Cheung, Lydia; Galimberti, Jaqueson K.; Vermeulen, Philip

Working Paper

Evidence on the determinants and variation of idiosyncratic risk in housing markets

ADB Economics Working Paper Series, No. 783

Provided in Cooperation with:
Asian Development Bank (ADB), Manila

APPENDIX

Evidence on the Determinants and Variation of Idiosyncratic Risk in Housing Markets

Lydia Cheung, Jaqueson K. Galimberti, and Philip Vermeulen

ADB Economics Working Paper No. 783
May 2025

A TECHNICAL APPENDIX

A.1 Repeat Sales Model

Compounding of (1) provides a model for the capital gain over k periods given by

$$p_{i,t+k} - p_{i,t} = a_{r,t,k} + b_{t,k} p_{i,t} - c_{t,k} + d_{t,k} \Psi X_{i,t} + e_{i,t,k}, \tag{A.1}$$

where $a_{r,t,k}$, $b_{t,k}$, and $d_{t,k}$ are functions of the parameters of equation (1), $c_{t,k}$ is a composite of median regional prices, $\widetilde{p}_{r,t}$, and $e_{i,t,k}$ is also a composite of the original idiosyncratic shocks, $\varepsilon_{i,t}$. Specifically, it can be shown that the parameters and error term of equation (A.1) evolve with the repeat sale interval ($k > 1$) according to

$$a_{r,t,k} = a_{r,t,k-1} (1 + \beta_{t+k-1}) + \alpha_{r,t+k-1}, \tag{A.2}$$
$$b_{t,k} = b_{t,k-1} (1 + \beta_{t+k-1}) + \beta_{t+k-1}, \tag{A.3}$$
$$c_{t,k} = c_{t,k-1} (1 + \beta_{t+k-1}) + \beta_{t+k-1} \widetilde{p}_{r,t+k-1}, \tag{A.4}$$
$$d_{t,k} = d_{t,k-1} (1 + \beta_{t+k-1}) + 1, \tag{A.5}$$
$$e_{i,t,k} = e_{i,t,k-1} (1 + \beta_{t+k-1}) + \varepsilon_{i,t+k-1}, \tag{A.6}$$

where the recursions depart from $a_{r,t,1} = \alpha_{r,t}$, $b_{t,1} = \beta_t$, $c_{t,1} = \beta_t \widetilde{p}_{r,t}$, $d_{t,1} = 1$, and $e_{i,t,1} = \varepsilon_{i,t}$. Note $d_{t,k}$ is derived under the assumption that changing house features have evolved according to an arithmetic average during the (unobserved) periods between repeat sales.

The coefficients α_t and β_t can be estimated from repeat sales data using nonlinear estimation methods on the system of equations characterized by one equation (A.1) for each possible pair of repeat sales periods, and the restrictions imposed by equations (A.2)-(A.5). Note the number of equations in this system grows with the number of periods in the sample (T) according to the binomial coefficient formula, $T!/(T-2)!2$.

A.2 Maximum Likelihood Estimation

The variances of the idiosyncratic shocks in equation (1) can be estimated by maximum likelihood. Assuming $\varepsilon_{i,t} \sim N\left(0, \sigma^2_{\varepsilon,i,t}\right)$, it follows from (A.6) that $e_{i,t,k} \sim N\left(0, \sigma^2_{e,i,t,k}\right)$, where

$$\sigma^2_{e,i,t,k} = \frac{1}{(1+\beta_{t+k})^2} \sum_{h=0}^{k-1} \sigma^2_{\varepsilon,i,t+h} \prod_{l=h+1}^{k} (1+\beta_{t+l})^2 . \tag{A.7}$$

Under normality, the probability density function of an estimated residual, from equation (A.1), is given by

$$f\left(e_{i,t,k}\right) = \frac{1}{\sigma_{e,i,t,k}\sqrt{2\pi}} \exp\left(-\frac{e^2_{i,t,k}}{2\sigma^2_{e,i,t,k}}\right), \tag{A.8}$$

and the corresponding log likelihood function for the sample of N estimated residuals is given by[1]

$$\ln LF\left(\left\{\sigma^2_{e,i,t,k}\right\}\right) = -\frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{N}\ln\sigma^2_{e,i,t_i,k_i} - \frac{1}{2}\sum_{i=1}^{N}\frac{e^2_{i,t,k}}{\sigma^2_{e,i,t_i,k_i}}. \tag{A.9}$$

Estimates of $\sigma^2_{\varepsilon,t}$ can be obtained by maximizing (A.9), using (A.7) to expand $\left\{\sigma^2_{e,i,t,k}\right\}$ in terms of $\left\{\sigma^2_{\varepsilon,i,t}\right\}$. To account for potential heteroskedasticity coming from regional variation, initial house price, and holding period, the variance of the idiosyncratic shocks can be estimated by setting $\sigma^2_{\varepsilon,i,t}$ according to (2). This only adds 2 + number of regions parameters (δ, ρ, and $\{\phi_r\}$) to the estimation of $\sigma^2_{\varepsilon,i,t}$, and a rescaling of $\sigma^2_{e,i,t,k}$ in equation (A.7) by $\exp\left(\phi_{r_i} + \delta\ddot{p}_i + \rho(k_i - 1)\right)$.

