Does the Potential to Merge Reduce Competition?*

Dirk Hack Barth† Bart Taub‡

May 25, 2018

Abstract

We study anti-competitive mergers in a dynamic model with noisy collusion. At each instant, firms either privately choose output levels or merge to form a monopoly, trading off the benefits of avoiding price wars against the costs of merging. The potential to merge decreases pre-merger collusion, as punishments effected by price wars are weakened. Thus, although anti-competitive mergers harm competition ex-post, the implication is that barriers and costs of merging due to regulation should be reduced to promote competition ex-ante.

JEL Classification Numbers: D43, L12, L13, G34.

Keywords: Competition, imperfect information, industry structure, market power, mergers.

---

*This research was carried out in part during Taub’s stay at ICEF, Higher School of Economics, Moscow, and he acknowledges gratefully financial support from the Academic Excellence Project “5-100” of the Russian Government.

†Boston University, CEPR, and ECGI.

‡Glasgow University.
1 Introduction

Without appealing to technological complementarities between merging firms, we demonstrate that mergers can be welfare-improving by increasing the competitiveness of the merging firms prior to the merger. This suggests as a policy conclusion that barriers and costs of merging should be reduced. We achieve this result by considering a fully fleshed out dynamic model of tacit collusion between the two firms prior to their merger, in which the barriers and costs of merging are a parameter of the model.

The reason for the result is that when firms merge they avoid a price war that would otherwise enforce their collusion. Merging weakens these potential punishments and concomitantly weakens the extent of cooperation that can be enforced.

Our model extends Sannikov’s (2007) continuous-time model of tacit collusion, which built on the discrete-time models of Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1986) In these models firms share a market and choose output levels on an ongoing basis. The firms would like to collude but neither firm can observe the actions of the other firm. Instead, they observe price, which is influenced by both firms, but which also is influenced by the noise in demand. As a result, firms cannot directly infer the action of the rival firm, but instead must estimate it.

Figure 1 illustrates the pro-competitive effect of mergers.1 The axes are the continuation values of two colluding firms. There are two manifolds in the figure. The outer manifold is maximal equilibrium manifold, $\partial E$, found by Sannikov in his example of tacit collusion. In an equilibrium, at any moment the continuation values of the firms lie on a point of the manifold and are perturbed along the manifold by Brownian shocks.

If the continuation values lie in the northeast part of the manifold, the firms are colluding; they are producing at reduced rates and are effectively sharing monopoly profits, with some reduction due to the difficulty of coordination due to the noisy perturbations. If the continuation values lie in the southwest part of the diagram, the firms are in a price war. At this point both firms are producing close to competitive amounts, driving their current profits down. The ongoing Brownian perturbations eventually push the continuation values back to the northeast collusion region.

The inner manifold depicts a merger equilibrium. When the firms merge they share monopoly

1Interpreting tougher regulatory oversight, such as passages of national leniency laws to detect cartels, as higher merger cost or noisier collusion, the results of Dong, Massa, and Zaldokas (2016) largely support our paper’s result.
profits, with the shares determined endogenously by the locus at which the manifold intersects the line depicting the monopoly profits attained by merging, less the cost of merging. The merger equilibrium manifold lies entirely inside the no-merger equilibrium manifold, and in the collusive region—the northeast part of the equilibrium manifold—the merger manifold lies to the southeast of the no-merger manifold, although the distinct difference between the manifolds is not visible in the figure. As this region is highly stable, this figure demonstrates that collusion is weakened by the potential to merge.

**Figure 1. Merger and no-merger equilibrium manifolds**

This figure plots the merger equilibrium manifold from Figure 3 and the no-merger equilibrium manifold from Figure 2. The merger equilibrium manifold is contained entirely within the no-merger equilibrium manifold. The value in the collusive region of the merger manifold is therefore below the value in the collusive region of the no-merger manifold, expressing the reduced punishments and reduced competition of the merger manifold.

To establish these ideas, we take the main elements of Sannikov’s model, modified to include the potential to merge, but we also modify the solution method; we use an alternative equilibrium notion that enables us to solve the model using ordinary stochastic calculus methods. We elaborate
1.1 Other implications

We also offer insights into the interactions between product market dynamics and mergers, which is consistent with empirical patterns. Because collusion is stable, we can explain the seemingly anomalous empirical finding that mergers often do not appear to increase market power: pre-merger collusion comes much closer to replicating monopoly output than it does to a pure duopoly outcome. Thus, when firms merge, little changes.

As firms trade off the marginal gains from avoiding fiercer product market competition against the gains from creating a monopoly net of fixed cost of merging at the instant of the merger, there are no announcement returns, which is also empirically consistent; this has been noted for example by (Eckbo, 1983). A number of empirical articles conduct announcement return (event) studies. Eckbo (1983), Stillman (1983), and Eckbo and Wier (1985) initiated the study of the impact of mergers on market power by using capital market data. Their studies build on the proposition that, in an efficient market, any merger-induced change in future profits of firms competing in the same product market as the merging firms goes hand in hand with merger-induced abnormal announcement returns to merging and rival firms. The implicit assumption is that firms will incorporate the change in market power into their product market strategies only at the time of the merger announcement and hence the market will only then impound the effect of anti-competitive mergers into stock prices of the merging firms and their rival firms. These studies and also more recent ones (see, e.g., Fee and Thomas (2004) and Shahrur (2005)) fail to find evidence of anti-competitive effects associated with horizontal mergers. It has been argued that this failure is attributable to the inability of the event study methodology to detect the impact of horizontal mergers on competition in the case of diversified firms (McAfee and Williams, 1988) or the deterrent effect of antitrust enforcement (Prager, 1992). In contrast to these papers, our paper shows that once the cost of merging is also taken into account, there should be no announcement returns due to anti-competitive effects of mergers. Moreover, the decline in competition at the time of a merger is greater for higher merger costs.

Interpreting the firm’s continuation value as its stock price, the model has implications for pre-merger returns for acquirer and target. Defining the acquiring firm as the firm with the larger merger share, this firm experiences a run-up in stock price for some time prior to a merger. However,
at some point closer to the merger boundary the more valuable (acquiring) firm starts experiencing negative pre-merger returns. At the same time, the less valuable (target) firm, which consistently loses value on the way out of the collusive market-sharing region, experiences positive returns prior to the announcement of the deal. These results are consistent with empirical evidence reported, e.g., by Schwert (2000), that returns to target shareholders around a takeover announcement are typically positive, whereas returns to bidding shareholders are typically negative. In our model negative pre-merger returns to bidding shareholders can arise simply because of the competition dynamics of the noisy collusion equilibrium.

Because of the stability of pre-merger collusion, in equilibrium firms spend most of their time successfully colluding. This is consistent with the relative rarity of mergers: Andrade, Mitchell, and Stafford (2001) find that aggregate merger activity relative to market capitalization (or firms in the Center for Research in Security Prices database) is, on average, in the range of 1%–2% (or 2%–4%).

1.2 Related theory

Our results depend on the specific structure of the model. Sannikov and Skrzypacz (2007) for example show a dynamic model similar in spirit to ours, that if information arrives continuously and firms are able to react quickly to that information, and the key assumption that the firms’ goods are perfect substitutes is maintained, then collusion breaks down. We evade this outcome by the assumption that the firms’ outputs are not perfect substitutes, as similarly assumed by Sannikov (2007). Our conclusion is similar, in that we find that collusion is weakened by the potential to merge.

Jovanovic and Braguinsky (2004) consider a model with a continuum of firms, which have managers who choose projects from a continuum. It sets up a market in which firms have good managers who seek to acquire firms with good projects (but bad managers, who are fired after the acquisition). It thus is about efficiency gains from acquisitions, but does not address the issue of

---

2 Asquith (1983) documented a similar switch in the direction of return growth near announcement dates.

3 It has often been argued that when the market believes that a takeover is less likely to be value-enhancing, an acquiring firm’s stock price declines just prior to the announcement of a deal. According to our dynamic model, this reasoning might be flawed or, at a minimum, misleading: the acquiring firm in our model experiences negative pre-merger returns simply because it loses its position as an incumbent relative to the less valuable position of the target as market entrant, as it moves toward the merger boundary, i.e., entirely due to the dynamics of the equilibrium.

4 Hoberg and Phillips (2010) find that mergers are less likely for firms that are more similar to rivals, which they call “competitive effect,” because there are more firms with similar restructuring options so the potential to merge is less valuable. In our model, this corresponds to higher merger costs and hence greater potential for punishments.
market power associated with mergers, the focus of our paper.\textsuperscript{5}

The paper by Nocke and Whinston (2010) constructs a theoretical model of the optimal dynamic regulation of mergers. Like the Jovanovic-Braguinsky paper, they set up a model with a collection of many firms, and in each period merger opportunities—in the sense that a merger can achieve market power, production efficiencies, or both—randomly occur, and the firms facing these opportunities must decide whether to apply for regulatory approval. The regulator must then decide whether to approve a merger, taking account of the direct consequences of the merger in the sense of improving competition, but also on the future incentives for firms to apply for approval. Firms have merger opportunities but do not interact or collude or otherwise interact prior to this; thus, a firm’s optimal decision problem is static, namely to propose a merger or not when the opportunity arises. Also, there is no assumption of information asymmetry—noise—and the resulting strategic behavior. Our paper is very different: the entire focus is on firm behavior in anticipation of a merger, and in the face of noise in a very strategic environment.

Bernile, Lyandres, and Zhdanov (2012) is also related to our paper. As Gowrisankaran, they assume there are two incumbents and a third firm that is a potential entrant. Also as in Gowrisankaran, the incumbents compete on price in an essentially static fashion, the dynamics arising from demand shocks and decisions to merge or enter. Using a real options approach in which demand shocks follow a continuous-time diffusion process, they assume that entry by the third firm has a fixed cost; increased demand from the demand shock process leads to an increased gain from merging and exercising market power from the merger, but this increases the attractiveness of entry by the entrant, which dilutes those profits. Thus, in their model there is an \textit{indirect} fixed cost of merging—the dilution of profits arising from entry; in our model we assume an exogenous fixed cost of merging. They also characterize the merger threshold using value-matching and smooth-pasting conditions as we do. However because the competition between firms is essentially static, they do not model dynamic collusion that enables firms to enforce collusion via punishments, in the spirit of our model.

Finally, Hackbarth and Miao (2012) also bears similarities with Bernile, Lyandres, and Zhdanov (2012) and with our paper. They explore mergers in a continuous-time dynamic setting in which there are many firms. Demand evolves according to a diffusion process. A merger combines the firms’ capital as well as changing the competitive environment. They characterize the equilibrium

\textsuperscript{5}Bernile and Lyandres (2015) find that operating efficiencies also influence announcement returns of mergers.
value of a firm in the post-merger environment, taking account of the reduced number of firms in that environment, and the resulting changed oligopoly equilibrium. They then view the potential to merge as a real option to be triggered when demand passes a threshold that justifies the fixed cost of merging. As with Bernile, Lyandres, and Zhdanov (2012), they assume full information, and also assume that the firms act myopically (i.e., optimize output as if in a static oligopoly), so that the firms do not collude dynamically.

The rest of the paper is organized as follows. Section 2 outlines the model, which is then solved in Section 3. Section 4 derives the model’s implications. Section 5 concludes. The appendices contain derivations, proofs and other technical results. Because we use a different technical approach than Sannikov, some of the appendices focus on the alternative derivations and demonstrate how our alternative approach yields the identical equilibrium structure of the model.

