Online Appendix

Do information sharing platforms facilitate coordination?
Evidence from the collapse of a cartel

Robert Clark
University of Toronto

Jean-François Houde
UW-Madison & NBER

October 24, 2026

This appendix accompanies the paper of the same title. Appendix A, the proofs of Lemma 1 and Propositions 1 and 2, and Appendix B, the tables discussed in the text, are in the paper; the sections below continue from C. Section, equation, proposition, table and figure numbers without a letter refer to the main text.

Contents

C The length of the war and the ordering of the cutoffs **2**

D Accidental missing reports **3**

E The non-collusive value of the reports **6**

F The effect of the investigation on margins and price adjustments **8**

G Additional tables and figures **11**

H Construction of the information-sharing variables **23**

H.1 Survey responses 23

H.2 Construction of $Share_{it}$ 24

H.3 Construction of $StopShare_i$ 25

C The length of the war and the ordering of the cutoffs

This section complements the proofs in Appendix A of the paper. It derives the gain from a secret delay, records a necessary condition on the length of the war, explains why the bound is long in the model and short in the industry, and shows that the optimal trigger in regime (A) sets a lower cutoff after a complaint. Notation is that of Appendix A.

The gain from a secret delay. During an increase the firms that have raised post the collusive margin, normalized to one, and the firm that has not posts $1-\Delta_p$ and sells the whole flow. Letting L hold the low posting for a fraction λ of the quarter's volume gives L the share $\lambda + s(1-\lambda)$, which equals $s+\tau$ at $\lambda = \tau/(1-s)$, and each follower $s(1-\lambda) = s_f$, so the transfer is funded equally by the followers. A complying follower earns $\pi^{\text{col}} = \bar{m}\, s_f(1-c)$ per period and nothing during a war. A follower that delays takes a share d of L's volume at the price $1-\Delta_p$ and earns $\bar{m}[s_f + d(1-\Delta_p)] - \bar{m}(s_f+d)c = \pi^{\text{col}} + \bar{m}\, d\,(1-\Delta_p - c)$, which is equation (2) of the paper.

The comparison has content only if collusion is sustainable at some δ. Letting $\delta \to 1$ in (IC-P), with $(1-\delta)/[\delta(1-\delta^W)] \to 1/W$, gives a necessary condition in every regime:

$$W\,(1-\alpha_0)\,s_f \;\geq\; (1-s_f)(1+W\alpha_0). \tag{C.1}$$

At α_0 near zero, the war must last at least $(1-s_f)/s_f$ periods, a little more than $n-1$.

The length of this bound reflects two simplifications of the model rather than a feature of the industry. First, a public deviation is assumed to capture the market for a whole period, and the period is the reporting interval, a quarter. Posted prices are seen at once, and a cut is matched within days, so the gain from a public cut is at most a week of volume. In a model with a weekly period and a quarterly reporting interval, the version described in the footnote to (2), the war that deters a public cut lasts a little more than $n-1$ weeks, not quarters. Second, the market is homogeneous, so a cut by one station draws the customers of every other station in the city. In practice a station's cut draws customers from the stations near it, and the relevant n is the number of those stations, a handful, rather than the number of stations in the city. Under both corrections a war of a few weeks deters a public cut, which is the length of the wars observed before the investigation (Section 3.1). The length of the war enters the constraint against the secret delay only through $\rho(\delta)$ and α_0, and the delay is the deviation the trigger is designed to catch.

Lemma C.1 (State-contingent cutoffs)**.** *In regime (A), suppose the optimal rule has κ_C and κ_N finite and (IC-S) binds at a single $d^* \in (0, \bar{d}]$. Then $\kappa_C < \kappa_N$.*

Proof. Write $z \equiv z(d^*) > 0$. Lowering κ_C raises α_0 at the rate $\phi_0\varphi(\kappa_C)$ and the war probability after d^* at the rate $\phi(d^*)\varphi(\kappa_C - z)$. The corresponding rates for κ_N are $(1-\phi_0)\varphi(\kappa_N)$ and $(1-\phi(d^*))\varphi(\kappa_N - z)$. At an interior optimum the two ratios of benefit to cost are equal (α_0 also enters the right sides of (IC-S) and (IC-P), but those terms are proportional to the derivative of

α_0 and leave the ratio unchanged). Using $\varphi(\kappa - z)/\varphi(\kappa) = \exp(\kappa z - z^2/2)$, this gives

$$\kappa_N - \kappa_C = \frac{1}{z} \log \left[\frac{\phi(d^*)}{\phi_0} \cdot \frac{1 - \phi_0}{1 - \phi(d^*)} \right] > 0, \tag{C.2}$$

since $\phi(d^*) > \phi_0$. If (IC-S) binds only as $d \to 0$, the same argument with $z \to 0$ gives $\kappa_N - \kappa_C = \phi'(0)/[\phi_0(1 - \phi_0)z'(0)] > 0$. □

D Accidental missing reports

In the model of Section 3.2 a member that shares supplies its volume with certainty, so a follower's volume is missing from the report only if the follower withheld it, and the trigger is restricted not to respond to the number of missing volumes. The first assumption is a simplification: in the competitive cities of the province, where no station has a collusive reason to hide its sales, about 2 to 3 percent of volume readings are missing for reasons outside the operator's control. The second is a restriction on the cartel's strategy, and in the model it costs the cartel nothing, since no volume is ever missing on the equilibrium path. This section illustrates what relaxing both assumptions does to a deviator's decision to report or withhold its volume. We close the section with the trade-off the cartel faces when choosing its trigger. Characterizing the optimal trigger with accidental missing reports is beyond the scope of this paper. Notation is otherwise that of Section 3.2 and Appendix A of the paper; symbols introduced here are local to this section, and $\bar{n} \equiv n - 1$ denotes the number of symmetric followers.

Accidents and the evidence in the count. Each follower's reading fails with probability ε in each period, independently across followers and periods and independently of conduct, of demand, and of the message. The Platform prints the identity of the followers whose volume is missing but not the reason. We approximate the number of accidents in a period by a Poisson variable with mean $\nu \equiv \bar{n}\varepsilon$, truncated at $\bar{n} - 2$ so that at least one follower volume is always published:

$$\pi_h = \frac{\nu^h/h!}{\sum_{k=0}^{\bar{n}-2} \nu^k/k!}, \qquad h \in \{0, 1, \ldots, \bar{n} - 2\}. \tag{D.1}$$

With $\varepsilon \approx 0.03$ and the 2002 station counts, ν is about 0.4 in Thetford Mines, 0.8 in Victoriaville, and 2.2 in Sherbrooke, and the probability of a quarter with no missing volume, $\pi_0 \approx e^{-\nu}$, is 0.67, 0.45, and 0.11. Let ℓ be the number of follower volumes missing from a report. On the equilibrium path ℓ is the number of accidents. If a follower withholds, ℓ is one plus the number of accidents among the others, which we also take to follow (D.1), an approximation that ignores the one fewer follower at risk. For every $\ell \geq 1$ the likelihood ratio of observing ℓ missing volumes after a withholding, relative to observing it on the path, is

$$\frac{\pi_{\ell-1}}{\pi_\ell} = \frac{\ell}{\nu}. \tag{D.2}$$

A count of missing volumes is evidence of a withholding only when it exceeds the number that accidents usually produce. In Sherbrooke, a single missing volume is more likely on the equilibrium path than after a withholding.

A trigger that responds to the count. The cartel observes ℓ and may use it. The critical value in (4) becomes $\kappa_{M\ell}$, one for each message M and each count ℓ, with the convention $\kappa_{M,\bar{n}-1} = \kappa_{M,\bar{n}-2}$ for the state that arises only when a deviator withholds in a period with $\bar{n}-2$ accidents.[1] We measure how much one more missing volume tightens the trigger by

$$\Gamma_\ell \equiv \max_M \big(\kappa_{M\ell} - \kappa_{M,\ell+1}\big), \qquad \ell \in \{0, \dots, \bar{n}-2\}, \tag{D.3}$$

with the convention that the difference of two infinite critical values of the same sign is zero, so that the trigger of Section 3.2 has $\Gamma_\ell = 0$ for every ℓ. Write $\alpha_{0\ell} \equiv \phi_0 \Phi(-\kappa_{C\ell}) + (1-\phi_0)\Phi(-\kappa_{N\ell})$ for the probability of a war in state ℓ when nobody deviates. The on-path frequency of price wars is then

$$\alpha_0 = \sum_{h=0}^{\bar{n}-2} \pi_h \, \alpha_{0h}. \tag{D.4}$$

A trigger that tightens with the count raises α_0, because accidents put the cartel in the tighter states on the equilibrium path. Unlike in Section 3.2, treating a missing volume as evidence is costly.

The deviator's two options. A delay of size d lowers L's log volume by $a(d) \equiv -\log\big(1 - d/(s+\tau)\big)$ and raises the deviator's by $b(d) \equiv \log\big(1 + d/s_f\big)$, where $s_f \equiv s - \tau/\bar{n}$ is a follower's share. Both are continuous and strictly increasing with $a(0) = b(0) = 0$. Write $\sigma_r \equiv \sigma_u\sqrt{1+1/r}$ for the standard deviation of Y when r follower volumes are published. Consider a period in which h of the other followers are missing by accident.

- *Report.* The count is $\ell = h$ and $r = \bar{n} - h$ volumes are published, including the deviator's.[2] Its inflated volume enters the average in (3) with weight $1/r$, so the standardized fall of the statistic is $z_R(d,h) \equiv \big[a(d) + b(d)/(\bar{n}-h)\big]/\sigma_{\bar{n}-h}$.
- *Withhold.* The count is $\ell = h+1$ and $r = \bar{n} - 1 - h$ volumes are published. Only L's lower volume moves the statistic, and the standardized fall is $z_W(d,h) \equiv a(d)/\sigma_{\bar{n}-1-h}$.