B SUPPLEMENTARY SAMPLE STATISTICS

Table B.1: Number of Observations by Region

Region	Repeat Sales	Houses	Resold 1x	Resold 2x	Resold >2x
Auckland	325,677	183,684	96,143	51,266	36,275
Bay of Plenty	70,002	38,975	19,979	10,981	8,015
Canterbury	140,602	79,479	41,710	22,195	15,574
Gisborne	8,897	4,614	2,148	1,332	1,134
Hawke's Bay	36,375	19,932	9,972	5,766	4,194
Manawatu-Wanganui	60,777	32,407	15,489	9,381	7,537
Marlborough	13,655	7,105	3,352	2,006	1,747
Nelson	14,361	7,823	3,932	2,195	1,696
Northland	25,746	14,767	8,005	3,968	2,794
Otago	59,688	32,025	15,798	8,948	7,279
Southland	33,020	15,535	6,392	4,315	4,828
Taranaki	29,525	15,572	7,479	4,385	3,708
Tasman	8,802	4,903	2,520	1,375	1,008
Waikato	101,635	54,822	27,039	15,494	12,289
Wellington	124,574	68,836	34,941	19,353	14,542
West Coast	5,055	3,082	1,780	833	469
Total	1,058,391	583,561	296,679	163,793	123,089

Source: Real Estate Institute of New Zealand.

[1] A slight abuse of notation here, where we let i index for the repeat sale observation, and t_i and k_i stand for observation i's period t and repeat sale interval k, respectively.

Table B.2: Summary Statistics of Property Characteristics

	Mean	Std Dev.	Min.	Max.	N.
Bedrooms	3.064	0.819	1	9	1,054,103
Bathrooms	1.432	0.693	0	6	541,315
Garages	1.464	0.871	0	7	1,026,847
Floor area (m^2)	143.106	65.163	1	885	1,029,021

Source: Authors' calculations.

Table B.3: Descriptive Statistics of Explanatory Variables

Variable	Mean	Std Dev.	Min.	Max.	N.
Interest rate$^{(a)}$	0.050	0.023	0.004	0.094	7,088,089
	[0.048]	[0.025]	[0.004]	[0.094]	[29]
Loan-to-value ratio$^{(a)}$	0.951	0.099	0.700	1.000	7,088,089
	[0.940]	[0.109]	[0.700]	[1.000]	[29]
Internet access, 2006, region$^{(b)}$	0.607	0.048	0.466	0.656	7,088,089
	[0.568]	[0.051]	[0.466]	[0.656]	[16]
Internet access, 2006, SA2$^{(b)}$	0.605	0.131	0.183	1.000	6,785,587
	[0.605]	[0.138]	[0.183]	[1.000]	[2,125]
Total population, national (million)$^{(b)}$	4.198	0.338	3.538	5.094	7,088,089
	[4.205]	[0.435]	[3.538]	[5.094]	[29]
Total population, region (million)$^{(b)}$	0.669	0.523	0.032	1.436	7,088,089
	[0.274]	[0.340]	[0.032]	[1.436]	[16]
New build consent area, national (million)$^{(b)}$	4.255	0.962	2.618	6.158	7,088,089
	[4.291]	[1.040]	[2.618]	[6.158]	[29]
New build consent area, region (million)$^{(b)}$	0.659	0.576	0.010	2.473	7,088,089
	[0.268]	[0.387]	[0.010]	[2.473]	[464]
Average house atypicality$^{(c)}$	0.758	0.037	0.637	0.919	6,834,835
N. sales, national (thousand)$^{(c)}$	82.375	17.089	55.505	119.701	7,088,089
N. sales, region (thousand)$^{(c)}$	12.785	10.682	0.230	39.798	7,088,089
N. sales, SA2$^{(c)}$	54.255	28.121	1.000	342.000	6,778,603

Notes: All variables are matched to the repeat sales data extended to cover every period between purchase and sale. Original statistics for the variables not derived directly from the repeat sales data are presented in brackets. SA2 stands for statistical area (level 2) definition from the New Zealand's 2018 Census of Population and Dwellings.
Sources: ^{a}Reserve Bank of New Zealand; ^{b}Stats New Zealand; ^{c}Authors' calculations based on data from Real Estate Institute of New Zealand.