2 The model

We consider two duopolistic firms that compete in an industry with differentiated products that are imperfect substitutes, where the two firms continuously take private actions (outputs). 6

We extend Sannikov’s construction by including boundary conditions associated with the merger. Given the dynamic and stochastic nature of the model, it is necessary to view the moment the firms merge as a stopping time, and the firms as choosing this as the optimal stopping time. Because both firms must choose the same optimal stopping time, there is a complication that goes beyond standard optimal stopping problems: the firms must somehow coordinate their stopping times.

As with ordinary optimal stopping problems, the coordination problem is solved by a smooth-pasting condition applied to each firm’s agency problem. In the agency construct, both firms naturally choose the same smooth-pasting point, and hence the same stopping time, thus solving the simultaneity problem. But a transformation of the model must first be carried out to enable the smooth-pasting problem to be appropriately expressed.

Our treatment of the game differs significantly from Sannikov’s treatment in that it is far less general, but, we feel, far more tractable. We make a significant assumption in that we focus only

---

6 The assumption that the products are differentiated is crucial: Sannikov and Skrzypacz (2007) demonstrate that if firms share the market for the same good, the game becomes degenerate in the sense that firms can perfectly observe each others actions in equilibrium.
on equilibria in which each firm chooses its action conditional on the action function of the rival firm, where that function maps the current public information into an action—even though the public information state is not itself directly payoff relevant. (We also focus only on cases with no public randomization.) It is then optimal for the firm to itself react to the public information state, implicitly defining its own function of the state to determine its actions. This is in contrast to Sannikov’s more general construction in which each firm’s continuation values are driven directly by the public information processes.

A key feature of the method is that it is not necessary to know the rival’s policy function before developing the necessary conditions for a firm’s optimal policy. With this approach we can solve the model using conventional stochastic calculus methods.

Despite our more narrow construction, we arrive at the same result as Sannikov in the sense that the equilibrium we find comprises the same boundary set that Sannikov found when the firms are not allowed to merge. However our method does not encompass the potential equilibria in the interior of the boundary set that Sannikov’s method characterizes when there is public randomization.

2.1 Actions and payoffs

Each firm $i = 1, 2$ continuously chooses an action $A^i_t \in \mathcal{A}_i \subseteq \mathbb{R}_+$ for all $t \in [0, \infty)$. In Sannikov (2007), firms’ outputs are discrete. We allow them to range over a continuum for analytical tractability, which enables us to take derivatives rather than analyze the optimal strategy by comparing adjacent actions. The firms observe the history of a vector of public signals (i.e., price increments) $dP_t$, which depend on the actions. The instantaneous prices of firms 1 and 2 before and after the merger are given by the levels (not increments) of the processes:

$$dP^1_t = (\Pi_1 - \beta_1 A^1_t - \delta_1 A^2_t) \, dt + \text{noise},$$  

and

$$dP^2_t = (\Pi_2 - \delta_2 A^1_t - \beta_2 A^2_t) \, dt + \text{noise}.$$  

Because the products are imperfect substitutes the actions of both firms influence the prices of both outputs.
The instantaneous payoff functions are the product of output and price increments: \( g_i(\cdot)dt = A_i dP^i \) for \( i = 1, 2 \). Using lowercase letters to denote choices of \( A_i \) and realizations \( dP^i \), we express the expected incremental payoffs of the firms as:

\[
g_1(a_1, a_2) dt = a_1 dp^1 = a_1 (\Pi_1 - \beta_1 a_1 - \delta_1 a_2) dt,
\]

and similarly for firm 2. Discounted profits are integrals of these instantaneous profits. We also note for future reference that the functions \( g_i \) are by construction continuous and twice continuously differentiable.

### 2.2 Information

Firms learn about each other’s actions by observing prices; the linearity of the prices in equations (1)–(2) means that the information processes can be isolated from those observations and expressed as a continuous process with independent and identically distributed increments. Before the merger, firms do not see each other’s actions, but they see a vector of signals \( X_t \) by inverting the price process vector:

\[
dX^1_t = \frac{(A^1_t + A^2_t) dP^2_t}{2 (\Pi_2 - \delta_2 A^1_t - \beta_2 A^2_t)} - \frac{(A^1_t - A^2_t) dP^1_t}{2 (\Pi_1 - \beta_1 A^1_t - \delta_1 A^2_t)} = A^1_t dt + \sigma_1 dZ^1_t,
\]

and

\[
dX^2_t = \frac{(A^1_t + A^2_t) dP^1_t}{2 (\Pi_2 - \delta_2 A^1_t - \beta_2 A^2_t)} - \frac{(A^2_t - A^1_t) dP^2_t}{2 (\Pi_1 - \beta_1 A^1_t - \delta_1 A^2_t)} = A^2_t dt + \sigma_2 dZ^2_t,
\]

where \( Z_t \) consists of two independent Brownian motions \( Z^1_t \) and \( Z^2_t \). The state space information vector \( \Omega \) is thus characterized by all possible paths of \( X_t \), and the public information filtration \( \mathcal{F}_t \) is adapted to \( X_t \).

Each firm treats the same vector of public information, \( X_t \), as its state variable, but with only one control, namely its own output \( A^i_t \); the other firm’s output function, \( \alpha^i(\cdot) : X_t \mapsto R \) is taken as fixed; this represents a deviation from Sannikov’s approach, as we earlier noted. We will comment further on the distinction between our method and Sannikov’s method later in the development of the model.
2.3 Monopoly stage game

When firms merge, they eliminate their respective information problems and share the monopoly profits available when they are able to act as a single firm via the merger. To define the gains from merging, we calculate the monopoly profit of the merged firm. This is straightforward because we assume that one of the consequences of the merger is that the noise is eliminated, and the firms obtain full-information monopoly profits.

After the merger, the full-information, constant-output monopoly solution applies:

\[
\begin{align*}
    a_1^* &= \frac{(\delta_1 + \delta_2) \Pi_2 - 2 \beta_2 \Pi_1}{(\delta_1 + \delta_2)^2 - 4 \beta_1 \beta_2}, \\
    a_2^* &= \frac{(\delta_1 + \delta_2) \Pi_1 - 2 \beta_1 \Pi_2}{(\delta_1 + \delta_2)^2 - 4 \beta_1 \beta_2}.
\end{align*}
\]  

(6)

The flow payoff function starting from any time \( t \) after the merger is then

\[
\pi_m = r \int_t^\infty e^{-r(s-t)} \sum_{i=1}^2 g_i(a_1^*, a_2^*) \, ds = \frac{(\delta_1 + \delta_2) \Pi_1 \Pi_2 - \beta_1 \Pi_1^2 - \beta_2 \Pi_2^2}{(\delta_1 + \delta_2)^2 - 4 \beta_1 \beta_2},
\]  

(7)

which will be shared between the firms; we elaborate on the sharing rule below.

2.4 The merger

The merger entails fixed costs: these can include substantial legal fees that are necessary to obtain regulatory approval, due diligence measures, the generation of asset valuations, investment bank fees, and so on. As a practical matter, costly post-merger physical changes can be necessary as well: when two airlines merge, one of the fleets will need to be repainted.

Pre-merger collusion delays incurring these costs. The firms understand that they will eventually merge, and play the pre-merger game with this understanding.

The merger takes place at an equilibrium stopping time, \( T_m \), that is individually rational and optimal from the perspective of each firm. Because it is an optimal stopping time there really are two stopping times, \( T_m^1 \) and \( T_m^2 \), one for each firm; it is then a requirement of the equilibrium that the two stopping times be equal.

At the time of the merger, the firms form a monopoly and share the monopoly profits, possibly
asymmetrically. Firm $i$’s share is defined as:

$$\xi^i(X_{T_m})(\pi_m - k),$$

where $\pi_m$ is the (deterministic) payoff from entering the monopoly in equation (7), $\xi^i(\cdot)$ is firm $i$’s share in the monopoly as a function of the vector of states $X_{T_m}$ at the time of the merger, with $\xi^1(X_{T_m}) + \xi^2(X_{T_m}) = 1$, and $k$ represents the fixed cost of merging. The sharing rule (8) says that the merger enables the two firms to shift away from dynamic duopoly payoffs to sharing monopoly payoffs net of the transaction costs of merging.\(^7\)

At the time of the merger, the value of continuing without merging and the value of merging will be the same. Thus, firm $i$’s continuation value at the moment of the merger is given by

$$W^i(X_{T_m}^i) = \xi^i(X_{T_m}^i)(\pi_m - k).$$

We assume that the announcement of the willingness to merge is publicly observable; although output decisions are by construction noisy, the merger signal is not. This is key because a firm’s refusal to merge at the appropriate moment would then open the way to an equilibrium punishment for refusal, should that be needed to sustain mergers in equilibrium. We discuss the possibility of alternative paths to the merger in Appendix Appendix H, showing that the needed punishments would not be renegotiation-proof and so cannot support equilibria.

2.4.1 Pre-merger continuation value

With the basic structure of the merger in hand we can state the objective of the firms prior to the merger. Define the pre-merger continuation value $W_i^i(\cdot)$ as the mapping, $W^i : \mathcal{X} \to \mathbb{R}_+$, from the state vector $X_t$ to firm $i$’s time $t$ payoff in the continuous-time game. Given the boundary condition (9), this yields for all $t < T_m^i$ the following continuation value function for firm $i$:

$$W^i(X_t) = \sup_{T_m^i, A^i(\cdot)} \mathbb{E} \left[ r \int_t^{T_m} e^{-r(s-t)} g_i(A^1_s, A^2_s) \, ds + e^{-r(T_m^i-t)} \xi^i(X_{T_m^i})(\pi_m - k) \right| \mathcal{F}_t \right],$$

\(^7\)In other words, the merger does not have any other synergies, such as better cost efficiency, economies of scale, efficient capital reallocation, etc., enabling us to focus on the effects of market structure alone.
where $\mathbb{E}[\cdot] \mid \cdot$ denotes the conditional expectation operator at time $t$ and with the state process in (4) and (5), and with

$$A_r^{-i} = \alpha^{-i}(X_i) \tag{11}$$

where $\alpha^{-i}(\cdot)$ is an element of the positive bounded functions on $\mathbb{R}^2$, taken as given by firm 1.\(^8\)

### 2.5 The formal problem and definition of equilibrium

The problem we consider is that of finding the maximal set of payoffs attainable in equilibrium in the repeated game between the two firms, subject to the constraint that players’ continuation values can never fall below the merger line. This is because continuing to play the collusion equilibrium nets the firms more profit than merging due to the fixed cost of merging, insofar as every equilibrium point above the merger line Pareto-dominates at least part of the merger line.

At the merger line, the continuation values are by definition equal in the merger and no-merger states and so the merging does not affect the instantaneous outcome. However the marginal impact of merging must also be accounted for, and this is expressed as the requirement that the shares garnered by each firm at the moment of the merger must be locally optimal for each, conditional on the other firm’s strategy. More formally, we can define the game as follows.

**Definition 1** A Markov merger game is a repeated game with stage game that is a tuple $(\mathcal{L}_i)_{i \in \{1,2\}}, (g_i)_{i \in \{1,2\}}$, where $(\mathcal{L}_i)_{i \in \{1,2\}}$ is the space of action functions of player $i$, and $g_i$ is the payoff of player $i$, and in addition a tuple

$$\{T^1_m, T^2_m, \xi^1(\cdot, \cdot), \xi^2(\cdot, \cdot), S^1(\cdot), S^2(\cdot)\},$$

such that at any moment $t$, firms choose a publicly observable signal $S^i_t$ from the set \{“do not merge,” “agree to merge with share $\xi(X^1_t, X^2_t)$”\} and where $T^1_m$ and $T^2_m$ are the stopping times defined as the first time the signal $S^i = \text{“merge”}$ is chosen by firm $i$, and such that if the firms agree to merge simultaneously, that is, $T^1_m = T^2_m = T_m$, then the merger takes place and the sharing rule is implemented with sharing rules $\xi^1(\cdot, \cdot)$ and $\xi^2(\cdot, \cdot)$.

