Does the Potential for Mergers Reduce Competition?∗

Dirk Hackbarth†  Bart Taub‡

October 5, 2017

Abstract

We study the effect of horizontal mergers in a dynamic model with noisy collusion. At each in-
stant firms either privately choose output levels or merge to create a monopoly. In equilibrium,
merging trades off the benefits of avoiding price wars against costs of merging, net of higher prof-
its due to lower competition. There are three results. First, mergers are optimal when collusion
fails (i.e., firms sufficiently deviate from a collusive regime). Second, long periods of collusion
are likely, because colluding is dynamically stable. Therefore, mergers are rare. Third, mergers
(and, in particular, lower merger costs) decrease pre-merger collusion, as punishments by price
wars are weakened. Thus, although anti-competitive mergers harm competition ex-post, 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.

∗We are grateful to Jacques Cremer, William Fuchs, Pavel Gapeev, Charles Kahn, Evgeny Lyandres, Jianjun
Miao, Mikhail Panos, Sergey Popov, Yuliy Sannikov, Andy Skrzypacz, Chester Spatt, Ajay Subramanian, Andriy
Zapchelnyuk, Mihail Zervos, and seminar participants at HSE in Moscow and the 2012 EFA Annual Meeting for useful
comments and suggestions. Bart Taub thanks ICEF, Higher School of Economics, Moscow for research support and
acknowledges financial support from the Academic Excellence Project “5-100” of the Russian Government.
†Boston University, Questrom School of Business, 595 Commonwealth Avenue, Boston, MA 02215, USA. Email:
dhackbar@bu.edu.
‡Glasgow University, Adam Smith Business School, University Avenue, Glasgow, UK. Email:
bart.taub@glasgow.ac.uk.
1 Introduction

The market power doctrine states that the concentration of output among firms in an industry is a good measure of the industry’s market power (Demsetz, 1973). The standard view is that anti-competitive mergers increase industry concentration and hence increase market power (Stigler, 1964), implying that anti-competitive mergers harm competition and hence need to be regulated.

This view fails to account for dynamics. Using a dynamic model of two tacitly colluding firms that have the potential to merge at some point in the future, we show that the prospect of merging leads them to be more competitive than if there is no potential to merge. The reason is that the potential to merge reduces the punishments that firms can mete out to each other for the perception that they have deviated from a regime of tacit collusion.\(^1\)

Our model builds on the work of Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1986) in discrete time and Sannikov (2007) in continuous time. The 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. In the model, the continuation values of the two firms satisfy a differential equation that determines an equilibrium manifold in discounted value space.\(^2\)

Our model departs from the Green-Porter and Sannikov models in that the firms can merge. A merger enables the firms to share information and coordinate outputs and thus to earn deterministic, higher profits. But because mergers are costly, they are optimal only when collusion fails sufficiently.

In a model without mergers that we construct as a benchmark, firms collude despite their inability to monitor output levels, because punishments, implemented via price wars, occur if lower prices can potentially arise from overproduction rather than noise. In equilibrium, price reductions arise due to noise only.

We then add mergers. If mergers are possible, then as in the benchmark model firms also collude before merging to avoid punishments by price wars, and in so doing avoid incurring the merger cost. However, if a price war is imminent, they choose to merge when the expected lost

\(^1\)There are no efficiency gains or scale economies from mergers in the model, so they are not the reason for the result.

\(^2\)A benefit of the continuous-time approach is that mergers are specified by a differential equation’s boundary conditions that determine both locus and shape of the equilibrium manifold. This determines the shares for each firm in the merged entity and firms’ equilibrium behavior in the pre-merger play when the firms may still tacitly collude. Sannikov and Skrzypacz (2007, 2010) use similar methods.
profits from the price war exceed the merger profits net of merger costs.

We solve for the equilibrium manifold and also for output dynamics on the equilibrium manifold. There are two aspects of dynamic behavior that are important: the first is the local stability of regions of the equilibrium manifold, and the second is the strength of that stability, which corresponds to the expected movement toward or away from a region of the equilibrium manifold. Due to the option to merge, the collusive region of the equilibrium manifold decreases in size. Collusion remains dynamically stable in the equilibrium with mergers; by contrast, the region of the equilibrium manifold close to the merger points is dynamically unstable; these two properties reinforce each other so that the dynamic optimizing behavior of the firms actually delays the onset of mergers (i.e., makes them rarer qualitatively) and results in the firms spending most of their pre-merger time successfully colluding. Mergers being rare is consistent with Andrade, Mitchell, and Stafford (2001) report 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%).

The possibility of anti-competitive mergers (even if rare) does not mean, however, that there is less competition. The potential for mergers increases competition, because they weaken potential punishments, so ex-ante they are pro-competitive. Our model thus suggests that using observed mergers as a concern for market power is not satisfactory, because mergers occur only in the rare states when collusion has failed sufficiently. Mergers may be the tip of an iceberg of collusion dynamics that might not have been fully recognized. As a consequence, it might be desirable for regulators to facilitate the ex-ante potential to merge by lowering frictions, such as barriers, costs, and legal expenses formally or informally placed by merger regulation.

The model also offers additional insights into the ways in which product market dynamics and mergers interact that are consistent with empirical patterns. First, the result that collusion is the predominant pre-merger condition in a dynamic model helps to explain the anomalous empirical finding that mergers often appear to not increase market power: the pre-merger situation is one of collusion, not competition, and this collusion comes much closer to replicating monopoly output than it does to a pure duopoly outcome. Thus, when firms merge, little changes, but that is because they have already anticipated the increased market power from merging in dynamically optimal pre-merger outputs, that is, the net gain after accounting for the merger cost is zero. In particular,

---

3Hoberg and Phillips (2010) document incidence of mergers is lower for firms that are more similar to rivals, which they call “competitive effect” because there are more firms with similar restructuring options so that the potential to merge is less valuable. In our model, this corresponds to higher merger costs and hence more potential for punishments.
the change in competition at the time of the merger is determined by the merger cost. In particular, a higher merger cost implies a higher reduction in competition at the time of the merger, because firms optimally compete longer by moving closer to the price war region if merger costs are higher.

Second, the net benefit from merging does not imply a discontinuity—a jump in value at the onset of the merger—only when the fixed cost is ignored. Firm values (or stock prices) gradually and smoothly incorporate firms’ optimal strategies to compete or to merge. So, there are no announcement returns, given that firms trade off the gains from avoiding fiercer product market competition against the gains from creating a monopoly net of the fixed cost of merging. This enables us to re-interpret the tests of the market power doctrine (Eckbo, 1983). Firms will optimally adjust their output strategies prior to the merger, and the conjectured change in profitability (or value) at the time of the merger is anticipated and hence there are no announcement returns.

Third, interpreting the firm’s continuation value state 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 the takeover announcement typically are positive whereas returns to bidding shareholders typically are negative.4 In our model negative pre-merger returns to bidding shareholders can arise simply because of the competition dynamics of the noisy collusion equilibrium.5

The rest of the paper is organized as follows. Section 2 discusses further related papers. Section 3 outlines the model, which is then solved in Section 4. Section 5 derives the model’s implications. Section 6 concludes. The appendices contain derivations, proofs and other technical results.

4 Asquith (1983) earlier documented a similar switch in the direction of return growth near announcement dates.
5 It has oftentimes been argued that when the market believes that a takeover is less likely to be value-enhancing, the acquiring firm’s stock price declines just prior to the announcement of the 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 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.
2 Related literature

The merger literature is rich and diverse, with numerous papers in both the theoretical dimension and the empirical dimension. We review papers from both literatures to provide context and contrasts with our model. Regarding the theoretical dimension, Gowrisankaran’s (1999) paper builds a dynamic model of mergers in which firms are static Cournot competitors before merging, but are capacity-constrained and can achieve larger capacity via investment and also via merging; the firms solve a dynamic game in which capacities are chosen using a dynamic programming approach.\(^6\)

Jovanovic and Braguinsky (2004) consider a model with a continuum of firms, and the managers attached to those firms, and a continuum of projects that the managers undertake. 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 the efficiency gains from acquisitions, and does not speak to the issue of market power associated with mergers, the focus of our paper.\(^7\)

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 the 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. The firms that have merger opportunities do not interact prior to this (specifically, they do not collude); thus, the optimal decision problem of the firm 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 the merger, and in the face of noise in a very strategic environment.

The paper by Bernile, Lyandres, and Zhdanov (2012) is related our paper too. They assume, as did Gowrisankaran, that 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 the demand shocks and the decisions to merge or enter. Using a real op-

\(^6\)Because the model focuses on static Cournot equilibria, firms do not install excess capacity as a punishment strategy; if such punishments were implemented, then from the perspective of our model the prospect of mergers would reduce collusion in capacity due to the weakening of punishments.

\(^7\)Bernile and Lyandres (2015) find indeed operating efficiencies also influence announcement returns of mergers.
tions 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 *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 allows firms to enforce collusion via punishments in the spirit of our model.

Finally, the article by Hackbarth and Miao (2012) also has 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 value of a firm in the post-merger environment, taking account of the reduced number of firms and their sizes 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 dynamically collude.

Regarding the empirical dimension, a number of articles perform event studies. Eckbo (1983), Stillman (1983), and Eckbo and Wier (1985) initiate studying 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 diver-
sified firms (McAfee and Williams, 1988) or the deterrent effect of antitrust enforcement (Prager, 1992). In contrast to these papers, this paper shows that once the cost of merging is taken into account too, 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 larger for larger merger costs.