C SUPPLEMENTARY RESULTS

Table C.1: Model Estimates of β_t

t	(1)	(2)	(3)	(4)
1992	-2.868	-2.896	-2.161	-1.979
	(0.427)	(0.431)	(0.738)	(0.608)
1993	-1.220	-1.182	-0.560	-0.896
	(0.339)	(0.344)	(0.610)	(0.519)
1994	-3.772	-3.713	-3.593	-3.938
	(0.310)	(0.314)	(0.552)	(0.460)
1995	-4.628	-4.619	-4.468	-4.090
	(0.318)	(0.321)	(0.569)	(0.456)
1996	-5.118	-5.045	-4.241	-5.085
	(0.315)	(0.318)	(0.557)	(0.433)
1997	0.631	0.730	0.805	0.705
	(0.358)	(0.363)	(0.619)	(0.462)
1998	1.083	1.083	0.998	1.053
	(0.375)	(0.377)	(0.647)	(0.475)
1999	0.944	0.893	1.281	0.946
	(0.371)	(0.371)	(0.654)	(0.464)
2000	1.363	1.370	0.949	1.216
	(0.346)	(0.346)	(0.607)	(0.445)
2001	0.198	0.101	1.119	-0.047
	(0.285)	(0.285)	(0.491)	(0.382)
2002	-1.969	-1.945	-1.450	-1.449
	(0.222)	(0.223)	(0.383)	(0.333)
2003	-6.793	-6.794	-7.331	-6.929
	(0.190)	(0.191)	(0.306)	(0.303)
2004	-9.002	-9.058	-9.223	-9.304
	(0.188)	(0.190)	(0.295)	(0.296)
2005	-6.984	-7.230	-7.450	-6.375
	(0.206)	(0.210)	(0.313)	(0.316)
2006	-5.366	-5.560	-6.136	-5.043
	(0.235)	(0.241)	(0.333)	(0.331)
2007	0.400	0.461	0.678	0.176
	(0.345)	(0.355)	(0.461)	(0.423)
2008	2.462	2.378	2.878	2.405
	(0.384)	(0.396)	(0.506)	(0.456)
2009	1.493	1.565	1.785	1.420
	(0.377)	(0.390)	(0.518)	(0.440)
2010	0.155	0.132	-0.004	0.047
	(0.382)	(0.395)	(0.507)	(0.450)
2011	-0.525	-0.472	-0.332	-0.566
	(0.350)	(0.364)	(0.438)	(0.423)
2012	-1.007	-0.960	-1.545	-0.947
	(0.322)	(0.338)	(0.392)	(0.397)
2013	-0.796	-0.819	-0.883	-0.694
	(0.308)	(0.326)	(0.369)	(0.397)
2014	-2.612	-2.462	-3.208	-2.630
	(0.302)	(0.318)	(0.359)	(0.395)
2015	-2.969	-2.838	-2.875	-2.844
	(0.283)	(0.294)	(0.344)	(0.372)

Continued on the next page

Table C.1. *Continued*

t	(1)	(2)	(3)	(4)
2016	-4.034	-4.206	-5.133	-4.064
	(0.308)	(0.311)	(0.374)	(0.394)
2017	-5.459	-5.420	-6.217	-5.536
	(0.336)	(0.336)	(0.387)	(0.424)
2018	-6.822	-6.832	-7.693	-6.675
	(0.352)	(0.352)	(0.393)	(0.442)
2019	-3.261	-3.223	-4.178	-3.227
	(0.395)	(0.395)	(0.422)	(0.509)
2020	-2.288	-2.244	-3.866	-1.761
	(0.453)	(0.453)	(0.477)	(0.591)
Sample	All	Excluding new dwelling	Excluding missing	Excluding hold < 2 years
R-squared	0.760	0.761	0.776	0.773
N. obs	1,058,391	1,026,613	522,616	947,089

Notes: These are estimates of β_t from equation 1. Robust standard errors are presented in brackets.
Source: Authors' calculations.

Table C.2: Estimates of Variances of Idiosyncratic Shocks

t	(1)	(2)	(3)	(4)
1992	1.788	1.731	1.631	1.695
	(0.062)	(0.062)	(0.096)	(0.234)
1993	1.663	1.634	1.576	1.534
	(0.048)	(0.048)	(0.074)	(0.191)
1994	1.593	1.540	1.488	1.413
	(0.043)	(0.042)	(0.063)	(0.142)
1995	1.534	1.483	1.384	1.584
	(0.041)	(0.040)	(0.063)	(0.119)
1996	1.432	1.397	1.323	1.203
	(0.037)	(0.036)	(0.060)	(0.099)
1997	1.490	1.462	1.397	1.419
	(0.042)	(0.042)	(0.064)	(0.100)
1998	1.318	1.287	1.198	1.077
	(0.040)	(0.039)	(0.059)	(0.100)
1999	1.599	1.550	1.593	1.508
	(0.046)	(0.046)	(0.080)	(0.103)
2000	1.466	1.450	1.422	1.156
	(0.043)	(0.043)	(0.069)	(0.105)
2001	1.780	1.753	1.699	1.762
	(0.047)	(0.047)	(0.073)	(0.106)
2002	2.075	2.051	1.992	2.018
	(0.044)	(0.044)	(0.068)	(0.097)
2003	1.688	1.675	1.613	1.563
	(0.033)	(0.033)	(0.047)	(0.086)
2004	1.498	1.469	1.397	1.435
	(0.030)	(0.030)	(0.043)	(0.073)
2005	1.363	1.336	1.309	1.304