Thus, a Markov merger game is similar to Sannikov’s game in the run-up to the merger, during

---

\(^8\) By taking the rival’s action $A^2_t$ as given, firm 1 views $A^1_t$ as a drift function of the history of the signals $X_{id}$, expressed as the filtration $\mathcal{F}_t$, but firm 1 does not take account of its indirect influence on the rival’s action via its effect on the signal. See Sannikov (2007), pp 1292-1293.
which time the firms can be thought of as sending the “do not merge” signal (or at least one of them). (The dissimilarity with Sannikov’s structure arises from the structure of the actions: firms choose actions conditional the rival firm’s policy function and the public information state, implicitly defining their own policy function as a function of that state.) At the moment of the merger, they both send the “merge” signal along with the choice of the sharing rule, and the merger takes place, and is irreversible.

We turn now to the definition of equilibrium.

**Definition 2** A Markov merger game equilibrium consists of:

- an action policy function pair \((\alpha^1(\cdot), \alpha^2(\cdot))\) and a sharing function pair \((\xi^1, \xi^2)\) that satisfy:
  \[
  A^i_t = \alpha^i(X^i_t) \text{ maximizes the expected discounted payoff of player } i \text{ given the strategy } \alpha^{−i}(X^i_t) \text{ of the rival firm after all public histories } X^i_t, \text{ where the payoff is defined in equation (10), and in addition,}
  \]
- firms merge only if both firms simultaneously play “agree to merge” with shares defined by \(\xi(X^1_{\mathcal{T}_m}, X^2_{\mathcal{T}_m})\);
- the stopping times \(\mathcal{T}^1_m\) and \(\mathcal{T}^2_m\) are optimal for firm 1 and firm 2 respectively;
- the stopping times are identical, that is \(\mathcal{T}^1_m = \mathcal{T}^2_m = \mathcal{T}_m\);
- the sharing rules are feasible, that is, they satisfy
  \[
  \xi^2(X_{\mathcal{T}_m}, X_{\mathcal{T}_m}) = 1 - \xi^1(X_{\mathcal{T}_m}, X_{\mathcal{T}_m}).\tag{12}
  \]

While the largest equilibrium set is driven by the following sharing rule along the merger line,

\[
W^2_{\mathcal{T}_m} = (\pi_M - k) - W^1_{\mathcal{T}_m},\tag{13}
\]

it is unclear how to ensure that, in equilibrium, the firms agree to merge *simultaneously*. Our solution procedure transforms the model so as to characterize optimality, as is commonly the case with optimal stopping problems, by a smooth-pasting condition at the moment of the potential
merger that holds for both firms. It is straightforward to show that the smooth-pasting condition is satisfied simultaneously; this is how we then demonstrate simultaneity.

3 Solution

In this section we derive the solution to the dynamic game using stochastic calculus, in two stages. Our stochastic calculus approach complements Sannikov’s (2007) geometric approach to deriving continuation values, firms’ strategies, and the equilibrium manifold. Our approach is a practical recipe for solving such models, and facilitates a direct economic interpretation of the boundary conditions. Our solution procedure has the following steps:

Stage 1  (i) solve the conventional profit maximization problem for each firm, taking the other firm’s action profile as given, using the public signal vector $X_t$ as the state vector, satisfying *incentive compatibility*;

(ii) use the solution of the firm optimization problem to state the continuation-value process for each firm, $W^i_t$, in terms of the state vector $X_t$, that is, *promise-keeping*;

(iii) demonstrate that simultaneous promise-keeping implies a singular volatility matrix, or *enforcement*, restricting the structure of the continuation-value processes locally to a one-dimensional manifold;

(iv) using the single-dimensionality of the enforcement manifold, implicitly map the state vector $X_t$ into the continuation-value vector $W_t$ using calculus arguments, so that a firm’s continuation value process $W^i_t$ is implicitly expressed as a function of the rival firm’s continuation value $W^{-i}_t$, i.e., implicitly construct the mapping $\mathcal{M} : \mathcal{X} \rightarrow \mathcal{R}$,

Stage 2  (i) pose the profit maximization problem for each firm as an agency problem with the rival firm’s transformed continuation value as the state process;

(ii) solve the principal’s optimal stopping problem using value-matching and smooth-pasting conditions and verify that an optimum is attained;

(iii) characterize the *self-generating* manifold stemming from the main differential equation implied by the simultaneous solution of the principal’s problem for both firms, and also the simultaneity of merger decisions, yielding an equilibrium;
(iv) verify that the optimal action $\alpha_i(\cdot)$ of firm $i$ is a function of the state $W^i$, and that this maps back into the function $\alpha^i(\cdot)$ that maps the state $X_t$ into an action.

3.1 First stage: Bellman equations when the states are noisy signals of actions

In the first stage we posit that the public signal vector $X_t$ comprises the state vector for the firms' value functions.

The required rate of return for the owners of firm 1 is the risk-free rate $r$. Thus, the Bellman equation for firm 1 in the continuation region of the continuous-time game is:

$$r W^1(X^1, X^2) dt = \max_{A^1} \left\{ \mathbb{E} \left[ r g_1(A^1, \alpha^2(X^1, X^2)) dt + dW^1(X^1, X^2) \right] \right\}, \quad (14)$$

Applying Ito’s lemma to expand the right-hand side of the Bellman equation and dropping the $dt$ terms, it is easy to verify that, with the observed signals as states, firm 1’s dynamic optimization problem is expressed in the following Hamilton-Jacobi-Bellman (HJB) equation:

$$r W^1(X^1, X^2) = \max_{A^1} \left\{ r g_1(A^1, A^2) + A^1 W^1_{X^1} + \alpha^2(X^1, X^2) W^1_{X^2} ight. \right.$$  

$$\left. + \frac{1}{2} \sigma_1^2 W^1_{X^1 X^1} + \frac{1}{2} \sigma_2^2 W^1_{X^2 X^2} \right\}, \quad (15)$$

where the arguments of $W^1$ have been suppressed on the right hand side to avoid clutter, and where the cross-partial terms have dropped out given that the noise terms are uncorrelated.

Invoking our assumption that the action space is a continuum, we can use a conventional derivative to generate the optimality condition:

$$rg_1 A^1 + W^1_{X^1} (X^1, X^2) = 0. \quad (16)$$

where we have made use of our assumption that $g_i$ is differentiable. We also note again that it is not necessary to know the rival’s policy function before deriving this condition.

Formulating the model using the history of public information as the driving process (Sannikov’s approach). Our key departure from Sannikov’s formulation is our positing that the public information process $X_t$ can be treated as a state variable. This approach depends on
the assumption that the rival firm chooses its actions based on the public information state; it is then optimal for the firm to react to the information state as well. In Sannikov’s formulation, the continuation values are direct functions of the entire history of the public information process. We can establish informally that the two approaches lead to the same first order conditions.

Using Sannikov’s approach, starting from the construction of the process in (14), define the continuation value process for firm 1 as

$$dW^1 = (rW^1_t - g^1(A^1_t, A^2_t))dt + W^1_{X^1_t}dX^1_t + W^2_{X^2_t}dX^2_t$$

that is, with the value process driven by the public information processes. Substituting from the definition of the $X_t$ process,

$$g^1(A^1_t, A^2_t)dt + W^1_{X^1_t}(A^1_t dt + \sigma_1 dZ^1_t) + W^2_{X^2_t}(A^2_t dt + \sigma_2 dZ^2_t)$$

$$= (g^1(A^1_t, A^2_t) + W^1_{X^1_t}A^1_t + W^2_{X^2_t}A^2_t)dt + (W^1_{X^1_t}\sigma_1 dZ^1_t + W^2_{X^2_t}\sigma_2 dZ^2_t)$$

Now substitute into (14), that is,

$$r W^1(X^1, X^2)dt = \max_{A^1} \left\{ E \left[ r g_1(A^1, A^2) dt + dW^1(X^1, X^2) \right] \right\},$$

$$= \max_{A^1} \left\{ E \left[ r g_1(A^1, A^2) dt + W^1_{X^1_t}dX^1_t + W^2_{X^2_t}dX^2_t \right] \right\}$$

$$= \max_{A^1} \left\{ E \left[ r g_1(A^1, A^2) dt + W^1_{X^1_t}A^1_t + W^2_{X^2_t}A^2_t dt \right. \right.$$ 

$$\left. + W^1_{X^1_t}(\sigma_1 dZ^1_t + W^2_{X^2_t}\sigma_2 dZ^2_t) \right\}$$

the incentive condition is then

$$r g_1 A^1 + W^1_{X^1_t}(X^1, X^2) = 0$$

which is identical to equation (16).

### 3.1.1 Promise-keeping

The envelope condition, that is, the substitution of the optimality condition (16) into the HJB equation (15), in combination with Ito’s lemma generate the stochastic process of the value states
of the firms. This is promise-keeping.

Lemma 1 The firms’ value states follow the processes

\[ dW^i = r(W^i - g_i) \, dt - \sigma_i r g_i A_i dZ^i_t + \sigma_{-i} W^i_{X_{-i}} dZ^{-i}_t. \]  \tag{21}

Proof: This follows from substituting the optimality conditions (16) and the HJB equation (15) into the Ito expansion of \( W^i_t \). \( \square \)

Equivalence with Sannikov’s approach to promise-keeping. We can substitute the optimality condition into the stochastic process for \( W^i_t \) constructed with Sannikov’s approach, that is, using (17), yielding

\[ dW^1 = r(W^1 - g_1) \, dt - \sigma_1 r g_1 A_1 dZ^1_t + \sigma_2 W^1_{X^2} dZ^2_t. \]  \tag{22}

This provides a way to express the idea that the continuation values are functions of the histories in a tractable way. Thus, we end up in the same place as with our state variable approach.

3.1.2 Characterizing the volatility matrix with calculus arguments

Combining the continuation value processes for the two firms and denoting the volatility matrix by \( B_t \) yields the vector process:

\[ dW_t = r (W_t - g(A_t)) dt + \begin{pmatrix} -\sigma_1 r g_1 A_1 & \sigma_2 W^1_{X^2} \\ \sigma_1 W^2_{X^1} & -\sigma_2 r g_2 A_2 \end{pmatrix} dZ_t \]

\[ = r (W_t - g(A_t)) dt + B_t dZ_t. \]  \tag{23}

The volatility matrix \( B_t \) contains cross-partial derivatives that we can partially characterize.

Proposition 1 The volatility matrix is singular.

Proof: The optimality conditions (16) can be multiplied to yield

\[ r^2 g_1 A_1 g_2 A_2 = W^1_{X^1} W^2_{X^2} \]  \tag{24}
Using the chain rule, we can write

\[ W_{X_1}^1 = W_{W_2}^1 W_{X_2}^2 \quad \text{and} \quad W_{X_2}^2 = W_{W_1}^2 W_{X_1}^1 \] (25)

and the condition (24) can then be written

\[
\begin{align*}
    r^2 g_{1A^1} g_{2A^2} - W_{X_2}^2 W_{X_1}^1 & = r^2 g_{1A^1} g_{2A^2} - W_{W_2}^1 W_{X_1}^1 W_{W_1}^2 W_{X_2}^1 \\
    & = r^2 g_{1A^1} g_{2A^2} - W_{X_1}^1 W_{X_2}^2 = 0
\end{align*}
\] (26)

This is the determinant of the volatility matrix \( B_t \), which is therefore singular as asserted. □

The singularity of the volatility matrix is the property of enforcement.

