

Reference Dependence in the Housing Market*

Steffen Andersen Cristian Badarinza Lu Liu
Julie Marx and Tarun Ramadorai[†]

July 21, 2020

Abstract

We model listing decisions in the housing market, and structurally estimate household preference and constraint parameters using comprehensive Danish data. Sellers optimize expected utility from property sales, subject to down-payment constraints, and internalize the effect of their choices on final sale prices and time-on-the-market. The data exhibit variation in the listing price-gains relationship with “demand concavity;” bunching in the sales distribution; and a rising listing propensity with gains. Our estimated parameters indicate reference dependence around the nominal purchase price and modest loss aversion. A new and interesting fact that the canonical model cannot match is that gains and down-payment constraints have interactive effects on listing prices.

*We thank Jan David Bakker, Nick Barberis, Richard Blundell, Pedro Bordalo, John Campbell, João Cocco, Joshua Coval, Stefano DellaVigna, Andreas Fuster, Nicola Gennaioli, Arpit Gupta, Adam Guren, Chris Hansman, Henrik Kleven, Ralph Koijen, Ulrike Malmendier, Atif Mian, Karthik Muralidharan, Tomek Piskorski, Claudia Robles-Garcia, Andrei Shleifer, Jeremy Stein, Ansgar Walther, Joshua White, Toni Whited, and seminar participants at the FCA-Imperial Conference on Household Finance, CBS Doctoral Workshop, Bank of England, King’s College Conference on Financial Markets, SITE (Financial Regulation), CEPR Household Finance conference, BEAM, Cambridge Virtual Real Estate Seminar, Bank of England/Imperial/LSE Household Finance and Housing conference, UC Berkeley, Rice University, UT Dallas, Southern Methodist University, University of Chicago, and University of Michigan for useful comments.

[†]Andersen: Copenhagen Business School, Email: san.fi@cbs.dk. Badarinza: National University of Singapore, Email: cristian.badarinza@nus.edu.sg. Liu: Imperial College London, Email: l.liu16@imperial.ac.uk. Marx: Copenhagen Business School, Email: jma.fi@cbs.dk. Ramadorai (Corresponding author): Imperial College London, Tanaka Building, South Kensington Campus, London SW7 2AZ, and CEPR. Tel.: +44 207 594 99 10, Email: t.ramadorai@imperial.ac.uk.

1 Introduction

Housing is typically the largest household asset, and mortgages, typically the largest liability (Campbell, 2006, Badarinza et al. 2016, Gomes et al. 2020). Decisions in the housing market are highly consequential, and are therefore a rich and valuable source of field evidence on households’ underlying preferences, beliefs, and constraints. An influential example is the finding that listing prices for houses rise sharply when their sellers face nominal losses relative to the initial purchase price, originally documented by Genesove and Mayer (2001), and reconfirmed and extended in subsequent literature (see, e.g., Engelhardt (2003), Anenberg, 2011, Hong et al. 2019, and Bracke and Tenreyro 2020). This finding has generally been accepted as *prima facie* evidence of reference-dependent loss aversion (Kahneman and Tversky, 1989, Köszegi and Rabin, 2006, 2007).

Mapping these facts back to underlying preference parameters requires confronting challenges not fully addressed by the extant literature. A rigorous mapping permitting quantitative assessment of parameter magnitudes requires an explicit model of reference-dependent sellers. A plausible model would incorporate additional realistic constraints, such as the fact that optimizing sellers’ listing decisions may be disciplined by demand-side responses.¹ Moreover, such a model would predict the behavior of a range of observables in addition to prices—which can be harnessed to accurately pin down parameters. For example, recent work assessing reference dependence in the field extracts information from transactions quantities (see, e.g., Kleven, 2016 and Rees-Jones, 2018), suggesting new moments to match in the residential housing market setting.

In this paper, we develop a new model of house selling decisions incorporating realistic housing market frictions. We structurally estimate the parameters of the model using a large and granular administrative dataset which tracks the entire stock of Danish housing, and the universe of Danish listings and housing transactions between 2009 and 2016,

¹Recent progress has been made on documenting the shape of housing demand (e.g., Guren, 2018), but it is important to understand how this affects inferences about the relationship between listing prices and sellers’ “potential gains”.

matched to household demographic characteristics and financial information. These rich data also yield several new facts about household decisions that we cannot match using canonical model features, making them targets for future theoretical work.

In our model, sellers face an extensive margin decision of whether to list, as well as an intensive margin choice of the listing price. Sellers maximize expected utility both from the final sale price of the property as well as (potentially asymmetrically) from any gains or losses relative to a fixed reference price, which we simply set to the nominal purchase price of the property. We adopt a standard piecewise linear formulation of reference-dependent utility, characterized by two parameters: η captures how gains are weighed relative to the utility of the final sale price, and λ captures the asymmetric disutility of losses, i.e., conventionally, when $\lambda > 1$, sellers are loss averse. Sellers enjoy additional “gains from trade” from successful sales, receive an outside option utility level otherwise, and face down-payment constraints à la Stein (1995). Sellers take into account how their choices affect outcomes, i.e., the probability of sale as well as the final sale price, given housing demand.

We summarize a few important insights from the model here. When sellers exhibit “linear reference dependence” ($\eta > 0$, i.e., gains and losses matter to sellers, but $\lambda = 1$, i.e., there is no asymmetry between gains and losses), optimal listing premia decline linearly with “potential gains” (the difference between the expected sale price and the reference price) accrued since purchase. Intuitively, such linearly reference-dependent sellers facing losses require a greater final sales price to elevate the total utility received from a successful sale above that of the outside option. This leads them to raise (lower) listing prices in the face of potential losses (gains).² In addition, if sellers are loss averse, with $\lambda > 1$, then optimal listing premia slope up more sharply when sellers face potential losses than when they face potential gains, reflecting the asymmetry in underlying preferences.

These predictions on listing premia are mirrored in the behavior of quantities. With

²In the trivial case of no reference dependence, i.e., when $\eta = 0$, the model predicts that optimal listing premia are simply flat in potential gains.

linear reference dependence, completed transactions more frequently occur at realized gains (when the final sales price exceeds the reference price) than at realized losses. Put differently, $\eta > 0$ implies a shift of mass to the right in the distribution of transactions along the realized gains dimension, relative to the distribution when $\eta = 0$. With loss aversion, there is, in addition, sharp bunching of transactions precisely at realized gains of zero, and a more pronounced shift of mass of transactions away from realized losses.

Reference dependence and loss aversion also affect the extensive margin. The model predicts that the propensity to list rises in potential gains if $\eta > 0$. When $\lambda > 1$, there is also a pronounced decline over the domain of potential losses. Accounting for the extensive margin decision additionally helps to clean up inferences on the intensive margin, which can otherwise be biased by the drivers of selection into listing.

This discussion suggests that mapping reduced-form facts to underlying preference parameters is straightforward, but several key confounds can interfere. For one, the model reconfirms an issue recognized in prior work (e.g., Genesove and Mayer, 1997, 2001), that downsizing aversion à la Stein (1995) is difficult to separate from loss aversion. Down-payment constraints on mortgages create an incentive for households to “fish” with higher listing prices, since household leverage magnifies declines in collateral value, severely compressing the size of houses into which households can move. This effect of household leverage strongly manifests itself in listing prices in the data, but we document significant independent variation with potential gains, allowing us to cleanly identify loss aversion.

Second, accurate measurement of sellers’ “potential gains” is important for our exercise. We confirm that the hedonic model that we employ to predict house prices in our main analysis fits the data with high explanatory power ($R^2 = 0.86$), and that our empirical work is robust to alternative house price prediction approaches. Third, relatedly, as Genesove and Mayer (2001), Clapp et al. (2018), and others note, variation in the unobservable property quality and potential under- or over-payment at the time of

property purchase are important sources of measurement error. As we describe later, we adopt a wide range of strategies to check robustness to this possible confound.³

Fourth, the shape of demand is very important for model outcomes. If sale probabilities respond linearly and negatively to higher listing prices (“linear demand”), there are material incentives to set low list prices to induce quick sales. However, Guren (2018) shows that U.S. housing markets are characterized by “concave demand,” i.e., past a point, lowering list prices does not boost sale probabilities, but does negatively impact realized sale prices; we confirm this finding in the Danish data.⁴ The model reveals that this can generate a nonlinear optimal listing price schedule even without any underlying loss aversion. Intuitively, in the face of linear demand, a seller with $\eta > 0$ and $\lambda = 1$ linearly lowers list prices with potential gains, focusing on inducing a swift sale. However, when facing concave demand, lowering list prices past a point is unproductive, leading to an observed “flattening out” in the optimal listing price schedule, which is then nonlinear even though $\lambda = 1$. A related and important observation from the model is that sharp demand responses to raising listing prices are associated with weaker listing price responses to losses, and vice versa.

Keeping these potential confounds in mind, we outline the main facts in the data. First, the listing price schedule has the characteristic “hockey stick” shape first identified by Genesove and Mayer (2001), rising substantially as expected losses mount, and virtually flat in gains. Our estimates are similar in magnitude to those in that paper despite the differences in location, sample period, and sample size.⁵ Second, listing premia vary considerably across regional housing markets in Denmark which exhibit varying degrees

³This includes estimation with property-specific fixed effects, applying bounding strategies previously proposed in the literature (Genesove and Mayer, 2001), utilizing an instrumental variables approach proposed by Guren (2018), and employing a Regression Kink Design (Card et al., 2015b)

⁴We also show using these data that there are substantial increases in the *volatility* of time on the market associated with higher listing premia, a new and important observation.

⁵In the original Genesove and Mayer sample of Boston condominiums between 1990 and 1997 [N=5,792], list prices rise between 2.5 and 3.5% for every 10% nominal loss faced by the seller. We find rises of 4.4 and 5.4% for the same 10% nominal loss in the Danish data of apartments, row houses, and detached houses between 2009 and 2016 [N=173,065].

of demand concavity. This variation is consistent with the model: steep listing premia responses to losses are observed in markets with weaker demand concavity, and vice versa. These regional moments provide additional discipline to our structural estimation exercise and help account for the demand-concavity confound. Third, we see sharp bunching in the sales distribution at realized gains of zero, and a significant shift in mass in the distribution of sales towards realized gains and away from realized losses. Fourth, we estimate listing propensities for the entire Danish housing stock of over 5.5 million housing units as a function of potential gains. There is a visible increase in the propensity to list houses on the market as potential gains rise, and the slope appears more pronounced over the potential loss domain than the potential gain domain.

Taken together, these facts appear consistent with underlying preferences that are both reference dependent and loss averse around the original nominal purchase price of the house. To more rigorously map these facts back to the model, we structurally estimate seven model parameters using seven selected moments from the data (including those described above) using classical minimum distance estimation in this exactly identified system. The resulting point estimates yield $\eta = 0.948$ (*s.e.* 0.344), meaning that gains count about as much as final prices for final utility, and $\lambda = 1.576$ (*s.e.* 0.570), a modest degree of loss aversion, lower than early estimates between 2 and 2.5 (e.g., Kahneman et al. 1990, Tversky and Kahneman, 1992), but closer to those in more recent literature (e.g., Imas et al. 2016 find $\lambda = 1.59$). The role of concave demand is important for these parameter estimates—in a restricted model in which we assume that demand is (counterfactually) linear, estimated $\eta = 0.750$ (*s.e.* 0.291) and $\lambda = 3.285$ (*s.e.* 0.867).⁶ This strongly reinforces a broader message (see, e.g., Blundell, 2017) that realistic frictions need to be incorporated when mapping reduced-form facts from the data to inferences about deeper underlying parameters, strengthening the case for applying a structural

⁶This also highlights that frictions in matching in the housing market are another important part of the explanation for the positive correlation between volume and price observed in housing markets, an original motivation for the mechanisms identified by both Stein (1995) and Genesove and Mayer (2001)—both of which our model incorporates.

behavioral approach (DellaVigna (2009, 2018)) to field evidence. Finally, the estimated parameters also reveal strong evidence of the down-payment channel originally identified by Stein (1995), reveal significant “gains from trade” from successful house listings, and highlight that there are substantial (psychological and transactions) costs associated with listing.

The model does a good job of matching the selected moments with plausible preference parameters. However as an out-of-sample exercise, we conduct a broader evaluation of how the model matches the entire surface of the listing premium along the home equity and gains dimensions. A novel pattern that we uncover, and that our model cannot match, is that home equity and expected losses have interactive effects on listing prices in this market. To be more specific, when home equity levels are low, i.e., when down-payment constraints are tighter, households set high listing prices that vary little around the nominal loss reference point. In contrast, households that are relatively unconstrained set listing prices that are significantly steeper in expected losses. Households’ listing price responses to down-payment constraints are also modified by their interaction with nominal losses. Mortgage issuance by banks in Denmark is limited to an LTV of no greater than 80%,⁷ and for households facing nominal losses since purchase, listing prices rise visibly as home equity falls below this down-payment constraint threshold of 20%. But for households expecting nominal gains, there is a strong upward shift in this constraint threshold (i.e., to values above 20%) in the *level* of home equity at which they raise listing prices. We discuss these findings and conjecture mechanisms to explain them towards the end of the paper; we view them as potential targets for future theoretical work.

The paper is organized as follows. Section 2 introduces the model of household listing behavior. Section 3 discusses the construction of our merged dataset, and provides descriptive statistics about these data. Section 4 introduces the moments that we use for structural estimation and uncovers new facts about the behavior of listing prices and

⁷We later describe the precise institutional features of the Danish setting, which permits additional non-mortgage borrowing at substantially higher rates for higher LTV mortgages.

listing decisions. Section 5 describes our structural estimation procedure, and reports parameter estimates. Section 6 describes validation exercises, and highlights areas where the model falls short in explaining features of the data. Section 7 concludes.

2 A Model of Household Listing Behavior

We develop a model in which a household (the “seller”), optimally decides on a listing price (the “intensive margin”), as well as whether or not to list a house (the “extensive margin”). The model framework can flexibly embed different preferences and constraints that have commonly been used to explain patterns in listing behavior. In this section we describe the main features and specific predictions of the model, which we later structurally estimate to recover key preference and constraint parameters from the data.

2.1 General Framework

The market consists of a continuum of sellers and buyers of residential property. There are two periods in the model: in period 0, some fraction of property owners receive a shock $\theta \sim \text{Uniform}(\theta_{\min}, \theta_{\max})$, and decide (i) whether or not to put their property up for sale, and (ii) the optimal listing price in case of listing. This “moving shock” θ can be thought of as a “gain from trade” (Stein, 1995), i.e., a boost to lifetime utility which sellers receive in the event of successfully selling and moving, which captures a variety of reasons for moving, including labor market moves to opportunity, or the desire to upsize arising from a newly expanded family. In period 1, buyers visit properties that are up for sale. If the resulting negotiations succeed, the property is transferred to the buyer for a final sale price. If negotiations fail, the seller stays in the property, and a receives a constant level of utility \underline{u} .

We seek to uncover the structural relationship between listing decisions and seller preferences and constraints. To sharpen this focus, we model buyer decisions and equilibrium

negotiation outcomes in reduced-form, and focus on recovering seller policy functions from this setup. In particular, let L denote the listing price set by the seller; \widehat{P} be a measure of the “expected” or “fundamental” property value;^{8,9} $\ell = L - \widehat{P}$ be termed the listing premium; let α denote the probability that a willing buyer will be found; and P denote the final sale price resulting from the negotiation between buyer and seller where $P(\ell) = \widehat{P} + \beta(\ell)$.