Withholding therefore does two things: it moves the cartel one state up, where the trigger may be tighter, and it removes the deviator's own volume from the statistic. The evidence that the

[1] Left unrestricted, the trigger could punish that state at no cost, which is the objection this section addresses, reintroduced by the truncation. The state has probability of order $\nu^{\bar{n}-2}/(\bar{n}-2)!$ and is negligible in the markets of interest.

[2] We ignore the deviator's own accident, which occurs with probability ε and makes a report equivalent to a withholding; it multiplies $\alpha_R(d) - \alpha_W(d)$ below by $1-\varepsilon$ and changes no sign.

deviator's own volume would add is

$$g_h(d) \equiv z_R(d,h) - z_W(d,h) = a(d)\Big(\frac{1}{\sigma_{\bar{n}-h}} - \frac{1}{\sigma_{\bar{n}-1-h}}\Big) + \frac{b(d)}{(\bar{n}-h)\,\sigma_{\bar{n}-h}}. \tag{D.5}$$

The first term is the precision effect: a statistic computed on $\bar{n}-h$ volumes is less noisy than one computed on $\bar{n}-1-h$. The second is the direct effect: the deviator's inflated volume pushes the average up. These are the two effects behind Lemma 1. Since $1/\sigma_r$ is increasing and concave in r and $1/(r\sigma_r)$ is decreasing, (a) for each h, g_h is continuous and strictly increasing in d with $g_h(0) = 0$, and (b) for each $d > 0$, $g_h(d)$ is strictly increasing in h: the deviator's own volume weighs more, and its absence costs more precision, when fewer other volumes are published.

For a count ℓ, a delay d, and a standardized fall z, the probability of a war is $B_\ell(z;d) \equiv \phi(d)\Phi(-\kappa_{C\ell}+z) + \big(1-\phi(d)\big)\Phi(-\kappa_{N\ell}+z)$, which is strictly increasing in z when at least one critical value in state ℓ is finite. Averaging over accidents, the probability of a war after a delay of size d is $\alpha_R(d) = \sum_h \pi_h B_h(z_R(d,h);d)$ if the deviator reports and $\alpha_W(d) = \sum_h \pi_h B_{h+1}(z_W(d,h);d)$ if it withholds. The gain from the delay is the same under both options, so the deviator picks the option with the lower probability of a war.

Lemma D.1 (The reporting decision of a deviator)**.** *Let accidents follow* (D.1)*, and let the trigger have critical values $\kappa_{M\ell}$ with at least one finite critical value in every state and $\kappa_{M,\bar{n}-1} = \kappa_{M,\bar{n}-2}$.*

(i) If $g_h(d) \geq \Gamma_h$ for every $h \in \{0,\dots,\bar{n}-2\}$, a follower that takes a secret delay of size d strictly prefers to withhold its volume. In particular, if the trigger does not tighten with the count ($\Gamma_h \leq 0$ for every h), it strictly prefers to withhold for every $d \in (0,\bar{d}]$.

(ii) If the trigger tightens with the count on average, $\sum_h \pi_h(\alpha_{0,h+1} - \alpha_{0h}) > 0$, there exists $\underline{d} > 0$ such that the follower strictly prefers to report for every $d \in (0,\underline{d})$.

(iii) For each h with $0 < \Gamma_h \leq g_h(\bar{d})$ let $\hat{d}_h \equiv g_h^{-1}(\Gamma_h)$, and set $\hat{d}_h = 0$ if $\Gamma_h \leq 0$. If $\hat{d} \equiv \max_h \hat{d}_h$ is well defined, the follower strictly prefers to withhold for every $d \in [\hat{d},\bar{d}]$ with $d > 0$. Each $\hat{d}_h$ is strictly increasing in the penalty Γ_h on $(0, g_h(\bar{d})]$ and strictly decreasing in the number of accidents h at a given penalty.

Proof. Write $\alpha_R(d) - \alpha_W(d) = \sum_h \pi_h T_h(d)$ with $T_h(d) \equiv B_h(z_R(d,h);d) - B_{h+1}(z_W(d,h);d)$. The follower strictly prefers to withhold if the sum is strictly positive and to report if it is strictly negative. Expanding B,

$$T_h(d) = \phi(d)\Big[\Phi\big(-\kappa_{Ch}+z_R\big) - \Phi\big(-\kappa_{C,h+1}+z_W\big)\Big] + \big(1-\phi(d)\big)\Big[\Phi\big(-\kappa_{Nh}+z_R\big) - \Phi\big(-\kappa_{N,h+1}+z_W\big)\Big],$$

with z_R and z_W evaluated at (d,h). For each message the bracket is non-negative if and only if $z_R - z_W = g_h(d) \geq \kappa_{Mh} - \kappa_{M,h+1}$, and strictly positive if the inequality is strict and at least one of the two critical values is finite.

(i) If $g_h(d) \geq \Gamma_h$, both brackets are non-negative, so $T_h(d) \geq 0$ for every h. At $h = \bar{n}-2$ the convention gives $\Gamma_{\bar{n}-2} = 0 < g_{\bar{n}-2}(d)$, the critical values of the two states coincide and at least one

is finite, and $\phi(d) \in (0,1)$, so $T_{\bar{n}-2}(d) > 0$. Since $\pi_{\bar{n}-2} > 0$, the sum is strictly positive. If $\Gamma_h \leq 0$ for every h, then $g_h(d) > 0 \geq \Gamma_h$ for every h and every $d > 0$.

(ii) At $d = 0$, $z_R = z_W = 0$ and $\phi(0) = \phi_0$, so the sum equals $\sum_h \pi_h(\alpha_{0h} - \alpha_{0,h+1})$, which is strictly negative by hypothesis. The sum is continuous in d, so it is strictly negative on some interval $(0, \underline{d})$.

(iii) By property (a), $\hat{d}_h$ is well defined when $0 < \Gamma_h \leq g_h(\bar{d})$ and $g_h(d) \geq \Gamma_h$ for every $d \geq \hat{d}_h$. At $d \geq \hat{d}$ the condition of (i) holds for every h. Monotonicity in Γ_h follows from (a); monotonicity in h follows from (b), since a larger h shifts g_h up at every $d > 0$, so it reaches a given level at a smaller d. □

The lemma says that a missing volume is a penalty that does not depend on the size of the delay, while a report is evidence that grows with it. A small delay is better hidden inside the report, because the deviator's own volume moves the average by little. A large delay is better hidden behind a missing volume, because the deviator's own volume would be conspicuous. The threshold falls with the number of accidents, because the deviator's own volume is more conspicuous when fewer other volumes are published. Lemma 1 is the special case in which the trigger does not respond to the count, $\Gamma_h = 0$ for every h, so that part (i) applies at every d. The difference is that in Section 3.2 that restriction was free, and here it is not: by (D.4), a trigger that penalizes missing volumes starts more wars on the equilibrium path.

Does the cartel want to respond to the count? Lemma D.1 holds for every trigger in the class, whether or not it is the one the cartel would choose. Which trigger it chooses is a separate question, and the answer is not obvious. Responding to the count buys deterrence against withholding, but every accident then raises the probability of a war that nobody provoked. In the markets of interest at least one volume is missing by accident in a third to nine tenths of the quarters, so a trigger that treats a missing volume as evidence is often in its tighter states. The added wars lower the value of the arrangement for every member and tighten the constraint against the public price deviation, and if they lower it enough the arrangement is no longer sustainable. A cartel that minimizes the frequency of wars subject to deterring both deviations therefore trades the deterrence it gains against withholding for wars after accidents, and the penalties it chooses may be small. We do not pursue the characterization here. What the paper uses is the weaker statement: under any trigger in the class, the delays that are bundled with withholding are the large ones, and under the restriction of Section 3.2 all of them are.

E The non-collusive value of the reports

This appendix derives the claim in Section 3.3 that the non-collusive value of the reports is positive and rises with the instability of shares. Symbols are local to this appendix. In particular, k is a cost constant and α is the frequency of periods in which the arrangement does not govern the allocation.

Setup. There are n stations in the market, and volume is (1). Write x_t for $\log m_t$ in deviation from its mean, and let

$$x_t = \rho\, x_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2), \qquad \rho \in (0,1), \tag{E.1}$$

with stationary variance $\sigma_x^2 = \sigma_\eta^2/(1-\rho^2)$. Shares are $\log s_{jt} = \log \bar{s}_j + \zeta_{jt}$, where $\bar{s}_j$ is the share under the prevailing conduct and ζ_{jt}, with variance σ_ζ^2, is reallocation between stations of the same market; the adding-up constraint $\sum_j s_{jt} = 1$ becomes $\sum_j \zeta_{jt} = 0$ in log-linear form (Assumption E.1).

Assumption E.1. $\sum_j \zeta_{jt} = 0$, and $(\zeta_{1t}, \ldots, \zeta_{nt})$ is exchangeable, independent across periods, and independent of m_t and of the u_{jt}.

Assumption E.2. A station orders next period's fuel before next period's prices are set, using the information available at the time of the order.

Under Assumption E.2 the ordering cost enters profit additively and does not affect the choice of price, so the derivation below is valid under any conduct, and conduct enters the value of the reports only through σ_ζ^2.

