Cartel Dating
Online Appendix

H. Peter Boswijk[*], Maurice J.G. Bun[†] and Maarten Pieter Schinkel[‡]

Appendix A: Proofs

Proof of Proposition 1. We first provide a detailed proof for $R = 1$ and then discuss the straightforward generalization to $R > 1$. Regarding the cartel dummy D_t we use the definition in (2). Following the standard asymptotic analysis of structural breaks (Perron, 1989), we assume that $T_B = \lambda_B T$ and $T_E = \lambda_E T$ with the break fractions λ_B and λ_E fixed numbers for all values of T.

We have for the average estimated but-for price

$$\text{plim}\,\frac{1}{T_E - T_B}\sum_{t=T_B+1}^{T_E}\widehat{\text{bfp}}_t = \text{plim}\,\frac{1}{T_E - T_B}\sum_{t=T_B+1}^{T_E}\text{bfp}_t + \text{plim}\,\frac{1}{T_E - T_B}\sum_{t=T_B+1}^{T_E}\left(\widehat{\text{bfp}}_t - \text{bfp}_t\right).$$

We will analyze the limiting behavior of the second right-hand side term in more detail. Noting that

$$\widehat{\text{bfp}}_{T_B} = \text{bfp}_{T_B} = p_{T_B},$$

in period $T_B + 1$ we have for the estimated and effective but-for prices:

$$\begin{aligned}\widehat{\text{bfp}}_{T_B+1} &= \hat{\gamma} p_{T_B} + \hat{\beta}' x_{T_B+1} + \hat{\alpha}_1, \\ \text{bfp}_{T_B+1} &= \gamma p_{T_B} + \beta' x_{T_B+1} + \alpha_1 + \varepsilon_{T_B+1}.\end{aligned}$$

Therefore, we can write for the prediction error

$$\begin{aligned}v_{T_B+1} &= \widehat{\text{bfp}}_{T_B+1} - \text{bfp}_{T_B+1} \\ &= (\hat{\gamma} - \gamma)p_{T_B} + (\hat{\beta} - \beta)' x_{T_B+1} + \hat{\alpha}_1 - \alpha_1 - \varepsilon_{T_B+1} \\ &= -\varepsilon_{T_B+1} + O_P(T^{-1/2}),\end{aligned}$$

[*] Amsterdam School of Economics and Tinbergen Institute, University of Amsterdam.
[†] Amsterdam School of Economics, University of Amsterdam.
[‡] Amsterdam School of Economics and ACLE, University of Amsterdam.

because estimation errors are $O_P(T^{-1/2})$ under Assumption 1 due to standard asymptotic theory. In period T_B+2 we have for the difference in estimated and effective but-for prices

$$\begin{aligned}
\widehat{\text{bfp}}_{T_B+2} - \text{bfp}_{T_B+2} &= \hat{\gamma}\widehat{\text{bfp}}_{T_B+1} - \gamma \text{bfp}_{T_B+1} + (\hat{\beta}-\beta)' x_{T_B+2} + \hat{\alpha}_1 - \alpha_1 - \varepsilon_{T_B+2} \\
&= \hat{\gamma} v_{T_B+1} + (\hat{\gamma}-\gamma)\,\text{bfp}_{T_B+1} + (\hat{\beta}-\beta)' x_{T_B+2} + \hat{\alpha}_1 - \alpha_1 - \varepsilon_{T_B+2} \\
&= \hat{\gamma} v_{T_B+1} + v_{T_B+2},
\end{aligned}$$

where for the prediction error v_{T_B+2} we have

$$v_{T_B+2} = -\varepsilon_{T_B+2} + O_P(T^{-1/2}).$$

In general, we have for $s = 1, 2, \ldots, T_E - T_B$

$$\begin{aligned}
\widehat{\text{bfp}}_{T_B+s} - \text{bfp}_{T_B+s} &= \hat{\gamma}^{s-1} v_{T_B+1} + \hat{\gamma}^{s-2} v_{T_B+2} + \ldots + \hat{\gamma} v_{T_B+s-1} + v_{T_B+s} \\
&= \sum_{j=0}^{s-1} \gamma^j v_{T_B+s-j} + O_P(T^{-1/2}) \\
&= -\sum_{j=0}^{s-1} \gamma^j \varepsilon_{T_B+s-j} + O_P(T^{-1/2}),
\end{aligned}$$

where the second line follows from $\hat{\gamma} = \gamma + O_P(T^{-1/2})$. Therefore,

$$\text{plim}\, \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \left(\widehat{\text{bfp}}_t - \text{bfp}_t \right) = -\,\text{plim}\, \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \sum_{j=0}^{s-1} \gamma^j \varepsilon_{T_B+s-j} = 0,$$

by Chebyshev's inequality, using $E[\varepsilon_t] = 0$ and $E[\varepsilon_t \varepsilon_s] = \sigma_N^2 + D_t(\sigma_C^2 - \sigma_N^2)$ for $s = t$, and 0 otherwise. This completes the proof for $R = 1$ because, due to stationarity (Assumption 1) we have

$$\text{plim}\, \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \text{bfp}_t = E\left[\text{bfp}_t\right].$$

Consider next multiple effective cartel periods ($R > 1$), i.e. $\mathbb{T}_C = \mathbb{T}_{C_1} \cup \ldots \cup \mathbb{T}_{C_R}$. Each set of effective cartel periods $\mathbb{T}_{C_r}$, $r = 1, \ldots, R$, has a known begin date T_{B_r} and end date T_{E_r}, so that $\mathbb{T}_{C_r} = \{T_{B_r}+1, \ldots, T_{E_r}\}$ with period length $T_{C_r} = T_{E_r} - T_{B_r}$. We assume that $T_{C_r} = \lambda_{C_r} T$ with the break fraction λ_{C_r} a fixed number for all values of T. The above derivations are then valid for each period, and we have

$$\begin{aligned}
\text{plim}\, \frac{1}{T_C} \sum_{t \in \mathbb{T}_C} \widehat{\text{bfp}}_t &= \sum_{r=1}^{R} \lim \frac{T_{C_r}}{T_C}\, \text{plim}\, \frac{1}{T_{C_r}} \sum_{t \in \mathbb{T}_{C_r}} \widehat{\text{bfp}}_t \\
&= \sum_{r=1}^{R} \frac{\lambda_{C_r}}{\lambda_C}\, \text{plim}\, \frac{1}{T_{C_r}} \sum_{t \in \mathbb{T}_{C_r}} \text{bfp}_t \\
&= E\left[\text{bfp}_t\right],
\end{aligned}$$

which completes the proof. □

Proof of Proposition 2. We first provide a detailed proof for $R = 1$ and then discuss the straightforward generalization to $R > 1$. Noting that $\text{bfp}_{T_B} = p_{T_B}$ we have for the overcharge

$$O_{T_B+s} = \left(p_{T_B+s} - \text{bfp}_{T_B+s}\right) = \frac{1-\gamma^s}{1-\gamma}\alpha_2,$$

for $s = 1, 2, \ldots, T_E - T_B$. Therefore, we have for the average effective overcharge

$$\begin{aligned}
\operatorname{plim} \bar{O} &= \operatorname{plim} \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} (p_t - \text{bfp}_t) \\
&= \lim \frac{1}{T_E - T_B} \sum_{s=1}^{T_E-T_B} \frac{1-\gamma^s}{1-\gamma}\alpha_2 \\
&= \frac{\alpha_2}{1-\gamma}.
\end{aligned}$$

Using $\widehat{O}_{1t}$ we have for the average estimated overcharge

$$\begin{aligned}
\operatorname{plim} \overline{O}_1 &= \operatorname{plim} \frac{1}{T_E - T_B} \sum_{s=1}^{T_E-T_B} \frac{1-\hat{\gamma}^s}{1-\hat{\gamma}}\hat{\alpha}_2 \\
&= \operatorname{plim} \frac{\hat{\alpha}_2}{1-\hat{\gamma}} \left(1 - \operatorname{plim} \frac{1}{T_E - T_B} \sum_{s=1}^{T_E-T_B} \hat{\gamma}^s\right) \\
&= \frac{\alpha_2}{1-\gamma}.
\end{aligned}$$

Using $\widehat{O}_{2t}$ we have for the average estimated overcharge

$$\begin{aligned}
\operatorname{plim} \overline{O}_2 &= \operatorname{plim} \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \left(p_t - \widehat{\text{bfp}}_t\right) \\
&= \operatorname{plim} \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} (p_t - \text{bfp}_t) - \operatorname{plim} \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \left(\widehat{\text{bfp}}_t - \text{bfp}_t\right) \\
&= \operatorname{plim} \bar{O},
\end{aligned}$$

which follows directly from Proposition 1.

Consider next multiple periods of cartel effects ($R > 1$). The above derivations are valid for each period and the probability limit of the overcharge $\frac{\alpha_2}{1-\gamma}$ is constant for each period. Therefore, when $R > 1$ we simply have

$$\operatorname{plim} \bar{O} = \sum_{r=1}^{R} \frac{\lambda_{C_r}}{\lambda_C} \frac{\alpha_2}{1-\gamma} = \frac{\alpha_2}{1-\gamma},$$

and similarly for $\bar{O}_1$ and $\bar{O}_2$. This completes the proof. □

Assumption $2(ii)$ **under Linear Demand.** Suppose model (1) and Assumption 1 hold. Consider the stylized linear demand function

$$Q_t = a + bp_t + \epsilon_t,$$

where $a > 0$ and ϵ_t is an error term with $E[\epsilon_t|p_t] = 0$. Then Assumption 2 (ii) holds when demand is downward sloping ($b \leq 0$) and x_t is strictly exogenous, i.e. $E[x_t\varepsilon_s] = 0$ for all $s, t = 1, \ldots, T$.

From model (1),

$$p_t = \frac{\alpha_1}{1-\gamma} + \alpha_2 \sum_{i=0}^{\infty} \gamma^i D_{t-i} + \beta \sum_{i=0}^{\infty} \gamma^i x_{t-i} + \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-i},$$

so that

$$Q_t = a + \frac{b\alpha_1}{1-\gamma} + b\alpha_2 \sum_{i=0}^{\infty} \gamma^i D_{t-i} + b\beta \sum_{i=0}^{\infty} \gamma^i x_{t-i} + b \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-i} + \epsilon_t.$$

Under Assumption 1 and strict exogeneity of x_t,

$$\begin{aligned} E\left[Q_t \varepsilon_{t-j}\right] &= E\left[\left(b\beta \sum_{i=0}^{\infty} \gamma^i x_{t-i} + b \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-i} + \epsilon_t\right) \varepsilon_{t-j}\right] \\ &= b\gamma^j E[\varepsilon_{t-j}^2] \\ &\leq 0, \end{aligned}$$

for $b \leq 0$, with equality if and only if $b = 0$, as $0 < \gamma < 1$.