A typical seller’s decision in period 0 can be written as:

$$\max_{s \in \{0,1\}} \left\{ (s) \max_{\ell} \underbrace{[\alpha(\ell) (U(P(\ell), \cdot) + \theta) + (1 - \alpha(\ell))\underline{u} - \varphi]}_{EU(\ell)} + (1 - s)\underline{u} \right\} \quad (1)$$

The seller decides on the extensive margin of whether ($s = 1$) or not ($s = 0$) to list, as well as the listing premium ℓ , to maximize expected utility from final sale of the property. For a listed property, there are two possible outcomes in period 1, which depend on ℓ . With probability $\alpha(\ell)$ the negotiation succeeds, and the seller receives utility from selling the property for an equilibrium price $P(\ell) = \widehat{P} + \beta(\ell)$. With probability $1 - \alpha(\ell)$ the listing fails, in which case the seller falls back to their outside option level of utility \underline{u} . In addition, owners who decide to list incur a one-time cost φ , which is sunk at the point of listing—all utility costs associated with listing (e.g., psychological “hassle factors”, search, listing and transaction fees) are captured by this single parameter.

When making these listing decisions, the seller takes $\alpha(\ell)$ and $\beta(\ell)$, i.e., the “demand”

⁸Guren (2018) assumes that the buyer’s expected value is given by the average listing price in a given zip code and year. This allows for more flexibility, allowing listing prices to systematically deviate from hedonic/fundamental property values across time and locations. We begin with a simpler benchmark, setting \widehat{P} to the fundamental/hedonic value of the house in the interests of internal consistency of the model. As we show later, this distinction does not play a significant role in our empirical work, as Denmark has a relatively homogenous and liquid housing market, and we show that the listing premium based on hedonic prices more strongly predicts a decrease in the probability of sale than the alternative based on average listing prices in a direct comparison in the online appendix.

⁹In the model solution and calibration exercise, we normalize \widehat{P} to 1. All model quantities can therefore be thought of as being expressed in percentages (which we later map to logs, relying on the usual approximation), to be consistent with the definitions of gains/losses and home equity employed in our empirical work.

functions, as given; we estimate these functions in the data as a reduced-form for equilibrium outcomes in the negotiation process in period 1, which the seller internalizes when optimizing utility. As in Guren (2018), we note that sufficient statistic formulas (Chetty, 2009) for equilibrium outcomes are mappings between sale probabilities $\alpha(\ell)$, final sale prices $P(\ell) = \widehat{P} + \beta(\ell)$, and listing premia ℓ . In particular, the realized premium $\beta(\ell)$ of the final sales price P over the expected property value \widehat{P} , and the probability of a quick sale $\alpha(\ell)$ arise from the seller’s listing behavior, and the subsequent negotiation process with the buyer. This assumption simplifies the model, and allows us to more closely focus on our goal, namely, extracting the underlying parameters of seller utility and constraints.¹⁰

The functions $\alpha(\ell)$ and $\beta(\ell)$ restrict the seller’s action space, and capture the basic tradeoff that sellers face: a larger ℓ can lead to a higher ultimate transaction price, but decreases the probability that a willing buyer will be found within a reasonable time frame.¹¹ These points capture the link between listing premia, final realized sales premia, and time-on-the-market or TOM originally detected by Genesove and Mayer (2001). In the remainder of the paper, we refer to these two functions $\alpha(\ell)$ and $\beta(\ell)$ collectively as *concave demand*, following Guren (2018), who documents using U.S. data that above average list prices increase TOM (i.e., they reduce the probability of final sale), while below average list prices reduce seller revenue with little effect on TOM. We find essentially the same patterns in the Danish data.

We next describe the components of $U(P(\ell), \cdot) = u(P(\ell), \cdot) - \kappa(P(\ell), \cdot)$, which allows us to nest a range of preferences $u(P(\ell), \cdot)$, including reference-dependent loss-aversion à la Kahneman and Tversky (1979) and Köszegi and Rabin (2006, 2007), as well as down-payment constraints $\kappa(P(\ell), \cdot)$ à la Stein (1995).

¹⁰As we describe later, we do allow for the seller to perceive $\alpha(\ell)$ differently from the (ex-post) estimated mapping function in the data by adding a parameter δ to the model, i.e., the seller maximizes subject to their perceived $(\alpha(\ell) + \delta)$ probability.

¹¹In our estimation, we define a *period* as equal to six months. In this case, the function $\alpha(\ell)$ captures the probability that the transaction goes through within a time frame of six months after the initial listing.

2.2 Reference-Dependent Loss Aversion

We adopt a standard formulation of reference-dependent loss averse preferences, writing $u(P(\ell), \cdot)$ as:

$$u(P(\ell), R) = \begin{cases} P(\ell) + \lambda\eta G(\ell), & \text{if } G(\ell) < 0 \\ P(\ell) + \eta G(\ell), & \text{if } G(\ell) \geq 0 \end{cases}. \quad (2)$$

In equation (2), the seller’s reference price level is R , which we simply assume is fixed, and in our empirical application, we set R to the original nominal purchase price of the property.¹² Realized gains $G(\ell)$ relative to this reference level are then given by $G(\ell) = P(\ell) - R$.

The parameter η captures the degree of reference dependence. Sellers derive utility both from the terminal value of wealth (i.e. the final price P realized from the sale), as well as from the realized gain G relative to the reference price R .

The parameter $\lambda > 1$ governs the degree of loss aversion. This specification of the problem assumes that utility is piecewise linear in nominal gains and losses relative to the reference point, with a kink at zero, and has been used widely to study and rationalize results found in the lab (e.g., Ericson and Fuster, 2011), as well as in the field (e.g., Anagol et al., 2018).

2.2.1 State Variables

In the model, seller decisions are determined by four state variables, namely, the moving shock θ , the hedonic value of the property \hat{P} , the reference point R , and the outside option level \underline{u} . To map model quantities more directly to estimates in the data, and to make our setup more directly comparable to extant empirical and theoretical literature, we calculate the seller’s expected or “potential” gains $\hat{G} = \hat{P} - R$ as a transformation of

¹²While this is a restrictive assumption, we find strong evidence to suggest the importance of this particular specification of the reference point in our empirical work. We follow Blundell (2017), trading off a more detailed description of the decision-making problem in the field against stronger assumptions that permit measurement of important underlying parameters.

two of the state variables.¹³ Realized gains $G(\ell)$ arise from their “potential” level \widehat{G} plus the markup/premium $\beta(\ell)$, i.e.:

$$G(\ell) = \widehat{G} + \beta(\ell).$$

The remaining two state variables θ and \underline{u} are unobserved, but only the wedge between them ($\underline{u} - \theta$) is relevant for the seller’s decision. Without loss of generality, we therefore set the outside option $\underline{u} = \widehat{P}$, which implies that absent any additional reasons to move ($\theta = 0$), and with costless and frictionless listings, the seller will be indifferent between staying in their home and receiving the hedonic value in cash. This assumption can equivalently be mapped onto a specification in which the seller does not receive any gains from moving, but experiences a $-\theta$ shock in the event of a failed sale (i.e. the outside option is then rewritten as $\underline{u} = \widehat{P} - \theta$).

We also note that the model implicitly specifies conditions on the relationship between \underline{u} and R . In the online appendix, we discuss this issue in detail. We show there that (i) assuming that R enters (or equals) the outside option (i.e., the consumption utility of households in the event of no sale) generates implausible predictions that we can reject in the data, (ii) if R is used by the seller to “rationally” forecast \widehat{P} (given our normalization of $\underline{u} = \widehat{P}$), the result is innocuous, and doesn’t affect any inferences from the model, and (iii) it is potentially possible to reinterpret the model as one of non-rational belief formation (i.e., the seller might view R as the “correct” outside option value), but it is potentially more difficult to rationalize several of the patterns we find in the data (i.e., bunching at just positive gains) with such a model of beliefs.

We next discuss selected predictions of the model to build intuition, and to guide our

¹³We capture listing behavior by studying the listing premium $\ell = L - \widehat{P}$, which is an innocuous normalization of the listing price L . One way to see this is to note that the regression $\underbrace{L - \widehat{P}}_{\ell} = \rho \underbrace{(\widehat{P} - R)}_{\widehat{G}}$

is equivalent to $L = (1 + \rho)\widehat{P} - \rho R$. We estimate a version of this regression in the online appendix and verify the original inferences of Genesove and Mayer (2001) using our sample.

choice of key moments of the data with which to structurally estimate key parameters.

2.2.2 Optimal Listing Premia

To begin with, consider only the intensive margin decision of the optimal choice of listing premium, and assume that $U(P(\ell), \cdot) = u(P(\ell), \cdot)$:

$$\max_{\ell} [\alpha(\ell) (u(P(\ell), \cdot) + \theta) + (1 - \alpha(\ell))\underline{u}] \quad (3)$$

The first-order condition which determines the optimal ℓ^* balances the marginal utility benefit of a higher premium (conditional on a successful sale) against the marginal cost of an increased chance of the transaction failing, and the consequent fall to the outside option utility level.

To aid interpretation, we analytically solve a version of the simple model in equation (3), under the assumption that demand functions $\alpha(\ell) = \alpha_0 - \alpha_1\ell$ and $\beta(\ell) = \beta_0 + \beta_1\ell$ are linear in ℓ (this is an assumption that we later relax to account for concave demand). In this case, the model yields an optimal listing premium schedule which is piecewise linear:

$$\ell^*(\widehat{G}) = \begin{cases} \frac{1}{2} \left(\frac{\alpha_0}{\alpha_1} - \frac{\beta_0}{\beta_1} - \frac{1}{\beta_1} \frac{\theta}{1+\eta} \right) - \frac{1}{2\beta_1} \frac{\eta}{1+\eta} \widehat{G}, & \text{if } \widehat{G} \geq \widehat{G}_0 \\ -\frac{\beta_0}{\beta_1} - \frac{1}{\beta_1} \widehat{G}, & \text{if } \widehat{G} \in (\widehat{G}_1, \widehat{G}_0) \\ \frac{1}{2} \left(\frac{\alpha_0}{\alpha_1} - \frac{\beta_0}{\beta_1} - \frac{1}{\beta_1} \frac{\theta}{1+\lambda\eta} \right) - \frac{1}{2\beta_1} \frac{\lambda\eta}{1+\lambda\eta} \widehat{G}, & \text{if } \widehat{G} \leq \widehat{G}_1, \end{cases} \quad (4)$$

where \widehat{G}_0 and \widehat{G}_1 are levels of potential gains determined by underlying model parameters.¹⁴

Figure 1 illustrates how equation (4) varies with the underlying parameters characterizing preferences.

In the case of no reference dependence ($\eta = 0$), utility derives purely from the terminal house price. In this case, the top left-hand plot shows that ℓ^* is unaffected by the reference

¹⁴We derive the equation explicitly in the online appendix.

price R .

In the case of linear reference dependence ($\eta > 0$, $\lambda = 1$), there is a negatively-sloped linear relationship between ℓ^* and \widehat{G} . In this case, R does not affect the marginal benefit of raising ℓ^* , but it does affect the marginal cost, as it affects the distance between \underline{u} and the achievable utility level in the event of a successful transaction. Intuitively, if the household can realize a gain (i.e., when R is sufficiently low), the utility from a successful sale rises. The resulting ℓ^* will therefore be lower, as the household seeks to increase the probability that the sale goes through. The opposite is true when the household faces a loss in the event of a completed sale (i.e., when R is sufficiently high), which consequently results in a higher ℓ^* .¹⁵

In the case of (reference dependence plus) loss aversion ($\eta > 0$, $\lambda > 1$), the kink in the piecewise linear utility function leads to a more complex piecewise linear pattern in ℓ^* , which determines the gains that sellers ultimately realize. There is a unique level of potential gains, \widehat{G}_0 , which maps to a realized gain of *exactly* zero (recall that $G(\ell^*) = \widehat{G} + \beta(\ell^*)$). Sellers with potential gains below \widehat{G}_0 want to avoid realizing a loss, meaning that they adjust ℓ^* upwards. However, this upward adjustment increases the probability of a failed sale. Beyond some lower limit \widehat{G}_1 , the costs in terms of the failure probability become unacceptably high relative to the benefit of avoiding a loss, and it becomes sub-optimal to aim for a realized gain of zero. The seller has no choice but to accept the loss at levels of $\widehat{G} < \widehat{G}_1$, but still sets a marginally higher listing premium for each unit loss beyond this point.

¹⁵As mentioned earlier, it is important to assume that households do not receive utility from simply living in a house that has appreciated relative to their reference point R , i.e. they do not enjoy utility from “paper” gains until they are realized. If this condition does not hold, the model is degenerate in that R is irrelevant both for the choice of the listing premium (intensive margin) and the decision to list (extensive margin). We demonstrate this result analytically in the online appendix.

2.2.3 Bunching Around Realized Gains of Zero

The model reveals that household listing behavior also has material implications for quantities. Loss-averse preferences show up in non-linearities in the schedule of ℓ^* along the \widehat{G} dimension, as well as on the likelihood of transaction completion, and the final price at which these transactions occur. This shows up as shifts in mass in the distribution of completed transactions along the G dimension, additional moments which allow us to pin down underlying utility parameters. In the simple version of the model (assuming linear demand) discussed above, the equation relating potential gains \widehat{G} with final realized gains G is:

$$G(\widehat{G}) = \begin{cases} \beta_0 + \frac{\beta_1}{2} \left(\frac{\alpha_0}{\alpha_1} - \frac{\beta_0}{\beta_1} - \frac{1}{\beta_1} \frac{\theta}{1+\eta} \right) + \left(1 - \frac{1}{2} \frac{\eta}{1+\eta} \right) \widehat{G} & \text{if } \widehat{G} > \widehat{G}_0, \\ 0 & \text{if } \widehat{G} \in [\widehat{G}_1, \widehat{G}_0], \\ \beta_0 + \frac{\beta_1}{2} \left(\frac{\alpha_0}{\alpha_1} - \frac{\beta_0}{\beta_1} - \frac{1}{\beta_1} \frac{\theta}{1+\lambda\eta} \right) + \left(1 - \frac{1}{2} \frac{\lambda\eta}{1+\lambda\eta} \right) \widehat{G} & \text{if } \widehat{G} < \widehat{G}_1. \end{cases} \quad (5)$$

The two bottom panels of Figure 1 illustrate how this relationship varies with underlying utility parameters.

When $\eta = 0$, sellers choose a constant listing premium ℓ^* , which results in a constant realized premium $\beta(\ell^*)$ of actual gains G over potential gains \widehat{G} (bottom left plot), meaning that the distribution of G is a simple parallel shift of the distribution of \widehat{G} (bottom right plot, the black dotted line becomes the purple line).

In the linear reference dependence model ($\eta > 0, \lambda = 1$), sellers with $\widehat{G} < 0$ choose relatively higher ℓ^* . This lowers the likelihood that willing buyers will be found, meaning that the likelihood of observing transactions in this domain of \widehat{G} is lower. However, if these transactions *do* go through, the associated G will be higher, shifting mass in the final sales distribution towards $G > 0$ (bottom right plot, the black dotted line becomes the green line).

The effect mentioned above is especially pronounced if sellers are loss averse, i.e., when $\lambda > 1$, in which case the model predicts bunching ($F(\widehat{G}_0) - F(\widehat{G}_1)$) in the final distribution of house sales precisely at $G = 0$ (bottom left plot, black line and bottom right plot black solid line), and greater mass in the distribution when $G > 0$, coming from even less mass when $G < 0$ (bottom right plot, the black dotted line becomes the black solid line).

In the discussion thus far, to build intuition about the effect of the underlying parameters characterizing preferences, we focused on the intensive margin, made several assumptions about the shape of demand, and assumed away other frictions and constraints. We next outline the predictions of the model in the broader case when we consider the extensive margin decision, and then turn to discussing two important potential confounds, namely, concave demand, and the effect of financial constraints.

