

# Online Appendix for: The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry

By JASON ALLEN AND ROBERT CLARK AND JEAN-FRANÇOIS HOUDE

## A: LARGE CANADIAN MERGERS

In this section we provide detailed information on the large Canadian mergers that are highlighted in Figure 1 of the paper.

*1993*

On January 1, 1993 TD Bank acquired, under duress, Central Guaranty Trust. Even though Central Guaranty Trust had a poor balance sheet, there was substantial interest by several financial institutions in acquiring its assets from Central Capital Corporation, which owned 87 per cent of trust. At the official auction TD won over the joint bid from National Bank, Canada Trust, and Montreal Trust. Given the conditions of the Central Guaranty Trust balance sheet (they had very high risk mortgages and commercial lending activities), the Canadian Deposit Insurance Corporation provided financial support to TD in the takeover. In terms of branches, TD acquired 156 locations. TD also inherited nearly 11 billion dollars in deposits.

Soon after TD Bank's acquisition of Central Guaranty Trust, the second largest acquisition (in terms of assets) in Canada occurred, with Royal Bank acquiring Royal Trust. In 1992 Royal Trust's parent company, Royal Trustco, had experienced liquidity issues, and in early 1993 announced it was looking for a buyer of the trust company. RBC's takeover of Royal Trust was announced on March 18, 1993 and consummated on September 1, 1993. The Royal Trust brand was well-known and well-respected in the financial industry. Furthermore, most people believed that RBC was a perfect match for Royal Trust. They shared the same name, colors, and both had distinguished histories. Royal Trust had 150 branches at the time of the acquisition, largely in Ontario and Quebec, but also with a significant presence in Alberta and British Columbia. According to Competition Bureau (2003), the RBC-Royal Trust merger was analyzed by the Bureau when it was first proposed. The Competition Bureau did not place any restrictions on the merger.

*1994*

The Competition Bureau also did not place restrictions on Bank of Nova Scotia's (BNS) acquisition of Montreal Trust on April 12, 1994. In this instance BNS acquired 9 billion dollars in deposits and 125 branches located mostly in Quebec and Ontario. Montreal Trust had experienced some losses in the early 1990s because of market value deficiencies in the investments and assets, but was considered a sound financial institution and a good purchase by BNS.

*1997*

On August 14, 1997 Bank of Nova Scotia acquired National Trust and Victoria and Grey Mortgage Corporation, without restrictions by the Competition Bureau. In this case BNS acquired nearly 12 billion dollars in deposits and 199 branches located in Quebec, Ontario, and Manitoba. Most of what is known about National Trust is confidential. We do know, however, that the transaction was

valued at approximately 1.25 billion dollars even though National Trust was considered a poorly run institution that had acquired an excessive number of small, failing trusts throughout the 1990s. It was largely the mis-management of the infrastructure that led National Trust to look for a buyer.

2000

The last bank merger to be approved in Canada was Toronto-Dominion Banks' (TD) acquisition of Canada Trust in 2000. The price tag was roughly 8 billion dollars (TD financed the purchase by issuing 700 million dollars in equity) and it resulted in 440 branches located across the country being acquired as well as a strong share of the mortgage market. The merger was analyzed by the Competition Bureau and allowed to be completed under minor conditions. For example, TD had to divest in some of its branches in three of the seventy-four markets defined by the Bureau (Kitchener-Waterloo-Cambridge-Elmira, Port Hope, and Brantford-Paris). TD also had to sell CT's MasterCard credit card business (they sold the consumer credit card business to Citibank in November 2000). TD was issuing Visa credit cards at the time of the acquisition, and it was not until 2009 that Canadian banks could sell both brands simultaneously.

## B: ESTIMATION STRATEGY

In Section 3 of the paper we explain our estimation strategy. Here we provide further details.

Our estimation focuses first on measuring the average treatment effect (ATE) of the merger on negotiated rates and on the probability of paying the posted rate (i.e. zero discount). We estimate the ATE using three econometric techniques: (i) linear difference-in-difference (DiD) estimated by OLS, (ii) the propensity-score matching DiD estimator developed by Heckman, Ichimura and Todd (1997), and (iii) the change-in-change (CiC) estimator proposed by Athey and Imbens (2006).

We think of the merger as an intervention ( $I_i$ ) that changes the rates negotiated between borrowers and lenders:

$$(B1) \quad \text{Rate}_i = I_i f^I(u_i, T_i) + (1 - I_i) f^N(u_i, T_i),$$

$\text{Rate}_i$  denotes the transaction mortgage rate,  $T_i$  is a before/after period indicator, and  $f^I(u_i, T_i)$  and  $f^N(u_i, T_i)$  correspond to the post- and pre-merger reduced-form pricing functions, respectively.

To measure residual price dispersion, we proceed by decomposing the observed margins into a deterministic function of borrowers' characteristics, and an idiosyncratic component  $m_i$ :

$$(B2) \quad \text{Margin}_i = \beta' \mathbf{X}_i + \mu_i^{\text{week}} + \theta_{G_i, T_i} + e_i = \beta' \mathbf{X}_i + m_i,$$

where  $\text{Margin}_i = \text{Rate}_i - \text{Bond}_i$ , and  $\theta_{G_i, T_i}$  is a group/period fixed-effect,  $G_i$  and  $T_i$  index the group (treatment or control) and time period (before or after the merger) for borrower  $i$ ,  $\mu_i^{\text{week}}$  is a closing-week fixed-effect, and  $\mathbf{X}_i$  is a vector of control variables. In the empirical analysis we mostly focus on the impact of the merger on the level and dispersion of *negotiated margins*,  $m_i$ . Note that we will sometimes refer to this as a residual price, especially when we relate our findings to the literature on residual price dispersion.