The example above assumes $E[\epsilon_t|p_t] = 0$, which effectively says that demand shocks ϵ_t do not affect prices p_t. When $E[\epsilon_t|p_t] \neq 0$, Assumption 2 (ii) may still hold under weak additional conditions, as the following example illustrates.[1] Consider the symmetric Cournot oligopoly model. The inverse demand function is

$$p_t = e - fQ_t + u_{pt},$$

where $e = -a/b$, $f = -1/b$ and $u_{pt} = -\epsilon_t/b$. Suppose that the marginal costs of production have mean c and can be expressed as $c + u_{ct}$, with u_{ct} a random cost shock. For ease of exposition, assume that the demand and cost shocks u_{pt} and u_{ct} are uncorrelated, mean-zero random variables with variances σ_p^2 and σ_c^2, respectively. Then the Cournot (competitive) equilibrium with n firms implies

$$p_{nt} = \frac{e + nc + nu_{ct} + u_{pt}}{n+1},$$

$$Q_{nt} = \frac{n}{n+1} \frac{e - c + u_{pt} - u_{ct}}{f},$$

[1] We are indebted to one of the anonymous referees for suggesting this illustrative example.

with p_{nt} and Q_{nt} the equilibrium prices and quantities, respectively. If we furthermore assume the cartel to behave as a monopolist, the joint-profit maximizing cartel (monopoly) price and quantity are

$$p_{mt} = \frac{e + c + u_{ct} + u_{pt}}{2},$$
$$Q_{mt} = \frac{e - c + u_{pt} - u_{ct}}{2f}.$$

In this market, the observed price is the Cournot (competitive) price p_{nt} if $D_t = 0$, and the cartel (monopoly) price p_{mt} if $D_t = 1$, that is

$$p_t = p_{nt} + D_t(p_{mt} - p_{nt}) = \begin{cases} p_{nt}, & D_t = 0, \\ p_{mt}, & D_t = 1. \end{cases}$$

The Cournot price can be written as

$$p_{nt} = \alpha_1 + \eta_t, \qquad \alpha_1 = \frac{e + nc}{n+1}, \qquad \eta_t = \frac{u_{pt} + nu_{ct}}{n+1},$$

while the difference between monopoly prices and Cournot prices, i.e. the overcharge, is

$$p_{mt} - p_{nt} = \alpha_2 + \upsilon_t, \qquad \alpha_2 = \frac{n-1}{2(n+1)}(e - c), \quad \upsilon_t = \frac{(n-1)(u_{pt} - u_{ct})}{2(n+1)}.$$

Summarizing, we can write for the observed price

$$\begin{aligned} p_t &= \alpha_1 + \alpha_2 D_t + \eta_t + D_t \upsilon_t \\ &= \alpha_1 + \alpha_2 D_t + \varepsilon_t, \end{aligned}$$

which is a stripped down version of model (1). Note that

$$\sigma_N^2 = \text{Var}\left(\varepsilon_t | D_t = 0\right) = \text{Var}(\eta_t) = \frac{\sigma_p^2 + n^2\sigma_c^2}{(n+1)^2},$$
$$\sigma_C^2 = \text{Var}\left(\varepsilon_t | D_t = 1\right) = \text{Var}(\eta_t + \upsilon_t) = \frac{\sigma_p^2 + \sigma_c^2}{4},$$

which implies different variances during cartel and non-cartel periods, in agreement with Assumption 1. Furthermore, we have

$$\text{Cov}\left(Q_t, \varepsilon_t | D_t = 0\right) = \text{Cov}\left(Q_{nt}, \eta_t\right) = \frac{n\left(\sigma_p^2 - n\sigma_c^2\right)}{f(n+1)^2},$$
$$\text{Cov}\left(Q_t, \varepsilon_t | D_t = 1\right) = \text{Cov}\left(Q_{mt}, \eta_t + \upsilon_t\right) = \frac{\sigma_p^2 - \sigma_c^2}{4f}.$$

The sign of these covariances clearly depends on the relative magnitude of the variances of the (inverse) demand shocks and cost shocks. When $\sigma_c^2 > \sigma_p^2$, Assumption 2 (*ii*) is satisfied. Analogously, specific additional conditions can be formulated for more complex models of oligopolistic competition. □

Proof of Theorem 1. We first provide a detailed proof for $R = 1$ and then discuss the straightforward generalization to $R > 1$. For the effective damage, we have

$$\begin{aligned}
\operatorname{plim} \frac{1}{T}\mathrm{CD} &= \operatorname{plim} \frac{1}{T} \sum_{t=T_B+1}^{T_E} (p_t - \mathrm{bfp}_t)\, Q_t \\
&= \operatorname{plim} \frac{1}{T} \sum_{s=1}^{T_E-T_B} \frac{1-\gamma^s}{1-\gamma} \alpha_2 Q_{T_B+s} \\
&= \frac{\alpha_2}{1-\gamma} \left(\frac{T_E - T_B}{T} \operatorname{plim} \frac{1}{T_E - T_B} \sum_{s=1}^{T_E-T_B} Q_{T_B+s} - \frac{1}{T} \sum_{s=1}^{T_E-T_B} \gamma^s Q_{T_B+s} \right) \\
&= \frac{\alpha_2}{1-\gamma} (\lambda_E - \lambda_B)\, Q_C. \qquad \text{(A.1)}
\end{aligned}$$

The final result follows from (15), together with the fact that $\sum_{s=1}^{T_E-T_B} \gamma^s Q_{T_B+s} = O_P(1)$.

Using $\widehat{O}_{1t}$ we have for the average estimated damage

$$\begin{aligned}
\operatorname{plim} \frac{1}{T}\widehat{\mathrm{CD}}_1 &= \operatorname{plim} \frac{1}{T_E - T_B} \sum_{s=1}^{T_E-T_B} \frac{1-\hat{\gamma}^s}{1-\hat{\gamma}} \hat{\alpha}_2 Q_{T_B+s} \\
&= \frac{\alpha_2}{1-\gamma} (\lambda_E - \lambda_B)\, Q_C,
\end{aligned}$$

which follows directly (A.1) and consistency of $\hat{\gamma}$. Using $\widehat{O}_{2t}$ we have for the average estimated damage

$$\begin{aligned}
\operatorname{plim} \frac{1}{T}\widehat{\mathrm{CD}}_2 &= \operatorname{plim} \frac{1}{T} \sum_{t=T_B+1}^{T_E} \left(p_t - \widehat{\mathrm{bfp}}_t\right) Q_t \\
&= \operatorname{plim} \frac{1}{T} \sum_{t=T_B+1}^{T_E} (p_t - \mathrm{bfp}_t)\, Q_t + \operatorname{plim} \frac{1}{T} \sum_{t=T_B+1}^{T_E} \left(\mathrm{bfp}_t - \widehat{\mathrm{bfp}}_t\right) Q_t \\
&= \operatorname{plim} \frac{1}{T}\mathrm{CD} - \operatorname{plim} \frac{1}{T} \sum_{s=1}^{T_E-T_B} \left(\widehat{\mathrm{bfp}}_{T_B+s} - \mathrm{bfp}_{T_B+s}\right) Q_{T_B+s}.
\end{aligned}$$

We will analyze the limiting behavior of the second term in more detail under Assumption 2. Proposition 1 implies that

$$\left(\widehat{\mathrm{bfp}}_{T_B+s} - \mathrm{bfp}_{T_B+s}\right) Q_{T_B+s} = -\sum_{j=0}^{s-1} \gamma^j \varepsilon_{T_B+s-j} Q_{T_B+s} + O_P(T^{-1/2}).$$

Assumption 2 implies that

$$-\sum_{j=0}^{s-1} E\left[\gamma^j \varepsilon_{T_B+s-j} Q_{T_B+s}\right] \geq 0,$$

with equality if and only if Assumption 2 (ii) holds with equality. Applying a LLN we find

$$\operatorname{plim}\frac{1}{T}\sum_{s=1}^{T_E-T_B}\left(\widehat{\text{bfp}}_{T_B+s}-\text{bfp}_{T_B+s}\right)Q_{T_B+s}\geq 0,$$

again with equality if and only if Assumption 2 (ii) holds with equality.

When there are multiple periods of cartel effects ($R>1$), the above derivations are valid for each period. When $R>1$ we have

$$\begin{aligned}\operatorname{plim}\frac{1}{T}\text{CD}&=\sum_{r=1}^{R}\operatorname{plim}\frac{1}{T}\sum_{t\in\mathbb{T}_{C_r}}(p_t-\text{bfp}_t)\,Q_t\\&=\sum_{r=1}^{R}\frac{\alpha_2}{1-\gamma}\lambda_{C_r}Q_C=\frac{\alpha_2}{1-\gamma}\lambda_C Q_C,\end{aligned}$$

because $\lambda_C=\sum_{r=1}^{R}\lambda_{C_r}$. Because for each period $\widehat{\text{CD}}_1$ is consistent for CD, we have

$$\operatorname{plim}\frac{1}{T}\widehat{\text{CD}}_1=\operatorname{plim}\frac{1}{T}\text{CD}.$$

Regarding $\widehat{\text{CD}}_2$ the weak inequality holds for each period, hence

$$\operatorname{plim}\frac{1}{T}\widehat{\text{CD}}_2\leq\operatorname{plim}\frac{1}{T}\text{CD},$$

which completes the proof. □

Proof of Lemma 1. It is helpful for further calculations on the OLS inconsistency to centre all regressors such they have mean zero. The reason is that we have a constant term in the model, so all variables can be taken in deviation from their sample average, affecting only the definition and interpretation of the intercept. If we let x_t^0 denote the original control variable, we then have

$$x_t=x_t^0-\bar{x}^0,$$

such that

$$m_x=\operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}x_t=0.$$

Moreover, we can redefine the cartel dummy variables D_t and d_t such that actually the measurement error $v_t=d_t-D_t$ has mean zero. If we let D_t^0 and d_t^0 denote the original 0-1 dummy variables, we define D_t and d_t as dummy variables in deviation from their sample mean:

$$\begin{aligned}D_t&=D_t^0-(\lambda_{01}+\lambda_{11}),\\d_t&=d_t^0-(\lambda_{10}+\lambda_{11}).\end{aligned}$$

The result is that, irrespective of the type of misclassification, the average measurement error is zero, i.e.

$$\frac{1}{T}\sum_{t=1}^{T} v_t = 0.$$

The reparametrized model is

$$\begin{aligned} p_t &= \alpha_1 + \alpha_2 D_t^0 + \beta x_t^0 + \varepsilon_t \\ &= \alpha + \alpha_2 D_t + \beta x_t + \varepsilon_t, \end{aligned}$$

so that the new intercept becomes

$$\alpha = \alpha_1 + (\lambda_{01} + \lambda_{11})\,\alpha_2 + \beta \bar{x}^0.$$

Analogously, the estimated model becomes

$$\begin{aligned} \hat{p}_t &= \hat{\alpha}_1 + \hat{\alpha}_2 d_t^0 + \hat{\beta} x_t^0 \\ &= \hat{\alpha} + \hat{\alpha}_2 d_t + \hat{\beta} x_t, \end{aligned}$$

with

$$\hat{\alpha} = \hat{\alpha}_1 + (\lambda_{10} + \lambda_{11})\,\hat{\alpha}_2 + \hat{\beta}\bar{x}^0.$$

Stacking the observations ($t = 1, \ldots, T$), we write the regression model to be estimated as

$$y = Z\theta + u,$$

where $y = (p_1, \ldots, p_T)'$ and $u = (u_1, \ldots, u_T)'$. Furthermore, $Z = (z_1, \ldots, z_T)'$ with $z_t = (x_t, 1, d_t)'$ and $\theta = (\beta, \alpha, \alpha_2)'$. The OLS estimator of the full parameter vector θ is equal to

$$\hat{\theta} = (Z'Z)^{-1}Z'y.$$

Taking the probability limit we have

$$\begin{aligned} \operatorname{plim}\hat{\theta} &= \theta + \left(\operatorname{plim}\frac{1}{T}Z'Z\right)^{-1}\operatorname{plim}\frac{1}{T}Z'u \\ &= \theta + \Sigma_{ZZ}^{-1}\Sigma_{Zu}. \end{aligned} \tag{A.2}$$

The vector $\Sigma_{ZZ}^{-1}\Sigma_{Zu}$ is the OLS inconsistency.