2.2.4 Extensive Margin

In the discussion thus far, we ignored the seller’s decision of whether or not to list. In the model, any force inducing a wedge between the expected utility from a successful listing and the outside option \underline{u} affects decisions along the intensive margin, but can also push the seller towards deciding that listing is sub-optimal. In particular, the model predicts that sellers with lower \widehat{G} are less likely to list. This clear prediction allows us to exploit the relationship of the listing propensity and \widehat{G} as an additional moment to inform structural estimation of underlying preference parameters.¹⁶

Another important observation here here is that modeling the extensive margin decision is also important to account for any selection effects that may drive patterns of observed intensive margin listing premia in the data, an issue that the prior literature (e.g., Genesove and Mayer, 1997, 2001, Anenberg, 2011, Guren, 2018) has been unable to control for as a result of data limitations. For example, if sellers that decide not to

¹⁶Bunching in the distribution of realized house sales captures *ex post*-negotiation outcomes, and extensive margin decisions capture sellers’ *ex ante* listing behaviour, i.e., these two moments are informative about different phases in the listing/selling decision.

list are more conservative (i.e., they set lower listing premia), and those who decide to list are more aggressive (i.e., setting higher listing premia) the resulting selection effect would lead to a higher observed non-linearity in listing premia around reference points that would bias parameter estimates and inferences conducted only using the intensive margin.¹⁷

The moving shock θ (which alters the distance between the outside option and the utility from a successful listing) is a key model component that helps to capture such selection effects. Conditional on the moving shock, the listing decision is a simple binary choice. This means that accounting for the distribution of shocks, as we do in the model, allows us to capture the variation in listing decisions and to calculate average listing premia in the population. These average listing premia incorporate the endogenous first-stage selection effects and can be mapped directly back to the data.

There are more subtle implications of the model linking the extensive and the intensive margins. High realizations of θ affect the listing decision, and push the seller towards setting higher listing premia. However, this force can move ℓ into regions of concave demand (which we discuss in detail in the next subsection) in which the response of buyers is more (or less) pronounced, because of nonlinearities in $\alpha(\ell)$. This in turn means that θ variation can affect the observed magnitude of the seller’s responses to \widehat{G} , smoothing and blurring the kinks in the model-implied ℓ^* profile. The online appendix illustrates this with a specific example, showing that the characteristic “hockey stick” shape of the average listing premium profile can result from averaging the three-piece-linear form of the listing premium profile in the case of $\lambda > 1$ across the distribution of θ .

2.3 Concave Demand

The demand functions $\alpha(\ell)$ and $\beta(\ell)$ are a critical determinant of listing behavior and the expected shape of ℓ^* in this model. This can be seen even in the simple case of the linear

¹⁷We thank Jeremy Stein for useful discussions on this issue.

demand functions posited earlier. Equation (4) shows that when the probability of sale is less sensitive to ℓ (i.e., when α_1 is lower), the marginal cost of choosing a larger listing premium is lower, and therefore the optimally chosen ℓ^* is higher. This intuition carries over to a case in which $\alpha(\ell)$ has the concave shape first identified by Guren (2018), and has important implications for the relationship between ℓ^* and \widehat{G} . Figure 2 graphically illustrates this mechanism, positing a concave shape for $\alpha(\ell)$ and considering the effect of varying $\alpha(\ell)$ around $\underline{\ell} = 0$, i.e., the point at which $L = \widehat{P}$ (solid and dashed red lines, right-hand plot).

The left-hand plot in Figure 2 documents the relationship between the optimal listing premium ℓ^* and \widehat{G} in the presence of concave demand. When $\widehat{G} > 0$, the seller’s incentive is to set ℓ^* low, since they are motivated to successfully complete a sale and capture gains from trade θ . However, in the presence of concave demand (i.e., as illustrated in the right-hand plot, horizontal $\alpha(\ell)$ when $\ell < \underline{\ell}$; combined with $P(\ell) = \beta_0 + \beta_1\ell$), lowering ℓ below $\underline{\ell}$ does not boost the sale probability $\alpha(\ell)$, but doing so does negatively impact the realized sale price $P(\ell)$. It is thus optimal for ℓ^* to “flatten out” at the level $\underline{\ell}$.

The tradeoff faced by sellers facing losses $\widehat{G} < 0$ is different—raising ℓ^* helps to offset expected losses, but lowers the probability of a successful sale. When demand concavity increases, i.e., $\alpha(\ell)$ is more steeply negative, the probability of a successful sale falls at a faster rate with increases in ℓ . Figure 6 illustrates this force—moving from the dashed $\alpha(\ell)$ schedule to the solid $\alpha(\ell)$ schedule in the right-hand plot in turn leads to dampening of the slope of ℓ^* in the left-hand plot. In the extreme case in which concave demand has an infinite slope around some level of the listing premium, rational sellers’ ℓ^* collapses to a constant—which would be observationally equivalent to the case in which sellers are not reference dependent at all ($\eta = 0$).

The main predictions from the model in this case are: First, the optimal ℓ^* in a linear reference-dependent model ($\eta > 0$, $\lambda = 1$) in the presence of concave demand exhibits a flatter slope in the domain $\widehat{G} > 0$ relative to the case of linear demand. This means

that the graph of ℓ^* against \widehat{G} can exhibit a characteristic “hockey stick” shape of the type detected by Genesove and Mayer (2001) even if there is no loss aversion, i.e., $\lambda = 1$. Second, the model predicts a tight link between the shape of $\alpha(\ell)$ and the slope of ℓ^* . We later use this insight to exploit cross-sectional variation in the concavity of demand across different segments of the Danish market to aid structural parameter identification.¹⁸ Third, while we have focused our discussion on how concave demand can generate a non-linear listing premium profile, it will also result in effects on transactions volume. That is, concave demand can result in additional shifts of mass towards positive values of realized gains, depending on the level of $\underline{\ell}$, though it will not be associated with sharp bunching of the type associated with loss aversion.

A subtle point here is that any change in the precise specification of the reference point R in the presence of loss aversion will change the location at which bunching is observed. Indeed, heterogeneity in reference points will make it hard to observe the precise location of bunching. To complicate matters further, variations in the level of $\underline{\ell}$ are a confound, potentially rendering it difficult to distinguish models with heterogeneous reference points from models with spatial or temporal variation in $\underline{\ell}$, the point at which demand concavity kicks in. We avoid this complexity in our setup by simply taking the stance that R is the nominal purchase price of the property and evaluating the extent to which we see bunching given this assumption. As we will later see, this turns out to be a reasonable assumption—we observe significant evidence in the data of bunching using this assumption about R , confirming its relevance to sellers.

2.4 Down-payment Constraints

A well-known confound for the estimation of preference parameters from listing premia (see, e.g., Genesove and Mayer (1997, 2001)) is the effect of down-payment constraints,

¹⁸For example, if $\eta = 0$ in this model, demand concavity does not affect the slope of the ℓ^* profile along the G dimension. In contrast, a high η leads to a high “pass-through” of demand concavity into optimal listing premia.

which we account for in the model through the function $\kappa(P(\ell), \cdot)$ (recall that $U(P(\ell), \cdot) = u(P(\ell), \cdot) - \kappa(P(\ell), \cdot)$). Let M denote the level of the household’s outstanding mortgage, and γ the required down-payment on a *new* mortgage origination. For a given price level $P(\ell)$, the “realized” home equity position of the household is $H(\ell) = P(\ell) - M$. Under the assumption that H is put towards the down payment on the next home, we can distinguish between constrained (i.e., downsizing-averse) households for which $H(\ell) < \gamma$, and unconstrained households for which $H(\ell) \geq \gamma$.

In the face of binding down-payment constraints, only unconstrained sellers can move to another property of the same or greater value. However, there are several ways in which households could relax these constraints despite legal restrictions on LTV at mortgage initiation (which, as we discuss later, are strictly set at 20% in Denmark). The first is for households to downsize to a less expensive home than $P(\ell)$, or indeed, to move to the rental market—either decision might incur a utility cost. The second is that households can engage in non-mortgage borrowing to fill the gap $\gamma - H(\ell)$. A common approach in Denmark is to borrow from a bank or occasionally from the seller of the property to bridge funding gaps between 80% and 95% loan-to-value (LTV); this is typically expensive.¹⁹ A third (usually unobservable) possibility is that households can bring additional funds to the table by liquidating other assets.²⁰ We therefore assume that violating the down-payment constraint does not lead the seller to withdraw the sale offer, assuming instead that the seller incurs a monetary penalty of μ per unit of realized home equity below the

¹⁹Danish households can borrow using “Pantebreve” or “debt letters” to bridge funding gaps above LTV of 80%. Over the sample period, this was possible at spreads of between 200 and 500 bp over the mortgage rate. For reference, see categories *DNRNURI* and *DNRNUPI* in the Danmarks Nationalbank’s statistical data bank.

²⁰In Stein (1995), M represents the outstanding mortgage debt net of any liquid assets that the household can put towards the down payment. The granular data that we employ allow us to measure the net financial assets that households can bring to the table to supplement realized home equity. We later verify using these data that our inferences are sensible when taking these additional funds into account.

constraint threshold:²¹

$$\kappa(P(\ell)) = \begin{cases} \mu(\gamma - H(\ell)), & \text{if } H(\ell) < \gamma \\ 0, & \text{if } H(\ell) \geq \gamma \end{cases}. \quad (6)$$

We turn next to describing the data and key estimated moments as a precursor to more rigorous structural estimation of the underlying parameters of the model.

3 Data

Our data span all transactions and electronic listings (which comprise the overwhelming majority of listings) of owner-occupied real estate in Denmark between 2009 and 2016. In addition to listing information, we also acquire information on property sales dates and sales prices, the previous purchase price of each sold or listed property, rich hedonic characteristics of each property, and a range of demographic characteristics of the households engaging in these listings and transactions, including variables that accurately capture households' financial position at each point in time. Furthermore, we merge the data on the entire housing stock captured in the Danish housing register with the listings data to assess the determinants of the extensive margin listing decision for all properties in Denmark over the sample period. This allows us to assess the fraction of the total housing stock that is listed, and to condition observed listing propensities on functions of the predicted sales price, such as the prospective seller's potential gains relative to the original purchase price, or the prospective seller's potential level of home equity in the property.

Our data link administrative datasets from various sources; all data other than the listings data are made available to us by Statistics Denmark. We briefly describe these data below; the online appendix contains detailed information about data sources, con-

²¹i.e.,

$$U(P(\ell)) = \begin{cases} u(P(\ell)) - \mu(\gamma - H(\ell)), & \text{if } H(\ell) < \gamma \\ u(P(\ell)), & \text{if } H(\ell) \geq \gamma \end{cases}.$$

struction, filters, and the process of matching involved in assembling the dataset.

3.1 Property Transactions and other Property Data

We acquire comprehensive administrative data on registered properties, property transactions, property ownership, and hedonic characteristics of properties from the registers of the Danish Tax and Customs Administration (SKAT) and the Danish housing register (Bygnings-og Boligregister, BBR). These data are available from 1992 to 2016. In our hedonic model, described later, we also include the (predetermined at the point of inclusion in the model) biennial property-tax-assessment value of the property that is provided by SKAT, which assesses property values every second year.^{22,23}

Loss aversion and down-payment constraints were originally proposed as explanations for the puzzling *aggregate* correlation between house prices and measures of housing liquidity, such as the number of transactions, or the time that the average house spends on the market. In the online appendix, we show the price-volume correlation in Denmark over a broader period containing our sample period. The plot looks very similar to the broad patterns observed in the US.

3.2 Property Listings Data

Property listings are provided to us by RealView (<http://realview.dk/en/>), who attempt to comprehensively capture all electronic listings of owner-occupied housing in Denmark. We link these transactions to the cleaned/filtered sale transactions in the official property registers. 76.56% of all sale transactions have associated listing data.²⁴ For each property

²²As we describe later, this is the same practice followed by Genesove and Mayer (1997, 2001); it does not greatly affect the fit of the hedonic model, and barely affects our substantive inferences when we remove this variable.

²³Tax-assessed property values are used for determining tax payments on property value. The appendix describes the property taxation regime in Denmark in greater detail including inheritance taxation; we simply note here that there is the usual “principal private residence” exemption on capital gains on real estate, and that property taxation does not have important effects on our inferences.

²⁴We more closely investigate the roughly 25% of transactions that do not have an associated electronic listing. 10% of these transactions can be explained by the different (more imprecise) recording of addresses

listing, we know the address, listing date, listing price, size, and time of any adjustments to the listing price, changes in the broker associated with the property, and the sale or retraction date for the property.

3.3 Mortgage Data

To establish the predicted/potential level of the owner’s home equity in each property at each date, we obtain data on the mortgage attached to each property from the Danish central bank (Danmarks Nationalbank), which collects these data from mortgage banks. The data are available annually for each owner from 2009 to 2016, cover all mortgage banks and all mortgages in Denmark and contain information on the mortgage principal, outstanding mortgage balance each year, the loan-to-value ratio, and the mortgage interest rate. If several mortgages are outstanding for the same property, we simply sum them, and calculate a weighted average interest rate and loan-to-value ratio for the property and mortgage in question.²⁵

3.4 Owner/Seller Demographics

We source demographic data on individuals and households from the official Danish Civil Registration System (CPR Registeret). In addition to each individual’s personal identification number (CPR), gender, age, and marital history, the records also contain a family identification number that links members of the same household. This means that we

in the listing data relative to the registered transactions data. The remaining 15% of unmatched transactions can be explained by: (a) off-market transactions (i.e., direct private transfers between friends and family, or between unconnected households); and (b) broker errors in reporting non-publicly announced listings (“skuffesager”) to boligsiden.dk. We find that on average, unmatched transactions are more expensive than matched transactions. Sellers of more expensive houses tend to prefer the skuffesalg option for both privacy and security reasons.

²⁵The online appendix provides a detailed description of several features of the Danish mortgage market including the conditions under which mortgages are assumable, as well as the effects of the Danish refinancing system (studied in greater detail in Andersen et al. (2020)) on sale and purchase incentives. These features do not materially impact our inferences.

can aggregate individual data on wealth and income to the household level.²⁶ We also calculate a measure of households' education using the average length of years spent in education across all adults in the household. These data come from the education records of the Danish Ministry of Education. We source individual income and wealth data from the official records at SKAT, which hold detailed information by CPR numbers for the entire Danish population.

3.5 Final Merged Data

We only keep transactions for which we can measure both nominal losses and home equity, and since the mortgage data run from 2009 to 2016, this imposes the first restriction on the sample. The sample is further restricted to properties for which we know both the ID of the owner, as well as that of the owner's household, in order to match with demographic information. Transactions data are available from 1992 to the present, meaning that we can only measure the purchase price of properties that were bought during or after 1992.²⁷ We exclude foreclosures (both sold and unsold),²⁸ properties with a registered size of 0, and properties that are sold at prices which are unusually high or low (below 100,000 DKK and above 20MM DKK in 2015, accounting for roughly 0.05% of the total housing stock in Denmark).²⁹ For listings that end in a final sale, we also drop within-family transactions, transactions that Statistics Denmark flag as anomalous or unusual, and transactions where the buyer is the government, a company, or an organization.³⁰

²⁶Households consist of one or two adults and any children below the age of 25 living at the same address.

²⁷In Appendix Table A.2 and Appendix Figure A.39 we further examine properties traded before 1992. Since these properties have no known purchase price, we match them to otherwise similar properties for which we know the purchase price, using two approaches that we describe in the online appendix, with a reasonable success rate. Figure A.39 shows that the main relationships that we find in the main dataset essentially hold in the matched sample using this approach.

²⁸The online appendix describes the Danish foreclosure process in detail.

²⁹We apply this filter to reduce error in our empirical work, because the market for such unusually priced properties is extremely thin, meaning that predicting the price using a hedonic or other model is particularly difficult.

³⁰We apply this filter, as company or government transactions in residential real estate are often conducted at non-market prices—for tax efficiency or evasion purposes in the case of corporations, and

We also restrict our analysis to residential households, in our main analysis dropping summerhouses and listings from households that own more than three properties in total, as they are more likely property investors than owner-occupiers.³¹

In the online appendix, we describe the data construction filters and their effects on our final sample in more detail. Once all filters are applied, the sample comprises 214,508 listings of Danish owner-occupied housing in the period between 2009 and 2016, for both sold (70.4%) and retracted (29.6%) properties, matched to mortgages and other household financial and demographic information.³² These listings correspond to a total of 191,843 unique households, and 179,262 unique properties. Most households that we observe in the data sell one property during the sample period, but roughly 9% of households sell two properties over the sample period, and roughly 1.5% of households sell three or more properties. In addition, we use the entire housing stock, filtered in the same manner as the listing data, comprising 5,540,376 observations of 807,666 unique properties to understand sellers' extensive margin decision of whether or not to list the properties for sale.

