Search Frictions and Idiosyncratic Price Dispersion in the US Housing Market

Nadia Kotova [1] Anthony Lee Zhang [2]

[1]Stanford GSB

[2]UChicago Booth

Prices of individual houses are highly volatile

- A large fraction of household wealth is in housing:
  - US house ownership rate is 64.8%.
  - Equity in own home constitutes 34% of the total net worth of the US population.

Prices of individual houses are highly volatile

- A large fraction of household wealth is in housing:
  - US house ownership rate is 64.8%.
  - Equity in own home constitutes 34% of the total net worth of the US population.
- Jordà, Schularick, Taylor (2019): Housing outperforms equity, same return (7%) but lower volatility (8% vs. 20%).

Prices of individual houses are highly volatile

- A large fraction of household wealth is in housing:
  - US house ownership rate is 64.8%.
  - Equity in own home constitutes 34% of the total net worth of the US population.
- Jordà, Schularick, Taylor (2019): Housing outperforms equity, same return (7%) but lower volatility (8% vs. 20%).
- But households do not hold a diversified real estate portfolio!

Prices of individual houses are highly volatile

- A large fraction of household wealth is in housing:
    - US house ownership rate is 64.8%.
    - Equity in own home constitutes 34% of the total net worth of the US population.
- Jordà, Schularick, Taylor (2019): Housing outperforms equity, same return (7%) but lower volatility (8% vs. 20%).
- But households do not hold a diversified real estate portfolio!
- Individual houses are subject to both volatility in average prices and large <u>idiosyncratic</u> risk.

Prices of individual houses are highly volatile

- A large fraction of household wealth is in housing:
  - US house ownership rate is 64.8%.
  - Equity in own home constitutes 34% of the total net worth of the US population.
- Jordà, Schularick, Taylor (2019): Housing outperforms equity, same return (7%) but lower volatility (8% vs. 20%).
- But households do not hold a diversified real estate portfolio!
- Individual houses are subject to both volatility in average prices and large idiosyncratic risk.
- Sources of housing idiosyncratic price dispersion (IPD) are not well understood.

This paper

Search frictions are an important driver of housing IPD

This paper

Search frictions are an important driver of housing IPD

1. Empirical results:
   - PD is countercyclical and seasonal.

This paper

Search frictions are an important driver of housing IPD

1. Empirical results:
   - PD is countercyclical and seasonal.
   - PD is strongly correlated with TOM and other market tightness measures.

This paper

Search frictions are an important driver of housing IPD

1. Empirical results:
   - PD is countercyclical and seasonal.
   - PD is strongly correlated with TOM and other market tightness measures.
2. Theory:
   - Construct a search-and-bargaining model to rationalize empirical results.

This paper

Search frictions are an important driver of housing IPD

1. Empirical results:
   - PD is countercyclical and seasonal.
   - PD is strongly correlated with TOM and other market tightness measures.
2. Theory:
   - Construct a search-and-bargaining model to rationalize empirical results.
   - Calibrate model to quantify tradeoffs facing agents.

Outline

1. Data
2. Measuring price dispersion
3. Empirical results
4. Model
5. Calibration
6. Conclusion

Data

Corelogic (2001–2017): Transaction prices & volumes, house characteristics

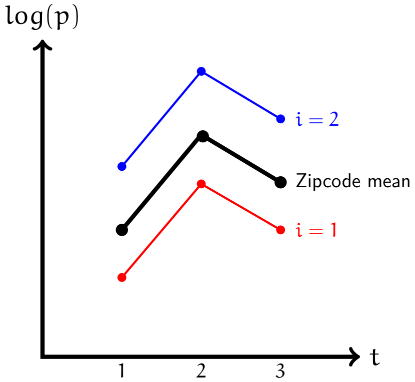
- Arms-length non-foreclosure transactions of single family residences with recorded sale price.

Zillow Research (2010–2017): County-month TOM, Zillow Home Value Index.

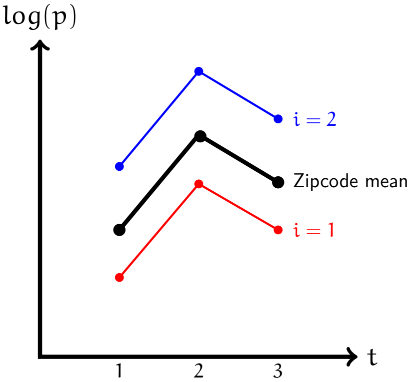
Realtor.com (2012–2017): Zip-month TOM.

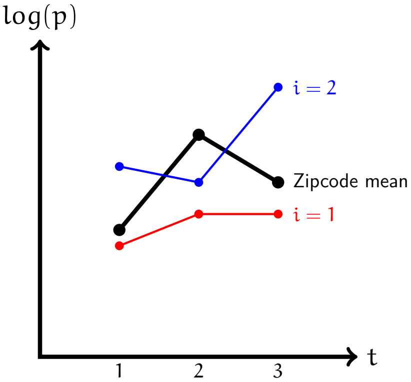
ACS Social Explorer (2012-2016): Demographic covariates.

