UNIVERSITY OF CAMBRIDGE

2019-05-17

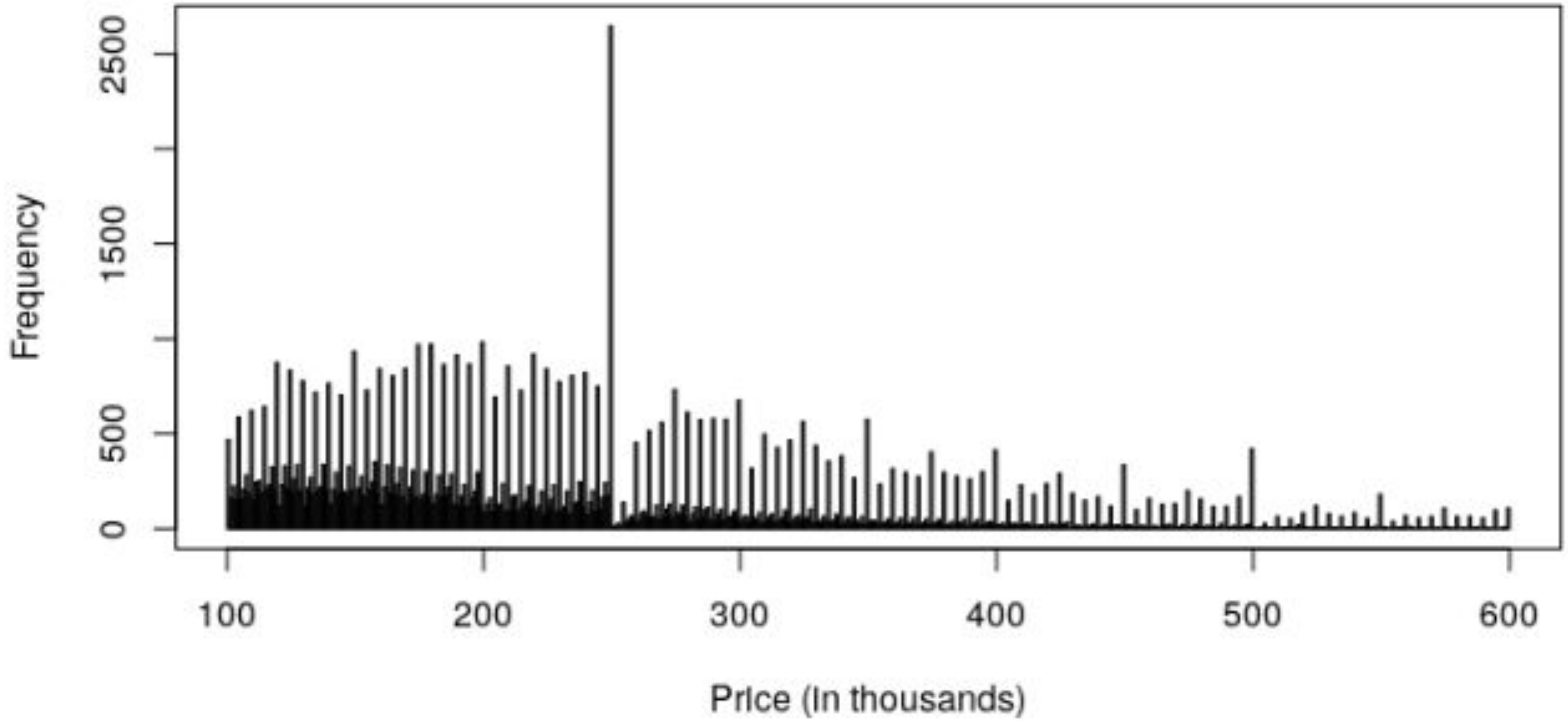
The Odd One(s) Out

Thies Lindenthal & Carolin Schmidt

htl24@cam.ac.uk
carolin.schmidt@zew.de

Two core observations motivate our paper

First: Observed transaction prices follow a "ruler distribution"



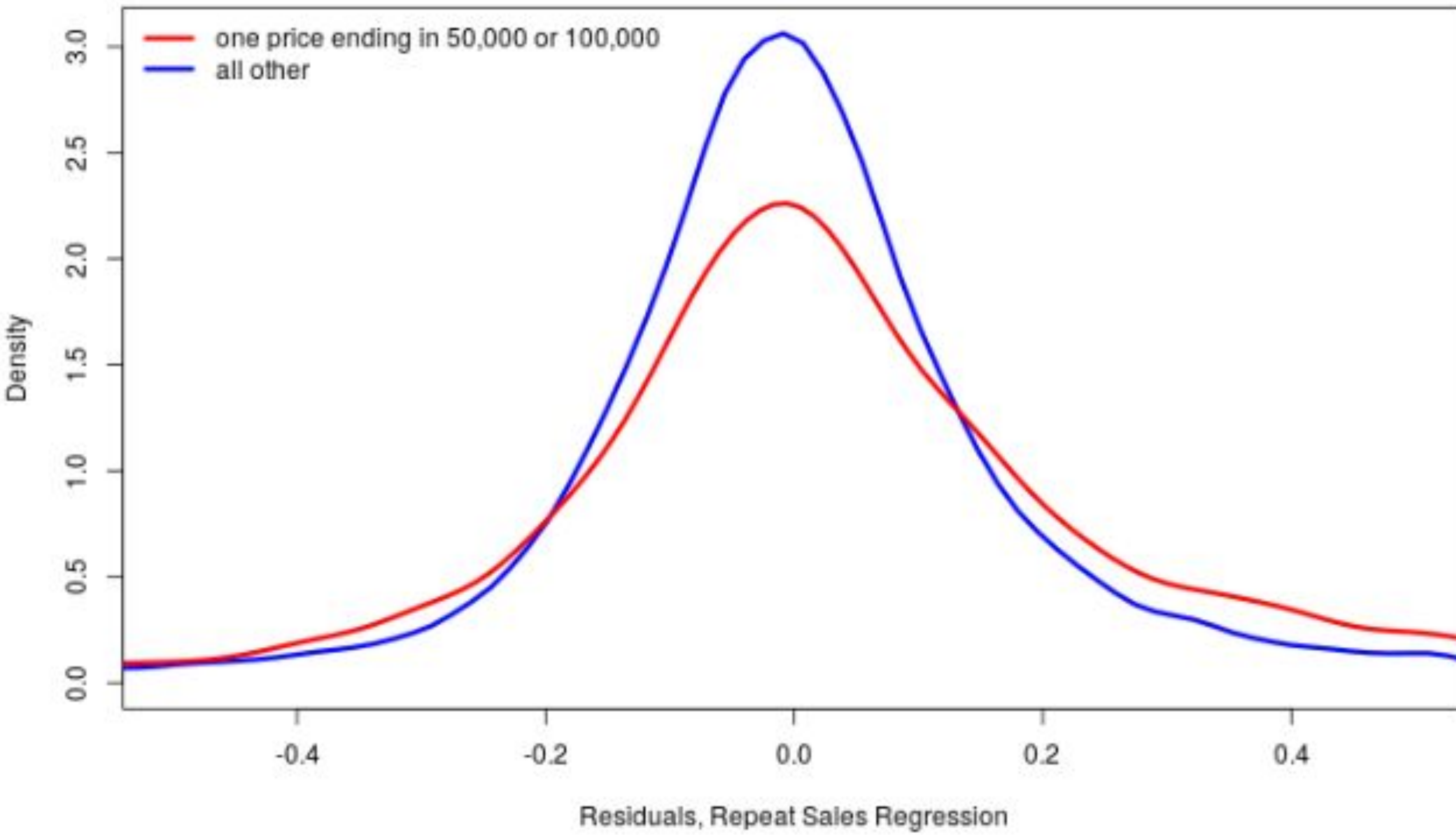
- Disproportionally high share for round prices (e.g. multiples of 10K/25K/50K)
- Prices get coarser with price levels (Ball et al., 1985; Thomas et al., 2010)

UNIVERSITY OF CAMBRIDGE

Second: Round prices are less precise...