Continued on the next page

Table C.2. *Continued*

t	(1)	(2)	(3)	(4)
	(0.029)	(0.029)	(0.043)	(0.071)
2006	1.173	1.146	1.087	1.010
	(0.025)	(0.025)	(0.034)	(0.065)
2007	1.480	1.438	1.441	1.190
	(0.033)	(0.033)	(0.045)	(0.068)
2008	1.285	1.258	1.220	1.127
	(0.033)	(0.033)	(0.041)	(0.079)
2009	1.226	1.189	1.121	1.045
	(0.037)	(0.037)	(0.045)	(0.087)
2010	1.282	1.271	1.272	1.229
	(0.042)	(0.043)	(0.051)	(0.095)
2011	1.206	1.198	1.202	0.953
	(0.038)	(0.039)	(0.046)	(0.096)
2012	1.369	1.352	1.360	1.275
	(0.041)	(0.042)	(0.049)	(0.091)
2013	1.391	1.368	1.430	1.188
	(0.041)	(0.042)	(0.050)	(0.091)
2014	1.425	1.396	1.453	1.298
	(0.037)	(0.038)	(0.047)	(0.083)
2015	1.337	1.318	1.371	1.097
	(0.032)	(0.032)	(0.037)	(0.072)
2016	1.246	1.221	1.204	1.094
	(0.034)	(0.033)	(0.040)	(0.072)
2017	1.199	1.180	1.163	1.120
	(0.035)	(0.035)	(0.038)	(0.072)
2018	1.164	1.144	1.148	0.948
	(0.034)	(0.033)	(0.034)	(0.067)
2019	1.447	1.429	1.357	1.357
	(0.041)	(0.040)	(0.039)	(0.069)
2020	1.639	1.611	1.474	1.563
	(0.061)	(0.061)	(0.051)	(0.097)
$\hat{\delta}$	-0.278	-0.286	-0.332	-0.275
	(0.008)	(0.008)	(0.011)	(0.008)
$\hat{\rho}$	-0.156	-0.152	-0.192	-0.117
	(0.003)	(0.003)	(0.005)	(0.003)
Sample	All	Excluding new dwelling	Excluding missing	Excluding hold < 2 years
Log likelihood (avg.)	0.189	0.193	0.257	0.142
N. obs	1,058,391	1,026,613	522,616	947,089

Notes: These are estimates of $\exp(\phi_{Akl})\sigma_t^2$ (scaled by 100x), δ, and ρ from equation 2. Note the estimates are scaled to Auckland region; see Table C.3 for the other regions' estimates of ϕ_r. Robust standard errors are presented in brackets.
Source: Authors' calculations.