The paper of Ashenfelter, Hosken, and Weinberg (2015), summarized in the article of Ashenfelter, Hosken, and Weinberg (2013a), finds evidence for price increases due to the merger of two large US beer brewing companies, Miller and Coors. They identify production efficiencies from the merger as stemming mainly from trucking distances from the breweries and can thus extract this effect when measuring price increases stemming from the increased market power flowing from the merger. They argue that efficiency gains will result in slow price increases, and this is what they observe. They also make a case for market power price increases but these seem small (on the order of a few per cent); our model also predicts that there will be price increases, but that they should be small.\textsuperscript{8}

The paper of Ashenfelter, Hosken, and Weinberg (2013b) also brings empirical evidence to bear on mergers. This paper looks at the household appliance maker Maytag’s acquisition by Whirlpool in 2005. This was in response to an attempted purchase by a Chinese firm. The main object of intuitive interest in the paper is Figure 3: using ovens as a benchmark, because ovens face a competitive market, they look at price increases of other appliances, such as dishwashers, dryers, and clothes washers where the merged firm had a greater market share, relative to ovens subsequent to the merger, but starting at the merger announcement in 2005. They find that for clothes washers there is a substantial price increase (\( \approx 40\% \)), but the price increase spans three years, gradually. It is thus difficult to discern whether this slow price increase is due to other effects; specifically, their study period straddles the financial crisis in 2008. Because their study period commences at the merger announcement date, the paper cannot compare pre-merger collusion and post-merger pricing.

3 Model

Our dynamic model with noisy collusion builds on Sannikov’s (2007) model applied to a duopoly. The key idea in our approach is that each firm behaves as a principal in an agency construct, treating the other firm as its agent, with the continuation value of the agent as the state variable.

\textsuperscript{8}To be more precise, we predict that the change in joint valuation will be zero in the purest version of the model, but output prices increase. This follows because the merger is a transition from duopoly to monopoly, resulting in aggregate output falling and output prices increasing, but the concomitant rise in discounted profit only just compensates for the cost of the merger.
for the principal; correspondingly, each firm at the same time behaves as an agent reacting to what is effectively a contract set out by the other firm.\textsuperscript{9}

The agency construction allows us to characterize the equilibrium via the firms’ continuation values mapping out a one-dimensional manifold, \( \mathcal{E} \), in the plane defined by the continuation values, that comprises the largest equilibrium manifold, just as in Sannikov’s model. Sannikov did not emphasize that the equilibrium characterized by this manifold is Markovian, that is, the movement along the manifold depends only on the current state of the continuation values.

This Markovian property means that the equilibrium state can be mapped one-to-one to the equilibrium vector of public information \((X^1, X^2)\) that is driven by the output decisions of the firms and by noise. We begin by treating this information as a state variable for the firms, and then carrying out a transformation or mapping of this state to the space of continuation values using stochastic calculus by way of our agency construction. This then maps out the equilibrium manifold \( \mathcal{E} \) via Sannikov’s main differential equation (Sannikov (2007), equation (24), p. 1309).

We add two elements to Sannikov’s construction. The first is our treatment of “enforcement,” that is, the restriction on the matrix of coefficients of the reduced-form noise processes that are the shocks to the continuation values along the equilibrium manifold. For a coherent equilibrium, the matrix of coefficients must be singular so that firms coordinate so as to move on the equilibrium manifold—this is what yields the “self-generating” property of the equilibrium manifold. The singularity property is essential for our refinement of the continuation values to determine the manifold.

The second additional element concerns the boundary conditions. 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 stopping time to be \emph{optimal}. However because both firms must choose the same \emph{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.

We remark that our solution method—informational state variables transformed to a continu-

\textsuperscript{9}Solving a dynamic contract via agents’ continuation values as state variables is an established solution technique in the dynamic contracting literature; see Miao and Zhang (2015) for a summary of the literature.
ation value state vector using stochastic calculus and an agency framework—has the potential be applied to other versions of the continuous-time Abreu, Pearce and Stacchetti model. Sannikov’s continuous-time approach made an intuitive, geometric interpretation of Abreu, Pearce and Stacchetti’s model possible. Aside from our direct interest in modeling mergers we also note that our approach makes further applications of their model tractable and practical.

3.1 Actions

In every stage game, there are two firms that take private actions and see public signals, which depend on their actions. Brownian motion distorts public signals, so that firms face an imperfect monitoring problem with regard to each other’s actions. Private action and public signals determine the firms’ payoffs (i.e., discounted values in a dynamic game). The stage game is played continuously at each moment in time $t \in [0, \infty)$. That is, time is continuous, and uncertainty is modeled by a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\{\mathcal{F}_t\}$.

We consider two firms that compete in an industry with differentiated products that are imperfect substitutes; i.e., a duopoly with imperfectly observable actions where two firms continuously take private actions (outputs) $A_t = (A^1_t, A^2_t)$. That is, firm $i = 1, 2$ chooses an action $A^i_t \in \mathcal{A}_i \subseteq \mathbb{R}_+$ for all $t \in [0, \infty)$.\footnote{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 solution will follow from an ordinary differential equation that characterizes the boundary of the equilibrium set, $\mathcal{E}(r)$, i.e., the continuation value pairs achievable by all public perfect equilibria of the continuous-time game.

The firms observe the history of a vector of public signals (i.e., price increments) $dP_t$, which depend on their actions $A^1_t$ and $A^2_t$. The instantaneous prices of firms 1 and 2 before and after the merger are given by the levels (not increments) of the processes:

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

(1)

and

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

(2)

Because the products are imperfect substitutes the actions of both firms influence the prices of both outputs.

The instantaneous payoff functions of the firms are the product of output and price increments: $g_i(\cdot) = A^i dP^i$ for $i = 1, 2$. Using lower case letters to denote choices of $A^i$ and realizations $dP^i,$
we express the payoffs as:

\[ g_1(a_1, a_2) = a_1 dp^1 = a_1 (\Pi_1 - \beta_1 a_1 - \delta_1 a_2), \tag{3} \]

and

\[ g_2(a_1, a_2) = a_2 dp^2 = a_2 (\Pi_2 - \delta_2 a_1 - \beta_2 a_2). \tag{4} \]

Discounted profits are integrals of these instantaneous profits.

### 3.2 Information

Firms learn about each other’s actions from observing prices; the linearity of the prices in (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 \( X_t \):

\[
dX^1_t = \frac{(A^1_t + A^2_t) dP^2_t}{2(\Pi_2 - \delta_2 A^1_t - \beta_2 A^2_t)} - \frac{(A^1_t - A^2_t) dP^1_t}{2(\Pi_1 - \beta_1 A^1_t - \delta_1 A^2_t)} = A^1_t dt + \sigma_1 dZ^1_t, \tag{5} 
\]

and

\[
dX^2_t = \frac{(A^2_t + A^1_t) dP^1_t}{2(\Pi_2 - \delta_2 A^1_t - \beta_2 A^2_t)} - \frac{(A^2_t - A^1_t) dP^2_t}{2(\Pi_1 - \beta_1 A^1_t - \delta_1 A^2_t)} = A^2_t dt + \sigma_2 dZ^2_t, \tag{6} 
\]

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 \( X_t \).

It can make sense to consider equilibria that are measurable with respect to the history of public signals \( (X_s)_{s \leq t} \), because this would then reflect the history of actions \( (a_s)_{s \leq t} \) that are payoff relevant; the current value of the public information \( X_t \) is not payoff-relevant per se.

We go a step further and consider only equilibria in which \( X_t \) is the state, which will map one-to-one onto the vector of value states. These equilibria include those that are characterized by the manifold that is the boundary of the set of all possible equilibria, and which is itself an equilibrium—the same manifold found by Sannikov in his example. In those equilibria, we derive an equivalence between \( X_t \) and the vector of continuation value states \( W_t \).

To begin, each firm will be modeled as facing the same vector of public information, \( X_t \), as its state variable, but with only one control, namely its own output \( a_it \); the other firm’s output history, \( (a_{-i,s})_{s \leq t} \) is taken as an exogenous stochastic process. We then map the continuation value process of the rival firm into a single state variable for the firm; because both firms undergo this
mapping, there is still effectively a vector of equilibrium states, $W_t$, that maps one-to-one with the informational state $X_t$.

### 3.3 Duopoly stage game

As in many other dynamic settings, there might be solutions to the repeated game that are simply the repetition of the Nash equilibrium of the stage game (see, e.g., Hackbarth and Miao (2012)). To focus on the cooperative equilibria, we follow Sannikov (2007) and rule out the stage game Nash equilibrium strategies. This in turn requires that we calculate those equilibria. The optimal full information Nash (duopoly) solutions are:

$$a_1^* = \frac{\delta_1 \Pi_2 - 2 \beta_2 \Pi_1}{\delta_1 \delta_2 - 4 \beta_1 \beta_2},$$  

and

$$a_2^* = \frac{\delta_2 \Pi_1 - 2 \beta_1 \Pi_2}{\delta_1 \delta_2 - 4 \beta_1 \beta_2}.$$  

Substituting equations (7) and (8) into equations (3) and (4) yields the following flow payoff functions:

$$\pi_{t,1}^d = r \int_{t}^{\infty} e^{-r(s-t)} g_1(a_1^*, a_2^*) ds = \frac{\beta_1 (\delta_1 \Pi_2 - 2 \beta_2 \Pi_1)^2}{(\delta_1 \delta_2 - 4 \beta_1 \beta_2)^2},$$  

and

$$\pi_{t,2}^d = r \int_{t}^{\infty} e^{-r(s-t)} g_2(a_1^*, a_2^*) ds = \frac{\beta_2 (\delta_2 \Pi_1 - 2 \beta_1 \Pi_2)^2}{(\delta_1 \delta_2 - 4 \beta_1 \beta_2)^2}.$$  

Anticipating, these will be the payoffs from the only pure strategy Nash equilibrium in the set $A^N$, which needs to be ruled out in equation (B.13).