Measuring price dispersion: intuition



Measuring price dispersion: intuition





Measuring price dispersion

- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z\left(x_i, t\right) + \epsilon_{it}$$

Measuring price dispersion

- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z(x_i, t) + \epsilon_{it}$$

- $\widehat{\epsilon}_{it}^2$, error term, is our house-level measure of idiosyncratic PD

Measuring price dispersion

- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z(x_i, t) + \epsilon_{it}$$

- $\widehat{\epsilon}_{it}^2$, error term, is our house-level measure of idiosyncratic PD
- Specification captures:
  - η_{zt}: Zipcode-month trend

Measuring price dispersion

- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z(x_i, t) + \epsilon_{it}$$

- $\widehat{\epsilon}^2_{it}$, error term, is our house-level measure of idiosyncratic PD
- Specification captures:
  - η_{zt}: Zipcode-month trend
  - γ_i: Time-invariant house quality (observed or unobserved)

Measuring price dispersion

- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z(x_i, t) + \epsilon_{it}$$

- $\widehat{\epsilon}_{it}^2$, error term, is our house-level measure of idiosyncratic PD
- Specification captures:
  - η_{zt}: Zipcode-month trend
  - γ_i: Time-invariant house quality (observed or unobserved)
  - $f_z(x_i, t)$: Time-varying effects of characteristics x_i

Measuring price dispersion

- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z(x_i, t) + \epsilon_{it}$$

- $\widehat{\epsilon}_{it}^2$, error term, is our house-level measure of idiosyncratic PD
- Specification captures:
  - η_{zt}: Zipcode-month trend
  - γ_i: Time-invariant house quality (observed or unobserved)
  - $f_z(x_i, t)$: Time-varying effects of characteristics x_i
- Concerns:
  - Time-varying effects of unobservables (e.g. construction quality, flood risk).

Measuring price dispersion

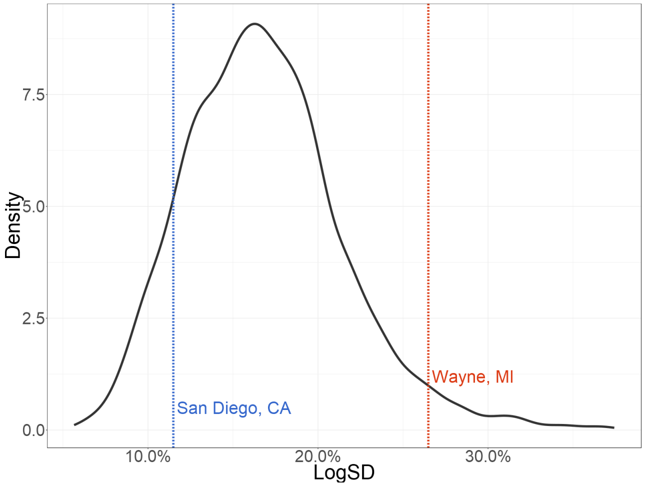
- For zip code z, house i w. characteristics x_i, month t, estimate:

$$\log(p_{it}) = \eta_{zt} + \gamma_i + f_z(x_i, t) + \epsilon_{it}$$

- $\widehat{\epsilon}^2_{it}$, error term, is our house-level measure of idiosyncratic PD
- Specification captures:
  - η_{zt}: Zipcode-month trend
  - γ_i: Time-invariant house quality (observed or unobserved)
  - $f_z(x_i, t)$: Time-varying effects of characteristics x_i
- Concerns:
  - Time-varying effects of unobservables (e.g. construction quality, flood risk).
  - Time-varying characteristics (e.g. renovations, depreciation).

Distribution of $\sqrt{\widehat{\epsilon}^2_{it}}$ across zipcodes

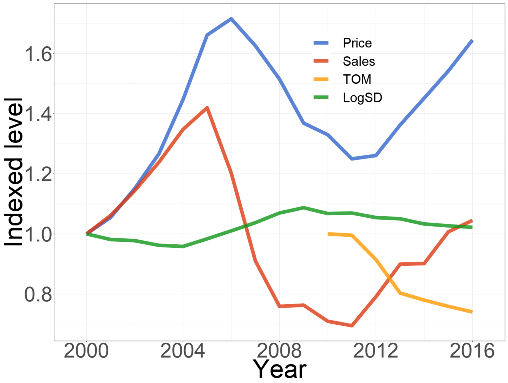
Mean:	16.8%
SD:	4.6%
P10:	11.3%
P90:	22.6%