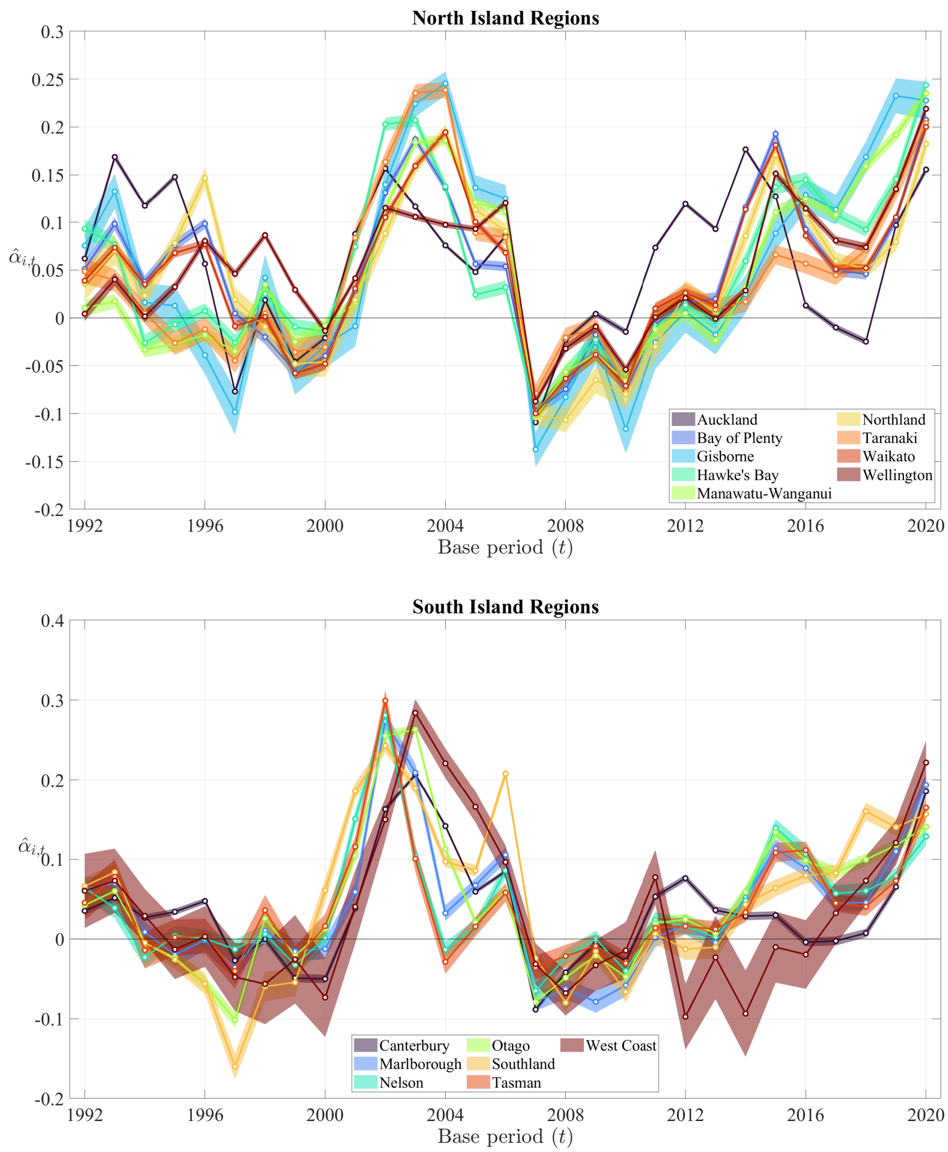
Table C.3: Estimates of Regional Shifts on Variances of Idiosyncratic Shocks

Region	(1)	(2)	(3)	(4)
Auckland	0	0	0	0
	–	–	–	–
Bay of Plenty	-0.121	-0.114	-0.096	-0.135
	(0.014)	(0.014)	(0.019)	(0.015)
Canterbury	-0.166	-0.166	-0.078	-0.179
	(0.011)	(0.011)	(0.016)	(0.012)
Gisborne	-0.046	-0.036	-0.064	-0.076
	(0.031)	(0.031)	(0.034)	(0.035)
Hawke’s Bay	-0.267	-0.260	-0.220	-0.284
	(0.018)	(0.019)	(0.026)	(0.021)
Manawatu-Wanganui	-0.048	-0.045	-0.028	-0.046
	(0.015)	(0.015)	(0.020)	(0.017)
Marlborough	-0.462	-0.460	-0.462	-0.481
	(0.034)	(0.035)	(0.050)	(0.039)
Nelson	-0.607	-0.599	-0.600	-0.647
	(0.031)	(0.031)	(0.043)	(0.035)
Northland	-0.009	-0.006	-0.174	-0.019
	(0.021)	(0.021)	(0.027)	(0.023)
Otago	0.115	0.120	0.100	0.112
	(0.014)	(0.014)	(0.020)	(0.015)
Southland	0.217	0.217	0.242	0.227
	(0.018)	(0.018)	(0.027)	(0.020)
Taranaki	0.107	0.101	0.023	0.118
	(0.018)	(0.018)	(0.024)	(0.020)
Tasman	-0.540	-0.543	-0.585	-0.544
	(0.041)	(0.041)	(0.052)	(0.045)
Waikato	-0.163	-0.156	-0.208	-0.166
	(0.012)	(0.012)	(0.017)	(0.013)
Wellington	-0.271	-0.266	-0.297	-0.295
	(0.013)	(0.013)	(0.018)	(0.014)
West Coast	0.443	0.445	0.417	0.469
	(0.049)	(0.050)	(0.041)	(0.051)
Sample	All	Excluding new dwelling	Excluding missing	Excluding hold < 2 years
N. obs	1,058,391	1,026,613	522,616	947,089

Notes: These are estimates of ϕ_r from equation 2 and taking Auckland as the base of reference. Robust standard errors are presented in brackets.
Source: Authors’ calculations.