### 3.4 Monopoly stage game

When firms merge, they open their doors to each other informationally and merge their managements. They therefore 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 prepare for this we must 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 monopoly solution applies:

$$a_1^* = \frac{(\delta_1 + \delta_2) \Pi_2 - 2 \beta_2 \Pi_1}{(\delta_1 + \delta_2)^2 - 4 \beta_1 \beta_2},$$
and

\[ a_2^* = \frac{(\delta_1 + \delta_2) \Pi_1 - 2 \beta_1 \Pi_2}{(\delta_1 + \delta_2)^2 - 4 \beta_1 \beta_2}. \]  

(12)

This implies the following flow payoff function starting from any time \( t \) after the merger, which is an exogenous or non-strategic outside option:

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

(13)

i.e., it will only be shared between the firms – based on pre-merger equilibrium dynamics – when they decide to end the dynamic game by agreeing to the terms of a merger.

### 3.5 The sequence of play

A fundamental property of our model is that firms gain from delaying the merger. The merger entails fixed costs: these include substantial legal fees necessary to obtain regulatory approval, 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: when two airlines merge, one of the firms’ aircraft will need to be repainted.

Pre-merger collusion delays incurring these costs. Thus, firms would be interested in agreeing to an interim period in which they collude, conditional on their discounted profits from collusion dominating the profits available from an immediate merger. However, because they have not yet merged, they cannot observe each other’s actions, so they must play a noisy collusion game prior to merging.

The firms are aware of the game and understand that they will eventually merge, and play the pre-merger game with this understanding. This results in an equilibrium characterized by a one-dimensional manifold in the space of continuation values that the firms stochastically traverse until they merge. Characterizing this manifold will be the main analytical task.

To give a preview of this construction, we first note that the merger results in the transition to monopoly, but the firms need to agree on a division of the monopoly profits between the shareholders of the two firms when they merge. The different potential divisions of the profits can be expressed as a line in the space of continuation values of the two firms; we will refer to this line as the merger line. We make no assumption that there is a pre-agreed split such as 50-50; rather, the split is itself endogenous to the equilibrium of the game, treating the merger line as a boundary.
The firms also need to agree on the duration of the collusion period before the merger; the merger will occur at some time that is a stopping time which must be the same for both firms, even though each firm must independently agree to merge at the same time as the other firm; we assume that each firm chooses an optimal stopping time, and that in equilibrium these stopping times exactly coincide.

We model the potential and decision to merge as follows: in addition to being able to decide on output at each moment of the pre-merger collusion period, we assume that each firm can at any instant announce its willingness to merge at that moment, but the merger does not take place until and unless both firms simultaneously announce their willingness to merge; once this happens the merger occurs and is irreversible. We also assume that the firms announce their willingness to merge when it is weakly optimal to merge. Thus, we don’t consider situations when the firms are willing to merge but can then back away from the merger.

Finally, 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 G, showing that the needed punishments would not be renegotiation-proof and so cannot support equilibria.

### 3.6 The merger

We are now ready to conceptually discuss the merger. At the time of the merger, the joint continuation values reach a threshold at which the value of joint value collusion and the value of full-information monopoly are the same, net of the fixed cost of merging. In addition, there must be no marginal benefit to either firm of deviating from the continuation both before the merger and also at the time of the merger.

The merger will take 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 they be equal:

$$T_m^1 = T_m^2 = T_m$$

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 (13), \(\xi^i(\cdot)\) is firm \(i\)'s share in the monopoly as a function 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 (15) says that the merger enables the two firms to shift away from the dynamic duopoly payoffs to sharing monopoly payoffs net of the transaction costs of merging.\(^{11}\)

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 is given by

\[
W^i(X^i_{T_m}) = \xi^i(X^i_{T_m})(\pi_m - k),
\]

Notice that we treat the stopping time as specific to the firm at this stage of the model construction. It is also important to take note that although we will ultimately demonstrate that a value matching condition applies to the merger, it is conceptually possible for the merger to take place via a jump, that is, the discounted value of the firm just prior to the merger might be different from the value of the merger. However this cannot be the case in equilibrium play, because the firms anticipate this jump, and so the discounted value of the firm must approach the merged value as the stopping time gets is approached.

### 3.6.1 Pre-merger continuation value

Define the pre-merger continuation value \(W^i_t(\cdot)\) as the mapping, \(W^i: \mathcal{X} \to \mathbb{R}_+\), from the state vector \(X_t\) to firm \(i\)'s time \(t\) payoff of the continuous-time game. Given the boundary condition (17), this yields for all \(t < T_{m}^i\) the following continuation value function for firm \(i\):

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

taking as given the rival firm’s action process \(A^i_t\), where \(\mathbb{E}[\cdot|\cdot]\) denotes the conditional expectation operator at time \(t\).\(^{12}\) Note again that each firm chooses its own \(T_{m}^i\), and therefore it must be

\(^{11}\) In other words, the merger does not have any other synergies, such as better cost efficiency, economies of scale, efficient capital reallocation, etc.

\(^{12}\) By taking the rival’s action \(A^2_t\) as given, firm 1 views \(A^2_t\) as a drift function of the history of the signals \(X_{it}\), 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.
established that in equilibrium the two stopping times are equal, that is, condition (14) holds in equilibrium.

3.7 The formal problem and definition of equilibrium

The problem we consider is to find the maximal set of payoffs attainable in equilibrium of the repeated game between the two firms, subject to the constraint that continuation values of players can never go below the merger line. The reason for this is that continuing to play the collusion equilibrium nets the firms more profit than merging due to the fixed cost of merging, in that 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 doesn’t 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 strategy of the other firm.

More formally, we can define the game as follows.

**Definition 1** A Markov merger game is repeated game as described on page 1292 of Sannikov (2007) with stage game that is a tuple \( \{ N, (A_i)_{i \in N}, (g_i)_{i \in N} \} \), where \( N \) is the set of players, \( (A_i)_{i \in N} \) is the action set of player \( i \), and \( g_i \) is the payoff of player \( i \), and in addition a tuple

\[
\{ T^1_m, T^2_m, \tilde{\xi}^1(\cdot, \cdot), \tilde{\xi}^2(\cdot, \cdot), S^1(\cdot), S^2(\cdot) \},
\]

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 \( \tilde{\xi}(W^1_t, W^2_t) \)” \} and where \( T^1_m \) and \( T^2_m \) are the stopping times defined as the first time the signal \( S^i_t = “merge” \) 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 \( \tilde{\xi}^1(\cdot, \cdot) \) and \( \tilde{\xi}^2(\cdot, \cdot) \).

Thus, a Markov merger game behaves just like Sannikov’s game in the run-up to the merger, during which time the firms can be thought of as sending the “do not merge” signal (or at least one of them). 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 note as an aside that the game is expressed notationally in terms of the sharing rule \( \tilde{\xi} \), which is a function of the continuation values \( W^1_t \) and \( W^2_t \), which is different from the sharing rule
\( \xi \) that was defined in equation (17), which is a function of the state vector \( X_t \). The equivalence of \( \xi \) and \( \tilde{\xi} \) follows from the one-to-one nature of the mapping between the state vector \( X_t \) and the vector of continuation values \( W_t \).

We now turn to the definition of equilibrium.

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

- an action profile \( (A^1, A^2) \) and a sharing function pair \((\tilde{\xi}^1, \xi^2)\) that satisfy: the action profile \( (A^1, A^2) \) is a PPE as defined in Sannikov (2007), that is, “A profile \( A = (A^1, A^2) \) is a PPE if, for \( i \in 1, 2 \) the profile \( A^i \) maximizes the expected discounted payoff of player \( i \) given the strategy \( A^{-i} \) of the opposing player after all public histories,” where the payoff is defined in equation (18), and in addition,

- firms merge only if both firms simultaneously play “agree to merge” with shares defined by \( \tilde{\xi}(W^1, W^2) \);

- the stopping times \( T^1_m \) and \( T^2_m \) are optimal for firm 1 and firm 2 respectively;

- the stopping times are identical, that is \( T^1_m = T^2_m = T_m \);

- the sharing rules are feasible, that is, they satisfy

\[
\tilde{\xi}^2(W^1(X_{T_m}), W^2(X_{T_m})) = 1 - \tilde{\xi}^1(W^1(X_{T_m}), W^2(X_{T_m})).
\]

We seek to characterize the largest equilibrium set such that the sharing rule is characterized by the merger line,

\[
W^2 = (\pi_M - k) - W^1;
\]

It is far from clear how to achieve the key requirement of equilibrium that 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. Of course the smooth pasting condition is a characteristic of the solution, so we provide a verification of its optimality.
Because there are two firms, it is not completely obvious how to specify the smooth-pasting condition however. Much of the subsequent technical development will be devoted to transforming the optimization problem in equation (18), so that the smooth-pasting condition can be clearly expressed. Once this solution is in place, we will establish that the smooth-pasting condition ensures (local) incentive compatibility at the merger point. That is, satisfying the smooth-pasting condition means that neither firm wants to locally deviate from the merger equilibrium. We then analyze the solution numerically.

4 Solution

In this section we derive the solution to the dynamic game using stochastic calculus, in two stages. Our stochastic calculus approach differs from Sannikov’s 2007 approach, which uses a geometric approach to derive continuation values, firms’ strategies, and the equilibrium manifold. Our approach provides a practical recipe for solving such models, and facilitates a direct economic interpretation of the boundary conditions.