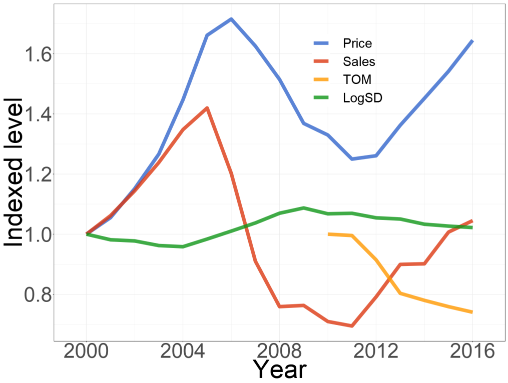
Summary of results

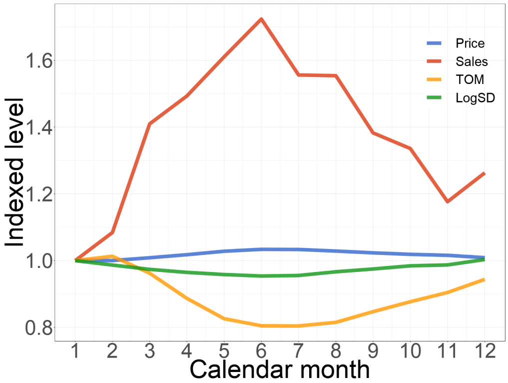
- IPD is countercyclical and seasonal
- In panel and cross-sectional specs, IPD is correlated with measures of market tightness: time-on-market, vacancy rates, migration rates, sales, prices

IPD is countercyclical and seasonal



IPD is countercyclical and seasonal





County-year panel regressions

	LogSD x 100					
	(1)	(2)	(3)	(4)	(5)	(6)
Log ZHVI	−1.056***					−0.834
	(0.369)					(0.768)
Log sales		−0.971***				−1.750***
		(0.194)				(0.541)
Time on market (months)			0.521***			0.170
			(0.162)			(0.175)
Vacancy rate				16.204***		10.804***
				(1.798)		(2.407)
Pop growth rate					−8.726***	−4.758
					(2.057)	(3.779)
County fixed effects	X	X	X	X	X	X
Year fixed effects	X	X	X	X	X	X
Sample period	2000-2016	2000-2016	2010-2016	2007-2016	2007-2016	2010-2016
N	10,366	10,366	2,516	5,807	5,284	2,492
Adjusted R^2	0.858	0.859	0.911	0.895	0.891	0.919

Zipcode cross-sectional regressions

	LogSD x 100						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time on market (months)	2.463***				1.910***	2.941***	2.345***
	(0.088)				(0.095)	(0.120)	(0.093)
Vacancy rate		15.335***			7.486***	2.121***	4.412***
		(0.829)			(0.852)	(0.770)	(0.757)
Pop growth			−1.729*		0.401	−1.130**	−1.095*
			(0.886)		(0.783)	(0.572)	(0.639)
Mean log price				−3.899***	−3.406***	−1.075***	−1.643***
				(0.209)	(0.193)	(0.222)	(0.199)
Controls	X	X	X	X	X	X	X
Fixed effects						State	CBSA
N	4,109	4,109	4,109	4,109	4,109	4,109	4,109
Adjusted R^2	0.542	0.496	0.455	0.497	0.580	0.797	0.732

Robustness checks

- Zipcode-year panel regressions and county cross-sectional regressions.
- Controlling for time-between-sales and times sold.
- Removing polynomial term.
- Zillow vs Realtor.com time-on-market.

Model

Stationary equilibrium search-and-bargaining model. 3 kinds of agents:

1. Buyers: exogeneously enter market, match with sellers to buy houses
2. Matched owners: receive separation shocks at rate λ_M
3. Sellers: match with buyers to sell and leave market

Model

Stationary equilibrium search-and-bargaining model. 3 kinds of agents:

1. Buyers: exogeneously enter market, match with sellers to buy houses
2. Matched owners: receive separation shocks at rate λ_M
3. Sellers: match with buyers to sell and leave market

Price dispersion arises from dispersion in buyer match quality and seller holding costs