3.6 Hedonic Pricing Model

To calculate potential gains \widehat{G} (and potential home equity \widehat{H}), we require a measure of the expected sale price \widehat{P} for each property-year in the data. To arrive at this measure, we estimate a standard hedonic pricing model on our sample of sold listings and use it to predict prices for the full sample of listed properties, including those that are not sold.³³

for eminent domain reasons in the case of government purchases, for example.

³¹Genesove and Mayer (2001) separately estimate loss aversion for these groups of homeowners and speculators. We simply drop the speculators in this analysis, choosing to focus our parameter estimation in this paper on the homeowners.

³²The data comprises 173,065 listings that have a mortgage, and 41,443 listings with no associated mortgage (i.e., owned entirely by the seller)—we later utilize these subsamples for various important checks.

³³Later in the paper, we also assess the extent to which gains, losses, and home equity determine the *decision* to list. We estimate a separate hedonic model on a larger data set, including unlisted properties, in order to conduct these additional tests.

The hedonic model predicts the log of the sale price P_{it} of all sold properties i in each year t :

$$\begin{aligned} \ln(P_{it}) = & \xi_{tm} + \beta_{ft} \mathbb{1}_{i=f} \mathbb{1}_{t=\tau} + \beta \mathbf{X}_{it} \\ & + \beta_{fx} \mathbb{1}_{i=f} \mathbf{X}_{it} + \Phi(v_{it}) + \mathbb{1}_{i=f} \Phi(v_{it}) + \varepsilon_{it}, \end{aligned} \quad (7)$$

where \mathbf{X}_{it} is a vector of property characteristics, namely $\ln(\text{lot size})$, $\ln(\text{interior size})$, number of rooms, bathrooms, and showers, a dummy variable for whether the property was unoccupied at the time of sale or retraction, $\ln(\text{the age of the building})$, dummy variables for whether the property is located in a rural area, or has been marked as historic, and $\ln(\text{distance of the property to the nearest major city})$. (Most property characteristics in \mathbf{X}_{it} are time-varying, which contributes to the accuracy of the model). ξ_{tm} are year cross municipality fixed effects (there are 98 municipalities in Denmark), and $\mathbb{1}_{i=f}$ is an indicator variable for whether the property is an apartment (denoted by f for flat) rather than a house.³⁴ $\Phi(v_{it})$ is a third-order polynomial of the previous-year tax assessor valuation of the property.³⁵ We interact the apartment dummy with time dummies, as well as with the hedonic characteristics and the tax valuation polynomial, to allow for a different relationship between hedonics and apartment prices.

When we estimate the model, the R^2 statistic equals 0.88 in the full sample.³⁶ The

³⁴In the online appendix, we also include cohort effects ξ_c in the hedonic regression, and continue to find robust evidence of all patterns uncovered in our empirical analysis, showing that intra-cohort variation in gains and losses is also associated with changes in listing premia.

³⁵Genesove and Mayer (1997, 2001) also consider tax assessment data in their hedonic model. Importantly, the tax assessment valuation is carried out before the time of the transaction, in some cases even many years before. Until 2013, the tax authority re-evaluated properties every second year. The assessment, which is valid from January 1st each year, is established on October 1st of the prior year. In the years between assessments, the valuation is adjusted by including local-area price changes. This adjustment has been frozen since 2013, recording such price changes as of 2011. Only in the case of significant value-enhancing adjustments to a house or apartment would a re-assessment have taken place thereafter—and once again, is pre-determined at the point of property sale.

³⁶The online appendix contains several details about the hedonic model and estimates. We also estimate the model in levels rather than logs, with an R^2 of 0.89. Moreover, the R^2 when we eliminate the tax assessor valuation from the hedonic characteristics is 0.77. To check the robustness of our results to the specification of the hedonic model, we also amend it in various ways as outlined in the appendix. Our results are qualitatively, and for the most part, quantitatively unaffected by these amendments.

large sample size allows us to include many fixed effects in the model, helping to deliver a better fit. This helps to ameliorate concerns of noise or unobserved quality in the measure \widehat{P} , an important concern when estimating the effects of both loss aversion and home equity (e.g., Genesove and Mayer, 1997, 2001, Anenberg, 2011, Clapp, et al., 2018). We also adopt a number of alternative approaches to deal with the important issue of unobserved quality and its effects on our inferences, as we later describe.

4 First Inferences about Model Parameters

In this section, we document patterns in listing premia and sales transactions volumes in the data in relation to measured G and \widehat{G} , and informally discuss how these patterns relate to the predictions of the model, especially regarding the primary parameters of interest η and λ . We also explore how the patterns in the data and possible inferences about underlying parameters vary when we account for three important factors. These are: (i) sellers' down-payment constraints, (ii) concave demand, and (iii) robustness to changes in measurement. Before turning to structural estimation that takes the model's predictions to the data more rigorously in the next section, we discuss the robustness of the patterns seen in the data to various estimation approaches and controls.

4.1 Listing Premia in the Data

Armed with the hedonic pricing model, we estimate listing premia in the data as $\ell = \ln L - \widehat{\ln P}$, where L is the reported initial listing price observed in the data.³⁷ Mean (median) ℓ is 12.7% (11.3%), and $\ell > 0$ (< 0) for 75% (25%) of the sample. We also estimate potential gains $\widehat{G} = \widehat{\ln P} - \ln R$, where R is set to the nominal purchase price of the property. Mean (median) \widehat{G} estimated in this way is 36% (28%), and 23% (77%) of

³⁷We confirm, estimating Genesove and Mayer's (2001) specifications on our data (see online appendix), that the coefficient on $\widehat{\ln P}$ in our data using their regression, controlling for a range of other determinants, is close to 1. We discuss below how our results are robust to using the alternative approach of Genesove and Mayer (2001), and discuss identification and measurement concerns in greater detail below as well.

property-years have $\widehat{G} < 0$ ($\widehat{G} > 0$). The online appendix plots the distributions of these and other variables.

In Figure 3 we plot the average observed listing premium (on the vertical axis) for each percentage bin of potential gains (on the horizontal axis). Sellers who hold properties that have appreciated (declined in value) since the initial purchase choose lower (higher) listing premia. Importantly, this negative relationship is visible not only in the potential loss domain (i.e., $\widehat{G} < 0$), but also across different values in the potential gain domain (i.e., $\widehat{G} > 0$). This is consistent with the predictions of a model with reference dependence $\eta > 0$. Moreover, as we move from the gain to the loss domain, the slope becomes much more pronounced, i.e., sellers react much more aggressively to every unit decrease in potential returns when $\widehat{G} < 0$. For potential gains in the neighbourhood of zero, this “hockey stick” pattern is consistent with the predictions of a model with loss aversion $\lambda > 1$. However, in the piecewise linear formulation that we consider, loss aversion also predicts a flattening out of the listing premium profile deeper into the loss domain, which is not visible in the plot.

While these patterns provide prima facie evidence of the underlying parameters of the seller’s utility, we must be wary of such inferences given the influence of three important confounding factors discussed above, namely: (i) concave demand, (ii) the extensive margin, which smooths out the locations of kinks, and can lead to selection effects, and (iii) sellers’ financial/down-payment constraints. Keeping these issues in mind, we next discuss additional evidence available from the analysis of transactions volumes.

4.2 Bunching of Realized Sales

Figure 4 plots the distribution of property sales across the dimension of realized gains ($\ln P - \ln R$)—each dot shows the empirical frequency of sales (y-axis) occurring in each 1 percentage point bin of realized gains (x-axis). We overlay on this plot (as a dotted line) the empirical frequency of realized sales (i.e., the same y-axis) occurring in each 1

percentage point of *potential* gains $\widehat{\ln P} - \ln R$ (i.e., a different x-axis). Observing the counterfactual is difficult in most settings which attempt to estimate loss aversion using bunching estimators (e.g., Rees-Jones 2018 cleverly extracts evidence of loss aversion from U.S. tax returns data, where it is difficult to measure “expected tax avoidance costs and benefits”). The distribution of sales with respect to pre-listing potential gains can serve as one possible counterfactual, as we describe in greater detail below.³⁸

Figure 4 shows significant bunching of transactions in the positive domain of realized gains G , with a sharp “spike” around $G = 0$, and with significant mass extending further into the domain $G > 0$.³⁹ While the spike is clearly evident, more information can be extracted about model parameters from the broader distribution of sales across realized gains, especially when we compare it to the distribution of sales across potential gains \widehat{G} . This is because in the model when $\eta > 0$, as mentioned earlier, the mapping between \widehat{G} and G occurs through the choice of ℓ^* , and the associated probability of sale. This mapping results in mass in the final sales distribution shifting towards sales with realized $G > 0$. In contrast, when $\eta = 0$, the model predicts that the distribution of G is simply a constant linear transformation of the distribution of \widehat{G} . The precise position of the pronounced jump in the distribution at $G = 0\%$, and the distribution of mass to the left and right of this point relative to the counterfactual are also informative about λ . When $\lambda > 1$, the model predicts a jump in the final distribution of house sales precisely at $G = 0$, additional mass in this distribution just to the right of this point, and relatively lower mass in the loss domain, to the left of $G = 0$. The pronounced bunching that we observe precisely at the point $G = 0$ also offers empirical support (which is essentially

³⁸We also use alternative approaches to measure this counterfactual density, following Chetty et al. (2011) and Kleven (2016), and fitting a flexible polynomial to the empirical frequency distribution. When doing so, we exclude bins near the threshold, and extrapolate the fitted distribution to the threshold, excluding one bin on each side of the zero gain bin, i.e. $j \in \{-1\%, 1\%\}$, with a polynomial order of 7. The results, reported in the Online Appendix, are robust to other polynomial orders and to variations of the excluded range, and generate similar (but less cleanly estimated) results on the excess bunching mass.

³⁹The plot also reveals a small but visible “hole” just to the left of $G = 0$, that may be evidence of a notch in preferences—an important additional feature of the data that we are currently investigating.

non-parametric, since it does not require reliance on a hedonic or other model) for the choice of R as the nominal purchase price (see Kleven, 2016, for a discussion of bunching at reference points).

4.3 Extensive Margin: Probability of Listing

As discussed earlier, understanding the seller’s decision of whether or not to list is important for at least two reasons. First, the model makes predictions about this decision, in addition to predicting patterns of listing premia and transactions volumes. Second, accounting for this decision helps to correct for possible selection effects that may drive patterns of observed intensive margin listing premia in the data. This is an issue that the prior literature (e.g., Genesove and Mayer, 1997, 2001, Anenberg, 2011, Guren, 2018) has been unable to control for as a result of data limitations.

To understand the decision to list, we turn to data on the total housing stock in Denmark, corresponding to 12, 565, 190 property-years in the data, once merged with the listings data. We compute the unconditional average annual listing propensity, which is 3.75% of the housing stock (corresponding to between 2% and 4% of the housing stock listed across sample years).⁴⁰ Figure 8 plots the listing propensity at each level of \widehat{G} , which comes from estimating $\widehat{\ln P}$ for all properties in Denmark for which we have data on the nominal purchase price R . The figure shows a mild, but visible increase in the probability of listing as \widehat{G} increases, which is evident when $\widehat{G} > 0$, but more pronounced when $\widehat{G} < 0$. This pattern is once again apparently consistent with levels of $\eta > 0$ and $\lambda > 1$.

⁴⁰We do not attempt to use the model to explain the average propensity to list, as this exercise is beyond the scope of this paper. It would require us to take a strong stance on the factors that drive the moving decision, which we currently summarize using our estimates of θ .

4.4 Confounding Factors

4.4.1 Down-payment Constraints and Home Equity

To account for the role of down-payment constraints, for each observation in the data, we calculate the seller’s potential home equity level $\widehat{H} = \widehat{\ln P} - \ln M$, where $\widehat{\ln P}$ is estimated using our hedonic model as before, and M is the outstanding mortgage balance reported by the household’s mortgage bank each year.⁴¹ Mean (median) \widehat{H} is 27% (25%), and 77% (23%) of property-years have $\widehat{H} < 0$ ($\widehat{H} \geq 0$). Modal \widehat{H} is around 22%, which is to be expected, as Denmark has a constraint on the issuance of mortgages—the Danish Mortgage Act specifies that LTV at issuance by mortgage banks is restricted to be 80% or lower.⁴² Clearly, \widehat{G} and \widehat{H} are jointly dependent on $\widehat{\ln P}$, but there are multiple other factors that influence this correlation, including the LTV ratio at origination (i.e., variation in initial down payments), and households’ post-initial-issuance remortgaging decisions. In the online appendix, we plot the joint distribution of \widehat{G} and \widehat{H} , and show that there is substantial variation in the four regions defined by $\widehat{G} \leq 0$ and $\widehat{H} \leq 0$, which permits identification of their independent impacts on listing decisions.⁴³

To assess the extent to which any variation in ℓ attributed to \widehat{G} might be confounded

⁴¹The online appendix plots the distributions of \widehat{G} and \widehat{H} in the data. Both \widehat{G} and \widehat{H} are winsorized at the 1 percentile point; \widehat{G} is also winsorized at the 99 percentile point. We winsorize \widehat{G} because of several large values of given the substantial time elapsed since the purchase of some properties in the data. We set \widehat{H} to 100% in cases in which households have substantial home equity ($\geq 60\%$), meaning that we consider households to be essentially unconstrained at high levels of home equity. This is necessary to avoid \widehat{H} levels greater than 1, given the log difference approach that we use to compute it. These filters make no material difference to our results—we confirm that our structural estimates are unaffected by these choices.

⁴²This constraint does not change over our sample period, though it must be noted that as mentioned earlier, households can engage in non-mortgage borrowing to effectively increase their LTV, but at substantially higher rates. The online appendix documents the changes in the Danish Mortgage Act over the 2009 to 2016 sample period. While the constraint does not move during this period, there are a few changes in the wording of the act, a change in the maximum maturity of certain categories of loans in February 2012 from 35 to 40 years, and the revision of certain stipulations on the issuance of bonds backed by mortgage loans. None of these materially affect our inferences.

⁴³The online appendix also contains a fuller discussion of additional evidence that we uncover which is consistent with households exhibiting aversion to downsizing. We are able to link sale transactions with future purchase transactions for a subset of households, and show that the future purchase is almost always of higher value than the sale.

by simultaneous variation in \widehat{H} , the top left plot in Figure 5 shows a 3-D representation of ℓ against both \widehat{G} and \widehat{H} in the data, averaged in bins of 3 percentage points. The plot reveals that ℓ declines in both \widehat{G} and \widehat{H} , consistent with the patterns previously identified in the literature. Unusually, given the large administrative dataset that we have access to, the plot captures the variation ℓ along both dimensions simultaneously, and clearly reveals both *independent* and *interactive* variation along both dimensions. To better see the independent variation, the dotted lines on the 3-D surface indicate two cross-sections in the data ($G = 0\%$ and $H = 20\%$), which we also use later for structural estimation. Clearly, the “hockey stick” profile of ℓ along the \widehat{G} dimension survives, controlling for \widehat{H} , and there is also a pronounced downward slope in ℓ along the \widehat{H} dimension, controlling for \widehat{G} . In terms of the interactive variation, Panel B of Figure 9 plots how the “marginals” of the listing premium vary as we vary the control variable in each case (i.e., \widehat{H} in the left plot and \widehat{G} in the right plot); we discuss these in more detail towards the end of the paper, where we also evaluate the extent to which we can match these relationships using the model.⁴⁴

4.4.2 Concave Demand

Using the underlying data on the time-on-the-market (TOM) that elapses between sale and listing dates, the left plot in Figure 6 calculates the probability of a house sale within six months (this maps to $\alpha(\ell)$ in the model), which we plot on the y-axis, as a function of ℓ on the x-axis.⁴⁵ To smooth the average point estimate at each level of ℓ , we use a simple nonlinear function which is well-suited to capturing the shape of $\alpha(\ell)$, namely, the generalized logistic function or GLF (Richards, 1959, Zwietering et al., 1990, Mead,

⁴⁴The online appendix reports sale transaction

frequencies (to show the degree of bunching) in a similar 3-D fashion. We confirm that regardless of the level of \widehat{H} , there is a visible shift of mass from the $\widehat{G} < 0$ domain to the $\widehat{G} > 0$ domain.