We can break down our solution procedure into the following steps:

Stage 1 (i) Solve the conventional profit maximization problem of each firm, taking the other firm’s action profile as given, using the public signal vector $X_t$ as the state vector;

(ii) Use the solution of the firm optimization to state the continuation value process for each firm in terms of the state vector $X_t$;

(iii) Use the requirement of local promise keeping or “enforcement,” which necessitates a singular volatility matrix, to restrict the structure of the continuation value processes;

(iv) Implicitly map the state vector $X_t$ into the continuation value vector $W_t$ using calculus arguments, so that the continuation value process of a firm is implicitly expressed as a function of the rival firm’s continuation value, that is, implicitly construct the mapping

$\mathcal{M}: \mathcal{X} \to \mathcal{R}$

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

(ii) Solve the principal’s optimal stopping problem using value matching and smooth pasting boundary conditions and verification of the optimality of smooth pasting;
(iii) Verification of equilibrium, namely Sannikov’s equation and also the simultaneity of merger decision.

The $X_t$ public information vector is a two-dimensional state that is influenced via the drift function by the actions of both firms. The end of our first stage of the solution procedure maps this two-dimensional state into a two-dimensional state of continuation values. We then consider the manifold defining the largest self-generating equilibrium set; as in Sannikov’s model, this manifold is one-dimensional; from the perspective of each firm, its problem reduces to having a one-dimensional state vector, namely the continuation value of the other firm. This is what allows us to use the smooth pasting condition as a way to characterize the stopping time, and also to ensure that the optimal stopping times of the two firms are in fact the same stopping time. Also, the dimension of the state vector that characterizes the equilibrium is preserved, that is, $W_t$ and $X_t$ have the same dimension, namely 2.

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

Part of our solution strategy will be to posit that the public signal vector $X_t$ comprises the state vector for the value functions of the firms.

The processes $X^1_t$ and $X^2_t$ are obviously states that are themselves functions of the whole histories, and that are controlled by the actions of the players. The current state $X_t$ can be attained via many different paths of actions, that is, histories; we are implicitly asserting that the only thing that matters is the current state, and thus that all those histories lead to the same value. This does not obviously hold generally.

In our formulation, the state vector $X_t$ summarizes the particular histories that lead to the equilibrium we are interested in, namely the largest equilibrium manifold, and for these histories, $X_t$ fully characterizes these histories.

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:

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

where, for brevity’s sake, firm 1’s maximization over $A^1$ suppresses the exclusion of its pure strategy Nash equilibrium in equation (7). 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 equation:

$$rW^1(X^1, X^2) = \max_{A^1} \left\{ rg_1(A^1, A^2) + A^1W_{X^1}^1(X^1, X^2) + A^2W_{X^2}^1(X^1, X^2) + \frac{1}{2}\sigma_1^2W_{X^1X^1}^1(X^1, X^2) + \frac{1}{2}\sigma_2^2W_{X^2X^2}^1(X^1, X^2) \right\}, \quad (22)$$

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:

$$rg_1 + W_{X^1}^1(X^1, X^2) = 0. \quad (23)$$

We now make use of the envelope condition, that is, the substitution of the optimality condition (23) into the Bellman equation (22), in combination with Ito’s lemma to generate the stochastic process of the firm’s value state.

**Lemma 1** The firms’ value states follow the processes

$$dW^1 = r(W^1 - g_1)dt - \sigma_1rg_1A^1dZ^1_t + \sigma_2W_{X^2}^1dZ^2_t. \quad (24)$$

and

$$dW^2 = r(W^2 - g_2)dt - \sigma_1W_{X^1}^2dZ^1_t + \sigma_2rg_2A^2dZ^2_t. \quad (25)$$

The proof uses standard ideas and is in Appendix Appendix A.

Combining the two firms yields the vector process:

$$dW_t = r(W_t - g(A_t))dt + \begin{pmatrix} -\sigma_1rg_1A^1 \\ \sigma_1W_{X^2}^1 \\ -\sigma_2rg_2A^2 \\ -\sigma_2W_{X^2}^2 \end{pmatrix} dZ_t, \quad (26)$$

where \(r(W_t - g(A_t))\) is the drift of the process. Denoting the volatility matrix by \(B\), we can write this more compactly as,

$$dW_t = r(W_t - g(A_t))dt + B_t dZ_t. \quad (27)$$

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

Our next step will be to transform this system and that we will do so by manipulating the partial derivatives \(W_{X^2}\).

Note that the marginal value of the firm’s state with respect to the other firm’s signal, \(W_{X^2}^i\), ends up determining the volatility. This is what Sannikov calls enforcement of equilibrium actions.
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 (27). We next turn to this element of the model. We will 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 implicitly we obtain $W^1$ as a 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 $W^1$.

4.1.2 Characterizing the volatility matrix with calculus arguments

The volatility matrix $B_t$ contains cross-partial derivatives that we can partially characterize.

**Proposition 1** In a Markovian equilibrium, the volatility matrix is singular.

**Proof:** The enforcement property must hold along the equilibrium path. This translates into the direction of movement of the Markovian state vector being restricted to be tangential to the equilibrium manifold. This in turn requires compressing the stochastic shocks into a single effective shock, with the direction of motion of that shock tangential to the manifold. This in turn requires a singular volatility matrix, such that the eigenvector of the subspace defined by the singular matrix is tangential to the manifold. The only detail is establishing that the tangency holds at each point of the manifold. But this is by construction: see equation (25) and later; the volatility matrix depends on the local state. ■

Because the volatility matrix is singular, the largest equilibrium set is a one-dimensional manifold in the space of continuation values with elements $(W^1, W^2)$.\(^{13}\) The slope of that manifold at any point is

$$\frac{dW^2}{dW^1}$$

(28)

This derivative can be expressed in two distinct ways using implicit differentiation:

$$\frac{dW^2}{dW^1} = \frac{W_{X^2}^2}{W_{X^1}^1} = \frac{W_{X^2}^2}{W_{X^1}^1}$$

(29)

\(^{13}\)We use the term manifold advisedly: technically speaking a manifold is a closed curve, but we believe the term is unambiguous in this setting.
Using the first equality we can write

\[ W_{X_1}^2 = W_{W_1}^2 W_{X_1}^1 \]  

(30)

Two of the derivatives, \( W_{X_1}^2 \) and \( W_{X_2}^1 \), appear in the volatility matrix \( B \) in equation (27), and the other two appear in the optimality conditions (see, e.g., equation (23)). Therefore, we can write:

\[ W_{X_1}^2 - g_{1A^1} A_1 = \frac{-g_{2A^2}}{W_{X_2}^1}. \]  

(31)

Because the volatility matrix must be singular, any joint evolution of the noise processes is then forced to move the equilibrium point along this lower-dimensional manifold. Therefore we can also assert that the determinant of the volatility matrix must be zero, which means:

\[ g_{1A^1} g_{2A^2} = W_{X_2}^1 W_{X_1}^2. \]  

(32)

which is equivalent to the previous equation.

4.2 Carrying out the transformation/mapping \( \mathcal{M} : X_t \mapsto W_t \)

In the next subsection we carry out the second stage of the solution. We will write down a new optimization for firm 1 in which the continuation value \( W^2 \) is a state variable for firm 1. We therefore will need to state the partial derivatives of \( W^1 \) with respect to the state variable \( W^2 \) in the Bellman equation. Substituting from the optimality condition (23) into equation (30) results in the transformations

\[ W_{X_1}^2 = -r W_{W_1}^2 g_{1A^1} = -r \frac{1}{W_{W_2}^1} g_{1A^1}, \]  

(33)

and similarly

\[ W_{X_2}^1 = -r W_{W_2}^1 g_{2A^2} \]  

(34)

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

We draw attention to the fact that equations (33) and (34) 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 \( X^2 \). Rather than develop a new and separate notation for these functions, we retain the notation \( W^i \) in the sequel, with the arguments of the functions evident from context.

4.3 The second-stage agency reformulation

We now re-solve the optimization problems of the firms, this time with each firm treating the other as an agent. Because the equilibrium movement of the continuation values is now restricted to a
one-dimensional manifold, it is possible to model each firm as having only one state variable along the manifold! In an explicit agency formulation, there would be an incentive constraint, which would simply express the requirement that the agent optimize in the face of the contract. But we have already solved that optimization problem, and so we treat the optimized continuation value simply as a state variable, with its dynamics already stated in equation (25). Once this is done we can express the problem as a standard optimal stopping problem that can be solved with a standard HJB equation and a smooth pasting condition. We can then establish that the equilibrium consists of the agency solutions simultaneously and consistently holding for both firms. In this section we solve the optimization problem and in the next section we construct the equilibrium.  

4.3.1 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 $X_i^t$ as an argument by substituting from equation (33) into equation (25), we can re-state the continuation value process for firm 2 in terms of $W^2_t$ as follows:

$$dW^2 = r(W^2 - g_2)dt - rg_2A^zdZ^2_t + (-r)\frac{1}{W^1_t}g_1A^1dZ^1_t.$$  

(35)

4.3.2 The merger boundary from the agency perspective

Define the share function $\tilde{\xi}(\cdot, \cdot)$ as a function of the continuation values, noting that we use separate notation to distinguish the function from the earlier function $\xi(\cdot, \cdot)$, as follows. First, we note that the two functions must be equivalent in the sense that

$$\tilde{\xi}(W^1(X^1, X^2), W^2(X^1, X^2)) = \xi(X^1, X^2)$$

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

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

(36)

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

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

(37)

\footnote{By using the already-solved value process to capture the agency approach, we are in essence using the “first-order” approach to agency—that is, using first order conditions. But we need to show that this is valid later on when we demonstrate that the smooth pasting condition holds, and that as a consequence the second order condition is also satisfied when each firm acts as the “agent.”}
4.3.3 The agency contract objective

Observe that the objective for firm 1 in equation (18), which satisfies the value-matching condition (17), expresses the sharing rule in terms of the state vector, \((X_{T_m}^1, X_{T_m}^2)\) at the stopping time \(T_m\).

