Soft information imbalance is bad for fair credit allocation

Cal Muckley[1] Tian Tao[1] Xiangyang Zhao[2]

[1]University College Dublin
[2]Renmin University of China

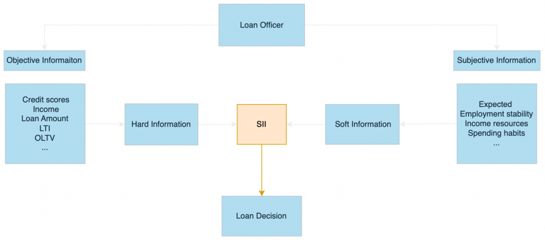
12th Workshop on Banking and Institutions

August 19, 2026

Motivation

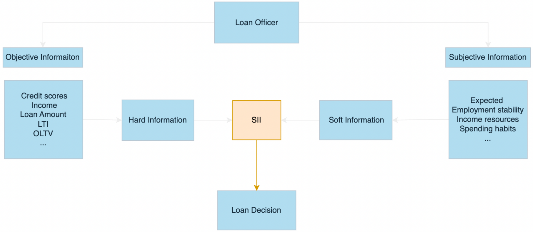
- **Unfair discrimination in lending**
  - Since the Fair Housing Act of 1968, there have been rafts of federal- and state-level fair lending legislation to prevent discrimination and ensure equal access to credit in the US.
  - Notwithstanding, minority (racial/ethnic) borrowers still face unfairly high denial rates compared to White applicants (Munnell et al., 1996; Lim et al 2023; Bhutta et al., 2025; Frame et al., 2025).
- **Consequences of unfair discrimination in lending**
  - Lower allocative efficiency, higher default risk, and increased financial instability (Vojtech et al., 2020).
  - Violates a civil right, e.g., the right to equal treatment (De Andres et al., 2021). Undermines social cohesion (Waters & Foster, 2021).

Motivation and Contribution: Loan decision schema extended for Soft Information Imbalance (SII)



- **Key reference:** Liberti & Petersen (2019) *Information: Hard and soft*, RCFS
- **Soft Information Imbalance (SII)** is the unequal utilization of soft information by loan officers across borrower groups: White (soft information) - Minority (soft information) > 0.

Motivation and Contribution: Loan decision schema extended for Soft Information Imbalance (SII)



Key References:

- Stiglitz and Weiss (1981), *Credit rationing in markets with imperfect information*, AER
- Wang (2020), *Screening soft information: Evidence from loan officers*, RAND J. Econ.
- Frame et al. (2025), *The impact of minority representation at mortgage lenders*, JF
- Liberti and Mian (2008) Estimating the effect of hierarchies on information use. RFS.
- Heider, Florian, Inderst, Roman. (2012). Loan prospecting. RFS.

Research Questions and Findings

Q1: Is there – across white and minority borrowers – SII in the home mortgage market?

- Loan decision models – R^2 differences. **Yes**
- Soft information utilization – R^2 differences; bank–county-year level. **Yes**
- Cross-sectional dispersion in contract terms; bank–county-year level. **Yes**

Q2: What are the loan-level consequences of SII for minority borrowers?

- Higher denial rates.
- Poorer loan performance.
- Loan pricing is invariant.

Q3: How can the adverse effects of SII be attenuated?

- Enhanced liquidity through internal capital reallocation; Loan officer diversity & training; loan applicant education.

Data Sources

- **Loan-level data (Home Mortgage Disclosure Act, HMDA, 2018–2023)**
  - Covers $\sim$90% of all U.S. mortgage originations.
  - Matched with Federal Deposit Insurance Corporation (FDIC) - Summary of Deposits (SOD) to identify local banks with physical branches.
  - Includes applicant demographics, income, loan amount, approval status.
- **Loan performance data (GSE)**
  - Publicly available loan-level information on mortgage performance, i.e., repayment and default.
  - Includes Debt-to-Income, Default, Loan-to-Value and Tract Credit Score