⁴⁵Mean (median) TOM in the data is 37 weeks (25 weeks). We pick six months in the computation of $\alpha(\ell)$ to match the median TOM observed in the sample. The online appendix shows the distribution of TOM, which is winsorized at 200 weeks, meaning that we view properties that spend roughly 4 years on the market as essentially retracted.

2017).⁴⁶ The solid line corresponds to this set of smoothed point estimates.

The right-hand plot in Figure 6 shows how $\ln P(\ell) - \widehat{\ln P}$, i.e., the “realized premium” of the final sales price over the hedonic value (which corresponds to the “markup” $\beta(\ell)$ in the model) varies with ℓ . The plot shows that $\beta(\ell)$ rises virtually one-for-one with ℓ when ℓ is low, but flattens out as ℓ rises. The solid line shows a simple polynomial fit of this relationship that we use in the model.

From the two plots, we can see that in Denmark low list prices appear to reduce seller revenue with little corresponding decline in time-on-the-market. This is virtually identical to the patterns detected by Guren (2018) in three U.S. markets, which he terms “demand concavity”.⁴⁷

This evidence of demand concavity serves as a confound for estimating λ , as described earlier. This is because the model predicts two possible and distinct sources of the differential slopes of ℓ^* across gains and losses. One is that in the presence of loss aversion (i.e., $\lambda > 0$), there are kinks in ℓ^* around $\widehat{G} = 0$, which can be smoothed into a differential slope by variation in θ . The second is buyer sensitivity to ℓ , i.e. the degree of demand concavity $\alpha(\ell)$. The top panel of Figure 6 illustrates this second mechanism in the model, which predicts that sellers set a steeper ℓ^* slope when $\widehat{G} < 0$ in markets where $\alpha(\ell)$ demand is *less* steeply sloped and vice versa. This predicts a tight correlation between the slope of $\alpha(\ell)$ and the slope of ℓ when $\widehat{G} < 0$, which cannot be seen in Figure 6, which is estimated using the entire dataset. To estimate the impact of demand concavity on the shape of the listing premium “hockey stick,” we therefore exploit regional variation across sub-markets of the Danish housing market.

To illustrate the predicted correlation between the shape of the listing premium

⁴⁶We describe the GLF in more detail in the online appendix. It is useful for our purposes as it is (i) bounded both from above and below, and it (ii) allows us to easily capture the degree of concavity observed in the data in a convenient way, through a single parameter. In our estimation of the parameters, we restrict the lower bound of the GLF to be equal to zero, to impose that the probability of sale asymptotically converges to 0 for arbitrary high levels of ℓ .

⁴⁷These plots also show that Danish sellers who set high ℓ suffer longer TOM, but ultimately achieve higher prices (i.e., high realized premia) on their house sales, confirming the original finding of Genesove and Mayer (2001), who analyze the Boston housing market between 1990 and 1997.

“hockey stick” and the degree of demand concavity (i.e., the shape of $\alpha(\ell)$) in the data, we separately estimate the slope of ℓ in the domain $\widehat{G} < 0$, as well as separate $\alpha(\ell)$ functions (in particular, the slope of $\alpha(\ell)$ when $\ell \geq 0$) in different local housing markets, namely, different municipalities of Denmark.⁴⁸

The bottom panel of Figure 7 shows results when we sort municipalities by their estimated demand concavity (i.e., the slope of $\alpha(\ell)$ when $\ell \geq 0$). The right-hand panel of the plot illustrates that there is indeed substantial variation in demand concavity across municipalities, showing municipalities in the top and bottom 5% of estimated demand concavity. The slope for municipalities with strong demand concavity (top 5%) lies between -1.4 and -1.1 , while the slope for municipalities with weak demand concavity (bottom 5%) lies between -0.1 and -0.3 . The left-hand plot in Figure 7 Panel A shows the corresponding figure for the relationship between $\hat{\ell}$ and \widehat{G} for these municipalities. Indeed, as the model predicts, markets with strong demand concavity exhibit a substantially weaker slope of ℓ in the domain $\widehat{G} < 0$ (-0.1 to -0.4) than markets with weak demand concavity (-0.5 to -0.9).⁴⁹ Towards the end of the paper, we describe a validation analysis that we undertake to confirm the model-predicted mechanism in the data using instruments for demand concavity.

⁴⁸Municipalities are a natural local market unit—there are 98 in Denmark, each of around 60,000 inhabitants, and with roughly 1,800 listings on average. We also re-do this exercise using shires, which are a smaller geographical delineation covering 80 listings on average as a cross-check.

⁴⁹For the purposes of our current investigation, we focus on the slope differentials, and to show these, Figure 7 normalizes sub-markets to have the same level of the listing premium. We also observe important differences between the *levels* of $\alpha(\ell)$ across these markets i.e., there are both “hot” and “cold” municipalities à la Ngai and Tenreyro (2014). Un-normalized plots in the online appendix reveal that the *level* of ℓ is lower when the level of $\alpha(\ell)$ is higher and vice versa; and consistent with Ngai and Tenreyro (2014), the levels of $\alpha(\ell)$ and $P(\ell)$ are strongly positively correlated across sub-markets.

4.5 Robustness

4.5.1 Time-series Variation

While it is reassuring that \hat{G} and \hat{H} exhibit independent variation in the data, it could well be the case that this variation is confined to one particular part of the sample period, i.e., driven by time-variation in aggregate Danish house prices. To check this, in the online appendix we plot the shares of the data in each of four groups of properties defined by $\hat{G} \leq 0$ and $\hat{H} \leq 0$, in each of the years in our sample. We find that aggregate price variation does shift the relative shares in each group across years, with price rises increasing the fraction of unconstrained winners ($\hat{G} > 0$ and $\hat{H} > 0$) relative to losing and constrained groups. However, the relative shares of all four groups are substantial and fairly stable over the sample period, alleviating concerns that different groups simply come from different time periods, i.e., the plots is reassuring that identification of any effects is likely to arise mainly from the cross-section rather than the time-series. We also verify that the inclusion of cohort and cohort-cross-municipality fixed effects in the hedonic model does not affect our inferences materially.

4.5.2 Bunching: Round Numbers and Holding Periods

In the online appendix, we verify that the bunching patterns documented earlier are robust to commonly expressed concerns in this literature (e.g., Kleven 2016, Rees-Jones 2018). We find that the spike in sales volumes at $G = 0$ and the patterns of excess mass relative to the counterfactual do not appear to be driven by bunching at round numbers, as they remain striking and visible when we exclude sales at prices ending in multiples of 10,000, 50,000, 100,000, and 500,000 DKK, which (cumulatively) affect roughly 20%, 17%, 5%, and 2% of all observations, respectively. We also show that these bunching patterns are robust when we split the sample into five groups (< 3 , $3 - 6$, $6 - 9$, $9 - 12$, > 12 years) based on the time between sale and purchase, i.e., the holding period of the

property. Except for the sub-sample with the longest holding period (> 12 years, 20% of the data), we find strong evidence of bunching. Finally, we also find strong evidence of bunching in all cases when we split the sample into quintiles based on the level of R , with quintile cutoffs ranging from around 658,000 DKK to 1.9MM DKK. Together, these checks assuage concerns that bunching could result from differences in the underlying characteristics of properties—for instance these tests suggest that it is implausible that bunching results from a combination of small properties with shorter holding periods clustering around $G = 0$, and larger properties with longer holding periods showing up at values of $G > 0$.

4.5.3 Unobserved Quality

An important and often-repeated concern in the literature is that the relationships that we observe between ℓ and \widehat{G} (and indeed $\alpha(\ell)$ and ℓ), can be spuriously affected by measurement error in the underlying model for \widehat{P} . In particular, if properties with $\widehat{G} < 0$ are deemed to be such as a result of underestimated \widehat{P} , we would also see higher listing premia for such properties, resulting in the hockey-stick shape that we observe. Moreover, such an issue could also upwardly bias the true (decreasing) relationship between the probability of a quick sale and ℓ , especially when $\ell > 0$, as houses with mismeasured high listing premia would be expected to transact faster.

We assess the robustness of our results to these concerns in a number of ways, all of which we describe in detail in the online appendix. First, we show that the relationships between ℓ , \widehat{G} , and $\alpha(\ell)$ are robust to a battery of changes to the underlying model used to estimate \widehat{P} . We do so in several ways. We employ a repeat sales model to difference out time-invariant unobserved property quality; we instrument variation in \widehat{P} using regional house price indices; we control for demographics, financial wealth, and further interactions in the hedonic model using granular data that have previously been unutilized in this manner, and which are potentially informative about the seller’s response to earlier under-

or over-payment on the property; we use the external tax assessor value of the property instead of our estimated hedonic model; and finally, we verify that our inferences hold even when we use out-of-sample estimated hedonic coefficients.⁵⁰

Second, we implement the bounding approach proposed in Genesove and Mayer (2001) to account for unobserved quality, and confirm that our inferences are robust to doing so.

Third, while the tests just described focus on showing robustness of the magnitudes of the nonlinear relationships observed in the data between ℓ , \widehat{G} , and $\alpha(\ell)$, we also document evidence in line with the key predictions from the model. That is, we are able to demonstrate that the observed nonlinearities are in fact discontinuous and sharp around the respective thresholds of $\widehat{G} < 0$ and $\ell > 0$, using a regression kink design (RKD) originally suggested by Card et al. 2015b and implemented e.g., by Landais, 2015, Nielsen et al. 2010, and Card et al. 2015a. In line with the identifying assumptions of this research design, we also show that property-and household-specific observable characteristics are smooth around the respective thresholds.

5 Structural Estimation

5.1 Moments in the Model

To match the data moments inside the model, we make a few assumptions. First, we simply use the estimated demand concavity $\alpha(\ell)$ and $P(\ell)$ shown in Figure 6 as two of these inputs. Second, we set $\gamma = 20\%$ according to Danish law. Third, we normalize all quantities in the model, setting the property’s fundamental value $\widehat{P} = 1$ and we set the outside option $u = \widehat{P}$. Fourth, we define the variables $\widehat{G} = \widehat{P} - R$ and $\widehat{H} = \widehat{P} - M$ as the model equivalents of potential gains and home equity in the data.

⁵⁰This last variation helps to assuage concerns of overfitting or mechanical correlation arising from our hedonic model being estimated using the sample of sold listings. The model fits relatively precisely out of sample, with R^2 ’s ranging between 0.80 to 0.88 when predicting between 1% to 50% of the data out-of-sample, and the patterns in the relationships between ℓ , \widehat{G} , and $\alpha(\ell)$ are robust to using the oos coefficient estimates.

Next, consider the set of parameters from the model:

$$\mathbf{x} = \left[\eta, \lambda, \delta, \mu, \theta_{\min}, \theta_{\max}, \varphi \right]'. \quad (8)$$

To obtain policy functions of state variables and parameters, we solve the model numerically, inputting grids of \widehat{G} and \widehat{H} , and yielding:

$$\left[s^*(\widehat{G}, \widehat{H}, \theta, \mathbf{x}), \ell^*(\widehat{G}, \widehat{H}, \theta, \mathbf{x}) \right] = \arg \max_{s \in \{0,1\}} \left\{ (s) \max_{\ell} \left\{ EU(\ell, \widehat{G}, \widehat{H}, \theta, \mathbf{x}) \right\} + (1-s)\underline{u} \right\}. \quad (9)$$

We then compute aggregates, i.e., averages in the population of listing probabilities, and average listing premia which account for the extensive margin decision:

$$S^*(\widehat{G}, \widehat{H}, \mathbf{x}) = \int s^*(\widehat{G}, \widehat{H}, \theta, \mathbf{x}) d\theta, \quad (10)$$

$$\mathcal{L}^*(\widehat{G}, \widehat{H}, \mathbf{x}) = \int_{s^*=1} \ell^*(\widehat{G}, \widehat{H}, \theta, \mathbf{x}) d\theta. \quad (11)$$

These functions then allow us to compute the set of seven model-implied moments $\mathbf{M}_m(\mathbf{x})^{7 \times 1}$ corresponding to the moments in the data $\mathbf{M}_d^{7 \times 1}$ described above.

The first moment is the average listing premium $\mathcal{L}^*(\widehat{G} = 0\%, \widehat{H} = 20\%, \mathbf{x})$. The second is a slope from regressing $\mathcal{L}^*(\widehat{G}, \widehat{H} = 20\%, \mathbf{x})$ on the grid of \widehat{G} for $\widehat{G} < 0$. The third is a slope from regressing $\mathcal{L}^*(\widehat{G} = 0\%, \widehat{H}, \mathbf{x})$ on the grid of \widehat{H} for $\widehat{H} < 20\%$.

We next propose a simple procedure to approximate the regional correlation moments (i.e., the relationship between variation in demand concavity and the slope of the listing premium) inside the model. Let $\kappa_{\widehat{G} < 0}$ be the slope from a regression of $\mathcal{L}^*(\widehat{G}, \widehat{H} = 20\%, \mathbf{x})$ on the grid of \widehat{G} for $\widehat{G} < 0$, and $\kappa_{\widehat{G} \geq 0}$ the analogous slope for $\widehat{G} \geq 0$ ($\kappa_{\widehat{G} < 0}$ and $\kappa_{\widehat{G} \geq 0}$ simply capture the slopes of the listing premium above and below potential gains of zero). Now consider a change $\tilde{\delta}$ in demand concavity. We re-compute each of the κ slopes for $\delta - \frac{\tilde{\delta}}{2}$ and $\delta + \frac{\tilde{\delta}}{2}$, which is a first-order approximation of the degree to which a change in

concave demand “passes through” to the slopes of \mathcal{L}^* above and below $\widehat{G} = 0\%$. The fourth and fifth moments inside the model are then given by $\frac{\kappa_{\widehat{G}<0}^+ - \kappa_{\widehat{G}<0}^-}{\delta}$ and $\frac{\kappa_{\widehat{G}\geq 0}^+ - \kappa_{\widehat{G}\geq 0}^-}{\delta}$.

The sixth moment measures bunching of transactions around realized gains of zero. To calculate this measure, we begin with a randomly generated sample of $N = 1,000$ draws of \widehat{G} from a uniform distribution with limits $(-50\%, +50\%)$. For each observation in the sample, we obtain the optimal aggregate listing premium \mathcal{L}^* for a level of home equity equal to 20% and the average level of the moving shock, and calculate realized gains as $G = P(\mathcal{L}^*) - R$. In addition, we model the likelihood that the transaction goes through by drawing a random number ϵ from a uniform distribution and including the observation in the final sample of transactions if $\epsilon < \alpha(\mathcal{L}^*)$. The measure of bunching is then given by the relative density of transactions in the positive vs. the negative domain, in the interval $[-5\%, +5\%]$.⁵¹

Finally, the seventh moment is given by the slope from a regression of $S^*(\widehat{G}, \widehat{H} = 20\%, \mathbf{x})$ on the grid of \widehat{G} , to match the corresponding extensive margin moment in the data.

5.2 Classical Minimum Distance Estimation

From the moments in the data and in the model, we calculate:

$$g(\mathbf{x}) = M_m(\mathbf{x}) - M_d.$$

Since the system is exactly identified, i.e., seven moments and seven parameters, we can estimate the structural parameters $\widehat{\mathbf{x}}$ simply as:

$$\widehat{\mathbf{x}} = \arg \min_{\mathbf{x}} g(\mathbf{x})'g(\mathbf{x}).$$

⁵¹We choose this slightly wider interval than in the data to avoid situations in which our results may be influenced by the grid sizes.