… larger deviations from fundamentals.

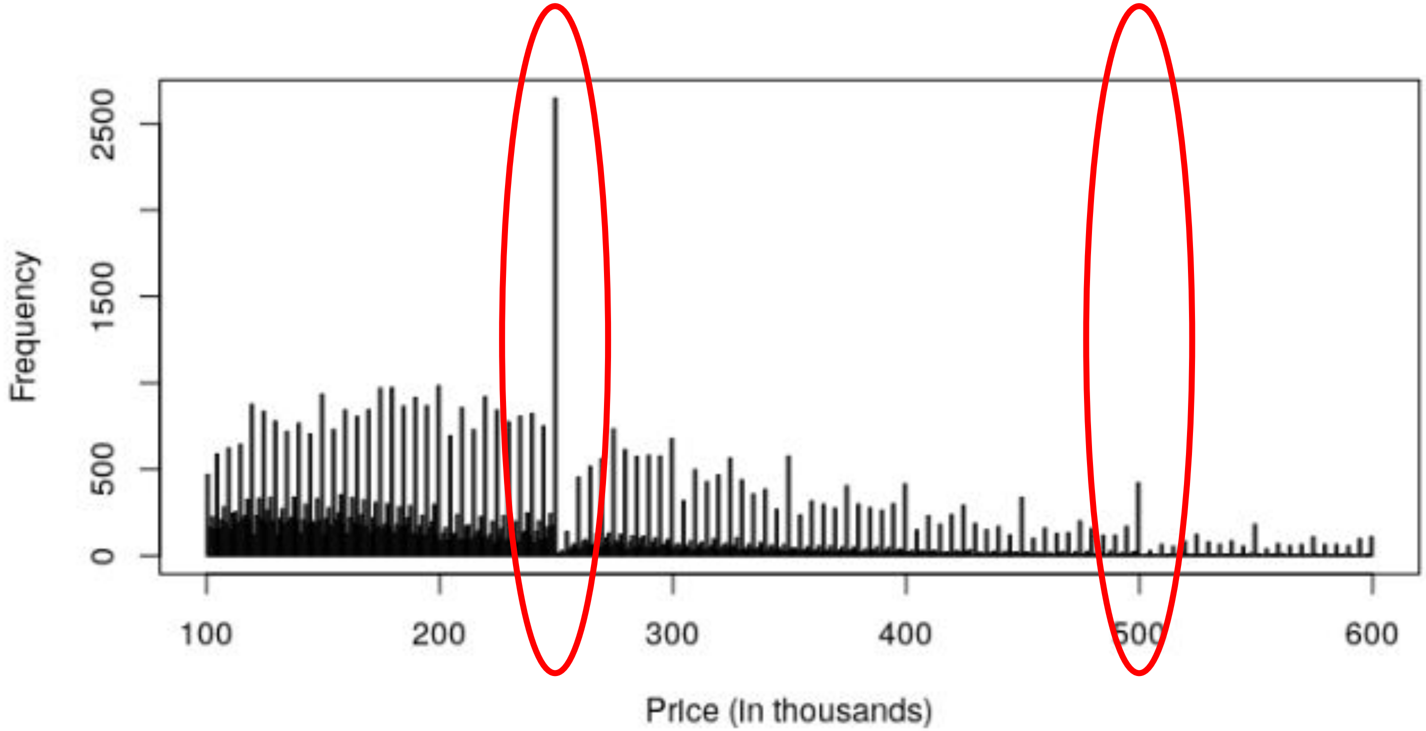
Distribution of residuals from repeat sales regression



This effect is robust, not driven by tax subsidy thresholds

These UK-specific price regions have been excluded from analysis

- **250K "Help to buy" scheme**
- **500K first time buyer stamp duty discount (since 2018)**



UNIVERSITY OF CAMBRIDGE

Negotiated sales prices, not asking prices

British way(s) of trading houses

- **Guide price set by seller**
- **Potential purchasers hand in sealed bids**
- **The seller is not bound to accept the highest offer,**
  - She can pick any (or none)
  - Estate agents often facilitate the trade and, often, strategically release information
- **In England and Wales, the terms of an offer remain subject to contract**
  - No-one is legally obliged to continue with the transaction until the formal contract has been signed and the parties have exchanged the contracts
  - Transactions take months
  - Both sides have to assess the risk of transaction falling through
- **In Scotland, transactions are binding earlier in the process**

What is so special about round prices?
Why do we care?

- **Direct applications**
    - Reliability of comparables when valuing individual buildings
    - Mass appraisal systems
    - Signalling in negotiations
    - Design decisions when developing

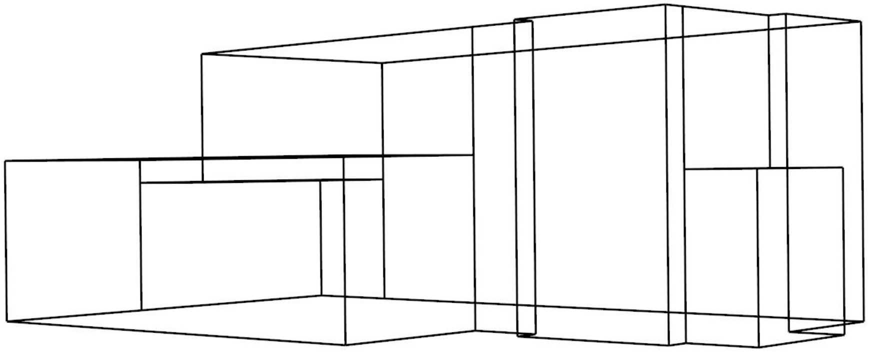
- **Round prices offer insights on human decision making**
    - When are we confident deciders? In which cases is it difficult to make a judgement?