The cost of ordering. Write $\mathcal{I}_{jt}$ for the station's information and $v \equiv \mathrm{Var}(\log q_{j,t+1} \mid \mathcal{I}_{jt})$. The station orders a litres and incurs the cost

$$k\, \mathbb{E}\big[(\log a - \log q_{j,t+1})^2 \mid \mathcal{I}_{jt}\big], \qquad k > 0. \tag{E.2}$$

Lemma E.1. *The optimal order is* $\log a^* = \mathbb{E}[\log q_{j,t+1} \mid \mathcal{I}_{jt}]$*, and the minimized cost is* $k\, v$*.*

Proof. Adding and subtracting $\mathbb{E}[\log q_{j,t+1} \mid \mathcal{I}_{jt}]$ in (E.2) gives $k[(\log a - \mathbb{E}[\log q_{j,t+1} \mid \mathcal{I}_{jt}])^2 + v]$. □

Forecast variance. Next period's log volume is $\log \bar{s}_j + \mathbb{E}\log m_t + x_{t+1} + \zeta_{j,t+1} + \log u_{j,t+1}$. Only x_{t+1} is forecastable, so $v = Q + \sigma_\zeta^2 + \sigma_u^2$ with $Q \equiv \mathrm{Var}(x_{t+1} \mid \mathcal{I}_{jt})$. A non-subscriber observes its own volume, which, taking ζ_{jt} Gaussian or restricting attention to linear forecasts, is a signal of x_t with noise $\zeta_{jt} + \log u_{jt}$ (up to a constant) and noise variance $\sigma_\zeta^2 + \sigma_u^2$. A subscriber observes every station's volume. Averaging the n signals removes the ζ_{jt} exactly by Assumption E.1, leaving noise variance σ_u^2/n.[3] With the stationary distribution of x_t as prior and a signal of noise variance σ^2,

$$Q(\sigma^2) = \rho^2\, \frac{\sigma_x^2\, \sigma^2}{\sigma_x^2 + \sigma^2} + \sigma_\eta^2, \qquad Q'(\sigma^2) = \rho^2\, \frac{\sigma_x^4}{(\sigma_x^2 + \sigma^2)^2} > 0. \tag{E.3}$$

Using the whole history of signals in place of one replaces Q by the steady-state filtering variance, which is also increasing in σ^2.

[3] If only $r < n$ stations report, with $r \geq 2$ (r counts stations here, not followers), the subscriber's noise variance is $\sigma_u^2/r + \sigma_\zeta^2(n-r)/(r(n-1))$. Then $G > 0$ still holds, and G is increasing in σ_ζ^2 if $\frac{n-r}{r(n-1)} Q'(\sigma_S^2) < Q'(\sigma_u^2 + \sigma_\zeta^2)$, where σ_S^2 is the subscriber's noise variance. The condition holds when coverage is close to full.

The value of subscribing. By Lemma E.1, the saving from subscribing is

$$G = k\Big[Q\big(\sigma_u^2 + \sigma_\zeta^2\big) - Q\big(\sigma_u^2/n\big)\Big]. \tag{E.4}$$

Proposition E.1 (The reports are worth more when shares move)**.** *$G > 0$, and G is strictly increasing in σ_ζ^2. Let the reallocation variance be $\sigma_{\zeta,c}^2$ in periods in which the arrangement governs the allocation and $\sigma_{\zeta,w}^2 > \sigma_{\zeta,c}^2$ in periods in which it does not, and let α be the frequency of the latter. Then the expected saving $(1-\alpha)G(\sigma_{\zeta,c}^2) + \alpha\, G(\sigma_{\zeta,w}^2)$ is strictly increasing in α.*

Proof. Since $\sigma_u^2 + \sigma_\zeta^2 > \sigma_u^2/n$ and $Q' > 0$, $G > 0$. The second term of (E.4) does not involve σ_ζ^2 and the first is strictly increasing in it. The expected saving is a convex combination of $G(\sigma_{\zeta,c}^2) < G(\sigma_{\zeta,w}^2)$ with weight α on the larger value. □

Evidence. Table B.2 in Appendix B of the paper compares the quarters before and after the warrants in the three prosecuted markets observed over the full sample period. Panels A and B report the dispersion of station volumes within a market and quarter, net of station and market-by-quarter fixed effects. This is the counterpart of σ_ζ^2. It rises in every market, by 0.034 to 0.079 log points on the standard deviation and by 0.033 to 0.117 on the interquartile range. Panel C reports the within-quarter standard deviation of the weekly margin, which rises by 1.1 to 1.8 cents per litre. The survey covers only the prosecuted markets, so Panels A and B have no control group.

F The effect of the investigation on margins and price adjustments

This appendix reports the city-level replication of Clark and Houde (2014) summarized in Section 2.2.

To investigate the impact of the collapse of the cartels, we estimate the following difference-in-differences specification:

$$y_{ct} = \alpha_c + \eta_t + \beta_1 \left(Cartel_c \times Post_t\right) + \beta_2 \left(OtherCollusive_c \times Post_t\right) + \varepsilon_{ct}, \tag{F.1}$$

where y_{ct} is the outcome variable (margin, margin dispersion, or price adjustment ratio) for city c in period t, α_c is a city fixed effect, η_t is a period fixed effect, $Cartel_c$ is an indicator for the four prosecuted cities, and $OtherCollusive_c$ is an indicator for the remaining cities in the collusive cluster. $Post_t$ equals one for the post-investigation period (May 2006 to March 2007). The estimation sample consists of two periods per city (pre-investigation and post-investigation), yielding 106 observations. Standard errors are clustered at the city level.

Table F.1 reports the results. Column (1) shows that the investigation caused a significant decline in median retail margins. In the cartel cities, margins fell by 2.03 cents per liter ($p < 0.05$) relative to competitive control markets, which is comparable to the 2.27 cent decline estimated at the station level in Clark and Houde (2014). The other potentially collusive cities also experienced a significant margin decline of 0.78 cents per liter ($p < 0.10$), smaller than in the cartel cities but

Table F.1: Effect of the Investigation on Margins and Price Adjustments

	(1) Margins (c/l)	(2) Margin dispersion	(3) Adjustment ratio
Cartel cities × Post	-2.030**	2.793***	-1.276**
	(0.937)	(0.911)	(0.525)
Other collusive × Post	-0.784*	0.708	-1.151***
	(0.410)	(0.538)	(0.355)
Observations	106	106	106
R^2	0.940	0.900	0.820
Period FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes

Notes: Each column estimates equation (F.1) with a different dependent variable. Margins are the median posted price minus the Régie price floor in cents per liter. Margin dispersion is the interdecile range (p90−p10) of weekly margins within a city-period, capturing the degree of price instability over time. The adjustment ratio is the ratio of median price increases to median price decreases, normalized by the corresponding ratio for the price floor. The estimation sample covers two periods: pre-investigation (January 2002 to May 2006) and post-investigation (May 2006 to March 2007). Robust standard errors clustered at the city level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

indicating that the disruption extended beyond the directly prosecuted markets. This could be because some of the potentially collusive cities featured elevated levels of instability prior to the collapse, or simply because they did not respond to the collapse, or even because our classification was imperfect and some of the cities are not in fact collusive.

Column (2) examines margin instability within each city, measured as the interdecile range (p90−p10) of weekly margins over the period. Under the cartel, margins in the prosecuted cities were stable from week to week. After the investigation the interdecile range rises by 2.79 cents per litre ($p < 0.01$), as margins swing between near zero during price wars and temporary spikes during failed attempts to restore coordination. The descriptive statistics show that the higher dispersion persists through 2007–2012, so the cartel cities did not settle into a new stable equilibrium. The other collusive-cluster cities also show an increase in margin dispersion (0.71), although the effect is not statistically significant ($p > 0.10$), suggesting that the transition to instability was most pronounced in the directly prosecuted markets.

Column (3) estimates the effect on the price adjustment asymmetry ratio. Cartel cities show a decline of 1.28 ($p < 0.05$), indicating a shift toward more symmetric price adjustments as the coordinated Edgeworth cycle pattern breaks down. The other potentially collusive cities exhibit a decline of similar magnitude (1.15, $p < 0.01$), consistent with the same breakdown of asymmetric pass-through in those markets. In Clark and Houde (2014), the corresponding station-level estimates were -1.63 for cartel cities and -0.97 for cyclical markets. The broadly similar magnitudes confirm that the city-level aggregate data replicate the main findings from the earlier station-level analysis.

These results show that the investigation disrupted collusive pricing in the cartel cities, with smaller effects in the other markets that priced like them before 2006. This disruption is the

necessary precondition for the information-sharing results of Sections 4 and 5. If the investigation did not alter competitive conduct, there would be no reason to expect changes in stations' sharing behavior.