The asymptotic variance of the parameters is given by:

$$\overline{avar}(\hat{\mathbf{x}}) = \left[\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} \overline{W} \frac{\partial g(\mathbf{x})}{\partial \mathbf{x}'} \right]^{-1},$$

where we set \overline{W} to the inverse of the normalized covariance matrix of moments \mathbf{x} . We consider both a simple (diagonal) case: $\overline{W}_{ii} = (\sigma_i^2/N_i)^{-1}$, as well as the (shire-clustered) bootstrap full covariance matrix. Finally, we make inferences about the parameter estimates using the asymptotic relationship:

$$\hat{\mathbf{x}} \rightarrow^d N(\mathbf{x}, \overline{avar}(\hat{\mathbf{x}})).$$

5.3 Parameter Estimates

Table 2 shows the estimated parameters and associated standard errors. The data favor a model of reference dependence with $\eta = 0.948$ with a degree of loss aversion $\lambda = 1.576$. This λ estimate is lower than that commonly considered in the early literature, which lies between 2 and 2.5 (e.g., Kahneman et al. 1990, Tversky and Kahneman, 1992), but is closer to estimates reported in more recent literature (e.g., Imas et al. 2016 finds a value of $\lambda = 1.59$).⁵²

The parameter $\mu = 1.060$ best matches the average $\hat{\ell}$ slope with respect to \hat{H} , i.e., there is an 106 bp penalty (expressed as a fraction of the mortgage amount) for every percent that H drops below $\gamma = 20\%$. This parameter can be contrasted with an average rate increase of roughly 50 bp on the whole loan if the household were to borrow an additional 10% in the unsecured Danish lending market.⁵³ The relatively larger number

⁵²Given how close the estimated η is to 1, we re-estimated a restricted version of the model where $\eta = 1$. Further details are discussed in the online appendix. We obtained similar estimates of $\lambda = 1.522$ (*s.e.* 0.479), $\mu = 1.158$ (*s.e.* 0.218), $\delta = -0.093$ (*s.e.* 0.0183), $\theta_{\min} = 0.235$ (*s.e.* 0.148), $\theta_{\max} = 1.052$ (*s.e.* 0.131) and $\varphi = 0.039$ (*s.e.* 0.025).

⁵³Households in this market face between 200-500 basis points increases in interest rates for every percentage point of borrowing in this market between 80 and 95 LTV over our sample period. Taking 450 bp as the point estimate within this range, at an 80% LTV an additional ten percent borrowing adds roughly 50 bp to the overall loan.

suggests that households in Denmark faced financial constraints preventing them from borrowing. In support of this, we find that the median household in our sample has *negative* net liquid financial wealth of roughly -9% , i.e., their unsecured debt is greater than their liquid financial assets (stocks, bonds, cash) by this amount.

We find that $\delta = -0.097$, which corresponds to a perceived relative reduction of the probability of sale of 9.7% , for a household listing at $\ell = 10\%$, and that the distribution $\theta \sim \text{Uniform}(\theta_{\min}, \theta_{\max})$ has parameters $\theta_{\min} = 0.217$ and $\theta_s = 1.005$. These “moving shocks” correspond to the present discounted value of future benefits from successfully selling and/or moving, and are on the order of 21.7% of the hedonic price for a household at the minimum of the distribution, and approximately equal to the entire hedonic value for a household at the maximum of the distribution. Finally, we find that the estimated “all-in” cost of listing is 3.7% of the hedonic value of the house.

Andrews et al (2017) argue that in method-of-moments estimation of the type that we use, it is often useful to understand the mapping from moments to estimated parameters. In the online appendix we propose a simple and less formal application of this idea, describing how each moment varies when we re-compute the model-implied moments varying each of the structural parameters by two standard deviations. This also provides useful intuition on the sources of identification in the data for each of the model’s parameters. We also evaluate the importance of correctly modelling demand concavity. We do so by adopting a naïve approach to estimation that eschews this important feature and simply assumes that demand is linear. To do so, we preserve the $P(\ell)$ function, but simply estimate a linear $\alpha(\ell)$ function, and re-estimate the parameters (apart from δ) under this assumption. We find that in the case of this restricted model, we estimate $\eta = 0.750$ with a degree of loss aversion $\lambda = 3.285$, a radical departure from the more realistic estimates that we extract when demand is permitted to be concave.

6 Validation and Open Questions

6.1 Interactions

The top panel of Figure 9 compares the 3-dimensional patterns of optimal listing premia in the data (left-hand plot) and the model (right-hand plot). The model matches the pronounced increase in $\hat{\ell}$ for $G < 0$, and the similar increase in $\hat{\ell}$ when \hat{H} declines. A striking feature of this plot is that it seems to indicate that the position of any reference point is not uniquely determined by \hat{G} or \hat{H} alone. As we briefly mentioned earlier, there is considerable variation in the slope of the relationship between $\hat{\ell}$ and both \hat{G} and \hat{H} that depends on the level of the other variable. Put differently, both in the data and in the model, it appears as if the effects of losses and constraints interact with one another, and that the factors affecting household behavior are neither one nor the other variable in isolation.

The bottom panel of Figure 9 explores these interaction effects in more detail. We plot selected cross-sections of the listing premium surface in the data, using a smooth function of the bins for ease of visualization as dashed lines, alongside their model equivalents as solid lines.⁵⁴ The left-hand plot in the bottom panel shows that there is a change in the slope of the $\ell - \hat{G}$ relationship as \hat{H} varies, and the right-hand plot, that there seems to be a change in the inflection point in the $\ell - \hat{H}$ relationship as \hat{G} varies. Note that the average level of ℓ in the data declines substantially as households become less constrained, and increases substantially as households become more constrained—this is simply the unconditional relationship between ℓ and \hat{H} , seen in a different way in the left-hand plot. What is more interesting is that controlling for this change in level, the *slope* of ℓ as a function of \hat{G} is affected by the level of \hat{H} . The important new fact is that down-payment-unconstrained households exhibit seemingly greater levels of reference

⁵⁴We simply use the GLF function for this purpose. The online appendix shows a plot of the actual bins in the data alongside the model-implied listing premia.

dependence along the gain/loss dimension, exhibiting a pronounced increase in the slope to the left of $\widehat{G} = 0$. In contrast, down-payment constrained households exhibit a flatter ℓ along the \widehat{G} dimension. The right-hand plot in the bottom panel of the figure shows the ℓ - \widehat{H} relationship, where again, the level differences reflect the ℓ - \widehat{G} relationship. Another interesting fact emerges—along the \widehat{H} dimension, while the slope around the threshold does not change, *the position of the kink in ℓ increases* with the level of past experienced gains.

These new facts appear to require a more intricate model of preferences and/or constraints than the literature has thus far proposed, which cannot be rationalized by our canonical model, which captures many of the forces thus far proposed in the literature. We briefly speculate on the possible types of models that may rationalize these findings here, with a view towards motivating theoretical work on a broader class of preference and constraint specifications.

One possible rationalization of the variation in the ℓ - \widehat{G} relationship with \widehat{H} is that the luxury of being unconstrained appears to cause more psychological motivations such as loss aversion to come to the fore. Put differently, unconstrained households seem constrained by their loss aversion à la Genesove and Mayer (2001), while constrained households respond to their real constraints by engaging in “fishing” behavior à la Stein (1995). It may also be that this finding can be rationalized by a more complex specification of reference points such as expectations-dependent reference points (e.g., Köszegi and Rabin, 2006, 2007, and Crawford and Meng, 2011).

Turning to the change in the position of the kink in the ℓ - \widehat{H} relationship as \widehat{G} varies, it appears as if a household’s propensity to engage in “fishing” behavior kicks in at a level of \widehat{H} that is strongly influenced by their expected \widehat{G} . One possible rationalization of this is that households facing nominal losses feel constrained at levels of home equity (i.e., $H = 20\%$) that would force them to downsize, while those expecting nominal gains may have in mind a larger “reference” level of housing into which they would like to upsize

(or indeed, a larger fraction of home equity in the next house). To achieve this larger reference level of housing, they begin “fishing” at levels of $H > 20\%$ in hopes of achieving the higher down payment on the new, larger house. To provide suggestive evidence on this story, in the online appendix we focus on a sample of 14,440 households for which we can find two subsequent housing transactions and mortgage down payment data. For this limited subsample, we show a binned scatter plot of the ℓ on the subsequently sold listing against the realized down payment on the subsequent house, controlling for the level of \widehat{H} on the subsequently sold listing. We find evidence that the down payment on the new house is correlated with ℓ , which, given our evidence of \widehat{G} predicting ℓ , is consistent with the idea that households shifting their reference level of housing on the basis of expected gains.

6.2 Demand Concavity, Housing Stock Homogeneity, and Listing Premia

Earlier, we documented how regional variation in demand concavity correlates with regional variation in the shape of the listing premium schedule. This relationship could be driven by a number of different underlying forces. For instance, demand may respond to primitive drivers of supply rather than the other way around—i.e., some markets may be populated by more loss-averse sellers, and buyer sensitivity to ℓ^* might simply accommodate this regional variation in preferences. Another possibility is that this regional relationship simply captures the different incidence of common shocks to demand and market quality.

Our model is partial equilibrium, and describes a different underlying mechanism for this correlation, namely, that sellers are optimizing in the presence of the constraints imposed by demand concavity. In order to understand whether the left-hand plot of Panel B of Figure 7 is potentially consistent with sellers responding to such incentives, we

implement an instrumental variables (IV) approach. Our IV approach is driven by the intuition that the degree of demand concavity is related to the ease of value estimation and hence price comparison for buyers. Intuitively, a more homogeneous “cookie-cutter” housing stock can make valuation more transparent, and should therefore increase buyers’ sensitivity to ℓ . That is, this intuition predicts that markets with high homogeneity should exhibit more pronounced demand concavity.

Our main instrument is the share of apartments and row houses listed in a given sub-market. Row houses in Denmark are houses of similar or uniform design joined by common walls, and apartments have less scope for unobserved characteristics such as garden sheds and annexes than regular detached houses.⁵⁵ As an alternative, we also use the distance (computed by taking the shire-level distance to the closest of the four cities, averaged over all shires in a given municipality) to the four largest cities in Denmark (Copenhagen, Aarhus, Odense, and Aalborg) as a measure of how rural a given market is, and how far away from cities people live on average. This alternative relies on the possibility that homogeneous housing units are more likely to be built in suburbs or in cities, rather than in the countryside.

In the case of both instruments, the identifying assumption is that these measures of homogeneity of the housing stock only affect the slope of $\hat{\ell}$ with respect to \hat{G} through their effect on $\alpha(\hat{\ell})$. To account for cross-market differences in model-predicted demand-side factors affecting the slope of ℓ with respect to \hat{G} and \hat{H} , we also include specifications which control for the average age, education length, financial assets, and income of sellers in a given sub-market.

Figure 7 on the right-hand side of Panel B shows strong evidence of the “first-stage” correlation, i.e., demand concavity on the y-axis against homogeneity measured by the share of apartments and row-houses in a given municipality on the x-axis, with each dot representing a municipality (more negative values of demand concavity mean a sharper

⁵⁵In the online appendix, we show pictures of typical row houses in Denmark.

slope of $\alpha(\ell)$ to the right of $\ell = 0$). Table 3 reports the results of the more formal IV exercise. Column 1 shows the simple OLS relationship between the slope of ℓ for $\widehat{G} < 0$ on demand concavity slope (slope of $\alpha(\ell)$ for $\ell \geq 0$) across municipalities,⁵⁶ with a baseline level of -0.407 . Column 2 uses the apartment-and row-house share as an instrument for demand concavity, and the just identified two-stage least squares (2SLS) specification yields a coefficient estimate of -0.520 . With both instruments (i.e., including the distance to the largest cities as well), the overidentified 2SLS specification gives a result of -0.504 without, and -0.346 with controls for average household characteristics in the municipality. The first-stage F-statistics are between 17 and 25, assuaging weak-instrument concerns (Stock and Yogo, 2002) and we cannot reject the null of the Hansen overidentification test of a correctly specified model and exogenous instruments at conventional significance levels.⁵⁷ These results appear to validate the mechanism that we propose in the model.

7 Conclusion

We structurally estimate a new model of house listing decisions on comprehensive Danish housing market data, and acquire new estimates of key behavioral parameters and household constraints from this high-stakes household decision. Underlying preferences seem well characterized by reference dependent around the nominal purchase price plus modest loss aversion, and there is also evidence of the important role of down-payment constraints on household behavior.

The model cannot completely match some new facts which we identify in the data, which we view as a new target for behavioral economics theory. Nominal losses and down-

⁵⁶Municipalities are required to have at least 30 observations where $\widehat{G} < 0$, leaving 95 out of 98 municipalities, but results are robust to keeping all municipalities.

⁵⁷These results are robust to using a logit model, different cutoffs ($\ell \geq 5, 10, 15\%$) for the demand concavity estimation, cuts of the data such as excluding the largest cities Copenhagen and Aarhus, and regressions at the shire level. These robustness checks are all available in the online appendix.

payment constraints interact with one another, in the sense that reference-dependent behavior is less evident when households are facing more severe constraints, and most pronounced for unconstrained households. Home equity constraints also appear to loom larger for households facing nominal losses. However, for households facing nominal gains, there is evidence consistent with an upward shift in the point at which they feel constrained. This could be explained by households resetting their desired size or quality of housing upwards in response to experienced gains.

In micro terms, this interaction between reference dependence and constraints could have implications for the way we model behavior. We tend to assume that agents optimize their (potentially behavioral) preferences subject to constraints, and in numerous models, agents may also wish to impose constraints on themselves to “meta-optimize” (Gul and Pesendorfer, 2001, 2004, Fudenberg and Levine, 2005, Ashraf et al. 2006, DellaVigna and Malmendier 2006). However, if constraints affect the incidence of behavioral biases, or indeed, if being in a zone that is more prone to bias affects the response to constraints, our models must of necessity become more complicated to accommodate such behavior. From a more macro perspective, reference dependence appears important for understanding aggregate housing market dynamics. The housing price-volume correlation tends to fluctuate, and especially during housing market downturns, prices and liquidity can move in lockstep. This has important implications for labor mobility, which responds strongly to housing “lock” (Ferreira et al., 2012, Schulhofer-Wohl, 2012). Interaction effects such as the effect of expected losses on the household response to constraints could also help to make sense of the seemingly extreme reactions of housing markets to apparently small changes in underlying prices, and help to inform mortgage market policy (Campbell, 2012, Piskorski and Seru, 2018).

References

- ANAGOL, S., V. BALASUBRAMANIAM, AND T. RAMADORAI (2018): “Endowment effects in the field: Evidence from India’s IPO lotteries,” *The Review of Economic Studies*, 85(4), 1971–2004.
- ANDERSEN, S., J. Y. CAMPBELL, K. M. NIELSEN, AND T. RAMADORAI (2018): “Inattention and Inertia in Household Finance: Evidence from the Danish Mortgage Market,” *Unpublished working paper*.
- ANDREWS, I., M. GENTZKOW, AND J. M. SHAPIRO (2017): “Measuring the sensitivity of parameter estimates to estimation moments,” *The Quarterly Journal of Economics*, 132(4), 1553–1592.
- ANENBERG, E. (2011): “Loss aversion, equity constraints and seller behavior in the real estate market,” *Regional Science and Urban Economics*, 41(1), 67–76.
- ASHRAF, N., D. KARLAN, AND W. YIN (2006): “Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines,” *The Quarterly Journal of Economics*, 121(2), 635–672.
- BADARINZA, C., J. Y. CAMPBELL, AND T. RAMADORAI (2016): “International Comparative Household Finance,” *Annual Review of Economics*, 8(1).
- BLUNDELL, R. (2017): “What have we learned from structural models?,” *American Economic Review: Papers & Proceedings*, 107(5), 287–92.
- BRACKE, P., AND S. TENREYRO (2019): “History dependence in the housing market,” *Bank of England Working Paper*.
- CALONICO, S., M. D. CATTANEO, AND R. TITIUNIK (2014): “Robust nonparametric confidence intervals for regression-discontinuity designs,” *Econometrica*, 82(6), 2295–2326.
- CAMPBELL, J. Y. (2006): “Household finance,” *The Journal of Finance*, 61(4), 1553–1604.
- (2012): “Mortgage market design,” *Review of Finance*, 17(1), 1–33.
- CARD, D., A. JOHNSTON, P. LEUNG, A. MAS, AND Z. PEI (2015a): “The effect of unemployment benefits on the duration of unemployment insurance receipt: New evidence from a regression kink design in Missouri, 2003–2013,” *American Economic Review*, 105(5), 126–30.
- CARD, D., D. S. LEE, Z. PEI, AND A. WEBER (2015b): “Inference on causal effects in a generalized regression kink design,” *Econometrica*, 83(6), 2453–2483.
- (2017): “Regression kink design: Theory and practice,” in *Regression discontinuity designs: Theory and applications*, pp. 341–382. Emerald Publishing Limited.