In the second-stage “agency” setup, with the state \(W^2\) for firm 1 and vice versa, this becomes:

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

(38)

taking as given the other firm’s control process \(A^2\).

It is possible to separate the problem of finding the optimal action \(A^1\) from the optimal stopping problem. To solve the optimal action problem we state the Hamilton-Jacobi-Bellman equation; the optimal stopping problem is then solved, using the optimal policy \(A^1\), via a boundary condition, namely the smooth pasting condition.

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 (35). The full statement of the HJB equation is then

\[
\max_{A^1} \left\{ r(g_1 - W^1) - r(g_2 - W^2)W^1_{W^2} + \frac{1}{2} \left( -rW^2 \cdot g_1 A^1 \right)^2 W^1_{W^2W^2} + \frac{1}{2} \left( rg_2 A^2 \right)^2 W^1_{W^2W^2} \right\}.
\]

(39)

with state process

\[
dW^2 = r(W^2 - g_2) dt - r g_2 A^2 dZ^2_t + (-r) \frac{1}{W^2} g_1 A^1 dZ^1_t.
\]

(40)

and with the boundary condition

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

(41)

Thus, everything is expressed in terms of the state \(W^2\). There is a similar expression for firm 2.

We summarize with the following proposition:

**Proposition 2**

(i) The two-step procedure in which the optimal continuation values \(W^i(X^i, X^2)\) as a function of the publicly observable signals reflects optimization with regard to the firm’s action \(a^i_t\) at time \(t\), followed by the transformation to the agency problem, satisfies incentive compatibility, as defined in Sannikov (2007), p. 1298 Proposition 2.

Proof: (i) Incentive compatibility is equivalent to optimization in the face of the constraints imposed by the equilibrium. The optimization is explicit here, namely the use of the first order condition to alter the continuation value; we just need to demonstrate that this is also equivalent to optimization relative to the second-stage state variables, \((W^1, W^2)\). But this follows from going backward from the second stage to the first stage, and noting that each firm optimizes its Bellman equation as reflected in equations (22) (and the equivalent for firm 2) and the optimality conditions (23) (and the equivalent for firm 2).

(ii) Promise-keeping follows from examining the differential form of equation (5) of Sannikov (2007), p. 1296 (without public randomization); this is then our equation (24) and (25).

4.3.4 Optimality: the smooth-pasting condition

There are three main conditions that comprise the solution of the optimization problem stated in (39)-(41). The first condition is the standard optimality condition: maximise (39) with respect to \(A^1\). However, rather than calculating this derivative explicitly, we will convert the problem into an equivalent maximisation problem (see equation (E.7)) and solve that problem instead.

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.

\[
W^1(W^2_{\mathcal{T}_m}) = \tilde{\xi}^1_{\mathcal{T}_m}(W^1_{\mathcal{T}_m}, W^2_{\mathcal{T}_m})(\pi_m - k)
\] (42)

The third condition is the smooth pasting condition. The smooth-pasting condition for firm 1 follows from differentiating the boundary function (37):

\[
W^1_{W^2}(W^2_{\mathcal{T}_m}) = \tilde{\xi}^1_{W^2_{\mathcal{T}_m}}(\pi_m - k) = -1,
\] (43)

with a similar condition for firm 2. (We again draw attention to our notation: \(W^2_{\mathcal{T}_m}\), viewed by firm 1 as a state variable, denotes firm 2’s continuation value evaluated at the stopping time \(\mathcal{T}_m\), whilst \(W^3_{W^2_{\mathcal{T}_m}}\) denotes the partial derivative of firm 1’s continuation value as a function of that state at the stopping time.)

We begin with a lemma about the smooth pasting condition. We show that the smooth pasting condition locally satisfies the second order condition for the firm solving the agency problem.

Lemma 2 The smooth pasting condition is necessary for a local optimum with respect to the action \(A^1\) at the merger point.
Proof: See Appendix Appendix E. ■

To recapitulate, we first solved each firm’s optimization in the face of the other firm’s fixed policy. One can interpret this as the firm optimizing in the face of a contract. This optimization yielded a value process for the firm. The rival firm also generates a continuation value process. The first firm now solves a second stage problem in which the rival firm’s value process is a state variable; this is equivalent to the firm optimizing the contract and imposing an incentive compatibility constraint on the rival firm. That second-stage optimization results in a Bellman equation that can then be re-stated as above.

Proposition 3 The smooth pasting condition implies equal stopping times: \( T_m^1 = T_m^2 \).

Proof: This is a key result, as it underscores the usefulness of the agency two-stage solution procedure. The proof follows from two observations. First, in order 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

\[
W_{1z}^1(W_{T_m}^2) = -1, \tag{44}
\]

for firm 1; inverting the equation yields

\[
W_{1z}^2(W_{T_m}^1) = -1, \tag{45}
\]

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 (43) 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 (43), 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 4** The action $A^i(\cdot)$ and $W^i(\cdot)$ that solve (39), (41), and smooth pasting condition (43), taking as given $A^{-i} \in \mathcal{A}$, solve the optimal action and stopping problem (38).

**Proof:** See Appendix Appendix E. ■

### 4.4 Equilibrium and characterization

An equilibrium of the game requires that firms choose optimal contracts in their role as principals, optimally react to the other firm’s contract in their role as agents, and that the contracts are identical; furthermore, they must agree on an identical stopping time. The agency approach allows 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 Appendix Appendix G.

We then have

**Proposition 5** The value functions that solve (39)-(43) for both firms simultaneously are a Markov Merger Equilibrium.

**Proof:** This is a direct consequence of satisfying Sannikov’s main differential equation, equation (14) of Sannikov (2007), and the smooth pasting condition, (43). ■
The reason for converting the optimization problem into the form in equation (E.7) is 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.

4.4.1 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. We follow Sannikov (2007) and 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:

\[ N(\theta) = (\cos(\theta), \sin(\theta)), \quad T(\theta) = (-\sin(\theta), \cos(\theta)). \] (46)

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

\[
\frac{d\theta}{d\ell} = \kappa, \quad \text{and} \quad \frac{d\ell}{d\theta} = \frac{1}{\kappa},
\] (47)

along the manifold. 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 main differential equation we use for our numerical simulations:

\[
\left( \frac{dW^1(\theta)}{d\theta} \right) = \left( -\frac{\sin(\theta)}{\kappa(\theta)} \right), \quad \frac{dW^2(\theta)}{d\theta} = \frac{T(\theta)}{\kappa(\theta)}. \]

We solve this ordinary differential equation numerically to determine the equilibrium manifold, \( \mathcal{E}(r) \), which provides a benchmark equilibrium. In case of a merger, we numerically solve for \( \mathcal{E}(r) \) subject to the boundary conditions (i.e., value-matching and smooth-pasting in equations (17) and (43)), 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.
5 Implications

5.1 The no-merger equilibrium

To begin our analysis, we solve for the equilibrium of the benchmark model without mergers. The benchmark model’s solution to the differential equation (48) is characterized by an equilibrium set, \( \mathcal{E}(r) \), that forms a manifold in the space of continuation values, \((W^1, W^2)\), as seen in Figure 1, where the continuation value of firm 1 (firm 2) is on the horizontal (vertical) axis. For the numerical analysis of the model’s solution, 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 \( r = 1.15 \). Using equations (7) and (8), note that 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 1 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. Observe that the firms have a region, corresponding to the northeast stretch of the equilibrium manifold, in which they engage in cooperative behavior in that their output levels are around 4 and hence highly collusive. We refer to this as the market sharing region. 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. Put differently, the figure shows that total output increases as we move from the midpoint of the collusive region to the midpoint of the price war region. Both firms have an incentive to overproduce because the winner of the overproduction-region gets rewarded by becoming dominant for some period of time in one of the two contestability regions that surround the price war region. Subsequently, outputs move to \((10, 0)\) or \((0, 10)\), so one firm becomes dominant in output and reaches its maximal value on the equilibrium manifold, which is the entrant and incumbent region, before both firms potentially move to the market sharing region or return to the contestability region.

In the upper right region of \( \mathcal{E}(r) \) in Figure 1, output is low for both firms in that it sums up to about 8. That is, the firms get close to the monopoly output, which is, according to equations (11) and (12), \( a^*_i = 3.75 \) for each of the two firms (or 7.5 for both firms). Their continuation values are...
This figure plots the no-merger equilibrium manifold (blue, solid line). Firm 2’s output choices are outside the equilibrium manifold. Note that, 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 \( r = 1 \).

Consequently higher in the dynamic duopoly’s market sharing equilibrium (e.g., 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 it is depicted by the (dotted) monopoly line for all (feasible) sharing rules in the unit interval. Clearly, it is unattainable in either the dynamic or the static duopoly. The northeast region is thus the Pareto frontier of \( \mathcal{E}(r) \) where at least one firm colludes by producing less than the static best response and letting the other firm be the production leader. Recall that we refer to this as the market sharing regime. 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 by moving away from the center of the market sharing region. In addition, the noisy
observations of firms’ actions make it difficult determine whether movement away from the market sharing region is due to a series of random shocks (i.e., noise) or due to less collusive output choices.