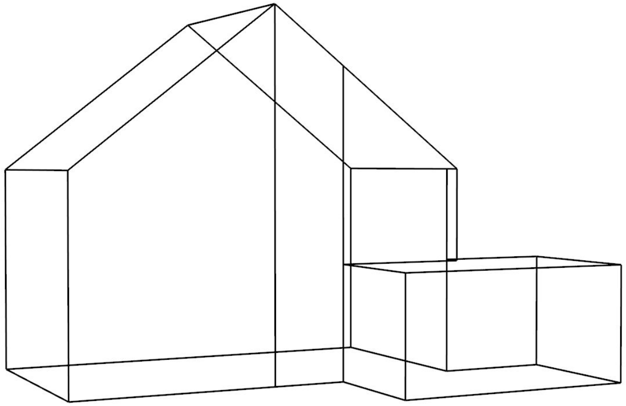
UNIVERSITY OF CAMBRIDGE

Value of aesthetics / architecture / preferences / beauty

Paper is part of a larger research theme on "human" side of property

- **"Beauty in the Eye of the Home-Owner: Aesthetic Zoning and Residential Property Values" (REE, 2017)**





Value = f(X, Shape, Shape neighbours,...)

“Machine Learning, Building Vintage and Property Values”
(Lindenthal, Johnson)

What is our contribution to the literature?

P(round price | sale) = f(buyer and seller factors, market factors, asset factors)

- **There is rich theoretical & empirical research on round vs. precise numbers**
  - general psychology, retail, negotiations, sport
- **Most research focuses on the psychological aspect of these numbers**
  - In real estate, round transaction prices investigated by Palmon et al. (2004) and Beracha and Seiler (2013)
- **We show that market conditions and asset characteristics influence the salience of these mental traits**
  - Heterogeneity of real estate creates variation in the likelihood of observing a round price
  - Some buildings are easier to value than others - the price reveals the relative difficulty
  - Reliability of observed prices is dynamic

UNIVERSITY OF CAMBRIDGE

Buyers and sellers

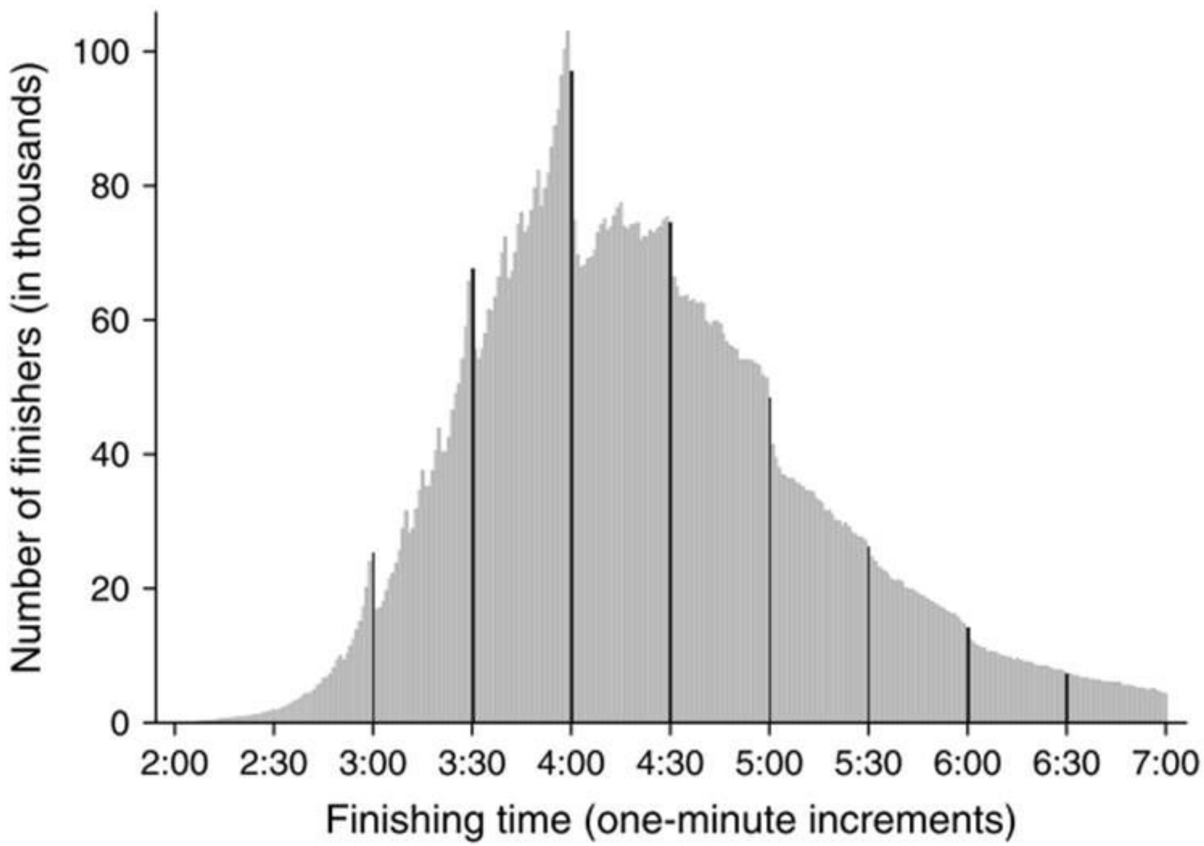
Amateurs and experts alike use mental shortcuts, consciously or unconsciously

- **Heaping (uncertainty or inability)**
    - Age (A'Hearn et al., 2009): dyscalculics tend to report their age as a multiple of five or other attractive numbers
    - Analyst forecasts (Herrmann and Thomas, 2005)
- **Conformist behaviour**
    - ½-carat diamonds sell at an 18% premium relative to diamonds slightly < ½ carat (Scott and Yelowitz, 2010)
- **Institutional rules, familiarity, efficiency**
    - Stock and commodity prices end in even/round numbers even though finer pricing is permitted (Osborne, 1962; Niederhoffer, 1966; Ball et al., 1985)
    - Simplification of financial record processing (Stevenson and Bear, 1970)
    - Round prices are easier to process (Tversky and Kahneman, 1973)
- **Digit preference**
    - Best drug dosage? Such that 20 drops of cough syrup three times a day are effective in ~90% of the cases (Herxheimer, 1991)

Sellers (and buyers?) want to achieve thresholds

Similar to motivation of Marathon runners (Allen et al., 2017)

- **For a seller, it is gratifying to exceed a mental mark, willing to push just a bit harder “on the last mile”**
- **What about buyers? Shouldn’t they have the opposite motivation?**
  - Buyers vs sellers markets?