Descriptive Statistics

Table 1: Summary Statistics

Variable	Mean	SD	25th	Median	75th
Panel A: HMDA Sample					
Denial	0.2665	0.4421	0.0000	0.0000	0.0000
Minority Denial	0.3613	0.4804	0.0000	0.0000	1.0000
White Denial	0.2452	0.4302	0.0000	0.0000	0.0000
Interest Rate (%)	4.2833	1.6436	3.0000	3.8750	5.0000
SII	0.0294	0.1556	-0.0321	0.0272	0.0906
Minority	0.1834	0.3870	0.0000	0.0000	0.0000
Male	0.6562	0.4750	0.0000	1.0000	1.0000
Loan-to-Value Ratio (%)	67.1811	22.1189	53.3300	72.6000	80.0000
Ln(Loan Amount)	5.0633	0.9995	4.4427	5.1059	5.7203
Loan-to-Income Ratio	2.2624	3.5751	0.8523	1.7568	2.9271
FHA Loan	0.0228	0.1491	0.0000	0.0000	0.0000
VA Loan	0.0097	0.0980	0.0000	0.0000	0.0000
FSA/RHS Loan	0.0024	0.0492	0.0000	0.0000	0.0000
Home Improvement	0.1797	0.3840	0.0000	0.0000	0.0000
Refinance	0.3829	0.4861	0.0000	0.0000	1.0000
Home Purchase	0.3135	0.4639	0.0000	0.0000	1.0000
Tract Minority Share (%)	31.7619	25.1374	12.4300	23.6200	44.7800
Tract-to-MSA Income Ratio (%)	123.6652	49.1300	91.8200	116.0000	147.0000
Tract Median Housing Age	37.9497	18.6331	23.0000	37.0000	51.0000
Panel B: HMDA–GSE Sample					
Default	0.0143	0.1189	0.0000	0.0000	0.0000
Loan-to-Value Ratio (%)	64.7262	19.8359	51.0000	66.9230	80.0000
Tract Credit Score	761.3121	76.4006	734.0000	772.0000	796.0000
Debt-to-Income Ratio (%)	33.8449	18.6773	26.0000	34.0000	42.0000

Research Question 1

Is there – across white and minority borrowers – SII in the home mortgage market?

Q1: Is there – across white and minority borrowers – soft information imbalance in the home mortgage market?

Table 2: Soft Information: Loan denial specification R^2 for minority borrowers and White borrowers.

Specification		Adjusted R^2		
Controls	Bank × County × Year FE	Minority sample	White sample	Minority sample – White sample
Spec (1): Borrower Demographics	Yes	0.1618	0.1146	0.0472***
Spec (2): Spec (1) + Financial Controls	Yes	0.4342	0.3737	0.0605***
Spec (3): Spec (2) + Loan Type	Yes	0.4344	0.3739	0.0605***
Spec (4): Spec (3) + Loan Purpose	Yes	0.4394	0.3761	0.0633***
Spec (5): Spec (4) + Tract Controls	Yes	0.4402	0.3765	0.0637***
Spec (6): Spec (4) + Tract Credit Score	Yes	0.4635	0.3846	0.0789***

Note: Follows Buchak *et al.* (2018), *JFE*.

$$Denial_i = \beta X_i + \delta_{bct} + \varepsilon_i$$

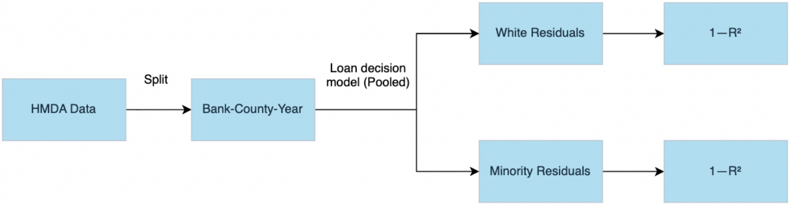
Q1: Empirical Specification

At the bank county year level, to examine how soft information is utilized by loan officers across minority and non-minority borrowers, we follow Haggendorf et al. (2022) to estimate an ordinary least squares (OLS) regression model:

$$(1 - R^2)_{bct} = \beta \cdot \text{Minority}_{bct} + \delta_{bct} + \varepsilon_{bct} \tag{1}$$

- Dependent variable: $(1 - R^2)_{bct}$, proxy for the degree of soft information usage.
- $\text{Minority}_{bct} = 1$ if borrower is racial/ethnic minority, 0 if White.
- δ_{bct}: bank–county–year fixed effects.
- ε_{bct}: error term.

Q1: Measurement of Soft Information



Q1: Is there – across white and minority borrowers – soft information imbalance in the home mortgage market?