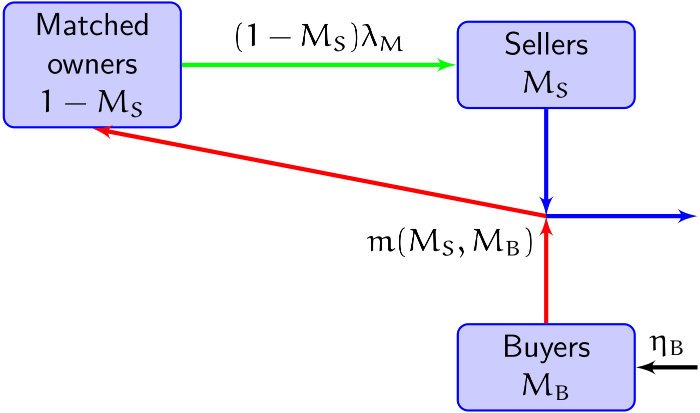
Agents and stationary flows

Matched owners
$1 - M_S$

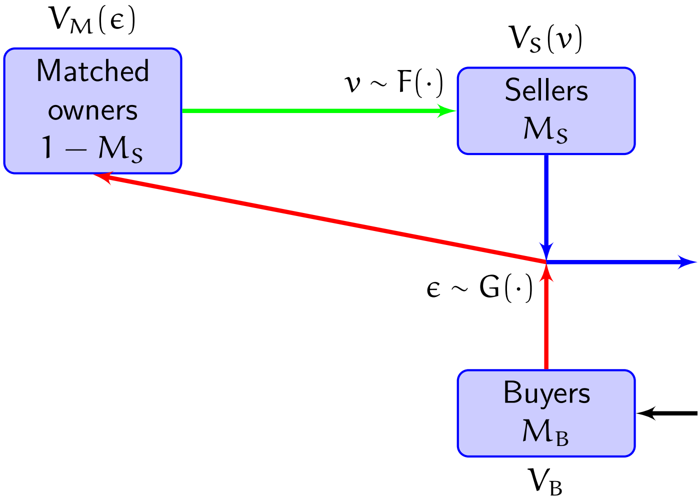
Sellers
M_S

Buyers
M_B

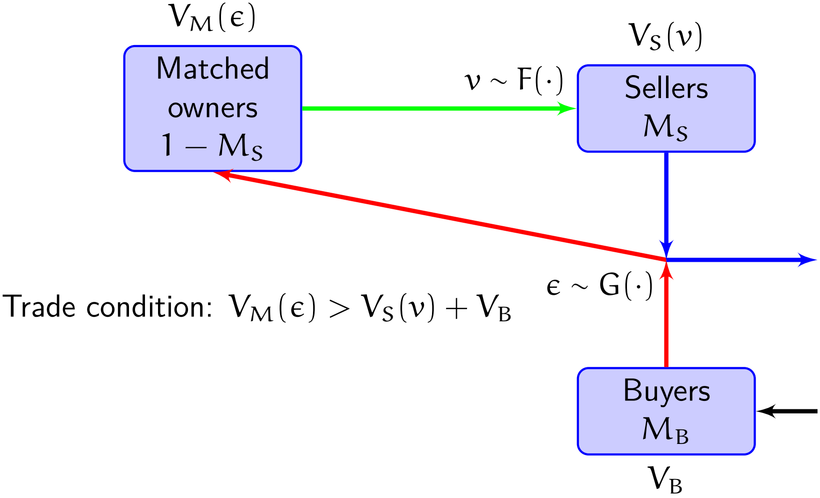
Agents and stationary flows



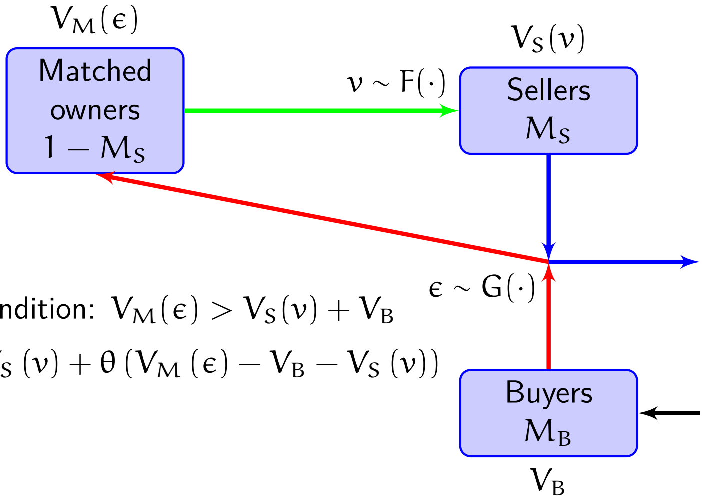
Agents and stationary flows



Agents and stationary flows



Agents and stationary flows



Trade condition: $V_M(\epsilon) > V_S(v) + V_B$

$$P(v, \epsilon) = V_S(v) + \theta\left(V_M(\epsilon) - V_B - V_S(v)\right)$$

Equilibrium conditions

Buyer, seller, and matched owner Bellman equations:

$$rV_B = \lambda_B \int\int_{\epsilon > \epsilon^*(\nu)} [(1-\theta)(V_M(\epsilon) - V_B - V_S(\nu))] \, dG(\epsilon) \, dF_{eq}(\nu)$$

$$rV_S(\nu) = \nu + \lambda_S \int_{\epsilon > \epsilon^*(\nu)} \theta (V_M(\epsilon) - V_B - V_S(\nu)) \, dG(\epsilon)$$

$$rV_M(\epsilon) = \epsilon + \lambda_M \left(\int V_S(\nu) \, dF(\nu) - V_M(\epsilon) \right)$$