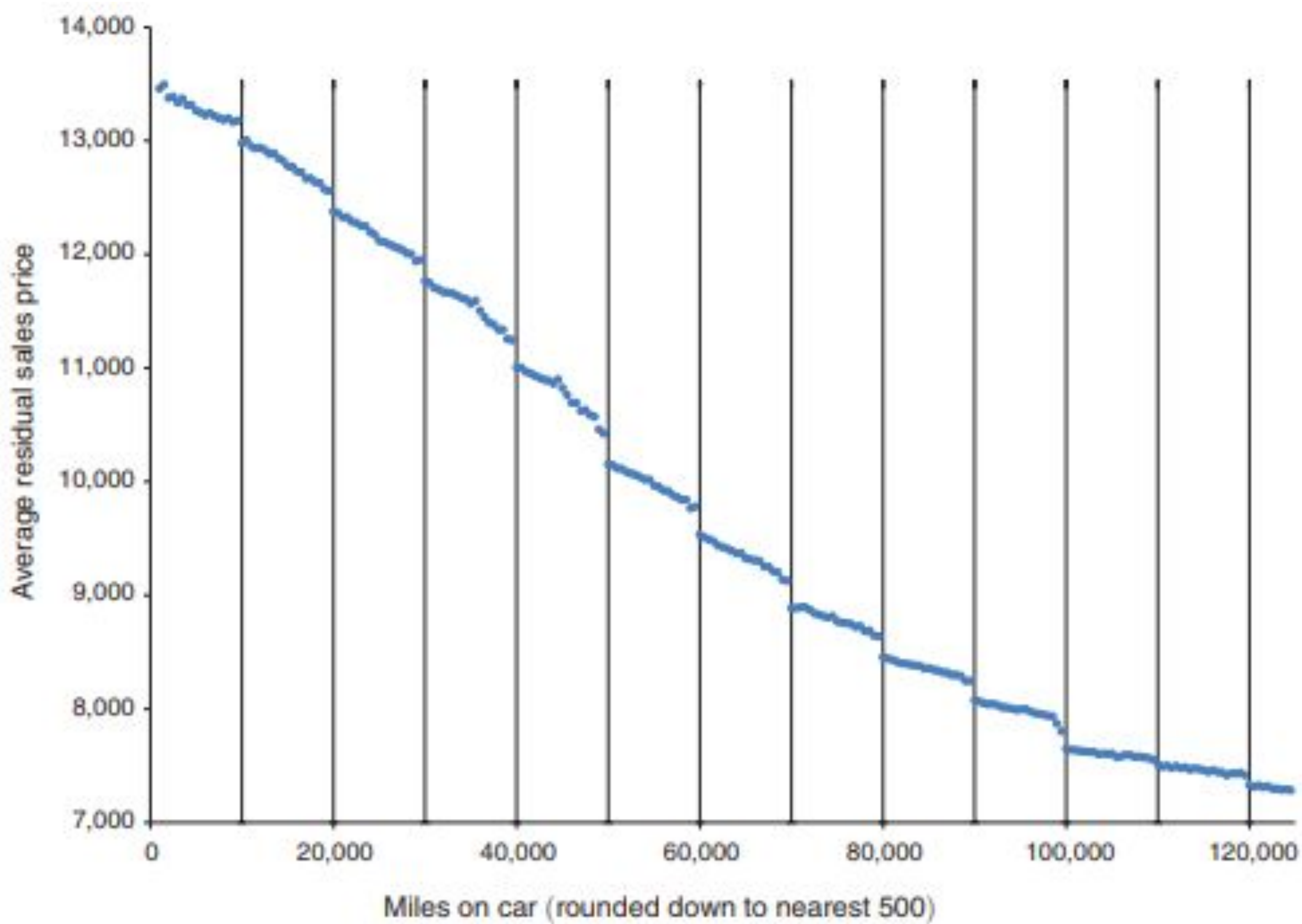
Some use coarse prices strategically

- **Cheap-talk model (Backus et al., 2015)**
    - Sellers advertising at round prices signal their willingness to negotiate (lower TOM)
    - Precise-price advertisers achieve higher final sales prices on average (higher TOM)
- **Negotiation efficiency hypothesis (Harris, 1991)**
- **Round list prices speed up transactions**
- **Palmon et al. (2004) on clustering in real estate prices**
    - List prices more often just-below-even ending, transaction prices more even-ending
    - NEH predicts even-ending transaction prices, especially when information on the property is scarce & costly to obtain
- **Psychological literature: coarse numbers signal uncertainty, precise numbers confidence and knowledge**
    - “One year” versus “365 days”
    - Yaniv and Foster (1995), Goldsmith et al. (2002), Zhang and Schwarz (2013), Mason et al. (2013)

A (hypothetical) blue book for homes...

… could explain discontinuities in price distributions

- **Left digit bias, due to limited information-processing ability (Lacetera et al., 2012)**



Uncertainty about marginal prices for attributes

High uncertainty for hedonic coefficients due to low number of comparables?

- **Liquidity (information on market)**
  - Low # comparables implies high quality uncertainty (Martel, 2018)
  - Everything else equal, lower liquidity and fewer comparables should lead to more round prices being observed.
- **Quality uncertainty (information on building)**
  - Prices cluster when asset values are uncertain (Ball et al., 1985; Binder, 2017)
    - Listings are notoriously vague. Square footage? Damp? Noise? Sitting tenants?
  - Similarly: firm valuation is more subjective and variable for young firms with a short earnings history (Baker and Wurgler, 2006)
- **Asset uniqueness (combination of both)**
  - How to quantify and value uncommon specifications?
  - Extreme values or interaction terms reduce # relevant comparables, driving up uncertainty
    - Value of detached house with garage in Romsey Town?

Uncommon combination (style and location)

Uncommon attributes

Empirical strategy
Can we predict the occurrence of round prices?

- **We cannot observe buyers' and sellers' characteristics or motivation**
  - Omit (for now)
- **Information on market liquidity from universe of sales (land registry)**
  - Sales are geocoded and have time stamps
  - Number of comparables in previous x months within y miles from each building
- **Information on asset uniqueness from limited set of hedonics and computer vision**
  - Derive additional variables from images
  - Model asset uniqueness directly

More comparables reduce odds of round prices

More information on the local market makes it easier to value a property

- **Probit regression on round price**
  - Controlling for location at postcode and streetlevel
  - Year
  - Hedonics: size/volume, new, vintage
  - Price band (50K buckets)
- **Price is defined as being “round” if it is a multiple of £25K**
  - Is 275,000 more round than 280,000?
- **Core variable of interest: # comparables**
  - Number of sales in same postcode in preceding 12 months as number of comps.