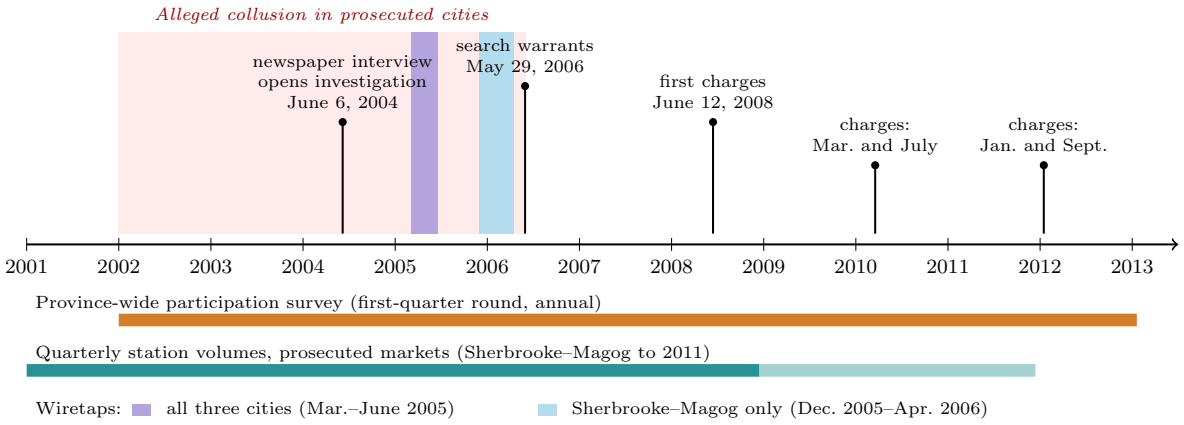


Figure G.1: Timeline of the cartel, the investigation, and the data coverage

G Additional tables and figures

Timeline and pre-investigation participation. Figure G.1 plots the timeline of the case and the coverage of the data sources described in Section 2 of the paper. Table B.1 in Appendix B of the paper describes participation in 2006, the year before the investigation, for the 1,653 stations observed across 29 markets. Column (1) reports the composition of the sample: branded stations affiliated with the major refiners are 44 percent, the low-price guarantee chain 24 percent, unbranded distributors 16 percent, retail chains 7 percent, and independent and unclassified stations, the omitted category in the regressions, 9 percent. Only 5 percent of the 2006 cross-section entered after 2002, and 23 percent exited before 2013. Column (2) is the linear probability model of $Share_{it}$ on station-type indicators and market fixed effects discussed in Section 2.4 of the paper. Column (3) adds the owner of the station's site, measured by the holder of the permit for its underground storage tanks in the registry of the Régie du bâtiment du Québec.[4] The omitted category is a site owned by a dealer, a co-operative, or a retail chain. The coefficients of the low-price guarantee chain and of the other branded stations fall from 0.34 to 0.28 when ownership is added, since most of these sites are owned by a refiner. The other station-type coefficients change little.

[4]We match each station in the Platform's survey to its site in the Régie's permit files for 2003 and 2013 by street address, and confirm the location with postal codes, municipality names before and after the 2002 municipal mergers, and geocoded distances. When a street was renamed or renumbered after the mergers, the Régie's file number, which identifies a site across years, links the two files. We use the 2003 holder, which dates ownership to the cartel period, and the 2013 holder for stations with no match in 2003. Of the 1,653 stations in the estimation sample, 1,578 are matched in 2003, 47 only in 2013, and 28 in neither year. The last group has its own indicator. We classify the holder by name as a refiner (Ultramar, Imperial Oil, Petro-Canada, Shell, or Irving), a regional distributor (for instance Pétroles Therrien or Philippe Gosselin & associés), a retail chain, a co-operative, or a dealer. The permit holder is a refiner for 47 percent of the sample and a regional distributor for 8 percent. The variable measures who owns the site and its tanks. It does not measure who sets the price. The owner can lease the station to a dealer who operates it and sets its retail price, so a refiner-owned site can be run by a lessee dealer.

Table G.1: Pre-trend test and HonestDiD sensitivity analysis: cartel cities

	(1) Cartel cities
Treatment effects:	
DiD estimate (baseline)	-0.079***
	(0.016)
Pre-trend diagnostics:	
Pre-investigation trend coef.	-0.0042
	(0.0064)
Joint pre-trend test (p-value)	0.934
Breakdown values:	
Short-run (2007)	0.20
Medium-run (2008–2010)	0.75
Long-run (2011–2013)	0.00†

Notes: Standard errors clustered at the market level in parentheses. Significance: * p<0.10, ** p<0.05, *** p<0.01. †: breakdown reported even though standard 95% CI marginally contains zero (upper bound within 0.01 of zero).

Pre-investigation trend coef.: estimate of γ from $Share_{it} = \alpha_i + \delta_t + \gamma\,(\text{year}_t - 2002) \cdot D_i + \varepsilon_{it}$, fit on pre-period observations (year $\leq$ 2006). α_i: station fixed effects; δ_t: year fixed effects; D_i: treatment indicator.

Joint pre-trend test: test that the pre-2006 event-study coefficients are zero. The event study is estimated relative to 2006 with market-clustered standard errors, as the relative-magnitudes method requires; the event-study figures in the text use 2002 as the reference year and cluster by station.

Breakdown M: smallest value of $\bar{M}$ under the relative-magnitudes restriction of Rambachan and Roth (2023) for which the 95% robust confidence interval for the post-period effect first contains zero. Larger values indicate greater robustness to post-treatment violations of parallel trends.

Table G.2: Participation rates by potentially collusive market and year

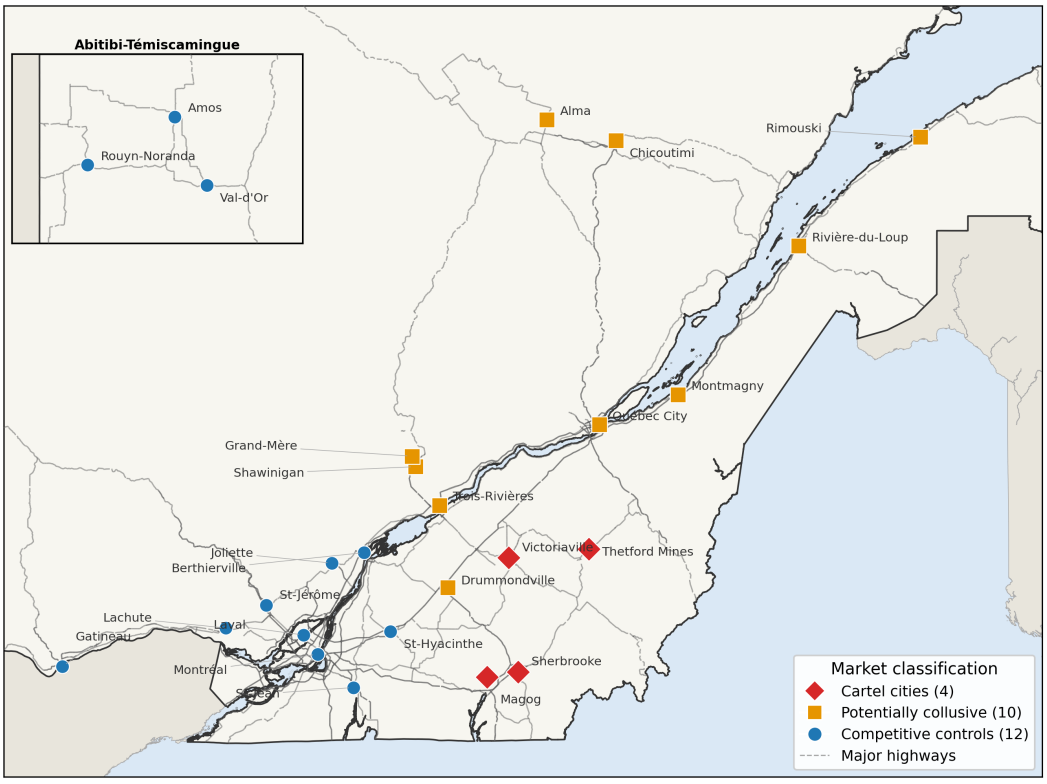
Market	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	N
Alma	0.81	0.80	0.81	0.80	0.74	0.78	0.78	0.78	0.82	0.82	0.80	0.80	19
Chicoutimi	0.84	0.90	0.91	0.89	0.88	0.84	0.81	0.79	0.82	0.82	0.83	0.82	86
Drummondville	0.94	0.91	0.86	0.82	0.79	0.77	0.80	0.79	0.79	0.75	0.78	0.83	33
Montmagny	0.86	0.86	0.86	0.86	0.86	0.86	0.71	0.71	0.71	0.71	0.71	0.83	7
Rimouski	0.67	0.72	0.72	0.67	0.71	0.69	0.62	0.62	0.56	0.44	0.62	0.69	17
Riviere-du-Loup	0.86	0.79	0.71	0.71	0.57	0.57	0.64	0.64	0.64	0.64	0.64	0.64	14
Shawinigan	0.85	0.92	0.92	1.00	1.00	0.92	0.83	0.83	0.83	0.83	0.83	0.70	14
Trois-Rivieres	0.96	0.96	0.96	0.96	0.96	0.92	0.80	0.84	0.93	0.93	0.95	0.95	49
All eight markets	0.87	0.88	0.88	0.87	0.85	0.82	0.78	0.78	0.80	0.79	0.81	0.82	239
Competitive cities	0.89	0.91	0.92	0.92	0.92	0.92	0.92	0.91	0.91	0.91	0.91	0.91	1041

Notes: Cells report the mean of $Share_{it}$ by survey year for each potentially collusive market outside Québec City. The last two rows pool the eight markets and report the competitive control group used in the event study of Figure 3. N is the number of stations observed in 2006. Grand-Mère is excluded throughout the analysis: it contributes a single station that never shares, so it carries no within-station variation. These raw series are reported in place of market-level event studies, which the market sizes cannot support.

Table G.3: Québec City: participation by brand group

Group	2002	2004	2006	2008	2010	2012	Stations (2006)
Discount chain 1	0.74	0.65	0.66	0.40	0.33	0.30	29
Discount chain 2	0.85	0.80	0.76	0.62	0.53	0.58	17
Major branded networks	0.94	0.97	0.96	0.97	0.97	0.94	186
All other stations	0.91	0.94	0.88	0.88	0.85	0.83	33

Notes: Cells report the mean of $Share_{it}$ by survey year for stations in the Québec City market, by brand group. The two discount chains are unbranded distributor networks; major branded networks are the five vertically integrated or franchised brands. The last column reports the number of stations observed in 2006. Group names are withheld under the paper's anonymity convention for the two discount chains.



Notes: Markets classified via k-medians clustering on pre-investigation (2002–2006) median margins, price adjustment asymmetry, and price stickiness. Montréal includes five sub-markets. Three Abitibi-Témiscamingue markets shown in inset. Boundaries: Statistics Canada, 2021 Census boundary files.