Regarding Σ_{Zu} in (A.2) we have:

$$\begin{aligned} m_{xu} &= \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} x_t u_t = 0, \\ m_u &= \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} u_t = 0, \end{aligned}$$

where $u_t = \varepsilon_t - \alpha_2 v_t$. Also we have

$$E\left[d_t u_t\right] = -\alpha_2 E\left[d_t v_t\right] \neq 0,$$

hence $m_{du} \neq 0$. Regarding Σ_{ZZ} in (A.2) we furthermore have:

$$\begin{aligned} m_x = m_d &= 0, \\ m_{xd} &= 0, \end{aligned}$$

where the last equality follows from Assumption 1. Under the assumptions, we have now:

$$\Sigma_{ZZ} = \begin{pmatrix} m_{xx} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & m_{dd} \end{pmatrix}, \qquad \Sigma_{Zu} = \begin{pmatrix} 0 \\ 0 \\ m_{du} \end{pmatrix},$$

hence the inconsistency simply becomes

$$\Sigma_{ZZ}^{-1} \Sigma_{Zu} = \begin{pmatrix} 0 \\ 0 \\ \frac{m_{du}}{m_{dd}} \end{pmatrix}.$$

The inconsistency for α and β is zero, while it is $\frac{m_{du}}{m_{dd}}$ for α_2. Evaluating the latter we have

$$\begin{aligned} m_{du} &= \operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} d_t u_t \\ &= \operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} d_t \left(\varepsilon_t - \alpha_2 \left(d_t - D_t\right)\right) \\ &= -\alpha_2 \left(m_{dd} - m_{dD}\right), \end{aligned}$$

hence we have

$$\begin{aligned} \operatorname{plim} \hat{\alpha}_2 &= \alpha_2 \frac{m_{dD}}{m_{dd}} \\ &= \alpha_2 \left(\frac{\lambda_{11} - \left(\lambda_{10} + \lambda_{11}\right)\left(\lambda_{01} + \lambda_{11}\right)}{\left(\lambda_{10} + \lambda_{11}\right)\left(1 - \lambda_{10} - \lambda_{11}\right)} \right). \end{aligned}$$

Define:

$$a = \lambda_{11}, \quad b = \lambda_{10} + \lambda_{11}, \quad c = \lambda_{01} + \lambda_{11},$$

then the attenuation bias $\mathrm{AB} = \operatorname{plim} \hat{\alpha}_2 / \alpha_2$ becomes

$$\mathrm{AB} = \frac{a - bc}{b(1-b)}.$$

We have $0 < b < 1$, hence $b(1-b) > 0$. Furthermore, because $0 \leq a \leq c < 1$ we have that

$$\mathrm{AB} = \frac{a - bc}{b(1-b)} \leq \frac{a - ba}{b(1-b)} = \frac{a}{b} \leq 1.$$

The equality signs only hold when both $\lambda_{01} = 0$ and $\lambda_{10} = 0$. Therefore with misdating we have that AB < 1, hence $\text{plim}\,\hat{\alpha}_2 < \alpha_2$.

Regarding $\hat{\alpha}_1$ we have

$$\begin{aligned}\text{plim}\,\hat{\alpha}_1 &= \alpha_1 + \text{plim}\,(\hat{\alpha} - \alpha) - (\lambda_{10} + \lambda_{11})\,\text{plim}\,\hat{\alpha}_2 + (\lambda_{01} + \lambda_{11})\,\alpha_2 - \text{plim}\,\bar{x}^0\,\text{plim}\left(\hat{\beta} - \beta\right)\\ &= \alpha_1 - (\lambda_{10} + \lambda_{11})\,\text{plim}\,\hat{\alpha}_2 + (\lambda_{01} + \lambda_{11})\,\alpha_2\\ &= \alpha_1 + \alpha_2\frac{\lambda_{01}}{\lambda_{00} + \lambda_{01}}.\end{aligned}$$

We therefore have

$$\text{plim}\,\hat{\alpha}_1 \geq \alpha_1,$$

with equality only when $\lambda_{01} = 0$. This completes the proof. □

Proof of Proposition 3.

(i) The average estimated but-for price for the DGP (16) is

$$\text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\widehat{\text{bfp}}_t = \text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\text{bfp}_t + \text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\left(\widehat{\text{bfp}}_t - \text{bfp}_t\right).$$

We will analyze the limiting behavior of the second right-hand side term in more detail. We can write for the probability limit of the average prediction error

$$\begin{aligned}\text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\left(\widehat{\text{bfp}}_t - \text{bfp}_t\right) &= \text{plim}(\hat{\beta} - \beta)\,\text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}x_t + \text{plim}\,(\hat{\alpha}_1 - \alpha_1) - \text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\varepsilon_t\\ &= \text{plim}\,(\hat{\alpha}_1 - \alpha_1)\,,\end{aligned}$$

noting that Lemma 1 proofs that $\text{plim}(\hat{\beta} - \beta) = 0$. The result now follows straightforwardly from Lemma 1.

(ii) From Lemma 1 we know that $0 < \text{plim}\,\hat{\alpha}_2 < \alpha_2$ always from which the result for $\overline{O}_1$ follows. Using $\widehat{O}_{2t}$ we have for the average estimated overcharge

$$\begin{aligned}\text{plim}\,\overline{O}_2 &= \frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\left(p_t - \widehat{\text{bfp}}_t\right)\\ &= \text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}(p_t - \text{bfp}_t) - \text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}\left(\widehat{\text{bfp}}_t - \text{bfp}_t\right)\\ &\leq \text{plim}\,\frac{1}{T_c}\sum_{t\in\mathbb{T}_c}(p_t - \text{bfp}_t)\,,\end{aligned}$$

where the equality holds only when $\lambda_{01} = 0$. Furthermore, note that

$$\begin{aligned}\operatorname{plim} \frac{1}{T_c} \sum_{t \in \mathbb{T}_c} (p_t - \mathrm{bfp}_t) &= \frac{\lambda_{10}}{\lambda_{10} + \lambda_{11}} \operatorname{plim} \frac{1}{T_{10}} \sum_{t \in \mathbb{T}_{10}} (p_t - \mathrm{bfp}_t) \\ &\quad + \frac{\lambda_{11}}{\lambda_{10} + \lambda_{11}} \operatorname{plim} \frac{1}{T_{11}} \sum_{t \in \mathbb{T}_{11}} (p_t - \mathrm{bfp}_t) \\ &= \frac{\lambda_{11}}{\lambda_{10} + \lambda_{11}} \alpha_2 \\ &\leq \alpha_2,\end{aligned}$$

where the equality holds only when $\lambda_{10} = 0$. Because in case of misdating λ_{01} and λ_{10} cannot be zero at the same time, we have

$$\operatorname{plim} \overline{O}_2 < \alpha_2.$$

This completes the proof. □

Proof of Theorem 2. We have for the effective damage

$$\begin{aligned}\operatorname{plim} \frac{1}{T} \mathrm{CD} &= \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_C} (p_t - bfp_t)\, Q_t \\ &= \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_C} \alpha_2 Q_t \\ &= \alpha_2 (\lambda_{01} + \lambda_{11})\, Q_C.\end{aligned}$$

Regarding the estimator $\widehat{\mathrm{CD}}_1$ we have

$$\begin{aligned}\operatorname{plim} \frac{1}{T} \widehat{\mathrm{CD}}_1 &= \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} \hat{\alpha}_2 Q_t \\ &= \operatorname{plim} \hat{\alpha}_2 \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} Q_t \\ &= \alpha_2 \left(\frac{\lambda_{11} - (\lambda_{10} + \lambda_{11})(\lambda_{01} + \lambda_{11})}{(\lambda_{10} + \lambda_{11})(1 - \lambda_{10} - \lambda_{11})} \right) \\ &\quad \times (\lambda_{10} Q_N + \lambda_{11} Q_C).\end{aligned}$$

Comparing $\widehat{\mathrm{CD}}_1$ with CD, when $Q_N \neq Q_C$ no systematic pattern emerges, i.e.

$$\operatorname{plim} \frac{1}{T} \widehat{\mathrm{CD}}_1 \gtreqless \operatorname{plim} \frac{1}{T} \mathrm{CD}.$$

When $Q_N = Q_C$, however, some algebra shows that

$$\begin{aligned}\operatorname{plim} \frac{1}{T} \widehat{\mathrm{CD}}_1 &= \operatorname{plim} \frac{1}{T} \mathrm{CD} - \alpha_2 \frac{\lambda_{01}}{1 - \lambda_{10} - \lambda_{11}} Q_C \\ &\leq \operatorname{plim} \frac{1}{T} \mathrm{CD}.\end{aligned}$$

When $\lambda_{01} = 0$ we furthermore have

$$\begin{aligned}
\operatorname{plim} \frac{1}{T}\widehat{\mathrm{CD}}_1 &= \alpha_2 \frac{\lambda_{11}}{\lambda_{10} + \lambda_{11}} \left(\lambda_{10} Q_N + \lambda_{11} Q_C\right) \\
&\geq \alpha_2 \lambda_{11} Q_C \\
&= \operatorname{plim} \frac{1}{T}\mathrm{CD},
\end{aligned}$$

where the exact equality holds when $Q_N = Q_C$ too. When $\lambda_{10} = 0$ we have

$$\begin{aligned}
\operatorname{plim} \frac{1}{T}\widehat{\mathrm{CD}}_1 &= \alpha_2 \lambda_{11} \frac{1 - \lambda_{01} - \lambda_{11}}{1 - \lambda_{11}} Q_C \\
&< \alpha_2 \lambda_{11} Q_C \\
&< \operatorname{plim} \frac{1}{T}\mathrm{CD}.
\end{aligned}$$

Regarding the $\widehat{\mathrm{CD}}_2$ estimator we have

$$\begin{aligned}
\operatorname{plim} \frac{1}{T}\widehat{\mathrm{CD}}_2 &= \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} \left(p_t - \widehat{\mathrm{bfp}}_t\right) Q_t \\
&= \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} \left(p_t - \mathrm{bfp}_t\right) Q_t - \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} \left(\widehat{\mathrm{bfp}}_t - \mathrm{bfp}_t\right) Q_t.
\end{aligned}$$