- CHETTY, R. (2009): “Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods,” *Annual Review of Economics*, 1(1), 451–488.
- CHETTY, R., J. N. FRIEDMAN, T. OLSEN, AND L. PISTAFERRI (2011): “Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records,” *The Quarterly Journal of Economics*, 126(2), 749–804.
- CLAPP, J. M., R. LU-ANDREWS, AND T. ZHOU (2018): “Controlling Unobserved Heterogeneity in Repeat Sales Models: Application to the Disposition Effect in Housing,” *University of Connecticut School of Business Research Paper*, (18-16).
- CRAWFORD, V. P., AND J. MENG (2011): “New York City cab drivers’ labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income,” *American Economic Review*, 101(5), 1912–32.
- DELLAVIGNA, S. (2009): “Psychology and economics: Evidence from the field,” *Journal of Economic Literature*, 47(2), 315–72.
- (2018): “Structural behavioral economics,” *National Bureau of Economic Research Working Paper*.
- DELLAVIGNA, S., AND U. MALMENDIER (2006): “Paying not to go to the gym,” *American Economic Review*, 96(3), 694–719.
- ENGELHARDT, G. V. (2003): “Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States,” *Journal of Urban Economics*, 53(1), 171–195.
- FERREIRA, F., J. GYOURKO, AND J. TRACY (2010): “Housing busts and household mobility,” *Journal of Urban Economics*, 68(1), 34–45.
- FERREIRA, F., J. GYOURKO, J. TRACY, ET AL. (2012): “Housing busts and household mobility: an update,” *Economic Policy Review*, (Nov), 1–15.
- FUDENBERG, D., AND D. K. LEVINE (2006): “A dual-self model of impulse control,” *American Economic Review*, 96(5), 1449–1476.
- GENESOVE, D., AND L. HAN (2012): “Search and matching in the housing market,” *Journal of Urban Economics*, 72(1), 31–45.
- GENESOVE, D., AND C. MAYER (2001): “Loss aversion and seller behavior: Evidence from the housing market,” *The Quarterly Journal of Economics*, 116(4), 1233–1260.
- GENESOVE, D., AND C. J. MAYER (1997): “Equity and time to sale in the real estate market,” *The American Economic Review*, 87(3), 255.
- GUL, F., AND W. PESENDORFER (2001): “Temptation and self-control,” *Econometrica*, 69(6), 1403–1435.

- (2004): “Self-control and the theory of consumption,” *Econometrica*, 72(1), 119–158.
- GUREN, A. M. (2018): “House price momentum and strategic complementarity,” *Journal of Political Economy*, 126(3), 1172–1218.
- HAHN, J., P. TODD, AND W. VAN DER KLAAUW (2001): “Identification and estimation of treatment effects with a regression-discontinuity design,” *Econometrica*, 69(1), 201–209.
- HONG, D., R. LOH, AND M. WARACHKA (2016): “Realization utility and real estate,” *Unpublished working paper*.
- IMAS, A., S. SADOFF, AND A. SAMEK (2016): “Do people anticipate loss aversion?,” *Management Science*, 63(5), 1271–1284.
- KAHNEMAN, D., J. L. KNETSCH, AND R. H. THALER (1990): “Experimental tests of the endowment effect and the Coase theorem,” *Journal of Political Economy*, 98(6), 1325–1348.
- KAHNEMAN, D., AND A. TVERSKY (1979): “Prospect theory: An analysis of decision under risk,” *Econometrica*, pp. 263–291.
- KLEVEN, H. J. (2016): “Bunching,” *Annual Review of Economics*, 8, 435–464.
- KŐSZEGI, B., AND M. RABIN (2006): “A model of reference-dependent preferences,” *The Quarterly Journal of Economics*, 121(4), 1133–1165.
- (2007): “Reference-dependent risk attitudes,” *American Economic Review*, 97(4), 1047–1073.
- LANDAIS, C. (2015): “Assessing the welfare effects of unemployment benefits using the regression kink design,” *American Economic Journal: Economic Policy*, 7(4), 243–78.
- MARZILLI ERICSON, K. M., AND A. FUSTER (2011): “Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments,” *The Quarterly Journal of Economics*, 126(4), 1879–1907.
- MCCRARY, J. (2008): “Manipulation of the running variable in the regression discontinuity design: A density test,” *Journal of Econometrics*, 142(2), 698–714.
- MEAD, R. (2017): *Statistical methods in agriculture and experimental biology*. Chapman and Hall.
- NEWAY, W. K., AND D. MCFADDEN (1994): “Handbook of Econometrics,” *Elsevier*, 4.
- NGAI, L. R., AND S. TENREYRO (2014): “Hot and cold seasons in the housing market,” *American Economic Review*, 104(12), 3991–4026.

- NIELSEN, H. S., T. SØRENSEN, AND C. TABER (2010): “Estimating the effect of student aid on college enrollment: Evidence from a government grant policy reform,” *American Economic Journal: Economic Policy*, 2(2), 185–215.
- PISKORSKI, T., AND A. SERU (2018): “Mortgage market design: Lessons from the Great Recession,” *Brookings Papers on Economic Activity*, 2018(1), 429–513.
- REES-JONES, A. (2018): “Quantifying loss-averse tax manipulation,” *The Review of Economic Studies*, 85(2), 1251–1278.
- RICHARDS, F. (1959): “A flexible growth function for empirical use,” *Journal of Experimental Botany*, 10(2), 290–301.
- SCHULHOFER-WOHL, S. (2012): “Negative equity does not reduce homeowners’ mobility,” *Federal Reserve Bank of Minneapolis Quarterly Review*, 35(1), 2–15.
- STEIN, J. (1995): “Prices and Trading Volume in the Housing Market: A Model With Down Payment Effects,” *The Quarterly Journal of Economics*, 110(2), 379–406.
- STOCK, J. H., AND M. YOGO (2002): “Testing for weak instruments in linear IV regression,” .
- TVERSKY, A., AND D. KAHNEMAN (1991): “Loss aversion in riskless choice: A reference-dependent model,” *The Quarterly Journal of Economics*, 106(4), 1039–1061.
- ZWIETERING, M., I. JONGENBURGER, F. ROMBOUTS, AND K. VAN’T RIET (1990): “Modeling of the bacterial growth curve,” *Appl. Environ. Microbiol.*, 56(6), 1875–1881.

Figure 1
Reference dependence and loss aversion

The figure illustrates how each specification of utility function is reflected in the sellers' optimal choice of listing premia. We plot a stylized version of listing premium profiles, for the case in which demand functions $\alpha(\ell)$ and $\beta(\ell)$ are linear and the household is not facing financing constraints. In the online appendix, we describe and solve an analytical version of this model.

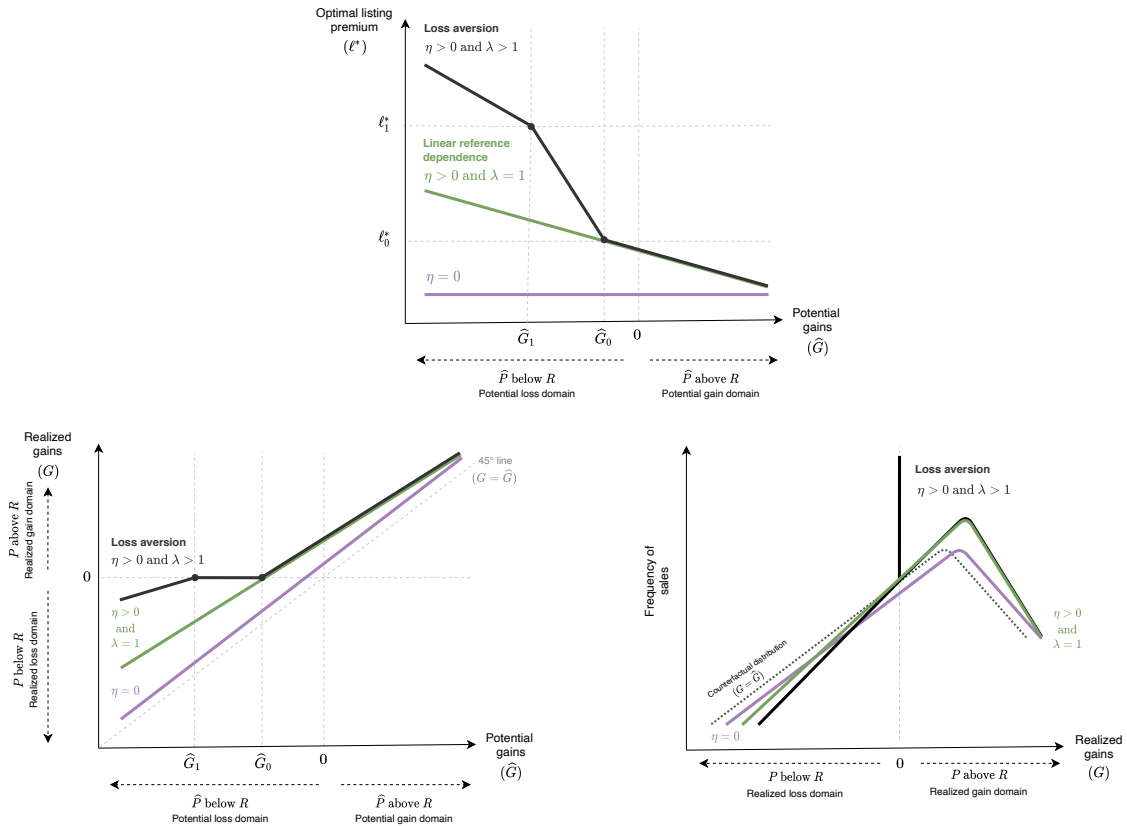


Figure 2
Concave demand

This figure illustrates the link between concave demand and the choice of optimal listing premia. We plot a stylized listing profile resulting from a case of pure reference dependence with no loss aversion ($\eta > 0$ and $\lambda = 1$). Since the probability of sale does not respond to listing premia set below a certain level $\underline{\ell}$, it is rational for sellers to not respond to the exact magnitude of the expected gain. A steeper slope of demand translates into a general flattening out of the listing premium profile.

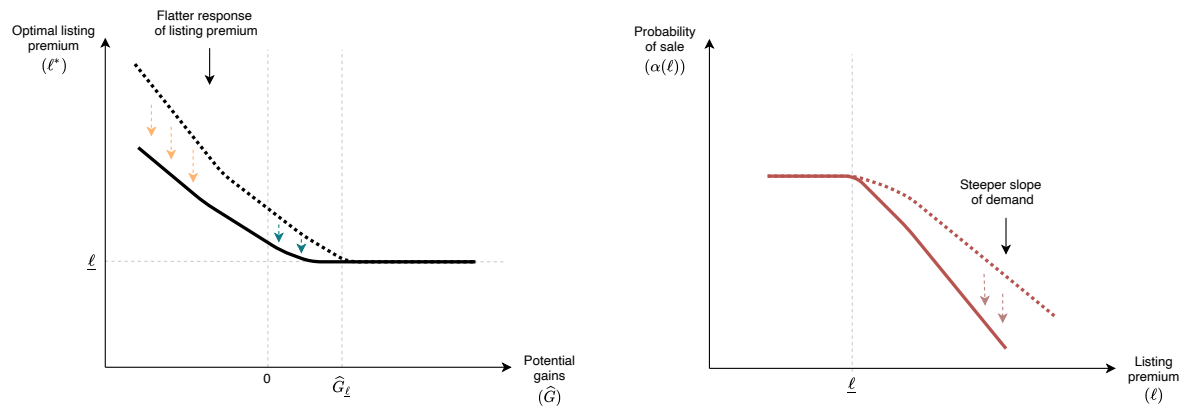


Figure 3
Listing premia and potential gains

The figure reports binned average values (in 1 percentage point steps) for the listing premium (ℓ) for different levels of potential gains (\hat{G}). The green line corresponds to a polynomial of order three, fitted in the positive domain of potential gains. The red line corresponds to an equivalent polynomial fit in the potential loss domain.

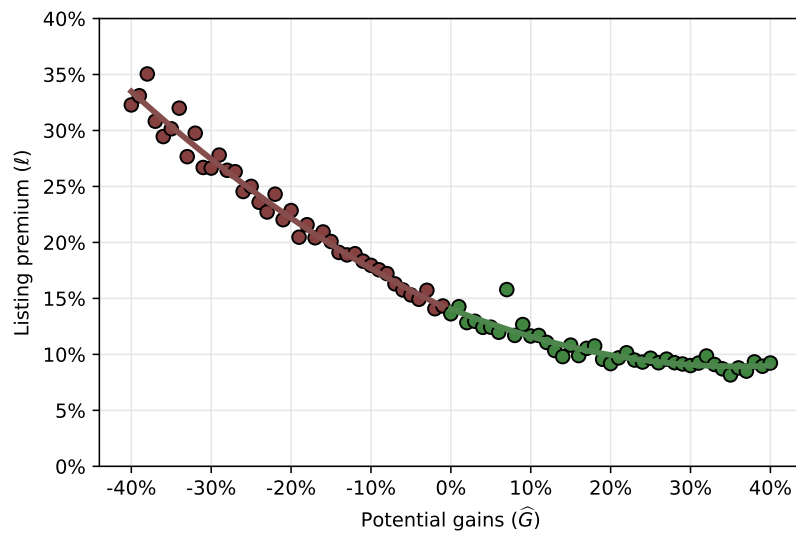


Figure 4
Bunching around realized gains of zero

The figure reports binned frequencies of observations (in 1 percentage point steps) for different levels of realized gains (G). The dotted line shows the counterfactual corresponding to the distribution of potential gains (\hat{G}) in the sample of realized sales.

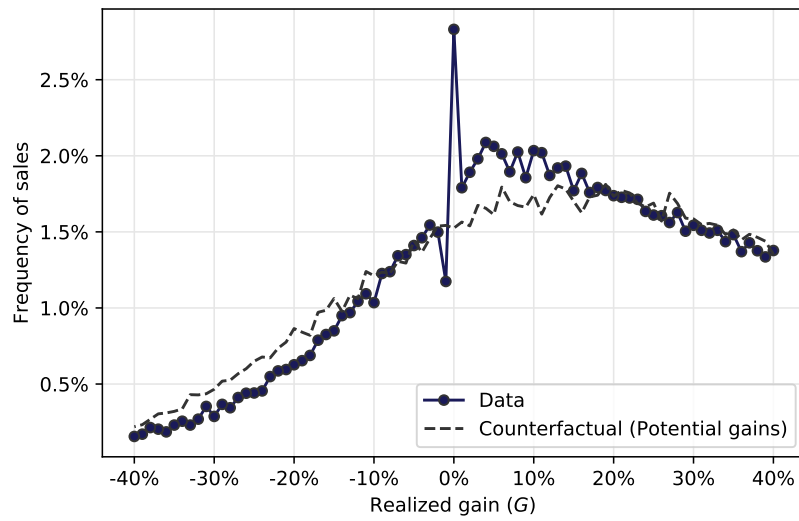


Figure 5
Gains vs. home equity

The figure reports binned average values (in 3% steps) for the listing premium (ℓ) along both levels of potential gains and home equity, and the observed frequency of sales along levels of realized gains and home equity. The dotted lines show the binned values for two cross-sections, where we condition on a home equity level of 20%, and a level of gains of 0%, respectively. We use these two representative cross-sections to generate the empirical moments used in structural estimation.