Trade cutoffs:

$$V_M(\epsilon^*(\nu)) = V_S(\nu) + V_B$$

Matching rates:

$$M_S \lambda_S = M_B \lambda_B = \alpha M_B^{\phi} M_S^{1-\phi}$$

Flow equality:

$$(1 - M_S) \lambda_M f(\nu) = \lambda_S M_S f_{eq}(\nu) (1 - G(\epsilon^*(\nu)))$$

$$G_{eq}(\epsilon) = \frac{\int_\nu \lambda_S M_S \left[\int_{\tilde{\epsilon} = \epsilon_0}^{\epsilon} 1(\tilde{\epsilon} > \epsilon^*(\nu)) \, dG(\tilde{\epsilon}) \right] dF_{eq}(\nu)}{\int_\nu \lambda_S M_S (1 - G(\epsilon^*(\nu))) \, dF_{eq}(\nu)}$$

$$(1 - M_S) \lambda_M = \eta_B$$

Price variance decomposition

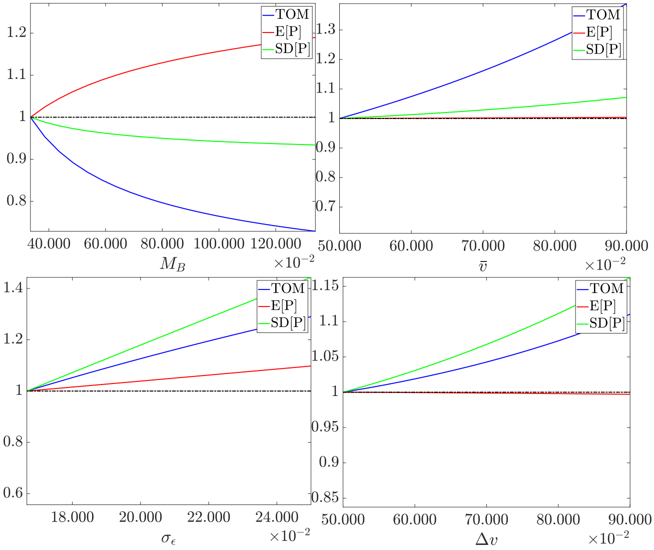
$$\underbrace{Var_{v\sim F(\cdot)}\left(V_S\left(v\right)\right)}_{\text{Seller holding utility}} + \underbrace{\left(\frac{\theta}{r+\lambda_M}\right)^2 \sigma_\epsilon^2}_{\text{Buyer match utility}}$$

Price variance decomposition

$$\underbrace{Var_{v \sim F(\cdot)}\left(V_S\left(v\right)\right)}_{\text{Seller holding utility}} + \underbrace{\left(\frac{\theta}{r+\lambda_M}\right)^2 \sigma_\epsilon^2}_{\text{Buyer match utility}}$$

$$V_S'\left(v\right) = \frac{TOM\left(v\right)}{rTOM\left(v\right)+\theta}$$

Comparative statics



Calibration: TOM-PD coef

Type	Coef
Yearly	0.695
Seasonal	0.667
Panel	0.521
Cross-sectional	1.498 - 2.941

Calibration: TOM-PD coef

Type	Coef
Yearly	0.695
Seasonal	0.667
Panel	0.521
Cross-sectional	1.498 - 2.941

- Calibrate the model in stationary eq. to match panel TOM-PD coef, PD & TOM levels, prices, volumes

Calibration: TOM-PD coef

Type	Coef
Yearly	0.695
Seasonal	0.667
Panel	0.521
Cross-sectional	1.498 - 2.941

- Calibrate the model in stationary eq. to match panel TOM-PD coef, PD & TOM levels, prices, volumes
- How large is the TOM-price tradeoff?

Calibration: TOM-PD coef

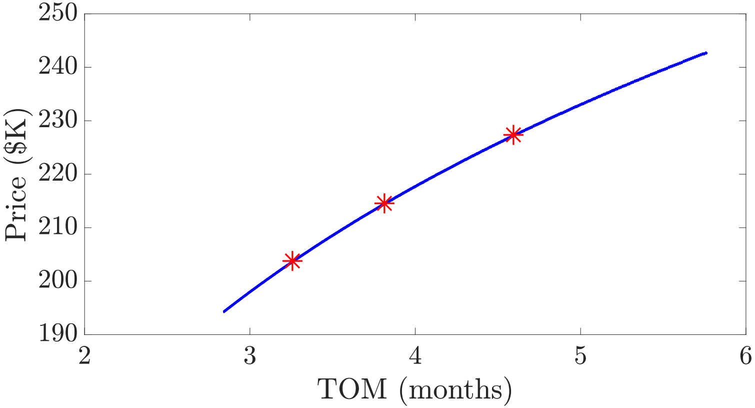
Type	Coef
Yearly	0.695
Seasonal	0.667
Panel	0.521
Cross-sectional	1.498 - 2.941

- Calibrate the model in stationary eq. to match panel TOM-PD coef, PD & TOM levels, prices, volumes
- How large is the TOM-price tradeoff?
- How does the tradeoff vary with market tightness?