To perform the estimation, we impose three assumptions on the residual margin function and the distribution of  $u_i$ . First, in order to compare treated and control markets, we assume that the function  $f^N(u, T)$  is common across treated and control consumers (hence the importance of using comparable markets with intermediate levels of competition). Second, residual margins are monotonically increasing functions of  $u_i$ . Ceteris paribus, a high  $u_i$  consumer pays a high rate, and receives a small discount. Finally, the distribution of unobserved heterogeneity is assumed to be constant over time, but can differ across control and treated markets. This assumption implies

that the effect of the merger could differ across the two groups, even though the residual margin function is common. We let  $H_1(u)$  and  $H_0(u)$  denote the distributions of  $u$  within treated and control areas, respectively.

All three assumptions are fairly standard in the difference-in-difference literature (e.g. Abadie (2005)), and are adapted to our context, following Athey and Imbens (2006), to reflect the fact that the errors are non-separable. Below we describe the three estimators used to characterize the average and distributional impact of the merger.

**Linear difference-in-difference estimator** We first use a linear difference-in-difference (DiD) approach to measure the average treatment effect (ATE) of the merger. In our context, the DiD model linearly approximates the margin function above, and estimates equations (B2) and (B1) jointly. Moreover, the linearity assumption allows us to control for location (FSA) fixed-effects. This leads to the following estimating equation:

$$(B3) \quad \text{Margin}_i = \bar{\alpha}I_i + \beta' \mathbf{X}_i + \mu_i^{\text{week}} + \mu_i^{\text{fsa}} + e_i,$$

where  $\mathbf{X}_i$  is a vector of control variables, and  $\mu_i^{\text{week}}$  and  $\mu_i^{\text{fsa}}$  are week and FSA fixed-effects.

An important concern is that the merger predominantly affected a small number of provinces, and therefore might be correlated with location-specific housing or financial variables affecting negotiated interest rates. For instance, house prices and the timing of purchases follow regional cycles and trends that could be confounded with the merger. Furthermore, local housing market trends may impact mortgage rates differently in treated and control markets. We consider two sets of control variables to address this question.

Our baseline specification includes controls that describe the financial and demographic characteristics of the contract, and the identity of lenders. We use the vector  $\mathbf{X}_i^{\text{base}}$  to denote these controls. In an effort to further control for confounding factors we also estimate  $\bar{\alpha}$  using a richer set of variables, labelled  $\mathbf{X}_i^{\text{trend}}$ , that control for heterogeneous observable trends across our control and treatment groups. We do so by interacting every consumer and contract characteristics in  $\mathbf{X}_i$  with  $T_i$ , our “post-merger” dummy variable. We also include in  $\mathbf{X}_i^{\text{trend}}$  province-level linear and quadratic trends. We present the empirical results using both  $\mathbf{X}_i^{\text{base}}$  and  $\mathbf{X}_i^{\text{trend}}$  specifications.

In addition to estimating the ATE for margins, we study the impact on the probability of paying the posted rate. Using the same two specifications, we replace  $p_i$  by an indicator variable equal to one if the consumer pays a rate greater than or equal to the posted-rate plus 10 bps.

**Matching estimator** An alternative approach to control for differences between control and treatment groups is to estimate  $\bar{\alpha}$  using a matching DiD estimator. Since we want to compare local markets along multiple dimensions, we implement a matching DiD estimator based on a propensity score function, and measure the degree of similarity between local markets using a Kernel smoothing function. In addition to the three assumptions imposed on  $u_i$  already mentioned, this estimator also imposes the restriction that treated and control neighborhoods have a significant overlap in terms of the probability of being affected by the merger. We therefore exclude neighborhoods in the top and bottom 2% of the propensity score distribution.

In order to control for borrower-specific observed covariates, we calculate the matching estimator using negotiated margins. This leads to a sequential estimator: (i) estimate  $m_i$  by OLS from equation B2, (ii) estimate the propensity score function via a Logit model (i.e. probability of having both  $A$  and  $B$  in a neighborhood), and (iii) estimate the ATE of the merger using the estimator of Heckman, Ichimura and Todd (1997). The propensity score controls for time-invariant demographic characteristics and branch concentration of each neighborhood.<sup>1</sup>

<sup>1</sup>The list of controls are demographic characteristics measured using 2001 census tables at the FSA level and branch location

This estimator relaxes the assumption that the function  $f^N(u_i, T_i)$  is common across control and treated local markets. Instead, we are implicitly assuming that neighborhoods with similar propensity scores have homogenous pricing functions.

**Change-in-change estimator** We apply the change-in-change (CiC) estimator proposed by Athey and Imbens (2006) to study the distributional impact of the merger on negotiated margins. In addition to the three assumptions above, we also assume that the support of  $u_i$  in the treatment group overlaps with the support of  $u_i$  in the control group. We impose this restriction by dropping the non-overlapping negotiated margins across the two groups for each time period.

In practice, we also assume that the observed characteristics of the contracts  $\mathbf{X}_i$  are independent of  $u_i$ . Like with the matching estimator, this allows us to perform the estimation using the negotiated margins distribution defined in equation B2, and to control for observable characteristics of the contracts in the parametric form suggested by equation B2.<sup>2</sup>

Under these assumptions, the change in the  $q^{th}$  percentile of the margin distribution in the control group identifies the change in the markup function that is strictly due to time. Using this logic, and the fact that the distributions are invertible (under the monotonicity assumption), we can recover the counter-factual distribution of prices in the treatment group for any point  $m_i$  in the common support:

$$(B4) \quad F_{1,1}^c(m_i) = F_{1,0} \left( F_{0,0}^{-1} (F_{0,1}(m_i)) \right),$$

where the superscript  $c$  identifies the counter-factual distribution, and  $F_{G,T}(m)$  is the CDF of margins in the sub-population  $(G, T)$ . Intuitively, we obtain the counter-factual distribution by transforming the observed negotiated margin distribution in the treatment group at time 0 (i.e.  $F_{1,0}(m)$ ) to mimic the change in the distribution observed in the control group. We then use the empirical counterparts of  $F_{1,1}^c(m)$  and  $F_{1,1}$  to calculate the impact of the merger on various moments of the margin distribution, such as the average, the standard-deviation and the inter-quartile range. We also recover an estimate of the effect of the merger for different percentiles  $q$ :

$$(B5) \quad \hat{\alpha}(q) = \hat{F}_{1,1}^{-1}(q) - \hat{F}_{0,1}^{-1} \left( \hat{F}_{0,0} \left( \hat{F}_{1,0}^{-1}(q) \right) \right),$$

where  $\hat{F}(\cdot)$  refers to the empirical distribution function (EDF), and  $\hat{F}^{-1}(\cdot)$  is the inverse of the EDF. Since we can normalize the scale of  $u_i$  to be between zero and one,  $\alpha(q)$  measures the estimated effect of the merger on a consumer of type  $u_i = q$ .