UNIVERSITY OF CAMBRIDGE

Probit estimates
Cambridge submarket

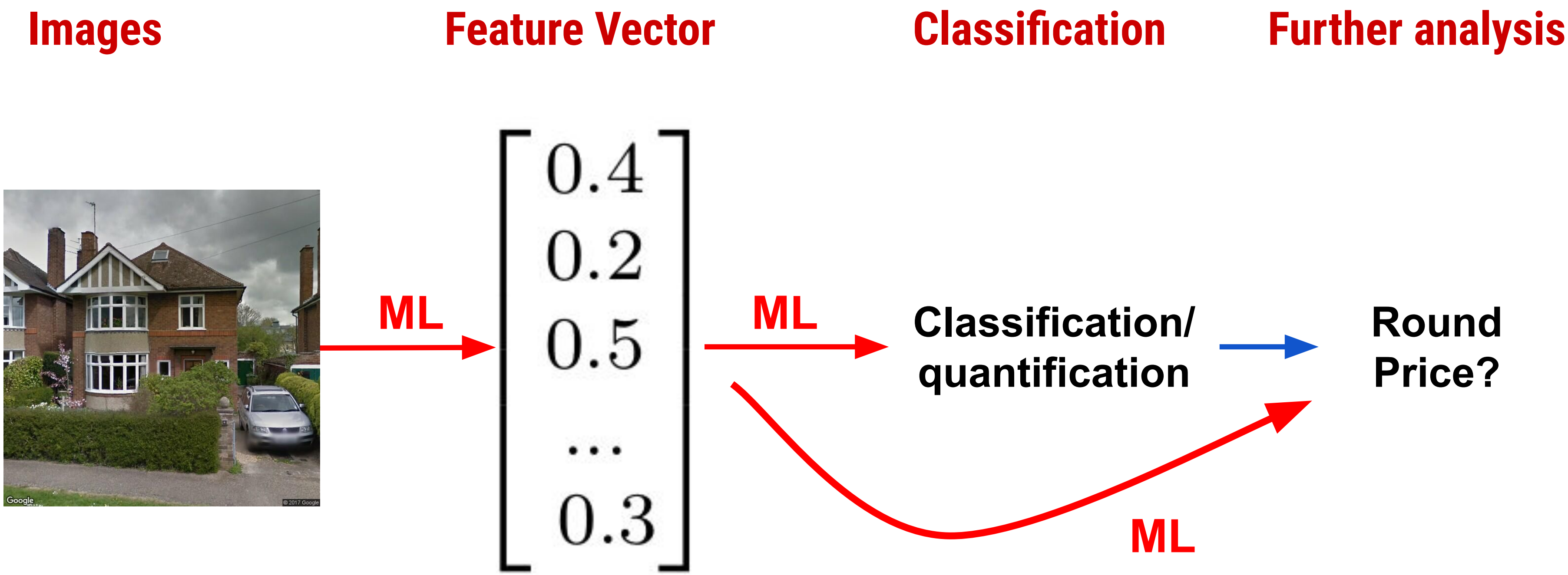
- **Control variables**
  - Hedonics & Vintage
  - Location
  - Year
  - Price band
- **Expected sign for # comps!**

	Dependent variable:
	round price
new-built	0.014
	(0.070)
ln(area, m^2)	0.133***
	(0.039)
ln(volume, m^3)	−0.004
	(0.005)
ln(num. comps)	−0.032**
	(0.015)
Constant	−1.389***
	(0.355)
Observations	24,035
Log Likelihood	−7,506.127
Akaike Inf. Crit.	15,222.250
Note:	*p<0.1; **p<0.05; ***p<0.01

UNIVERSITY OF CAMBRIDGE

From computer vision to economic analysis

Deriving additional variables / model uniqueness directly

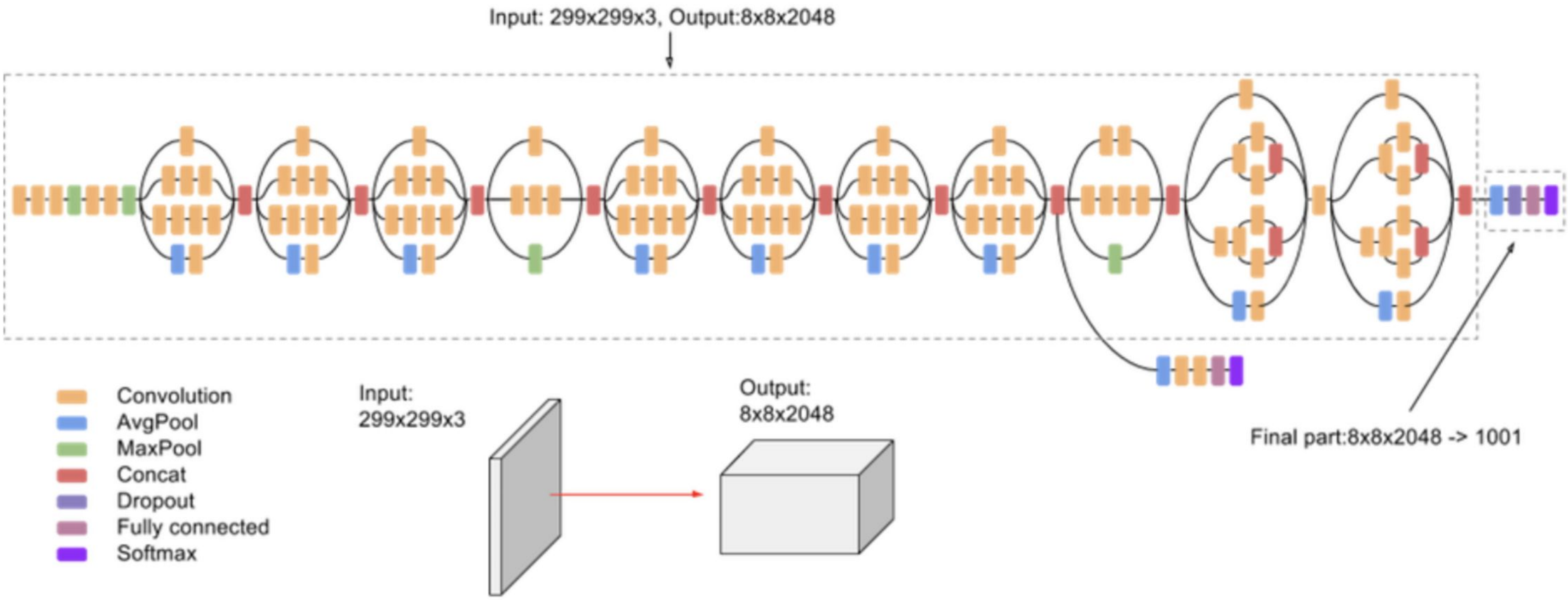


UNIVERSITY OF
CAMBRIDGE

Computer vision, off the shelf

Deep convolutional neural network to obtain feature vectors

- **Pre-trained *Inception v3* model in Tensorflow API**
  - Convolutional: Exceptionally suitable to detect era specific details such as window styles, ratios, brickwork, ratios
  - Freely available & frequently used
- **Penultimate layer is 2048 dimensional feature vector**



DNN Design

- **Testing many specifications**
- **Compromise across geographic scope and # variables**

Layer	*Type*	*Output Shape*	*Activation*	*# Parameters*
		UK: Base		
1	Dense	None, 37	ReLu	2,331
2	Dense	None, 18	ReLu	684
3	Dense	None, 1	Sigmoid	19
total				3,034
		Cambridge: Base		
1	Dense	None, 10	ReLu	280
2	Dropout	None, 10		0
3	Dense	None, 1	Sigmoid	11
total				291
		Cambridge: Vint.		
1	Dense	None, 18	ReLu	828
2	Dropout	None, 18		0
3	Dense	None, 1	Sigmoid	19
total				847
		Cambridge: Vint. & Inception		
1	Dense	None, 837	ReLu	1,752,678
2	Dropout	None, 837		0
3	Dense	None, 1	Sigmoid	838
total				1,753,516

Notes: The training data for Cambridge are matrices of size $4,400 \times 27$ (Base) $4,400 \times 45$ (Vint.) $4,400 \times 2093$ (Vint. & Inception).