We will analyze the limiting behavior of both the first and second terms in more detail. Regarding the first term we have

$$\begin{aligned}
\operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} \left(p_t - \mathrm{bfp}_t\right) Q_t &= \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_C} \left(p_t - \mathrm{bfp}_t\right) Q_t \\
&\quad + \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_{10}} \left(p_t - \mathrm{bfp}_t\right) Q_t - \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_{01}} \left(p_t - \mathrm{bfp}_t\right) Q_t \\
&= \operatorname{plim} \frac{1}{T}\mathrm{CD} - \alpha_2 \lambda_{01} Q_C \\
&\leq \operatorname{plim} \frac{1}{T}\mathrm{CD},
\end{aligned}$$

where equality holds only if $\lambda_{01} = 0$. Regarding the second term we have

$$\widehat{\mathrm{bfp}}_t - \mathrm{bfp}_t = \left(\hat{\alpha}_1 - \alpha_1\right) + \left(\hat{\beta} - \beta\right) x_t - \varepsilon_t,$$

and using the results from Lemma 1 and Assumption 2 we can write:

$$\begin{aligned}
\operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} \left(\hat{\alpha}_1 - \alpha_1\right) Q_t &= \operatorname{plim} \left(\hat{\alpha}_1 - \alpha_1\right) \operatorname{plim} \frac{1}{T} \sum_{t \in \mathbb{T}_c} Q_t \\
&= \alpha_2 \frac{\lambda_{01}}{\lambda_{00} + \lambda_{01}} \left(\lambda_{10} Q_N + \lambda_{11} Q_C\right) \\
&\geq 0,
\end{aligned}$$

$$\text{plim}\,\frac{1}{T}\sum_{t\in\mathbb{T}_c}\left(\hat{\beta}-\beta\right)x_tQ_t = \text{plim}\left(\hat{\beta}-\beta\right)\text{plim}\,\frac{1}{T}\sum_{t\in\mathbb{T}_c}x_tQ_t$$
$$= 0,$$

$$\text{plim}\,\frac{1}{T}\sum_{t\in\mathbb{T}_c}\varepsilon_tQ_t \leq 0.$$

Collecting terms we find

$$\text{plim}\,\frac{1}{T}\sum_{t\in\mathbb{T}_c}\left(\widehat{\text{bfp}}_t - \text{bfp}_t\right)Q_t \geq 0.$$

Equality results only when price and quantity are uncorrelated and $\lambda_{01} = 0$. In all other cases a strict inequality follows. This completes the proof. □

Proof of Lemma 2. It is helpful for further calculations on the OLS inconsistency to invoke the result that without loss of generalization we can assume that x_t has mean zero, hence

$$m_x = \text{plim}\,\frac{1}{T}\sum_{t=1}^{T}x_t = 0.$$

Moreover, we can redefine the cartel dummy variable such that actually the measurement error $v_t = d_t - D_t$ has mean zero. If we let D_t^0 and d_t^0 denote the original $0-1$ dummy variables, such that $D_t = D_t^0 - (\lambda_E - \lambda_B)$ and $d_t = d_t^0 - (\lambda_e - \lambda_b)$, then

$$\begin{aligned} p_t &= \alpha_1 + \alpha_2 D_t^0 + \beta' x_t + \gamma p_{t-1} + \varepsilon_t \\ &= \alpha + \alpha_2 D_t + \beta' x_t + \gamma p_{t-1} + \varepsilon_t, \end{aligned}$$

so that the new intercept becomes

$$\alpha = \alpha_1 + (\lambda_E - \lambda_B)\,\alpha_2.$$

Analogously, the estimated model becomes

$$\begin{aligned} \hat{p}_t &= \hat{\alpha}_1 + \hat{\alpha}_2 d_t^0 + \hat{\beta}' x_t + \hat{\gamma} p_{t-1} \\ &= \hat{\alpha} + \hat{\alpha}_2 d_t + \hat{\beta}' x_t + \hat{\gamma} p_{t-1}, \end{aligned}$$

with

$$\hat{\alpha} = \hat{\alpha}_1 + (\lambda_e - \lambda_b)\,\hat{\alpha}_2.$$

Stacking the observations ($t = 1, \ldots, T$), we write the regression model to be estimated as

$$y = Z\theta + u,$$

where $y = (p_1, \dots, p_T)'$ and $u = (u_1, \dots, u_T)'$. Furthermore, $Z = (z_1, \dots, z_T)'$ with $z_t = (p_{t-1}, x_t, 1, d_t)'$ and $\theta = (\gamma, \beta, \alpha, \alpha_2)'$. The OLS estimator of the full parameter vector θ is equal to

$$\hat{\theta} = (Z'Z)^{-1}Z'y.$$

Taking the probability limit we have

$$\begin{aligned}\operatorname{plim}\hat{\theta} &= \theta + \left(\operatorname{plim}\frac{1}{T}Z'Z\right)^{-1}\operatorname{plim}\frac{1}{T}Z'u \\ &= \theta + \Sigma_{ZZ}^{-1}\Sigma_{Zu}. \qquad \text{(A.3)}\end{aligned}$$

The vector $\Sigma_{ZZ}^{-1}\Sigma_{Zu}$ is the OLS inconsistency.

Regarding Σ_{Zu} in (A.3) we have:

$$\begin{aligned}m_{xu} &= \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} x_t u_t = 0, \\ m_u &= \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} u_t = 0,\end{aligned}$$

where $u_t = \varepsilon_t - \alpha_2 v_t$. Also we have

$$E\left[d_t u_t\right] = -\alpha_2 E\left[d_t v_t\right] \neq 0,$$

hence $m_{du} \neq 0$. Also there is a non-zero correlation between p_{t-1} and u_t. Exploiting the stationarity assumption by repeated substitution we can write

$$p_t = \beta\sum_{s=0}^{\infty}\gamma^s x_{t-s} + \frac{\alpha}{1-\gamma} + \alpha_2\sum_{s=0}^{\infty}\gamma^s D_{t-s} + \sum_{s=0}^{\infty}\gamma^s\varepsilon_{t-s},$$

so we have (exploiting the fact that ε_t has no autocorrelation, and zero correlation with lagged x_{t-s-1} by Assumption 1) that

$$\begin{aligned}E\left[p_{t-1}u_t\right] &= -\alpha_2^2 E\left[v_t\sum_{s=0}^{\infty}\gamma^s D_{t-1-s}\right] \\ &= -\alpha_2^2\left(v_t D_{t-1} + \gamma v_t D_{t-2} + \gamma^2 v_t D_{t-3} + \dots\right) \\ &\neq 0,\end{aligned}$$

so $m_{p_{-1}u} \neq 0$. Under the assumptions, we have now:

$$\Sigma_{ZZ} = \begin{pmatrix} m_{pp} & m_{xp_{-1}} & m_p & m_{dp_{-1}} \\ m_{xp_{-1}} & m_{xx} & 0 & 0 \\ m_p & 0 & 1 & 0 \\ m_{dp_{-1}} & 0 & 0 & m_{dd}\end{pmatrix}, \qquad \Sigma_{Zu} = \begin{bmatrix} m_{p_{-1}u} \\ 0 \\ 0 \\ m_{du}\end{bmatrix}.$$

After some algebra we find for the inconsistency

$$\Sigma_{ZZ}^{-1}\Sigma_{Zu} = \frac{1}{\det\left(\Sigma_{ZZ}\right)} \left[\begin{array}{c} m_{xx}\left(m_{dd}m_{p_{-1}u} - m_{du}m_{dp_{-1}}\right) \\ -m_{xp_{-1}}\left(m_{dd}m_{p_{-1}u} - m_{du}m_{dp_{-1}}\right) \\ -m_{p}m_{xx}\left(m_{dd}m_{p_{-1}u} - m_{du}m_{dp_{-1}}\right) \\ -m_{dp_{-1}}m_{xx}m_{p_{-1}u} - m_{du}(\sigma_{xp_{-1}}^2 - \sigma_x^2\sigma_p^2) \end{array} \right], \tag{A.4}$$

where $\det\left(\Sigma_{ZZ}\right) > 0$ and we define $\sigma_p^2 = m_{pp} - m_p^2$. Furthermore, because we assume $m_x = 0$, we have $m_{xx} = \sigma_x^2$ and $m_{xp_{-1}} = \sigma_{xp_{-1}}$.

We now have to evaluate all separate terms in (A.4). We always have

$$\begin{aligned} m_{dd} &= \operatorname{plim}\frac{1}{T}\sum d_t^2 \\ &= \left(1 - \lambda_e + \lambda_b\right)\left(\lambda_e - \lambda_b\right). \end{aligned}$$

Furthermore, to evaluate $m_{p_{-1}u}$ we note that

$$\begin{aligned} m_{p_{-1}u} &= \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} p_{t-1}u_t \\ &= -\alpha_2^2 \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\left(v_tD_{t-1} + \gamma v_tD_{t-2} + \gamma^2 v_tD_{t-3} + \ldots\right) \\ &= -\alpha_2^2 \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\sum_{s=0}^{\infty}\gamma^s v_t D_{t-1-s} \\ &= -\alpha_2^2 \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\sum_{s=0}^{\infty}\gamma^s v_t\left(D_t + D_{t-1-s} - D_t\right) \\ &= -\frac{\alpha_2^2}{1-\gamma}\operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} v_tD_t + \alpha_2^2 \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\sum_{s=0}^{\infty}\gamma^s v_t\left(D_t - D_{t-1-s}\right) \\ &= -\frac{\alpha_2^2}{1-\gamma}\operatorname{plim}\frac{1}{T}\sum_{t=1}^{T} v_tD_t. \end{aligned}$$

The final equality holds because $v_t\left(D_t - D_{t-1-s}\right) \propto \frac{s+1}{T}$, because $D_t - D_{t-1-s}$ has nonzero values in (two times) $\frac{s+1}{T}$ observations only. Therefore,

$$\begin{aligned} \sum_{s=0}^{\infty}\gamma^s v_t\left(D_t - D_{t-1-s}\right) &\propto \sum_{s=0}^{\infty}\gamma^s\frac{s+1}{T} \\ &= \frac{1}{T}\sum_{s=0}^{\infty}\left(s+1\right)\gamma^s \\ &= \frac{1}{T\left(1-\gamma\right)^2}, \end{aligned}$$

which is of order $O(T^{-1})$ only.

For the same reason we have

$$\begin{aligned} m_{dp_{-1}} &= \operatorname{plim} \frac{1}{T} \alpha_2 \sum_{t=1}^{T} \sum_{s=0}^{\infty} \gamma^s d_t D_{t-1-s} \\ &= \alpha_2 \operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=0}^{\infty} \gamma^s d_t \left(D_t - (D_t - D_{t-1-s}) \right) \\ &= \frac{\alpha_2}{1-\gamma} \operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} d_t D_t. \end{aligned}$$

The precise magnitude of the separate terms will depend on the type of break misdating. The four possible misdating scenarios for the break dates are:

$$\begin{aligned} &\text{Case 1: } T_b < T_B,\ T_e < T_E \\ &\text{Case 2: } T_b < T_B,\ T_e > T_E \\ &\text{Case 3: } T_b > T_B,\ T_e < T_E \\ &\text{Case 4: } T_b > T_B,\ T_e > T_E \end{aligned} \tag{A.5}$$

In Case 1 the cartel is formally dated to begin and end too early—including a formal cartel period that entirely preceeds the cartel's effects. In Case 2 the formal cartel period encompasses the effective cartel period and includes non-cartel periods too. Case 3 is a legally too narrowly defined period, for which there is indication competition authorities conservatively do. Case 4 is the mirror image of Case 1. Compared to the misclassification set up introduced in (18), Case 2 is equal to the special situation that $\lambda_{01} = 0$, while Case 3 amounts to $\lambda_{10} = 0$. For Cases 1 and 4 both $\lambda_{01} > 0$ and $\lambda_{10} > 0$.