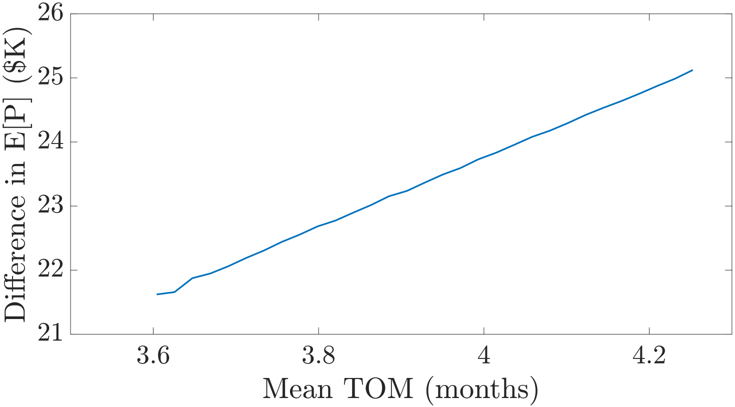
Results for 2010 (PRELIMINARY)

- Monthly values: $v \sim U[-11.3K\$, -5K\$]$;
- SD in monthly sellers values $\approx \$1,830$;
- Approx. 14% of PD comes from seller values.

E[P] and TOM of sellers with different ν for fixed M_B (PRELIMINARY)



Difference in E[P] between ν 75$^{\text{th}}$ and 25$^{\text{th}}$ percentiles (PRELIMINARY)



Conclusion

- Idiosyncratic house price dispersion is:
  - Counter-cyclical and seasonal,
  - Correlated in panel and cross-section with market tightness measures.
- Construct a model in which IPD comes from traders' value heterogeneity, amplified by market frictions
- Calibrate model to data to quantify tradeoffs agents face
- In progress: try to obtain welfare implications of search frictions?

Thank you!

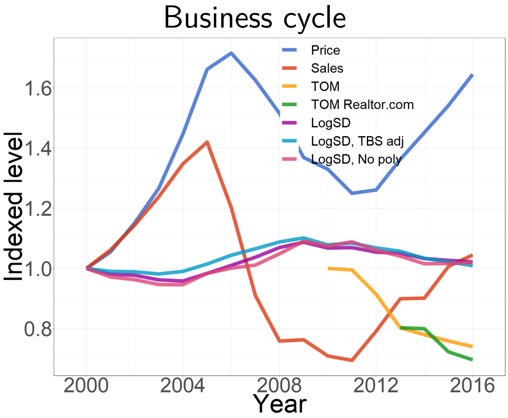
Zipcode-year panel regressions

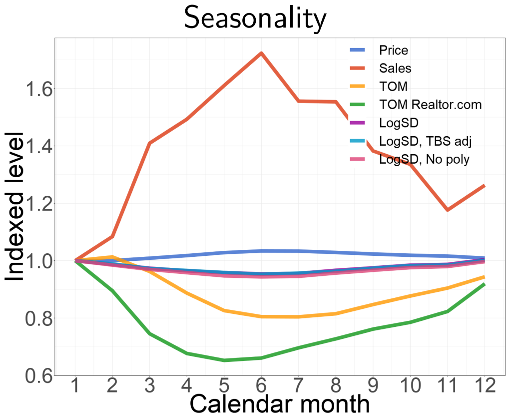
	LogSD x 100			
	(1)	(2)	(3)	(4)
Log ZHVI	−1.201***			−1.703***
	(0.098)			(0.352)
Log sales		−0.862***		−0.178
		(0.120)		(0.169)
Time on market (months)			0.263***	0.243***
			(0.071)	(0.073)
Zip fixed effects	X	X	X	X
Year fixed effects	X	X	X	X
Sample period	2000-2016	2000-2016	2013-2016	2013-2016
N	52,061	52,061	12,257	12,257
Adjusted R^2	0.848	0.850	0.924	0.924

County cross-sectional regressions

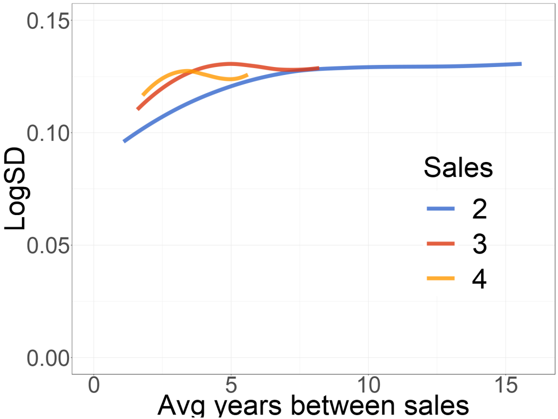
	LogSD x 100				
	(1)	(2)	(3)	(4)	(5)
Time on market (months)	1.019***			0.694**	1.481***
	(0.336)			(0.336)	(0.409)
Vacancy rate		16.726***		16.820***	11.240***
		(3.128)		(3.899)	(3.330)
Mean log price			−4.121***	−4.535***	−3.625***
			(0.630)	(0.898)	(1.004)
Controls	X	X	X	X	X
Fixed effects					State
N	299	473	473	299	299
Adjusted R^2	0.443	0.461	0.477	0.510	0.717