UNIVERSITY OF CAMBRIDGE

Training on balanced training set
First for the UK (100K sample)

- **For the UK, we have basic hedonics only - but # comps!**
- **Same number of round/non-round sales in training**
  - Out of samples test realistic (using unseen data, ~11% round)

	observed: not round	*observed: round*
Unbalanced out-of-sample test, UK		
prediction: not round	72,764	3,292
prediction: round	20,452	3,492
recall	0.78	0.51
precision	0.96	0.15
F_1-score	0.86	0.23

- F_1-score: 2 (recall * precision) / (recall + precision)

UNIVERSITY OF CAMBRIDGE

Zooming in on Cambridge

Basic hedonic variables (area/volume) don't boost predictive power much

- **Precision for "round" improves, recall does not**

	observed: not round	*observed: round*
Unbalanced out-of-sample test, UK		
prediction: not round	72,764	3,292
prediction: round	20,452	3,492
recall	0.78	0.51
precision	0.96	0.15
F_1-score	0.86	0.23
Unbalanced out-of-sample test, Cambridge, base		
prediction: not round	3,691	334
prediction: round	748	247
recall	0.83	0.43
precision	0.92	0.25
F_1-score.2	0.87	0.31

Add more information derived from images

Can we spot the odd ones out?

- **Base line**

+

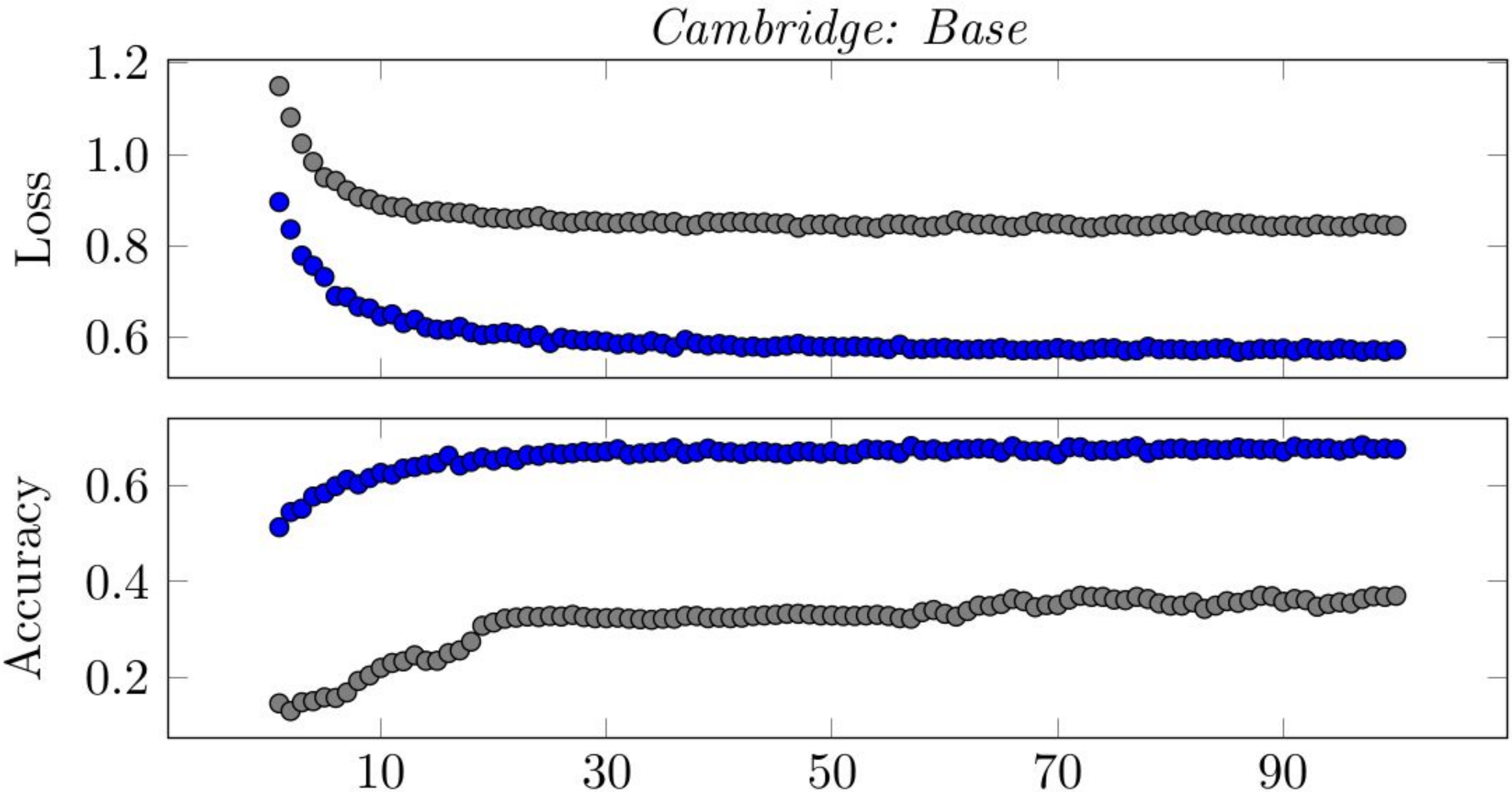
- **Vintage Classifications**

+

- **“Raw” feature vectors**

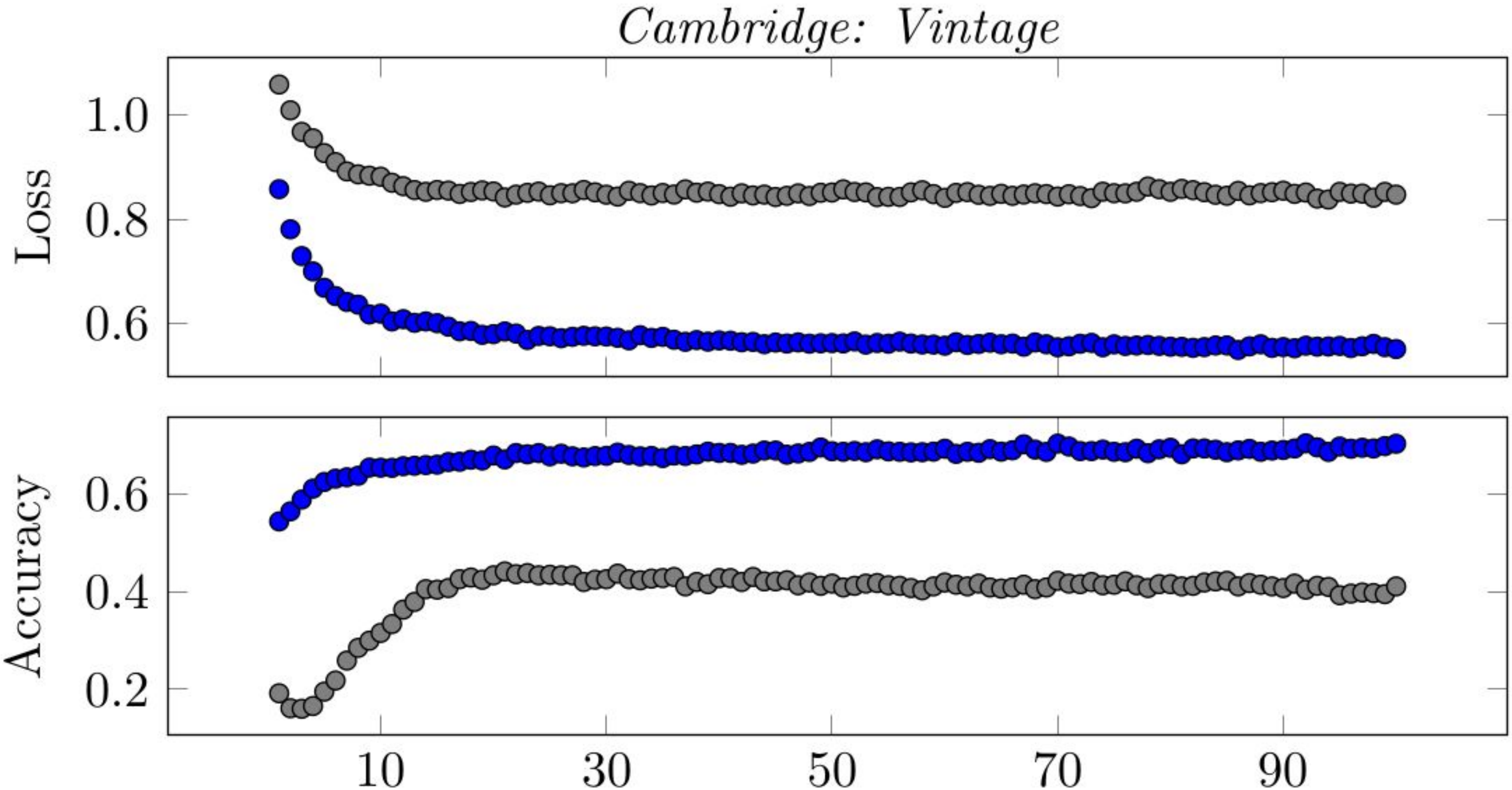
	observed: not round	*observed: round*
Unbalanced out-of-sample test, Cambridge, base		
prediction: not round	3,691	334
prediction: round	748	247
recall	0.83	0.43
precision	0.92	0.25
F_1-score.2	0.87	0.31
Unbalanced out-of-sample test, Cambridge, vintage estimates		
prediction: not round	3,402	261
prediction: round	1,037	320
recall	0.77	0.55
precision	0.93	0.24
F_1-score.1	0.84	0.33
Unbalanced out-of-sample test, Cambridge, Vintage and Inception vec.		
prediction: not round	3,134	269
prediction: round	1,305	312
recall	0.71	0.54
precision	0.92	0.19
F_1-score	0.80	0.28