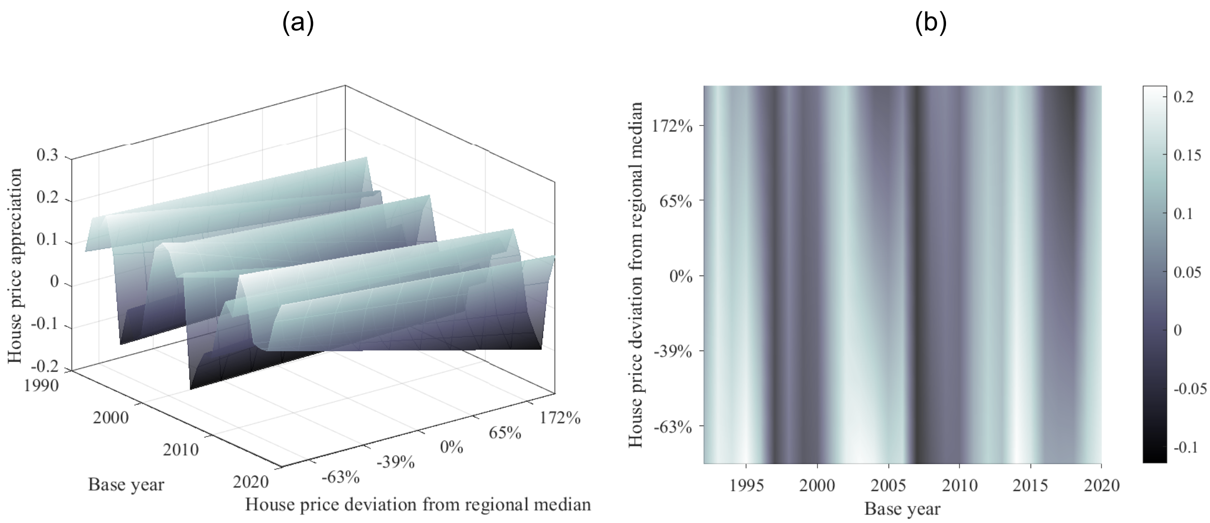
D ADDITIONAL FIGURES

Figure D.1: Regional Capital Gains Estimates



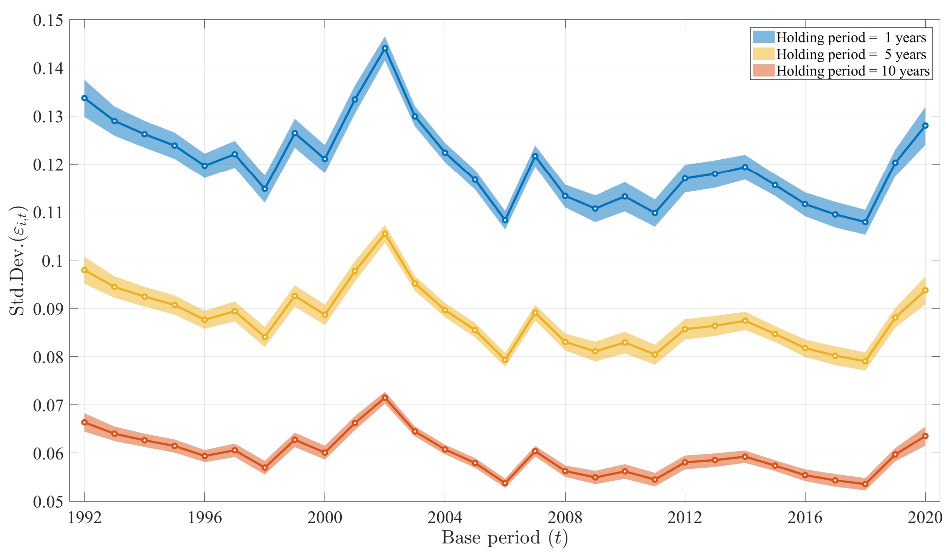
Notes: The lines depict estimates of the average (log) annual capital gains for a median house in the corresponding region. The shaded areas indicate the corresponding 90% confidence intervals.
Source: Authors' calculations.

Figure D.2: Predicted House Price Appreciation for Auckland



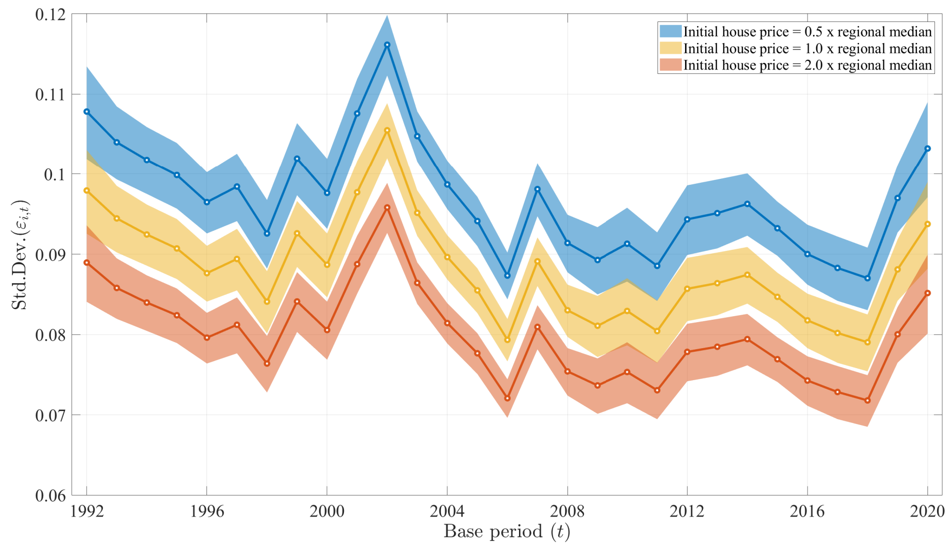
Notes: Predictions based on estimates of $\alpha_{Akl,t}$ and β_t from Equation 1 assuming no changes in house features. Panel (b) is the top view of (a).
Source: Authors' calculations.