**Corollary 1** The continuation value process maps out a one-dimensional manifold.

**Proof:** The volatility matrix of the continuation value vector process is singular by Proposition 1, so the error process vector is mapped into a single effective stochastic process. Increments to this single process are added to the evolution determined by the drift functions, which is along a one-dimensional manifold. □

### 3.1.3 Strategy for transforming the \( X_t \) state to the \( W_t \) state

In the next step, we analyze the volatility matrix \( B \); we have not yet pinned down the cross-coefficients in the volatility matrix, that is, the optimality-determined values for \( W_{X_2}^1 \) and \( W_{X_1}^2 \), in equation (23). We use a simple strategy: the cross-partial derivatives \( W_{X_2}^1 \) and \( W_{X_1}^2 \) are slopes of the value-function vector \( W \) with respect to the state vector \( X \), but when we combine them we obtain slopes of \( W^1 \) in terms of \( W^2 \) and vice versa, and thus we obtain \( W^1 \) as an implicit function of \( W^2 \). This is roughly analogous to finding the slope of an indifference curve by taking the ratio of the marginal utilities. This will mesh with our strategy of formulating the second-stage problem, in which we express \( W^2 \) as a stochastic process that is the state for \( W^1 \).

Substituting from the optimality condition (16) into equation (25) results in the transformations,

\[
    W_{X_1}^2 = -r W_{W^1}^2 g_{1A^1} = -r \frac{1}{W_{W^2}^1} g_{1A^1},
\] (27)
and, similarly,
\[ W_{X^2}^1 = -r W_{W^2}^1 g_{2A^2}. \]  

(28)

We will use these transformations to construct the second-stage problem.

We draw attention to the fact that equations (27) and (28) implicitly redefine the value functions \( W^1 \) and \( W^2 \) as functions of \( W^2 \) and \( W^1 \) respectively, and no longer as direct functions of the states \( X^1 \) and \( X^2 \). To distinguish this transformed system we use a tilde notation, that is, \( \tilde{W}_t^i = \tilde{W}_t^i(W_{-i}^t) = W_t^i(X_t) \), and similarly \( \tilde{\xi}_t^i(W_t^i) = \xi_t^i(X_t) \), and so on; this approach has the usual abuse of notation in the sense that \( \tilde{W}_t^i \) denotes a process, whilst \( W_t^i(X_t) \) is a function of the process \( X_t \).

3.2 The second-stage agency reformulation

In the first stage formulation we generated the first order conditions for each firm, taking the other firm’s policy rule as fixed. From firm \( i \)'s perspective, firm \( -i \)'s policy rule can be viewed as a kind of contract against which firm \( i \) chooses its own actions. Knowing this, firm \( -i \) wants to choose the optimal contract. The second stage optimization solves this problem.

The optimal contract problem is potentially very complicated, but it turns out to be straightforward in this setting: simply optimize the action, taking as the state variable the value process of the rival firm, conditional on the rival firm’s optimization against the contract; we already solved this problem in stage 1. Thus, firm 2 finds its optimal contract with a state process (21), \( i = 2 \), and symmetrically firm 2 finds its optimal contract with the state process in (21), \( i = 1 \). Contracts that are consistent across firms are then an equilibrium.

3.2.1 The state equation

To formulate the equivalent agency problem we first characterise the state variable process. Normalizing \( \sigma_1^2 = \sigma_2^2 = 1 \) and \( \sigma_{12} = 0 \), and eliminating \( X_t^1 \) as an argument by substituting from equation (27) into equation (21), the continuation value process for firm 2 in terms of \( W_t^2 \) is:
\[ d\tilde{W}_t^2 = r(\tilde{W}_t^2 - g_2) dt - r g_{2A^2} dZ_t^2 + (-r) \frac{1}{W_t^1} g_{1A^1} dZ_t^1. \]  

(29)
3.2.2 The merger boundary from the agency perspective

For firm 1, the equation for its share of the net payoff from merging at time $t = T_m$ is given by:

$$
\tilde{W}_{T_m}^1(\tilde{W}_{T_m}^2) = (\pi_m - k) - \tilde{W}_{T_m}^2(\tilde{W}_{T_m}^1) = \tilde{\xi}^1(\tilde{W}_{T_m}^1, W_{T_m}^2)(\pi_m - k).
$$

(30)

Solving for the share of firm 1, $\tilde{\xi}^1$, yields:

$$
\tilde{\xi}^1(\tilde{W}_{T_m}^1, \tilde{W}_{T_m}^2) = 1 - \frac{\tilde{W}_{T_m}^2(\tilde{W}_{T_m}^1)}{\pi_m - k} \quad \text{and therefore} \quad \tilde{\xi}^1_{W_{T_m}^2} = -\frac{1}{\pi_m - k}
$$

(31)

3.2.3 The agency contract objective

In the second-stage “agency” setup the objective for firm 1 becomes

$$
\tilde{W}_t^1(\tilde{W}_t^2) = \sup_{T_m, A^1(\cdot)} \mathbb{E} \left[ r \int_t^{T_m} e^{-r(s-t)} g_1(A^1_s, \tilde{\alpha}^2(\tilde{W}_s)) \, ds \
+ e^{-r(T_m-t)} \tilde{\xi}(\tilde{W}_{T_m}^1, \tilde{W}_{T_m}^2)(\pi_m - k) \right],
$$

(32)

with the promise-keeping-constrained continuation value process of the rival firm in equation (29) as the state, and with the boundary condition

$$
\tilde{W}_T^1(\tilde{W}_T^2) = \tilde{\xi}^1(\tilde{W}_{T_m}^1, W_{T_m}^2)(\pi_m - k).
$$

(33)

and taking as given the other firm’s control process $A^2 = \alpha^2(\tilde{W}_t)$. There is a similar expression for firm 2.

We can separate the problem of finding the optimal action $A^1$ from the optimal stopping problem. That is, the HJB equation solves the optimal action problem and the smooth-pasting condition pins down the optimal stopping problem taking as given the optimal policy $A^1$.

The first step in this process is to write the Hamilton-Jacobi-Bellman optimization problem for firm 1, using $W^2$ as a state with state equation (29). The full statement of the firm-1 HJB equation is then

$$
0 = \max_{A^1} \left\{ r(g_1 - \tilde{W}^1) - r(g_2 - \tilde{W}^2)\tilde{W}^1_{W^2} + \frac{1}{2} \left( -r \tilde{W}^2_{W^1} g_1 A^1 \right)^2 \tilde{W}^1_{W^2 W^2} + \frac{1}{2} (rg_2 A^2)^2 \tilde{W}^1_{W^2 W^2} \right\}.
$$

(34)
This equation differs from the first-stage Bellman equation, equation (15). It is however consistent with (15) in that it takes the solution of (15) as an implicit constraint.

### 3.2.4 Optimality: the smooth-pasting condition

There are three main conditions that comprise the solution to the optimization problem stated in (34)-(33). The first condition is the standard optimality condition: maximize (34) with respect to $A^1$. However, rather than calculating this derivative explicitly, we will convert the problem into an equivalent maximization problem and solve that problem instead; we defer this to a later stage.

The second condition is the value-matching condition: at the merger point, the continuation value is identical to the share of the monopoly value less the cost of merging:

$$\tilde{W}_1^{\pi_m} = \tilde{\xi}_1^{T_m}(\tilde{W}_1^{T_m}, \tilde{W}_2^{T_m})(\pi_m - k).$$

The third condition is the smooth-pasting condition. The smooth-pasting condition for firm 1 follows from differentiating the boundary function (31):

$$\hat{W}_1^{1}(\tilde{W}_2^{T_m}) = \hat{\xi}_1^{T_m}(\tilde{W}_1^{T_m}, \tilde{W}_2^{T_m})(\pi_m - k).$$

The third condition is the smooth-pasting condition. The smooth-pasting condition for firm 1 follows from differentiating the boundary function (31):

$$\hat{W}_1^{1}(\tilde{W}_2^{T_m}) = \hat{\xi}_1^{T_m}(\tilde{W}_1^{T_m}, \tilde{W}_2^{T_m})(\pi_m - k) = -1,$$

with a similar condition for firm 2.\(^9\) We begin with a lemma about the smooth-pasting condition.

**Lemma 2** The smooth-pasting condition is necessary for an optimum with respect to the action $A^1$.

**Proof:** See Appendix Appendix E. □

We can then state a key result.

**Proposition 2** The smooth-pasting condition implies equal stopping times: $T_m^1 = T_m^2$.

---

\(^9\)We again draw attention to our notation: $\tilde{W}_2^{T_m}$, viewed by firm 1 as a state variable, denotes firm 2’s continuation value evaluated at the stopping time $T_m$, whilst $\hat{W}_1^{1}(\tilde{W}_2^{T_m})$ denotes the partial derivative of firm 1’s continuation value as a function of that state at the stopping time.
Proof: The proof follows from two observations. First, for the value-matching condition to be met, that is, for the terminal point to be on the merger line, the value-matching condition is necessarily met for both firms simultaneously. Second, the smooth-pasting condition (in the stage 2 agency formulation) entails the condition

\[ \tilde{W}_1^1(\tilde{W}_2^1) = -1, \]

for firm 1; inverting the equation yields

\[ \tilde{W}_1^2(\tilde{W}_2^1) = -1, \]

which is the smooth-pasting condition for firm 1. Thus, satisfying the smooth-pasting formula for firm 1 necessarily satisfies the smooth-pasting formula for firm 2. □

Although the smooth-pasting condition in equation (36) is straightforward to state once we have transformed the problem to agency form, it is worthwhile contemplating the economic meaning of the condition. As the firms are driven to the merger line by the realizations of the noise, they stay on the equilibrium manifold by trading current payoffs against future “promise-keeping” payoffs. Indeed, our agency reformulation makes this trade-off explicit, in the sense that each firm sees itself as a principal offering this trade-off to the other firm via an equilibrium “contract” that accounts for the fact that the other firm optimizes against this contract (i.e., it is incentive-compatible). We note that the smooth-pasting condition (36), which expresses incentive compatibility at the merger point, reflects—like Sannikov’s (2007) incentive-compatibility condition (9) in what for us is the pre-merger play—the trading of utility between the two firms. However, the rate of exchange is fixed by the slope of the merger line. We summarize with the following proposition:

Proposition 3 The action \( A^i(\cdot) \) and \( \tilde{W}^i(\cdot) \) that solve (34), (33), and smooth-pasting condition (36), taking as given \( A^{-i} = \tilde{\alpha}^{-i}(\tilde{W}_i) \), solve the optimal action and stopping problem (32).

Proof: See Appendix Appendix E. □
3.3 Equilibrium and characterization

To reach an equilibrium in the game, firms must choose optimal contracts in their role as principals, optimally reacting to the other firm’s contract in their role as agents, and the contracts must be identical; furthermore, they must agree on an identical stopping time. The agency approach enables us to do this in a direct and tractable way. We begin with a lemma.

Lemma 3 For any manifold that is self-generating, satisfies the value-matching and smooth-pasting conditions, and which lies entirely above the merger line, it is not Pareto-improving to merge prior to reaching the merger line.