Yearly and seasonal robustness





Effect of time-between-sales and times sold



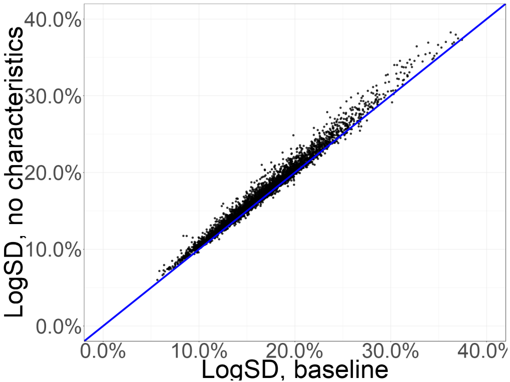
TBS adj county-year panel regressions

	LogSD x 100					
	(1)	(2)	(3)	(4)	(5)	(6)
Log ZHVI	-0.590^{***}					-0.776^{*}
	(0.217)					(0.467)
Log sales		-0.673^{***}				-1.204^{***}
		(0.124)				(0.304)
Time on market (months)			0.285^{**}			0.056
			(0.130)			(0.141)
Vacancy rate				9.333^{***}		6.688^{***}
				(1.115)		(1.567)
Pop growth rate					-6.884^{***}	-3.236
					(1.283)	(2.472)
County fixed effects	X	X	X	X	X	X
Year fixed effects	X	X	X	X	X	X
Sample period	2000-2016	2000-2016	2010-2016	2007-2016	2007-2016	2010-2016
N	10,286	10,286	2,512	5,793	5,271	2,490
Adjusted R^2	0.886	0.888	0.923	0.912	0.909	0.927

TBS adj zipcode cross-sectional regressions

	LogSD x 100						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time on market (months)	1.727***				1.260***	2.155***	1.681***
	(0.070)				(0.076)	(0.091)	(0.070)
Vacancy rate		11.785***			6.601***	2.277***	4.088***
		(0.645)			(0.674)	(0.585)	(0.569)
Pop growth			−0.784		0.790	−0.377	−0.348
			(0.690)		(0.620)	(0.435)	(0.480)
Mean log price				−2.950***	−2.639***	−0.337**	−0.969***
				(0.163)	(0.153)	(0.169)	(0.150)
Controls	X	X	X	X	X	X	X
Fixed effects						State	CBSA
N	4,109	4,109	4,109	4,109	4,109	4,109	4,109
Adjusted R^2	0.524	0.494	0.452	0.493	0.564	0.806	0.749

Effect of polynomial term



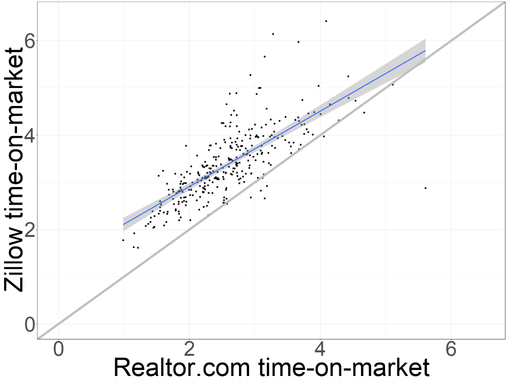
No poly county-year panel regressions

	LogSD x 100					
	(1)	(2)	(3)	(4)	(5)	(6)
Log ZHVI	−1.308***					−2.592***
	(0.434)					(0.925)
Log sales		−1.010***				−1.476**
		(0.226)				(0.663)
Time on market (months)			0.339*			0.082
			(0.182)			(0.200)
Vacancy rate				19.166***		12.286***
				(1.942)		(3.067)
Pop growth rate					−10.240***	−4.827
					(3.027)	(5.387)
County fixed effects	X	X	X	X	X	X
Year fixed effects	X	X	X	X	X	X
Sample period	2000-2016	2000-2016	2010-2016	2007-2016	2007-2016	2010-2016
N	10,366	10,366	2,516	5,807	5,284	2,492
Adjusted R^2	0.819	0.820	0.894	0.855	0.849	0.902

No poly zipcode cross-sectional regressions

	LogSD x 100						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time on market (months)	2.413***				1.779***	2.880***	2.229***
	(0.090)				(0.097)	(0.126)	(0.097)
Vacancy rate		16.383***			9.066***	3.386***	5.474***
		(0.839)			(0.869)	(0.808)	(0.787)
Pop growth			−1.340		0.757	−0.777	−0.774
			(0.901)		(0.799)	(0.600)	(0.664)
Mean log price				−3.899***	−3.450***	−0.962***	−1.623***
				(0.212)	(0.197)	(0.233)	(0.207)
Controls	X	X	X	X	X	X	X
Fixed effects						State	CBSA
N	4,109	4,109	4,109	4,109	4,109	4,109	4,109
Adjusted R^2	0.548	0.514	0.468	0.509	0.588	0.789	0.727