Figure D.3: Idiosyncratic Risk Estimates by Holding Period



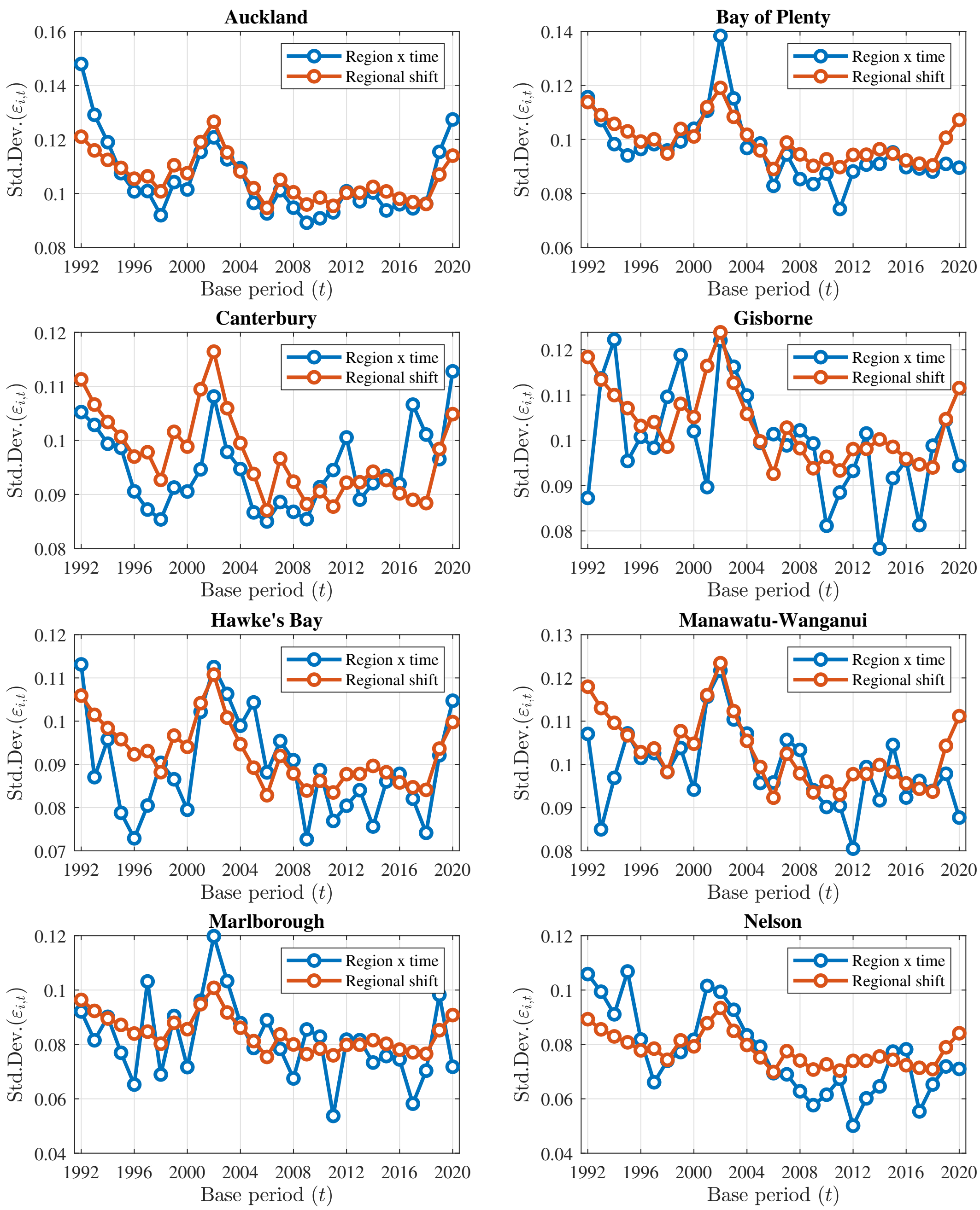
Notes: This is obtained by maximum likelihood from the residuals of the repeat sales model, assuming the idiosyncratic shocks are distributed as a Normal with variance given by equation 2. The estimates are for a median-priced house in Auckland region. The confidence bands are based on robust standard errors and calculated at a 90% confidence level.
Source: Authors' calculations.