We will provide detailed derivations for Cases 1 and 2 only. In Case 1 ($T_b < T_B < T_e < T_E$) we have for the measurement error:

$$v_t = \begin{cases} \lambda_b - \lambda_e - \lambda_B + \lambda_E, & t \leq T_b, \\ 1 - \lambda_e + \lambda_b - \lambda_B + \lambda_E, & T_b < t \leq T_B, \\ \lambda_b - \lambda_e - \lambda_B + \lambda_E, & T_B < t \leq T_e, \\ \lambda_b - \lambda_e - 1 - \lambda_B + \lambda_E, & T_e < t \leq T_E, \\ \lambda_b - \lambda_e - \lambda_B + \lambda_E, & t > T_E. \end{cases}$$

We then have

$$
\begin{aligned}
m_{du} = & -\alpha_2 \operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} d_t v_t \\
= & -\alpha_2 \left[\lambda_b \left(\lambda_b - \lambda_e \right) \left(\lambda_b - \lambda_e - \lambda_B + \lambda_E \right) \right. \\
& + \left(\lambda_B - \lambda_b \right) \left(1 - \lambda_e + \lambda_b \right) \left(1 - \lambda_e + \lambda_b - \lambda_B + \lambda_E \right) \\
& + \left(\lambda_e - \lambda_B \right) \left(1 - \lambda_e + \lambda_b \right) \left(-\lambda_e + \lambda_b - \lambda_B + \lambda_E \right) \\
& + \left(\lambda_E - \lambda_e \right) \left(\lambda_b - \lambda_e \right) \left(-1 - \lambda_e + \lambda_b - \lambda_B + \lambda_E \right) \\
& \left. + \left(1 - \lambda_E \right) \left(\lambda_b - \lambda_e \right) \left(\lambda_b - \lambda_e - \lambda_B + \lambda_E \right) \right] \\
= & -\alpha_2 \left[\left(\lambda_B - \lambda_b \right) \left(1 - \lambda_e + \lambda_b \right) + \left(\lambda_e - \lambda_b \right) \left(\lambda_E - \lambda_e \right) \right].
\end{aligned}
$$

Furthermore, we have that

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} d_t D_t = & \lambda_b \left(\lambda_b - \lambda_e \right) \left(\lambda_B - \lambda_E \right) \\
& + \left(\lambda_B - \lambda_b \right) \left(1 - \lambda_e + \lambda_b \right) \left(\lambda_B - \lambda_E \right) \\
& + \left(\lambda_e - \lambda_B \right) \left(1 - \lambda_e + \lambda_b \right) \left(1 - \lambda_E + \lambda_B \right) \\
& + \left(\lambda_E - \lambda_e \right) \left(\lambda_b - \lambda_e \right) \left(1 - \lambda_E + \lambda_B \right) \\
& + \left(1 - \lambda_E \right) \left(\lambda_b - \lambda_e \right) \left(\lambda_B - \lambda_E \right) \\
= & \lambda_e - \lambda_B + \left(\lambda_e - \lambda_b \right) \left(\lambda_B - \lambda_E \right),
\end{aligned}
$$

hence we find that

$$
m_{dp_{-1}} = \frac{\alpha_2}{1 - \gamma} \left(\lambda_e - \lambda_B + \left(\lambda_e - \lambda_b \right) \left(\lambda_B - \lambda_E \right) \right).
$$

Together with

$$
\frac{1}{T} \sum_{t=1}^{T} D_t^2 = \left(1 - \lambda_E + \lambda_B \right) \left(\lambda_E - \lambda_B \right),
$$

we find

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} v_t D_t = & \frac{1}{T} \sum_{t=1}^{T} d_t D_t - \frac{1}{T} \sum_{t=1}^{T} D_t^2, \\
= & \lambda_e - \lambda_B + \left(1 - \lambda_E + \lambda_B + \lambda_e - \lambda_b \right) \left(\lambda_B - \lambda_E \right),
\end{aligned}
$$

and we therefore have

$$
m_{p_{-1}u} = -\frac{\alpha_2^2}{1 - \gamma} \left[\lambda_e - \lambda_B + \left(1 - \lambda_E + \lambda_B + \lambda_e - \lambda_b \right) \left(\lambda_B - \lambda_E \right) \right].
$$

Inspecting signs, we obviously have $m_{dd} > 0$. Furthermore, assuming the cartel effect $\alpha_2 > 0$ and given $T_b < T_B < T_e < T_E$ we find that $m_{du} < 0$ and $m_{dp_{-1}} > 0$ because

$$\begin{aligned}\lambda_e - \lambda_B + (\lambda_e - \lambda_b)(\lambda_B - \lambda_E) &< \lambda_E - \lambda_B + (\lambda_e - \lambda_b)(\lambda_B - \lambda_E) \\ &= (\lambda_E - \lambda_B)(1 + \lambda_b - \lambda_e) \\ &> 0.\end{aligned}$$

Also we find $m_{p_{-1}u} > 0$ because

$$\begin{aligned}&(\lambda_E - \lambda_B)(1 - \lambda_E + \lambda_B + \lambda_e - \lambda_b) - (\lambda_e - \lambda_B) \\ >&(\lambda_E - \lambda_B)(1 - \lambda_E + \lambda_B + \lambda_e - \lambda_B) - (\lambda_e - \lambda_B) \\ =&(1 - \lambda_E + \lambda_B)(\lambda_E - \lambda_e) \\ >&0,\end{aligned}$$

where the second line follows from the fact that $\lambda_B > \lambda_b$. Collecting terms we therefore find for Case 1 that

$$m_{dd}m_{p_{-1}u} - m_{du}m_{dp_{-1}} > 0.$$

Hence, we can write for the direction of the first three elements in the OLS inconsistency (A.4):

$$\begin{aligned}&\text{plim}\,\hat{\gamma} > \gamma, \\ &\text{plim}\,\hat{\beta} \lesseqgtr \beta, \quad \text{if } m_{xp_{-1}} \gtreqless 0, \\ &\text{plim}\,\hat{\alpha} < \alpha,\end{aligned}$$

where the last inequality holds as prices are positive, hence $m_{p_{-1}} > 0$. Finally, the last element in the inconsistency is

$$\begin{aligned}\text{plim}\,(\hat{\alpha}_2 - \alpha_2) &= \frac{-m_{dp_{-1}}m_{xx}m_{p_{-1}u} - m_{du}(\sigma^2_{xp_{-1}} - \sigma^2_x\sigma^2_p)}{\det(\Sigma_{ZZ})} \\ &= -\sigma^2_x\frac{m_{dp_{-1}}m_{p_{-1}u} - m_{du}\sigma^2_p(1 - \rho^2_{xp_{-1}})}{\det(\Sigma_{ZZ})} \\ &< 0,\end{aligned}$$

because $m_{dp_{-1}} > 0$, $m_{p_{-1}u} > 0$ and $m_{du} < 0$.

We find exactly the same qualitative results for the Cases 2, 3 and 4. Because Case 2 will turn out to be most important we briefly state the results for this case. In Case 2, i.e. $T_b < T_B$, $T_e > T_E$, we have for the measurement error:

$$v_t = \begin{cases}\lambda_b - \lambda_e - \lambda_B + \lambda_E, & t \leq T_b, \\ 1 - \lambda_e + \lambda_b - \lambda_B + \lambda_E, & T_b < t \leq T_B, \\ \lambda_b - \lambda_e - \lambda_B + \lambda_E, & T_B < t \leq T_E, \\ 1 - \lambda_e + \lambda_b - \lambda_B + \lambda_E, & T_E < t \leq T_e, \\ \lambda_b - \lambda_e - \lambda_B + \lambda_E, & t > T_e.\end{cases}$$

We then have:

$$
\begin{aligned}
m_{du} &= -\alpha_2 \left(1 - \lambda_e + \lambda_b\right) \left(\lambda_B - \lambda_E + \lambda_e - \lambda_b\right), \\
m_{dp_{-1}} &= \frac{\alpha_2}{1-\gamma} \left(\lambda_E - \lambda_B\right) \left(1 - \lambda_e + \lambda_b\right), \\
m_{p_{-1}u} &= -\frac{\alpha_2^2}{1-\gamma} \left(\lambda_E - \lambda_B\right) \left(\lambda_b - \lambda_e + \lambda_E - \lambda_B\right).
\end{aligned}
$$

Inspecting signs, we find that $m_{du} < 0$ as

$$
\begin{aligned}
\left(1 - \lambda_e + \lambda_b\right) \left(\lambda_B - \lambda_E + \lambda_e - \lambda_b\right) &= \left(1 - \left(\lambda_e - \lambda_b\right)\right) \left(\lambda_e - \lambda_b - \left(\lambda_E - \lambda_B\right)\right) \\
&> 0,
\end{aligned}
$$

because in Case 2 $\lambda_e - \lambda_b > \lambda_E - \lambda_B$. Furthermore, it is obvious that $m_{dp_{-1}} > 0$ and $m_{p_{-1}u} > 0$. Collecting terms we therefore find the same qualitative results compared with Case 1.

For the original intercept α_1, we obtain the following. Defining

$$
c = \frac{m_{dd} m_{p_{-1}u} - m_{du} m_{dp_{-1}}}{\det\left(\Sigma_{ZZ}\right)},
$$

we can write

$$
\begin{aligned}
\operatorname{plim}\left(\frac{\hat{\alpha}}{1-\hat{\gamma}} - \frac{\alpha}{1-\gamma}\right) &= \frac{\operatorname{plim}\left(\hat{\alpha} - \alpha\right) + \frac{\alpha}{1-\gamma} \operatorname{plim}\left(\hat{\gamma} - \gamma\right)}{1 - \gamma - \operatorname{plim}\left(\hat{\gamma} - \gamma\right)} \\
&= \frac{-m_p m_{xx} c + \frac{\alpha}{1-\gamma} m_{xx} c}{1 - \gamma - m_{xx} c} \\
&= \left(\frac{m_{xx} c}{1 - \gamma - m_{xx} c}\right) \left(\frac{\alpha}{1-\gamma} - m_p\right) \\
&= 0.
\end{aligned}
$$

It can also be shown that

$$
\left(\lambda_E - \lambda_B\right) \frac{\alpha_2}{1-\gamma} - \left(\lambda_e - \lambda_b\right) \operatorname{plim} \frac{\hat{\alpha}_2}{1-\hat{\gamma}} = \begin{cases} 0, & T_b < T_B, T_e > T_E, \\ > 0, & \text{otherwise.} \end{cases} \tag{A.6}
$$

Therefore, given that $\operatorname{plim} \hat{\gamma} > \gamma$ it is easily seen that for Case 2

$$
\begin{aligned}
\operatorname{plim} \hat{\alpha}_1 &= \operatorname{plim} \hat{\alpha} - \left(\lambda_e - \lambda_b\right) \operatorname{plim} \hat{\alpha}_2 \\
&= \alpha_1 \frac{\operatorname{plim}\left(1 - \hat{\gamma}\right)}{1-\gamma} \\
&< \alpha_1.
\end{aligned}
$$

Numerical simulations confirm that it holds for the other cases as well. □

Proof of Proposition 4.