Well-behaved training curves
(core hedonics, liquidity measures)



Most training done after ~20 epochs

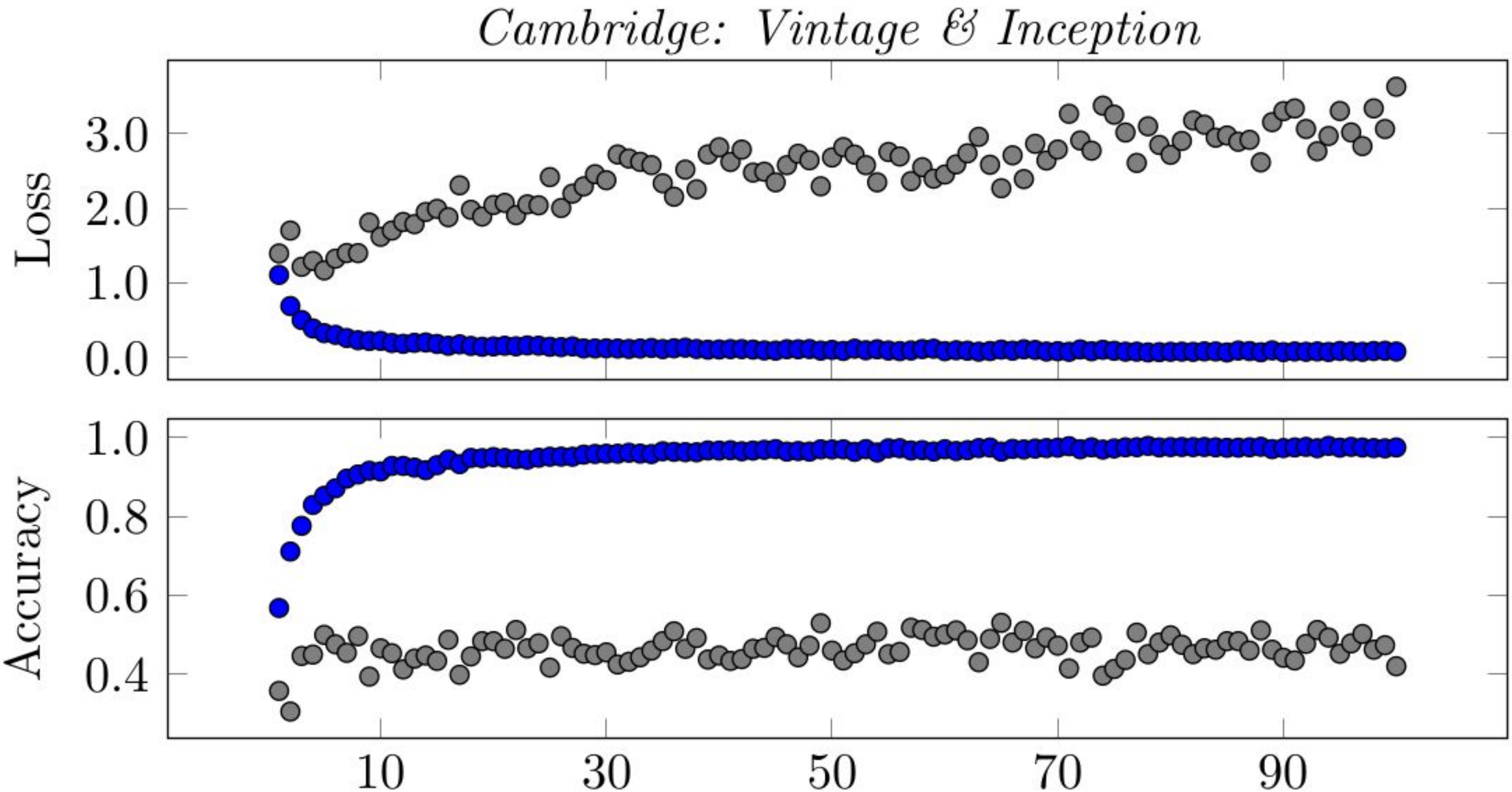
Adding vintage of house and neighbouring buildings



UNIVERSITY OF CAMBRIDGE

Oops. Overfitting.

Clearly not optimal.