Observe that it is the upper left segment of the equilibrium manifold in Figure 1 where 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. While firm 2 chooses actions closer to its (static) duopoly response, firm 1 needs strong incentives to “stay out.” To reward firm 1 for staying out, firm 2 accommodates, and to punish firm 1 for deviating, firm 2 fights — the entrant and incumbent regime. Moreover, at the lower right and upper left, output is very asymmetric: at the lower right for example, firm 1’s output is high (i.e., 10) and 2’s is low (i.e., 0). In the lower right segment of equilibrium manifold, firms thus display similar strategies with the roles of firm 1 and 2 reversed, namely firm 1 is the incumbent and firm 2 is the entrant.

Finally, in the contestability and price war regimes, output increases from around 13 to 20, and continuation values decline for both firms. For example, above the center of the price war region (i.e., intersection of $E(r)$ with the 45 degree line), firm 1 is acting passively by producing close to its static best responses, while firm 2 aggressively overproduces. At this point, firm 2 is rewarded for overproducing by being able to drive firm 1 out of the market. This is the contestability regime, which will be the precursor of the endogenous merger regime in the next figure. At the intersection of $E(r)$ with the 45 degree line, firms engage in a price war in that both firms aggressively overproduce, i.e., outputs are (10,10) and substantially exceed the (static) duopoly outputs of (5,5), which leads continuation values to drop well below (25,25). The firms have incentives to do so because the firm that looks more aggressive will come out as the winner of the price war (i.e., continues to fight by overproducing while the other firm accommodates by underproducing). The winner of the price war gains in the contestability region, becomes the incumbent in the adjacent entrant and incumbent regime, and subsequently gains relatively more from the collusive play in the adjacent market sharing regime.

5.2 Merger equilibrium

We continue our analysis by solving the differential equation (48) for the equilibrium manifold, $E(r)$, with mergers by incorporating the boundary conditions for value-matching and smooth-pasting in equations (17) and (43). In this solution of the dynamic game, firms also select dynamically optimal actions that are not static best responses. As a result, merger incentives arise endogenously when
firms sufficiently deviate from their collusive actions. As illustrated in Figure 5, in the presence of an anti-competitive merger, the equilibrium manifold is significantly affected by the restructuring opportunity. 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 (36), 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 lie more on the lower right section of the line; conversely, if firm 2 captures more of the gains, then the merger point is more on the upper left section of the line.

There are two potential states in which a merger can take place, and they apportion different shares to the firms. The merger line smoothly pastes to two optimally determined merger points of the equilibrium manifold. We refer to the range between these two endpoints as the endogenous merger regime, which is, unlike the other regimes, an absorbing one. Therefore, the merger line presents an absorbing boundary for the equilibrium manifold.

Figure 2 establishes that firms’ equilibrium behavior anticipates the impending merger, as seen both in terms of the dynamically optimal output choices and in terms of the resulting continuation values. In the equilibrium with mergers, the two firms are able to keep their outputs lower in a coordinated way than in the no-merger equilibrium. In particular, the equilibrium manifold with mergers is entirely contained inside the original no-merger equilibrium manifold in Figure 1. 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 in the equilibrium with mergers. The punishments are weaker because the opportunity to merge eliminates the severe punishments in the price war regime. The low payoffs from the price war regime are “clipped away” by the endogenous mergers at the endpoints of the new (merger) equilibrium manifold. However, this weakens maximally possible punishments, which makes the entrant and incumbent region less attractive. In this sense, the dynamically optimal actions are more muted (i.e., involve less fighting and accommodating) in the merger equilibrium.

16It is therefore more precise to state the stopping time $T_m$ as $T_{m_1} \wedge T_{m_2}$, where $m_1$ and $m_2$ are the merger points.

17One might expect that as the cost of merging shrinks, the pure monopoly payoff manifold would be attained in the limit. However, this is not so: the weakening of punishments means that the maximum collusive profit shrinks as the merger cost shrinks. Moreover, below a threshold minimum but strictly positive cost (in our example around 8.1), the merger equilibrium no longer exists; we conjecture that firms simply merge immediately below this threshold.
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. Note that, 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 $r = 1$.

Observe that 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 2 compared to Figure 1. 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 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 to the previous figure’s no-merger equilibrium manifold, total output tends to be lower in the worst stages of the dynamic game. As a result of, on average, more collusive output choices (which generate, on average, higher continuation values in the presence of mergers) both firms are better off, even if they never merge. However, the maximally attainable continuation values in the entrant and incumbent regions are slightly lower (around 80 instead of 83), because the threat of punishments is weaker in the presence of mergers. In sum, allowing for anti-competitive mergers implies pronounced periods of pre-merger collusion and leads to less overproduction (i.e., attenuated behavior in the entrant and incumbent regime and potentially complete avoidance of the price war region).

We illustrate the increased competition stemming from mergers in Figure 3. 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. Recalling that this region is highly stable, this demonstrates our main point, that collusion is weakened by the prospect of mergers.

In practice, merger gains (or payoffs) are also typically split asymmetrically among the merging firms. 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. (We will therefore refer to the former as the target and the latter as the acquirer when we analyze merger returns.) This provides an explanation for asymmetric equity shares based on the product market’s regime in the dynamic game (i.e., output strategies). 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. Notably, this asymmetry is not driven by any inherent asymmetry in the demand functions, noise parameters, or other parameters, which we have assumed away precisely to make this point. The asymmetry in equity shares of the merged entity is therefore driven, in general, by dynamic optimization and, in particular, by the state of product market competition that the firms have attained as a result of cumulative play of the noisy duopoly game.
This figure plots the merger equilibrium manifold from Figure 2 and the no-merger equilibrium manifold from Figure 1. 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.

5.3 Collusion and the dearth of mergers

As a next step in the analysis, we examine the stability of the merger and the no-merger equilibria.\textsuperscript{18} There are two aspects of dynamic behavior that are important: the first is the local stability of regions of the equilibrium manifold, and the second is the strength of that stability, which corresponds to the expected movement away from a region of the equilibrium manifold. 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 in equation (27), indicates the strength of stability. The direction of the arrows informs about the local stability of each region.

\textsuperscript{18}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 $r = 1.$

Given that the arrows point away from each other in the price war region of Figure 4, 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 node in the contestability region nearer the price war or merger node.\textsuperscript{19} Comparing the stability diagrams, in both figures the instability of the contestability region makes it likely that the firms

\textsuperscript{19}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.
Figure 5. Merger equilibrium manifold and stability

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 $r = 1$.

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.\footnote{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.}

Importantly, the area close to the merger points in Figure 5 is unstable. As a result, dynamic optimizing behavior of the firms actually delays the onset of mergers (i.e., makes them rare). As the firms approach one of the merger points, they exhibit the tendency to revert back to the collusive region of the equilibrium manifold. This is also illustrated by Figures 6 and 7, which display simulated time-series samples of play for 5,000 periods in the no-merger and merger cases.
respectively. In each figure, the period is on the horizontal axis and the polar coordinate angle, \( \theta \), characterizing the region of the equilibrium manifold, is on the vertical axis. Thus, collusion corresponds to \( \theta = \frac{\pi}{4} \approx 0.7854 \), and the full price war punishment region to \( \theta = \frac{5\pi}{4} \approx 3.9270 \) or \(-\frac{3\pi}{4} \approx -2.3562\). It is apparent that in the no-merger case of Figure 6, divergence from the collusion region into the punishment region is rare, and also brief. In the merger model, mergers occur at the angle corresponding to the punishment regimes, (i.e., at the merger points \( \theta = \frac{5\pi}{4} \) or \(-\frac{3\pi}{4}\)). It is evident from Figure 7 that in case of mergers collusion is also highly stable, with only rare movements towards the merger angles, although the weakening of punishments increases the frequency of these movements slightly.\(^{21}\)

**Figure 6. History of collusion and punishment: No-Merger Case**

![Figure 6](image_url)

This figure plots a simulated history of the non-merger equilibrium for 5,000 draws. The vertical axis indexes \( \theta \), the polar-coordinate angle corresponding to points on the equilibrium manifold. Collusion corresponds to \( \theta = \frac{\pi}{4} \), and the full price war punishment region to \( \theta = \frac{5\pi}{4} \) or \(-\frac{3\pi}{4}\). It is evident that excursions away from collusion into the punishment states are extremely rare and brief. 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 \( r = 1 \).

Taken together, the dynamic game’s analysis reveals that, 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

\(^{21}\)A histogram of occupancy times for all of the regimes, corresponding to ranges of the angles \( \theta \), would be so lopsided toward the collusion state that little insight would be drawn from it. Also, to simulate the merger equilibrium we follow the equilibrium as if firms could de-merge once they have merged; however, a sequence of independent histories, using starting points distributed around the equilibrium manifold and terminating at the merger, would similarly point to the stability of collusion, but even more strongly.
This figure plots a simulated history of the merger equilibrium for 5,000 draws. The vertical axis indexes $\theta$, the polar-coordinate angle corresponding to points on the equilibrium manifold. Collusion corresponds to $\theta = \frac{\pi}{4}$, and the merger points are $\theta = \frac{3\pi}{4}$ or $-\frac{3\pi}{4}$. It is evident that excursions away from collusion and toward the merger points are extremely rare. 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 $r = 1$.

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, e.g., Andrade, Mitchell, and Stafford (2001) point out that stronger antitrust laws and stricter enforcement 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.