Proof: Because the merger announcement is public, the firms could mutually agree to merge prior to attaining the merger line, that is, they could agree to jump to some point on the merger line prior to attaining it via evolution along the manifold. By hypothesis, the manifold lies above the merger line, so for at least one of the firms the jump to the merger would reduce its continuation value and it would be not be individually rational to agree to the merger.

We remark that it is possible to construct self-generating manifolds that satisfy the value-matching and smooth-pasting conditions, but which lie entirely below the merger line; these manifolds fail as equilibria precisely because it is Pareto-improving to jump to the merger line rather than evolve toward it via the self-generating manifold. We explore this in more detail in Appendix H.

Proposition 4 The value functions that solve (34)-(36) for both firms simultaneously constitute a Markov merger equilibrium.

Proof: This is a direct consequence of satisfying Sannikov’s main differential equation, equation (14) of Sannikov (2007), which we derive in Appendices B, C, and D, and the smooth-pasting condition (36).

We convert the optimization problem into the form it takes in equation (E.7) to facilitate the expression of the equilibrium manifold in terms of polar coordinates, which in turn makes the numerical solution of the model more straightforward. We turn to this agenda next.
3.3.1 Converting the ODE into geometric form

The nonlinearity of the model forces us, as it did Sannikov, to solve the model numerically in the analysis of practical examples. Like Sannikov (2007), we adopt a reformulation of the second-stage optimized Bellman equations. We express the differential equation for the equilibrium manifold in polar coordinates to facilitate the computation of numerical solutions in the next section. The details of these derivations, which were not provided by Sannikov, are presented in the appendix.

The normal and tangent to the manifold are, respectively:

\[ \mathbf{N}(\theta) = (\cos(\theta), \sin(\theta)) \], \quad \text{and} \quad \mathbf{T}(\theta) = (-\sin(\theta), \cos(\theta)). \tag{39} \]

As demonstrated in Appendix B, we can then express the value process in polar coordinate form: the curvature \( \kappa(W) \) of the equilibrium manifold, which is determined by the transformed form of the optimized Bellman equations of the firms, is the derivative of the polar coordinate with respect to movement along the equilibrium manifold. For example, \( \kappa \) would be zero if the manifold were locally a straight line. Thus, angle \( \theta \) and arc length \( \ell \), we have along the manifold that:

\[ \frac{d\theta}{d\ell} = \kappa, \quad \text{and} \quad \frac{d\ell}{d\theta} = \frac{1}{\kappa}, \tag{40} \]

The left-hand side derivative can be expressed in terms of the derivatives of the continuation values using polar coordinates, which leads us to the differential equation we solve numerically:

\[ \begin{pmatrix} \frac{dW^1(\theta)}{d\theta} \\ \frac{dW^2(\theta)}{d\theta} \end{pmatrix} = \begin{pmatrix} \frac{\sin(\theta)}{\kappa(\theta)} \\ \frac{\cos(\theta)}{\kappa(\theta)} \end{pmatrix} = \mathbf{T}(\theta) \frac{\kappa(\theta)}{\kappa(\theta)}. \tag{41} \]

We solve this ordinary differential equation numerically to determine the equilibrium manifold, \( \partial \mathcal{E}(r) \), which provides a benchmark equilibrium. If a merger occurs, we numerically solve for \( \partial \mathcal{E}(r) \) subject to the boundary conditions (i.e., value-matching and smooth-pasting in equations (9) and (36)), which are not present in Sannikov (2007). These boundary conditions for the merger will have non-trivial effects on the firms’ pre-merger strategies and values, which we study in the next section.
3.3.2 The impact of the merger on the equilibrium set

Our next task is to prove that the equilibrium manifold shrinks with the merger cost $K$, that is, that the potential to merge results in a boundary condition that translates into reduced punishments and therefore reduced collusion. We begin with notation. Appropriating Sannikov’s notation, we denote the boundary of the equilibrium manifold, $\partial E^K_M$ for the merger model, and $\partial E_{NM}$ for the no-merger model, with

Proposition 5 The Markov merger equilibrium set is strictly contained inside the no-merger Markov equilibrium set and shrinks with the merger cost $K$, that is, for $K' > K$,

\[ \partial E^K_M \subset \partial E^{K'}_M \subset \partial E_{NM} \] (42)

The proof is in Appendix Appendix F.

4 Implications

4.1 The no-merger equilibrium

To begin our numerical analysis, we solve for the equilibrium of the benchmark model without mergers. The benchmark model’s solution to the differential equation (41) is characterized by an equilibrium set, $\partial E(r)$, that forms a manifold in the space of continuation values, $(W^1, W^2)$, as seen in Figure 2. We assume a baseline environment with symmetric demand functions and the following parameter values: $\Pi_1 = 30$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 2$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1$.\(^{10}\)

The static Nash equilibrium in the duopoly stage game is $(5, 5)$ in the baseline environment. This generates continuation values of $\pi_{d,i} = 50$ for each firm $i = 1, 2$ (or 100 for both firms).

Figure 2 displays the equilibrium manifold in the benchmark case without mergers along with firm 2’s output choices. Due to the symmetric demand functions, we can rotate firm 2’s output choices around the 45 degree line to obtain firm 1’s output choices. In the northeast stretch of the equilibrium manifold the firms cooperate, with output levels around 4 and hence are highly

\(^{10}\)Our symmetric example differs from Sannikov’s (2007) asymmetric example, the parameter values of which are $\Pi_1 = 25$, $\Pi_2 = 30$, $\beta_1 = 2$, $\beta_2 = 2$, $\delta_1 = 1$, $\delta_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $r = 1.5$. While the baseline parameter choices could be motivated in detail, we omit this for the sake of brevity. The model’s results and implications vary quantitatively but not qualitatively with parameters.
collusive. On the opposite side of the equilibrium manifold, that is, in the southwest stretch, they engage in a price war, in which output levels are around 10 and hence highly non-collusive.

In the upper right region of \( \partial E(r) \) in Figure 2, output is low for both firms in that it sums up to about 8. That is, the firms approach the monopoly output, which is, according to equations (6), \( a_i^* = 3.75 \) for each of the two firms (or 7.5 for both firms). Their continuation values are consequently higher, (around (56,56) at the midpoint of the market-sharing region) than in the static duopoly’s Nash equilibrium, which corresponds in terms of continuation values to (50,50), depicted by \( N \) in the figure. The monopoly value of the two firms is \( \pi_m = 112.5 \) and is depicted by the (dotted) monopoly line for all (feasible) sharing rules in the unit interval. Clearly, the monopoly
value is unattainable in either the dynamic or the static duopoly. It is evident that, in this region, when a firm’s continuation value increases, its market share also increases. Therefore, firms are tempted to overproduce, moving away from the center of the market-sharing region.

In the upper left segment of the equilibrium manifold in Figure 2, firm 2 obtains the maximal continuation value of almost 83, while the continuation value of firm 1 equals about 25. At that point, firm 1 underproduces, while firm 2 overproduces relative to the duopoly and monopoly quantities. Output is very asymmetric: at the lower right, for example, firm 1’s output is high (i.e., 10) and firm 2’s is low (i.e., 0). In the lower right segment of the equilibrium manifold, firms thus display similar strategies with the roles of firm 1 and 2 reversed, namely with firm 1 the incumbent and firm 2 as the entrant.

At the intersection of $\partial E(r)$ with the 45 degree line, firms engage in a price war in that both firms aggressively overproduce. Their output levels are (10,10) and substantially exceed the (static) duopoly outputs of (5,5), which leads continuation values to drop well below (25,25).

4.2 Merger equilibrium

We continue our analysis by solving the differential equation (41) for the equilibrium manifold, $\partial E(r)$, with mergers by incorporating the boundary conditions for value-matching and smooth-pasting in equations (9) and (36). When the merger occurs, both firms share (net of the merger cost) the value of the resulting monopoly stage game without imperfect information. This yields the merger line in equation (30), which corresponds to the monopoly line, $\pi_m$, minus the cost of merging, $k$, and is represented by the red, dashed line in the figure for all feasible sharing rules in the unit interval. If firm 1, for example, captures more of the merger gains, then the merger point will more likely lie on the lower right section of the line; conversely, if firm 2 captures more of the gains, then the merger point is more likely to be on the upper left section of the line.

Figure 3 illustrates how firms anticipate the impending merger. As we demonstrated in Proposition 5 and in our discussion of Figure 1 in the introduction, the equilibrium manifold with mergers is entirely contained inside the original no-merger equilibrium manifold in Figure 2. This means that some of the collusion profits attainable in the no-merger equilibrium are not attainable in the merger equilibrium, while some of the non-collusion costs (due to potential price wars) are avoided in the merger equilibrium. Intuitively, this stems from the weaker punishments inherent
This figure plots the merger equilibrium manifold (blue, solid line), the monopoly line (black, dotted line), and the merger line (red, dashed line) for a merger cost of \( k = 24 \). Firm 2’s output choices are outside the equilibrium manifold. Due to the symmetric demand functions, we can rotate firm 2’s output choices around the 45 degree line to obtain firm 1’s output choices. The static Nash equilibrium’s output choices of \((5,5)\) are depicted by \( N \) in terms of the continuation values of \((50,50)\). We use the baseline environment in which \( \Pi_1 = 30, \Pi_2 = 30, \beta_1 = 2, \beta_2 = 2, \delta_1 = 2, \delta_2 = 2, \sigma_1 = 1, \sigma_2 = 1, \) and \( r = 1 \).

in the equilibrium with mergers. The punishments are weaker because the opportunity to merge eliminates the severe punishments in the price-war regime.

The (upper right) market-sharing region of the equilibrium manifold, which now reflects the possibility of a merger, is slightly less stretched out in Figure 3 compared with Figure 2. The firms trade off being the production leader in this noisy duopoly against being punished for deviating. As in the no-merger equilibrium, total output stays low at around 8, which is again very close to the optimal monopoly output of 7.5. In other words, the market-sharing regime, in which firms’ optimal output levels are highly collusive, is similarly large relative to one without an anti-competitive merger, as in the previous figure. Moreover, as the entrant and incumbent regime is
approached at the upper left region, total output increases to 10 and finally to 12 and 13 in the contestability region. But then total output declines slightly again to 12 just before the merger line is smoothly pasted to the equilibrium manifold. Thus, compared with the previous figure’s no-merger equilibrium manifold, total output tends to be lower in the worst stages of the dynamic game.

In practice, merger gains (or payoffs) are also typically split asymmetrically between the merging firms. The model predicts this: the firm that is being punished in the contestability region gets a smaller share of the merged entity’s value, because it has a smaller continuation value and hence it appears to be taken over by the overproducing firm that has a larger continuation value in the contestability region. It is therefore reasonable to designate them target and acquirer. The firm that overproduces at the right time will be rewarded by the larger share in the merged entity if the merger boundary is reached. This asymmetry is not driven by any inherent asymmetry in the demand functions, noise parameters, or other parameters, which are all symmetric: it is driven solely by the state of product market competition that the firms have attained as a result of cumulative play of the noisy duopoly game.

4.3 Collusion and the dearth of mergers

We next examine the stability of the no-merger and merger equilibria. The arrows in Figures 4 and 5 provide information about the stability of the regions. The length of the arrows, which corresponds to the volatility-scaled drift of the continuation value process represented in equation (23), indicates the strength of stability. The direction of the arrows conveys information about the local stability of each region.