Table 3: Soft Information Usage for Minority Borrowers

	Dependent Variable: Soft Information Usage (1-R^2)		
	1- adjust R^2	1- quasi-R^2	1- quasi-R^2 (weighted)
	(1)	(2)	(3)
Minority	-0.0399***	-0.0429***	-0.0288***
	(0.0057)	(0.0035)	(0.0041)
Bank × County × Year FE	Yes	Yes	Yes
Observations	75,960	75,960	75,960
Adjusted R^2	0.2263	0.3993	0.4831

Q1: Is there – across white and minority borrowers – soft information imbalance in the home mortgage market?

- **Conceptual Framework:**
  - Skrastins & Vig (2019, RFS): The **loss of information** reduces the dispersion of lending decisions, leading to more standardized loan contracts.
- **Empirical Implementation:**
  - At the bank county year level, we measure information using two dispersion-based indicators of loan quantities:
    - Interquartile Range (IQR) of Debt.
    - Standard Deviation of Debt.
  - Higher values of these metrics indicate greater informational content in lending.
- **Research Hypothesis:**
  - The loan contracts of **minority borrowers** exhibit **lower lending dispersion** compared to those of **White borrowers**.

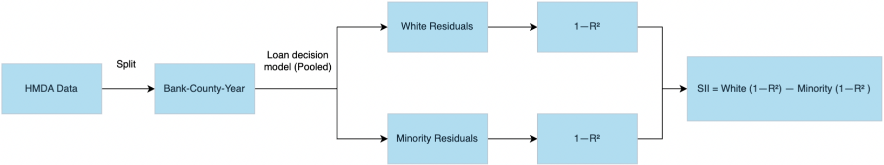
Q1: Is there – across white and minority borrowers – soft information imbalance in the home mortgage market?

Table 4: Soft Information: Minority Borrowers and Loan Contract Dispersion

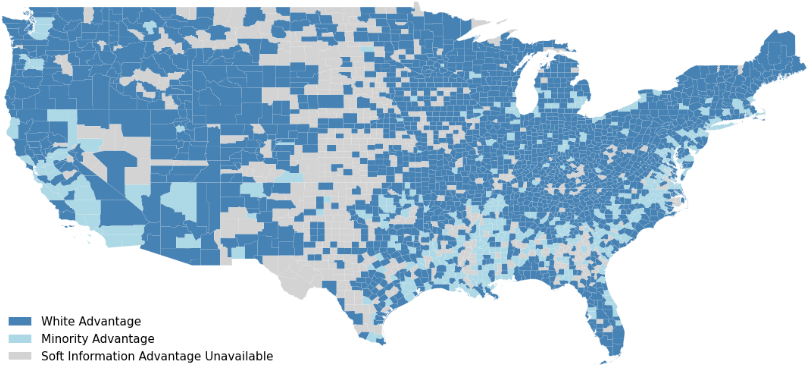
	Dependent Variable: Loan Contract Dispersion	
	$\ln(\sigma_{\text{loan}_{b,c,t}})$	$\ln(\text{IQR}_{b,c,t})$
	(1)	(2)
Minority	-0.0488***	-0.2107***
	(0.0030)	(0.0135)
Bank × County × Year FE	Yes	Yes
Adjusted R^2	0.2734	0.2167
Observations	139,469	153,206

What are the loan-level consequences of SII for minority borrowers?

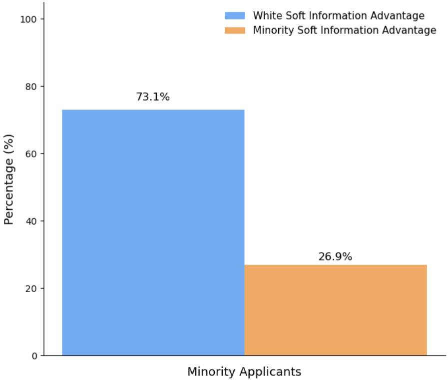
Q1: Measurement of Soft Information Imbalance (SII)



White Soft Information Advantage in 73% of Counties



Where do minority applicants apply for loans? 73.1 % at White soft information advantaged bank branches



We focus in this paper on White soft information advantaged bank branches (full sample = robustness test)