Existing merger theory largely builds on static considerations in the sense that, at each instance, both theory and practice would contemplate a merger the moment it occurs. In contrast, merger incentives arise endogenously in our model dynamically when firms sufficiently deviate from their collusive actions, which in turn influences pre-merger strategies. Consequently, the impact of mergers on rival firms’ profits is nil or negative in relation to changes in industry concentration as in Eckbo (1985) who doubts that the firms were already acting monopolistically prior to the merger.

22Recently, Ahern and Sosyura (2011) provide support for a dynamic view by examining pre-merger strategic manipulation of news releases that affect the firm’s stock prices and hence the outcome of merger negotiations.
Similarly, Eckbo (1983) concludes that these results are inconsistent with the market power doctrine and consistent with the efficiency argument. He argues: “Thus, if mergers typically take place to realize efficiency gains, we cannot conclude that the ‘synergy’ effect is expected to produce a significant expansion of the merging firm’s share of the market along with an increase in industry rate of output. If scale economies are involved, then these seem on average to be insufficient to make the rivals worse off. Furthermore, the same evidence contradicts the argument that the merging firms were expected to initiate a (monopolistic) ‘predatory’ price war after consummation of the merger.”

From a methodological standpoint, our dynamic model casts doubt on the tests that lead to rejection of the market power doctrine. In fact, 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:

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 key insights from the model analysis help to better understand several important 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.
5.4 Pre-merger stock returns

Before considering pre-merger stock returns, it is helpful to consider stock prices throughout the game as a function of the computational state variable that defines the location of the two firms on the equilibrium manifold. Using again the baseline parameter values, Figure 8 plots the pre-merger continuation values, or stock prices, of firm 1 (blue, solid line) and firm 2 (red, dashed line) against the polar coordinate angle $\theta$. The merger cost increases from the top-left panel ($k = 24$) to the top-right panel ($k = 34$), and further from the bottom-left panel ($k = 44$) to the bottom-right panel ($k = 67.2$). There are several interesting patterns for the firms’ pre-merger stock prices. First, the payoffs at the endpoints, where the firms are merging (i.e., at the left and right margins of the plots), are generally asymmetric across the firms. Only in the lower right panel, where the merger cost is highest, we see that the payoffs (and hence shares in the emerged entity) are not asymmetric (see discussion of merger shares in Section 4.2). We will therefore refer to the less valuable firm as the target and the more valuable firm as the acquirer when we analyze merger returns. Second, the firms’ stock prices are equal in the middle of the market sharing regime at the angle $\theta = \frac{\pi}{4} \approx 0.7854$, where the vertical axis crosses the horizontal axis. Also striking is that the stock price of the acquiring firm first rises, as expected, but then gradually declines as it gets sufficiently close to the merger boundary. Finally, the range of continuation values increases as the fixed cost of merging rises.

Recalling that the bottom-right panel corresponds to the no-merger equilibrium from Figure 1, the analysis indicates that there are notable differences between firms’ price patterns in the merger and in the no-merger equilibrium. For example, the acquisition price can be higher than the stock prices most closely observed before the announcement of the deal. This is just another way of saying that there is locally a positive pre-merger return for targets in our model, which leads us to our next implication regarding pre-merger returns.

To study pre-merger return dynamics, we approximate instantaneous returns with the local percentage change in value at each point on the equilibrium manifold. We then plot these quantities against the polar coordinate angle $\theta$ in Figure 9. The inner range of $\theta$ represents the normal operation of the firms: locally, one firm gains while the other loses as they approach the center of the cooperation region which lies on the 45 degree line in the previous figures. At the outer extremes, the merger is imminent. It is apparent that the larger firm, which we loosely label the acquiring firm, exhibits locally negative pre-merger returns, and the acquired firm has locally positive pre-merger returns. In addition, the figure reveals that the acquiring (acquired) firm first experiences an extended...
This figure plots the pre-merger continuation values (or stock prices), $W_i$, of firm 1 (blue, solid line) and firm 2 (red, dashed line). Firm 1 acquires 2 when moving clockwise (on the left) but firm 2 acquires 1 when moving counterclockwise (on the right) away from the center of the merger equilibrium manifold of most collusive outputs (i.e., from $\theta = 0.79$ or 45 degrees). The merger cost increases from the top-left panel ($k = 24$) to the bottom-right panel ($k = 67.2$). We use the baseline environment in which $\Pi_1 = 30, \Pi_2 = 30, \beta_1 = 2, \beta_2 = 2, \delta_1 = 2, \delta_2 = 2, \sigma_1 = 1, \sigma_2 = 1, \text{ and } r = 1$.

### Figure 9. Pre-merger stock returns

This figure plots the instantaneous pre-merger stock returns, $dW_i/W_i$, to firm 1 (blue, solid line) and firm 2 (red, dashed line). Firm 1 acquires 2 when moving clockwise (on the left) but firm 2 acquires 1 when moving counterclockwise (on the right) away from the center of the merger equilibrium manifold of collusive outputs (i.e., from $\theta = 0.79$ or 45 degrees). The merger cost increases from the top-left panel ($k = 24$) to the bottom-right panel ($k = 67.2$). We use the baseline environment in which $\Pi_1 = 30, \Pi_2 = 30, \beta_1 = 2, \beta_2 = 2, \delta_1 = 2, \delta_2 = 2, \sigma_1 = 1, \sigma_2 = 1, \text{ and } r = 1$. 
Consistent with these findings, Schwert (2000) and Andrade, Mitchell, and Stafford (2001) document that the average \([-20, \text{Close}]\) return to target firms is positive but the average \([-20, \text{Close}]\) return to acquiring firms is negative over the 1973–98 period. Note that because we do not model idiosyncratic noise, our returns do not reflect any systematic risk factors by assumption. Moreover, we do not model any other incentives for merging, such as economies of scale, removal of entrenched management, behavioral biases to overinvest by making acquisitions, co-insurance and diversification, etc. The economic magnitude of our pre-merger returns is therefore expected to be moderate.\(^{23}\) While the stock returns to targets tend to be larger than predicted by our theory, the negative stock returns to acquirers follow more closely the ones observed in practice but are uniquely produced by a dynamic duopoly with imperfect information.\(^{24}\)

6 Conclusion

We studied mergers in a model using Sannikov’s continuous-time version of the noisy collusion dynamics models of Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1986). Two competing firms in an industry know that they will merge when conditions are right, but that the merger will incur significant fixed costs. A merger provides monopoly profits to the participating firms and also resolves imperfect information issues. We show that output strategies and firm valuations must account for the longer term potential for mergers, even if the merger is remote from the current state, which in turn influences merger decisions and product market competition in important ways.

A corollary is that regulators who wish to preclude collusion must first understand its dynamic implications. Our model demonstrates that collusion will in general exceed the collusion seemingly warranted at the time of the merger, because regulators do not necessarily forecast impending mergers; indeed firms prefer to hide that information. Moreover, mergers will be rare because of the stability of the collusive region and the local instability of the merger points. For practical purposes, mergers will be phantoms. But when they occur, firms will be colluding less than normally—which serves to identify and validate anti-competitive mergers from a regulatory standpoint. Crucially, the potential for anti-competitive mergers generates stable collusion outcomes long before they occur.

Counterintuitively, the potential to merge reduces the discounted profits from merging, because

\(^{23}\) Bradley and Sundaram (2006) also study merger activity in the 1990s. One of their findings is that the importance of “agency costs” or “hubris theory” is perhaps overstated because they apply only to a small minority of cases.

\(^{24}\) The results in Schwert (2000) and Andrade, Mitchell, and Stafford (2001) are consistent with other studies, such as Jensen and Ruback (1983), Jarrell, Brickley, and Netter (1988), and Moeller, Schlingemann, and Stulz (2005).
the weakening of punishments resulting from the option to merge reduces pre-merger collusion, thus improving net welfare. Indeed, the results suggest that reducing regulatory barriers to mergers, which in our framework would translate as a reduction of the fixed cost of merging, increases welfare as it weakens punishments and reduces pre-merger collusion.

We close by noting areas for future research in this class of dynamic models. First, we have restricted attention to two firms. It would be of compelling interest to extend our analysis to three or more firms, as we could characterize the impact of mergers on non-merging rivals. This would be key because many of the empirical tests of the impact of mergers rest on measuring the impact on those non-merging firms. Moreover, the theoretical literature (for example beginning with Salant, Switzer, and Reynolds (1983)) has found 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. The reason for these counterintuitive results is that a fully fleshed out model must include the endogenous response of the non-merging competitors, and this can vitiate the profits of the merging firms. Perry and Porter (1985) examined this result further with a more detailed treatment of the allocation of costs in the merged firm and moved the conclusion back in the classical direction. Farrell and Shapiro (1990) established that quantity competition in the 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), examined 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.

Second, as we showed, the model is related to agency. While we (as does the literature) treat firms as black boxes that are able to hide information, one might reinterpret this as a more 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.
References


Ashenfelter, Orley C., Daniel S. Hosken, and Matthew C. Weinberg, 2013a, Did US beer mergers cause a price increase?, *VoxEU*.