(i) Recall that

$$t \leq T_b : \widehat{\text{bfp}}_t = p_t,$$
$$t > T_b : \widehat{\text{bfp}}_t = \hat{\gamma}\widehat{\text{bfp}}_{t-1} + \hat{\beta}x_t + \hat{\alpha}_1,$$

and

$$\text{bfp}_t = \gamma \text{bfp}_{t-1} + \beta x_t + \alpha_1 + \varepsilon_t,$$

where in terms of the parameter vector $\theta = (\gamma, \beta, \alpha, \alpha_1)'$, we have $\hat{\alpha}_1 = \hat{\alpha} - (\lambda_e - \lambda_b)\hat{\alpha}_2$ and $\alpha_1 = \alpha - (\lambda_E - \lambda_B)\alpha_2$. We have for the average estimated but-for price

$$\begin{aligned}
\text{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} \widehat{\text{bfp}}_t &= \text{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} \text{bfp}_t + \text{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} \left(\widehat{\text{bfp}}_t - \text{bfp}_t \right) \\
&= \text{plim} \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \text{bfp}_t + \text{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} \left(\widehat{\text{bfp}}_t - \text{bfp}_t \right).
\end{aligned}$$

We will analyze the limiting behavior of the second term in more detail. In period $T_b + 1$ we can write for the prediction error

$$\begin{aligned}
v_{T_b+1} &= \widehat{\text{bfp}}_{T_b+1} - \text{bfp}_{T_b+1} \\
&= (\hat{\gamma} - \gamma) \text{bfp}_{T_b} + \hat{\gamma} \left(p_{T_b} - \text{bfp}_{T_b} \right) + (\hat{\beta} - \beta) x_{T_b+1} + \hat{\alpha} - \alpha - \varepsilon_{T_b+1}.
\end{aligned}$$

Note that standard asymptotic theory for method of moments estimators gives the following large sample distribution of the OLS estimator

$$\sqrt{T} \left(\hat{\theta} - \theta - \theta^* \right) \xrightarrow{d} \mathcal{N}(0, V),$$

where $\theta^* = \Sigma_{ZZ}^{-1} \Sigma_{Zu}$ is the inconsistency. Therefore, we can write $\hat{\theta} - \theta = \theta^* + O_P(T^{-1/2})$. The implied inconsistency in α_1 is

$$\alpha_1^* = \alpha^* - (\lambda_e - \lambda_b)\alpha_2^* + (\lambda_E - \lambda_B - (\lambda_e - \lambda_b))\alpha_2.$$

Also note that we have:

$$\text{bfp}_{T_b} = \begin{cases} p_{T_b}, & T_b \leq T_B, \\ < p_{T_b}, & T_b > T_B, \end{cases}$$

so that:

$$\hat{\gamma} \left(p_{T_b} - \text{bfp}_{T_b} \right) = \begin{cases} 0, & T_b \leq T_B, \\ (\gamma + \gamma^*) \left(p_{T_b} - \text{bfp}_{T_b} \right) + O_P(T^{-1/2}), & T_b > T_B. \end{cases}$$

Therefore we can write:

$$v_{T_b+1} = \begin{cases} \gamma^* \text{bfp}_{T_b} + \beta^* x_{T_b+1} + \alpha_1^* - \varepsilon_{T_b+1} + O_P(T^{-1/2}), & T_b \leq T_B, \\ \gamma^* \text{bfp}_{T_b} + \beta^* x_{T_b+1} + \alpha_1^* - \varepsilon_{T_b+1} + O_P(1) & T_b > T_B. \end{cases}$$

In period $T_b + 2$ we have for the difference in estimated and true but-for prices

$$\begin{aligned} \widehat{\text{bfp}}_{T_b+2} - \text{bfp}_{T_b+2} &= \hat{\gamma}\widehat{\text{bfp}}_{T_b+1} - \gamma \text{bfp}_{T_b+1} + (\hat{\beta} - \beta) x_{T_b+2} + \hat{\alpha}_1 - \alpha_1 - \varepsilon_{T_b+2} \\ &= \hat{\gamma} v_{T_b+1} + (\hat{\gamma} - \gamma) \, \text{bfp}_{T_b+1} + (\hat{\beta} - \beta) x_{T_b+2} + \hat{\alpha}_1 - \alpha_1 - \varepsilon_{T_b+2} \\ &= \hat{\gamma} v_{T_b+1} + v_{T_b+2}. \end{aligned}$$

Regarding the prediction error v_{T_b+2} we find

$$v_{T_b+2} = \gamma^* \text{bfp}_{T_b+1} + \beta^* x_{T_b+2} + \alpha_1^* - \varepsilon_{T_b+2} + O_P(T^{-1/2}),$$

irrespective of the precise timing of T_b and T_B. In general we find

$$\begin{aligned} \widehat{\text{bfp}}_{T_b+s} - \text{bfp}_{T_b+s} &= \hat{\gamma}^{s-1} v_{T_b+1} + \hat{\gamma}^{s-2} v_{T_b+2} + \ldots + \hat{\gamma} v_{T_b+s-1} + v_{T_b+s} \\ &= \sum_{j=0}^{s-1} \hat{\gamma}^j v_{T_b+s-j} \\ &= \sum_{j=0}^{s-1} (\gamma + \gamma^*)^j \, v_{T_b+s-j} + O_P(T^{-1/2}), \end{aligned}$$

where

$$v_{T_b+s-j} = \gamma^* \text{bfp}_{T_b+s-j-1} + \beta^* x_{T_b+s-j} + \alpha_1^* - \varepsilon_{T_b+s-j} + O_P(T^{-1/2}), \quad j = 0, \ldots, s-2.$$

By repeated substitution we can write

$$\text{bfp}_t = \frac{\alpha_1}{1 - \gamma} + \beta \sum_{i=0}^{\infty} \gamma^i x_{t-i} + \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-i},$$

hence

$$E\left[\text{bfp}_t\right] = \frac{\alpha_1}{1 - \gamma},$$

because $E\left[x_t\right] = 0$ and $E\left[\varepsilon_t\right] = 0$. Defining

$$c = \frac{m_{dd} m_{p_{-1}u} - m_{du} m_{dp_{-1}}}{\det\left(\Sigma_{ZZ}\right)},$$

we can write

$$\begin{aligned} \text{plim}\left(\frac{\hat{\alpha}}{1 - \hat{\gamma}} - \frac{\alpha}{1 - \gamma}\right) &= \frac{\text{plim}\left(\hat{\alpha} - \alpha\right) + \alpha \, \text{plim}\left(\hat{\gamma} - \gamma\right)/(1 - \gamma)}{1 - \gamma - \text{plim}\left(\hat{\gamma} - \gamma\right)} \\ &= \frac{-m_p m_{xx} c + m_{xx} c \alpha/(1 - \gamma)}{1 - \gamma - m_{xx} c} \\ &= 0, \end{aligned}$$

because $m_p = \alpha/(1-\gamma)$.

Using $\gamma^* = m_{xx}c$ and $\alpha^* = -m_p m_{xx} c$ once more we get

$$\begin{aligned}
E\left[v_{T_b+s-j}\right] &= E\left[\gamma^* \mathrm{bfp}_{T_b+s-j-1} + \beta^* x_{T_b+s-j} + \alpha_1^* - \varepsilon_{T_b+s-j}\right] + O(T^{-1}) \\
&= m_{xx}c\left(\frac{\alpha}{1-\gamma} - (\lambda_E - \lambda_B)\frac{\alpha_2}{1-\gamma}\right) - m_p m_{xx} c \\
&\quad - \left((\lambda_e - \lambda_b)\,\mathrm{plim}\,\hat{\alpha}_2 - (\lambda_E - \lambda_B)\,\alpha_2\right) + O(T^{-1}) \\
&= -\,\mathrm{plim}\,(\hat{\gamma} - \gamma)\,(\lambda_E - \lambda_B)\,\frac{\alpha_2}{1-\gamma} - (\lambda_e - \lambda_b)\,\mathrm{plim}\,\hat{\alpha}_2 \\
&\quad + (\lambda_E - \lambda_B)\,\alpha_2 + O(T^{-1}) \\
&= (\lambda_E - \lambda_B)\,\frac{\alpha_2}{1-\gamma}(1 - \mathrm{plim}\,\hat{\gamma}) - (\lambda_e - \lambda_b)\,\mathrm{plim}\,\hat{\alpha}_2.
\end{aligned}$$

From (A.6) we see that:

$$\mathrm{plim}\,\hat{\alpha}_2 = \begin{cases} \dfrac{\lambda_E - \lambda_B}{\lambda_e - \lambda_b}\dfrac{\alpha_2}{1-\gamma}(1 - \mathrm{plim}\,\hat{\gamma}), & T_b < T_B, T_e > T_E, \\ < \dfrac{\lambda_E - \lambda_B}{\lambda_e - \lambda_b}\dfrac{\alpha_2}{1-\gamma}(1 - \mathrm{plim}\,\hat{\gamma}), & \text{otherwise.} \end{cases} \tag{A.7}$$

Using this we find:

$$E\left[v_{T_b+s-j}\right] = \begin{cases} O(T^{-1}), & T_b < T_B, T_e > T_E, \\ O(1) > 0, & \text{otherwise.} \end{cases}$$

The difference in estimated and true but-for price is

$$\widehat{\mathrm{bfp}}_{T_b+s} - \mathrm{bfp}_{T_b+s} = \sum_{j=0}^{s-1} (\gamma + \gamma^*)^j\, v_{T_b+s-j} + O_P(T^{-1/2}),$$

with:

$$E\left[\sum_{j=0}^{s-1} (\gamma + \gamma^*)^j\, v_{T_b+s-j}\right] = \begin{cases} O(T^{-1}), & T_b < T_B, T_e > T_E, \\ O(1) > 0, & \text{otherwise.} \end{cases}$$

Hence

$$\begin{aligned}
&\mathrm{plim}\,\frac{1}{T_e - T_b}\sum_{s=1}^{T_e - T_b}\left(\widehat{\mathrm{bfp}}_{T_b+s} - \mathrm{bfp}_{T_b+s}\right) \\
=\ &\mathrm{plim}\,\frac{1}{T_e - T_b}\sum_{s=1}^{T_e - T_b}\sum_{j=0}^{s-1} (\gamma + \gamma^*)^j\, v_{T_b+s-j} = \begin{cases} 0, & T_b < T_B, T_e > T_E, \\ > 0, & \text{otherwise,} \end{cases}
\end{aligned}$$

which completes the proof of (i).

(ii) Noting that $\mathrm{bfp}_{T_B} = p_{T_B}$ we have for the overcharge

$$O_{T_B+t} = \left(p_{T_B+t} - \mathrm{bfp}_{T_B+t}\right) D_t = \frac{1-\gamma^t}{1-\gamma}\alpha_2 D_t,$$

which for $t \to \infty$ simplifies to $\alpha_2/(1-\gamma)$. Therefore, we have for the average effective overcharge

$$\begin{aligned}
\operatorname{plim} \bar{O} &= \operatorname{plim} \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} (p_t - \mathrm{bfp}_t) \\
&= \lim \frac{1}{T_E - T_B} \sum_{t=T_B+1}^{T_E} \frac{1-\gamma^t}{1-\gamma} \alpha_2 \\
&= \frac{\alpha_2}{1-\gamma}.
\end{aligned}$$

We have for the average estimated overcharge

$$\begin{aligned}
\operatorname{plim} \overline{O}_2 &= \operatorname{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} \left(p_t - \widehat{\mathrm{bfp}}_t\right) \\
&= \operatorname{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} (p_t - \mathrm{bfp}_t) - \operatorname{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} \left(\widehat{\mathrm{bfp}}_t - \mathrm{bfp}_t\right).
\end{aligned}$$

Using Proposition 4 (i) we find that

$$\begin{aligned}
\operatorname{plim} \overline{O}_2 &= \operatorname{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} (p_t - \mathrm{bfp}_t), \quad T_b < T_B, T_e > T_E, \\
&< \operatorname{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} (p_t - \mathrm{bfp}_t), \quad \text{otherwise.}
\end{aligned}$$