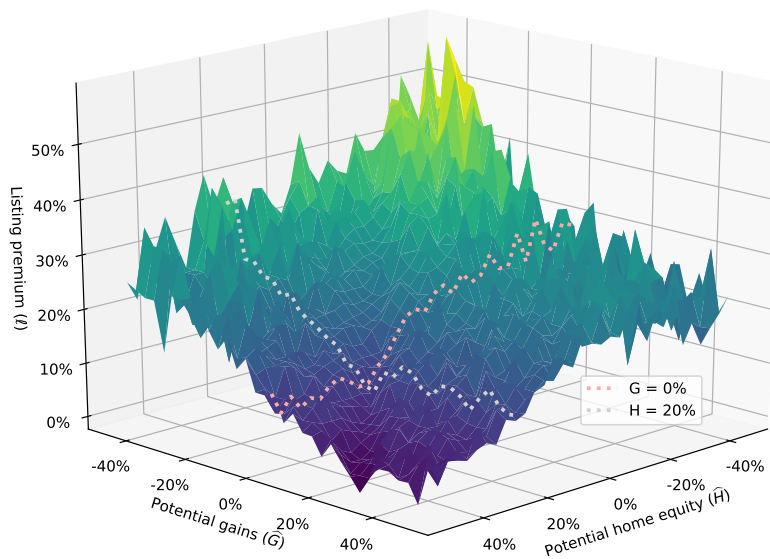


Figure 6
Concave demand in the data

The left-hand side of the figure reports the average probability of sale within six months $\alpha(\ell)$ across 1 percentage point bins of the listing premium in the sample. The solid line indicates fitted values corresponding to a generalized logistic function (GLF). The right-hand side of the figure shows the average realized premium $\beta(\ell)$ across bins of the listing premium. The solid line indicates fitted values corresponding to a polynomial of order three.

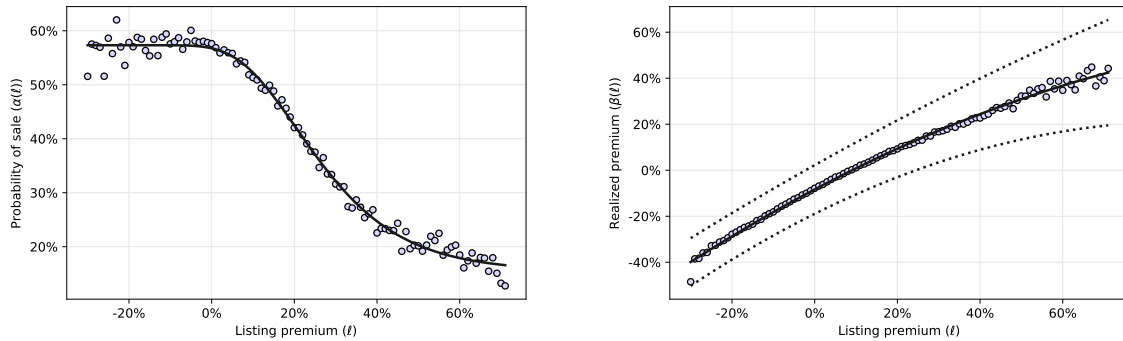
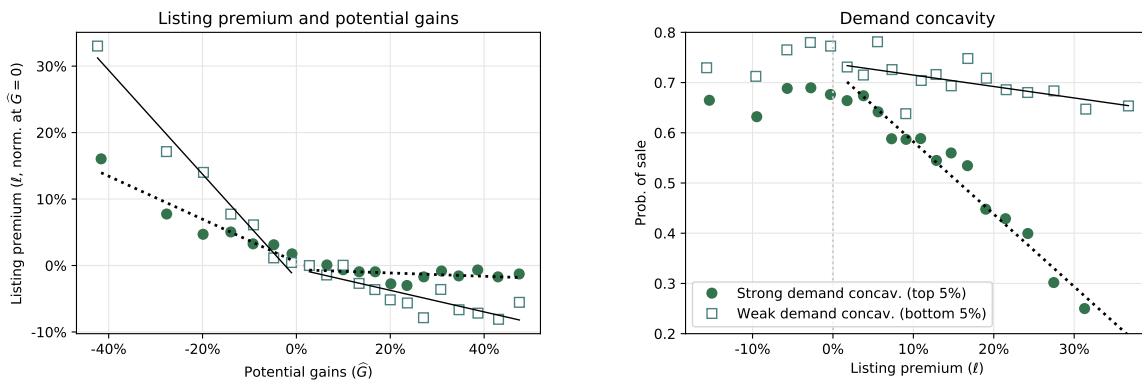


Figure 7
 Listing premium-gain slope and demand concavity

Panel A shows the listing premium over gains (left-hand side) and demand concavity (right-hand side) patterns. We sort municipalities by the estimated demand concavity, using municipalities in the top and bottom 5% of observations. Demand concavity is estimated as the slope coefficient of the effect of the listing premium on the probability of sale within six months, for $\ell > 0$. For better illustration of the main effect, we adjust the quantities measured to have the same level at $G = 0\%$ and $\ell = 0\%$ respectively. The left-hand side of Panel B shows the correlation between the estimated listing premium slope and demand concavity across municipalities using a binned scatter plot with equal-sized bins. The right-hand side of Panel B shows a binned scatter plot of the correlation between the main instrument, the share of listed apartments and row houses in a given municipality, and demand concavity in a binned scatter plot with equal-sized bins.

Panel A



Panel B

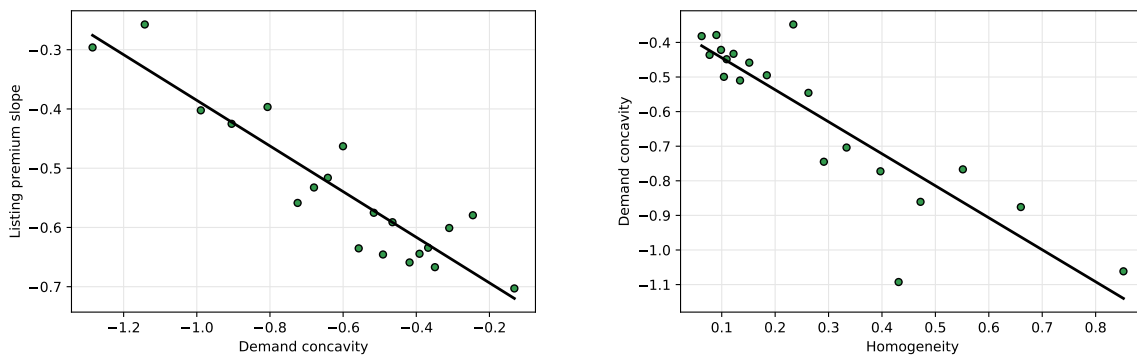


Figure 8
Extensive margin

The figure reports the average yearly probability of listing a property for sale. We first calculate the potential gain level for each unit in the stock of properties in Denmark, for each year covered by our sample of listings. We then divide the number of properties which have been listed for sale by the number of total property \times year observations in the stock of properties, for each 1 percentage point bin of potential gains.

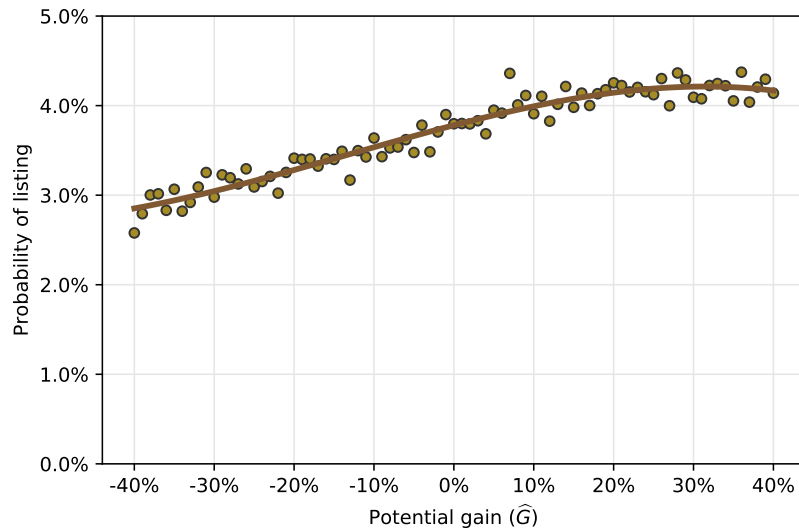
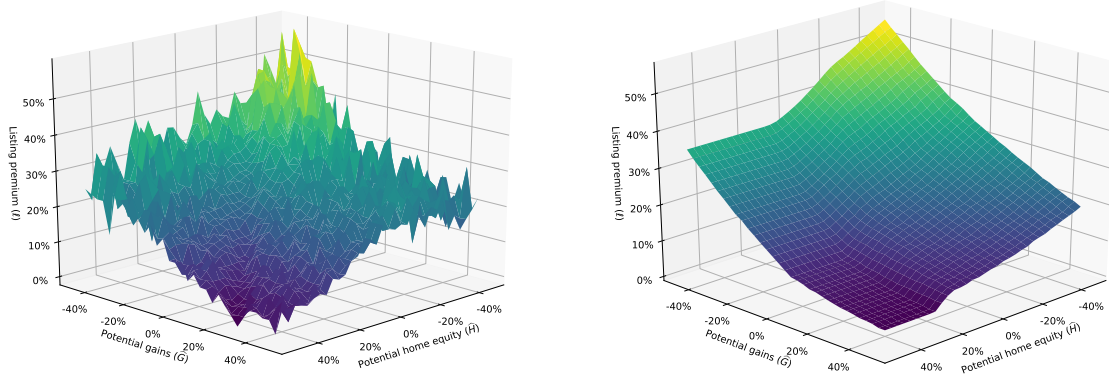


Figure 9
Model fit

Panel A reports listing premia by potential gains and home equity, both in the data and in the model. We use the set of seven estimated parameters to evaluate average quantities in the model, accounting for the extensive margin decision of whether to list the property for sale or not. Panel B illustrates the model fit for conditional listing premia profiles, conditioning on different levels of potential gains and home equity. Dotted lines indicate observations in the data (for which we report fitted values using generalized logistic functions) and solid lines their model-implied counterparts.

Panel A



Panel B

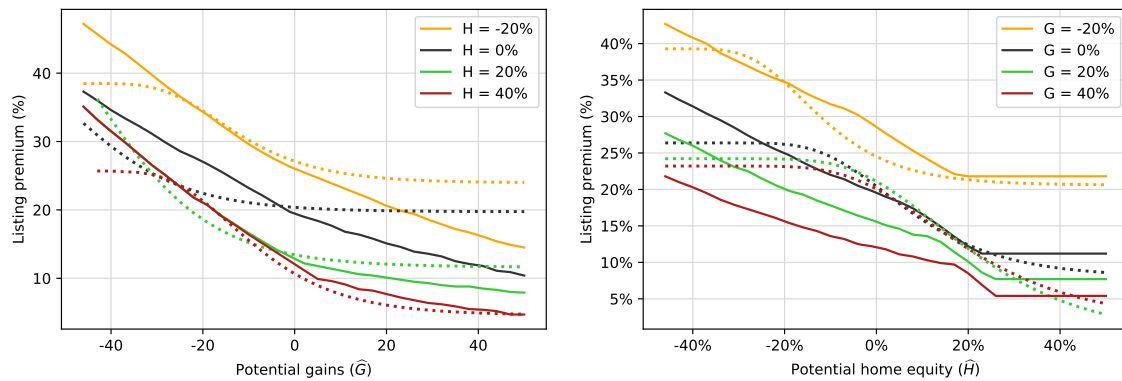


Table 1

Overview of moments and other estimates from the data

The table reports estimated empirical moments in the data. The first two capture the level and the slope of the listing premium with respect to the seller's level of potential gains, for $\widehat{G} > 0\%$, conditional on a home equity level of $\widehat{H} = 20\%$. The third moment is the slope of the listing premium with respect to potential home equity, for $\widehat{H} < 20\%$, conditional on gains of $\widehat{G} = 0\%$. The fourth and fifth moments are obtained as slope coefficients from cross-sectional regressions by municipality. For each municipality, we compute the slope $\ell - \widehat{G}$ for $\widehat{G} < 0\%$ and $\widehat{G} \geq 0\%$ respectively, as well as the concavity of demand (i.e. the slope $\alpha - \ell$ for $\ell > 0$). The sixth moment is the slope of the listing probability with respect to the potential gains, conditional on a home equity level of $\widehat{H} = 20\%$. The final moment captures the bunching of transactions around realized gains of 0%, calculated as the relative frequency of transactions in the [0,3%] interval of realized gains, relative to the [-3%,0] interval. In parentheses, we report bootstrap standard errors, clustered at the shire level. *, **, *** indicate statistical significance at the 10%, 5% and 1% confidence levels, respectively.

1.	Level of ℓ for $\widehat{G} = 0\%$	0.106***	(0.005)
2.	Slope $\ell - \widehat{G}$ for $\widehat{G} < 0\%$	-0.490***	(0.047)
3.	Slope $\ell - \widehat{H}$ for $\widehat{H} < 20\%$	-0.333***	(0.030)
4.	Cross-sectional slope $\ell - \widehat{G} - \alpha$ for $\widehat{G} < 0\%$	-0.407***	(0.065)
5.	Cross-sectional slope $\ell - \widehat{G} - \alpha$ for $\widehat{G} \geq 0\%$	-0.122**	(0.043)
6.	Slope of list. prob. by \widehat{G}	0.005**	(0.002)
7.	Bunching above $G = 0\%$	0.302***	(0.050)

Table 2
Estimated parameters

The table reports structural parameter estimates obtained through classical minimum distance estimation. We recover concave demand $\alpha(\ell)$ and $P(\ell)$ from the data and set the down-payment constraint $\gamma = 20\%$. In parentheses, we report standard errors based on the estimated bootstrap variance-covariance matrix in the data, clustered at the shire level. *, **, *** indicate statistical significance at the 10%, 5% and 1% confidence levels, respectively.

η	=	0.948***	(0.344)
λ	=	1.576***	(0.570)
μ	=	1.060***	(0.107)
δ	=	-0.097***	(0.009)
θ_{\min}	=	0.217	(0.165)
θ_{\max}	=	1.005***	(0.197)
φ	=	0.037	(0.011)

Table 3

Listing premium-slope over gains and demand concavity slope regressions

This table reports regression results for the relationship between the listing premium slope over gains and demand concavity. The dependent variable in all regressions is the slope of the listing premium over $\widehat{G} < 0$ across municipalities.⁵⁸ Column 1 reports the baseline correlation with the demand concavity slope across municipalities using OLS. Column 2 reports the 2-stage least squares regression instrumenting demand concavity with the apartment- and row-house share. Columns 3 and 4 report the overidentified 2SLS regression with both instruments, row-house and apartment share and average distance to city, without and with household controls (age, education length, net financial assets and log income), respectively. In parentheses, we report bootstrap standard errors, clustered at the shire level. *, **, *** indicate statistical significance at the 10%, 5% and 1% confidence levels, respectively.

	OLS	2SLS		
	(1)	(2)	(3)	(4)
		<i>Single IV</i>	<i>Overidentified</i>	
Demand concavity	-0.407*** (0.067)	-0.520*** (0.111)	-0.504*** (0.087)	-0.346 (0.259)
Household controls				✓
Observations	95	95	95	95
R^2	0.432			
First-stage F-stat		35.96	16.94	25.376
Hansen J-stat (p-val)			0.175	0.199