## C: ROBUSTNESS ANALYSIS

In this section we provide further details to Section 3.1, on the robustness of the average merger effect to alternative estimation methods, and sample selection. We present robustness results with respect to the choice of controls, the size of local neighborhoods, and different event windows. We present results for transaction margins and discounting.

In Table C1 we present the average effect of the merger estimated with alternative sets of control variables. Our baseline specification is reproduced in column (2) and the trend results are reproduced in column (3). Column (1) shows that failure to control for characteristics of the transaction

data and include: average house value and income to house value ratio, fraction of households that rent, average age, inter-province migration ratio, fraction of people with a university degree, number of dwellings, pre-merger branch-HHI, and number of lenders pre-merger.

<sup>2</sup>This is the approach suggested by Athey and Imbens (2006) (pages 465-466). See Linton, Massoumi and Whang (2005) for a related two-step procedure.

Table C1—: Alternative linear difference-in-difference specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Merger	0.0623 <sup>a</sup> (0.0188)	0.0607 <sup>a</sup> (0.0183)	0.0719 <sup>a</sup> (0.0242)	0.0500 <sup>b</sup> (0.0251)	0.0564 <sup>a</sup> (0.0193)	0.0602 (0.0374)
Observations	18,121	18,121	18,121	18,121	15,770	13,216
R-squared	0.344	0.408	0.416	0.417	0.415	0.420
Location (FSA) FE	✓	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓	✓
Control group	Full	Full	Full	Full	A	None
X Controls		✓	✓	✓	✓	✓
D × X & Province trends			✓	✓		
D × Province FE				✓		
% Treated	0.628	0.628	0.628	0.628	0.722	0.862

The dependent variable is margins. Significance levels: <sup>a</sup>  $p < 0.01$ , <sup>b</sup>  $p < 0.05$ , <sup>c</sup>  $p < 0.1$ . Standard errors are clustered at the location level (FSA).

does not lead to a substantial change in the ATE. In column (4) we introduce provincial fixed effects interacted with the post-merger dummy variable. Since a large fraction of observations are clustered in a few provinces this specification adds substantial strain to the data. However, we still estimate an ATE of 5 bps that is statistically significantly different from zero at the 5% confidence level. The point estimates remain fairly constant across all four specifications, which suggests that our conclusions are not sensitive to the choice of controls.

In columns (5) and (6) we modify the definition of the control group. First we restrict the control group to markets in which only bank *A* is present, and then to markets where neither *A* nor *B* are present. The “only *A*” specification raises the fraction of treated observations to 72%, but leaves the point estimate largely unchanged (i.e. 5.64 versus 6.04 bps). In contrast, in the “no *A* or *B*” specification the point estimate is equal to 6.02 bps, but the standard-error nearly doubles. This decrease in precision is due to the fact that the fraction of observations belonging to the control group goes down from 43% to less than 14%. Despite this, the results from these two specifications suggest that the magnitude of the estimated average merger effect is not driven by changes occurring in the control markets, such as the re-branding of bank *A* post-merger.

Table C2 repeats the same robustness exercise as Table C1 but for discounting instead of margins. The baseline specification is reproduced in column (2) and the trend specification in column (3). The point estimates remain relatively constant for all specifications, including in columns (5) and (6) where we modify the control groups.

Our second set of robustness tests considers the impact of different neighborhood sizes. Our main specification assumes that consumers shop within a 5KM radius around the centroid of their FSA’s. Increasing this threshold tends to raise the number of FSA’s directly affected by the merger, at the cost of including lenders that are not considered by the average consumer (falsely treated neighborhoods). Using a smaller threshold reduces the number of treated neighborhoods, and includes in the control group areas that are affected by the merger. Therefore, over-estimating or under-estimating the size of consumers’ choice-sets exacerbates the measurement error problem in our treatment variable, and biases our results towards zero. The estimation results, reproduced in Table C3a and C4a, confirm this intuition. The effect of the merger is statistically different from zero for most neighborhoods, but is smaller and less precisely estimated for smaller and larger

Table C2—: Zero Discount Robustness to alternative specifications

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Merger	0.0638 <sup>a</sup> (0.0166)	0.0646 <sup>a</sup> (0.0154)	0.0477 <sup>a</sup> (0.0180)	0.0525 <sup>a</sup> (0.0187)	0.0603 <sup>a</sup> (0.0162)	0.0805 <sup>b</sup> (0.0353)
Observations	18,121	18,121	18,121	18,121	15,770	13,216
R-squared	0.118	0.181	0.189	0.190	0.179	0.179
Location FE	✓	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓	✓
Control group	Full	Full	Full	Full	A	None
X Controls		✓	✓	✓	✓	✓
D × X & Province trends			✓	✓		
D × Province FE				✓		

radii. This is particular true when we control for province-level trends in the larger neighborhood specifications, since the number of provinces present in both treated and control groups shrinks (i.e. it becomes harder to distinguish between the merger effect and other province-level trends).

Finally, we consider different event windows and present the results of a falsification exercise in Table C3 for margins and Table C4 for discounting. Recall that throughout the paper we use contracts signed at most one year before or after the official merger date. In columns (1) and (3) we compare our main specification with a shorter window of six months and a longer window of eighteen months. The results suggest that the effect of the merger was strongest within six months of the announcement, but diminished over time, possibly due to the closing of branches.