The arrows point away from each other in the price-war region of Figure 4, indicating that it is unstable. On the other hand, the arrows point towards each other in the collusive market-sharing region, indicating that it is stable. Even though the collusive region is smaller in Figure 5 than in Figure 4, because the option to merge weakens the punishments that enforce collusion and thus weakens collusion, it remains highly stable in the merger case. In addition, observe that there are two nodes where the stability flips between the collusion node and the merger node: the unstable one in the entrant and incumbent region of the equilibrium manifold, and the additional stable

\footnote{Sannikov (2007) studies stability for the partnership example (see his Figure 2), but not for the duopoly example.}
This figure plots the no-merger equilibrium manifold (blue, solid line). The gray arrows indicate the stability of the dynamic game, where the length of each arrow is the scaled drift of the value state vector. The static Nash equilibrium’s output choices of \((5,5)\) are depicted by \(N\) in terms of the continuation values of \((50,50)\). We use the baseline environment in which \(\Pi_1 = 30, \Pi_2 = 30, \beta_1 = 2, \beta_2 = 2, \delta_1 = 2, \delta_2 = 2, \sigma_1 = 1, \sigma_2 = 1,\) and \(r = 1\).

As will be apparent in the simulations of the model, the extra stable node in the contestability region has little impact on the actual dynamics of the equilibrium. However, if the merger state is approached, then in terms of corporate practice, this corresponds to mergers being “imminent” or “anticipated” just before they are announced. For example, Edmans, Goldstein, and Jiang (2012) and Cornett, Tanyeri, and Tehranian (2011) provide empirical evidence for this anticipation.

The area close to the merger points in Figure 5 is unstable; this instability, in concert with the stability of the collusion region, delays the onset of mergers: they will be rare.

\(^{12}\) However, if the merger state is approached, then in terms of corporate practice, this corresponds to mergers being “imminent” or “anticipated” just before they are announced. For example, Edmans, Goldstein, and Jiang (2012) and Cornett, Tanyeri, and Tehranian (2011) provide empirical evidence for this anticipation.

\(^{13}\) Comparing the stability diagrams, in both figures the instability of the contestability region makes it likely that the firms will get back to cooperating if they stray into this region. In the unlikely event that the cusp in the contestability region is crossed, a price war (or a merger if there is the potential for it) is unstable, thus making a price war (or a merger) unlikely.
This figure plots the merger equilibrium manifold (blue, solid line), the monopoly line (black, dotted line), and the merger line (red, dashed line) for a merger cost of \( k = 24 \). The gray arrows indicate the stability of the dynamic game, where the length of each arrow is the scaled drift of the value state vector. The static Nash equilibrium’s output choices of \((5,5)\) are depicted by \( N \) in terms of the continuation values of \((50,50)\). We use the baseline environment in which \( \Pi_1 = 30, \Pi_2 = 30, \beta_1 = 2, \beta_2 = 2, \delta_1 = 2, \delta_2 = 2, \sigma_1 = 1, \sigma_2 = 1, \) and \( r = 1 \).

As we demonstrated in Proposition 5, as the merger cost falls, the equilibrium manifold changes generally. It flattens, reflecting that the firms are increasingly acting like a shadow monopoly in terms of output, with the main issue being the equity shares in the merged entity. Outsiders unaware of the potential for a merger attempting to value the companies would find output choices diminished relative to the theoretical prediction of the static Nash equilibrium. In addition, regulators would find greater collusion than would seem warranted by that same benchmark. This collusion will be strongest when the merger is most remote. For practical purposes, the merger will be a phantom, seemingly unrelated and hidden from the firms’ current actions. While Andrade, Mitchell, and Stafford (2001), for example, point out that stronger antitrust laws and stricter en-
forcement have provided challenges for anti-competitive mergers, this model’s solution implies that the dearth of market power increasing mergers need not imply more competition in a dynamic duopoly game, which is designed for the companies to compete.

The dynamic model thus suggests that tests pointing to rejection of the market power doctrine might be ill-posed: there is not much to deter if the anti-competitive effects of horizontal mergers are anticipated in merging and rival firms’ product market strategies prior to merger announcements (or likely challenges by regulators). Consistent with our dynamic model’s insight, Eckbo (1992) even concludes the following on p. 1005:

While the U.S. has pursued a vigorous antitrust policy towards horizontal mergers over the past four decades, mergers in Canada have until recently been permitted to take place in a virtually unrestricted antitrust environment. The absence of an antitrust overhang in Canada presents an interesting opportunity to test the conjecture that the rigid market share and concentration criteria of the U.S. policy effectively deters a significant number of potentially collusive mergers. The effective deterrence hypothesis implies that the probability of a horizontal merger being anti-competitive is higher in Canada than in the U.S. However, parameters in cross-sectional regressions reject the market power hypothesis on samples of both U.S. and Canadian mergers.

Judging from the Canadian evidence, there simply isn’t much to deter.

In sum, the model is consistent with several regularities in the mergers and acquisitions literature that have heretofore led to rejection of the market power doctrine. The solution suggests an alternative interpretation of the literature’s empirical tests. According to our dynamic model with the possibility of anti-competitive mergers, it is not surprising but rather inevitable that the evidence for the market power doctrine is weak when using capital market data and short-term announcement return methods to gauge changes in competition (or concentration) that have already taken place prior to the announcement return window when firms optimize dynamically.

5 Conclusion

We have studied mergers in a dynamic noisy collusion model, building on the models of Green and Porter (1984), Abreu, Pearce, and Stacchetti (1986), and especially Sannikov (2007). At each instant, firms either privately choose output levels or merge, which trades off benefits of avoiding
price wars against the costs of merging. Mergers are optimal when collusion fails. Long periods of collusion are likely, because colluding is dynamically stable. Therefore, mergers are rare. Lower merger costs decrease pre-merger collusion, as punishments by price wars are weakened. This suggests that, although anti-competitive mergers harm competition ex-post, barriers and costs of merging due to regulation should potentially be reduced to promote competition ex-ante.

We close by noting areas that warrant future research in this class of dynamic models. First, we have restricted attention to two firms. Extending our analysis to three or more firms would be informative about the impact of mergers on non-merging rivals as examined in many of the empirical studies. Moreover, Salant, Switzer, and Reynolds (1983), for example, show that the presence of three or more firms in an industry can deter mergers, and that mergers can be welfare-enhancing, even in the absence of scale economies or synergies. Perry and Porter (1985) examine this result further with a more fine-grained treatment of the allocation of costs in the merged firm and moved the conclusion back in the classical direction. Farrell and Shapiro (1990) establish that quantity competition in a post-merger industry raises prices if there are no scale economies or synergies, but still find cases where mergers are deleterious to potential merging firms. Other researchers, such as Deneckere and Davidson (1985) and Gaudet and Salant (1992), analyze welfare and policy implications in extensions of these models, again finding some counterintuitive results. The technical challenge in expanding the model to multiple firms is significant, however, in that equilibrium manifolds would reside in higher-dimensional spaces, with a concomitant increase in the computational difficulty of numerical solutions, such as inverting the volatility matrix.

Second, as we show, the model is related to agency. While we (as do many studies) treat firms as black boxes that are able to hide information, one might reinterpret this as more like standard agency construct in which managers hide information from rival firms. This might be expanded to incorporate explicit agency elements into the continuous-time game. With agency explicit, a merger might not eliminate all information asymmetries: we could ask whether the increase in market power effected by the merger is strengthened or weakened, and how pre-merger collusion is affected.
Appendix A Derivation of the value process in the first stage

Here is the derivation of Lemma 1.

**Proof:** We establish the result for firm 1. We first apply Ito’s lemma to $W^1(X^1_t, X^2_t)$ in generate the stochastic continuation value process of the state:

$$
dW^1 = (A^1 W^1_{X^1} + A^2 W^1_{X^2} + \frac{1}{2} \sigma^2_1 W^1_{X^1 X^1} + \frac{1}{2} \sigma^2_2 W^1_{X^2 X^2}) dt + \sigma_1 W^1_{X^1} dZ^1_t + \sigma_2 W^1_{X^2} dZ^2_t. \quad (A.1)
$$

Notice the resemblance of the terms in the drift to the stage-game payoffs in the Bellman equation. Substituting equation (15) into (A.1) yields a simpler expression for the continuation value process:

$$
dW^1 = (r W^1(X^1, X^2) - r g_1(A^1, A^2)) dt + \sigma_1 W^1_{X^1} dZ^1_t + \sigma_2 W^1_{X^2} dZ^2_t. \quad (A.2)
$$

We further modify this equation by using the optimality condition (16) to eliminate the $W^1_{X^1}$ term, replacing $W^1_{X^1}$ with $-r g_1 A^1$ (i.e., the envelope condition):

$$
dW^1 = (r W^1(X^1, X^2) - r g_1(A^1, A^2)) dt - \sigma_1 r g_1 A^1 dZ^1_t + \sigma_2 W^1_{X^2} dZ^2_t. \quad (A.3)
$$

Dropping the arguments, we find that $W^1$ evolves according to:

$$
dW^1 = r (W^1 - g_1) dt - \sigma_1 r g_1 A^1 dZ^1_t + \sigma_2 W^1_{X^2} dZ^2_t. \quad (A.4)
$$

This eliminates the explicit influence of the state variable $X^1_t$ from the equation. □
Appendix B  Converting to polar coordinates

We express the differential equation for the equilibrium manifold in polar coordinates to facilitate computation of numerical solutions. Repeating the introduction of polar coordinate notation from the main text, the normal and tangent to the equilibrium manifold are:

\[ N(\theta) = (\cos(\theta), \sin(\theta)), \quad T(\theta) = (-\sin(\theta), \cos(\theta)). \]  \hspace{1cm} (B.5)

We can express the first-order cross-partial of \( W^1 \) in trigonometric form:

\[ W^2_1 \frac{dW^2}{dW^1} = -\frac{\cos(\theta)}{\sin(\theta)}. \]  \hspace{1cm} (B.6)

Hence the second-stage Bellman equation (34) becomes:

\[
0 = \max_{A^1} \left\{ r(g_1 - W^1) - r(g_2 - W^2) \left( -\frac{\sin(\theta)}{\cos(\theta)} \right) \right.
\]
\[
+ \frac{1}{2} \left( \frac{-r \cos(\theta)}{\sin(\theta)} g_1 A^1 \right)^2 W^1_{W^2W^2} + \frac{1}{2} \left( r g_2 A^2 \right)^2 W^1_{W^2W^2} \left\}, \right. \hspace{1cm} (B.7)
\]

which then leads to:

\[ W^1_{W^2W^2} = -\max_{A^1} \left\{ \frac{(g_1 - W^1) - (g_2 - W^2) \left( -\frac{\sin(\theta)}{\cos(\theta)} \right)}{r \left( \frac{-r \cos(\theta)}{\sin(\theta)} g_1 A^1 \right)^2 + (g_2 A^2)^2} \right\}. \hspace{1cm} (B.8)\]

After some algebra, the equation can be restated as follows:

\[ W^1_{W^2W^2} = -\max_{A^1} \left\{ \frac{1}{\cos(\theta)} \frac{(\cos(\theta)(g_1 - W^1) - (g_2 - W^2) (-\sin(\theta)))}{r \cos(\theta)^2 \left( \left( \frac{g_1 A^1}{\sin(\theta)} \right)^2 + \left( \frac{g_2 A^2}{\cos(\theta)} \right)^2 \right)} \right\}, \hspace{1cm} (B.9)\]

or

\[ W^1_{W^2W^2} = -\max_{A^1} \left\{ \frac{1}{\cos(\theta)^3} \frac{\cos(\theta)(g_1 - W^1) + \sin(\theta)(g_2 - W^2)}{r \left( \frac{g_1 A^1}{\sin(\theta)} \right)^2 + \left( \frac{g_2 A^2}{\cos(\theta)} \right)^2} \right\}. \hspace{1cm} (B.10)\]
The numerator term is $N(g - W)$, and the denominator term is $r|\phi|^2$, just as in Sannikov’s formula. Notice that this equation has a curvature on the left-hand side. The fact that it is a curvature will later be used in the numerical solution of the model. We repeat the exercise with firm 2 and obtain:

$$W_{w^1w^1}^2 = -\max_{A^2} \left\{ \frac{1}{\sin(\theta)^3} \cos(\theta)(g_1 - W^1) + \sin(\theta)(g_2 - W^2) \right\}.$$  \(\text{(B.11)}\)

Notice that the denominators in equations (B.10) and (B.11) are the same.