Combining this with the fact that

$$\begin{aligned}
\operatorname{plim} \frac{1}{T_e - T_b} \sum_{t=T_b+1}^{T_e} (p_t - \mathrm{bfp}_t) &= \frac{\alpha_2}{1-\gamma} \quad T_b > T_B, T_e < T_E, \\
&< \frac{\alpha_2}{1-\gamma}, \quad \text{otherwise,}
\end{aligned}$$

completes the proof. □

Proof of Theorem 3. We have for the effective damage

$$\begin{aligned}
\operatorname{plim} \frac{1}{T} \mathrm{CD} &= \operatorname{plim} \frac{1}{T} \sum_{t=T_B+1}^{T_E} (p_t - \mathrm{bfp}_t) Q_t \\
&= \operatorname{plim} \frac{1}{T} \sum_{t=T_B+1}^{T_E} \frac{1-\gamma^t}{1-\gamma} \alpha_2 Q_t \\
&= \lim \frac{T_E - T_B}{T} \alpha_2 Q_C \lim \frac{1}{T} \sum_{t=T_B+1}^{T_E} \frac{1-\gamma^t}{1-\gamma} \\
&= \frac{\alpha_2}{1-\gamma} (\lambda_E - \lambda_B) Q_C.
\end{aligned}$$

We have for the estimated damage

$$
\begin{aligned}
\operatorname{plim} \frac{1}{T}\widehat{\mathrm{CD}}_2 &= \operatorname{plim} \frac{1}{T} \sum_{t=T_b+1}^{T_e} \left(p_t - \widehat{\mathrm{bfp}}_t\right) Q_t \\
&= \operatorname{plim} \frac{1}{T} \sum_{t=T_b+1}^{T_e} \left(p_t - \mathrm{bfp}_t\right) Q_t + \operatorname{plim} \frac{1}{T} \sum_{t=T_b+1}^{T_e} \left(\mathrm{bfp}_t - \widehat{\mathrm{bfp}}_t\right) Q_t \\
&= \operatorname{plim} \frac{1}{T}\mathrm{CD} - \operatorname{plim} \frac{1}{T} \sum_{t=T_b+1}^{T_e} \left(\widehat{\mathrm{bfp}}_t - \mathrm{bfp}_t\right) Q_t.
\end{aligned}
$$

We will analyze the limiting behavior of the second term in more detail under Assumption 1. Proposition 4 (i) implies that

$$
\left(\widehat{\mathrm{bfp}}_{T_b+s} - \mathrm{bfp}_{T_b+s}\right) Q_{T_b+s} = Q_{T_b+s} \sum_{j=0}^{s-1} \left(\gamma + \gamma^*\right)^j v_{T_b+s-j} + O_P(T^{-1/2}),
$$

where

$$
v_{T_b+s-j} = \gamma^* \mathrm{bfp}_{T_b+s-j-1} + \beta^* x_{T_b+s-j} + \alpha_1^* - \varepsilon_{T_b+s-j} + O_P(T^{-1/2}), \quad j = 0, \ldots, s-2.
$$

We note that

$$
\mathrm{bfp}_{t-1} = \frac{\alpha_1}{1-\gamma} + \beta \sum_{s=0}^{\infty} \gamma^s x_{t-1-s} + \sum_{s=0}^{\infty} \gamma^s \varepsilon_{t-1-s}.
$$

Due to $m_x = 0$, we can define the linear projection coefficients of a regression of bfp_{t-1} on a constant and x_t as m_p and $\frac{m_{xp_{-1}}}{m_{xx}}$ respectively. Therefore, we can write

$$
\begin{aligned}
\mathrm{bfp}_{t-1} &= m_p + \frac{m_{xp_{-1}}}{m_{xx}} x_t + \eta_t, \\
\eta_t &= \beta \sum_{s=0}^{\infty} \gamma^s x_{t-1-s} - \frac{m_{xp_{-1}}}{m_{xx}} x_t + \sum_{s=0}^{\infty} \gamma^s \varepsilon_{t-1-s},
\end{aligned}
$$

where by construction $E\left[\eta_t x_t\right] = 0$. This leads to

$$
E\left[\mathrm{bfp}_{T_b+s-j-1} Q_{T_b+s}\right] = \frac{m_{xp_{-1}}}{m_{xx}} E\left[x_{T_b+s-j} Q_{T_b+s}\right] + \frac{\alpha_1}{1-\gamma} E\left[Q_{T_b+s}\right],
$$

where we assumed that the projection error η_t is uncorrelated with Q_{T_b+s}.

Using $\gamma^* = m_{xx}c$, $\beta^* = -m_{xp_{-1}}c$ and $\alpha^* = -m_p m_{xx} c$ we get

$$
\begin{aligned}
E\left[v_{T_b+s-j} Q_{T_b+s}\right] &= E\left[\left(\gamma^* \mathrm{bfp}_{T_b+s-j-1} + \beta^* x_{T_b+s-j} + \alpha_1^* - \varepsilon_{T_b+s-j}\right) Q_{T_b+s}\right] + O(T^{-1}) \\
&= m_{xx} c \frac{m_{xp_{-1}}}{m_{xx}} E\left[x_{T_b+s-j} Q_{T_b+s}\right] \\
&\quad + m_{xx} c \left(\frac{\alpha}{1-\gamma} - \left(\lambda_E - \lambda_B\right) \frac{\alpha_2}{1-\gamma}\right) E\left[Q_{T_b+s}\right] \\
&\quad - m_{xp_{-1}} c E\left[x_{T_b+s-j} Q_{T_b+s}\right] \\
&\quad - m_p m_{xx} c E\left[Q_{T_b+s}\right] \\
&\quad - \left(\left(\lambda_e - \lambda_b\right) \operatorname{plim} \hat{\alpha}_2 - \left(\lambda_E - \lambda_B\right) \alpha_2\right) E\left[Q_{T_b+s}\right] \\
&\quad - E\left[\varepsilon_{T_b+s-j} Q_{T_b+s}\right] + O(T^{-1}).
\end{aligned}
$$

Collecting terms, this leads to

$$E\left[v_{T_b+s-j}Q_{T_b+s}\right] = E\left[Q_{T_b+s}\right]\left\{(\lambda_E-\lambda_B)\frac{\alpha_2}{1-\gamma}(1-\operatorname{plim}\hat{\gamma})-(\lambda_e-\lambda_b)\operatorname{plim}\hat{\alpha}_2\right\}$$
$$- E\left[\varepsilon_{T_b+s-j}Q_{T_b+s}\right]+O(T^{-1}).$$

Using (A.7) we find:

$$E\left[v_{T_b+s-j}Q_{T_b+s}\right] = \begin{cases} -E\left[\varepsilon_{T_b+s-j}Q_{T_b+s}\right]+O(T^{-1}), & T_b<T_B, T_e>T_E, \\ -E\left[\varepsilon_{T_b+s-j}Q_{T_b+s}\right]+O(1), & \text{otherwise,} \end{cases}$$

where the $O(1)$ remainder term is positive. The difference in estimated and true damage per period is

$$\left(\widehat{\mathrm{bfp}}_{T_b+s}-\mathrm{bfp}_{T_b+s}\right)Q_{T_b+s} = \sum_{j=0}^{s-1}(\gamma+\gamma^*)^j\, v_{T_b+s-j}Q_{T_b+s}+O_P(T^{-1/2}),$$

hence

$$\begin{aligned} &\operatorname{plim}\frac{1}{T_e-T_b}\sum_{s=1}^{T_e-T_b}\left(\widehat{\mathrm{bfp}}_{T_b+s}-\mathrm{bfp}_{T_b+s}\right)Q_{T_b+s} \\ =&\operatorname{plim}\frac{1}{T_e-T_b}\sum_{s=1}^{T_e-T_b}\sum_{j=0}^{s-1}(\gamma+\gamma^*)^j\, v_{T_b+s-j}Q_{T_b+s} \\ =&-\sum_{j=0}^{s-1}(\gamma+\gamma^*)^j\, E\left[\varepsilon_{T_b+s-j}Q_{T_b+s}\right] \\ \geq&0, \end{aligned}$$

when $T_b<T_B, T_e>T_E$ and under Assumption 2. In the other cases a strict inequality sign holds. This shows that:

$$\begin{aligned} \operatorname{plim}\frac{1}{T}\widehat{\mathrm{CD}}_2 &\leq \operatorname{plim}\frac{1}{T}\mathrm{CD}, \quad T_b<T_B, T_e>T_E \\ &< \operatorname{plim}\frac{1}{T}\mathrm{CD}, \quad \text{otherwise,} \end{aligned}$$

which completes the proof. □

Appendix B: Simulation Results

For the Monte Carlo study, data have been generated according to DGP (3) with $\varepsilon_t \sim$ i.i.d. $N(0, \sigma_\varepsilon^2)$. Explanatory variables are lagged prices p_{t-1} and, for simplicity, a single cost factor x_t. The cartel dummy is defined as in (2) and set at $T_B = \frac{1}{3}T$ and $T_E = \frac{2}{3}T$. The explanatory variable x_t follows an AR(1) model

$$x_t = \rho x_{t-1} + v_t, \tag{A.8}$$

where $v_t \sim$ i.i.d. $N(0, \sigma_v^2)$ independent of ε_t, i.e. the cost factor is assumed to be strictly exogenous.

In order to investigate the actual size of various structural break inference procedures, data have been generated under the null hypothesis $H_0 : \alpha_2 = 0$. All experiments have a sample of $T = 100$ observations and the number of replications is 5,000. Without loss of generality we set $\alpha_1 = 100(1 - \gamma)$, so that the average price level in the simulations is 100. We normalized with respect to the variance of the disturbance term σ_ε^2. We furthermore choose $\gamma \in \{0.1, 0.5, 0.9\}$ and $\rho \in \{0.1, 0.5, 0.9\}$. These values roughly correspond to a low, intermediate and high degree of serial correlation in the time series p_t and x_t.

To facilitate the comparison of simulation results across experiments, some important design parameters are held fixed. We always set β such that the long-run effect of x on p is unity, i.e. we specify $\beta = 1 - \gamma$. Furthermore, σ_ε^2 is chosen such that the signal-to-noise ratio of the model, defined as

$$\text{SNR} = \frac{\text{Var}(p_t - \varepsilon_t)}{\text{Var}(\varepsilon_t)}, \tag{A.9}$$

does not change between experiments. Assuming $\sigma_\varepsilon^2 = 1$ for DGP (3), Kiviet (1995) derives the following relation between σ_v^2 and SNR and other model parameters

$$\sigma_v^2 = \frac{1}{\beta^2}\left[\text{SNR} - \frac{\gamma^2}{1-\gamma^2}\right]\frac{(1-\gamma^2)(1-\rho^2)(1-\gamma\rho)}{1+\gamma\rho}. \tag{A.10}$$

We choose SNR = 9 across experiments corresponding with a population R^2 of 0.9.

The correct dynamic specification (3) has been estimated in all experiments, hence OLS estimators are consistent, and analyze actual rejection probabilities of sup F-tests and double maximum (UD and WD) tests proposed by Bai and Perron (1998, 2003, 2006).[2] The nominal significance level in the simulations is 5% always. For the sup F-tests, the null hypothesis is no break versus k breaks where we experimented with $k = \{1, 2, 3\}$. The finite sample properties of the various test procedures have been investigated for trimming parameter $\mu = h/T = 0.15$, which implies that a minimum of 15 observations is used in any partition of the data, given that $T = 100$.[3]

[2]We used EViews 9 for all calculations.

[3]We obtained similar size and power properties result for $\mu = 0.05$.