Q2: Denial and Soft Information Imbalance

To test, in White Information Advantaged bank branches, whether Minority applicants are more likely to be denied in bank–county branches where soft information usage is more imbalanced, we estimate the following model:

$$\text{Denial}_{ibct} = \alpha \cdot \text{Minority}_{ibct} + \beta \cdot (\text{Minority}_{ibct} \times \text{SII}_{bct}) + \mathbf{X}_{it}\gamma + \delta_{bct} + \varepsilon_{ibct} \quad (2)$$

- Denial_{ibct}: indicator $= 1$ if loan application i is denied by bank b in county c at year t
- Minority_{ibct}: borrower race/ethnicity indicator.
- SII_{bct}: soft information imbalance at bank–county–year level.
- $\mathbf{X}_{it}$: loan-level controls (income, loan/income, gender, loan type, etc.).
- δ_{bct}: bank–county–year fixed effects.

Table 5: Soft Information Imbalance and Loan-Level Mortgage Decisions

	Dependent Variable: Denial				
	(1)	(2)	(3)	(4)	(5)
Minority × SII	0.1566***	0.1501***	0.0581***	0.0597***	0.0585***
	(0.0179)	(0.0167)	(0.0108)	(0.0106)	(0.0108)
Minority	0.0874***	0.0894***	0.0457***	0.0460***	0.0473***
	(0.0075)	(0.0069)	(0.0049)	(0.0049)	(0.0050)
Loan-to-Income Ratio			0.0045***	0.0045***	0.0047***
			(0.0006)	(0.0007)	(0.0006)
Log Loan Amount			-0.1023***	-0.1020***	-0.0918***
			(0.0113)	(0.0113)	(0.0086)
Loan-to-Value Ratio			0.1282***	0.1317***	0.1336***
			(0.0197)	(0.0190)	(0.0140)
FHA Loan				-0.0155	-0.0074
				(0.0113)	(0.0117)
VA Loan				-0.0640***	-0.0596***
				(0.0172)	(0.0163)
FSA/RHS Loan				0.0131	0.0192*
				(0.0108)	(0.0111)
Home Improvement					0.0760***
					(0.0119)
Refinance					0.0109
					(0.0129)
Male Dummy	No	Yes	Yes	Yes	Yes
Age Bin Dummy	No	Yes	Yes	Yes	Yes
DTI Ratio Dummy	No	No	Yes	Yes	Yes
Bank × County × Year FE	Yes	Yes	Yes	Yes	Yes
Observations	7,399,123	7,399,123	7,399,123	7,399,123	7,399,123
Adjusted R^2	0.1079	0.1123	0.3915	0.3917	0.3948

Based on the model in Frame et al. (2025) *JF*

Q2: Instrumental Variable Strategy

Our identification strategy exploits geographical and temporal variation in sunshine exposure (Cort´es et al. 2016 JFE) to generate plausibly exogenous variation in loan officers' interpretation of soft information.

IV approach:

- 2SLS using proportion of **perfectly sunny weekdays** as exogenous instrument.
- Sunshine $\Rightarrow$ loan officers' mood $\Rightarrow$ interpretation of soft information.
- Instrument relevance test requires co-variation within b*c*t between SunnyDays and SII when minority status varies.

Data construction:

- NOAA ISD: ~14,000 active weather stations (24h/day*365 days/year).
- County-year measure = share of weekdays with 0 cloud cover (8am–4pm); Exclude counties >50 miles from nearest station.

Q2: Instrumenting for Soft Information Imbalance with Variation in Local Sunshine

Table 6: Instrumenting for Soft Information Imbalance with Variation in Local Sunshine

	First Stage	Second Stage
	$\widehat{\text{Minority} \times \text{SII}}$	Denial
	(1)	(2)
Minority × SunnyDays	0.4359***	
	(0.0352)	
$\widehat{\text{Minority} \times \text{SII}}$		0.0675*
		(0.0391)
Control Variables in Table 5 Column (5)	Yes	Yes
Bank × County × Year FE	Yes	Yes
Observations	6,295,779	6,295,779
Adjusted R^2	0.6033	0.3298
First-stage F-statistic	153.02	

Q2: Data Privacy Legislation as a Shock to Soft Information Usage

Data Privacy Legislation (DPL) in 2018

- Enhances data security, integrity and transparency: informs consumers about how data are collected, used, and shared.