Figure G.2: Québec retail gasoline markets by competitive classification

Table G.4: Crosswalk between Régie survey cities and Platform markets

Platform market	Stations (2006)	Régie survey cities
Cartel		
Magog	12	Magog
Sherbrooke	62	Sherbrooke
Thetford Mines	12	Thetford Mines
Victoriaville	22	Victoriaville
Québec City		
Québec City	265	Levis, Québec (Beauport), Québec (Charlesbourg), Québec (Québec), Québec (Sainte-Foy), Saint-Augustin-de-Desmaures
Other potentially collusive		
Alma	19	Alma
Chicoutimi	86	Saguenay (Chicoutimi), Saguenay (La Baie)
Drummondville	33	Drummondville
Grand-Mere	1	Shawinigan (Grand-Mere)
Montmagny	7	Montmagny
Rimouski	17	Rimouski
Riviere-du-Loup	14	Riviere-du-Loup
Shawinigan	14	Shawinigan
Trois-Rivieres	49	Trois-Rivieres
Montreal-area control		
Laval	92	Laval
Montreal (centre)	236	Mont-Royal, Montreal (Saint-Laurent), Montreal (Verdun), Montreal (centre)
Montreal (east)	118	Montreal (Anjou), Montreal-Nord, Repentigny
Montreal (north)	90	Blainville, Mirabel, Saint-Eustache, Terrebonne
Montreal (south shore)	214	Brossard, Chateauguay, Longueuil
Montreal (west)	90	Beaconsfield / Kirkland, Dorval / DDO, Vaudreuil-Dorion
Other competitive control		
Amos	8	Amos
Berthierville	8	Berthierville
Hull	63	Gatineau, Gatineau (Hull)
Joliette	19	Joliette
Lachute	10	Lachute
Rouyn-Noranda	12	Rouyn-Noranda
Saint-Hyacinthe	26	Saint-Hyacinthe
Saint-Jean	16	Saint-Jean-sur-Richelieu
Saint-Jerome	28	Saint-Jerome
Val-d'Or	11	Val-d'Or

Notes: "City" denotes a Régie de l'énergie survey unit; "market" denotes a Platform survey unit. The classification of Section 2.2 is estimated on the Régie city panel: 53 cities with sufficient coverage, of which 4 are cartel cities, 18 potentially collusive, and 31 competitive (Table 1). Classified cities are assigned to Platform markets by name. The table lists the 30 matched markets, which form the sharing-analysis sample: 1,654 stations in 2006; Grand-Mère contributes a single station and is dropped as a singleton in estimation, which yields the 1,653 stations and 29 markets of Table B.1 in the paper. Several cities can map into one market, so the 18 potentially collusive cities correspond to Québec City plus nine other potentially collusive markets, with two cities (Maniwaki and Mont-Joli) unmatched. One of those nine, Grand-Mère, contains a single station and is excluded from the analysis of Section 4.3. Seventeen Platform markets have no classified Régie city and are excluded from the sharing analysis: Baie-Comeau, Chandler, Dolbeau, Forestville, Gaspé, Granby, Iberville, L'Assomption, Mascouche, Mistassini, Roberval, Sept-Îles, Sorel, Saint-Félicien, Saint-Georges-de-Beauce, Tracy, and Valleyfield. Four classified cities have no Platform market and enter the price analysis only: Maniwaki, Mont-Joli, Saint-Lin-Laurentides, and Ville-Marie.

Table G.5: Summary statistics: quarterly volume panel, cartel cities

	Victoriaville		Thetford Mines		Sherbrooke	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Panel A: Sales volume						
Daily gasoline volume (litres)	6,397	4,346	4,264	2,020	6,685	3,505
Log daily volume	8.54	0.69	8.25	0.46	8.64	0.66
Volume reported (direct read)	0.80	0.40	0.86	0.35	0.83	0.38
Panel B: Station characteristics						
Gasoline pumps	8.45	4.25	6.84	3.68	11.14	7.16
Islands	2.03	0.88	1.95	1.01	2.45	1.15
Convenience store	0.80	0.40	0.50	0.50	0.66	0.47
Car wash	0.15	0.36	0.14	0.34	0.21	0.41
Repair shop	0.66	0.47	0.50	0.50	0.41	0.49
Self-service	0.35	0.48	0.50	0.50	0.56	0.50
Stations in the same network	2.47	1.21	2.31	0.78	9.13	4.84
Panel C: Brand affiliation						
Low-price guarantee chain	0.09	0.29	0.24	0.43	0.24	0.43
Other major brand	0.37	0.48	0.58	0.49	0.51	0.50
Unbranded distributor	0.27	0.44	0.02	0.12	0.18	0.39
Retail chain	0.04	0.21	0.08	0.27	0.04	0.20
Independent and other	0.22	0.42	0.08	0.27	0.02	0.14
Panel D: Local market structure						
Rivals within 500 m	1.18	1.09	0.88	0.68	1.11	1.13
Rivals within 1 km	3.22	2.26	2.11	0.76	2.22	1.56
Stations in the market	22.99	2.44	12.46	0.50	62.67	3.27
Stations	27		14		71	
Quarters	32		32		32	
Station-quarters	727		398		2,000	
Coverage	2001–2008		2001–2008		2001–2008	

Notes: Station-quarter observations from the Platform's quarterly survey of the three cartel markets with coverage over the full sample, 2001Q1 to 2008Q4, the window used in the volume analysis of Section 5. Sherbrooke is surveyed through 2011, but we do not use the later quarters. Magog is excluded because it enters the survey in 2005q4. Daily gasoline volume is the cumulative volume of the four gasoline grades recorded between two meter readings, divided by the number of days elapsed; diesel is excluded. Volume is reported only when the surveyor obtained a direct reading, and the third row of Panel A gives the fraction of station-quarters for which this is the case. Means for volume are taken over reported observations. Convenience store, car wash, repair shop, self-service and the brand indicators are dummies. Rival counts are the numbers of competing stations within the stated radius, measured from station coordinates.

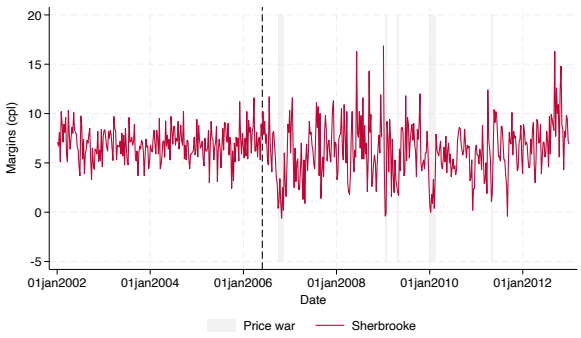
Table G.6: Summary Statistics: Volume Event Study Sample

	Mean	Std. Dev.	Min	Max
Panel A: Dependent Variable and Station Characteristics				
Log Volume	8.548	0.656	5.783	9.969
Number of Pumps	9.529	6.459	2.000	32.000
Number of Islands	2.222	1.057	1.000	6.000
Convenience Store	1.414	1.256	0.000	3.000
Car Wash	0.165	0.371	0.000	1.000
Repair Shop	0.485	0.500	0.000	1.000
Network size	6.543	5.045	1.000	17.000
	500 m		1 km	
	Mean	Std. Dev.	Mean	Std. Dev.
Panel B: Information Sharing Variables				
1(Neighbor Stopped Sharing)	0.126	0.332	0.211	0.408
1(Neighbor Never Sharing)	0.112	0.315	0.239	0.426
Panel C: Neighborhood Composition				
Number of Rivals	1.108	1.121	2.536	1.742
Number of Same-Brand Stations	0.024	0.153	0.150	0.358
Number of Distributor Brands	0.206	0.450	0.493	0.727
Number of Major Brands	0.724	0.836	1.670	1.196

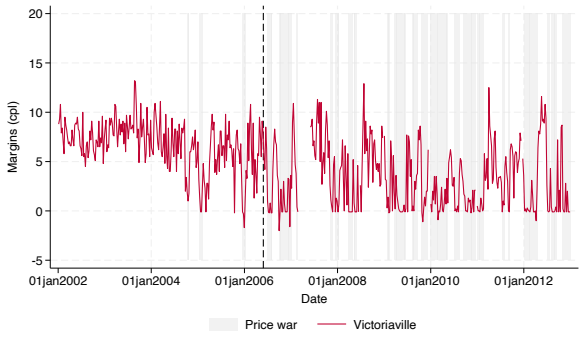
Notes: Sample includes cartel-market stations with valid volume data, excluding Magog and stations that stopped sharing with the Platform. Neighborhood variables are measured within 500 m and 1 km radii. The sample contains 2280 station-quarter observations, 93 unique stations, and 31 quarters. "Neighbor Stopped Sharing" indicates the presence of at least one neighbor that stopped sharing its volume with the Platform after 2006. "Neighbor Never Sharing" indicates the presence of at least one neighboring station that does not share data in the survey year and is not a stop-sharing station. The sample covers 2001Q1 to 2008Q4, excluding the 2006Q1 round.



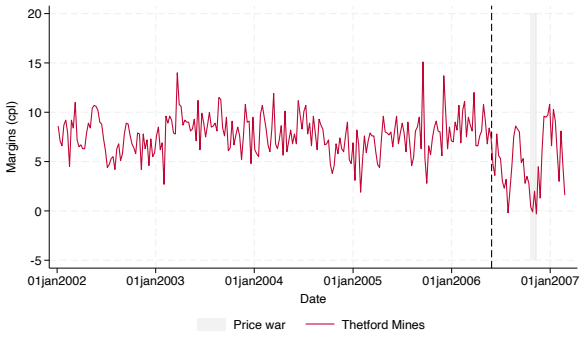
(a) Québec City



(b) Sherbrooke



(c) Victoriaville



(d) Thetford Mines

Figure G.3: Weekly retail margins and price-war episodes

Notes: Weekly retail margins (cents per liter) defined as the posted retail price minus the Régie price floor. Shaded regions identify price-war episodes using a two-step classification: seed periods of at least two consecutive weeks with margins ≤ 2 cents per liter, extended forward to include subsequent weeks with margins ≤ 4 cents per liter.

