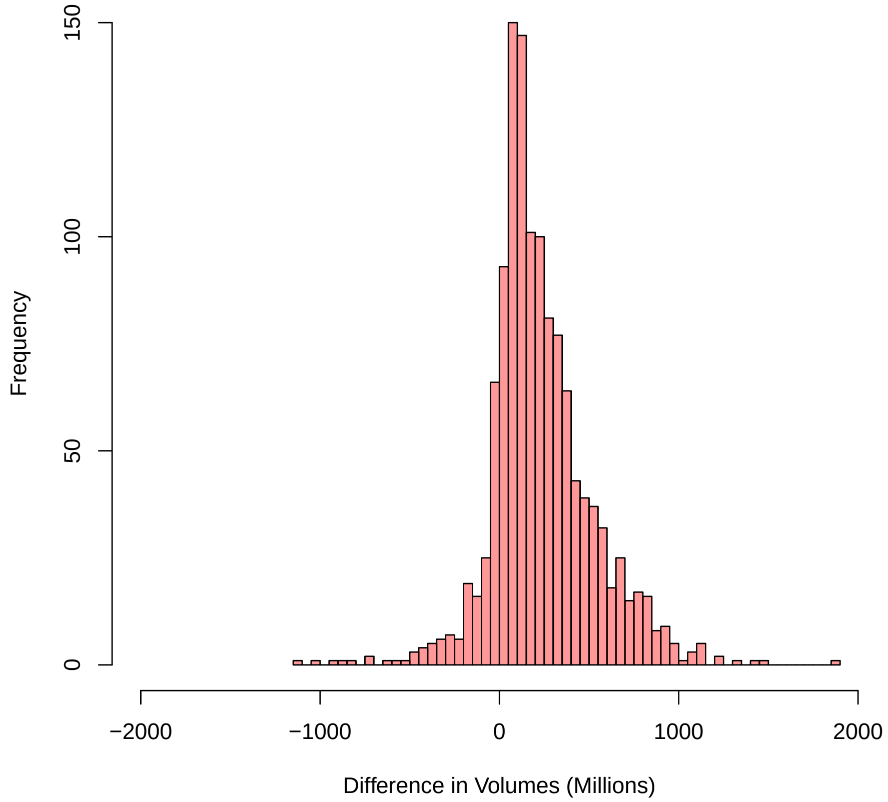

Figure 6: Implied Bias in Size (Millions), Q1 2005 - Q4 2011.
Implied bias of the Baseline Furfine algorithm output vs. refinement (Baseline - kNN refinement).