Impact on Information Production

- Gupta et al. (2023) show that DPL increases compliance costs for using soft information.
- Additional cost is expected to fall disproportionately on the processing of white borrowers' soft information (there's more of it)

Hypothesis

- DPL reduces soft information imbalance, leading to lower relative denial rates for minority borrowers.

Empirical Design: Difference-in-Differences

Following Hagendorff et al. (2022), we propose two specifications:

DPL and Soft Information Utilisation:

$$(1 - R^2)_{bct} = \alpha \cdot \text{Minority}_{bct} + \beta \cdot (\text{Minority}_{bct} \times \text{Post}_{st}) + \delta_{bct} + \varepsilon_{bct} \quad (3)$$

DPL and Minority Denial Rates:

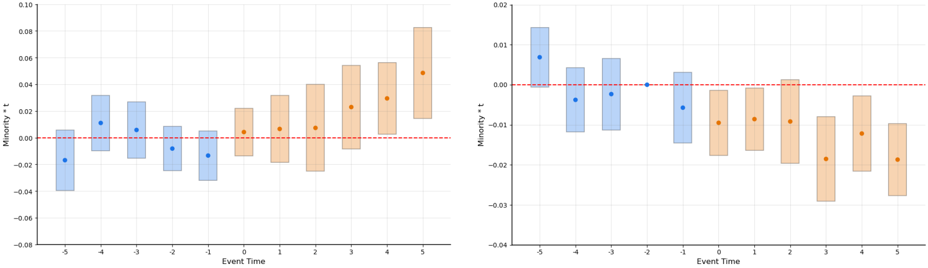
$$\text{Denial}_{ibct} = \alpha \cdot \text{Minority}_{it} + \beta \cdot (\text{Minority}_{it} \times \text{Post}_{st}) + \mathbf{X}'_{it}\gamma + \delta_{bct} + \varepsilon_{ibct} \quad (4)$$

Q2: Data Privacy Legislation, Soft Information Imbalance and Minority Denial Rates

Table 7: Difference in Differences: The Effect of Data Privacy Legislation on Soft Information Imbalance and Minority Denials

	Soft Information Usage (1-quasiR^2)		Denial
	Unweighted	Weighted	
	(1)	(2)	(3)
Minority × Post	0.0161**	0.0155**	-0.0279***
	(0.0063)	(0.0077)	(0.0038)
Minority	-0.0457***	-0.0327***	0.0501***
	(0.0062)	(0.0054)	(0.0040)
Control Variables in Table 5 Column (5)	No	No	Yes
Bank × County × Year FE	Yes	Yes	Yes
Observations	18,982	18,982	2,778,155
Adjusted R^2	0.4442	0.5261	0.4136

Q2: Parallel Trends



(a) Soft Information Usage **(b) Minority Denial**

Q2: Additional Robustness Tests

- Control for pseudo loan officer fixed effects, e.g., ethnicity/ race.
- Using all bank branches not just those with a white soft information advantage.
- Soft information affect on loan denial for minorities.
- High- versus Low-Hierarchy Banks.
- Demand composition robustness: Apply Propensity Score 1:1 Matching (PSM).
- Alternative clusterings of standard errors.
- Winsorize outliers in soft information estimates at the 1% and 3% levels.
- A minimum threshold ($>$50 at bank-county-year level) of minority applications.
- Stability over time - Exclude Covid-19.

Additional analyses on loan outcomes: loan performance & interest rates

Soft Information Imbalance and Mortgage Default

- **Empirical Background:**
  - Under taste-based discrimination (Becker, 1957; Bhutta et al 2025), loan officers apply stricter lending standards to minority applicants, resulting in lower default rates.
  - Under statistical based discrimination (Phelps, 1972), loan officers impute mean minority creditworthiness in lending to minority borrowers, ultimately this gives higher default rates.
    - Eg. failure to use soft information on minority loan applicants.
- **Research Hypothesis:**
  - Minority borrowers are more likely to default at bank branches where **soft information is more imbalanced**.