Table G.7: Pre-trend test and HonestDiD sensitivity analysis: potentially collusive markets

	(1) Other PC excl. QC	(2) Other PC excl. QC, Drum. & RdL
Treatment effects:		
DiD estimate (baseline)	-0.052***	-0.050***
	(0.006)	(0.005)
Pre-trend diagnostics:		
Pre-investigation trend coef.	-0.0083	-0.0007
	(0.0080)	(0.0052)
Joint pre-trend test (p-value)	0.683	0.896
Breakdown values:		
Short-run (2007)	0.30	0.55
Medium-run (2008–2010)	0.10	0.45
Long-run (2011–2013)	0.00	0.00†

Notes: Standard errors clustered at the market level in parentheses. Significance: * p<0.10, ** p<0.05, *** p<0.01. †: breakdown reported even though standard 95% CI marginally contains zero (upper bound within 0.01 of zero).
Pre-investigation trend coef.: estimate of γ from $Share_{it} = \alpha_i + \delta_t + \gamma\,(\text{year}_t - 2002) \cdot D_i + \varepsilon_{it}$, fit on pre-period observations (year $\leq$ 2006). α_i: station fixed effects; δ_t: year fixed effects; D_i: treatment indicator.
Joint pre-trend test: test that the pre-2006 event-study coefficients are zero. The event study is estimated relative to 2006 with market-clustered standard errors, as the relative-magnitudes method requires; the event-study figures in the text use 2002 as the reference year and cluster by station.
Breakdown M: smallest value of $\bar{M}$ under the relative-magnitudes restriction of Rambachan and Roth (2023) for which the 95% robust confidence interval for the post-period effect first contains zero. Larger values indicate greater robustness to post-treatment violations of parallel trends.

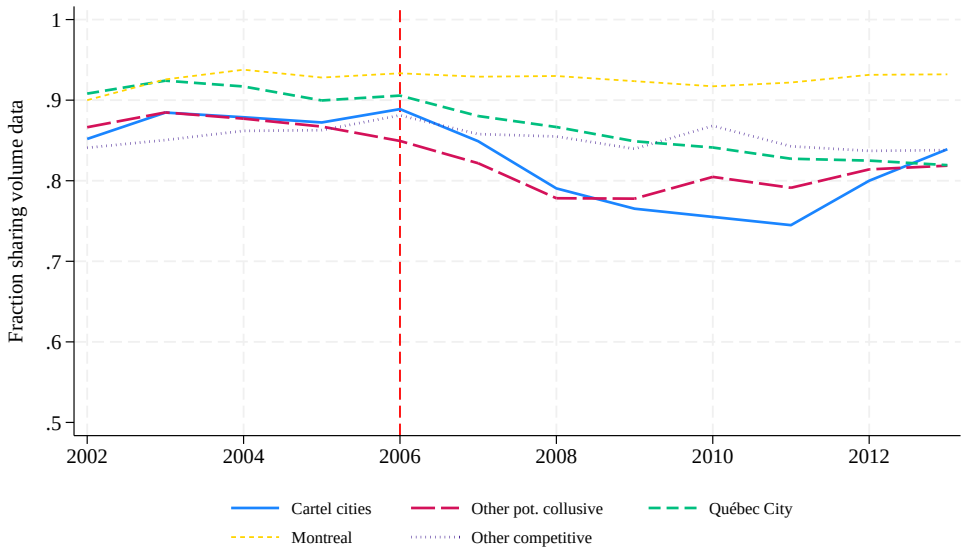


Figure G.4: Raw participation rates by market group
Notes: Mean of $Share_{it}$ by survey year for the five market groups. The dashed vertical line marks the May 2006 search warrants; the survey is a first-quarter snapshot, so 2007 is the first post-investigation observation.

Table G.8: Pre-trend test and HonestDiD analysis under alternative information sharing classification (Specification 1): identified set under bounded violations of parallel trends

	(1) Cartel cities	(2) Québec City‡	(3) Other PC excl. QC	(4) Other PC excl. QC & Drum.	(5) Other PC excl. QC, Drum. & RdL
Treatment effects:					
DiD estimate (baseline)	-0.073***	-0.041***	-0.037***	-0.037***	-0.032***
	(0.019)	(0.003)	(0.008)	(0.009)	(0.006)
Pre-trend diagnostics:					
Pre-investigation trend coef.	-0.0049	-0.0051	-0.0077	-0.0058	-0.0022
	(0.0047)	(0.0029)	(0.0064)	(0.0065)	(0.0045)
Joint pre-trend test (p-value)	0.123	0.020	0.345	0.317	0.402
Breakdown values:					
Short-run (2007)	0.25	0.00	0.00†	0.00†	0.00†
Medium-run (2008–2010)	0.60	0.40	0.15	0.25	0.40
Long-run (2011–2013)	0.00	0.55	0.00	0.00†	0.00†

Notes: Standard errors clustered at the market level in parentheses. Significance: * p<0.10, ** p<0.05, *** p<0.01. †: breakdown reported even though standard 95% CI marginally contains zero (upper bound within 0.01 of zero). ‡: Québec City sample restricted to 2004 onward; see Chaves et al. (2025) on the documented transition from price wars to uniform pricing in 2003–2004.
Pre-investigation trend coef.: estimate of γ from $Share_{it} = \alpha_i + \delta_t + \gamma\,(\text{year}_t - 2002) \cdot D_i + \varepsilon_{it}$, fit on pre-period observations (year ≤ 2006). α_i: station fixed effects; δ_t: year fixed effects; D_i: treatment indicator.
Joint pre-trend test: test that the pre-2006 event-study coefficients are zero. The event study is estimated relative to 2006 with market-clustered standard errors, as the relative-magnitudes method requires; the event-study figures in the text use 2002 as the reference year and cluster by station.
Breakdown M: smallest value of $\bar{M}$ under the relative-magnitudes restriction of Rambachan and Roth (2023) for which the 95% robust confidence interval for the post-period effect first contains zero. Larger values indicate greater robustness to post-treatment violations of parallel trends.

Table G.9: Pre-trend test and HonestDiD analysis under alternative information sharing classification (Specification 2): identified set under bounded violations of parallel trends

	(1) Cartel cities	(2) Québec City‡	(3) Other PC excl. QC	(4) Other PC excl. QC & Drum.	(5) Other PC excl. QC, Drum. & RdL
Treatment effects:					
DiD estimate (baseline)	-0.079***	-0.037***	-0.052***	-0.054***	-0.050***
	(0.016)	(0.003)	(0.006)	(0.007)	(0.005)
Pre-trend diagnostics:					
Pre-investigation trend coef.	-0.0042	-0.0026	-0.0083	-0.0058	-0.0007
	(0.0064)	(0.0029)	(0.0080)	(0.0081)	(0.0052)
Joint pre-trend test (p-value)	0.934	0.063	0.683	0.728	0.896
Breakdown values:					
Short-run (2007)	0.20	0.10	0.30	0.25	0.55
Medium-run (2008–2010)	0.75	0.30	0.10	0.20	0.45
Long-run (2011–2013)	0.00†	0.35	0.00	0.00†	0.00†

Notes: Standard errors clustered at the market level in parentheses. Significance: * p<0.10, ** p<0.05, *** p<0.01. †: breakdown reported even though standard 95% CI marginally contains zero (upper bound within 0.01 of zero). ‡: Québec City sample restricted to 2004 onward; see Chaves et al. (2025) on the documented transition from price wars to uniform pricing in 2003–2004.
Pre-investigation trend coef.: estimate of γ from $Share_{it} = \alpha_i + \delta_t + \gamma\,(\text{year}_t - 2002) \cdot D_i + \varepsilon_{it}$, fit on pre-period observations (year ≤ 2006). α_i: station fixed effects; δ_t: year fixed effects; D_i: treatment indicator.
Joint pre-trend test: test that the pre-2006 event-study coefficients are zero. The event study is estimated relative to 2006 with market-clustered standard errors, as the relative-magnitudes method requires; the event-study figures in the text use 2002 as the reference year and cluster by station.
Breakdown M: smallest value of $\bar{M}$ under the relative-magnitudes restriction of Rambachan and Roth (2023) for which the 95% robust confidence interval for the post-period effect first contains zero. Larger values indicate greater robustness to post-treatment violations of parallel trends.

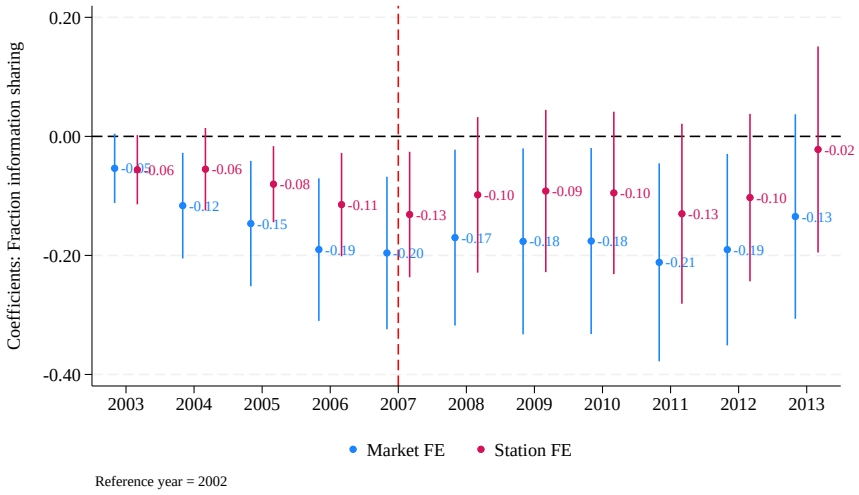


Figure G.5: Effect of the cartel collapse on information sharing: Drummondville

Notes: Each series plots the estimated coefficients $\hat{\beta}_y$ from equation (5), estimated separately with market fixed effects and station fixed effects. The reference year is 2002. The dashed vertical line marks the year of the investigation (2006). The control group consists of non-collusive markets. Standard errors clustered at the station level.