In columns (4) and (5) we consider two alternative merger dates: six months before and six months after the actual date. For margins, in both cases, we fail to find any evidence of rate increase, suggesting that our estimates are not confounded with the presence of unobserved factors affecting the treated neighborhoods within a year of the merger. In the case of discounting, there coefficient on minus 6 months is statistically significant in the baseline specification. For the probability of receiving a discount, therefore, it is likely that there is some unobservable that is confounded with the ATE when we do not control for trends.

Another way of looking for pre-intervention differences in trends between the two groups is to plot the evolution of transaction rates over time. In Figure C1a and C1b we plot the mean and median differences in the transaction margins between treated and control groups. Each dot represents a coefficient  $\theta_t$  from the following regression equation:

$$(C1) \quad \text{Margin}_i = \sum_{t=1}^T \theta_t 1(i \in t) \times G_i + \sum_{t=1}^T \beta_t 1(i \in t) \times G_i + u_i,$$

where  $t$  is a 30-day period before or after the merger,  $1(i \in t)$  is a dummy variable equal to one if contract  $i$  was negotiated within period  $t$ , and  $G_i$  is a dummy variable equal to one if contract  $i$  originates from the treatment group. The coefficient  $\theta_t$  therefore measures the difference between treated and control groups for period  $t$ . We estimate this equation by OLS to measure the mean difference, and by quantile-regression to estimate the median difference. The confidence intervals around the mean differences are calculated by clustering the standard-errors at the FSA level,

Table C3—: Robustness analysis results on the average merger effect

(a) Alternative neighborhood sizes

	(1)	(2)	(3)	(4)	(5)
	3KM	4KM	5KM	6KM	7KM
Baseline	0.0405 <sup>b</sup> (0.0184)	0.0529 <sup>a</sup> (0.0198)	0.0607 <sup>a</sup> (0.0183)	0.0478 <sup>b</sup> (0.0198)	0.0304 (0.0204)
With trends	0.0242 (0.0213)	0.0496 <sup>c</sup> (0.0262)	0.0719 <sup>a</sup> (0.0242)	0.0414 (0.0258)	0.0173 (0.0266)
Observations	18,121	18,121	18,121	18,121	18,121
% Treated	0.479	0.589	0.628	0.706	0.728

(b) Alternative event windows and falsification tests

	(1)	(2)	(3)	(4)	(5)
	6-Months	1-Year	18-Months	+ 6 Months	- 6 Months
Baseline	0.0969 <sup>a</sup> (0.0279)	0.0607 <sup>a</sup> (0.0183)	0.0172 (0.0152)	-0.0135 (0.0254)	0.0166 (0.0177)
With trends	0.0677 <sup>c</sup> (0.0391)	0.0719 <sup>a</sup> (0.0242)	0.0444 <sup>b</sup> (0.0200)	-0.0250 (0.0343)	0.0100 (0.0217)
Observations	8,040	18,121	27,978	17,296	17,148

Dependent variable is margins. Significance levels: <sup>a</sup> p<0.01, <sup>b</sup> p< 0.05, <sup>c</sup> p< 0.1. All specifications include week and location (FSA) fixed effects, these are included in  $X_i^{base}$ . Trend specifications also include the variables in  $X_i^{trend}$ . Standard errors are clustered at the location level (FSA).

and the quantile regression confidence intervals are obtained by bootstrapping. The two figures also present the weighted average of every coefficient before and after the merger, weighted by the sample size within each period. Finally, the number of months post-merger is smaller than 12 since we drop contracts for which the closing date is 90 days after the merger date.

The results of this exercise illustrate very well the average and quantile effects of the merger. The OLS result shows a 6 to 7 basis point increase in the average margin. This change was sudden and lasted until the 9th month after the merger date. Similarly, the median difference increased by 10 basis points post merger, and lasted 7 months. Importantly, both figures clearly show that the margin difference between the two groups before the merger was very stable and statistically significant. It became closer to zero and statistically indistinguishable from zero post-merger.

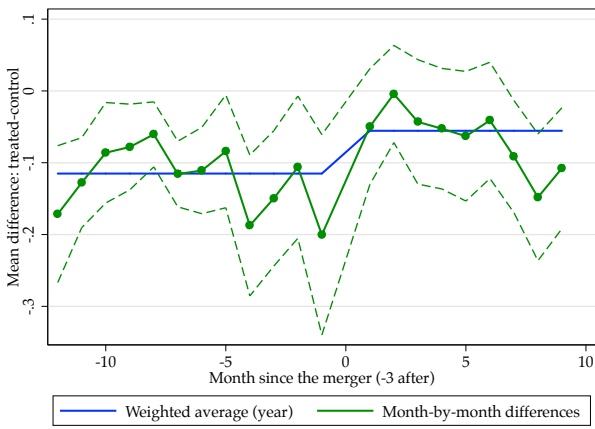
The fact that the negotiated margins eventually reverted to their pre-merger levels is consistent with the results reported in Table C4 concerning the larger time window. We conjecture that this is due to the fact that the merging firm started closing duplicate branches about a year after the merger.

Table C4—: Robustness analysis results on the average merger effect: Dependent variable is zero discount