Figure D.4: Idiosyncratic Risk Estimates by Initial House Price



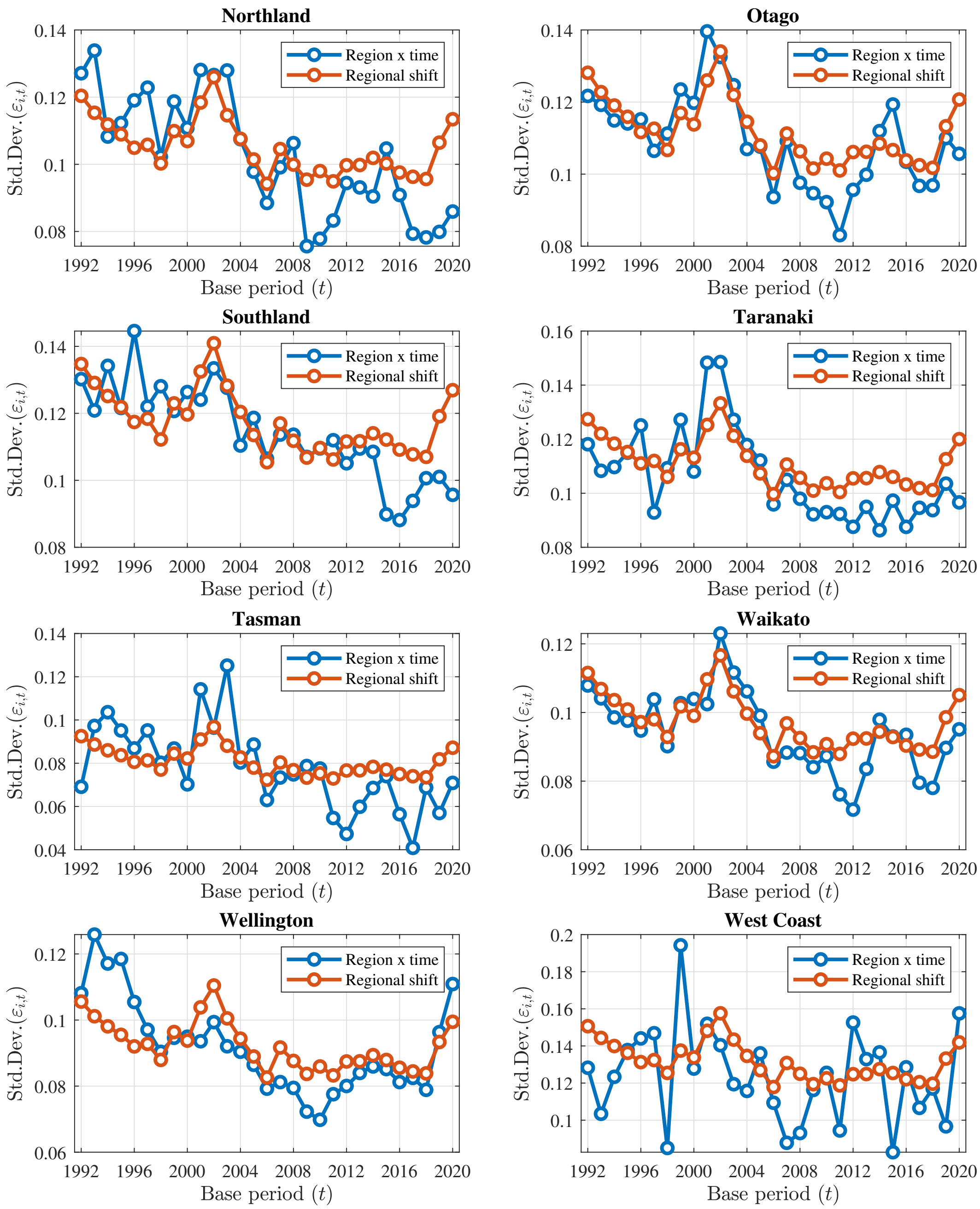
Notes: This is obtained by maximum likelihood from the residuals of the repeat sales model, assuming the idiosyncratic shocks are distributed as a Normal with variance given by equation 2. The estimates account for the effect of the initial house price, relative to the regional median. The estimates are for Auckland region and a holding period of 5 years between repeat sales. The confidence bands are based on robust standard errors and calculated at a 90% confidence level.
Source: Authors' calculations.

Figure D.5: Comparative of Regional Risk Estimates (1/2)



Notes: The plots compare implied risk estimates for each region, allowing for these to vary by region and time (blue lines) versus only by a regional shift (orange).
Source: Authors' calculations.

Figure D.6: Comparative of Regional Risk Estimates (2/2)



Notes: The plots compare implied risk estimates for each region, allowing for these to vary by region and time (blue lines) versus only by a regional shift (orange).
Source: Authors' calculations.