As shown in Appendix D, the second-order partial derivatives of the continuation values are weighted expressions of the curvature of the equilibrium manifold in the direction of the normal vector, $\kappa(W)$:

$$\cos(\theta)^3W^1_{w^1w^2} = \sin(\theta)^3W^2_{w^1w^1} \equiv \frac{1}{2} \kappa(W).$$  \(\text{(B.12)}\)

We can add the two curvature values in equations (B.10) and (B.11) and denote $A = A_1 \times A_2$ to obtain an expression for the curvature:

$$\kappa(W) = \max_{A \in A \setminus A^N} \frac{2N(g - W)}{r|\phi|^2},$$  \(\text{(B.13)}\)

which is Sannikov’s (2007) optimality equation. Note that the maximization in equation (B.13) is over both $A^1_t$ and $A^2_t$ (excluding the set of pure strategy Nash equilibria, $A^N$, as mentioned earlier). This is innocuous here because the numerator and denominator are each additively separable in $A^1_t$ and $A^2_t$, so separate maximization for each firm taking its turn as the “principal” is satisfied. ■
Appendix C  Why maximizing the Bellman equation is equivalent to maximizing a ratio in the curvature ODE

Having established that the “agency” optimization problem in equation (34) is equivalent to the ODE in equation (B.10) for the curvature of the equilibrium set boundary $\partial \mathcal{E}$, we need to show why the optimization of the Bellman equation is equivalent to the optimization of the ratio in the ODE.

Consider the abstract problem:

$$\max_x \{ f(x) + A g(x) \}.$$ (C.14)

The first-order condition is:

$$f'(x) + A g'(x) = 0,$$ (C.15)

or:

$$A = -\frac{f'}{g'}.$$ (C.16)

Now consider the maximization problem:

$$\max_x \frac{f(x)}{g(x)}.$$ (C.17)

The first-order condition can be written as:

$$\frac{f}{g} = \frac{f'}{g'},$$ (C.18)

and therefore, at the maximum, we have that:

$$\max_x \frac{f}{g} = \frac{f'}{g'}.$$ (C.19)

Therefore,

$$A = -\frac{f'}{g'} = -\max_x \frac{f}{g}.$$ (C.20)

Thus, the maximization of the ratio generates the same optimum (adjusted for the sign) as the Bellman equation. ■
Appendix D  Curvature equality

We want to show that:

\[ \cos(\theta)^3 W_{W^2W^2}^1 = \sin(\theta)^3 W_{W^1W^1}^2. \]  \hspace{1cm} (D.1)

To begin, note that:

\[ \frac{d}{dW^1} W^1_{W^2} = W^1_{W^2} \frac{dW^2}{dW^1} = -W^1_{W^2} \cos(\theta) \sin(\theta). \]  \hspace{1cm} (D.2)

This is equal to:

\[ \frac{d}{dW^1} \frac{1}{W^2_{W^1}} = -\frac{1}{(W^2_{W^1})^2} W^2_{W^1W^1} = -W^2_{W^1W^1} \left( \frac{\sin(\theta)}{\cos(\theta)} \right)^2. \]  \hspace{1cm} (D.3)

Equating the two terms and performing algebra yields the result. ■
Appendix E  Optimality of the smooth-pasting condition

In this section we prove Lemma 2 and Proposition 3, which establish optimality of the smooth-pasting condition. Our strategy differs a bit from the more well-known approaches such as Dixit (1993); we will show that if the smooth-pasting condition is satisfied then the second-order condition associated with the Bellman equation is locally satisfied at the boundary point characterised by the the smooth-pasting condition. That is, we provide a verification of the sufficiency of local optimality at the merger point. More concretely, the steps in this demonstration are as follows:

(i) Calculate the second-order condition for the Bellman equation from the optimality condition;
(ii) Evaluate the “ratio” version of the Bellman equation at the value-matching point, using the value-matching condition and also the smooth-pasting condition, resulting in a reduced-form expression for the second partial derivative $W_1^1 W_2^1 W_2^1$;
(iii) Substitute this reduced-form expression for $W_1^1 W_2^1 W_2^1$ into the second-order condition, as well as the value-matching condition, establishing that the second-order condition is negative.

Proof: (Of Lemma 2) Commencing step (i), the reprise of the Bellman equation is

$$0 = \max_{A_1} \left\{ r(g_1 - W^1) - r(g_2 - W^2) W_1^1 W_2^2 + \frac{1}{2} \left( -r W_1^2 g_1 A_1 \right)^2 W_2^1 W_2^2 + \frac{1}{2} \left( r g_2 A_1^2 \right)^2 W_1^1 W_2^2 \right\}. \quad (E.4)$$

with optimality condition

$$r g_1 A_1 - r g_2 A_1 W_1^1 W_2 + \left( -r W_2^2 g_1 A_1 \right) g_1 A_1^2 + \left( r g_2 A_1^2 \right) g_2 A_1^2 \right) W_1^1 W_2^2 = 0 \quad (E.5)$$

The second-order condition can make use of the quadratic structure of $g_1$ and $g_2$: the third derivatives are zero, so we have

$$r g_1 A_1 - r g_2 A_1 W_1^1 W_2 + r^2 \left( -W_1^2 (g_1 A_1^2)^2 + (g_2 A_1^2)^2 \right) W_1^1 W_2^2$$

Now we can substitute the smooth-pasting condition: $W_1^1 W_2^2 = -1$:

$$r g_1 A_1 + r g_2 A_1 W_1^1 W_2 + r^2 \left( -W_1^2 (g_1 A_1^2)^2 + (g_2 A_1^2)^2 \right) W_1^1 W_2^2$$

38
The same holds for the other player: $W_{W^1}^2 = -1$:

$$rg_{A^1,A^1} + rg_{A^2,A^1} + r^2 \left( (g_{A^1,A^1})^2 + (g_{A^2,A^1})^2 \right) W_{W^1W^2}^1$$

Now recall the structure of the stage game payoff function $g_1$:

$$g_1(a_1,a_2) = a_1 (\Pi_1 - \beta_1 a_1 - \delta_1 a_2) ,$$

and similarly for $g_2$, so that

$$g_{A^1,A^1} = -2\beta_1 - \delta_1 \quad g_{A^2,A^1} = -\delta_2 ,$$

which are both negative by assumption.

Carrying out step (ii), we next transform the system of Bellman equations to ratio form to isolate the second partials. Appendix Appendix C demonstrates by straightforward algebra that the maximization of the following ratio is equivalent to the original maximization in equation (34):

$$W_{W^1W^2}^1 = - \max_{A^1} \left\{ \frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r}W_{W^2}^1}{r \left( (W_{W^1A^1}^2)^2 + (g_{A^2})^2 \right)} \right\} . \quad (E.6)$$

Substituting $\frac{1}{W_{W^1}^1}$ for $W_{W^2}^2$ yields:

$$W_{W^1W^2}^1 = - \max_{A^1} \left\{ \frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r}W_{W^2}^1}{r \left( (\frac{1}{W_{W^1}^1} - g_{A^1})^2 + (g_{A^2})^2 \right)} \right\} , \quad (E.7)$$

which, along with the first-order condition in $A^1$, is an ordinary differential equation (ODE) in $W^1$. Thus, we have converted the Bellman equation from a partial to an ordinary differential equation.

Finally, carry out step (iii), substituting for $W_{W^1W^2}^1$ from the ratio reformulation of the Bellman
equation in equation (E.7), evaluated at the value-matching and smooth-pasting point:

\[
W_{W_1W_2}^1 = - \frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r}W_{W_2}^1}{r \left( \frac{1}{W_{W_2}^1 g_{1A_1}} \right)^2 + (g_{2A_2})^2} = - \frac{(g_1 - W^1) + (g_2 - W^2)}{r \left( \frac{1}{(g_{1A_1})^2 + (g_{2A_2})^2} \right)} \tag{E.8}
\]

If we can demonstrate that the numerator of this expression is negative or zero then we will have demonstrated that the second-order condition holds at the smooth-pasting point.

**Lemma 4** At the smooth-pasting point, we have that:

\[(g_1 - W^1) + (g_2 - W^2) = 0.\]

**Proof:** The expressions \((g_1 - W^1)\) and \((g_2 - W^2)\) are the drifts of the continuation value processes for \(W^1\) and \(W^2\), respectively (see equation (23)). At the smooth-pasting point, these drifts necessarily are equal and of the opposite sign in order to point along the merger line, which has a slope of \(-1\), and they therefore sum to zero. (Equivalently, the curvature at the smooth-pasting point, \(W_{W_1W_2}^1\), is zero.) □

This completes step (iii), establishing the result. □

We can use this result to prove global optimality.

**Proof:** (of Proposition 3) The optimality of the main action \(A_1^t\) as the solution of the HJB equation (34) follows directly from conventional optimal control considerations. Therefore the optimal value state path satisfies equation (B.10); further derivations lead to the differential equation (41), combined with equation (Appendix B), as demonstrated in Appendix B. Sannikov (Sannikov (2007) Theorem 2, p. 1309) establishes that any optimal path must satisfy this differential equation system.

The remaining issue is to provide two boundary conditions for the second-order differential equation, (B.10), that the optimal path necessarily satisfies; the value-matching condition (33) provides one boundary condition. By Lemma 2, the smooth-pasting condition is locally optimal. Suppose that an alternative optimal path exists that does not satisfy the smooth-pasting condition. In that case, there is a kink at the merger boundary, and therefore local optimality cannot be
satisfied. Any equilibrium path must satisfy optimality, and therefore any equilibrium path satisfies smooth pasting.

The kink is “one-sided:” the local curvature at the kink point is positive infinity, but cannot be negative infinity because the differential equation can approach the merger line from above. Therefore it is only necessary only to observe that the violation of the relevant inequality rests on the positive infinity property. □

We remark that our results have not established uniqueness of the equilibrium path. We know from numerical experiments that two manifolds that satisfy the smooth-pasting condition exist, with a smaller one fully contained within the larger one that we have analyzed here.

The “standard” proof of the optimality of the smooth-pasting point such as in Dixit (1993) uses a Taylor expansion of the solution of a second-order differential equation, exploiting the structure of the solution stemming from the assumption of fixed coefficients. Our model does not lead to an equation with fixed coefficients, so it is not possible to use Dixit’s approach directly. Notice, however, that the smooth-pasting condition is locally optimal at the smooth-pasting point; by continuity of the solution that obeys the main differential equation this optimality argument must also hold in a neighborhood of the smooth-pasting point.
Appendix F  Proof of the inclusion result, Proposition 5

Proof: The proof has three main parts.