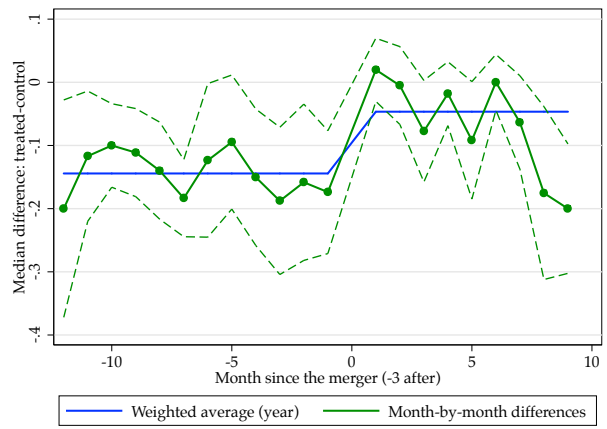
(a) Alternative neighborhood sizes					
	(1)	(2)	(3)	(4)	(5)
	3KM	4KM	5KM	6KM	7KM
Baseline	0.0474 <sup>†</sup> (0.0146)	0.0617 <sup>†</sup> (0.0157)	0.0646 <sup>†</sup> (0.0154)	0.0615 <sup>†</sup> (0.0174)	0.0522 <sup>†</sup> (0.0182)
With trends	0.0254 (0.0168)	0.0414 <sup>b</sup> (0.0183)	0.0477 <sup>†</sup> (0.0180)	0.0296 (0.0208)	0.0129 (0.0204)
Observations	18,121	18,121	18,121	18,121	18,121
% Treated	0.479	0.589	0.628	0.706	0.728
† p<0.01, <sup>b</sup> p<0.05, * p<0.1					
(b) Alternative event windows and falsification tests					
VARIABLES	(1)	(2)	(3)	(4)	(5)
	6-Months	1-Year	18-Months	+ 6 Months	- 6 Months
Baseline	0.0617 <sup>†</sup> (0.0229)	0.0646 <sup>†</sup> (0.0154)	0.0382 <sup>†</sup> (0.0128)	0.00264 (0.0214)	0.0658 <sup>†</sup> (0.0168)
with trends	0.0675 <sup>b</sup> (0.0288)	0.0477 <sup>†</sup> (0.0180)	0.0202 (0.0156)	0.0111 (0.0262)	0.0258 (0.0219)
Observations	8,040	18,121	27,978	17,296	17,148
† p<0.01, <sup>b</sup> p<0.05, * p<0.1					

Significance levels: <sup>a</sup> p<0.01, <sup>b</sup> p< 0.05, <sup>c</sup> p< 0.1. All specifications include week and location (FSA) fixed effects, these are included in  $X_i^{base}$ . Trend specifications also include the variables in  $X_i^{trend}$  Standard errors are clustered at the location level (FSA).





(a) Mean differences



(b) Median differences

Figure C1. : Results for the month-by-month difference in transaction margins across treated and control groups

D: ADDITIONAL TABLES

In this section we present additional tables.

Table D1—: Definition of Household / Mortgage Characteristics

Name	Description
FI	Type of lender
Source	Identifies how lender generated the loan (branch, online, broker, etc)
Income	Total amount of the borrower(s) salary, wages, and income from other sources
TDS	Ratio of total debt service to income
Duration	Length of the relationship between the borrower and FI
R-status	Borrowers residential status upon insurance application
FSA	Forward sortation area of the mortgaged property
Market value	Selling price or estimated market price if refinancing
Dwelling type	10 options that define the physical structure
Close	Closing date of purchase or date of refinance
Loan amount	Dollar amount of the loan excluding the loan insurance premium
Premium	Loan insurance premium
Purpose	Purpose of the loan (purchase, port, refinance, etc.)
LTV	Loan amount divided by lending value
Price	Interest rate of the mortgage
Term	Represents the term over which the interest rate applies to the loan
Amortization	Represents the period the loan will be paid off
Interest type	Fixed or adjustable rate
<i>CREDIT</i>	Summarized application credit score (minimum borrower credit score).

Some variables were available only from one of the mortgage insurers.

Table D2—: Average Treatment Effect of the Merger with Covariates

VARIABLES	Margin		Zero discount	
	Baseline	With trend	Baseline	With trend
Merger	0.0607 <sup>a</sup> (0.0183)	0.0719 <sup>a</sup> (0.0242)	0.0646 <sup>a</sup> (0.0154)	0.0477 <sup>a</sup> (0.0180)
log of Monthly income (X 100K)	0.150 <sup>a</sup> (0.0181)	0.123 <sup>a</sup> (0.0237)	0.109 <sup>a</sup> (0.0165)	0.0909 <sup>a</sup> (0.0220)
Log of house price (X 100K)	-0.310 <sup>a</sup> (0.0194)	-0.320 <sup>a</sup> (0.0237)	-0.220 <sup>a</sup> (0.0171)	-0.231 <sup>a</sup> (0.0219)
Log of other debt (X 1K)	-0.0511 <sup>a</sup> (0.00759)	-0.0493 <sup>a</sup> (0.00985)	-0.0310 <sup>a</sup> (0.00665)	-0.0364 <sup>a</sup> (0.00883)
FICO $\geq$ 600	-0.225 <sup>a</sup> (0.0354)	-0.235 <sup>a</sup> (0.0435)	-0.197 <sup>a</sup> (0.0305)	-0.198 <sup>a</sup> (0.0402)
Status: Renter	-0.0101 (0.0113)	-0.0112 (0.0159)	-0.00935 (0.0101)	-0.0221 (0.0140)
Status: Parents	-0.0547 <sup>a</sup> (0.0190)	-0.0541 <sup>b</sup> (0.0274)	-0.0517 <sup>a</sup> (0.0155)	-0.0630 <sup>a</sup> (0.0225)
0.85 $\leq$ LTV $\leq$ 90	0.0509 <sup>a</sup> (0.0113)	0.0404 <sup>a</sup> (0.0150)	0.0257 <sup>b</sup> (0.0103)	0.0212 (0.0133)
0.90 $\leq$ LTV $\leq$ 0.95	0.0650 <sup>a</sup> (0.0147)	0.0358 <sup>c</sup> (0.0204)	0.0410 <sup>a</sup> (0.0130)	0.0279 (0.0176)
I(Min.down)	0.118 <sup>a</sup> (0.0114)	0.107 <sup>a</sup> (0.0155)	0.0874 <sup>a</sup> (0.0112)	0.0824 <sup>a</sup> (0.0142)
Switcher	-0.0899 <sup>a</sup> (0.0129)	-0.0488 <sup>b</sup> (0.0189)	-0.0564 <sup>a</sup> (0.0123)	-0.0119 (0.0185)
Broker	-0.0615 <sup>a</sup> (0.0141)	-0.0378 <sup>b</sup> (0.0185)	-0.0945 <sup>a</sup> (0.0122)	-0.0904 <sup>a</sup> (0.0176)
R-squared	0.408	0.416	0.181	0.189

Significance levels: <sup>a</sup> p<0.01, <sup>b</sup> p< 0.05, <sup>c</sup> p< 0.1. Number of observations ins 18,121.