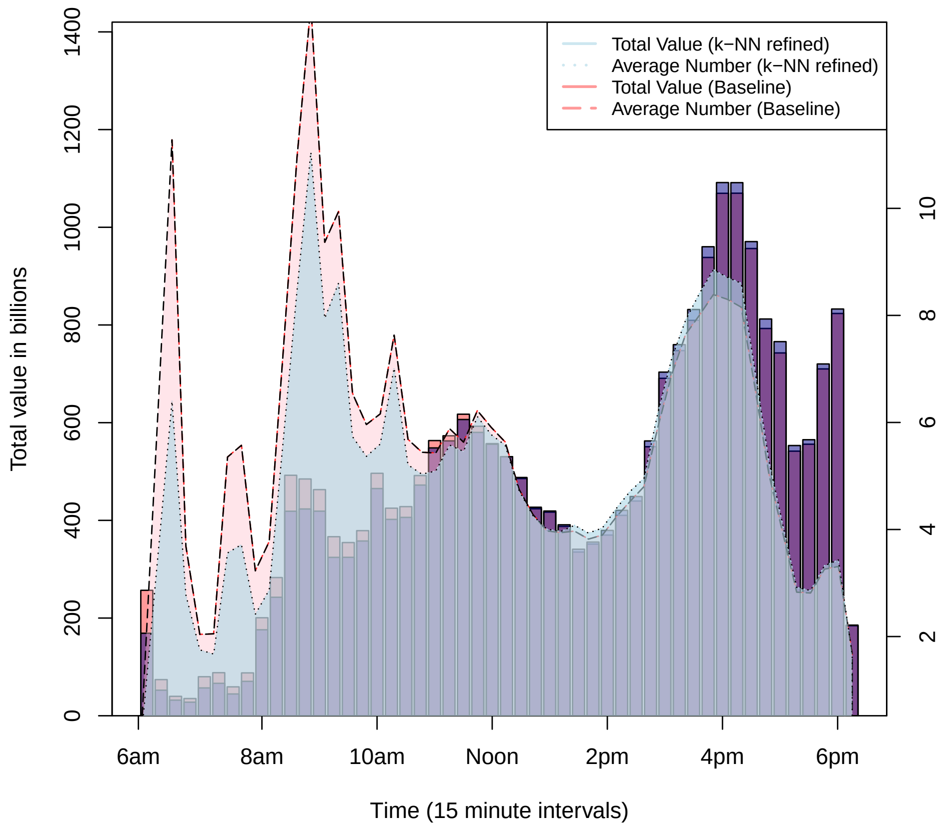

Figure 7: Implied Bias in the joint Value x Time Distribution, Q1 2005 - Q4 2011.
The barplot corresponds to the total value of all loans identified by the Baseline/kNN algorithm within a given 15 minute window, while the shaded line plots correspond to the number of transactions in that same window (right axis). The figure depicts the bias of the baseline algorithm towards a larger value/volume of loans earlier in the day versus the refinement.

		Mean($\widehat{FDR}$)	Pr($\widehat{FDR} > x$) $x = 0.2$	$x = 0.1$	$x = 0.05$	$x = 0.01$	Percent in samples
Baseline Algorithm	Within a week	11.96	11.80	53.45	88.42	97.10	59.87
	Between 1 and 2 weeks	9.43	5.84	35.81	81.35	97.03	58.27
	Between 2 weeks and a month	8.35	2.41	27.33	79.40	97.01	80.27
	Greater than a month	7.35	0.42	6.21	96.09	99.58	95.60
(k-NN) Refined Algorithm	Within a week	9.27	3.72	35.25	86.82	96.96	59.20
	Between 1 and 2 weeks	7.33	1.04	15.78	77.76	96.54	57.87
	Between 2 weeks and a month	6.66	0.58	9.33	73.94	96.50	80.07
	Greater than a month	5.99	0.07	1.26	82.19	99.72	95.47

Table 4: Sensitivity of Estimation of the Baseline Furfine Algorithm False Discovery Rate Within Fixed Sampling Windows of Repayment Days, Q1 2005 - Q4 2011.
With the exception of the last column, the top/bottom four rows decompose the summary statistics of the estimated false discovery rate for the Baseline/(k-NN) Refined Algorithm respectively (corresponding to Row 1 of Table 1/Table 3 in the paper) on the basis of the resampled days with repayment days falling within the specified fixed window of the principal loan day d_s. For instance, the false discovery rates summarized in Row 1 (in percentages) is computed on the basis of the subset of the resampled repayment days that fall within 7 days (business or otherwise) of the base comparison day, that is, $\widehat{FDR}^{FW}_{d_s} = \frac{\sum_{d_r} D^0_{d_s,d_r} \mathbf{1}(|d_s - d_r| \leq 7)}{\frac{\sum_{d_r} \mathbf{1}(|d_s-d_r|\leq 7)\mathbf{1}(D^0_{d_s,d_r}>0)}{\sum_{d_r} \mathbf{1}(|d_s-d_r|\leq 7)} D_{d_s,d'_s}}$. Rows 2/6 provides the same statistic for greater than 7 days, less than 14. Rows 3/7 is the window of $15 - 30$ days and Rows 4/8 is 30+ days. The last column provides the percent of the B^N samples for which that fixed window was populated (e.g. 60% of the computed $B^N = 1,500$ false discovery rate statistics included resampled dates that were within 7 calendar days of each other).

	Mean($\widehat{FDR}$)	Pr($\widehat{FDR} > x$)				B
		$x = 0.2$	$x = 0.1$	$x = 0.05$	$x = 0.01$	
Full Sample - 100	7.75	0.00	6.00	100.00	100.00	100
Full Sample - 200	8.27	1.00	13.00	100.00	100.00	200
Full Sample - 500	8.06	0.40	11.80	99.60	100.00	500
Full Sample - 1000	8.01	0.30	11.90	99.70	100.00	1000
Full Sample - 1500	8.02	0.33	11.33	99.53	100.00	1500

Table 5: Sensitivity of Full Sample False Discovery Rate to different B^N, Q1 2005 - Q4 2011.

References

Arjani, Neville and Darcey McVanel. 2006. "A Primer on Canada's Large Value Transfer System." Tech. rep., Bank of Canada.

Storey, John D. 2002. "A direct approach to false discovery rates." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (3):479–498.

———. 2003. "The positive false discovery rate: A Bayesian interpretation and the q-value." *Annals of Statistics* 31 (6):2013–2035.

Storey, John D and Robert Tibshirani. 2001. "Estimating the positive false discovery rate under dependence, with applications to DNA microarrays." URL `http://genomics.princeton.edu/storeylab/papers/dep.pdf`.

Tran, Lanh, Ba Chu, Chunfeng Huang, and Kim P. Huynh. 2014. "Adaptive permutation tests for serial independence." *Statistica Neerlandica* 68 (3):183–208. URL `http://ideas.repec.org/a/bla/stanee/v68y2014i3p183-208.html`.