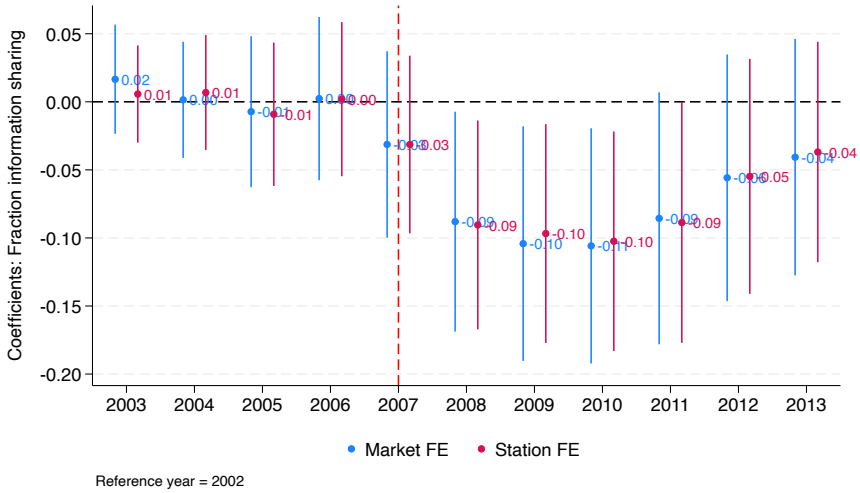


Figure G.6: Effect of the cartel collapse on information sharing under alternative information sharing classification (Specification 1): cartel cities

Notes: Each series plots the estimated coefficients $\hat{\beta}_y$ from equation (5), estimated separately with market fixed effects and station fixed effects. The reference year is 2002. The dashed vertical line marks the year of the investigation (2006). The control group consists of non-collusive markets. Standard errors clustered at the station level.

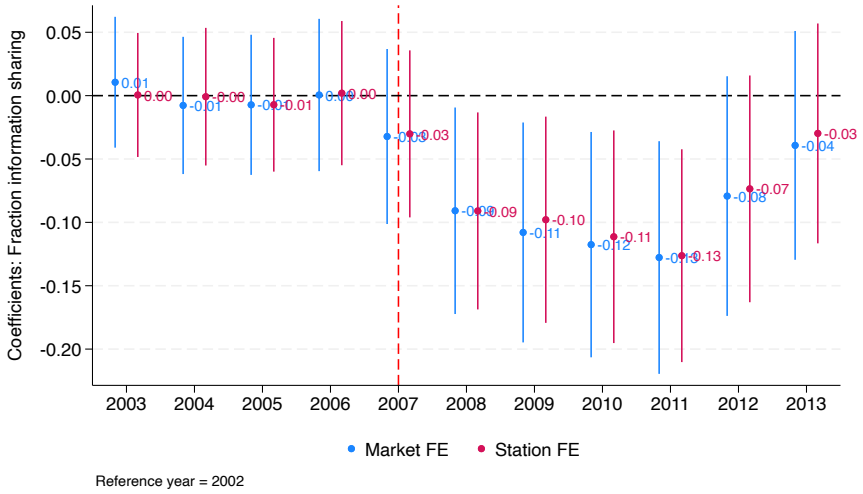


Figure G.7: Effect of the cartel collapse on information sharing under alternative information sharing classification (Specification 2): cartel cities

Notes: Each series plots the estimated coefficients $\hat{\beta}_y$ from equation (5), estimated separately with market fixed effects and station fixed effects. The reference year is 2002. The dashed vertical line marks the year of the investigation (2006). The control group consists of non-collusive markets. Standard errors clustered at the station level.

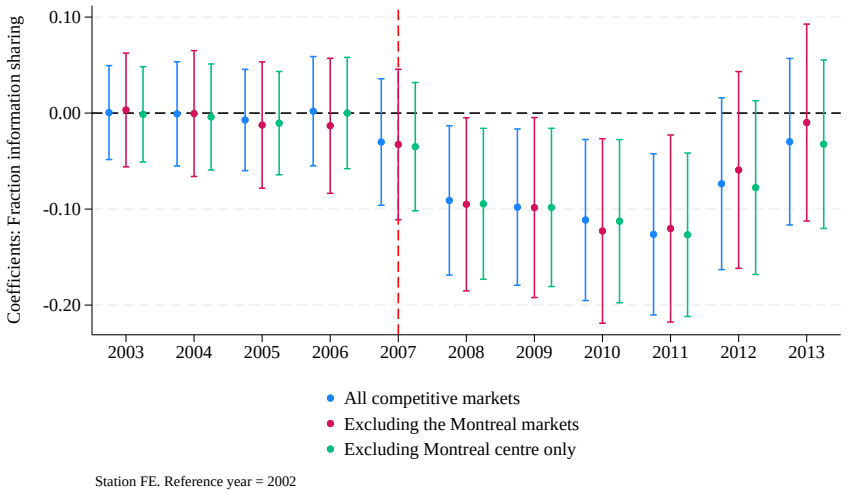


Figure G.8: Effect of the cartel collapse on information sharing: cartel cities, control group without Montreal

Notes: Each series plots the estimated coefficients $\hat{\beta}_y$ from equation (5) with station fixed effects, for three control groups: all competitive markets, as in Figure 2 (1,041 stations in 2006); the competitive markets excluding the six Montreal markets, Montreal centre, Rive-Sud, Est, Ouest, Nord, and Laval (201 stations); and the competitive markets excluding Montreal centre only (805 stations). The reference year is 2002. The dashed vertical line marks the year of the investigation (2006). Standard errors clustered at the station level.

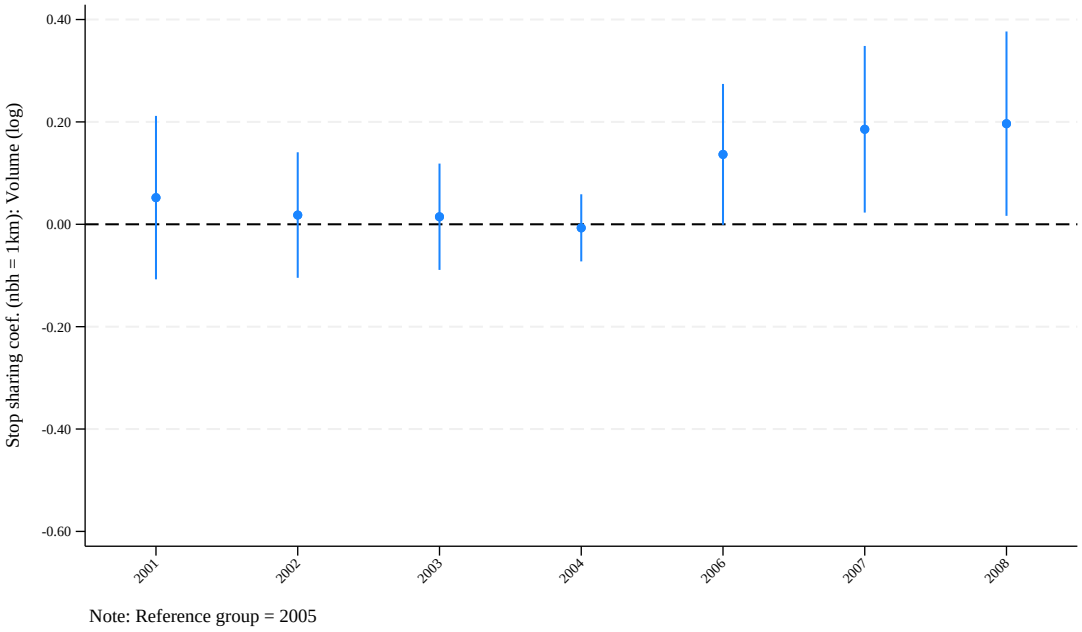
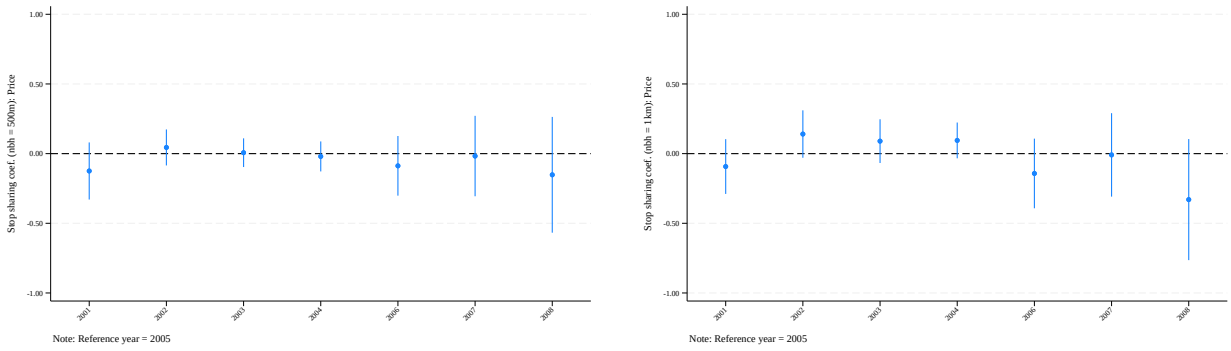


Figure G.9: Volume event study: effect of stop-sharing neighbors (1 km)

Notes: Coefficients from a regression of log quarterly volume on year interactions with **1**(Neighbor stop sharing within 1 km), controlling for never-sharing neighbor year interactions, neighborhood composition, station amenities, and station, brand, and market×quarter fixed effects. The sample runs from 2001 to 2008 and the reference year is 2005. Standard errors clustered at the station level.