E: THEORY APPENDIX

In this section we present further details on the model presented Section 4 of the paper.

In the interior, the solution to the effort-choice problem in equation (F3) gives rise to search effort and matching probability functions that are decreasing in  $u_i$  and increasing in  $n$ :

$$(E1) \quad e(u_i, n) = \left[ (u_i/\lambda\Delta(n))^{-1/(1+\lambda)} - 1 \right]$$

$$(E2) \quad s(u_i, n) = \begin{cases} 1 - (u_i/\lambda\Delta(n))^{\lambda/(1+\lambda)} & \text{if } \lambda\Delta(n) > u_i \\ 0 & \text{otherwise,} \end{cases}$$

where  $\lambda = \gamma/(1 - \gamma)$ .

To study the effect of a merger we consider the effect of  $u_i$  and  $n$  on the equilibrium margin:

$$\pi_1(u_i, n) = 1 + e(u_i, n) \left( 1 - \frac{1}{1+\lambda} + \frac{1}{1+\lambda} \right) - \frac{1}{1+\lambda} + \frac{1}{1+\lambda} = 1 + e(u_i, n) > 0$$

$$\pi_2(u_i, n) = \frac{u_i(e(u_i, n) + 1)}{\Delta(n)(1+\lambda)} \Delta'(n) - \frac{u_i(e(u_i, n) + 1)}{\Delta(n)(1+\lambda)} \Delta'(n) - s(u_i, n) \Delta'(n) = -s(u_i, n) \Delta'(n) < 0.$$

Therefore price is increasing in  $u_i$ , and decreasing in  $n$ . We can also use equation (E2) to determine the threshold level  $\bar{u}(n)$ , associated with positive search effort:  $e(u_i) = 0$  if  $u_i > \lambda\Delta(n)$ . If the search function  $s(u_i, n)$  is such that  $s(u_i, n) = 0$  for all  $u_i > \bar{u}$ , the marginal effect of changing  $n$  is zero for consumers at the top of the price distribution.

Using these two results leads to our first prediction: A merger raises prices for all  $u_i < \bar{u}$ :  $\pi(u_i, n - 1) - \pi(u_i, n) > 0$  for all  $u_i < \bar{u}(n)$ .

We are also interested in the effect of the merger along the distribution. From the above results we can calculate the following cross-partial derivatives:

$$(E3) \quad \frac{\partial^2 \pi(u_i, n)}{\partial u \partial n} = \frac{\partial^2 \pi(u_i, n)}{\partial n \partial u} = \frac{e(u, n)}{\Delta(n)(1+\lambda)} \Delta'(n) > 0.$$

From these we see that because the marginal effect of  $n$  is increasing in  $u_i$ , the effect of a merger is larger (more negative) at the bottom of the price distribution:

$$(E4) \quad 0 \leq \pi(u_{q_1}, n - 1) - \pi(u_{q_1}, n) \leq \pi(u_{q_2}, n - 1) - \pi(u_{q_2}, n), \quad \text{if } q_2 < q_1.$$

Since the function is continuous and the cross-partial derivatives are strictly positive (in the interior), we can establish the following effect of a merger on price dispersion as measured by the interquartile range (IQR):

$$\text{sign}(\text{IQR}(n - 1) - \text{IQR}(n)) = \text{sign} \left( \frac{1}{2} \int_{u_{25}}^{u_{75}} \frac{\partial \pi(u_i, n - 1)}{\partial u} du_i - \frac{1}{2} \int_{u_{25}}^{u_{75}} \frac{\partial \pi(u_i, n)}{\partial u} du_i \right) \leq 0$$

The second equality comes from the fact that:

$$(E5) \quad \pi(u_{75}, n) = \pi(u_{25}, n) + \int_{u_{25}}^{u_{75}} \frac{\partial \pi(u_i, n)}{\partial u} du_i,$$

and its sign is negative since from equation (E3):

$$(E6) \quad \frac{\partial \pi(u_i, n)}{\partial u} \geq \frac{\partial \pi(u_i, n-1)}{\partial u}.$$

In this section we derive and estimate a model that incorporates asymmetric information, as discussed in Section 4.3 of the paper.

Consider a model with  $n+1$  lenders, in which negotiation takes place over three stages. In the first stage, consumers receive a take-it-or-leave-it offer  $m_i$  from a random lender. In the second stage, if the initial offer is rejected, consumers put forth a search effort  $e_i$  at cost  $\kappa_i$  to gather additional quotes. The number of quotes that consumers obtain is stochastically determined by their search effort: with probability  $s(e_i)$  they are randomly matched with  $n$  banks, and with probability  $1-s(e_i)$  they receive 2 quotes.<sup>3</sup> In the final stage, competition takes place, and consumers choose the lowest price offer. We characterize each stage sequentially, starting with the final competition game.

**Competition stage** We model the competition stage as an English auction between at most  $n$  lenders. Lenders are ex-ante identical with a common consumer-specific marginal cost  $c_i$ , but face a mean-zero additive idiosyncratic cost shock  $\epsilon_j$  that is privately observed after consumers are matched with lenders.

The game has a unique dominant-strategy equilibrium: banks are willing to offer up to their privately observed cost  $c_i + \epsilon_j$ , and the most efficient bank wins the contract by offering a rate equal to the second lowest cost,  $c_{(2)} = c_i + \epsilon_{(2)}$ , where  $\epsilon_{(2)}$  is the second order statistic of  $\{\epsilon_1, \dots, \epsilon_k\}$ . Let  $E(m^*|\tilde{n}) = c_i + E(\epsilon_{(2)}|\tilde{n})$  denote the expected second-stage transaction price, where  $\tilde{n}$  is the number of quotes generated (either two or  $n$ ). Given our matching assumption, the gain from searching is summarized by the expected cost difference between obtaining two or  $n$  quotes:  $\Delta(n) = E(\epsilon_{(2)}|2) - E(\epsilon_{(2)}|n) > 0$ . This function is increasing in  $n$ .