Soft Information Imbalance and Mortgage Default

Table 10: Soft Information Imbalance and Loan-Level Mortgage Performance

	Dependent Variable: Default (1 Month)				3 Months
	(1)	(2)	(3)	(4)	(5)
Minority × SII	0.0423**	0.0426**	0.0423**	0.0426**	0.0510*
	(0.0124)	(0.0124)	(0.0124)	(0.0124)	(0.0263)
Minority	-0.0020	-0.0029**	-0.0020	-0.0028**	-0.0047***
	(0.0013)	(0.0012)	(0.0013)	(0.0012)	(0.0013)
Loan-to-Income Ratio		0.0002		0.0002	0.0019***
		(0.0004)		(0.0004)	(0.0005)
Log Loan Amount		-0.0003		-0.0003	-0.0035**
		(0.0012)		(0.0012)	(0.0014)
Loan-to-Value Ratio		0.00003		0.00003	0.0003**
		(0.00004)		(0.00004)	(0.0001)
Credit Score		-0.0002***		-0.0002***	-0.0004***
		(0.00002)		(0.00002)	(0.00007)
Debt-to-Income Ratio		0.0002***		0.0002***	0.0007***
		(0.00004)		(0.00004)	(0.0001)
Refinance			0.0001	0.0005	0.0008
			(0.0002)	(0.0003)	(0.0022)
Bank × County × Year FE	Yes	Yes	Yes	Yes	Yes
Observations	55,773	55,758	55,773	55,758	55,820
Adjusted R^2	0.0607	0.0654	0.0607	0.0654	0.1112

Soft Information Imbalance and Loan Pricing

- **Background:**
  - Cortés et al. (2016, JFE): Loan pricing is primarily determined by computerized bank algorithms that rely on hard information, with relatively limited input from loan officers.
  - Liberti & Petersen (2019, RCFS): Soft information is difficult to convert into hard information without any loss of information.
- **Research Hypothesis:**
  - Soft information imbalance **does not significantly affect loan pricing** for minority borrowers.

Soft Information Imbalance and Loan Pricing

Table 8: Soft Information Imbalance and Loan Pricing

	Dependent Variable: Interest Rate
	(1)
Minority × SII	-0.0684
	(0.1003)
Minority	-0.0471***
	(0.0082)
Control Variables in Table 5 Column (5)	Yes
Bank × County × Year FE	Yes
Observations	2,139,369
Adjusted R^2	0.4870

How can the adverse effects of SII be attenuated?

Q3: The Role of Liquidity in Soft Information Imbalance

- **Empirical Background:**
  - Dewally (2014, JBF): During the financial crisis, liquidity contraction reduced credit supply, while borrowers with stronger relationship-based soft information obtained more loans.
  - Álvarez-Román (2024): Under climate-induced liquidity shocks, local banks lend more to borrowers with soft information advantage

- **Research Hypothesis:**
  - When **bank branch liquidity** tightens, banks rely more on soft information. Minority borrowers, **due to soft information disadvantage**, are **more adversely affected** by local liquidity shortages.

Q3: Mitigating the Adverse Effects of Soft Information Imbalance

Table 10: Heterogeneity Analysis

	Dependent Variable: Denial		
	(1)	(2)	(3)
Minority × SII × Liquidity	-0.0100*		
	(0.0058)		
Minority × SII × High Minority Share		0.0215	
		(0.0231)	
Minority × SII × Low Income			0.1399***
			(0.0164)
Minority × Liquidity	0.0040		
	(0.0025)		
Minority × High Minority Share		-0.0274***	
		(0.0051)	
Minority × Low Income			-0.0019
			(0.0056)
Minority × SII	0.0596***	0.0476***	0.0339***
	(0.0110)	(0.0120)	(0.0107)
Minority	0.0469***	0.0529***	0.0476***
	(0.0051)	(0.0053)	(0.0049)
Control Variables in Table 5 Column (5)	Yes	Yes	Yes
Bank × County × Year FE	Yes	Yes	Yes
Observations	7,380,413	7,399,123	7,399,123
Adjusted R^2	0.3950	0.3949	0.3950

Conclusion

1. We address the question: why do minority borrowers continue to face discrimination in the home mortgage market?
2. We show that soft information imbalance leads to unfair credit allocation and worse loan performance, but that loan pricing is invariant.
3. We identify a liquidity channel which can mitigate the adverse effects of soft information imbalance.
   - Additional channels may include (i) increased recruitment of minority loan officers, (ii) increased training of loan officers and (iii) education of minority loan applicants.

Thank You for Your Attention!

All feedback is most welcome :-)