(a) 500 m (b) 1 km

Figure G.10: Price event study: effect of stop-sharing neighbors on the posted-price snapshot

Notes: Coefficients from the same specification as Figures 4 and G.9, with the posted price of unleaded gasoline on the survey date (cents per litre) replacing log volume as the dependent variable: year interactions with **1**(Neighbor stop sharing), controlling for never-sharing neighbor year interactions, neighborhood composition, station amenities, and station, brand, and market×quarter fixed effects. The sample runs from 2001 to 2008 and the reference year is 2005. Standard errors clustered at the station level.

H Construction of the information-sharing variables

This appendix documents the construction of the station-year sharing indicator $Share_{it}$ and the cross-sectional indicator $StopShare_i$ introduced in Section 2.4. It lists the survey codes that record why a volume reading was not obtained, states the rules that map them into the indicators, and describes two alternative codings used in the sensitivity analysis. Specification 0 is the baseline used in the main text. Specification 1 replaces the inferential rules used to classify ambiguous survey responses with a strict literal-text rule. Specification 2 retains the baseline classification of $Share_{it}$ but widens the post-period window used to construct $StopShare_i$.

H.1 Survey responses

Table H.1: Volume-status codes in the survey, by period

	2002–2006		2007–2013	
Volume-status code	N	Share	N	Share
BLANK (direct reading)	8399	0.885	9998	0.854
REFUSED	753	0.079	1242	0.106
RESETS	200	0.021	200	0.017
SOME ACTUAL VOLUMES	92	0.010	139	0.012
PUMP PROBLEMS	29	0.003	71	0.006
PUMPS SEIZED	13	0.001	46	0.004
R	1	0.000	3	0.000
SOME REFUSED	0	0.000	2	0.000
FIRST READS	0	0.000	1	0.000
POWER OUTAGE	0	0.000	1	0.000
SITE JUST PICKED UP	3	0.000	0	0.000
FIRST READ	1	0.000	0	0.000
SOME ACTUAL VOLUME	2	0.000	0	0.000
Total	9493		11703	

Notes: Station-year counts of the volume-status code recorded by the provider's surveyors, pooled across all markets. A blank code indicates a valid direct meter reading. The share of station-years coded REFUSED rises from 7.9 percent before the investigation to 10.6 percent after.

For each station–year observation, the Platform records a status code (`st1cds`) and a free-text explanation string. The status code takes four values: blank (a valid volume reading was obtained), `A` and `E` (no reading was obtained; the file does not document the distinction between the two, and we treat them identically), and `T` (station temporarily closed, typically for renovation). A fifth value, `C`, marks a permanent closure, and these station–years are dropped before classification. The explanation field is populated mainly when the status is non-blank, although a small number of valid reads carry quality annotations (`M` or `S`); these annotations are stripped prior to classification and do not affect any indicator. The substantive explanation strings that drive classification fall into three groups:

- **Explicit refusal:** `REFUSED`, `SOME` (partial or selective refusal), and the one-letter shorthand `R`.
- **Non-strategic:** `FIRST READ`, `FIRST READS`, `SITE JUST PICKED UP`, `POWER OUTAGE`.
- **Ambiguous mechanical:** `RESETS`, `PUMP PROBLEMS`, `PUMPS SEIZED`.

The first group documents an explicit refusal to share volume data. We treat `SOME` as a refusal because the two codes are used by the same stations: of the 927 station-years at stations that carry a `SOME` flag in some year, 640, or 69 percent, belong to stations that also carry a `REFUSED` flag in another year. The second group identifies reasons for a missing record that are unrelated to the operator's decision. The third group describes equipment failures. An isolated failure is plausibly genuine, but a mechanical excuse that repeats across annual survey visits, or that appears next to a documented refusal, is more likely to cover a decision to withhold. The rules below treat the third group accordingly.

H.2 Construction of $Share_{it}$

Let $NonShare_{it} = 1 - Share_{it}$. A station–year is classified as $NonShare_{it} = 1$ if the station can be plausibly inferred to have withheld its volume data, and $NonShare_{it} = 0$ otherwise. The three specifications differ only in how the ambiguous mechanical codes are classified.

Specification 0 (baseline). A station–year is assigned $NonShare_{it} = 1$ if either of the following holds.

1. *Explicit refusal.* The explanation contains `REFUSED` or `SOME`, or is the shorthand `R`, and no reading was obtained.
2. *Ambiguous mechanical with adjacency.* The explanation is one of `RESETS`, `PUMP PROBLEMS`, or `PUMPS SEIZED`, no reading was obtained, and at least one of the following adjacency conditions holds:
   (a) *Two-year run rule.* The same station's prior or next observation in the panel also has one of these three ambiguous codes.
   (b) *Confirmed-refusal adjacency rule.* The same station's prior or next observation in the panel is itself classified as $NonShare = 1$ under rule (1) or rule (2a). Rule (2b) is evaluated after rule (2a), so that any reclassification from the two-year run rule propagates one step further.

All other station–years are assigned $NonShare_{it} = 0$. This includes valid reads, the four non-strategic codes, temporary closures (status `T`), and isolated single-year occurrences of the ambiguous mechanical codes.

The rationale for the adjacency logic is that a single-year `RESETS` (or related) entry, surrounded on both sides by valid reports, is most naturally read as a genuine equipment issue. Equipment

failures are unlikely to persist across annual survey visits, so a run of two or more such years is coded as non-sharing. An ambiguous year adjacent to a documented refusal is coded as non-sharing as well, which captures stations that alternate between an outright refusal and a mechanical excuse. Any ambiguous year that survives both rules is an isolated one surrounded by clean reads, and it is coded as sharing.

Specification 1 (explicit refusals only). A station–year is assigned $NonShare_{it} = 1$ if the explanation contains `REFUSED` or `SOME`, or is the shorthand `R`, and no reading was obtained. All other station–years, including every occurrence of the ambiguous mechanical codes, are assigned $NonShare_{it} = 0$.

Specification 1 thus drops the inferential layer entirely and uses only the literal content of the explanation field. It is conservative in the sense that it counts only stations whose non-cooperation is explicitly documented in the survey response.

Specification 2 (extended post window). The $NonShare_{it}$ classification is identical to specification 0. Specifications 0 and 2 differ only in the post-period window used to construct $StopShare_i$ (described below).

Table H.2: Mapping of survey responses to $NonShare_{it}$ across specifications

Survey response	Spec 0	Spec 1	Spec 2
Valid read (`st1cds` blank)	0	0	0
`REFUSED` / `SOME` / `R`, no reading obtained	1	1	1
`FIRST READ` / `SITE JUST PICKED UP` / `POWER OUTAGE`	0	0	0
`st1cds` = `T` (temporarily closed)	0	0	0
`RESETS` / `PUMP PROBLEMS` / `PUMPS SEIZED`, no reading obtained:			
isolated single year	0	0	0
two or more consecutive years	1	0	1
adjacent to a confirmed refusal	1	0	1

H.3 Construction of $StopShare_i$

The cross-sectional indicator $StopShare_i$ identifies stations that cooperated in every pre-period observation but failed to cooperate in at least one post-period observation. The construction rule is identical across the three specifications:

$$StopShare_i \;=\; \mathbf{1}\Big[\,\overline{Share}_i^{\,\text{pre}} \;=\; 1 \;\wedge\; \overline{Share}_i^{\,\text{post}} \;<\; 1\,\Big],$$

where $\overline{Share}_i^{\,\text{pre}}$ and $\overline{Share}_i^{\,\text{post}}$ denote the mean of $Share_{it}$ over the station's pre-period and post-period observations, respectively. Two auxiliary rules apply in every specification: stations whose last observed year is before 2006 are assigned $StopShare_i = 0$, and stations for which either the pre- or post-period mean cannot be computed (because the station has no observations in that window)

are assigned $StopShare_i = \cdot$ (missing). Under the baseline specification the indicator is defined for 2,028 stations and equals one for 86 of them, including 13 of the 121 stations in the cartel cities.

The pre-period is the same in all three specifications: every observation in year ≤ 2006. The specifications differ in the post-period:

- **Specifications 0 and 1:** post-period = year $\in \{2007, 2008, 2009\}$.
- **Specification 2:** post-period = year > 2006, with no upper bound.

The three-year window used in specifications 0 and 1 corresponds to the period immediately following the May 2006 disclosure of the Bureau's investigation, before the partial recovery of cooperation documented in Section 4.1. Specification 2 widens the window to include the medium and long run. A station that resumes cooperation in 2007–2009 but withdraws again in 2010 or later is therefore classified as $StopShare_i = 1$ under specification 2 but as $StopShare_i = 0$ under specifications 0 and 1. Table H.3 summarizes.

Table H.3: Pre- and post-period windows for $StopShare_i$

Specification	Pre-period	Post-period
Spec 0 (baseline)	year ≤ 2006	$\{2007, 2008, 2009\}$
Spec 1 (explicit refusals only)	year ≤ 2006	$\{2007, 2008, 2009\}$
Spec 2 (extended post window)	year ≤ 2006	year > 2006

References

Clark, R. and J.-F. Houde (2014). The effect of explicit communication on pricing: Evidence from the collapse of a gasoline cartel. *The Journal of Industrial Economics 62*, 191–228.