UNIVERSITY OF CAMBRIDGE

Reverse regressions: Putting black box into context

Adding the ML classification as another regressor

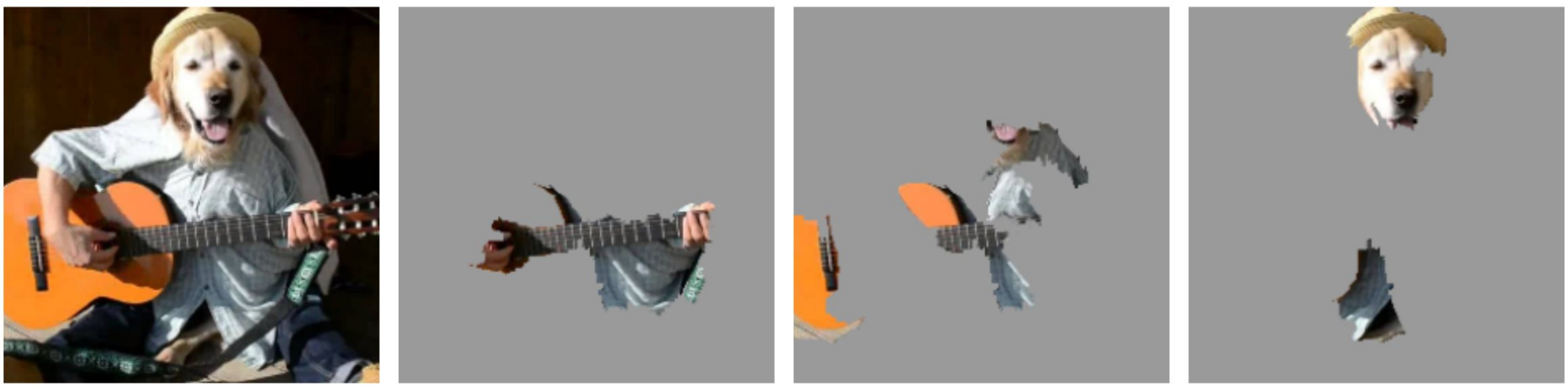
$$ln\left(\frac{RoundPrice_i}{1 - RoundPrice_i}\right) = \alpha + \beta_{Classifier_C} Classification_{C,i} + \sum_{h=1}^{H} \beta_h H_{h,i} + \epsilon_i$$

	Base	*Vint.*	*Vint. & Incep.*	*Base*	*Vint.*	*Vint. & Incep.*
Constant	-2.403^{***}	-2.568^{***}	-2.455^{***}	-3.718^{***}	-4.193^{***}	-4.409^{***}
	(0.057)	(0.064)	(0.064)	(0.580)	(0.582)	(0.571)
ML Classification	1.294^{***}	1.392^{***}	1.024^{***}	1.252^{***}	1.365^{***}	0.971^{***}
	(0.093)	(0.091)	(0.089)	(0.096)	(0.093)	(0.090)
Type: Detached				0.082	0.154	-0.033
				(0.168)	(0.169)	(0.166)
Type: Semi-det.				0.099	0.172	-0.041
				(0.112)	(0.114)	(0.110)
ln(floorplate)				0.355^{***}	0.386^{***}	0.514^{***}
				(0.134)	(0.134)	(0.131)
ln(# comp. + 1)				-0.076	-0.020	-0.053
				(0.060)	(0.060)	(0.058)
Observations	5,020	5,020	5,020	5,020	5,020	5,020
Log Likelihood	$-1{,}709$	$-1{,}682$	$-1{,}733$	$-1{,}698$	$-1{,}671$	$-1{,}722$
Akaike Inf. Crit.	3,421	3,368	3,471	3,408	3,354	3,455

Next steps

Boost sample size & open the black box (a tiny bit)

- **Broaden ML sample beyond Cambridge**
  - Focus on buildings that had round transactions
- **Black box, "Why should I trust you?" Ribeiro, Singh & Guestrin (2016)**
  - Which features influence the classifier most?



(a) Original Image (b) Explaining *Electric guitar* (c) Explaining *Acoustic guitar* (d) Explaining *Labrador*

Figure 4: Explaining an image classification prediction made by Google's Inception neural network. The top 3 classes predicted are "Electric Guitar" ($p = 0.32$), "Acoustic guitar" ($p = 0.24$) and "Labrador" ($p = 0.21$)

Why build like this? K.I.S.S.?

Burntwood Manor, Staffordshire (by Taylor Wimpey)

References

A'Hearn, B., Baten, J. and Crayen, D. (2009). Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital. *Journal of Economic History* 69(3), 783–808.

Allen, E., Dechow, P., Pope, D. and Wu, G. (2017). Reference-Dependent Preferences: Evidence from Marathon Runners. *Management Science* 63(6), 1657–1672.

Baker, M. and Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *Journal of Finance* 61(4), 1645–1680.

Ball, C., Torous, W. and Tschoegel, A. (1985). The Degree of Price Resolution: The Case of the Gold Market. *Journal of Futures Markets* 5(1), 29–43.

Beracha, E. and Seiler, M. (2013). The Effect of Listing Price Strategy on Transaction Selling Prices. *Journal of Real Estate Finance and Economics* 49(2), 237–255.

Binder, C. (2017). Measuring uncertainty based on rounding: New method and application to inflation expectations. *Journal of Monetary Economics* 29, 1–12.

UNIVERSITY OF CAMBRIDGE

References

Goldsmith, M., Koriat, A. and Weinberg-Eliezer, A. (2002). Strategic regulation of grain size in memory reporting. *Journal of Experimental Psychology: General* 131(1), 73–95.

Harris, L. (1991). Stock Price Clustering and Discreteness. *Review of Financial Studies* 4(3), 389–415.

Herrmann, D. and Thomas, W. (2005). Rounding of Analyst Forecasts. *Accounting Review* 80(3), 805–823.

Herxheimer, A. (1991). How much drug in the tablet? *The Lancet* 337, 346–348.

Lacetera, N., Pope, D. and Sydnor, J. (2012). Heuristic Thinking and Limited Attention in the Car Market. *American Economic Review* 102(5), 2206–2236.

Martel, Jordan (2018). Quality Uncertainty in Housing Markets. Working paper.

Mason, M., Lee, A., Wiley, E. and Ames, D. (2013). Precise offers are potent anchors: Conciliatory counteroffers and attributions of knowledge in negotiations. *Journal of Experimental Social Psychology* 49(4), 759–763.

References

Niederhoffer, V. (1966). A New Look at Clustering of Stock Prices. *Journal of Business* 39(2), 309–313.

Osborne, M. (1962). Periodic Structure in the Brownian Motion of Stock Prices. *Operations Research* 10(3), 345–379.

Palmon, O., Smith, B. and Sopranzetti, B. (2004). Clustering in Real Estate Prices: Determinants and Consequences. *Journal of Real Estate Research* 26(2), 115–136.

Ribeiro, M., Singh, S. and Guestrin, C. (2016). “Why should I trust you?” Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

Scott, F. and Yelowitz, A. (2010). Pricing anomalies in the market for diamonds: evidence of conformist behavior. *Economic Inquiry* 48(2), 353–368.

Stevenson, R. and Bear, R. (1970). Commodity Futures: Trends of Random Walks? *Journal of Finance* 25(1), 65–81.

References

Thomas, M., Simon, D. and Kadiyal, V. (2010). The Price Precision Effect: Evidence from Laboratory and Market Data. *Marketing Science* 29(1), 175–190.

Tversky, A. and Kahneman, D. (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology* 5, 207–232.

Yaniv, I. and Foster, D. (1995). Graininess of Judgment Under Uncertainty: An Accuracy–Informativeness Trade-Off. *Journal of Experimental Psychology: General* 124(4), 424–432.

Zhang, Y. and Schwarz, N. (2013). The power of precise numbers: A conversational logic analysis. *Journal of Experimental Social Psychology* 49(1), 944–946.

UNIVERSITY OF CAMBRIDGE