Realtor.com vs Zillow time-on-market



County regressions with Realtor.com TOM

	LogSD x 100			
	(1)	(2)	(3)	(4)
Realtor.com time on market	−0.043	−0.075	1.372***	1.041
	(0.193)	(0.196)	(0.271)	(0.633)
Log ZHVI		−2.470***		
		(0.772)		
Vacancy rate		5.661**		−1.898
		(2.464)		(6.424)
Daily list frac		31.003		196.011***
		(20.025)		(68.795)
Log sales		0.526		
		(0.448)		
Pop growth rate		−11.302*		
		(6.479)		
County fixed effects			X	X
Year fixed effects			X	X
Sample period	2012-2016	2012-2016	2013-2016	2013-2016
N	2,346	2,041	467	110
R^2	0.942	0.951	0.560	0.836

Zipcode cross-sectional regressions with heterogeneity controls

	Time on market (months)				LogSD x 100	
	(1)	(2)	(3)	(4)	(5)	(6)
Time on market (months)					1.910***	1.498***
					(0.095)	(0.101)
Norm SD yr built		0.214***		0.174***		0.625***
		(0.010)		(0.010)		(0.069)
Norm SD sqft			0.188***	0.124***		0.309***
			(0.011)	(0.011)		(0.073)
Vacancy rate	4.100***	4.250***	3.936***	4.114***	7.486***	9.339***
	(0.124)	(0.118)	(0.120)	(0.117)	(0.852)	(0.857)
Pop growth	−0.231*	−0.454***	−0.130	−0.345***	0.401	−0.176
	(0.128)	(0.122)	(0.124)	(0.121)	(0.783)	(0.777)
Mean log price	−0.255***	−0.302***	−0.267***	−0.301***	−3.406***	−3.668***
	(0.031)	(0.030)	(0.030)	(0.029)	(0.193)	(0.192)
Controls	X	X	X	X	X	X
N	4,109	4,109	4,109	4,109	4,109	4,109
Adjusted R^2	0.437	0.494	0.475	0.509	0.580	0.593

Calibration: parametric assumptions

Assume:

- $\epsilon \sim \epsilon_0 + exp(\sigma_\epsilon)$;
- $\nu \sim U[\bar{\nu} - \Delta_\nu, \bar{\nu} + \Delta_\nu]$.

Set:

- r=1.052;
- $m = M_S^{0.16} M_B^{0.84}$ (Genesove and Han 2012);
- $\theta = 0.5$.

Calibration: procedure

- Each year is a steady state equilibrium.
- Fix Δ_ν.
- For each year t, we match the following moments exactly:
  - Average level of PD (\$): σ_ϵ^t
  - Sales volume: λ_m^t
  - Average price (\$): ϵ_0^t, $\bar{\nu}^t$
  - Average number of house visits by buyers (Genesove and Han 2012): $\epsilon_0^t - \bar{\nu}^t$, M_B^t
  - Time on market: M_B^t.

Calibration: procedure

Recall: tight theoretical relationship between Δ_v and $\text{corr}(\text{PD}, \text{TOM})$.

Calibrate Δ_v by matching $\text{corr}(\text{PD}, \text{TOM})$ in model to data.

How to get $\text{corr}(\text{PD}, \text{TOM})$ from data?

- Multiple estimates, lower will imply smaller dispersion in seller values.
- We use the panel coefficient, which is the smallest.

How to get predicted $\text{corr}(\text{PD}, \text{TOM})$ from the model?

- For each year, create a grid of TOM's to match x-sectional distribution in data.
- Run a pooled regression of simulated PD on TOM with year FE to match the coefficient in the data.

Calibration: results

	2010	2011	2012	2013	2014	2015	2016
Moments:							
PD	0.387	0.387	0.387	0.387	0.387	0.387	0.387
TOM	0.330	0.328	0.302	0.265	0.257	0.251	0.245
$\text{corr}(\text{PD}, \text{TOM})$	0.289	0.289	0.289	0.289	0.289	0.289	0.289
Sales volume	0.035	0.035	0.039	0.045	0.045	0.050	0.052
Average price	2.152	2.022	2.040	2.204	2.349	2.495	2.662
House visits	9.96	9.96	9.96	9.96	9.96	9.96	9.96
Calibrated parameters:							
$\bar{v}$	0.871	0.728	0.636	0.623	0.730	0.841	0.977
δ_v	1.110	1.110	1.115	1.126	1.130	1.132	1.140
ϵ_0	0.489	0.347	0.309	0.406	0.543	0.678	0.838
λ_ϵ	1.508	1.504	1.463	1.410	1.400	1.386	1.376
λ_m	0.037	0.035	0.040	0.046	0.046	0.051	0.053
M_B	0.701	0.662	0.770	0.895	0.902	1.012	1.054