The bootstrap version of the various testing procedures uses a standard non-parametric resampling scheme. First, we obtained the OLS estimator allowing for breaks. Second, a random sample is taken from the empirical distribution of the OLS residuals. Third, the bootstrapped dependent variable is calculated according to equation (3). In the bootstrap scheme, we kept the values of exogenous regressors as before. For the lagged dependent variable regressor, the first observation on the dependent variable is kept also as before and pseudo values for the remaining observations are constructed iteratively. Fourth, we estimated the model and calculated the various test statistics from the resampled data.

Table B.1: Size of nominal 5% sup F- and double max tests

γ	ρ	sup $F(1)$	sup $F(2)$	sup $F(3)$	UD max	WD max
			asymptotic tests			
0.1	0.1	0.054	0.048	0.046	0.054	0.049
0.1	0.5	0.052	0.051	0.050	0.053	0.054
0.1	0.9	0.059	0.060	0.061	0.060	0.061
0.5	0.1	0.057	0.054	0.056	0.058	0.059
0.5	0.5	0.058	0.059	0.063	0.060	0.064
0.5	0.9	0.067	0.074	0.091	0.074	0.092
0.9	0.1	0.136	0.211	0.288	0.187	0.288
0.9	0.5	0.140	0.232	0.324	0.204	0.318
0.9	0.9	0.195	0.322	0.458	0.301	0.484
			bootstrap tests			
0.1	0.1	0.056	0.047	0.041	0.051	0.041
0.1	0.5	0.054	0.048	0.046	0.051	0.043
0.1	0.9	0.053	0.050	0.045	0.053	0.042
0.5	0.1	0.053	0.048	0.042	0.050	0.042
0.5	0.5	0.053	0.054	0.049	0.053	0.047
0.5	0.9	0.052	0.050	0.047	0.050	0.040
0.9	0.1	0.056	0.052	0.051	0.055	0.044
0.9	0.5	0.055	0.053	0.051	0.054	0.045
0.9	0.9	0.057	0.057	0.057	0.056	0.044

Note. The table displays empirical rejection frequencies under the null hypothesis $\alpha_2 = 0$, using a nominal 5% significance level, of the sup F- and double max tests for structural breaks. The top panel provides rejection frequencies of the tests using asymptotic critical values, the bottom panel gives the results for the bootstrap implementation. The sample size is $T = 100$, and the trimming parameter is $\mu = 0.15$.

Repeating steps 2 to 4 of the resampling scheme B times, together with the calculated test statistic on the original data $(B + 1)$ generated realizations of the test statistic. A size-corrected test was constructed with these $(B + 1)$ realizations by using the appropriate quantile of the bootstrap distribution as critical value. The number of bootstrap replications $B = 199$.

Table B.1 shows actual rejection probabilities of asymptotic and bootstrap sup F-tests and double maximum tests. Autoregressive dynamics are varied in both p_t and x_t to analyze their relevance for the accuracy of asymptotic approximations. Asymptotic tests performed well under the null hypothesis when there is only small or moderate autoregressive dynamics in the dependent variable, i.e. $\gamma \in \{0.1, 0.5\}$. Rejection frequencies do not exceed 0.10, which is reasonably close to the nominal level of 0.05. However, these tests become oversized when $\gamma = 0.9$. In this case, actual rejection frequencies are also increasing with the persistence in the regressor x. In addition, size distortions of sup F-tests increase with the number of breaks specified under H_1. In general, size distortions of asymptotic tests are largest when both p and x are highly serially correlated ($\gamma = \rho = 0.9$), corroborating the simulation results for univariate models.[4] In contrast, bootstrap tests are always size correct, irrespective of the true value of the autoregressive dynamics. Therefore, in practice one should favour the bootstrap version of the break tests always.

We also analyzed power of the various asymptotic and bootstrap break tests. Data have been generated as before, but now under the alternative hypothesis $H_1 : \alpha_2 \neq 0$. A 20% overcharge, i.e. $\alpha_2 = 0.2\alpha_1$, is in line with findings reported in Connor and Lande (2008) and the European Commission's 2013 *Practical Guide*. Table B.2 shows size-corrected power of asymptotic and bootstrap sup F-tests and double maximum tests. The power properties of all tests are satisfactory, but power decreases when persistence in the data becomes larger, as expected. Generally, for this DGP with two structural breaks the power of the sup $F(1)$ test for a single break is lower than the power of tests for multiple breaks. This corroborates the simulation evidence of Bai and Perron (2006), who find that an underspecification of the number of breaks leads to a power loss in finite samples.

To analyze the sensitivity of these results, in Table B.3 we report power for a lower bound 10% overcharge, i.e. $\alpha_2 = 0.1\alpha_1$. Power is still satisfactory for small to moderate values of γ, but in case of highly persistent time series a deterioration of rejection frequencies can be seen compared to the case of a 20% overcharge. This is as expected, because with a relatively low overcharge it is more difficult to distinguish between serial correlation and structural change.

Finally, we analyze the finite sample accuracy of OLS inference conditional on the estimated break dates. We assume throughout that the correct number of breaks has been

[4] Diebold and Chen (1996), O'Reilly and Whelan (2005). Unreported simulation results show the same pattern for asymptotic sequential sup F tests.

Table B.2: Size-corrected power, 20% overcharge

γ	ρ	$\sup F(1)$	$\sup F(2)$	$\sup F(3)$	UD max	WD max
			asymptotic tests			
0.1	0.1	0.544	1.000	1.000	1.000	1.000
0.1	0.5	0.484	1.000	1.000	1.000	1.000
0.1	0.9	0.296	1.000	1.000	1.000	1.000
0.5	0.1	0.699	1.000	1.000	1.000	1.000
0.5	0.5	0.662	1.000	1.000	1.000	1.000
0.5	0.9	0.537	1.000	0.999	1.000	1.000
0.9	0.1	0.647	0.999	0.998	0.997	0.998
0.9	0.5	0.628	0.997	0.996	0.995	0.995
0.9	0.9	0.473	0.947	0.956	0.951	0.969
			bootstrap tests			
0.1	0.1	0.419	1.000	1.000	1.000	1.000
0.1	0.5	0.347	1.000	1.000	1.000	1.000
0.1	0.9	0.208	1.000	1.000	1.000	1.000
0.5	0.1	0.600	1.000	1.000	1.000	1.000
0.5	0.5	0.566	1.000	1.000	1.000	1.000
0.5	0.9	0.404	1.000	0.999	1.000	1.000
0.9	0.1	0.546	0.999	0.993	0.991	0.991
0.9	0.5	0.509	0.994	0.985	0.984	0.981
0.9	0.9	0.364	0.928	0.924	0.904	0.916

Note. The table displays empirical rejection frequencies under the alternative hypothesis with $\alpha_2 = 0.2\alpha_1$, using 5% critical values, of the structural break tests analyzed in Table B.1.

estimated, which in the experiments is always equal to two breaks. In each replication of an experiment, we base the definition of the cartel dummy on the estimated break dates. In Tables B.4 and B.5 we report the bias and Monte Carlo standard deviation of the OLS estimator of the cartel effect α_2. Furthermore, we show the actual rejection frequency of the corresponding nominal 5% t-test of $H_0 : \alpha_2 = \alpha_{2,0}$ with $\alpha_{2,0}$ the true value.

In case of a 20% overcharge (Table B.4) the bias of the coefficient estimator is of moderate magnitude and is small compared to the Monte Carlo standard deviation. Note that the bias of the OLS estimator of α_2 can have either sign depending on the parameters of the model. Actual rejection frequencies indicate that OLS inference is somewhat conservative for moderate values of γ, while overrejection is seen for experiments with high γ. In general, rejection frequencies are reasonably close to the nominal level of 0.05. Regarding a

Table B.3: Size-corrected power, 10% overcharge

γ	ρ	$\sup F(1)$	$\sup F(2)$	$\sup F(3)$	UD max	WD max
			asymptotic tests			
0.1	0.1	0.990	1.000	1.000	1.000	1.000
0.1	0.5	0.947	1.000	1.000	1.000	1.000
0.1	0.9	0.467	1.000	1.000	1.000	1.000
0.5	0.1	0.909	1.000	1.000	1.000	1.000
0.5	0.5	0.847	1.000	1.000	1.000	1.000
0.5	0.9	0.486	0.999	0.999	1.000	1.000
0.9	0.1	0.100	0.701	0.627	0.488	0.592
0.9	0.5	0.095	0.659	0.587	0.459	0.544
0.9	0.9	0.077	0.498	0.421	0.315	0.414
			bootstrap tests			
0.1	0.1	0.950	1.000	1.000	1.000	1.000
0.1	0.5	0.837	1.000	1.000	1.000	1.000
0.1	0.9	0.355	1.000	1.000	1.000	1.000
0.5	0.1	0.827	1.000	1.000	1.000	1.000
0.5	0.5	0.751	1.000	1.000	1.000	1.000
0.5	0.9	0.366	0.998	0.999	1.000	1.000
0.9	0.1	0.066	0.573	0.476	0.423	0.287
0.9	0.5	0.055	0.526	0.345	0.297	0.184
0.9	0.9	0.043	0.356	0.393	0.346	0.223

Note. The table displays empirical rejection frequencies under the alternative hypothesis with $\alpha_2 = 0.1\alpha_1$, using 5% critical values, of the structural break tests analyzed in Table B.1.

10% overcharge (Table B.5), finite sample accuracy deteriorates as expected. We conclude that, at least in this set of experiments, the finite sample properties of OLS inference are satisfactory.

References

Connor, J. M., & Lande, R. H. (2008). Cartel Overcharges and optimal cartel fines. In Waller, S. W. (Ed.). *Issues in Competition Law and Policy, Vol. 3, ABA Section of Antitrust Law*, 2203–2218.

Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics, 68*, 53–78.

Table B.4: OLS inference, 20% overcharge

γ	ρ	α_2	bias $\hat{\alpha}_2$	sd $\hat{\alpha}_2$	rf $t_{\hat{\alpha}_2}$
0.1	0.1	18	−0.329	0.981	0.027
0.1	0.5	18	−0.298	1.047	0.026
0.1	0.9	18	−0.165	1.258	0.010
0.5	0.1	10	−0.099	0.645	0.031
0.5	0.5	10	−0.080	0.673	0.029
0.5	0.9	10	−0.054	1.128	0.021
0.9	0.1	2	0.143	0.369	0.080
0.9	0.5	2	0.122	0.452	0.089
0.9	0.9	2	−0.088	0.907	0.157

Note. The table displays the Monte Carlo bias, standard dervi-ation (sd) and t-test rejection frequencies (rf) of the estimator of α_2 based on estimated cartel dates. The nominal size of the t-test is 5%, and the true value of the overcharge satisfies $\alpha_2 = 0.2\alpha_1$. The sample size is $T = 100$, and the trimming parameter is $\mu = 0.15$.

Table B.5: OLS inference, 10% overcharge

γ	ρ	α_2	bias $\hat{\alpha}_2$	sd $\hat{\alpha}_2$	rf $t_{\hat{\alpha}_2}$
0.1	0.1	9	−0.217	0.446	0.052
0.1	0.5	9	−0.206	0.487	0.051
0.1	0.9	9	−0.183	0.984	0.042
0.5	0.1	5	−0.082	0.384	0.052
0.5	0.5	5	−0.068	0.408	0.049
0.5	0.9	5	−0.075	0.865	0.059
0.9	0.1	1	0.277	0.371	0.146
0.9	0.5	1	0.270	0.410	0.155
0.9	0.9	1	0.238	0.559	0.191

Note. The table displays the same results as Table B.4, but with $\alpha_2 = 0.1\alpha_1$.