**Search effort stage** At this stage, consumers incur a search cost  $\kappa_i$  to gather additional quotes. The search cost has a fixed and variable component:

$$(F1) \quad \kappa_i = u_i(1 + e_i) + \eta_i.$$

We assume that the marginal cost of effort  $u_i$  is publicly observed by both parties, and that consumers privately observe the sunk cost  $\eta_i$  of gathering extra quotes. We assume that  $\eta_i$  is distributed according to an exponential distribution with a mean denoted by  $\bar{\eta}$ .

We use a simple matching technology to describe the number of quotes that consumers receive. Search effort  $e$  increases the probability  $s(e)$  of receiving the maximum number of quotes  $n$ . Consumers putting forth zero effort are automatically matched with two lenders, and we use a Pareto distribution to characterize the matching probability function:

$$(F2) \quad s(e) = 1 - (1 + e)^{-\lambda}.$$

Consumers choose an optimal level of search effort to minimize the sum of the expected transaction price and the search cost. Importantly, this effort level depends only on public information, since  $\eta_i$  is a fixed-cost. A consumer facing  $n$  possible options chooses an optimal effort level that minimizes the net borrowing cost:

$$(F3) \quad r(u_i, n) \equiv \min_{e \geq 0} u_i \cdot (1 + e) + c_i + E(\epsilon_{(2)}|2) - s(e)\Delta(n) = c_i + \pi(u_i, n).$$

The solution to this problem gives rise to effort and matching probability functions that are decreasing in  $u_i$  and increasing in  $n$ . Moreover, the optimal effort level exhibits a threshold property:

<sup>3</sup>We assume that the initial lender does not participate in the competition stage. We also assume that consumers get a minimum of two quotes in order to generate a finite price even for consumers with high search-costs.

consumers with marginal costs larger than a threshold  $\bar{u}(n)$  invest zero effort.

The search decision depends on the reserve value of consumers net of the private-valued search cost:  $r(u_i, n) - \eta_i$ . In particular, if  $m^0$  is the initial offer received in the first stage, the search probability is given by the following expression:

$$(F4) \quad H(m^0|u_i, n) = 1 - \exp\left(-\frac{1}{\bar{\eta}} \max(0, m^0 - r(u_i, n))\right).$$

**Initial offer** In the first stage, information is asymmetric. The initial lender therefore chooses  $m^0$  to maximize its expected profits:

$$(F5) \quad \max_{m^0} (m^0 - c_i)(1 - H(m^0|u_i, n)).$$

Since the private-value search cost  $\eta_i$  cannot be negative, the search probability is positive only if the initial quote is larger than consumers' reservation value,  $r(u_i, n)$ . In the interior, the optimal quote takes the form of a constant markup over  $c$ :

$$(F6) \quad m_i^0 = c_i + \frac{1 - H(m_i^0|u_i, n)}{h(m_i^0|u_i, n)} = c_i + \bar{\eta},$$

where  $h(m_i^0|u_i, n) = \partial H(m_i^0|u_i, n)/\partial m^0 = \frac{1}{\bar{\eta}}(1 - H)$ . This constant markup-rule is specific to the exponential search cost distribution. Under more general functional forms, the initial offer would depend on  $n$  and  $u_i$ , but the predictions of the model would be remain qualitatively unchanged.

If the interior solution is less than the reservation value, the lender can increase its initial offer without increasing the probability of losing the consumer. Therefore, the optimal initial offer is the maximum of  $c + \bar{\eta}$  and the reservation value of consumers:

$$(F7) \quad m_i^0 = \begin{cases} c_i + \bar{\eta} & \text{If } \bar{\eta} > \pi(u_i, n), \\ r(u_i, n) & \text{Else.} \end{cases}$$

where  $\pi(u_i, n)$  is the markup over  $c$  given by consumers' reservation values.

Similarly, the equilibrium search probability takes the following piece-wise form:

$$(F8) \quad H(u_i, n) = \begin{cases} 1 - \exp\left(-\frac{1}{\bar{\eta}}(\bar{\eta} - \pi(u_i, n))\right) & \text{If } \bar{\eta} > \pi(u_i, n), \\ 0 & \text{Else.} \end{cases}$$

Incorporating asymmetric information can therefore lead to search for a segment of the market. In particular, a lender will find it optimal to prevent search entirely for consumers with high observable search costs (i.e. high profit margins). In contrast, it will offer a relatively high price to consumers with good negotiation skills (i.e. low  $u_i$ ), at the risk of losing these consumers to competing lenders. This prediction is very similar to the results found by Chatterjee and Lee (1998), in which the value of the next best alternative (rather than the search cost) is privately known to the buyer.

**Functional form assumptions** To solve the model we impose the following distributional assumptions on the shocks entering the model:

- Sunk cost of search:  $\eta_i \sim \text{Exp}(1/\bar{\eta})$ ,
- Marginal cost of effort:  $u_i \sim \text{Exp}(1/\bar{u})$ ,

- Idiosyncratic cost shocks:  $\epsilon_{ij} \sim U[-\sigma_\epsilon, \sigma_\epsilon]$ ,
- Common lending cost residual:  $c_i \sim N(\bar{c}, \sigma_c^2)$ .

Because of the presence of multiple shocks, we cannot derive comparative statics to predict the effect of mergers along the distribution of rates. However, conditional on  $c_i$ , the model predicts that only consumers with low  $u_i$  will search in equilibrium. Consumers with effort costs above this level will pay a price that reflects their reservation value  $r(u_i, n)$ . For those consumers, the same comparative statics results derived for the model with complete information hold here.