(i) In the first part we observe that the boundary of the equilibrium set $E$, $\partial E$, is described by the same ordinary differential equation, equation (41), for any merger model with merger cost $K$ and for the no-merger case. If the inclusion were violated there would be some point $(\tilde{W}^1, \tilde{W}^2)$ at which the differential equation would necessarily hold identically for both $K$ and $K'$ at $(\tilde{W}^1, \tilde{W}^2)$. However non-inclusion would necessarily imply that the local differential would be different at $(\tilde{W}^1, \tilde{W}^2)$, a contradiction. Thus, the merger manifold must lie either entirely inside the no-merger manifold or entirely outside it.

(ii) Now consider the case where the merger manifold is entirely inside the no-merger manifold, that is, $\partial E^K_M \subset \partial E^N_M$. To begin, note that the equilibrium manifold is locally characterized by the angle $\theta$, the slope of the manifold for each $\theta$, and the curvature $\kappa$. The curvature $\kappa$ is characterized by equation (B.13), which we reproduce here:

$$\kappa(\tilde{W}) = \max_{A \in A \setminus A^N} \frac{2N(g - \tilde{W})}{r|\phi|^2} \quad (F.9)$$

As seen from equations (B.10) and (B.11), the key part of this equation is the numerator term

$$-\cos(\theta)(g_1 - \tilde{W}^1) - \sin(\theta)(g_2 - \tilde{W}^2) \quad (F.10)$$

We will use the standard convention for measuring the angle $\theta$, that is, $\theta = 0$ along the $x$-axis.

- Suppose that we solve the differential equation for the manifold with initial point at the middle of the collusion point. In that case, $\theta = \frac{\pi}{4}$, so that $\cos(\theta) = \sin(\theta)$, resulting in equal weights on the numerator term in (F.10). Given that the smooth-pasting manifold must lie inside the no-merger manifold, if we reduce $\tilde{W}_1$ and $\tilde{W}_2$, that is, reduce the initial point of the differential equation, then the absolute value of the curvature must increase. Therefore the path of the differential equation will move away from the no-merger path toward the interior.

- Next, consider $\theta = -\frac{\pi}{4}$, corresponding to the “contestability” region. In this region $\sin(\theta) = -\cos(\theta)$, so that there is a negative weight, $\sin(\theta)$, on the $(g_2 - W^2)$ term, and a positive
weight, \( \cos(\theta) \), on the \((g_1 - W^1)\) term. Now suppose we increase \( \tilde{W}_2 \) and decrease \( \tilde{W}_1 \) along the negative-45 degree line, so that the sum \( \tilde{W}^1 + \tilde{W}^2 \) is constant. In that case, the negative weight on \( \tilde{W}_1 \) and the positive weight on \( \tilde{W}^2 \) means that the net effect of the movement is to increase the curvature. Therefore as we move inward, the curvature increases.

This argument implicitly holds the actions \( A^i \) fixed; however, recall that the curvature is defined by the formula in (B.13), which has a maximization with respect to the vector \( A \) on the right hand side. Therefore, adjustments of \( A \) that are driven by our thought experiments with \( \tilde{W} \) will not affect our argument, as the effects of \( A \) drop out due to the envelope condition.

It is clear that this logic holds for all other angles: moving toward the center increases the curvature. The result is that the manifold “curls up” near the merger line, so that a tangency to the merger line becomes possible; this is apparent in the numerical simulations presented in the figures.

The remaining possibility is that the inclusion is reversed. (This possibility is illustrated in Figure 10 of Appendix G; clearly it is possible for the ODE to have a solution lying outside the no-merger manifold.)

We can see that the tangency condition cannot be satisfied locally where the no-merger boundary \( \partial E_{NM} \) crosses the merger line. By continuity of the derivative of the manifold satisfying the ODE at that point, it cannot flip over and become tangent.

As we move away from that point, the logic we developed above for the “inside” manifold, establishes that the slope continues to flatten, so it moves away from a potential tangency, and the smooth pasting condition therefore cannot be satisfied.

However, we need to rule out the possibility that the smooth pasting conditions are satisfied, but pointing the wrong way. However this would require that the curvature of the manifold boundary reverse itself, at some locus, which in turn would imply that the ODE (48) was locally linear at the locus where this reversal occurs, that is, the local second derivative would be zero. But the structure of the ODE precludes this: it would imply that the differential equation is “stuck” at the zero-curvature point (see equation (41)).

(iii) The third part of the proof is to establish that the first inclusion holds, namely that increasing the merger cost \( K \) increases the equilibrium set. But this case is similar to the previous one, in that it would require that the merger manifold would not be tangent as we decrease \( K \). □
Appendix G  A remark on social welfare

The continuation values in the model are discounted profits, so we can infer that when these profits increase, the firms are colluding to a greater extent, and consumer surplus is concomitantly reduced. Thus, social welfare is inversely related to the firms’ joint profits.

The smooth-pasting equilibrium has lower firm profits when the firms are in the collusion phase than they do if they can never merge, that is, the equilibrium manifold moves in the southwesterly direction relative to the no-merger manifold.

However, the merger cost itself is part of the deadweight loss.\(^{14}\) How do we incorporate this dead weight loss in the accounting? The answer is that we can ignore it, because the firms’ discounted profits (their continuation values) in the collusion phase are still close to the monopoly line, despite the discounted merger cost. Thus, this deadweight cost is, from the firms’ perspectives, just an alternative cost like the cost they would face if they entered into the price war in the no-merger model.

\(^{14}\)We are grateful to Mikhail Panov for raising this question.
Appendix H  Refusals to merge, punishment and jumps

In this appendix\textsuperscript{15} we explore the impact of assumptions about the structure of the underlying game, especially the details of the moment of the merger, on outcomes. We focus on one issue: the consequences of the fact that the decision to merge is observable to the rival firm. We emphasize that our discussion is informal.

One of the key assumptions in the main text is that the firms are committed to merge when the value states attain a point on the merger line. Merging at that point is a binary decision, and each firm can observe whether or not the other firm has merged with it because in the event of the merger they subsequently share monopoly profits.

Because the decision by each firm to merge is observable, if the firms do not commit to the merger in advance, the possibility that a firm would refuse to merge can also be admitted as a strategy, and the continuation of the game must be specified in that case. Because the refusal to merge is observable, the structure of the game at that moment and the associated continuation are very different from the noisy collusion game that has been played up to that point.

If one firm refuses to merge, then, given the noncooperative environment of the game, it is appropriate to consider how the other firm would subsequently punish it. One can broadly describe this punishment: it would be to revert to an equilibrium with the worst possible outcome for the firm that has refused to merge.

The punishment phase must be an equilibrium. Given that the continuation game would not differ from the collusion game in the key respect of having actions that are obscured by noise, the punishment phase would be constituted from the same elements as the pre-merger (or pre-refusal) game. If such an equilibrium exist, we refer to it as a refusal-punishment (R-P) equilibrium.

What we will establish here is that credible punishment phases cannot in fact themselves be equilibria, and so alternative equilibria that are sustained by such punishments cannot exist.

\textbf{The refusal-punishment construction}

For a refusal-punishment equilibrium to function, there must be agreement between the firms about the punishments in the punishment game. These punishments, and in essence the structure of the

\textsuperscript{15}The ideas and contributions of Mikhail Panov, Yuli Sannikov, and Andrzej Skrzypacz led to the discussion in this appendix.
continuation game, must be agreed to in a pre-play phase of the game the firms play in advance of the noisy collusion game. Thus, there are really three phases of the overall game: (i) the pre-play phase in which the punishments that will be coordinated upon in the event of a refusal to merge; (ii) the pre-merger phase of Sannikov’s noisy duopoly game; (iii) the punishment phase in the event that a firm refuses to merge, which itself must be an equilibrium of a pre-merger Sannikov game, including the potential for repeated refusal-punishment phases. Thus, in this final stage, any boundary conditions entailed by the potential for future mergers, or the lack thereof, must be delineated.

In the sequel we first examine the simplest case: that if a merger has been refused and punishment invoked, there is no further potential to merge: the punishment is permanent.

If there is no potential to merge, then the game reverts to the collusion game described by Sannikov (2007): the value states of the firms stochastically and continuously transit around an egg-shaped manifold with payoffs that are strictly bounded by the monopoly line. The initial point on the manifold is determined by the extreme punishment: if firm 1 (with value state on the horizontal axis) has refused to merge, the worst continuation on the no-merger manifold is the leftmost point on it (point $B$ in the figure); similarly if firm 2 (with value state on the vertical axis) has refused to merge, the continuation commences on the lowest point on the manifold (point $B'$ in the figure).

Punishments alter the boundary conditions of the pre-merger game. At the merger point, the punishment must be weakly worse than the worst state at the moment of the merger. Because the no-merger manifold is fixed in size and location, the leftmost and bottom-most punishment points on the manifold are fixed. This in turn dictates the corresponding upper and lower merger points on the merger line (points $A$ and $A'$ in Figure 6). Finally, because the equilibrium pre-merger manifold must be continuous and obey the underlying main differential equation, the pre-merger manifold must land on these two points, and is thus fully and uniquely determined.\footnote{Of course in equilibrium the punishments are not carried out. If they were carried out, the equilibrium manifold would then have to terminate on the punishment point, and this would in turn eliminate the original manifold from consideration.}

Because the landing points are determined by potential punishments, the smooth-pasting condition no longer determines the boundary conditions. The resulting manifold is semicircular and is not tangent at the landing points. We will refer to this manifold as the refusal-punishment (R-P) manifold. In addition to not being tangent at the merger line, the R-P manifold lies entirely between the no-merger manifold and the monopoly line. This is in contrast to the smooth-pasting
In this figure one of the firms can refuse to merge; thereafter it is assumed that no merger is possible. If a firm refuses to merge, say at point $A$, then a punishment is initiated via a jump to point $B$, with symmetric possibilities at points $A'$ and $B'$. The punishment manifold is simply the original no-merger manifold, with play evolving along this manifold as in the original Sannikov no-merger model.

There are two ingredients in this construction. The first is that the refusal and punishment are observable, unlike the pre-merger play, and also unlike the no-merger manifold that the firms jump to if the punishment is carried out.

The second is that the punishment entails a jump. This is key: in the other stages of the game, jumps are not possible because actions are confounded by noise. But the observability of the refusal means that the game is no longer in its noisy mode at that instant, and a more conventional, static, full-information game can be played. Because there is a jump if the punishment is invoked, the punishment is far more severe than the incremental punishments of the pre-merger game. As the cost shrinks, the size of the jump interval actually increases, making the jump punishments even stronger. This is why the R-P manifold achieves better cooperation than the no-merger game alone can achieve, if it is an equilibrium.
The punishments are not renegotiation-proof

The above reasoning breaks down for a clear reason: the punishment phase of the equilibrium would entail the firms traversing the punishment manifold that lies below the merger line. Every point on the punishment manifold has the property that it is Pareto-dominated by some point on the merger line. One of the firms could propose to move to such a point from the punishment manifold by merging, and the other firm could accept this proposal and improve its payoff. This logic works because the proposal to merge, unlike the production action, is publicly observable. Therefore the punishment manifold is not renegotiation proof, so in turn the punishment cannot be an equilibrium. This leaves only the manifolds that satisfy the smooth-pasting property as potential equilibria: because the smooth-pasting manifolds lie above the merger line, it is not Pareto-improving to immediately jump to the merger line. Thus, the R-P manifolds are ruled out.
References