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 order to generate the stochastic continuation value process of the state:

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

Notice the resemblance of the terms in the drift to the stage-game payoffs in the Bellman equation. Substituting from the Bellman equation (22) into (A.1) yields the following simpler expression for the continuation value process:

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

We further modify this equation by using the optimality condition (23) to eliminate the $W^1_{X^1_t}$ term, replacing $W^1_{X^1_t}$ with $-r g_1(A^1)$ (i.e., the envelope condition):

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

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

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

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:

$$N(\theta) = (\cos(\theta), \sin(\theta)),$$
$$T(\theta) = (-\sin(\theta), \cos(\theta)). \quad (B.5)$$

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

$$W^2_{W^1} = \frac{dW^2}{dW^1} = -\frac{\cos(\theta)}{\sin(\theta)}. \quad (B.6)$$
Hence the second-stage Bellman equation (39) becomes:

\[
\max_{A} \left\{ r \left( g_1 - W^1 \right) - r(g_2 - W^2) \left( -\frac{\sin(\theta)}{\cos(\theta)} \right) \right. \\
\left. + \frac{1}{2} \left( -r \frac{\cos(\theta)}{\sin(\theta)} A_1^1 \right)^2 W_{W^2W^2}^1 + \frac{1}{2} \left( r g_2 A^2 \right)^2 W_{W^2W^2}^1 \right\},
\]

which then leads to:

\[
W_{W^2W^2}^1 = \max_{A_1^1} \left\{ \frac{(g_1 - W^1) - (g_2 - W^2) \left( -\frac{\sin(\theta)}{\cos(\theta)} \right)}{r \left( \left( -\frac{\cos(\theta)}{\sin(\theta)} g_1 A_1^1 \right)^2 + \left( g_2 A^2 \right)^2 \right)} \right\},
\]

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

\[
W_{W^2W^2}^1 = \max_{A_1^1} \left\{ \frac{1}{\cos(\theta)} \left( \cos(\theta) (g_1 - W^1) - (g_2 - W^2) \left( -\sin(\theta) \right) \right)}{r \left( \left( \frac{g_1 A_1^1}{\sin(\theta)} \right)^2 + \left( g_2 A^2 \right)^2 \right)} \right\},
\]

or

\[
W_{W^2W^2}^1 = \max_{A_1^1} \left\{ \frac{1}{\cos(\theta)^3} \frac{\cos(\theta) (g_1 - W^1) + \sin(\theta) (g_2 - W^2)}{r \left( \left( \frac{g_1 A_1^1}{\sin(\theta)} \right)^2 + \left( g_2 A^2 \right)^2 \right)} \right\}.
\]

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:

\[
W_{W^1W^1}^2 = \max_{A_2^2} \left\{ \frac{1}{\sin(\theta)^3} \frac{\cos(\theta) (g_1 - W^1) + \sin(\theta) (g_2 - W^2)}{r \left( \left( \frac{g_1 A_1^1}{\sin(\theta)} \right)^2 + \left( g_2 A^2 \right)^2 \right)} \right\}.
\]

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

As shown in Appendix 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, \( \kappa(W) \):

\[
\cos(\theta)^3 W_{W^2W^2}^1 = \sin(\theta)^3 W_{W^1W^1}^2 \equiv \frac{1}{2} \kappa(W).
\]

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:

\[
\kappa(W) = \max_{A_1^1, A_2^2} \frac{2 N (g - W)}{r |\phi|^2},
\]

46
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 (39) is equivalent to the ode in equation (B.10) for the curvature of the equilibrium set boundary $\partial E$, 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:

$$\max_x \{ f(x) + A g(x) \}.$$  \hspace{1cm} (C.14)

The first-order condition is:

$$f'(x) + A g'(x) = 0,$$  \hspace{1cm} (C.15)

or:

$$A = -\frac{f'}{g'}.$$  \hspace{1cm} (C.16)

Now consider the maximization problem:

$$\max_x \frac{f(x) g(x)}{g(x)}.$$  \hspace{1cm} (C.17)

The first-order condition can be written as:

$$\frac{f}{g} = \frac{f'}{g'},$$  \hspace{1cm} (C.18)

and therefore, at the maximum, we have that:

$$\max_x \frac{f}{g} = \frac{f'}{g'}.$$  \hspace{1cm} (C.19)

Therefore,

$$A = -\frac{f'}{g'} = -\max_x \frac{f}{g}.$$  \hspace{1cm} (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:

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

To begin, note that:

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

This is equal to:

\[ \frac{d}{dW^1} \frac{1}{W_{W^1}^2} = -\frac{1}{(W_{W^1}^2)^2} W_{W^1 W^1}^2 = -W_{W^1 W^1}^2 \left( \frac{\sin(\theta)}{\cos(\theta)} \right)^2. \]  \hfill (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 4, which establish optimality of the smooth pasting condition. Our strategy will be a bit different 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 smooth pasting condition. That is, we provide a verification of sufficiency of local optimality at the merger point. More concretely, the steps in this demonstration will be to

(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 \( W_{W^1 W^2}^1 \);

(iii) Substitute this reduced form expression for \( W_{W^1 W^2}^1 \) into the second-order condition, as well as the value-matching condition, establishing that the second-order condition is negative.

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

\[ \max_{A^1} \left\{ r(g_1 - W^1) - r(g_2 - W^2)W_{W^2 W^2}^1 + \frac{1}{2} \left( -r W_{W^1 g_1 A^1}^2 \right)^2 W_{W^1 W^1 W^2 W^2}^1 + \frac{1}{2} \left( rg_{2A^2} A^2 \right)^2 W_{W^1 W^2 W^2}^1 \right\}. \]  \hfill (E.4)

with optimality condition

\[ rg_{1A^1} - rg_{2A^1} W_{W^2 W^2}^1 + \left( -r W_{W^1 A^2}^2 g_1 A^1 \right) g_{1 A^1 A^1} + \left( r g_{2A^2} A^2 \right) g_{2A^2 A^1} W_{W^1 W^2 W^2}^1 = 0 \]  \hfill (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

$$rg_1 A_1^1 - rg_2 A_1^1 W_{W^2}^1 + r^2 \left(-W_{W^1}^2 (g_1 A_1^1)^2 + (g_2 A_2^1 A_1^1)^2\right) W_{W^2 W^2}^1$$

Now we can substitute the smooth pasting condition: $W_{W^2}^1 = -1$: This step shows that it is a verification process, and there should be some discussion prior to this step, and also a reminder at this step.

$$rg_1 A_1^1 + rg_2 A_1^1 + r^2 \left(-W_{W^1}^2 (g_1 A_1^1)^2 + (g_2 A_2^1 A_1^1)^2\right) W_{W^2 W^2}^1$$

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

$$rg_1 A_1^1 + rg_2 A_1^1 + r^2 \left[(g_1 A_1^1)^2 + (g_2 A_2^1 A_1^1)^2\right] W_{W^2 W^2}^1$$

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

$$g_1 A_1^1 = -2\beta_1 - \delta_1 \quad g_1 A_2^1 A_1^1 = -\delta_2$$

which are both negative by assumption.

Carrying out step (ii), we next transform the system of Bellman equations to ratio form in order 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 (39):

$$W_{W^2 W^2}^1 = \max A^1 \left\{ \frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r} W_{W^2}^1}{r \left((W_{W^1}^2 g_1 A_1^1)^2 + (g_2 A_2^2)^2\right)} \right\}, \quad (E.6)$$

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

$$W_{W^2 W^2}^1 = \max A^1 \left\{ \frac{(g_1 - W^1) - r(g_2 - W^2)\frac{1}{r} W_{W^2}^1}{r \left(\frac{1}{W_{W^2}^1} g_1 A_1^1 + (g_2 A_2^2)^2\right)} \right\}, \quad (E.7)$$

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 $W_1$ from the ratio reformulation of the Bellman equation in equation (E.7), evaluated at the value matching and smooth pasting point:

$$W_1^1 W_2^2 = \frac{(g_1 - W^1) - r(g_2 - W^2) \frac{1}{r} W_1^1}{r \left( \frac{1}{W_2^2} g_{1,A}\right)^2 + (g_{2,A})^2} = \frac{(g_1 - W^1) + (g_2 - W^2)}{r \left( \frac{1}{r} g_{1,A}\right)^2 + (g_{2,A})^2}$$ (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

$$(g_1 - W^1) + (g_2 - W^2) = 0$$

at the smooth pasting point.

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 (40)). At the smooth pasting point, these drifts necessarily are equal and of 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, $W_1^1 W_2^2$, is zero.) ■

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

We can use this result to prove global optimality.

Proof: (of Proposition 4) The optimality of the main action $A_{1,t}$ as the solution of the HJB equation (39) follows directly from conventional optimal control considerations. Therefore the optimal value state path satisfies equation (B.10); further derivations lead to the differential equation (48), combined with equation(Appendix B), as demonstrated in Appendix 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 (41) 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.

A remark is that 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 only from above. Therefore it is only necessary 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. However, notice that the smooth pasting condition is locally optimal at the smooth pasting point; by continuity of the solution that obeys the main differential equation (which would be easy to establish) this optimality argument must also hold in a neighborhood of the smooth pasting point. (The fixed coefficients assumption would hold locally in a neighbourhood of the fixed point and could thus be applied locally and then possibly expanded globally using continuity.)

It is important to note that Dixit’s approach uses a local argument but establishes the global optimality of the the smooth pasting condition. Our argument so far is purely local. Indeed, numerical experiments establish that there are multiple solutions for self-generating manifolds that satisfy the smooth pasting condition. ■

Appendix F  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 more, 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. How do we incorporate this deadweight 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.

Appendix G  Refusals to merge, punishment and jumps

In this appendix 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 in particular: 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, then 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 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.

It is key that the punishment phase must itself 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. Should such an equilibrium exist, we will refer to it as a refusal-punishment

We are grateful to Mikhail Panov for raising this question.

The ideas and contributions of Mikhail Panov, Yuliy Sannikov, and Andrzej Skrzypacz led to the discussion in this appendix.
(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.

**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 continuation game, must be agreed 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 in the future, 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).

The potential punishment alters 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 by the parameters of the model, 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 10). 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.}

**Figure 10. The refusal-punishment equilibrium**

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.

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 manifold, which lies entirely inside the no-merger manifold.

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.