For consumers receiving an initial quote equal to  $c_i + \bar{\eta}$ , the effect of the merger depends on their search decision. Since the gain from search is weakly increasing in  $n$ , the search probability is lower post-merger. This leads to three distinct merger effects on prices:

- 1) **Market-power effect:** Consumers who search before and after the merger pay higher prices due to an increase in market power dictated by cost differences across lenders (i.e.  $\sigma_\epsilon$ ). These consumers tend to pay relatively low rates before the merger.
- 2) **Search effect:** Consumers who change their search decision (i.e. from search to non-search) experience a price increase proportional to the difference between the average search cost ( $\bar{\eta}$ ) and the pre-merger average markup.
- 3) **Price-discrimination effect:** Consumers who accept the initial quote before and after the merger end up paying the same rate, and are unaffected by the merger.

These three effects produce a highly non-linear merger effect along the distribution of consumer types  $u_i$ . However, the presence of the cost-shock  $c_i$  smooths the distributional impact of the merger. Overall, the effect of the merger tends to have an inverted-u shape. It is highest around the middle of the rate distribution where the search effect is more important, lower at the bottom of the distribution where consumers are mostly affected by the increase in market power, and close to zero at the top of the distribution for consumers paying their reservation value. Depending on the parameter values, the model can therefore predict an increase or decrease in the dispersion of prices.

**Estimation and simulation results** Unlike with the complete information model, the search effort distribution cannot be inverted from the price distribution. Instead the distribution of prices is a mixture of four distributions, which prevents us from using the same nested fixed-point NLS estimator.

We estimate the model parameters using simulated method of moments. More specifically, we take 1,000 draws from the joint distribution of shocks, and evaluate the distribution of rates before and after the merger. Using these two distributions, we compute 10 moments: (i) average and standard-deviation of rates pre-merger, (ii) average effect of the merger, (iii) change in the standard-deviation of rates, (iv) percentile changes (i.e. 10%, 25%, 50%, 75% and 90%), and (v) the fraction of consumers searching more than one lender. We repeat this operation 100 times in order to compute the predicted average change in these moments.

We match the model predictions with the estimated effects of the merger obtained using the CiC estimator (i.e. Tables 3 and 4 in the paper). In addition, we use the average search probability from a national survey conducted by Ipsos-Reid. Overall, excluding consumers living in the most urban areas of the countries (i.e. more than 1M population), this survey estimates that 54.6% of new borrowers gather more than one quote.<sup>4</sup>

<sup>4</sup>Given that in the reduced-form we focus on medium-sized markets, i.e. those with 4-8 lenders, excluding the more populated and thus higher search probability markets makes sense. —The search probability in markets with over 1 million people is 64%.



The estimated parameters are presented in the second column of Table F1. The parameter estimates are not directly comparable to the complete information specification due to introduction of a new cost shock  $c_i$ , and the presence of an additional fixed search cost. Overall, the asymmetric information model predicts that firms competing at the auction stage have more homogenous costs, which lead to a lower market-power effect of the merger (i.e. 6.3 bps instead of 11 bps).

The average marginal cost of effort is also significantly lower (i.e. 12.6 bps instead of 35 bps). However, because of the additional fixed cost, on average, consumers face significantly higher search cost in this model. Indeed, the average sunk cost  $\bar{\eta}$  corresponds to 80 bps (median = 56 bps). Somewhat counter-intuitively, this large fixed-cost is needed to rationalize the high search probability. With a lower value for  $\bar{\eta}$ , fewer consumers would receive an initial quote equal  $c_i + \bar{\eta}$ , which would raise the fraction of non-searchers.

The last two columns compare the fit of the model with the moments used for estimation. Overall, the model accurately reproduces the magnitude of the reduced-form results obtained with the CiC estimator. This is particularly true for the quantile treatment effects, and the aggregate pre-merger moments. The model under-estimates the average effect by 1 bps; mostly because it predicts a smaller effect at the bottom of the price distribution.

Finally, the asymmetric and full information models have similar implications with respect to the relationship between market-power increase and the average treatment effect of the merger. Table F1 shows that the average effect of the merger predicted by the model underestimates the actual increase in market power by 24%. In contrast, the complete information model shows that the ATE is almost 50% lower than the pure market-power increase. While the magnitudes differ, both models clearly suggest that the ATE under-estimates the actual increase in market power caused by the merger.

Table F1—: Estimation and simulation results

	Parameters estimates	Moments	CiC est.	Model sim.
$\mu_c$	1.397	Search prob	0.546	0.552
$\sigma_c$	0.051	Avg. margin (Before)	1.720	1.721
$\sigma_\epsilon$	0.461	Std-dev margin (Before)	0.470	0.473
$\bar{u}$	0.126	ATE	0.057	0.048
$\lambda$	1.976	$\Delta$ Std-dev	-0.029	-0.025
$\bar{\eta}$	0.810	$\Delta$ P-10	0.091	0.053
		$\Delta$ P-25	0.076	0.081
$E(p^* n-1) - E(p^* n)$	0.063	$\Delta$ Median	0.084	0.090
Simulations	1000	$\Delta$ P-75	-0.001	0.003
GMM	0.061	$\Delta$ P-90	-0.004	0.002

REFERENCES

- Abadie, Alberto.** 2005. “Semiparametric Difference-in-Differences Estimators.” *Review of Economic Studies*, 72: 1–19.
- Athey, Susan, and Guido W. Imbens.** 2006. “Identification and Inference in Nonlinear Difference-in-Difference Models.” *Econometrica*, 74: 431–497.
- Chatterjee, Kalyan, and Ching Chyi Lee.** 1998. “Bargaining and Search with Incomplete Information about Outside Options.” *Games and Economic Behavior*, 22: 203–237.
- Competition Bureau.** 2003. “The Merger Enforcement Guidelines as Applied to a Bank Merger.” Technical report.
- Heckman, James J, Hidehiko Ichimura, and Petra Todd.** 1997. “Matching as an econometric evaluations estimator: Evidence from evaluating a job training programme.” *Review of Economic Studies*, 64: 605–654.
- Linton, Oliver, Esfandiar Massoumi, and Yoon-Jae Whang.** 2005. “Consistent Testing for Stochastic Dominance Under General Sampling Schemes.” *Review of Economic Studies*, 72: 735–765.